UNIVERSITY OF Southampton

# University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Stuart James Blakemore (2017) "Clinical and biological implications of genomic lesions in Chronic Lymphocytic Leukaemia", University of Southampton, Cancer Sciences Unit, Faculty of Medicine, PhD Thesis, pagination.

Data: Blakemore SJ (2017) Clinical and biological implications of genomic lesions in Chronic Lymphocytic Leukaemia. URI [dataset]

# UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Cancer Sciences Unit

Volume 1 of 1

**Clinical and biological implications of genomic lesions in Chronic Lymphocytic Leukaemia**

by

**Stuart James Blakemore**

Thesis for the degree of Doctor of Philosophy

November 2017

**UNIVERSITY OF SOUTHAMPTON**

# ABSTRACT

FACULTY OF MEDICINE

Cancer genomics

Thesis for the degree of Doctor of Philosophy

**CLINICAL AND BIOLOGICAL IMPLICATIONS OF GENOMIC LESIONS IN CHRONIC LYMPHOCYTIC LEUKAEMIA**

Stuart James Blakemore

The aim of this thesis was to interrogate the application of targeted genomics in the clinical management of Chronic Lymphocytic Leukaemia (CLL), and to explore inhibitory compounds of recurrently mutated genes in CLL. The genomic landscape of the LRF CLL4 clinical trial (n=500) was conducted using a custom 25 gene targeted re-sequencing NGS panel, identifying 903 variants (1.8/patient). *SF3B1* was identified as the most recurrently mutated gene in the cohort (26%), with the mutation landscape and subclonal architecture assessed for all 25 genes, identifying previously observed and novel associations. Clinical statistical survival analysis of the CLL4 TruSeq mutation data was undertaken for overall and progression-free survival, using traditional techniques and supervised machine learning tools. *SF3B1* was identified to associate with OS independently of multiple CLL biomarkers in a multivariate model, and as an important feature by machine learning approaches. In addition, subclonal *TP53* mutations predicted poor OS and PFS in cases treated with chlorambucil, and co-mutated/deleted del(11q) and *BIRC3* were found to independently predict for a poor PFS. Since *SF3B1* was found to be the most recurrently mutated gene in CLL4, and it associated with poor OS, it was selected for *in vitro* targeting using splicing inhibitors (Spliceostatin A and Meayamycin B) in CLL cells. Splicing inhibitors elicited substantial apoptosis, with Spliceostatin A inducing downregulation of Mcl-1 at the protein and RNA level, as well as acting synergistically with venetoclax in the context of micro-environmental support. Therefore, direct inhibition of Mcl-1 in CLL cells was undertaken using the Mcl-1 specific inhibitor UMI-77, eliciting apoptosis at micro molar concentrations. However, UMI-77 did induce re-sensitisation to the level of SSA in the context of CLL microenvironment support, suggesting that inhibition of Mcl-1 alone may not be sufficient to overcome BCL-2 inhibitor resistance in CLL. This work has demonstrated that the application of targeted genomic screening can offer added value to CLL patient management where the monitoring of *SF3B1* mutations can identify additional adverse disease events. In relation to targeting drugs to the spliceosome pathway in CLL, where SF3B1 is a major component, the current work demonstrated effective re-sensitisation of CLL cells to combination inhibitor exposure. Together, this research has provided a greater understanding of the clinical utility of gene mutation screening and challenges surrounding development of the next generation of therapies to circumvent molecular pathways that CLL cells use to proliferate and drive disease progression.

# Table of Contents

**Table of Contents** ..... **i**

**List of Tables** ..... **ix**

**List of Figures** ..... **xi**

**DECLARATION OF AUTHORSHIP** ..... **xvii**

**Acknowledgements** ..... **xix**

**Definitions and Abbreviations** ..... **21**

**Chapter 1: Introduction and Literature Review** ..... **29**

1.1 Introduction ..... 29

1.2 Literature Review ..... 31

1.2.1 Cancer ..... 31

1.2.2 Hallmarks of cancer ..... 32

1.2.3 Genotoxic stress and cancer ..... 33

1.2.4 Haematological malignancies ..... 36

1.2.5 Chronic Lymphocytic Leukaemia ..... 37

1.2.6 Disruption of intracellular signalling pathways in CLL ..... 46

1.2.7 Next Generation Sequencing ..... 57

1.3 Aims ..... 61

**Chapter 2: Methodology** ..... **63**

2.1 Laboratory methods ..... 63

2.1.1 Patient cohorts ..... 63

2.1.2 Recovering CLL patient cells ..... 64

2.1.3 Live cell count and sample concentration calculation ..... 65

2.1.4 DNA extraction ..... 66

2.1.5 RNA extraction ..... 67

2.1.6 Nucleic acid assessment ..... 68

2.1.7 Whole Genome Amplification (WGA) ..... 70

2.1.8 Primer design ..... 71

2.1.9 Sanger sequencing ..... 71

2.1.10 Droplet digital PCR ..... 72

2.1.11 Reverse Transcription PCR (RT-PCR) ....74
2.1.12 Illumina TruSeq Custom Amplicon Next Generation Sequencing ....75
2.1.13 Thermo Fisher Ion Torrent Next Generation Sequencing ....77
2.1.14 Oxford Gene Technologies (OGT) SureSelect Next Generation Sequencing ....79
2.1.15 *In vitro* drug treatment ....80
2.1.16 B-cell Receptor stimulation ....80
2.1.17 Flow cytometry ....81
2.1.18 Protein extraction ....83
2.1.19 Immunoblotting ....84
2.1.20 Immunoblotting data analysis ....85
2.2 Statistical/analytical ....87
2.2.1 Checking TruSeq panel design capture regions in UCSC Genome Browser87
2.2.2 Individual amplicon read depth analysis using Qlucore ....87
2.2.3 Bioinformatics pipeline ....88
2.2.4 Additional disease associated mutation annotation ....89
2.2.5 IGV mutation validation ....89
2.2.6 R ....90
2.2.7 IBM SPSS ....93
2.2.8 Microsoft Excel ....93
2.2.9 GraphPad Prism ....94
**Chapter 3: The mutational landscape of the LRF UK CLL4 trial cohort ....95**
3.1 CLL4 TruSeq gene panel design ....101
3.1.1 Initial design of the CLL4 TruSeq panel ....101
3.1.2 Re-designs of the CLL4 TruSeq panel ....101
3.2 Sample curation, library preparation and sequencing ....103
3.3 CLL4 TruSeq study sequencing coverage analysis ....105
3.3.1 Identifying the distribution of sequencing coverage in the CLL4 TruSeq study per MiSEQ run ....105
3.3.2 Correlating the mean sequencing coverage between MiSEQ runs at the amplicon level ....107

3.3.3 Identifying the distribution of sequencing coverage in the CLL4 TruSeq study per gene ....110
3.3.4 Individual amplicon performance analysis per patient....112
3.3.5 Biological and clinical implications of missing data: Insight using *TP53*115
3.4 Bioinformatics and mutation data processing of the CLL4 TruSeq study....119
3.4.1 Handing of variant call files (.vcf)....121
3.4.2 Removing recurrent false positives from the mutation data....121
3.4.3 Filtering strategy to retain high quality, clinically and biologically relevant mutations....121
3.4.4 *In silico* mutation validation using Integrated Genomics Viewer....122
3.4.5 Assessment of minor subclones using orthogonal sequencing and method agreement analysis....122
3.4.6 Dealing with regions of high percentage GC content ....124
3.4.7 Additional annotation and filtering to ensure high quality variants ....124
3.5 Concordance between TruSeq and previous CLL4 studies....125
3.5.1 CLL4 TruSeq study identified more *TP53* variants than Gonzalez *et al.* (2011), with increased sensitivity....126
3.5.2 CLL4 TruSeq study identified more *SF3B1* variants than Oscier *et al.* (2012), with increased sensitivity....129
3.5.3 CLL4 TruSeq study identified more *NOTCH1* variants than Oscier *et al.* (2012), with increased sensitivity....132
3.6 The mutation landscape in CLL4 ....135
3.6.1 Overview of the mutation landscape in CLL4: *SF3B1* found to be the most recurrently mutated gene in CLL4....135
3.6.2 Distribution of mutation type in the full cohort and stratified by IGHV mutational status: Transitions events predominate in CLL4....139
3.6.3 Mutation landscape in U-CLL and M-CLL subsets of CLL4....141
3.6.4 Overview of mutation landscape in *SF3B1* mutated cases of CLL4....144
3.6.5 Mutated *NBEAL2* and *CTBP2* found to be recurrently mutated in CLL4: First reported case of recurrent incidence in CLL ....147

3.7 Identification of mutated pathways in CLL4 .....150

3.7.1 One mutated intracellular signalling pathway .....150

3.7.2 More than one mutated intracellular signalling pathway.....152

3.8 Statistical association analysis of recurrently mutated genes in CLL4 .....154

3.9 Subclonal architecture of CLL4.....157

3.10 Discussion and Future Directions.....161

3.10.1 CLL4 biological material .....161

3.10.2 Sequencing data quality and coverage of CLL4 TruSeq panel.....162

3.10.3 Bioinformatics and mutation data processing of CLL4 sequencing data164

3.10.4 Orthogonal sequencing and molecular validation of CLL4 mutations ..164

3.10.5 Available data from previous CLL4 publications .....165

3.10.6 Mutation landscape of CLL4.....166

3.10.7 Future work.....167

3.10.8 Conclusions .....168

**Chapter 4: Clinical and statistical survival analysis of the LRF UK CLL4 TruSeq study169**

4.1 Comparison of the CLL4 TruSeq cohort to full the CLL4 clinical trial .....173

4.2 The CLL4 TruSeq custom panel has greater sensitivity but reduced specificity in comparison to previous CLL4 mutation publications.....175

4.2.1 Specificity higher in previous CLL4 study data for capturing OS events in all genes .....175

4.2.2 Comparable specificity between technologies for capturing *NOTCH1* PFS events.....177

4.3 Analysis of the CLL4 time to event survival data.....179

4.3.1 Median time to event in CLL4 remains consistent over time.....181

4.3.2 OS time to event data from 2012 and PFS time to event data from 2011 identify the most variables associated with survival .....182

4.3.3 Mutated genes in CLL4 TruSeq data have proportional and non-proportional hazards.....185

4.3.4 Splitting of OS and PFS survival analysis by median event time identifies additional survival associations.....189

4.3.5 Final remarks of overview of the CLL4 time to event survival data .....193

4.4 Univariate Survival Analysis of CLL4 recurrently mutated genes and clinical biomarkers ..... 195

4.4.1 Overview of variables associated with significant impact on OS in all CLL4 cases ..... 197

4.4.2 Overview of variables associated with significant impact on PFS in all CLL4 cases ..... 199

4.4.3 Impact of the DNA Damage Response and inflammatory/BCR signalling pathway genes on survival in CLL4 ..... 201

4.4.4 The RNA splicing and metabolism genes *SF3B1* and *DDX3X* are associated with poor OS and PFS, respectively ..... 215

4.5 Multivariate Survival Analysis in CLL4 ..... 219

4.5.1 *SF3B1* and *NRAS* found to independently predict poor OS in CLL4 ..... 219

4.5.2 Co-mutated/deleted del(11q) and *BIRC3* independently predicts poor PFS in CLL4 ..... 221

4.5.3 Random Survival Forests analysis in CLL4 ..... 222

4.6 Discussion and Future Directions ..... 227

4.6.1 Sensitivity Specificity analysis ..... 228

4.6.2 Time to event survival data time points, time-dependent covariates and SurvSplit functions ..... 228

4.6.3 Univariate survival analysis ..... 229

4.6.4 Multivariate survival analysis ..... 233

4.6.5 Future directions ..... 235

4.6.6 Conclusions ..... 235

**Chapter 5: Elucidating the clinical utility and mechanism of action of splicing inhibition in CLL ..... 237**

5.1 SSA induces apoptosis in CLL cells in a dose-, time-, and caspase-dependent manner ..... 241

5.1.1 SSA induces apoptosis in CLL cells in a dose- and time-dependent manner ..... 241

5.1.2 SSA induces apoptosis in CLL cells in a caspase-dependent manner ..... 241

5.2 Sensitivity of CLL cells to SSA induced apoptosis ....243
5.3 SSA treatment causes intron retention and modifies *MCL-1* and *SF3B1* RNA splicing in Ramos cells ....245
5.4 SSA treatment reduces MCL-$1_L$ protein expression in Ramos cells ....247
5.5 MCL-$1_L$ overexpression in Ramos cells induces resistance to SSA ....249
5.6 SSA induces intrinsic apoptosis in Ramos cells in a dose- and time-dependent manner ....251
5.7 SSA induces cell death in Eμ-TCL1 cells *in vitro* and Bim may play a role in SSA sensitivity in Eμ-TCL1 cells ....253
5.8 MMB acts analogously to SSA in CLL ....255
5.9 The effect of microenvironmental support on SSA induced apoptosis ....257
5.10 SSA combination treatment with BCL-2 inhibitors in the context of microenvironmental stimulation ....259
5.11 Direct targeting of MCL-1 using the selective small-molecule MCL-1 inhibitor UMI-77 in CLL cells ....261
5.12 UMI-77 induces apoptosis in a dose-, caspase- and time-dependent manner ....263
5.12.1 UMI-77 induces apoptosis in a dose-dependent manner in CLL cells ....263
5.12.2 UMI-77 induces apoptosis in a caspase-dependent manner in CLL cells ....263
5.12.3 UMI-77 induces apoptosis in a time-dependent manner in CLL cells ....265
5.13 Sensitivity of CLL cells to UMI-77 induced apoptosis ....267
5.14 The effect of microenvironmental stimulation on UMI-77 induced apoptosis ....269
5.15 UMI-77 combination treatment with BCL-2 inhibitors in the context of microenvironmental stimulation ....270
5.15.1 UMI-77 combination treatment with venetoclax in the context of CD40L/IL-4 stimulation ....270
5.15.2 UMI-77 combination treatment with venetoclax in the context of BCR stimulation ....272
5.16 Discussion & Future Directions ....275
5.16.1 Splicing inhibition ....275
5.16.2 Mcl-1 inhibition ....277
5.16.3 Future work ....279
5.16.4 Conclusions ....279

**Chapter 6: Discussion and future directions ..... 281**

6.1 Discussion ..... 281

6.2 Future Directions ..... 285

6.2.1 Targeted drug compound and Clustered Regularly Interspaced Short Palindromic Repeats (CRIPSR) screening libraries of all mutations in recurrently mutated genes in CLL, and their impact *in vitro* and *in vivo* 285

6.2.2 *In vitro* and *in vivo* testing of novel mono and combination therapies in specific CLL mutated subsets based on the compound drug and CRISPR library screening results ..... 286

6.3 Final Conclusion ..... 287

**Appendices ..... 289**

A.1 Table of Design 1 CLL4 TruSeq panel manifest file ..... 289

A.2 Correlation analysis of the sequencing coverage in Design 2 ..... 295

A.3 Correlation analysis of the sequencing coverage in Design 3 ..... 296

A.4 Mean sequencing coverage per amplicon, per gene for the CLL4 TruSeq panel Design 2 and 3 ..... 297

A.5 Sequence identity between *CTBP2* and *CTBP2* pseudogenes using the Pei *et al.* (2012)(241) dataset ..... 298

A.6 Mapping of *CTBP2* pseudogenes ..... 299

A.7 MAF file of high confidence variants (n=903) ..... 300

A.8 Table of all CLL, U-CLL and M-CLL power calculations ..... 306

A.9 Waterfall plots of U-CLL and M-CLL ..... 307

A.10 Breakdown of 3 mutated intracellular signalling pathways ..... 308

A.11 Significant associations from CLL4 Fisher's Exact testing ..... 309

A.12 Significant variables – Schoenfeld residuals ..... 311

A.13 Cohort characteristics of SSA/MMB & UMI-77 Southampton CLL cases ..... 312

A.14 SSA reduces cell viability in CLL cells at low nanomolar concentrations in comparison to chlorambucil ..... 313

A.15 The effect of microenvironmental signals on venetoclax induced apoptosis ..... 315

**List of References ..... 317**

# List of Tables

**Table 1 RT-PCR master mix**.....74

**Table 2 Applied R packages and functions**.....92

**Table 3 CLL4 TruSeq Panel Design, Coverage, and Priority Gene Status**.....102

**Table 4 Stratification of CLL4 panel genes into intracellular signalling pathways**.....150

**Table 5 Breakdown of 2 mutated intracellular signalling pathways**.....152

**Table 6 CLL4 TruSeq vs. full CLL4 trial: cohort characteristics**.....173

**Table 7 All significant and powered CLL univariate OS survival analysis variables**.....198

**Table 8 All significant and powered CLL univariate PFS survival analysis variables**.....200

**Table 9 Multivariate Cox PH model for OS in CLL4**.....220

**Table 10 Multivariate Cox PH model for PFS in CLL4**.....221

# List of Figures

**Figure 1 Timeline of the history of cancer** ..... 31

**Figure 2: Hallmarks and enabling characteristics of cancer** ..... 32

**Figure 3 The mutation burden of various cancers** ..... 35

**Figure 4 Haematological malignancies in the context of B cell development** ..... 36

**Figure 5 Overview of the components of the CLL tumour microenvironment** ..... 40

**Figure 6 Overview of BCR downstream signalling pathways** ..... 50

**Figure 7 Activation of major spliceosome** ..... 55

**Figure 8 Live cell count using a Neubauer haemocytometer** ..... 65

**Figure 9 Agarose gel image** ..... 68

**Figure 10 Nanodrop DNA profile showing the concentration and absorbance ratios** ..... 69

**Figure 11 Sanger Sequencing trace** ..... 72

**Figure 12 Droplet digital PCR analysis** ..... 73

**Figure 13 Annexin V flow cytometry** ..... 82

**Figure 14 Bioinformatics pipeline of TruSeq data** ..... 88

**Figure 15 Example of for looping for univariate Cox PH modelling** ..... 91

**Figure 16 COSMIC CLL recurrently mutated gene frequencies** ..... 96

**Figure 17 Number of NGS CLL publications in PubMed separated by NGS technique** ..... 97

**Figure 18 Flow diagram of the sample sequencing process** ..... 104

**Figure 19 Mean amplicon coverage distribution with more than 100 reads** ..... 106

**Figure 20 CLL4 TruSeq Design 1 correlation coverage analysis** ..... 108

**Figure 21 Overview of sequencing coverage correlation matrix analysis** ..... 109

**Figure 22 Mean coverage per gene** ..... 111

**Figure 23 Heat map analysis of coverage data** .....113

**Figure 24 Clinical analysis of missing data in *TP53*** .....117

**Figure 25 Flow diagram output of variant filtering strategy** .....120

**Figure 26 Bland-Altman plots of minor subclone validation** .....123

**Figure 27 Analysis of concordance between Gonzalez *et al.* (2011) and TruSeq for *TP53* mutations** .....127

**Figure 28 Molecular validation of *TP53* mutations** .....128

**Figure 29 Analysis of concordance between Oscier *et al.* (2012) and TruSeq *SF3B1* mutations** 130

**Figure 30 Molecular validation of *SF3B1* mutations** .....131

**Figure 31 Analysis of concordance between Oscier et al. (2012) and TruSeq *NOTCH1* mutations** .....133

**Figure 32 Molecular validation of *NOTCH1* mutations** .....134

**Figure 33 Mutation Landscape of CLL4** .....136

**Figure 34 Frequency of CLL4 clinical trial FISH CNVs** .....137

**Figure 35 Mutation Lolliplots of *SF3B1* and *BRAF* variants** .....138

**Figure 36 Distribution of mutation type** .....140

**Figure 37 *IGHV* mutational status co-occurrence analysis** .....142

**Figure 38 Mutation Lolliplots of *KRAS, RPS15* and *XPO1*** .....143

**Figure 39 Mutated *SF3B1* co-occurrence analysis** .....145

**Figure 40 Association analysis between *SF3B1* and *POT1* mutated CLL** .....146

**Figure 41 *NBEAL2* mutations in CLL4** .....148

**Figure 42 Lolliplot of mutations found in *CTBP2* (RIBEYE) in CLL4** .....149

**Figure 43 Overview of mutated intracellular signalling pathways in CLL4** .....151

**Figure 44 Heat-maps of co-mutated intracellular signalling pathways** .....153

**Figure 45 CLL mut driver association analysis** .....156

**Figure 46 Inferred clonal evolution of CLL4** ..... 160

**Figure 47 Sensitivity Specificity analysis of *TP53, SF3B1,* and *NOTCH1* for OS** ..... 176

**Figure 48 Sensitivity Specificity analysis of TP53, SF3B1, and NOTCH1 for PFS** ..... 177

**Figure 49 Overview of survival data time point analysis strategy** ..... 180

**Figure 50 Comparison of CLL4 survival data time points for OS and PFS** ..... 182

**Figure 51 Overview of univariate OS Kaplan Meier survival analysis over time** ..... 183

**Figure 52 Overview of univariate PFS Kaplan Meier survival analysis over time** ..... 184

**Figure 53 Scaled Schoenfeld residuals for OS** ..... 187

**Figure 54 Scaled Schoenfeld residuals for PFS** ..... 188

**Figure 55 SurvSplit of OS data** ..... 190

**Figure 56 SurvSplit of PFS data** ..... 192

**Figure 57 Overview of Univariate Survival Analysis Process** ..... 196

**Figure 58 *TP53* survival analysis in the context of del(17p)** ..... 202

**Figure 59 Subclonal *TP53* mutations with impact on OS and PFS in CHL treated CLL** ..... 204

**Figure 60 Longitudinal analysis of clonal evolution in *TP53* mutated cases** ..... 206

**Figure 61 del(11q) + *BIRC3* and biallelic *ATM* survival analysis** ..... 208

**Figure 62 Overview of impact of *KRAS, NRAS,* and *BRAF* mutations on OS in CLL4** ..... 210

**Figure 63 OS survival analysis of mutually exclusive *KRAS, NRAS,* and *BRAF* mutations** ..... 212

**Figure 64 Subclonal *EGR2* mutations significantly associate with poor OS and PFS in CLL4** ..... 214

**Figure 65 Distinct impact of clonal and subclonal *SF3B1* mutations on OS by *IGHV* mutational status** ..... 216

**Figure 66 Impact of *DDX3X* mutations on PFS** ..... 218

**Figure 67 RSF model of OS in CLL4 - Minimal depth vs. VIMP** ..... 224

**Figure 68 RSF model of PFS in CLL4 - Minimal depth vs. VIMP** ..... 226

**Figure 69 Chemical Structures and targets of Splicing Inhibitors** ..... 238

**Figure 70 SSA induces apoptosis in CLL cells in a dose- and caspase-dependent manner** .....242

**Figure 71 Sensitivity of CLL cells to SSA induced apoptosis** .....244

**Figure 72 SSA treatment causes intron retention and modifies *MCL-1*$_{L/S}$ and *SF3B1* RNA splicing in Ramos cells** .....246

**Figure 73 SSA treatment reduces Mcl-1$_L$ protein expression in Ramos cells**.....248

**Figure 74 MCL-1$_L$ overexpression in Ramos cells induces resistance to SSA induced apoptosis**250

**Figure 75 SSA induces intrinsic apoptosis in Ramos cells in a dose- and time-dependent manner** .....252

**Figure 76 SSA induces cell death in Eμ-TCL1 cells *in vitro* that may be dependent on Bim expression**.....254

**Figure 77 MMB induces apoptosis in CLL cells in a dose-dependent manner** .....255

**Figure 78 The effect of immobilised αIgM stimulation on SSA induced apoptosis** .....257

**Figure 79 SSA in combination with the Bcl-2 family inhibitor venetoclax augments apoptosis in CLL cells following immobilised αIgM stimulation** .....260

**Figure 80 UMI-77 induces apoptosis in a dose- and caspase-dependent manner in CLL cells**264

**Figure 81 UMI-77 induces apoptosis in a time-dependent manner in CLL cells** .....266

**Figure 82 Sensitivity of CLL cells to UMI-77 induced apoptosis** .....267

**Figure 83 The effect of CD40L/IL4 and αIgM stimulation on UMI-77 induced apoptosis** .....269

**Figure 84 UMI-77 in combination with venetoclax differentially augments apoptosis in CLL cells following CD40L/IL-4 stimulation** .....271

**Figure 85 UMI-77 in combination with venetoclax differentially augments apoptosis in CLL cells following immobilised αIgM stimulation** .....273

# DECLARATION OF AUTHORSHIP

I, Stuart James Blakemore, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Clinical and biological implications of genomic lesions in Chronic Lymphocytic Leukaemia

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
    - Young E, Noerenberg D, Mansouri L, Ljungström V, Frick M, Sutton L-A, **Blakemore S** et al. EGR2 mutations define a new clinically aggressive subgroup of chronic lymphocytic leukemia. Leukemia [Internet]. 2016;(August):1–8. Available from: http://www.nature.com/doifinder/10.1038/leu.2016.359
    - Larrayoz M, **Blakemore SJ**, Dobson RC, Blunt MD, Rose-Zerilli MJJ, Walewska R, et al. The SF3B1 inhibitor spliceostatin A (SSA) elicits apoptosis in Chronic Lymphocytic Leukemia cells through downregulation of Mcl-1. Leukemia [Internet]. Nature Publishing Group; 2015;(April):1–28. Available from: http://www.nature.com/doifinder/10.1038/leu.2015.286
    - Parker H, Rose-Zerilli MJ, Larrayoz M, Clifford R, Edelmann J, **Blakemore S**, et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. Leukemia. 2016;30(11):2179-86.
    - Ljungstrom V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, et al. Whole expme sequencing in relapsig chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. Blood. 2016;127(8):1007–16.
    - Larrayoz M, Rose-Zerilli MJ, Kadalayil L, Parker H, **Blakemore S**, Forster J, et al. Non-coding NOTCH1 mutations in chronic lymphocytic leukemia; their clinical impact in the UK CLL4 trial. Leukemia. 2017;31(2):510-4.
    - Pandzic T, Larsson J, He L, Kundu S, Ban K, Akhtar-Ali M, et al. Transposon Mutagenesis Reveals Fludarabine Resistance Mechanisms in Chronic Lymphocytic Leukemia. Clin Cancer Res [Internet]. 2016;22(24):6217–27. Available from: http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-15-2903

Signed:

Date: .....

# Acknowledgements

I would like to thank my supervisory panel Professor Jon Strefford, Dr Andy Steele and Professor Mark Cragg for all their help and advice throughout my project, but especially in the last months before submission. An additional thank you goes to Mark for giving me the opportunity before I started my PhD to volunteer in his lab, as that really was the springboard I needed to get into research, and I am eternally grateful. This brings me on to the 3 Ms, Dr Matt Carter, Dr Mat Rose-Zerilli and Dr Marta Larrayoz Ilundain. You have all supported me whenever I have needed it throughout this project, and without that support I would not be where I am today. Special mentions should also go to everyone in the Cancer Genomics group (especially Helen!), as well as my national and international collaborators via the CLL4 TruSeq project (especially Dr Ruth Clifford!). For those of my collaborators I haven't met yet, I will endeavour to do so, coming to a conference near you...

Then comes the family… The amount of encouragement and support I receive from you all is incredible, and you put up with me talking science at you. I promise I do try to explain things in a way that you understand, which usually involves peppermills and placemats to represent various parts of genes.

Suze and Andy: Allowing me to use your house as a place of refuge and relaxation has been incredible. Thank you!

Caz and Grant: Your words of wisdom at times of stress (yes, you too Grant!) have always come at a perfect time, and supported me in ways only Caz (being a CBT therapist) would know!

Mum: Coming to Southampton whenever we can fit in time for each other and just chatting about all the things I would love to chat about with you on a more regular basis. It has kept me sane! Oh, and the washing!

Dad and Sarah: Providing me the support to do as I needed to when I was at my lowest during this project was everything that you ever needed to do for me. Thank you so much.

Nora: Wir haben diesem Tag für fast vier Jahren gewartet... Danke für das Sein eine unglaubliche Freundin, du bist die beste! Nächste Halt, DEUTSCHLAND!

Friends: Those of you who are also studying for a PhD/have finished, thanks for not talking about it so much when we hang out. The tchoukball lot, thanks for reminding me at every possible moment that I am still a student and should go and get a real job.

# Definitions and Abbreviations

| Abbreviation | Meaning |
|---|---|
| .csv | Comma-separated values |
| .fcs | Flow Cytometry Standard |
| .tiff | Tagged image file format |
| .txt | Text file |
| A | Adenine |
| ALL | Acute lymphoblastic leukaemia |
| ANOVA | Analysis of variance |
| APC | Allophycocyanin |
| AV | Annexin V |
| B-cell | B lymphocyte |
| BAM | Binary alignment/map |
| BCR | B-Cell Receptor |
| BED | Browsable extensable data |
| BH3 | BCL2 homology 3 |
| BLAST | Basic Local Alignment Search Tool |
| BR | Broad range |
| BTK | Bruton Tyrosine Kinase |
| C | Cytosine |
| CC | Cell cycle |
| CCF | Cancer cell fraction |
| CD40L | CD40 ligand |

| | |
|---|---|
| cDNA | Complimentary deoxynucleic acid |
| CE-SCAA | Capillary Electrophoresis Single-Stranded Conformation Analysis |
| CHL | chlorambucil |
| Clk | Cdc2-like kinase |
| CLL | Chronic Lymphocytic Leukaemia |
| CNVs | Copy number variants |
| COSMIC | Catalogue of somatic mutations in cancer |
| CR | Complete response |
| CR/NodPR | Complete response/nodular partial response |
| CRC | Colorectal cancer |
| ddPCR | Droplet Digital Polymerase Chain Reaction |
| DDR | DNA damage response |
| DMEM | Dulbecco's Modified Eagle Medium |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxnucleic acid |
| DOR | Diagnostic odds ratio |
| DP | Read depth |
| dsDNA | Double stranded DNA |
| DTT | Dithiothreitol |
| ECL | Enhanced Chemiluminescent substrate |
| EDTA | Ethylenediaminetetraacetic acid |
| ERIC | European research initiative in CLL |
| ESE | Exonic splicing enhancers |

| | |
|---|---|
| EtOH | Ethanol |
| F | fludarabine |
| FACS | Flow Activated Cell Sorting |
| FAM | Caroboxyflourescein |
| FC | fludarabine + cyclophosphamide |
| FCS | Fetal Calf Serum |
| FDR | False discovery rate |
| FIR | FUSE-binding protein interacting repressor |
| FISH | Fluorescence in situ Hybridisation |
| FITC | Fluorescein isothiocyanate |
| FNR | False negative rate |
| FPR | False positive rate |
| FPU | Filter plate assembly unit |
| FUSE | Far upstream element |
| G | Guanine |
| GEF | Guanine exchange factor |
| GPS | Grey platelet syndrome |
| HBSS | Hank's Balanced Salt Solution |
| HEX | Hexachloro-fluorescein |
| HR | Hazard ratio |
| HRM | High resolution melt |
| HSC | Haematopoietic Stem Cell |
| HTS | High-throughput-screening |

| | |
|---|---|
| IARC | International agency for research on cancer |
| IC50 | Inhibitory concentration 50 |
| ID | Identification number |
| IGHV | Immunoglobulin Heavy Chain Variable Region |
| IgM | Immunoglobulin M |
| IGV | Integrated genomics viewer |
| IL4 | Interleukin-4 |
| IWCLL2008 | International Workshop on Chronic Lymphocytic Leukaemia 2008 |
| LOVD | Leiden open variation database |
| LR | Likelihood ratios |
| M-CLL | Mutated CLL |
| mAb | Monoclonal antibody |
| Mb | Megabase |
| MBL | Monoclonal B-cell Lymphocytosis |
| MDR | Minimally deleted region |
| MMB | Meayamycin B |
| MOPS | 3-(N-morpholino)propanesulfonic acid |
| NaOH | Sodium hydroxide |
| NCI | National Cancer Institute |
| NGS | Next Generation Sequencing |
| NHL | Non-Hodgkin Lymphoma |
| NodPR | Nodular partial response |
| NR/PD | No response/progressive disease |

| | |
|---|---|
| NTC | No template control |
| OB | Oligonucleotide/oligosaccharide-binding |
| OOB | Out of bag rate |
| OS | Overall survival |
| PARP | Poly (ADP Ribose) Polymerase |
| PBMCs | Peripheral Blood Mononuclear Cells |
| PCR | Polymerase Chain Reaction |
| PDX | Patient derived xenografts |
| PFS | Progression-free survival |
| PH | Proportional hazards |
| PI | Propidium Iodide |
| PI3K | Phosphatidylinositol-3-Kinase |
| PolyPhen | Polymorphism phenotyping |
| QVD | Q-VD-OPh |
| RBH | Royal Bournemouth hospital |
| RF | Random forests |
| RNA | Ribodeoxynucleic acid |
| RS | Richter's syndrome |
| RSF | Random survival forests |
| RT-PCR | Reverse transcription Polymerase Chain Reaction |
| SEM | Standard error of the mean |
| SGH | Southampton general hospital |
| SIFT | Sorting intolerant from tolerant |

| | |
|---|---|
| siRNA | Short interfering ribonucleic acid |
| SLL | Small Lymphocytic Lymphoma |
| SMI | Small molecule inhibitors |
| SNP | Single nucleotide polymorphisms |
| SNV | Single Nucleotide Variant |
| SPSS | Statistical package for the social sciences |
| SR | Serine/arginine |
| SSA | Spliceostatin A |
| T | Thymine |
| TAE | Tris base, acetic acid, and EDTA |
| TAM | Tumour Associated Macrophage |
| TBS | Tris-buffered saline |
| TNR | True negative rate |
| TPR | True positive rate |
| TTFT | Time-to-first treatment |
| U-CLL | Unmutated CLL |
| UCSC | University of California, Santa Cruz |
| UMI | Unique molecular identifier |
| UV | Ultra-violet |
| VAF | Variant allele frequency |
| VCF | Variant call format |
| VIMP | Variable importance |
| WCL | Whole cell lysate |

| | |
|---|---|
| WES | Whole exome sequencing |
| WGA | Whole Genome Amplification |
| WGS | Whole genome sequencing |
| WHO | World Health Organisation |
| WT | Wild type |
| ZVAD | Z-VAD.fmk |

# Chapter 1: Introduction and Literature Review

## 1.1 Introduction

Cancer causes millions of deaths each year worldwide (1), presenting fundamental challenges to national health services. In the United Kingdom, approximately fifty percent of cancer patients succumb to their disease annually (2), therefore identifying innovative ways to improve patient outcome through basic and translational research is critical to provide affected individuals with a prolonged, healthy life.

Accounting for one percent of all cancer diagnoses and around forty percent of all leukaemias in the United Kingdom annually (3), with concordant incidence data across the developed world, Chronic Lymphocytic Leukaemia (CLL) is a mature B lymphocyte (B-cell) malignancy largely affecting the aged population (4). Characterised by its remarkable clinical heterogeneity, CLL is defined by its homology to the immunoglobulin heavy-chain variable (IGHV) locus, molecular cytogenetics, as well as its immunophenotype and expression of intra- and extracellular markers (5). A plethora of genomic studies, conducted largely in institutional cohorts, have identified recurrently mutated genes that cluster into distinct biological signalling pathways (6). Mutated *TP53, ATM, SF3B1,* and *NOTCH1* have repeatedly been identified and associated with a poor prognosis across studies, with the subclonal architecture and clonal evolution of these genes in the context of therapy beginning to be elucidated (7–9). Furthermore, putative candidate genes have been unearthed (*FBXW7, EGR2, NFKBIE, CHD2, RPS15,* and *SETD2,* amongst others), each with variable supportive evidence towards their prognostic relevance, and identified in aggressive CLL subsets (10–15). Despite these advances, only the screening of *TP53* mutations has been incorporated into standard clinical practice, therefore further evidence that screening CLL patients for these recurrently mutated genes is informative, is required. Furthermore, although targeted therapies have become a part of the frontline clinical arsenal, the inclusion of Small Molecule Inhibitors (SMIs) that target these mutated genes is withstanding.

This thesis presents the mutational landscape of patients enrolled onto the UK LRF CLL4 clinical trial (CLL4) at a subclonal resolution, using a 25-gene targeted Next Generation Sequencing (NGS) panel. The clinical implications of these genetic events in this cohort have been explored, using traditional survival modelling as well as machine learning survival tools, confirming amongst other findings, the clinical significance of mutated *SF3B1* on prognosis in CLL. Finally, this thesis characterises the targeted inhibition of SF3B1 in CLL cells *in vitro*, presenting its mechanism of action via abrogated *MCL-1* splicing, and comparing splicing inhibition with direct targeting of MCL-1 using a BCL-2 Homology-3 (BH3)-mimetic.

The 25-gene panel consists of genes previously identified to be recurrently mutated in CLL and/or CLL subsets, as well as prospective genes which have only been shown to be mutated in CLL at low frequencies. To understand the impact that these genes may have on disease pathogenesis and therapy in CLL, an overview of cancer and the role of genotoxic stress in cancer will be discussed. Following an introduction into haematological malignancies, the aetiology and clinical management of CLL is discussed, outlining the importance of the tumour microenvironment alongside the genomic landscape. In addition, the types of treatment available to CLL patients is described, identifying why the validation of prospective and prognostic markers is still required, and that although the therapeutic options available have improved survival rates, the discovery of new drugs will improve patient management and lead to a more personalised therapy for each patient. Furthermore, the associated intracellular signalling pathways implicated in CLL pathogenesis will be discussed. Finally, the modern technologies which capture this genomic information will be presented, to explain the potential effect these mutations are having at a cellular level.

## 1.2 Literature Review

### 1.2.1 Cancer

The basis of modern cancer theory begins with the German pathologist Rudolf Virchow and the French bacteriologist Louis Pasteur, who both postulated in the 19th Century that cancer was caused by a "breakdown of order within cells" and by the "invasion of a foreign organism", respectively (16). Since these early observations, of which both can be attributed to different types of cancer, research scientists have defined cancer as having specific cellular processes that the malignant cells either hijack or turn off in the setting of genetic aberrations and the tumour microenvironment. As shown in **Figure 1**, critical events in cancer research, patient treatment and scientific technological advances over the last two hundred years have vastly improved the options available to cancer patients.

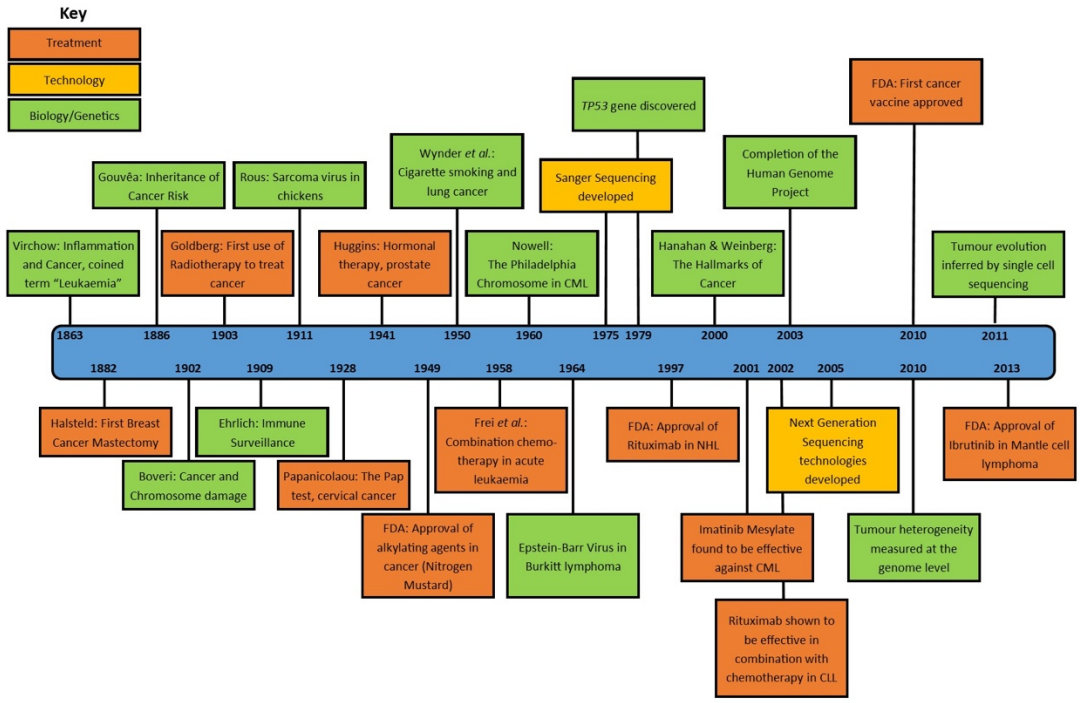



**Figure 1 Timeline of the history of cancer**

Timeline of important technological, clinical and scientific advances in relation to cancer over the last two hundred years. Timeline created by the author using information from online resources from the American Society of Clinical Oncology (251) and the American Cancer Society (252).

### 1.2.2 Hallmarks of cancer

The biology of cancer can be characterised as a process of disrupted homeostasis, with the underpinning mechanisms about which these regular biological processes are disrupted being complex. These mechanisms are termed the 'hallmarks of cancer' as coined by Hanahan and Weinberg (17,18), describing how the sustenance of proliferative signalling, evasion of growth suppressors, activation of invasion and metastasis, facilitation of replicative immortality, induction of angiogenesis and resistance of cell death all contribute to carcinogenesis. Furthermore, Hanahan and Weinberg (18) have described potential additional hallmarks with the deregulation of cellular energetics and avoidance of immune destruction, which when considered altogether in the context of the 'enabling characteristics' of genomic aberrations and tumour-promoting inflammation, continues to define the complexity of cancer (17,18). It is these genomic aberrations (caused by genotoxic stress) and their consequences on gene function and patient prognosis in CLL that is the focus of this thesis.

**Figure 2: Hallmarks and enabling characteristics of cancer**

The established and emerging hallmarks of cancer as reviewed by Hanahan & Weinberg (2011) (18). Hallmarks are displayed around the circumference of the circle. Arrows indicate these hallmarks are important for tumour cell development (grey cells – different cell shapes indicate the cellular heterogeneity between cancer cells) and growth.

### 1.2.3 Genotoxic stress and cancer

In cancer, germline and somatic genetic events can contribute to carcinogenesis. Germline events (such as the mutations in *BRCA1/2* genes in breast cancer), are an example of how perturbed function of a protein over a lifetime can increase an individual's risk of disease. However, somatic genetic events occur as a dysfunction of DNA replication during the process of cell division, or due to DNA damage caused by exogenous and endogenous sources (19). These two processes can lead to changes at the chromosome level as well as down to a single nucleotide change.

#### 1.2.3.1 Sources of genotoxic stress

Exogenous sources of DNA damage have been defined by the World Health Organisation (WHO), which fall into specific subgroups; lifestyle risk factors (tobacco, dietary, alcohol consumption), infection (chronic hepatitis, human papilloma virus), environmental pollution (drinking water, indoor/ambient air), occupational carcinogens, radiation (X-rays, ultra-violet [UV]) (20). Endogenous sources of DNA damage are produced as by-products of metabolic processes, and occur at a higher rate than exogenous DNA damage. Endogenous sources include; oxidative DNA damage (reactive oxygen species [ROS], lipid peroxidation, products derived from DNA oxidation (base propenols), oestrogen metabolites, alkylating agents (*S*-adenosylmethionine), DNA hydrolysis, hydrolytic deamination, carbonyl stress (from lipid peroxidation and glycation) (21).

#### 1.2.3.2 Impact of genotoxic stress on the genome

These sources of genotoxic stress induce DNA damage, however the result of the damage if unresolved by the DNA damage response (see section **1.2.6.1**) can be widespread. Somatically acquired copy number changes involve the: deletion, duplication, reciprocal translocation, inversion, and insertion of large portions of genetic material. This can lead to the loss or gain of gene function, which could be sufficient to modulate the homeostasis of the affected cell, allowing abrogation of the associated signalling pathway. These types of genetic events were the first to be described in cancer, since they can be observed via microscopic karyotyping at the chromosome level.

Additionally, smaller changes of the genome can occur, namely insertions or deletions (known as indels), and single nucleotide variants (SNVs), commonly referred to as mutations. The nomenclature surrounding somatic mutations is associated with several factors; whether the mutations lead to a change in the type of nucleotide base in the sequence, if the nucleotide change infers an amino acid change and whether the mutation causes the early inclusion of a stop codon. Transition mutations, those where the nucleotide base change is purine to purine (A>G) or pyrimidine to pyrimidine (C>T), are the least likely candidates to contribute to genomic instability,

as most these mutations do not lead to an amino acid change. Transversion mutations, on the other hand, are formed by purine to pyrimidine (A>C, A>T, G>C, G>T, C>A, C>G, T>A, T>G) or vice versa changes, largely resulting in missense or nonsense mutations. Missense and nonsense mutations are examples of nonsynonymous mutations, with the former describing a mutation that leads to an amino acid change and the latter describing a mutation that leads to the early formation of a stop codon. Synonymous mutations are those that do not lead to an amino acid change, however that does not mean that this mutational class does not have any effect on biology, as these mutations have been shown to cause aberrant splicing events (22).

Furthermore, these events can lead to loss of heterozygosity (LOH), which infers the loss of function in tumour suppressors, essential members of signalling pathways which regulate the appropriate transcription of downstream target genes (23). LOH can result in deleterious point mutations, those that result in a stop codon or non-synonymous mutations which prevent correct protein folding, and chromosomal deletions, where swathes of genes in a specific chromosomal segment are not translated into protein because the DNA template has been lost. This can lead to a loss of intracellular signalling control, potentiating continuous activation of the pathway. Proto-oncogenes are powerful signalling pathway members, which when mutated leading to prevention of deactivation or after a translocation event which leads to continuous transcription, have the capacity to drive carcinogenesis (24).

Cancers are not affected by mutation processes (exogenous or endogenous sources) in the same way, with each cancer having a unique set of signatures (25). These mutational signatures were found to be associated with a variety of features, such as cytidine deaminases (APOBEC), which function to maintain the pyrimidine pool during DNA replication, and tend to be pan-cancer, as well as exogenous sources such as tobacco, which were found to be cancer specific. These signatures also outline the distribution of mutation burden across cancer. The signature is also impacted by the mutation burden, defined as the number of somatic mutations found per megabase of the genome. The study of the genome at base pair resolution could not be possible without the recent advances in sequencing technologies, therefore an overview of this is presented in **1.2.7**.

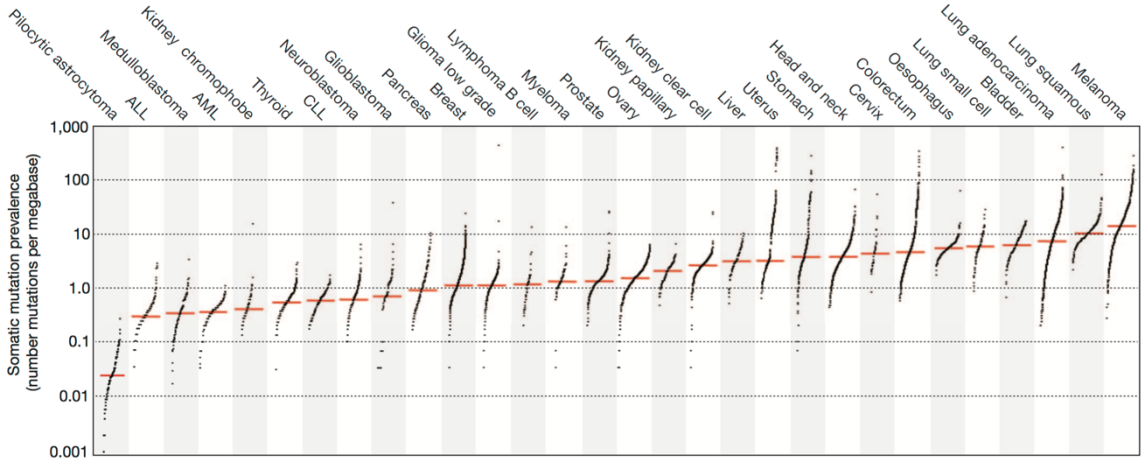



**Figure 3 The mutation burden of various cancers**

Graph representing the somatic mutation burden per megabase across a variety of cancers. Taken from Alexandrov *et al.* (2013) (25)

### 1.2.4 Haematological malignancies

Virchow is also attributed with the first identification of leukaemia, a group of haematological malignancies that are derived from the myeloid and lymphoid lineages originating from a haematopoietic stem cell (HSC) compartment. Haematological malignancies are stratified by the WHO and others into four subgroups; leukaemia, Hodgkin lymphoma, Non-Hodgkin lymphoma and myeloma (26). Leukaemia is the malignant generation of circulating leukocytes, with disease sub-types presenting as B cell, T cell or myeloid cell malignancies which can be acute or chronic in nature, largely dependent on the differentiation state of the affected founder cell (27). Lymphomas are tumours of the lymphatic system and associated organs, and myeloma develops from plasma cells.

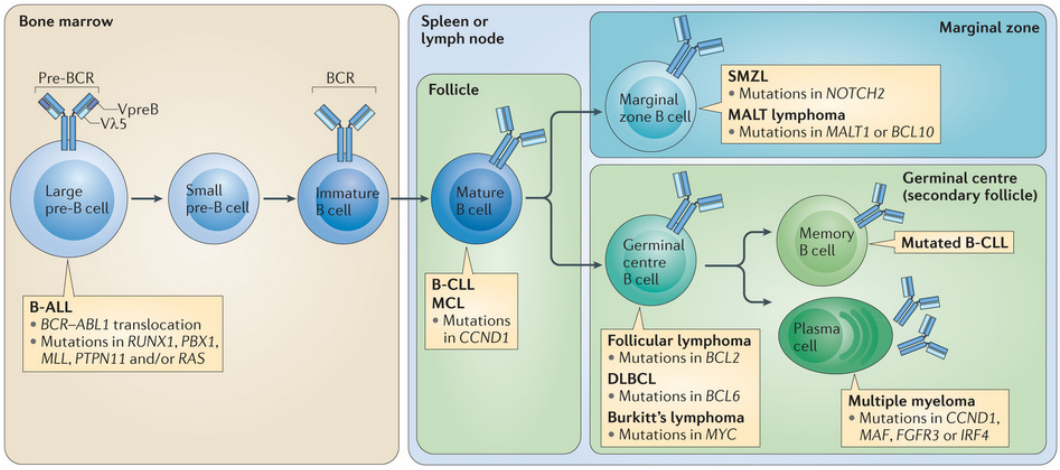



**Figure 4 Haematological malignancies in the context of B cell development**

B cell malignancies arise from various stages of B cell development, with characteristic genetic aberrations. Figure taken from Rickert *et al.* (2013) (253).

### 1.2.5 Chronic Lymphocytic Leukaemia

CLL is a Cluster of Differentiation 19 (CD19), CD23 and CD5 positive B cell malignancy (28) without a defining chromosomal aberration (29), instead it is a myriad of chromosomal abnormalities (30), recurrent mutations (31–33) and B cell specific characteristics (34–37) that infer prognostic clinical relevance (28,38). Forming a sub group of Non-Hodgkin Lymphoma (NHL) (39) and simultaneously a leukaemia, CLL shares a natural history with Monoclonal B-cell Lymphocytosis (MBL) (40), Small Lymphocytic Lymphoma (SLL) (38) and Richter's Syndrome (RS) (41).

#### 1.2.5.1 CLL genetics

CLL was first observed to carry trisomy events of chromosome 12, through traditional karyotyping experiments (42). Since then molecular genetic techniques, with ever increasing sensitivity, have been used to interrogate the genomes of CLL patients. Introduction of fluorescence *in situ* hybridisation (FISH) (43) provided a more high-throughput approach to analyse CNVs (through increased automation of analysis), also culminating in their clinical application. Definition of the minimally deleted regions (MDR) of CNVs in CLL through these techniques, allowed for the identification of genes which could be responsible for CLL pathogenesis via FISH and sequencing experiments. This led to the detection of *TP53* within del(17p) events, *ATM* in del(11q) events, as well as *DLEU1* and *DLEU2* in del(13q) events. These four CNVs represent the most recurrent in CLL, and have all been purported to be initiation events in CLL pathogenesis.

##### 1.2.5.1.1 del(17p)

Deletions in the short arm of chromosome 17 were already known to lead to a poor prognosis in CLL prior to sequencing, but this new resolution offered the opportunity to assess the impact of mutations on a specific gene. It was quickly observed, that mutations in the DNA-binding and activation domains of *TP53* were deleterious, with reduced p53 activity, as well as being found to reduce the outcome in CLL patients independently of del(17p) events. Furthermore, it was observed that mutations in *TP53* often co-occur with del(17p) events (>80%) (44), further supporting the importance of *TP53* in CLL pathogenesis, since LOH is a typical event in carcinogenesis. Del(17p) events occur in 5-8% of patients prior to treatment, and show marked resistance to immuno-/chemotherapeutic regimens (45). Mutations in *TP53* occur in 4-37% of CLL patients, and also associate with poor prognosis, even in the absence of del(17p) events (46). Furthermore, mutations in *TP53* associate with increased mutation burden.

#### 1.2.5.1.2 del(11q)

Deletions of the long arm of chromosome 11 can be found in approximately 25% of chemotherapy-naïve patients with advanced disease stage, and 10% of patients with early stage disease (32,47). These deletions frequently encompass band 11q23 harbouring the gene *ATM*, which encodes for the proximal DNA damage response kinase ATM. A subset of approximately 40% of patients carrying a del(11q) event display inactivating mutations of the second *ATM* allele and these cases show a poor chemotherapy response, like what has been described for *TP53*-aberrated CLL (48). In addition, patients carrying del(11q) events typically show rapid progression, and reduced overall survival (30). Furthermore, mutations and deletions in *BIRC3* have been reported in the context of del(11q) (49), however their importance has not been fully elucidated in a direct comparison with *ATM.* As for *TP53*, disabling *ATM* mutations are enriched in chemotherapy-treated patients, suggesting that an inactivation of the pro-apoptotic DDR is selected for in CLL (32,47). In addition, germline mutations in *ATM* have also been observed to impact response to chemotherapy.

#### 1.2.5.1.3 del(13q)

Deletions on the long arm of chromosome 13, specifically involving band 13q14 is the single most frequently observed cytogenetic aberration in CLL, occurring in approximately half of all cases. A del(13q14) event identified in a patient without other CNVs is typically characterised by a benign disease course. In addition to the genes *DLEU1* and DLEU2, the miRNAs miR-15a and 16-1 have been identified to be in the MDR of del(13q) (50). The pathophysiological role of these miRNAs is further underscored by the phenotype of genetically engineered mice carrying a targeted deletion of the mir-15a/16-1 locus in combination with a deletion of the non-coding RNA gene *DLEU2*. These animals develop a monoclonal B-cell lymphocytosis-like disorder, CLL and lymphoma, suggesting that the miRNAs 15a and 16-1 indeed play a role in CLL tumorigenesis (51).

#### 1.2.5.1.4 +(12)

Lastly, trisomy 12 is observed in 10-20% of CLL patients. However, the genes involved in the pathogenesis of CLLs carrying a trisomy 12 are largely unknown. Furthermore, prognostic relevance of trisomy 12 remains a matter of debate (44).

#### 1.2.5.2 CLL biology

##### 1.2.5.2.1 IGHV mutational status

In addition to CNVs, the B-cell maturation stage has been observed to be important in CLL biology, pathogenesis, and outcome. The impact on outcome inferred by maturation stage was first observed by Freda Stevenson in Southampton, where CLL patients were found to carry either mutated immunoglobulin variable heavy chain (IGHV) sequences (lack of homology to the germline sequence, referring to memory B-cells), or unmutated IGHV sequences (high homology to the germline sequence, referring to mature B-cells yet to differentiate into marginal zone or germinal centre B-cells). When stratified by these criteria, it was observed that patients with unmutated IGHV sequences (U-CLL) had a significantly inferior survival to mutated IGHV (M-CLL) patients (35). This led to intensive research into the IGHV sequences of CLL patients, further identifying that there is substantial stereotypy in a proportion of the CLL population (37), suggesting that antigen and environmental factors might have a role in CLL pathogenesis. These stereotypes have been further stratified into BCR subsets, of which certain subsets have been associated with further reduced survival, and aggressive disease features.

#### 1.2.5.3 CLL tumour microenvironment

The CLL tumour microenvironment is a collection of immune cells within the secondary lymphoid organs of CLL patients that provide proliferative support to the cycling CLL cells. These proliferation centres (or pseudofollicles) are composed of: Nurse-like cells (NLCs), T lymphocytes (T-cells) and Natural Killer (NK) cells, endothelial cells, bone marrow stromal cells (BMSCs), as well as follicular dendritic cells (FDCs) (52). These different cell types all contribute to the release of various ligands, chemokines, and cytokines, which influence the homing capacity, cellular motility, activation, and proliferation of CLL cells and the tumour microenvironment itself (52). The CLL tumour microenvironment has been implicated as a feature of disease pathogenesis, with overexpression of anti-apoptotic proteins a key feature of treatment resistance in the secondary lymphoid organs of CLL patients (53). Two of the most important supportive features of the CLL microenvironment are from T cells and nurse-like cells/M2-like tumour associated macrophages (TAM). Secretion of CD40L and IL-4 from T cells, which activates NF-κB signalling. Antigen is purported to be carried by nurse-like cells, which leads to BCR receptor activation. BCR activation leads to a complex downstream cascade, which can increase proliferation through MAPK/ERK signalling.

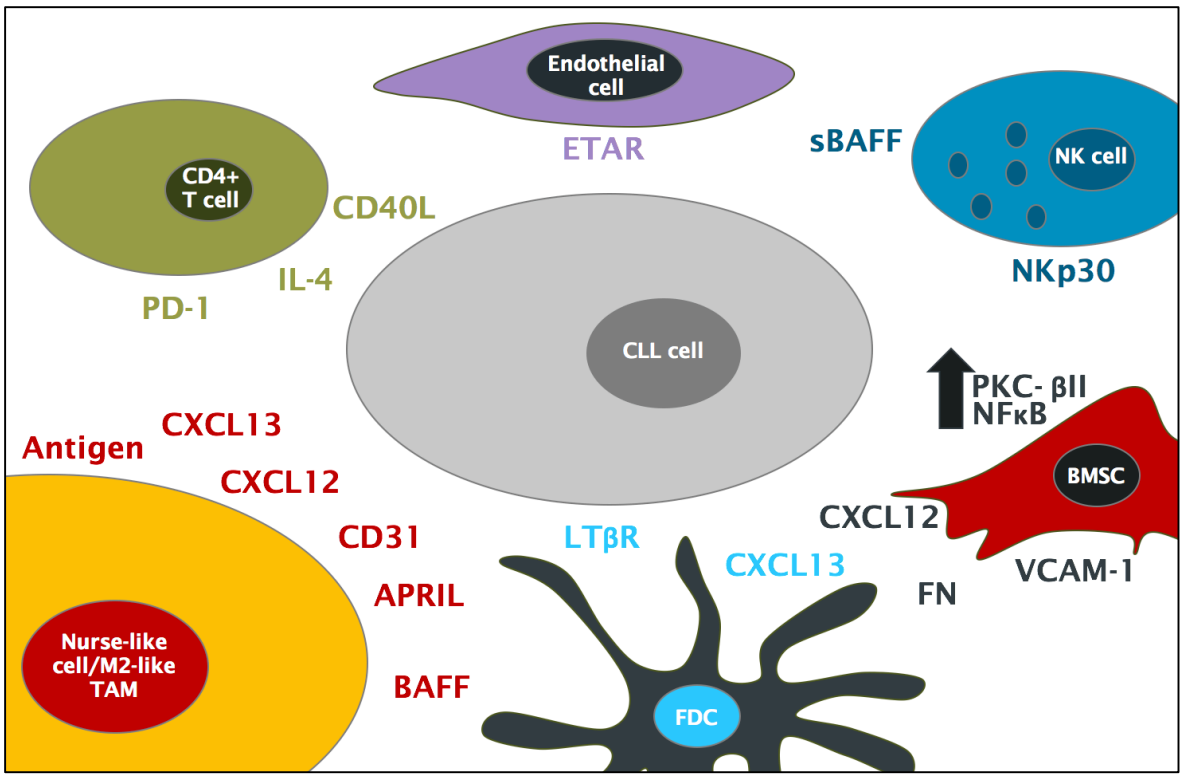



**Figure 5 Overview of the components of the CLL tumour microenvironment**

Representation of the cell types of the CLL tumour microenvironment. Figure information reproduced using ten Hacken and Burger (2015) as source material (52).

#### 1.2.5.4 Diagnosis of CLL

The diagnosis of CLL is often a fortuitous event, usually identified after a routine blood test showing that the patient has a lymphocytosis. On attendance to a CLL clinic, a further blood test is taken to assess the clonality and immunophenotype of the disease, diagnosing patients as having CLL if they have a population of clonal B cells surpassing 5000/μL of peripheral blood (28). Furthermore, they are palpated for organomegaly (which can be confirmed via a CT scan) and lymphadenopathy. Histological assessment of the leukaemic cells can also be carried out to check for the presence of smudge cells (CLL cells with low levels of the cytoskeletal protein Vimentin (54)). The immunophenotype screen varies dependent on the diagnostic lab, however the co-expression of the T-cell marker CD5 and the B-cell markers CD19, CD20 and CD23 is required (28).

Diagnosis of SLL is attributed to patients who do not reach the lymphocytosis threshold needed to diagnose CLL, but do present with splenomegaly or lymphadenopathy (55). However, there are clinical issues surrounding the potential 'over-diagnosis' of SLL, as there are currently no criteria for the diagnosis of SLL from biopsied lymph nodes extracted due to other medical conditions (55). The diagnosis of MBL is given when patients, who are by all other means healthy, have a clonal B cell population that like in the diagnosis of SLL is not greater than 5000 cells/μL of peripheral blood (56). MBL patients are stratified into three main groups depending on their clonal B cell population immunophenotype. CLL-like MBL as the title suggests shares its immunophenotype with CLL, with patients of this subgroup progressing to CLL at around 1% per annum, and makes up most MBL diagnoses. The remaining 20% of cases are either CD5+ or CD5- respectively, and have characteristics more related to marginal zone lymphomas than CLL (56). In a combined WES/WGS study on CLL and MBL from the Spanish Genome Consortium, MBL was shown to have a comparable mutation burden to CLL(57). Finally, clinical MBL (or high count MBL) cases seem to be more likely to progress to CLL than low count MBL(58).

#### 1.2.5.5 Clinical Staging of CLL

Once diagnosis has been ascertained, clinical staging can be implemented using the accumulation of basic diagnostic features based on two widely accepted methods, the Rai (59) and Binet clinical staging systems. These systems are a financially accessible, quick but crude way of prognostically stratifying patients. Rai clinical staging was initially stratified into five patient groups but has been reduced to three that now encompass low-risk, intermediate-risk and high-risk disease. Lymphocytosis without the presence of organomegaly defines low-risk disease, with the addition of organomegaly contributing to intermediate-risk disease. High-risk disease includes the aforementioned characteristics but also the inclusion of thrombocytopenia or disease related anaemia, with a platelet threshold of $<1x10^8$/μL or haemoglobin level of <110g/L, respectively

(28). Binet clinical staging is also constructed from three discrete disease levels termed Stage A, B and C. Unlike the Rai system, the Binet system considers organomegaly at all levels, however it defines the stage based on the number of affected sites and the inclusion of anaemic and thrombocytopenic blood counts, and this system does not consider lymphocytosis. Stage A and B disease have a platelet threshold of $\geq 1x10^8/\mu L$ or haemoglobin level of ≥110g/L, with Stage A patients having a maximum of two sites of organomegaly and Stage B patients having three or more. Stage C disease requires a platelet threshold of $<1x10^8/\mu L$ or haemoglobin level of <110g/L much like Rai high-risk disease, however this stage is attributed to patients regardless of their organomegaly status (28).

#### 1.2.5.6 Prognostication of CLL

Unlike other haematological malignancies, CLL patients do not share a single mutation or chromosomal aberration which defines the disease (29). Instead, an array of abnormalities which stratify the disease into diverse clinically relevant subsets presents itself. Patients with better prognostic markers and those under the "watch and wait" clinical mantra have long term overall survival (OS), rarely requiring therapy with CLL not contributing to their eventual cause of death. Meanwhile, those with poor prognostic markers are commonly found to relapse and have refractory disease. These chromosomal abnormalities have been the cornerstone of prognostication in CLL since 2000 when Döhner and colleagues (30) stratified patients for OS with del(17p), del(13q), del(11q) and trisomy 12 aberrations using FISH. The poorest prognosis comprises the del(17p) locus where the *TP53* gene is also located, which considering the role of p53 in cancer (60) provides a rationale for why these CLL patients have an aggressive disease and respond poorly to standard chemotherapy. CLL patients with only a del(13q) abnormality infer the best prognosis, with a superior OS in comparison to patients with a normal karyotype (30). The chromosomal aberrations with an intermediate prognosis in regards to OS are del(11q), which is decreased when found concomitantly with *ATM* mutations (61) (termed biallelic *ATM*), impairing the DNA damage response in these patients (48), and trisomy 12 (+(12)), which are associated with *NOTCH1* mutations and upregulation of integrin signalling (62).

In the following years, understanding around treatment regimens for these distinct groups along with other prognostic markers (such as CD38 (36), ZAP70 (34), Immunoglobulin Variable Heavy chain Gene (IGHV) mutational status (35)) has been sought. IGHV mutational status has been an integral prognostic marker in CLL since its discovery in 1999 (35). Unmutated IGHV sequences have very few changes to the germline sequence, suggesting that they have not experienced SHM, whilst mutated IGHV sequences are not homologous to the germline sequence. Mutated IGHV patients have a generally better prognosis whilst unmutated IGHV genes is associated with

more aggressive disease and poorer response to therapy (34), although the genomic complexity of CLL can confound these associations. Expression of cellular markers such as CD38 and ZAP-70 have both been implicated with aggressive disease (34,36), however these have yet to be shown to play important functional roles in CLL.

Certain CLL patients who have been diagnosed and their prognostic track identified via the aforementioned methods are further confronted with the prospect that their disease could transform into RS, a disease similar to Diffuse Large B Cell Lymphoma (DLBCL). The transformation frequency varies from 2% to more than 15% depending on the characteristics of the cohort studied and the length of time between sequential sampling (63). These patients have a dismal outcome, therefore identifying characteristics which associate with risk of transformation could improve clinical management.

#### 1.2.5.7 Treatment of CLL

The clinical management of CLL only currently offers one curative treatment, allogenic stem cell transplantation, however the success of this treatment requires young medically fit recipients, a combination which is relatively rare in CLL. Therefore, chemotherapeutic agents, whether as a monotherapy or a part of a combinatorial regimen have pervaded in CLL treatment strategies for decades. Consisting of:

- Alkylating agents, DNA cross-linkers (64) – chlorambucil, cyclophosphamide
- Purine analogues, compounds which are incorporated into DNA and disrupt DNA repair (65). – fludarabine, pentostatin

The alkylating agent chlorambucil (CHL) is an exceptional example of continued clinical utility in CLL with its moderate toxicity and oral availability, providing a cheap accessible option to treatment for unfit or infirm and elderly patients (66). Along with alkylating agents, purine analogues have been a staple treatment modality in CLL, with fludarabine (F) as a stand out drug which continues to be used in the clinic, also with its combination with cyclophosphamide (FC).

The UK LRF CLL4 clinical trial was a phase III randomised comparison of CHL, F, and FC, which was approved by the Medical Research and Ethics Committee (MREC) in 1998. In total, 777 previously untreated patients were enrolled onto the first randomisation, with the primary trial endpoint being Overall Survival (OS), and secondary endpoints as response rate, progression-free survival (PFS), toxic effects and quality of life (67). The second randomisation was formed of two arms, either local clinician guided treatment, or Differential Staining Cytotoxicity (DiSC) assay guided treatment (*in vitro* patient cell drug sensitivity assay), organised by the Royal United Hospital in Bath (68). Although no significant difference was observed for the primary endpoint, FC provided

the best overall response (94%), followed by fludarabine (80%) and then chlorambucil (72%). FC also performed best for PFS, with significantly improved results in comparison to fludarabine alone or chlorambucil (FC=36% 5 year PFS vs. 10% for fludarabine and chlorambucil). Furthermore FC was most beneficial for all age groups and prognostic groups (*IGHV* mutation status and cytogenetic sub-groups), however patients treated with FC or fludarabine alone spent more days in hospital and had more neutropenia than in comparison to chlorambucil treated patients (67).

To date there have been over 20 publications on the LRF CLL4 clinical trial, with studies associating recurrent gene mutations in *SF3B1* and *NOTCH1* with a poorer prognosis in comparison to patients who are wild type for these genes (7), and the role of *ATM* mutations on the prediction of survival of 11q deleted CLL patients (61). CLL4 has long OS and PFS follow up, with more than ten years of OS and 9 years of PFS data. Although the treatment modalities used in this trial are no longer routinely used, CLL4 offers a well characterised cohort of CLL patients. Considering a large proportion of patients are too fragile for frontline chemotherapy combined with monoclonal antibodies or monotherapy kinase inhibitors, further genetic characterisation of patients who respond well to chlorambucil could identify a subgroup of patients which may not require these expensive treatments. Furthermore, the availability of established biomarker data combined with putative genomic biomarkers using TruSeq makes the LRF CLL4 clinical trial an informative cohort in CLL for prognostic outcome analysis.

These chemotherapeutic regimens were the frontline approach to CLL, until the introduction of monoclonal antibodies (mAbs). Rituximab (anti-CD20 mAb), in combination with FC (FCR), was the first immuno-chemotherapy to be introduced in the clinical management of medically fit CLL patients (69), since it not only increased the PFS of CLL patients, but also their OS (70). Rituximab in combination with chemotherapy was initially implemented in B cell lymphomas, where impressive response rates were observed. CD20 expression in CLL is generally low, so originally it was not considered as an important candidate for CLL therapy. However, clinical trial phase I and II studies identified that higher doses in combination with chemotherapeutics were synergistic, leading to the phase III clinical trial which led to the roll out of FCR as the frontline regimen in CLL. The CLL8 trial compared FC to FCR in untreated CLL, finding increased benefit with the addition of rituximab to the FC regimen (69). Furthermore, in a mutation screening of *TP53, SF3B1,* and *NOTCH1* from a subset of patients from the CLL8 trial*, NOTCH1* mutated cases were identified to receive the added benefit from rituximab. FCR has led to the identification of a subset of patients which are “functionally cured”, patients with M-CLL and indolent disease features. However, patients with aggressive disease (such as *TP53*ab) do not respond well to FCR, therefore other options targeting CLL biology are required for these patients.

Another mAb used in the clinical management of CLL is alemtuzumab, which binds to CD52, a molecule with unknown function which is expressed on lymphoid cells, but not haematopoietic stem cells. Alemtuzumab is largely used as a monotherapy in the relapse/refractory setting in CLL, where it is efficacious in *TP53* mutated/deleted cases (71). Furthermore, combinations with FC and alemtuzumab (FCA) have been attempted in CLL, but did not improve response in comparison to FCR.

Although the introduction of mAbs, either as mono-therapy or in combination with chemotherapeutics, have improved not only response rates but also OS, CLL remains incurable in most patients. More recently, the introduction of SMIs based on the biology of CLL has been undertaken, with impressive response rates. This has been epitomised by the success of the Bruton Kinase (BTK) inhibitor ibrutinib and the Phosphatidylinositol-3-Kinase (PI3K) inhibitor idelalisib, which both target molecules downstream of the B Cell Receptor (BCR). Ibrutinib has been particularly efficacious in older patients as a monotherapy (72), however they require continuous treatment to prevent relapse. Furthermore, idelalisib was found in a phase II clinical trial to provide added benefit in relapse/refractory cases in combination with rituximab (73). Both ibrutinib and idelalisib have been licenced for second line use in patients. Besides BCR inhibition, several other SMIs have been identified based on CLL biological features, most pertinently the BCL-2 inhibitor venetoclax. Venetoclax has been observed to be efficacious in patients who have already received frontline therapy, especially in del(17p) cases (74). Recently, a phase II clinical trial has been started to assess the combination of venetoclax and ibrutinib. The trial, sponsored by the M.D. Anderson Cancer Centre, will be completed in 2024 and will assess whether this combination is efficacious in the relapse/refractory setting, as well as in treatment naïve patients with high risk genetic characteristics.

### 1.2.6 Disruption of intracellular signalling pathways in CLL

Discretely clustered into biologically significant signalling pathways, pathogenic events such as point mutations, chromosomal aberrations and overexpression of genes in CLL impinge on the normal healthy function of the relevant pathway, leading to disrupted homeostasis. Targeting critical pathways associated with survival, CLL associated abnormalities have been identified in DNA Damage Response (DDR), cell cycle control, RNA metabolism, chromatin modification, BCR signalling, Wnt signalling, Notch signalling and NF-κB signalling(31,75). Summarised below are the functions of these diverse pathways, contextualising their importance and function in normal biology and CLL pathology.

#### 1.2.6.1 DNA Damage Response Pathway

The DDR pathways are induced by internal and external cellular damage that cause single stranded breaks (SSBs) and double stranded breaks (DSBs). These pathways are responsible for the faithful repair of the DNA or the targeting of the cell for destruction. SSBs are caused by direct stimuli such as Reactive Oxygen Species (ROS) or indirectly by the error prone Base Excision Repair (BER) pathway for example, where either the five prime or three prime strand of the DNA has been damaged(76). DSBs are the result of two SSBs within proximity of each other leading to segregation of the DNA on both strands, caused by unresolved SSBs and external clastogenic sources.

The SSB repair pathway (SSBR) is initiated from direct stimuli by a Poly ADP-Ribose Polymerase (PARP) family member, with SSBs induced by BER sensed by the Polβ polymerase, a BER pathway member (76). After damage binding, end processing of the directly and indirectly induced break are conducted by protein complexes that both contain AP endonuclease (APE1). However, the component members of the protein complexes required for gap filling and ligation of the SSB differ, with direct breaks requiring Proliferating Cell Nuclear Antigen (PCNA) and indirect breaks requiring X-Ray Repair Cross-Complementing 1 (XRCC1), amongst others (76).

Unlike SSBs, DSBs are sensed and signals transduced using a comparable set of molecules, following which their repair is either conducted via Non-Homologous End Joining (NHEJ) or Homologous Recombination (HR). The MRE11/RAD50/NBS1 complex (Meiotic Recombination 11, RAD50, Nijmegen Breakage Syndrome 1) is the primary sensor of DSBs, and is responsible for activation of the DSB transducer Ataxia Telangiectasia Mutated (ATM) (77). Phosphorylation of downstream effectors by ATM leads to a plethora of cellular responses, from NHEJ and HR to cell cycle arrest, transcriptional regulation and apoptosis (78).

The application of NHEJ and HR in response to DSBs is dependent on the cell cycle position where the damage occurred (NHEJ: $G_0/G_1$, HR: $S/G_2$)(79). NHEJ is an error prone DNA damage repair mechanism in which the broken three prime and five prime ends are cleaned up and ligated following which polymerisation occurs without a complementary strand, commonly leading to inaccurate repair (80). HR requires a complementary strand from a sister chromatid, with both ends of the break modified to have a three-prime overhang, forming a D-loop. Subsequently, Synthesis Strand-Dependent Annealing (SSDA) or DSB Repair (DSBR) pathways can be implemented to repair the DSB. The SSDA pathway unloops the D-loop allowing adherence to ssDNA after DNA synthesis, whilst the DSBR pathway instigates DNA synthesis, ligases and in some instances Holliday junctions (a branched nucleic acid structure containing four double-stranded arms) with specific endonucleases with Non-crossover and Crossover events available (81).

#### 1.2.6.1.1 DDR pathway in CLL

Disruption of *TP53* and *ATM,* key components of the DDR and disrupted either by mutation or chromosomal deletion in CLL, constitutes two of the high risk patient groups, with poor OS and complications in response to therapy (48,46). Furthermore, mutations in genes; *POT1*, *CHK1, BRCA1* and *CHD2* have been observed, confirming the DDR as an important pathway implicated in CLL pathogenesis (32,47). *SF3B1* mutated patients without concurrent *TP53/ATM* mutation also appear to have impaired DDR activity, with these 'sole' *SF3B1* mutants showing impaired p53/ATM transcriptional activity and apoptosis in response to DNA damaging agents (82).

### 1.2.6.2 Apoptosis

The apoptotic programme is responsible for the execution of death inducing signals, received after stimulation from a variety of factors. Characterised by the intrinsic and extrinsic pathways (83), apoptosis is a homeostatic function which in normal B-cell biology is important for the selective deletion of inappropriate clones during Somatic Hypermutation (SHM). Intrinsic apoptosis is centralised around the mitochondria, B-Cell Lymphoma 2 (BCL2) family proteins and Caspases. BH3 only proteins such as BCL2 Interacting Mediator of cell death (Bim) and Noxa activate BCL2-associated X protein (Bax) and BCL2-antagonist/killer (Bak), key pro-apoptotic molecules which induce Mitochondrial Outer Membrane Permeabilisation (MOMP), leading to the release of Cytochrome *c*, formation of the apoptosome with Apoptotic Peptidase Activating Factor 1 (APAF1) and inactive Caspase 9, followed by Caspase dimerization and activation and subsequent activation of the effector Caspases 3 and 7 (84). This pathway is regulated by BCL2 family anti-apoptotic proteins such as BCL2, $BCL_{XL}$, MCL1 and A1 which inhibit Bax and Bak. Extrinsic apoptosis signals via death receptors such as CD95 which sequesters Fas-Associated Death Domain (FADD) and Caspase 8 molecules upon ligand binding, leading to activation of Bid

(BH3-only protein BH3-interacting domain death agonist) and MOMP following a pathway homologous to intrinsic apoptosis (85).

#### 1.2.6.2.1 Apoptosis pathway in CLL

The apoptotic machinery is implicated in the biology and targeted treatment of CLL. Overexpression of BCL2 is associated with poor response to traditional therapeutic options and a reduction in OS (86,87). Furthermore, differential protein expression of BH3 only protein Noxa and pro-survival protein MCL1 in the lymph nodes and periphery of CLL patients affects the response of these CLL cells to proteasome inhibition *in vitro* (88). Targeted treatment of members of the apoptotic machinery in CLL include BCL2 inhibitors and BH3 mimetics, which have both been shown to be efficacious in *in vitro* and *in vivo* studies of B-cell malignancies, with promising results in the clinical trials setting (89,90). However, apoptosis genes are seldom mutated in CLL.

### 1.2.6.3 Cell Cycle Control

Culminating in the faithful duplication of the entire genome and cellular organelles from one cell to two daughter cells, the cell cycle is the process which prepares every cell to divide. Highly regulated in conjunction epigenetic modifications (91) and telomeres (92), the cell cycle is divided into four distinct phases, of which progression from one phase to the next is controlled by Cyclins and Cyclin Dependent Kinases (CDKs) (93). Cyclins are activatory molecules which bind CDKs to form Maturation-Promoting Factors (MPFs). Once bound, this perpetuates the transcription of cell cycle phase specific genes through phosphorylation and subsequent activation of transcription factors. This leads to the expression of key components which facilitate progression through the next cell cycle checkpoint, simultaneously instigating a feedback signal to prevent inappropriate progression signals when they are no longer required. Each phase of the cell cycle ($G_1$, S, $G_2$, Mitosis, Cytokinesis, [$G_0$]) requires a unique combination of Cyclins and CDKs to perform their function, however the borders of these phases and cell cycle checkpoints are also critical for genotoxic stress, with implications in the DDR and apoptotic pathways (93).

#### 1.2.6.3.1 Cell cycle control pathway in CLL

Targeting cell cycle regulators to treat CLL may be advantageous, especially considering the difficulty of clearing CLL cells from the CLL microenvironment niche, the proliferative centres (94). Metformin is a candidate drug in CLL, inducing apoptosis and inhibiting cell cycle activation when co-cultured with stimulation factors such as CD40 ligand *in vitro* (95). In addition, *KRAS, NRAS,* and *BRAF* have all been found to be mutated in CLL at low levels (57,96). Furthermore, CDK inhibitors (CDKi) have been shown to be efficacious in CLL, with flavopiridol showing promising results in

phase II clinical trials (97), and pre-clinical evidence that CDKI-73 is synergistic with fludarabine *in vitro* (98).

#### 1.2.6.4 B-cell Receptor (BCR) Signalling

The BCR is a multi-pathway, membrane immunoglobulin receptor that binds to antigen, leading to downstream activation of Akt, PI3K, NF-κB, JAK/STAT and ERK signalling, amongst others (99). This antigen binding leads to phosphorylation of the BCR's co-receptors CD79a and CD79b, recruiting multiple kinases and adapter proteins forming the BCR signalosome (99). The signalosome is constructed of the receptor itself, tyrosine kinases such as Syk or Lyn, transmembrane adapter proteins such as CD19, and downstream signalling enzymes such as PI3K. The signalosome functions to propagate downstream signalling as well as to amplify the antigen activated BCR signal through the formation of complexes with other membrane bound molecules (100), leading to the aggregation of many signalosomes, called lipid rafts (101). Downstream BCR signalling pathways are implicated in proliferation, survival, differentiation, apoptosis, and anergy, with these downstream cascades also being modulated based on B cell differentiation maturation status, the longevity of the signal, the antigen that originally induced the signal, and additional signalling from other receptors sharing the same intracellular signalling pathway members.

##### 1.2.6.4.1 BCR signalling in CLL

BCR signalling is abrogated in various ways in CLL, which makes affected members of the BCR signalling pathway so attractive to target clinically. Survival stratification of patients by IGHV gene somatic hyper-mutation exemplifies the importance of BCR signalling, as IGHV encodes the IgM immunoglobulin receptor that forms the BCR (35). Overexpression of kinases associated with BCR activation, such as Syk and Lyn, has been identified in CLL, which propagates amplification of downstream signalling (102,103). Furthermore, amplification of members of these downstream signalling pathways has also been reported, further inferring the importance of the BCR in CLL pathogenesis. Indeed, inhibition of these molecules has shown clinical efficacy, with GS-1101 (PI3K inhibitor) (104), ibrutinib (105) and fostamatinib (Syk inhibitor) (106) all showing promise in the clinical trials setting.

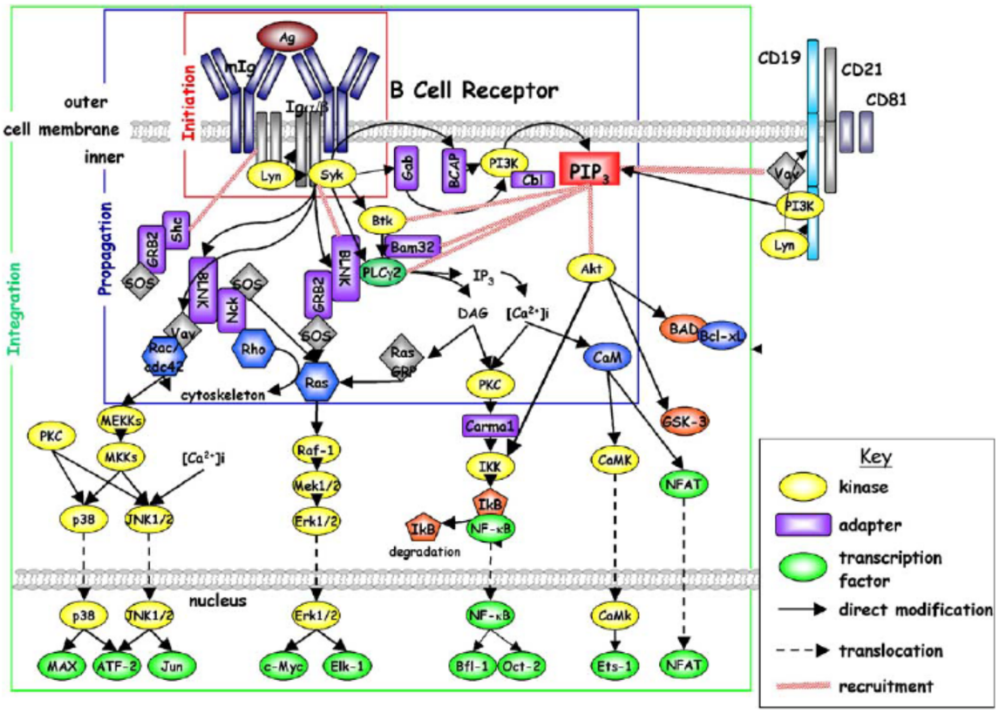



**Figure 6 Overview of BCR downstream signalling pathways**

B-cell receptor intracellular signalling cascades. Yellow proteins represent kinases, purple adapter molecules, and green transcription factors. Figure taken from Dal Porto *et al.* (2004).

### 1.2.6.5 Notch Signalling

Notch is a developmental intracellular signalling pathway and an example of 'juxtacrine signalling' (contact dependant signalling), which is implicated in prenatal and postnatal cell fate decisions in many cell lineages(107,108). The role of NOTCH in B-cell development is considered to be diverse, ranging from direct implication in HSC self-renewal (109), osteoblast cell regulation of HSCs via the Jagged1 NOTCH ligand (110), to complete redundancy of NOTCH in the bone marrow HSC niche (111). NOTCH signalling works through the presentation of the receptor to the cell membrane after processing steps in the golgi called S1. After ligand binding, the extracellular portion of the NOTCH receptor is cleaved by ADAM metalloproteases (S2), followed by cleavage by the γ-secretase presenilin complex, this classifies the S3 stage. Completion of S3 releases the NOTCH Intracellular Domain (NICD), which then either translocates to the nucleus to perform CBF/Suppressor of Hairless/Lag2 (CSL)-dependent transcription of target genes with the coactivator Mastermind, or CSL-independent actions that affect other signalling pathways such as Wnt signalling to effect cell assignation (112–115).

#### 1.2.6.5.1 Notch signalling in CLL

Multiple members of the NOTCH1 and NOTCH2 signalling pathway have been shown to be abrogated in CLL. Overexpression of NOTCH2 has been hypothesised to be responsible for the

overexpression of the key CLL biomarker CD23 (116). Mutations in the E3 ligase *FBXW7* have also been identified (10), potentially mimicking what is seen in T-cell Acute Lymphoblastic Leukaemia (T-ALL) where constitutive activation of NOTCH signalling is achieved in T-ALL patients who carry a mutation in this gene (117). However, the most recurrent and clinically important NOTCH signalling gene abrogated in CLL is *NOTCH1*. Largely observed as indel or non-coding mutation events (57), *NOTCH1* mutations associate with poor survival and can co-occur with trisomy 12 in CLL patients (7,118). Targeting the NOTCH signalling pathway has been shown *in vitro* to be advantageous, with γ-secretase inhibitors inducing apoptosis and resulting in the downregulation of expression of non-canonical NF-κB pathway members associated with cell death (119).

#### 1.2.6.6 Wnt Signalling

The canonical Wnt signalling pathway is a major regulator of intracellular β-Catenin, a molecule which is essential for propagating Wnt ligand responses, through the T-Cell Factor/Lymphoid Enhancer Factor (TCF/LEF) transcription factors (120). β-Catenin activity is tightly controlled in canonical Wnt signalling via the destruction complex, a group of proteins of which one of the members has serine/threonine kinase activity (Glycogen Synthase Kinase β3 [GSKβ3]), which phosphorylates β-Catenin targeting it for destruction by the proteasome (121). In the archetypal model of Wnt, β-Catenin is only free to accumulate and translocate into the nucleus in complex with TCF/LEF in the presence of Wnt ligand – Frizzled-LRP5/6 receptor binding (122) , which sequests Dishevelled (Dsh), Axin and GSKβ3 to the membrane stopping formation of these destruction complex members with Adenomatous Polyposis Coli (APC) and Casein Kinase 1α (CK1α), amongst others. Once in a complex with TCF/LEF and in coalition with nuclear Dsh, β-Catenin coactivates the transcription of target genes (123). In the absence of Wnt ligand the destruction complex can form allowing the GSKβ3's serine/threonine kinase activity to phosphorylate β-Catenin which labels it for destruction via the Ubiquitin Proteasome System (UPS) (124), allowing the transcriptional repressor Groucho to bind to TCF/LEF, repressing the transcription of target genes.

##### 1.2.6.6.1 Wnt signalling in CLL

Overexpression of Wnt pathway members has been identified in CLL in comparison to normal B cells, suggesting that this pathway is implicated in CLL pathogenesis (125). Furthermore, mutations in Wnt pathway members have been identified in CLL, however not at a frequency per member to be statistically significant, although grouped analysis suggested the opposite, with higher mutation frequencies observed (31). Inhibition of Wnt signalling using small molecule inhibitors has shown efficacy in pre-clinical testing, with compounds inducing apoptosis *in vitro* and *in vivo* (126).

#### 1.2.6.7 NF-κB signalling

NF-κB inflammatory signalling pathway is associated with the immune system and development. NF-κB is a transcription factor, which forms as a dimer to translocate to the nucleus to transcribe target genes. The NF-κB signalling pathway is downstream of single transmembrane receptors such as the BCR, T-Cell Receptor (TCR), Toll-like Receptors (TLRs), and Tumour Necrosis Factor Receptors (TNFRs) amongst others, with two major subgroups; canonical and non-canonical signalling (127). Canonical NF-κB signalling occurs downstream of many of the previously mentioned receptor families and requires the activation of NEMO, a regulatory protein which binds with Inhibitor of κB Kinase (IKK) α and β, leading to phosphorylation of Inhibitor of κB (IκB) and culminating in the activation of the NF-κB transcription factor. Non-canonical NF-κB signalling occurs in the context of specific TNF family members and is exclusive of NEMO.

##### 1.2.6.7.1 NF-κB signalling in CLL

In CLL NF-κB activity has been shown to be constitutively high and induced via the non-canonical pathway and stimulation via CD40 (128). Mutations of pathway members have also been identified in CLL, with *BIRC3* and *EGR2* recurrently mutated in specific disease sub-groups, predicted to be poor prognostic markers for these patients (12,49).

#### 1.2.6.8 RNA splicing and metabolism

The process of propagating gene to protein expression is achieved via protein synthesis, which involves multiple types of biological machinery, initiated from complex signalling pathways upstream in response to specific stimuli which orchestrates the selection of context dependent genes to be expressed. These machines are large, multi-component conglomerates which are dynamic in action, allowing a single gene transcript to be used to generate numerous protein isoforms (129). Protein synthesis begins with transcription to create a nascent RNA transcript, following which this transcript is matured through the removal intronic sequences termed RNA splicing, and then further processed before being transported out of the nucleus to be translated into protein.

The spliceosome is constructed from small nuclear ribonucleic proteins (snRNPs), which are formed themselves from specific sets of proteins, phosphatases and RNA molecules that are required to complete their designed function within the RNA splicing pathway (129,130). In mammalian systems, there are two types of spliceosome, the major spliceosome which conducts 99.5% of all splicing events and the minor spliceosome which is required for the remaining 0.5%

(129,131). The major spliceosome contains the U1, U2, U4, U5 and U6 snRNP subunits, whilst the minor spliceosome is constructed of the U11, U12 U4acac, U5 and U6acac snRNPs (129,132). The minor spliceosome shares the U5 snRNP with the major spliceosome, as well as specific proteins within other snRNPs such as SF3B1, an essential splicing factor in the U2 snRNP which is also found in the U12 snRNP (133). Assembly of the spliceosome occurs *de novo* on each intron of the pre-mRNA concomitantly whilst the residual pre-mRNA is being transcribed (129), therefore it is important to note that although RNA splicing is a stepwise process to remove an intron, it does not necessarily remove the introns chronologically intron by intron.

Five major complexes are formed from these snRNPs throughout the RNA splicing pathway, occurring in a stepwise manner: Prespliceosome (complex A), Precatalytic spliceosome (complex B), Activated spliceosome (complex B*), Catalytic step 1 spliceosome (complex C) and Post-spliceosomal complex (129). The Prespliceosome is composed of the U1 and U2 snRNP, which historically has been represented as being recruited to the pre-mRNA intron in a stepwise manner (U1 followed by U2). Shcherbakova *et al.*(2013) (130) however have shown using elegant studies in yeast that on a panel of intron containing yeast pre-mRNAs U1 and U2 can bind both before or after each other, and that this does not affect their ability to splice, as measured by colocalisation of the U4, U5, U5 tri-snRNP, which denotes the formation of the Precatalytic spliceosome. Depending on the size of the intron, the Prespliceosome can be formed via cross-intron and cross-exon complexes, the latter of which is called Exon definition (134), and has recently been shown to contain the U4, U5, U6 tri-snRNP and can be activated into a Precatalytic like complex (135).

After complex B formation, removal of the U1 and U4 snRNPs is necessary to generate complex B*, the Activated spliceosome which can then complete the first of two transesterification steps needed to remove the intron. This is achieved through 5’ splice site (5’ SS) adenosine residue attack, leading to the Catalytic step 1 spliceosome via phosphatase activation. 3’ splice site (3’ SS) adenosine residue attack can then ensue, resulting in the Post-spliceosomal complex which now contains the excised intron formed as a Lariat structure and a ligated exon pair. This complex can then be broken down to release the snRNPs and the Lariat structure completing the RNA splicing cycle on that intron(129).

SF3B1 (SAP155 or SF3B155) is a splicing factor which forms part of the SF3b subunit in the U2snRNP, binding the branch point sequence (BPS), an essential process for spliceosome assembly(133). The *SF3B1* gene resides on chromosome 2q33.1, forming a 155kDa protein with three isoforms. The SF3B1 protein is constructed of an N-terminal domain which has binding specificities to other splicing factors (U2AF65 (136), p14 (137), Cyclin E (138), NIPP1 (139)) and a C-terminal domain containing 22 Huntingtin, Elongation Factor 3 (EF3), regulatory A subunit

(65kDa) of Protein Phosphatase 2A (PP2A), Target of Rapamycin 1 (TOR1) (HEAT)-like repeats, which shows 75% identity between *Schizosaccharomyces pombe* and humans and acts to stabilise the U2snRNP by wrapping around it (140). Consolidating its importance in splicing further than binding the BPS is its phosphorylation during the Catalytic step 1 spliceosome stage, which is necessary to conduct the second transesterification step to excise the Lariat formed exon (141).

#### 1.2.6.8.1 RNA splicing in CLL

In CLL multiple genes involved in RNA splicing and transport have been observed as being abrogated (31,57,47). Mutations in *SF3B1* are the most common within these pathways, with this gene mutated in other cancers as well as being a marker of poor prognosis in CLL (7). Exclusively identified in the HEAT-like repeats, mutations of *SF3B1* are often sub-clonal and associated with BCR stereotypy (142), poor prognosis (7), sub-clonality (75), and with a potential role in AS in CLL (143). Located to exons 14, 15 and 16, the array of point mutations leading to a synonymous amino acid change have yet to be assigned individual functional significance, even though the K700E mutation on exon 15 has been shown to be the most recurrent (7). Recent work to elucidate its potential effect on disease has been conducted in uveal malignant melanoma (144) with a small additional study conducted in CLL (143). They report that *SF3B1* mutated CLL patients cluster together in their RNA analysis for the AS of particular genes (143). What makes this more interesting is that in uveal melanoma and myeloid dysplastic syndromes (MDS) *SF3B1* mutations infer a good prognosis, therefore it provides further questions as to how *SF3B1* mutations lead to a poor prognosis in CLL even though they share not just the same mutations, but their effects on particular genes via aberrant AS. Targeting SF3B1 using SMIs has been shown to be clinically relevant in vitro (145), however it's *in vivo* efficacy is yet to be fully ascertained for use in CLL.

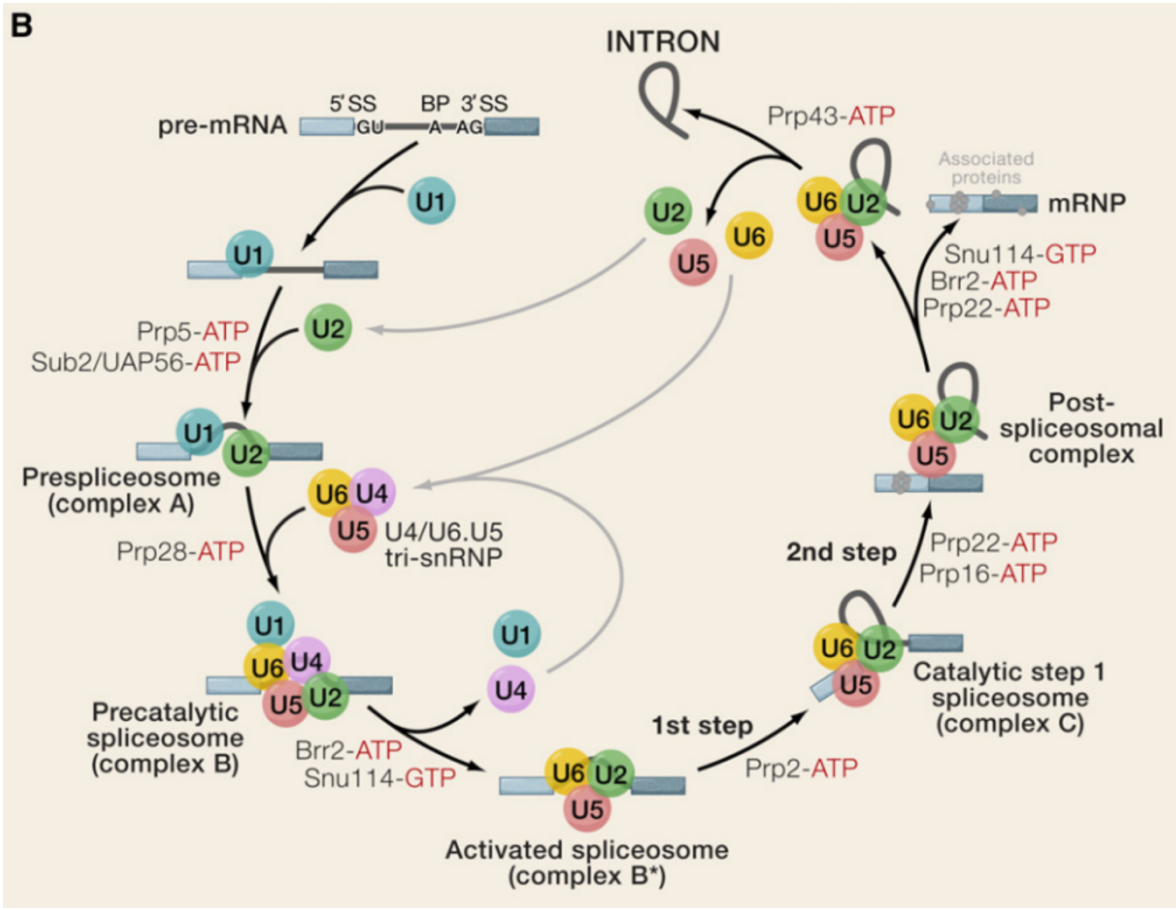



**Figure 7 Activation of major spliceosome**

Removal of introns by the major spliceosome. Formation of the complexes A, B, B*, C, the post-spliceosomal complex, and the released intron shown above, with the inclusion of the relevant snRNPs. Figure taken from Wahl *et al.* (2009).

#### 1.2.6.9 Chromatin remodelling

Chromatin is a protein and DNA conglomerate which condenses DNA and is the core component structure of chromosomes. Two types of chromatin exist, heterochromatin which is highly condensed and euchromatin which is less condensed and in this state can be transcribed. Chromatin is formed by solenoids, a string of nucleosomes, which are 147bp segments of DNA that have been wrapped around a histone octamer (146). Histones are positively charged DNA associated proteins which contribute to higher order DNA structure. The histone octamer contains H2A, H2B, H3, and H4 histones in duplicate, and these are subject to remodelling by various enzymes which change the structural state of the nucleosome, namely; acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, deamination and proline isomerisation (147). Furthermore, these modifications have been shown to be important in many cellular contexts, affecting the ability of specific genes to be transcribed (146,147).

##### 1.2.6.9.1 Chromatin remodelling in CLL

In CLL, abrogation of chromatin remodelling has been observed through the mutation and deletion of genes. An example of gene mutation in CLL is *CHD2* (15), a helicase binding protein which when studied *in vivo* using heterozygous mutation inferred a shorter lifespan largely due to formation of a lymphoid malignancy. Deletion events in *SETD2* (57), a H3K36 trimethyltransferase which have been previously described in other types of leukaemia (148) have also been observed in CLL and its role in CLL pathogenesis has been explored by our laboratory, where *SETD2* aberration was observed as an early loss-of-function event linked to aggressive disease (13).

### 1.2.7 Next Generation Sequencing

Also termed second-generation sequencing, NGS is a high-throughput sequencing approach which sequences millions of DNA fragments in a parallel fashion (149). The methodologies differ between NGS approaches; comprising of pyrosequencing, sequencing by synthesis and sequencing by ligation (150,151). These methodologies all follow core processes to prepare DNA for NGS; shearing or digestion of DNA into appropriately sized fragments, editing of DNA fragment ends to ligate platform specific adaptors, and attachment of DNA fragment-adaptor pairs to indices (complementary sequences which attach the DNA fragment-adaptor pairs to a solid surface) (150–152).

Massively parallel sequencing by synthesis (SBS) involves the hybridisation of fragment DNA molecules to a flow cell bound oligonucleotide (oligo), via complementary base pairing. After which, polymerisation of the fragment strand and bridging of this strand to another flow cell bound oligo occurs, leading to further polymerisation, culminating in the production of forward and reverse reads incorporated with the two types of flow cell bound oligos, termed bridge amplification. This process is repeated in a massively parallel fashion for all the fragment molecules in the sequencing reaction, forming clusters of reads on the flow cell. These oligo-bound forward and reverse reads are then subjected to fluorescently tagged dNTPs, which competitively bind to the fragment molecule template, emitting their fluorescence once bound on excitation from a light source. SBS chemistries are typically sequenced on Roche 454 and Illumina sequencers (HiSEQ and MiSEQ). Other types of SBS sequencers that function by emitting hydrogen ions rather than by fluorescence excitation have also been designed (Ion Torrent, Thermo Fisher).

Pyrosequencing is a type of DNA sequencing by synthesis, with incorporation of nucleotides during polymerisation and the release of pyrophosphate (PPi) leading to the emittance of light after PPi conversion to ATP (153). Pyrosequencing experiments are typically sequenced on the Genome Sequencer (FLX/Roche).

Sequencing by ligation (SBL) in a massively parallel fashion requires the hybridisation of fragment DNA molecules to a bead, which is then covalently bound to a flow cell. Using a DNA ligase with multiple di-base probes which are fluorescently tagged, template strands are read and synthesised multiple times, with the ligase shifting the start position by one base upon each round of sequencing. SBL is conducted largely using the Applied Biosystems SOLiD sequencer. These NGS techniques can be applied to genome wide as well as targeted approaches, the latter of which and the various types on offer are presented below.

#### 1.2.7.1 Targeted re-sequencing technologies

Targeted re-sequencing NGS platforms are utilised to answer genetic hypotheses formulated from previous WGS/WES gene discovery experiments. The major advantage to this approach is that it is possible to sequence at a significantly higher read depth than WES and WGS experiments, increasing the confidence in mutation calling and the opportunity to identify subclonal mutations. Furthermore, due to the considerable difference in the amount of captured genomic data in comparison to WGS/WES, targeted re-sequencing is cheaper and is more simplistic in terms of bioinformatic processing. Targeted re-sequencing platforms come in a variety of types dependent on the required downstream application, of which the major types will be discussed.

#### 1.2.7.2 Illumina Nextera

Available on two sequencing platforms (Roche 454 and Illumina), Nextera is a rapid, low DNA input NGS platform, with samples ready for sequencing two hours after beginning library preparation (152). The preparation of DNA ready for sequencing centres around a hyperactive transposase, which catalyses the integration of excised transposons into DNA targets in a near random fashion. These Transposome$^{TM}$ complexes are then fragmented and tagged simultaneously, generating a tagged DNA fragment library in a single reaction (152).

#### 1.2.7.3 Agilent Haloplex

Haloplex generated DNA libraries can be sequenced by either Illumina or Ion Torrent technologies. DNA fragmentation is achieved via 16 restriction enzymes, after which the DNA fragments are hybridised to oligonucleotide probes, purified and ligated into circularised DNA. Once prepared, these circular DNA targets are amplified by PCR which is then taken forward for sequencing (154).

#### 1.2.7.4 Illumina TruSeq

TruSeq is a PCR based targeted resequencing technology offering a maximum of 1536 PCR amplicons, with the potential of capturing up to 650kb of genomic sequence. TruSeq panel performance depends on the amplicon design as well as the number of patients sequenced per run. Samples are identified via unique combinations of barcoded indices which are tagged to the primers. Data is generated by SBS.

TruSeq custom amplicon libraries are generated by stepwise hybridisation of oligonucleotides, subsequent removal of unbound oligonucleotides, extension-ligation of the remaining bound oligonucleotides and PCR amplification. After PCR amplification, these are normalised to 4nM ready for library pooling and MiSEQ sequencing.

#### 1.2.7.5 Thermo Fisher Ion Torrent Ion AmpliSeq

Ion Torrent™ semi-conductor sequencing uses a multi-million well chip, which individually measure the release of hydrogen ions on incorporation of a nucleotide to a DNA base on a massively parallel scale. Library preparation involves the fragmentation of DNA and random attachment to individual beads, which are washed across the chip, leading to an individual bead in each well.

## 1.3 Aims

This thesis aims to contributing in depth understanding to the following hypotheses:

1. Application of a custom targeted re-sequencing panel to the LRF UK CLL4 clinical trial identifies mutations at subclonal resolution, and offers insight into the mutational landscape and clonal architecture of the disease
2. That traditional methods of survival analysis of the LRF UK CLL4 TruSeq mutation dataset, as well as application of machine learning survival tools, identify survival associations, including at clonal and subclonal level, and offer evidence into how these variables predict outcome in the context of CLL4 clinical and biological biomarkers
3. That using the genomic landscape data and survival data to select a target gene for *in vitro* inhibition, a suitable SMI can be sourced allowing for the identification of the mechanism of action in CLL cells, with the context of clinical questions pertaining to treatment resistance observed in CLL patients

# Chapter 2: Methodology

The following sections contain all the methods conducted with associated data in Chapters 3-5. The methods are stratified into two sections, laboratory (including rudimentary analysis where appropriate) and statistical/analytical.

## 2.1 Laboratory methods

### 2.1.1 Patient cohorts

Patient cohorts varied depending on the study, and were detailed in each results chapter. In summary, patient samples were obtained from;

1. The UK CLL4 Leukaemia Research Fund clinical trial. The trial was conducted between 1999 and 2004, and randomly assigning patients to first-line treatment with chlorambucil (CHL), fludarabine (F) or fludarabine plus cyclophosphamide (FC). Samples were taken prior to the initiation of therapy. All patients were diagnosed using standard morphologic and immunophenotypic criteria. Informed consent was obtained from all patients in accordance with the Helsinki declaration, and this study was approved by national or regional research ethics committees. Although the CLL4 clinical trial assessed traditional chemotherapy, a plethora of clinical data has been collected, over a decade of follow up, and allowing important associations to genomic features to be made.
2. Previously untreated CLL patients recruited in the lymphoproliferative disorder study at time of initial evaluation at the Department of Haematology of the Southampton University Hospital Trust from January 2001 to May 2015. Diagnosis of CLL was according to the 2008 International Workshop on Chronic Lymphocytic Leukaemia (IWCLL2008)/National Cancer Institute (NCI) criteria and confirmed by a flow cytometry "Matutes score" in all cases. The lymphoproliferative disorder study was approved by the institutional review boards at the University of Southampton (228/02/t). All patients provided informed consent prior to inclusion in the study.
3. Previously untreated patients diagnosed and treated at the Royal Bournemouth Hospital. Diagnostic Fluorescence *In-Situ* hybridization (FISH) was performed on all cases using commercially available probes (LSI TP53 (17p13.1); D12Z3 (centromere12), D13S25 (13q14.3), LSI ATM (11q22.3)) according to the manufacturer's instructions (Vysis from Abbott Laboratories, Berkshire, UK). 6q21 was analysed separately (courtesy of Dr S Stilgenbauer). 200 cell nuclei were examined and scored for each probe by two independent investigators. The cut-off points for defining loss were > 5 for 11q, 13q14,

17p13.1 and 6q21, and > 3 for trisomy 12. ZAP70 and CD38 expression was determined as previously described, where 10 and 30 positive cells were classed as positive, respectively. IGHV genes were sequenced as previously described and a cut-off of >98 germ-line homology was taken to define the unmutated sub-set.

### 2.1.2 Recovering CLL patient cells

Preparation of CLL samples prior to and during experimental treatment was conducted in a laminar flow hood ("hood") using aseptic techniques. The hood, and all items prior to entry into it were cleaned with 70% ethanol (EtOH).

1. Cryogenic vials containing Peripheral Blood Mononuclear Cells (PBMCs) extracted from the whole blood of CLL patients and stored in freeze media (FCS, 10% DMSO) (conducted by Ian Tracy on behalf of the CLL group), were removed from liquid nitrogen storage and placed on ice for transfer to the tissue culture room
2. PBMCs were defrosted in a 37°C water bath, during which period 10mL of pre-warmed culture media (DMEM, 10% FCS, 1% L-Glutamine, 1% Penicillin Streptomycin; "R10 media") was transferred to a 15mL tube and labelled with the sample identification number
3. Post sample thawing, the PBMCs were transferred from the cryogenic vial to the 15mL tube using a 1000μL pipette, ensuring to slowly release the cells into the 15mL tube down the wall of the tube. Following which, the cryogenic vial was washed with 1mL of R10 media and transferred into the 15mL tube. This step is conducted to remove the DMSO from the culture media, as it is toxic
4. Samples were then centrifuged at 1500rpm for 5 minutes, following which the supernatant was eluted into Virkon disinfectant, and the cell pellet resuspended in 3mL R10 media
5. In the 15mL tubes, samples were incubated for 1 hour in a 37°C 5% $CO_2$ incubator, to allow cell recovery. After which, the cells were ready for counting and downstream applications

### 2.1.3 Live cell count and sample concentration calculation

1. CLL PBMCs were counted using a Neubauer chamber haemocytometer, with a glass cover slip placed over the chamber
2. CLL PBMCs were mixed 1:1 with Trypan blue (viability stain) in a microcentrifuge tube, with 10µL of this mixture transferred into one of the chambers
3. Once prepared, the haemocytometer was transferred to a light microscope to view the cells and count them. After placing on the stage and focusing, the live CLL cells (non-Trypan blue stained) were counted using one box from each of the outer four quadrants (**Figure 8A**), only considering cells inside the box, as well as cells sitting on the left and upper boundary panes (**Figure 8B**).
4. This cell number was then used to calculate the sample concentration (1x10$^6$/mL), using the following equation:

$$sample\ concentration = \frac{(number\ of\ cells\ [n]\ x\ 10{,}000\ x\ dilution\ factor\ [2])}{number\ of\ boxes\ [4]}$$

5. For all experiments, the cell concentration was set at 1x10$^7$/mL. To achieve this, the sample concentration was divided by 10 (i.e. 10x10$^6$) and multiplied by the volume of R10 media they were resuspended in (3mL), to provide the volume required to obtain this concentration in mL

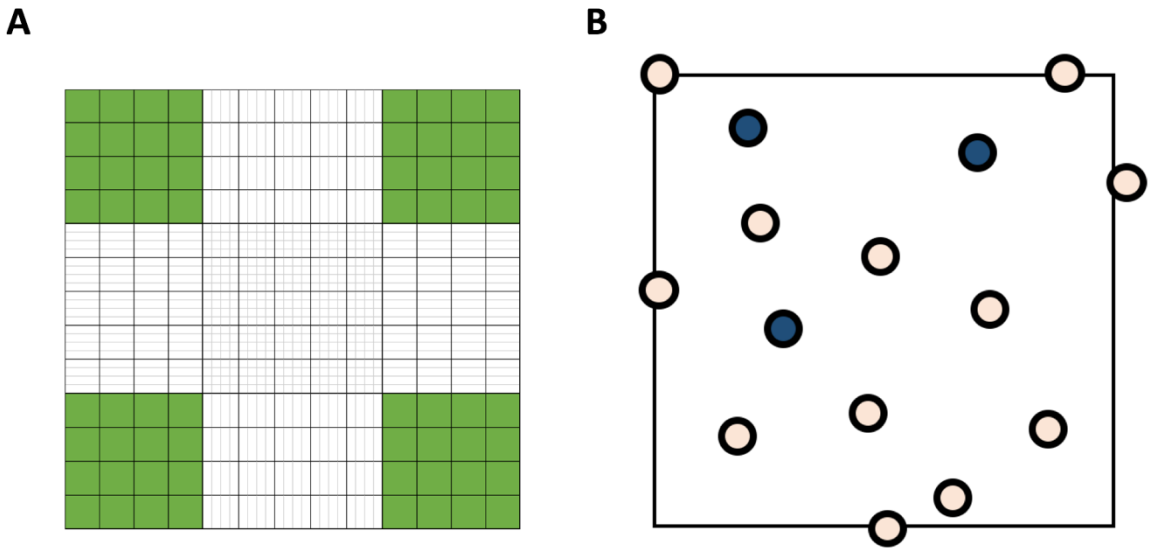



**Figure 8 Live cell count using a Neubauer haemocytometer**

**A** Example of the Neubauer chamber. Green boxes represent the one box per outer quadrant used to count CLL PBMCs. **B** Example of a cell count inside one of the green boxes from A. Cells are only included if they are; within the box boundary or placed on the upper or left boundary line, and alive (i.e. not stained with Trypan blue). The cell count for this box would therefore be 10 live cells. Blue = Trypan blue stained, dead, light orange = live cell.

### 2.1.4 DNA extraction

Genomic DNA was needed for applications such as WGA, PCR, Sanger sequencing, ddPCR, and NGS.

DNA was extracted using the DNeasy mini kit (Qiagen). The general protocol is as follows:

1. AW1 and AW2 were made up with 25µl and 30µl 100% ethanol, respectively.
2. Up to $5x10^6$ cells (per column) were centrifuged for 5 minutes at 300g and resuspended in 200µl PBS.
3. 20µl proteinase K and 4µl of RNase A (100mg/ml) was added and incubated at room temperature for 2 minutes; then 200µl buffer AL was added, mixed by vortexing and incubated at 56ºC for 10 minutes.
4. 200µl 100% ethanol was added and mixed by vortexing
5. The reaction mixture was pipetted into a DNeasy mini spin column placed in a 2ml collection tube and centrifuged for 1 minute at 6000g. The flow through and collection tube were discarded
6. The column was placed in a new collection tube. 500µl buffer AW1 was added to the column and centrifuged as above. The flow through and collection tube were discarded
7. The column was placed in a new collection tube. 500µl buffer AW2 was added to the column and centrifuged for 3 minutes at 20,000g to dry the membrane. The column was carefully removed from the collection tube to prevent any ethanol carryover. The flow through and collection tube were discarded
8. The column was placed in a new collection tube and 200µl buffer AE was added directly to the membrane. It was incubated at room temperature for 1 minute then centrifuged for 1 minute at 6000g. This step was repeated using the same elutant. The extracted DNA was stored at 4°C or -20°C

### 2.1.5 RNA extraction

RNA extracted from target cells may be used to investigate the effect of treatment of inhibitor compounds

RNA was extracted using the RNeasy mini kit (Qiagen). The protocol is as follows:

1. The DNase stock solution was prepared by suspending the lyophilized DNase 1 in 550µl RNase free water and gently mixing. 10µl aliquots were stored at -20°C
2. Buffer RLT was supplemented with 10µl β-mercaptoethanol per 1ml, in the fume hood. This solution was stable for 1 month.
3. 4 volumes 100% ethanol were added to buffer RPE to obtain a working solution.
4. To lyse the cells, 350µl buffer RLT was added to $<5x10^6$ pelleted cells. 600µl buffer RLT was added to a maximum of $10^7$ cells.
5. To completely homogenise the cells, they were passed through a 20-gauge needle fitted to an RNase free syringe, five times.
6. One volume 70% ethanol was added and mixed by pipetting. 700µl was applied to a column and centrifuged at 10,000g for 15 seconds at room temperature. The flow through was discarded. If the sample volume exceeded 700µl, successive volumes were applied to the column and centrifuged as above
7. 350µl buffer RW1 was applied to the column and centrifuged as above. The flow through was discarded
8. 70µl buffer RDD was added to a 10µl aliquot of DNase 1 solution, and mixed by inverting the tube. This solution was applied directly onto the RNeasy silica-gel membrane and incubated for 15 minutes at room temperature
9. 350µl buffer RW1 was added to the column and centrifuged as above. The column was then transferred to a new 2ml collection tube
10. 500µl buffer RPE was applied to the column and centrifuged as above. The flow through was discarded. Another 500µl buffer RPE was added and centrifuged at 10,000g for 2 minutes to dry the membrane
11. The column was transferred to a new 1.5ml collection tube and centrifuged at 10,000g for 1 minute
12. 50µl RNase free water was added to the membrane and incubated at room temperature for 10 minutes. The column was spun at 10,000g for 1 minute. This step was repeated once. The RNA was stored at -80°C

### 2.1.6 Nucleic acid assessment

After DNA or RNA extraction, the nucleic acids were assessed to check the quality and quantity before being applied for downstream applications.

#### 2.1.6.1 Gel electrophoresis

Gel electrophoresis was used to check DNA quality and PCR amplification. As the phosphate backbone of DNA is negatively charged, it can be loaded onto a semi permeable gel and an electrical current applied that pushes the DNA towards a positive anode. As it migrates through the gel, fragments of DNA separate by size with large high molecular weight fragments moving more slowly than those with a smaller molecular weight. An intercalating dye is used to bind the DNA and fluoresce under ultra-violet (UV) light, allowing an assessment of DNA quality to be made. DNA comprised of different sized fragments will pass through the gel at different times creating a ladder like image. Genomic DNA is comprised of very large molecules that will not migrate far, creating a discrete band. A smear indicates DNA degradation or digestion **(Figure 9)**.

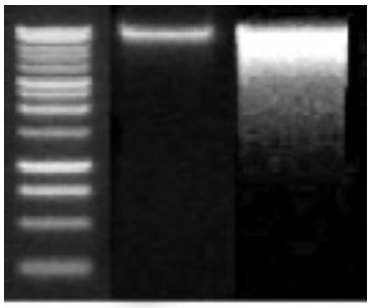


**Figure 9 Agarose gel image**

DNA ladder (left lane), genomic DNA (middle lane), and degraded DNA (right lane)

To make a 1% agarose gel:

1. 1g agarose powder (Promega) and 100ml 1xTAE were combined and heated in the microwave for 2 minutes. After cooling slightly, 5µl SafeView was added and the gel poured into a casting tray.
2. 2µl 6x DNA loading buffer (Promega) was added to 1-8µl DNA, depending on the expected concentration, and made up to 10µl with nuclease free water. The sample was loaded into one of the wells. Loading buffer contains glycerol which pulls the DNA to the bottom of the well, and dyes to allow the sample to be tracked down the gel. A 1kb DNA ladder was also loaded for reference (Promega).
3. The gel was run for 40-60 minutes at 120V
4. A UV transilluminator was used to visualise the gel and a photo was taken of the gel

### 2.1.6.2 DNA quantitation by nanodrop

Nano drop is a spectrophotometer that assesses the amount of DNA/RNA and contaminants present in a sample by measuring the amount of light that it absorbs. The sample is pipetted onto a fibre optic cable (the receiving cable) and a second fibre optic cable (the source cable) is bought into contact with the sample, causing the liquid to bridge the gap between the fibre optic ends. A pulse of light originating in the source cable is passed through the sample. When a photon encounters a DNA molecule it is absorbed and the intensity of light reaching the receiving cable is reduced and measured. DNA absorbs UV light at a wavelength of 260nm, proteins absorb light at 280nm and 230nm. Other contaminates such as carbohydrates also absorb at 230nm. Absorbance is measured at these three wavelengths, allowing the concentration and purity of the DNA sample to be determined (**Figure 10**).

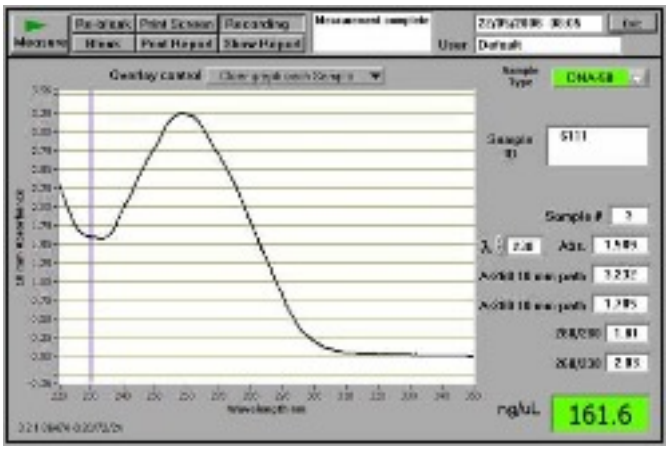



**Figure 10 Nanodrop DNA profile showing the concentration and absorbance ratios**

To quantify DNA using Nanodrop:

1. With the sampling arm in the down position the NanoDrop-1000 software was started and the nucleic acid application selected.
2. NanoDrop-1000 was calibrated by applying 1.5µl nuclease free water to the lower measurement pedestal, lowering the sampling arm, and selecting blank.
3. The water was wiped from both pedestals. 1.5µl nuclease free water was measured and a flat base-line was returned. Both pedestals were wiped.
4. 1.5µl of each DNA sample was measured. Each profile was saved after measurement and the pedestals were wiped.
5. The software provided a concentration in ng/µl. This is calculated in the following way:

   Optical density (absorbance reading at 260nm) x 50 (1 absorbance unit at 260nm= 50µg/µl DNA)

6. DNA purity is calculated by dividing the absorbance at 260nm by the absorbance at 280nm and at 230nm. This ratio should fall between 1.5 and 2.0. A ratio <1.5 indicates high contamination.

#### 2.1.6.3 DNA quantification by Qubit (Qubit™ 3.0)

Qubit fluorometric quantitation incorporates fluorescent dyes selectively into DNA, so increasing the sensitivity and specificity compared to spectrophotometer methods, that can be compounded by contaminants and nucleotides. DNA concentration is measured by comparing the fluorescent signal to the signals generated by a range of known concentration standards. Qubit was used to assess DNA quantity prior to and during any post amplification quantification steps in the NGS methods. The double stranded DNA (dsDNA) broad range assay kit was used for genomic DNA while the dsDNA high sensitivity assay kit was used for amplified DNA.

1. 2µl DNA sample was transferred to a Qubit tube.
2. 10µl of standard 1 and standard 2 were put into separate tubes.
3. The qubit working solution was prepared by diluting the dsDNA reagent 1:200 in dsDNA buffer. DNA samples and standards were made up to 200µl with the working solution.
4. Samples were vortexed and briefly centrifuged then incubated in the dark for 2 mins to incorporate the fluorescence dye
5. Samples were analysed on the Qubit V3.0 for DNA concentration

### 2.1.7 Whole Genome Amplification (WGA)

Genomic DNA was amplified using the illustra GenomiPhi V2 DNA amplification Kit. This allows highly uniform amplification over the entire genome, so that locus representation remains extremely close to the original template. WGA DNA can be used in downstream applications such as sequencing, and prevents the depletion of precious archived material.

1. 1µl genomic DNA (10ng/µl) was transferred to new plate/tube
2. 9µl sample buffer was added to the DNA and heated to 95°C for 3 minutes
3. Samples were incubated on ice whilst master mix was prepared (for multiple samples): 9µl reaction buffer and 1µl enzyme mix per sample
4. Thermocycling conditions were: 30°C for 1.5 hours, then 65°C for 10 minutes (enzyme inactivation)
5. WGA DNA was diluted to a working concentration (~10ng/µl; 4µl WGA DNA added to 96µl $H_2O$) and stored in the freezer

### 2.1.8 Primer design

Sequences for the genomic regions of interest were extracted from Ensembl (http://www.ensembl.org/index.html) (GRCh37/hg19 genome assembly). The sequence was uploaded to the on-line Primer3 (http://frodo.wi.mit.edu/ ) software, which designs primers based on specific parameters e.g. product size and primer Tm. Sanger sequencing primers had an optimum annealing temperature of 55°C, and the product length was between 450-500bp long. PCR Blat (BLAST-like alignment tool) (http://genome.ucsc.edu/cgi-bin/hgPcr?org=Human&db=hg18&hgsid=248670511) was used to confirm primers align with the correct genomic location and the online oligo calculator; http://www.basic.northwestern.edu/biotools/oligocalc.html was used to ensure primers would not self-anneal in primer-dimers or hairpin structures. Primers were purchased from Eurogentec.

### 2.1.9 Sanger sequencing

Sanger sequencing is used to detect the presence of mutations in genomic DNA, relative to a reference sequence. Primers were designed as described in section **(2.1.8)**.

#### 2.1.9.1 PCR amplification

The following Master mix was prepared: 10µl GoTaq master mix (Promega), 0.35µl each of 10µM forward and reverse primers, 4µl DNA (10ng/µl), 5.3µl $H_2O$. Samples were placed in the thermocycler under the following touchdown conditions: 95°C for 10 minutes, followed by 10 cycles of 30 seconds at 95°C, 30 seconds at 60°C (decreasing by 1°C per cycle) and 30 seconds at 72°C, followed by 30 cycles of 30 seconds at 95°C, 60 seconds at 56-60°C (depending on primer Tm) and 30 seconds at 72°C, ending with 72°C for 10 minutes and a 4°C hold.

#### 2.1.9.2 PCR purification

The ExoSap protocol allows the enzymatic removal of excess nucleotides and primers from PCR reactions. The exonuclease I removes leftover primers, while the Shrimp Alkaline Phosphatase removes any remaining dNTPs.

1. The following master mix was prepared

| Reagents | 1x |
|---|---|
| **Exonuclease I** | 0.025µL |
| **Shrimp Alkaline Phosphatase** | 0.250µL |
| **MilliQ water** | 9.725µL |
| **Final Volume** | 10µL |

2. 10µL of the 1x ExoSAP mix was added to each PCR sample and incubated at 37˚C for 30 minutes and then 95˚C for 5 minutes.

4. Samples were stored at 4˚C or -20˚C until needed.

#### 2.1.9.3 Sequencing

10µl of PCR product (10ng/µl) and 10µl of primers (3.2pg/µl) were sent to Source Bioscience (Oxford) for Sanger sequencing.

#### 2.1.9.4 Sequencing analysis

The raw sequence data was then visualised, aligned and interpreted using SeqMan (DNASTAR Lasergene) software to confirm if any mutations were present within the sequence, by a comparison to the wild-type (WT) or reference genomic sequence **(Figure 11)**.

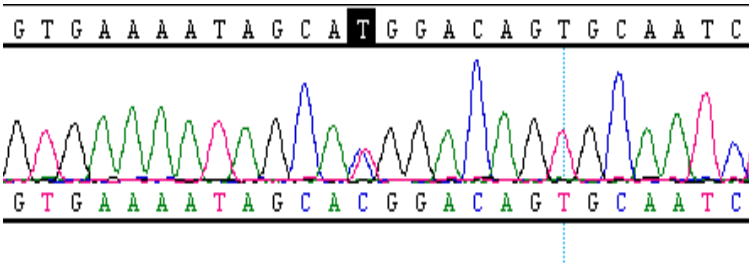



**Figure 11 Sanger Sequencing trace**

Each base is represented by a peak (A, T, C and G are displayed as green, red, blue and black, respectively). The reference sequence is displayed above the trace and the sequence of the sample is displayed below. In the centre of the trace, two overlapping peaks can be seen; a red 'T' (highlighted in black in the reference sequence) and a blue 'C'. This sample contains a T>C mutation at this locus, in approximately 50% cells (determined by the equal size of the two peaks)

### 2.1.10 Droplet digital PCR

Droplet digital PCR (ddPCR) can detect mutant DNA present in 0.1%, in a WT background; some studies have reported an improved sensitivity of just 0.0005%. The test and reference assays contain a single set of primers and two competitive probes to detect the WT and mutant alleles (labelled with HEX and FAM, respectively). A modified 5' fluorescent moiety oligonucleotide probe is quenched by a 3' quencher moiety that when bound to a specific target DNA, is cleaved by the 5' to 3' nuclease activity of the *Taq* polymerase, releasing the 5' fluorescent nucleotide and so generating a signal(155).

1. The following master mix (dependant on the gene of interest) was prepared:

a. NOTCH1:10µl 2x ddPCR supermix was mixed with 1µl 20x NOTCH1 p.P2514fs*4 (FAM) mutant assay, 1µl 20x NOTCH1 WT Human (HEX) and 3µl $H_2O$
b. SF3B1: 11µl 2x ddPCR supermix was mixed with 0.55µl 40x SF3B1 custom assay (AHGS8IT) and 4.95µl $H_2O$

2. 5.5µl DNA sample (10ng/µl) was added to the mastermix
3. 20ul of sample was transferred to the middle row of the droplet generator (DG8™) cartridge and 70µl droplet generator oil was pipetted into the first row.
4. A gasket was hooked over the cartridge then loaded onto the droplet generator
5. Oil and sample were drawn up through the microfluidic channels and droplets were generated and placed into the droplet well
6. 40µl of the emulsion was transferred carefully to a PCR plate, which was heat sealed and put on the thermocycler under the following conditions; 95°C for 10 minutes (enzyme activation); 40 cycles (with a ramp rate of 2°C/sec) of 30 seconds at 95°C (denaturation) and 60 seconds at 55°C (annealing/extension); followed by 10 minutes at 98°C (denaturation)
7. The plate was then transferred to the QX200 droplet reader; this aspirates the sample into singular droplets and detects and reads each droplet to determine if it is carrying a WT or mutant (or both) allele.
8. Data analysis was performed using the Quantasoft™ software (Bio-Rad). This generates four distinct clusters of droplets; 1) Negative droplets containing no DNA template, 2) WT droplets, 3) mutant droplets, 4) WT and mutant containing droplets. The concentration of the WT and mutant are calculated using Poisson statistics based on the number of droplets in each cluster **(Figure 12)**.

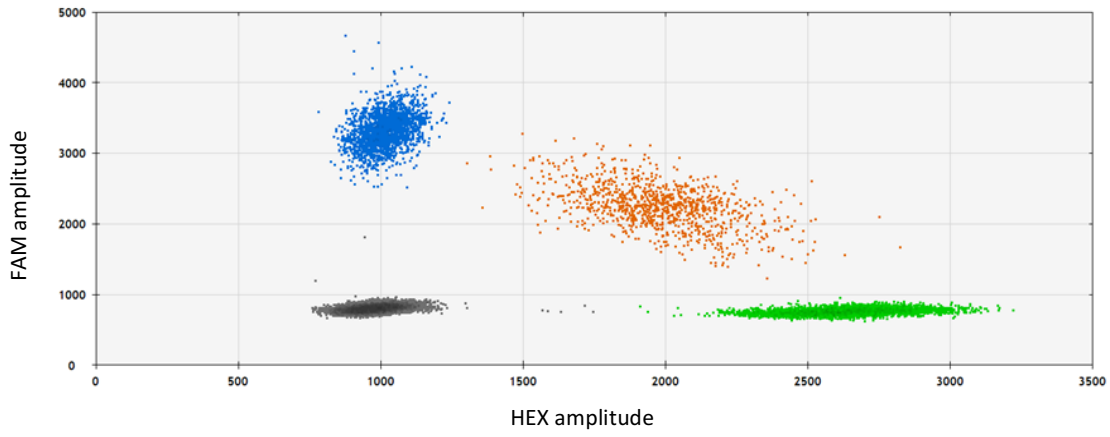



**Figure 12 Droplet digital PCR analysis**

Four distinct clusters of groups are generated; negative droplets (grey), wildtype (green), mutant (blue) and heterozygous (orange).

### 2.1.11 Reverse Transcription PCR (RT-PCR)

RT-PCR is used to determine if a sequence of DNA is being transcribed into mRNA. RNA samples are reversed transcribed into cDNA, which is subsequently amplified by a PCR reaction. The PCR products are run on an agarose gel. The presence of a band indicates that the sequence of interest has been transcribed.

RT-PCR was performed on 1 µg of total RNA isolate using the Improm™II RT-PCR kit (Promega, UK). All RT-PCR reaction set-ups contained positive (Kanamycin RNA), Reverse Transcription negative (RT-ve) and no-template controls (NTC) to control for reagent failure, internal and external gDNA contamination respectively.

1. 2µl RNA, 1µl Random primers and 2µl $dH_2O$ were mixed and denatured at 70°C for 5 minutes on a thermocycler to remove any secondary RNA structures. The samples were immediately put onto ice.
2. The following master mix **(Table 1)** was prepared and added to the primer/template;

**Table 1 RT-PCR master mix**

| Reagents | Kanamycin Reaction | RT-ve Reaction | Experimental/ NTC Reaction |
|---|---|---|---|
| **Nuclease-Free water** | 4.2 µl | 8.8 µl | **7.3 µl** |
| **5 x ImPromII™ Reaction buffer** | 4.0 µl | 4.0 µl | **4.0 µl** |
| **$MgCl_2$ (Final conc 1.5 mM)** | 4.8 µl | 1.2 µl | **1.2 µl** |
| **dNTP mix** | 1.0 µl | 1.0 µl | **1.0 µl** |
| **RNasin Ribonuclease Inhibitor (20 U)** | 0.0 µl | 0.0 µl | **0.5 µl** |
| **ImPromII™ RTase** | 1.0 µl | 0.0 µl | **1.0 µl** |
| **Final Volume** | **15 µl** | **15 µl** | **15 µl** |

3. Tubes were transferred to the thermocycler programme and heated to 25°C for 5 mins, 42°C for 1 hr, 70°C for 15 mins and then finally cooled and held at 20°C. The cDNA was diluted to a working concentration (sample >20ng/µl 1:7 dilution and samples <20ng/ µl 1:5 dilution) and then stored at -20°C until needed.

### 2.1.12 Illumina TruSeq Custom Amplicon Next Generation Sequencing

#### 2.1.12.1 Hybridisation of the oligo pool

1. DNA quantity and quality was assessed using an agarose gel and Qubit. DNA was diluted to either 5ng/µl or 25ng/µl, with the low input method having a maximum input of 15µl, and the high input having a maximum input of 10µl. All further steps were conducted in the same way regardless of DNA quantity input.
2. 5µl CAT was added to the gDNA. 35µl OHS2 was added and pipet mixed.
3. The plate was sealed with an adhesive aluminium foil, centrifuged (1000 x g) for 1 minute and placed on the Hybex heating block for 1 minute at 95°C.
4. The temperature was reduced to 40°C and the plate left to incubate for 80 minutes

#### 2.1.12.2 Removal of unbound oligos

5. The filter plate assembly unit (FPU) was assembled. 45µl SW1 was added to each sample well and centrifuged at 2400 x g for 10 minutes.
6. The sample plate was removed from the heating block and centrifuged for 1 minute at 1000x g.
7. The entire sample was transferred to the FPU plate and centrifuged at 2400 x g for 2 minutes
8. 45µl SW1 was added to each sample well and centrifuged for 2 minutes (2400 x g). this step was repeated.
9. The flow through was removed, 45µl of UB1 was added to each sample and centrifuged for 4 minutes at 2400 x g.

#### 2.1.12.3 Extension-ligation of bound oligos

A DNA polymerase extends from the upstream oligo through the targeted region, followed by ligation to the 5' end of downstream oligo using DNA ligase, resulting in the formation of products including the regions of interest and amplification sequences.

10. 45µl ELM4 was added to each sample well on the FPU plate. The plate was sealed with an adhesive aluminium foil seal, then incubated at 37°C for 45 minutes. The plate was centrifuged for 4 minutes at 2400 x g.
11. 25µl of 50mM NaOH was added to each sample well, pipetted mixed 5 times then incubated at room temperature for 5 minutes. Pipette mixing was repeated and 20µl of sample was transferred to a new plate.

#### 2.1.12.4 PCR amplification

Indexing PCR primers used to amplify and multiplex samples can be found in the example manifest file in appendix **A.1**.

12. 4µl of appropriate i5 primer was added to each sample followed by 4µl of appropriate i7 primer.
13. The following master mix was prepared (per sample): 0.58µl TDP1 and 29.2µl PMM2, pipette mixed 20 times. 22µl PCR master mix was transferred to each well, the plate was sealed and centrifuged for 1 minute at 1000 x g.
14. The plate was placed on thermocycler under the following conditions; 95°C for 3 minutes, 24 cycles of 95°C for 30 seconds, 66°C for 30 seconds and 72°C- 60 seconds, 72°C for 5 minutes, and a 10°C hold.

#### 2.1.12.5 PCR clean up

15. Following amplification, the plate was centrifuged for 1 minute at 1000 x g.
16. 5µl PCR product was loaded into an agarose gel (125V for 40 minutes) to confirm the library successfully amplified.
17. 45µl (room temperature) AMPure XP beads were added to a MIDI plate. The remaining PCR product was transferred to the MIDI plate and the plate was sealed.
18. The plate was shaken for 2 minutes, then removed and incubated for 10 minutes at room temperature followed by incubation on a magnetic stand for 2 minutes.
19. The supernatant was removed whilst the plate was on the magnetic stand. 200µl 80% ethanol was added, incubated for 30 seconds and then removed. This wash was removed.
20. The samples were air dried for 10 minutes, then 30µl EBT was added to each sample well. The plate was sealed and shaken for 2 minutes, then removed from the shaker and incubated for 2 minutes at room temperature.
21. The plate was placed on the magnetic stand for 2 minutes and 23µl of the EBT was then transferred onto a new plate.
22. Qubit HS DNA quantification was performed and the libraries were diluted to 4nM accordingly

#### 2.1.12.6 Library pooling and MiSEQ sample loading

23. 5µl of each 4nM normalised sample was combined into a single tube (PAL), the lid was secured with Parafilm and then incubated at 95°C for 5 minutes.
24. The PAL was removed and put on a shaker for 5 minutes (1800rpm)

25. 5µl PAL was transferred to a new tube (20pM PAL) with 5µl 0.2M NaOH, vortex mixed and centrifuged then incubated at room temperature for 5 minutes.
26. The library was diluted to 20pM with 990µl HT1.
27. 375µl 20pM PAL was transferred to a new tube (DAL). 219µl HT1 and 6µl of 12.5pM PhiX was added to give a concentration of 12.5pM. The DAL was vortex mixed and centrifuged then incubated at 95°C for 5 minutes.
28. The DAL was immediately put in an ice bath for ~10 minutes then loaded onto a defrosted MiSeq V2 500 cycle reagent cartridge. The cartridge was loaded into the Illumina MiSeq along with a prepared flow cell and the library was sequenced.

### 2.1.13 Thermo Fisher Ion Torrent Next Generation Sequencing

#### 2.1.13.1 DNA target amplification preparation

1. A maximum volume of 7.5µL of DNA (1-100ng) was combined with 5µL 5X Ion AmpliSeq HiFi mix.
2. This was mixed thoroughly before 5µL pipetted into two wells in a 96-well plate, followed by 5µL of each primer pool into each well, and then sealed ready for target amplification.

#### 2.1.13.2 Target amplification

3. Targets amplified on a thermocycler under the following conditions: 99°C for 2 minutes, 21 cycles of 99°C for 15 seconds, 60°C for 4 minutes, and a 10°C hold.

#### 2.1.13.3 Combination of target amplification reactions

4. The plate seal was carefully removed and the 2 PCR products per sample were combined

#### 2.1.13.4 Partial amplicon digestion

5. 2µL of FuPa reagent was added to each well, followed by the 96-well plate being sealed, vortexed, and centrifuged
6. The plate was placed on thermocycler under the following conditions: 50°C for 10 minutes, 55°C for 10 minutes, 60°C for 20 minutes, and a 10°C hold.

#### 2.1.13.5 Adapter ligation and purification

7. 4µL Switch solution was added, followed by 2µL IonCode Adapters, and then 4µL DNA Ligase.

8. The 96-well plate was sealed, vortexed, and centrifuged prior to being placed on a thermocycler under the following conditions: 22°C for 30 minutes, 68°C for 5 minutes, 72°C for 5 minutes, and a 10°C hold.
9. The seal was carefully removed from the 96-well plated, followed by the addition of 45μL of AMPure XP beads to each library, with the bead suspension thoroughly mixed by pipetting.
10. The bead suspension was incubated for 5 minutes at room temperature.
11. The plate was placed on a magnetic stand for 2 minutes, followed by removal of the supernatant
12. 150μL of 70% EtOH was added, with the plate orientated on opposing sides of the magnet stand to mix the beads in the supernatant.
13. Following which steps 10-12 were repeated
14. All supernatant was removed, allowing the beads to air dry for 5 minutes at room temperature

#### 2.1.13.6 Library amplification

15. The plate was removed from the magnet stand, then 50μL of Platinum PCR SuperMix HiFi and 2μL Library Amplification Primer Mix was added to each bead pellet.
16. The plate was sealed, vortexed, and centrifuged.
17. The plate was returned to the magnet stand for 2 minutes at room temperature, followed by transferal of 50μL of the supernatant to a new 96-well plate
18. The new plate was sealed, and placed on a thermocycler under the following conditions: 98°C for 2 minutes, 5 cycles of 98°C for 15 seconds, 64°C for 60 seconds, and a 10°C hold.

#### 2.1.13.7 Purification of amplified library

19. 25μL of AMPure XP beads were added to each well and pipette mixed, followed by a 5-minute room temperature incubation.
20. The plate was placed on a magnet stand for 5 minutes at room temperature
21. The supernatant was transferred to a new 96-well plate
22. Steps 19-20 were repeated, but with 60μL AMPure XP beads were added at the start instead of 25μL
23. 150μL of 70% EtOH was added, with the plate orientated on opposing sides of the magnet stand to mix the beads in the supernatant, then the supernatant removed.
24. Step 23 was repeated

25. All supernatant was removed, allowing the beads to air dry for 5 minutes at room temperature
26. The plate was removed from the magnet stand, followed by the addition of 50μL of Low TE, with the beads dispersed
27. The plate was sealed, vortexed, and centrifuged, followed by 2 minutes on the magnet stand at room temperature
28. An aliquot of each library was quantified using Qubit
29. Library normalisation and sequencing conducted using the standard Ion Torrent protocol

### 2.1.14 Oxford Gene Technologies (OGT) SureSelect Next Generation Sequencing

250ng of genomic DNA was prepared for all samples screened and sent to OGT for sequencing analysis of *TP53, ATM, SF3B1,* and *NOTCH1*. All sequencing analysis and bioinformatics was conducted using their industry pipeline. Visualisation of mutations called in the OGT data was conducted as described **(2.2.5)**.

### 2.1.15 *In vitro* drug treatment

1. All drugs were resuspended in the appropriate vehicle and stored in the appropriate conditions: with short term stocks stored at -20°C, and long term stocks stored at -80°C
2. When required, drugs were taken from -20°C and placed on ice until use in the hood
3. Drugs were added either at ten times the final concentration, with the volume fixed at one tenth of the cell treatment volume (experimental: Spliceostatin A [SSA], Meayamycin B [MMB], venetoclax, UMI-77), or by adding the required volume of drug to reach the final concentration, considering the cell treatment volume (support: CD40 ligand [CD40L], Interleukin-4 [IL-4], Q-VD-OPh [QVD]). For the highest dose of the former, and for the selected dose of the latter, the following equation was used:

$$drug\ concentration\ calculation = \frac{(desired\ concentration\ [10X\ or\ final]\ x\ desired\ total\ volume\ )}{stock\ concentration}$$

4. Following this, serial dilutions of the experimental drugs was conducted, by adding the appropriate amount of R10 media with the selected drug concentration (either 1:1, or at fixed volumes to generate specific concentrations)
5. Treatment of the cells was conducted for 6, 12, 18, 24, 48, or 72 hours as described in the chapter sub-section for each experiment
6. For support treatments, CLL cells were pre-treated for either 30 minutes (QVD) or 6 hours (CD40L/IL-4) prior to experimental drug treatment
7. Vehicle control treatment was conducted using the same concentration of vehicle in R10 media as the highest dose of the experimental drug for that experiment

### 2.1.16 B-cell Receptor stimulation

CLL cells were stimulated using immobilised anti-IgM for all BCR associated experiments.

1. Immobilised anti-IgM treatment of CLL cells
    a. Cells were treated at a ratio of 2:1 beads:cell, which translates to 3μL of beads per 100μL of cells
    b. Once calculated, the appropriate volume of beads was transferred from the stock microcentrifuge tube to a new tube ready for washing
    c. To remove the Sodium Azide, the beads were washed three times in 1mL R10 media using a magnet, before being resuspended in its original volume of R10 media
        i. The tube was placed in a magnet rack

ii. Wait for all beads to move to the wall of the microcentrifuge tube
iii. Whilst the tube was in the rack, the supernatant was removed, followed by the addition of 1mL R10 media
iv. The tube was then removed from the rack and resuspended via pipetting
v. The above steps were repeated two more times prior to resuspending the beads in their original volume

d. Once the beads were prepared, they were placed on ice prior to addition to the cells, making sure that the bead suspension was entirely mixed

### 2.1.17 Flow cytometry

Flow cytometry was used to assess the cell viability of CLL cells and the Ramos cell line after *in vitro* treatment with inhibitor compounds using Annexin V(AV) FITC/ Allophycocyanin (APC) (specifically for experiments with bead stimulation) and Propidium Iodide (PI). Flow cytometry experiments were set up in a final volume of 100μL of cells including treatment in a 96 well plate. An example flow cytometry plot is provided in **Figure 13A**, with an additional example of using AV conjugated to APC when assessing apoptosis in the context of immobilised αIgM bead stimulation, where AV-FITC cannot be used due to the beads also fluorescing in that channel (**Figure 13B**). For AV/PI staining (**Figure 13A**), The lower left quadrant represents viable cells, the lower right quadrant represents AV positive PI negative cells (so called "early apoptotic"), and the upper right quadrant represents AV positive PI positive cells (so called "late apoptotic"). No events generally occur in the upper left quadrant, since PI cannot enter cells with intact membranes. For AV (APC) staining (**Figure 13B**), cells to the left of the horizontal line are viable, whilst to the right of the line are apoptotic. The multiple populations within the apoptotic population represent the binding of 0 (lowest population), 1 (2$^{nd}$ lowest), and 2 (highest, discrete) beads binding each cell. Downstream analysis of all AV flow cytometry experiments used the percentage of viable cells to calculate the impact of SMIs on cells *in vitro*.

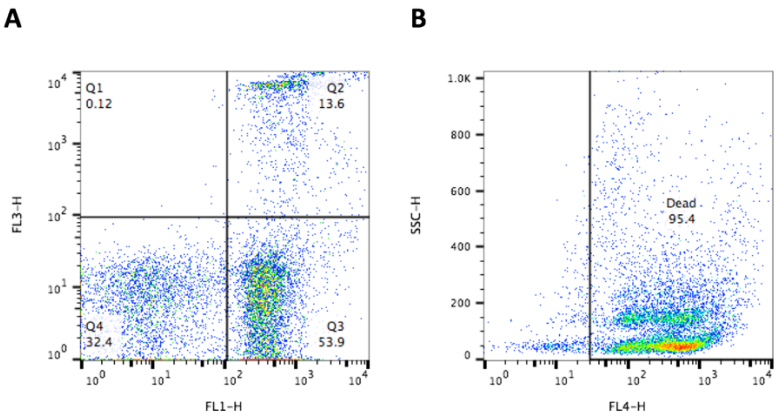



**Figure 13 Annexin V flow cytometry**

**A** Example flow cytometry plot of CLL cells treated with 5nM Spliceostatin A for 24 hours, processed by AV/PI flow cytometry. FL1-H represents the fluorescent channel which captures FITC, FL3-H represents the fluorescent channel which captures PI. **B** Example flow cytometry plot of CLL cells pre-stimulated with immobilised αIgM beads for 6 hours, then treated with 100nM venetoclax for 24 hours, processed by AV flow cytometry. FL-4 represents the fluorescent channel which captures APC, SSC represents the complexity of the cells.

To prepare the cells for flow cytometric analysis:

1. 250μL of Hank's Balanced Salt Solution (HBSS) was added to each well containing cells, and transferred into a pre-labelled FACS tube, following which cells were centrifuged at 1500rpm for 5 minutes
2. Whilst the cells were being centrifuged, a mastermix to resuspend the cells including the appropriate AV components was made (300μL/FACS tube: AV-FITC/PI=AV binding buffer 30μL, $dH_2O$ 270μL, AV-FITC 1μL, PI 0.5μL. AV-APC=AV binding buffer 30μL, $dH_2O$ 270μL, AV-APC 1μL)
3. Once the centrifugation process had finished, the supernatant was discarded (into a container with Virkon), and cells were resuspended into 300μL of the appropriate AV mastermix before being placed on ice prior to processing
4. Samples were processed using a BD FACSCalibur instrument attached to a Macintosh analysis computer. The BD CellQuest™ Pro software was used for; instrument settings; acquisition plot generation; and data acquisition and storage (as .fcs files)

#### 2.1.17.1 Raw flow cytometry data analysis

Analysis of raw flow cytometry data was conducted using FlowJo 10.2, with raw values exported to Microsoft Excel for data normalisation followed by input into GraphPad Prism for statistical analysis and graph generation **(**see **2.2.9** for further details**)**.

1. FlowJo analysis
    a. The .fcs files saved during the flow cytometry experiment were loaded into FlowJo
    b. The non-treated sample per patient was opened, revealing a FSC vs. SSC graph
    c. Using the dropdown menus on each axis, the graph was changed to show either AV-FITC (FL-1) or AV-APC (FL-4) on the *x* axis, and PI (FL-3) on the *y* axis in the case of AV-FITC, or kept on SSC in the case of AV-APC
    d. Cell populations were gated to identify the percentage of viable cells in each treatment group, with the gate kept in the same position for all groups within an experiment
2. Data normalisation
    a. Viable cell percentages were loaded into Microsoft Excel, with all values normalised to the vehicle control using the following formula:

$$\frac{Treatment\ group\ \%\ viable\ cells}{Vehicle\ control\ \%\ viable\ cells}\ x\ 100$$

    b. These values were calculated per condition and per treatment before being prepared for analysis in GraphPad Prism

### 2.1.18 Protein extraction

Extracted proteins were used to determine the impact of drug inhibitor treatment on apoptotic pathway members

Cells stored in DMSO were removed from the -80°C freezer and thawed on ice.

1. 500µl of the thawed cells were transferred to a 1.5ml Eppendorf tube and 500µl of cold HBSS was added.
2. Samples were centrifuged at 4°C for 5 minutes at 2,500rpm. The supernatant was removed and discarded.
3. 500µl of HBSS was added and centrifuged as before. The supernatant was removed.
4. 70µl of lysis buffer (100µl whole cell lysate (WCL), 1µl PMFS (17.4mg/ml DMSO) and 1µl of protease inhibitor was added and pipet mixed.
5. The sample was centrifuged at 4°C at 13,000 rpm for 10 minutes to pellet the nucleic acids.
6. The lysate was transferred to a new tube and stored at -20°C until needed.

### 2.1.19 Immunoblotting

1. 20µl protein sample and 10µl 4x Loading Dye (7.5µl Loading dye (Invitrogen) and 2.5µl DTT (1M) (Sigma)) were mixed.
2. Samples were heated to 95°C for 5 minutes
3. The white strip and the comb were removed from the gel.
4. The gel tanks were assembled and the gel running buffer was made: 50 ml 20X stock MOPS (Invitrogen) + 950ml $dH_2O$
5. The central reservoir was filled with running buffer and 500µl antioxidant (Invitrogen) was added and mixed using a Pasteur pipette. Any large bubbles were removed from the buffer surface.
6. The outer reservoirs were filled with running buffer and the gel wells were washed 3 times using a syringe.
7. 30µl sample was loaded into the wells. One lane was loaded with Novex® Sharp Pre-stained Protein Standard (LC5800) and 6µl rainbow marker. The gel was run for 1 hour (150v 75mA 12.5W).
8. Transfer buffer was made (20ml 20X trans Buffer at 4°C, 80ml Methanol, 300ml $dH_2O$ and 400µl Antioxidant) and poured into 2 containers with 5 sponges and 1 Hybond (Amersham) membrane.
9. The gel was removed from the tank and the front casing was opened.
10. A wet piece of 8.5cm x 6cm 3MM filter paper (Whatman) was placed over the gel and it was turned over.
11. The gel spatula was used to remove the back casing. A small amount of the transfer buffer was put onto the gel and the wet Hybond was placed directly onto the gel. A second 3MM filter paper soaked in transfer buffer was put on top of the Hybond.
12. A plastic pipette was rolled over the top of the gel, Hybond and 3MM paper to remove any bubbles which can reduce protein transfer to the membrane.
13. The wet sponges were arranged in the blotting cassette and the front cover was placed on the blotting rig and placed into the tank.
14. The blotting cassette was clamped into place and filled with transfer buffer until the blotting pads were covered. The outer reservoirs were filled with water to cool the blotting cassette.
15. The gel was run for 1h 30mins at 25V, 125mA, 15.0W.
16. The Hybond membrane was removed and put in $dH_2O$ in a tray, stained for 30 seconds with Ponceau S (Sigma) and rinsed in $dH_20$.

17. The membrane was then washed in TBS tween for 5 minutes to remove the Ponceau S.
18. The membrane was wrapped around the inside of a falcon tube with 5ml VETO and 250µl FCS for 30mins on the roller to block non-specific Ab binding.
19. The solution was removed and replaced with fresh 5ml VETO and 250ul FCS along with the appropriate antibody. This was left to incubate overnight (>16 hrs) at room temperature.
20. The following day the membrane was removed and washed in TBS tween for 40 minutes (with 4 buffer changes) to remove non-specific AB.
21. The secondary antibody was added into 10ml VETO and 0.5ml FCS
22. The membrane was incubated for 2 hours, then washed in TBS tween for 45mins, changing the wash buffer at least 4 times
23. 350µl of Reagent 1 (peroxidase solution) was mixed with 350ul reagent 2 (luminol enhancer solution).
24. The membrane was blotted and ECL was added. It was incubated for 1 min and wrapped in cling film. The membrane was imaged.

### 2.1.20 Immunoblotting data analysis

1. Densitometry
    a. Analysis of sample protein expression levels was conducted using ImageJ 1.50i
    b. The .TIFF images taken during imaging were loaded into the program, and orientated if required so that the bands across the entirety of the membrane were level (Image > Transform > Rotate > select number of degrees of rotation)
    c. Using the box drawing function, a box was drawn around the protein band of interest, and checked for correct sizing against other bands for the same protein across the membrane
    d. Firstly, a measurement was taken (>command >M) of the background intensity (used to normalise the protein bands against)
    e. Next, each protein band for a lane of the membrane was measured for each patient, with all measurements copied into Microsoft Excel for loading control normalisation and protein expression analysis
2. Normalisation
    a. Raw values from ImageJ were loaded into Microsoft Excel
    b. The background intensity was subtracted from the raw value of each protein band

c. Following which, two approaches of additional normalisation were undertaken depending on the protein of interest
    i. For PARP, the cleaved PARP value was divided by the total PARP value to create a cleaved PARP ratio
    ii. For all other proteins, the protein of interest value was divided by the relevant lane loading control protein value
d. These normalised values were then taken forward for analysis in GraphPad Prism

## 2.2 Statistical/analytical

### 2.2.1 Checking TruSeq panel design capture regions in UCSC Genome Browser

To check the captured amplicon regions between TruSeq panel designs, visualisation in UCSC Genome Browser was conducted using the following process:

1. A BED file of the amplicon genomic start and end positions, including the chromosome number was generated. This was conducted in Excel then saved as a .txt document
2. This was then loaded into the UCSC Genome Browser by going to their website: https://genome.ucsc.edu/ , clicking on Genome Browser, then selecting "add custom track"
3. This was repeated for each TruSeq design, following which all three designs could be viewed as tracks within the Genome Browser at any genomic location searched within the tool

### 2.2.2 Individual amplicon read depth analysis using Qlucore

The individual amplicon read depth per patient was scored based on the following cut offs: 1) Mutation data was not discernible (≤50DP), 2) Only clonal mutation data was discernible, since the poor coverage was associated with high GC content, and the amplicon covers a hot spot CLL mutation (51-100DP) 3) Amplicon performance was sufficient to capture mutations down to 2.8% VAF (101-500DP), 4) Amplicon performance was sufficient to capture mutations down to 2.4% VAF (501-1000DP), 5) Amplicon performance was sufficient to capture mutations down to 1% VAF (>1000DP). This was then visualised in Qlucore:

1. Dataframe of coded data made in Excel from raw amplicon coverage files for each TruSeq panel design
2. Dataframe loaded into Qlucore as a .csv file, following the upload wizard to present the data
3. Selecting the heat map view, hierarchical clustering based on amplicon and patients to identify similar groups
4. Visual grouping of clustered amplicons to identify which genes were represented
5. Data exported as .txt files for processing in Excel

### 2.2.3 Bioinformatics pipeline

The pipeline was composed of two separate analyses, one specifically designed for point mutations and the other for indels. For point mutations, BAM files from BaseSpace were put through VCF Somatic, followed by the in-house Oxford CLL literature database. This database was manually curated and contains all known published CLL mutations up to 2014. Known variants and novel variants were then separately run through ANNOVAR and annotated VCF Somatic, with known variants which remain in the analysis taken forward for the in-house coverage VAF algorithm. This in-house algorithm compares the VAF of all variants identified against its read depth, with variants which pass this step achieving the required read depth for its VAF. Novel variants post annotated VCF Somatic go through further filters based on functional components, before also going through the coverage VAF algorithm. For indels, the BAM files from BaseSpace were processed through Stampy, after which the analysis was duplicated through VCF Platypus and VCF Pindel, with VCF Platypus analysis following the same process for novel point mutations **(Figure 14)**.

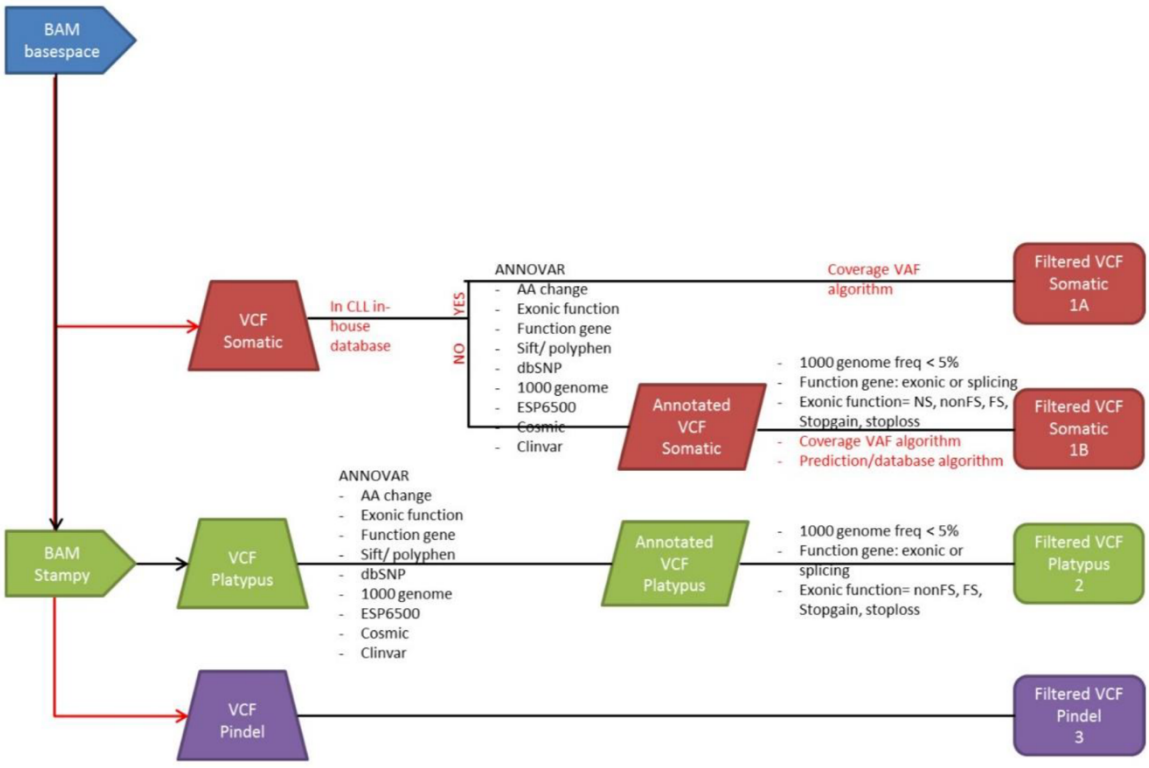



**Figure 14 Bioinformatics pipeline of TruSeq data**

Bioinformatics pipeline designed to analyse data from the CLL TruSeq panel. Stratified into three downstream analyses, the raw data from the MiSeq is the combination of two alignments to detect a maximum of SNVs and indels. Somatic is a variant caller detecting low VAF variants (down to 1%). Stampy and Platypus used to detect indels, with Pindel used to detect indels potentially missed by Stampy and Platypus.

### 2.2.4 Additional disease associated mutation annotation

WGS and WES supplemental mutation data from Landau *et al.* (2015)(75) and Puente *et al.* (2015)(57), were combined into a BED file (chromosomal start and end position of a variant, and the reference and alternate allele nucleotide) defining somatically acquired CLL mutations published in the last two years, removing duplicates. A BED file was also generated from the TruSeq dataset, with concatenated information including CLL4 ID, Hugo gene ID and amino acid change. These two BED files were then 'intersected' using BEDTools(156). This generated an output BED file of variants in the TruSeq dataset that were located previously found to be somatically acquired.

### 2.2.5 IGV mutation validation

Integrated Genomics Viewer (IGV) was stored locally for faster computational processing. For each variant identified in the 'passed' bioinformatics pipeline, IGV was used to visualise the sequencing data for that patient. This was achieved by loading the appropriate patient's '.bam' file, which contains the aligned sequencing data for that patient. The chromosome position of the variant was then searched for, followed by zooming in on this location to be able to see the sequenced region at basepair resolution.

To identify whether a variant was real or a false positive, a clear set of criteria was implemented. The following yes/no decision tree was followed:

1. Was the pipeline identified variant observed in IGV? (yes=take forward to next criteria, no=discard as false positive),
2. Was the variant located at the end of an amplicon read? (yes=discard as false positive, no=take forward to next criteria),
3. Did the surrounding basepairs have randomly distributed SNVs? (yes=discard "noise around the variant", no=take forward to next criteria),
4. Were there multiple low level variants at the same chromosomal position? (yes= discard as false positive, no= take forward to next criteria).
5. All mutations that passed these criteria were defined as IGV validated, to be taken forward for further functionally relevant filtering.
6. Screenshots of every variant analysed were taken and stored in PowerPoint.

### 2.2.6 R

All statistical approaches were conducted using publicly available packages. All analysis was conducted using R version 3.3.0

#### 2.2.6.1 Dataframe preparation and loading in R

1. Binary dataframe of cases and variables created in Microsoft Excel, saving as a .csv file for data loading
2. Dataframe loaded using the "<- read.csv(file.choose(), header = T)" command, where a Finder window opens to allow .csv file selection.
3. Each variable was coded as a vector by calling the dataframe title and denoting the column number in square brackets next to it.

   CLL4_ID <- dataframe[1]
4. Each variable was also coded as a numeric, using the following code:

   CLL4_ID_N<-as.numeric(dataframe$CLL4_ID)

#### 2.2.6.2 For looping in R

For looping allows the sequential testing of columns in a data matrix, if the columns to be tested and the sequence parameters are correctly defined. An example of for looping is presented in **Figure 15**.

1. This was achieved using the following code:

   for(i in seq(start,end,frequency)){

   a. Where "i" is the variable tested on each loop of the code
   b. seq() denotes how the for loop should proceed, with: "start" being a numerical value representing the first column to be tested, "end" being a numerical value representing the last column to be tested, and "frequency" being a numerical value defining what sequence the for loop should proceed (i.e. 1 = every column, 2 = every other column)
   c. The open brace bracket ({) opens the for loop to be applied to a function or package, with a closed brace bracket (}) used to finalise the for loop.

```
558 ####################Univariate Cox#####################
559
560 #For loop, i defined as going from column start to end moving 1 column each time. Square brackets in code used to define what to apply for loop to
561 for(i in seq(2,10,1)){
562
563   #Print function to assign each gene name above the Log Rank test output
564   print(paste("Univariate Cox for gene:", colnames(Thesis_Survival)[i]))
565
566   Allvariablesfit <- coxph(Surv(PFS_Rand_Months_11_N, PFS_Censor_11_N) ~ as.numeric(unlist(Thesis_Survival[i])))
567
568   print(summary(Allvariablesfit, conf.int = 0.95))}
569
```

**Figure 15 Example of for looping for univariate Cox PH modelling**

For loop of univariate Cox PH models. Thesis_Survival is the name of the dataframe. Allvariablesfit was the name of the model that was made in each loop, overwriting the analysis for the previous variable. The printed summary of the model is conducted for each loop, providing the output of the Cox PH analysis. Green text is not considered as code.

#### 2.2.6.3 Saving plots produced in R

Plots were saved in two ways, depending on whether for looping was employed or not.

1. When using for looping, images were recursively saved within the for loop using the jpeg() function, where the file deposit information of each image is written within the parentheses.
2. For all other plots, they were loaded into the plot area of R, where each image was exported by clicking “export” followed by saving as an image to the desired folder location.

#### 2.2.6.4 R packages

All packages were installed from the Comprehensive R Archive Network (CRAN) repository using the "install.packages("package.name")" function. All tests conducted were using each package's associated Vignette document, which outlines the uses and features of each package. A list of packages used and the tools used within each package can be found in **Table 2.**

**Table 2 Applied R packages and functions**

| R Package | Functions | Purpose |
|---|---|---|
| corrplot | corrplot | Correlation matrix analysis |
| fisher.test | fisher.test | Fisher's Exact Test |
| survival | Surv, survfit, survdiff, coxph, cox.zph, survSplit | 1-3 = Kaplan Meier analysis, 4 = Cox PH model, 5-6 = Time dependent covariate analysis |
| forestmodel | forest_model | Forest plot of multivariate Cox PH models |
| forestplot | forestplot | Forest plot of combined features from multivariate Cox PH models |
| randomForestSRC | rfsrc | Random Forest Survival analysis |
| party | ctree | Recursive partitioning of Random Forest model |
| ipred | gg_minimal_vimp | Minimal depth vs. VIMP analysis from Random Forest model |

### 2.2.7 IBM SPSS

Pairwise Kaplan Meier analysis was conducted in SPSS using version 23. A dataframe was loaded into the data view without headers, and in the variable view the header and variable information was input. All variables were described as nominal. Pairwise Kaplan Meier plots were produced in the following manner:

1. Analyse > Survival > Kaplan Meier options were selected to open the Kaplan Meier analysis window
2. Months since Randomisation was placed in the Time box, whilst the censor variable was set in the Status box with event defined as a single value of 1
3. The variable being tested was placed in the Factor box
4. In the Compare Factor options, Log Rank was selected, followed by selection of Pairwise over strata
5. In the options window, Survival Table(s) was de-selected, and Survival plots selected, following which the plot was produced.
6. Pairwise table data was extracted for analysis

### 2.2.8 Microsoft Excel

Further to the raw data analysis described above, Microsoft Excel was used for chi-squared contingency tables, and calculation of the false discovery rate of methods involving multiple hypothesis testing.

#### 2.2.8.1 Chi-squared Contingency Tables in Excel

1. The observed values of each test were input into the appropriate sized table (2x2, 2x3), with the sum of each column or row combination calculated using the (SUM=) function.
2. The expected values of each test were calculated by multiplying each representative column value by its associated row value, then dividing the total sum of all values within each test.
3. Finally, the function =CHISQ.TEST was used using the observed and expected values to generate a *P* value, with any value $P<0.01$ considered highly significant, and $P<0.05$ significant.

#### 2.2.8.2 False Discovery Rate (FDR) in Excel

1. All the *P* values from the relevant statistical tests (generated from Chi-squared or Fisher's exact tests) were tabulated and sorted from smallest to largest, and given a ranking value
2. Using the Benjamini Hochberg method, the *Q* value was calculated, by dividing the rank of each *P* value by the total sum of tests, followed by multiplying by the set *P* value significance threshold of $P = 0.01$.
3. Each pair of *Q* and *P* values were tested to check whether *Q* was larger than *P*, with only $Q>P$ held as highly significant, with a secondary threshold of $P<0.01$ used for other interactions

#### 2.2.8.3 2-drug fractional method

Originally devised to assess the synergistic interactions between chemotherapeutic compounds in Acute Lymphoblastic Leukaemia (ALL) samples *in vitro,* the 2-drug fractional method is a simplistic way of calculating the observed and expected contributions of combination treatments (157). This method was used to assess whether drug combinations *in vitro* were synergistic or additive, and was achieved using the following steps:

1. The observed normalised viable cell percentages for mono and combination treatments were calculated as described in **(2.1.17.1)**
2. The expected combination viable cell % was calculated by multiplying both mono treated viable cell % values, then dividing this value by 100
3. Combination treatments were considered synergistic when the observed viable cells % value was lower than the expected viable cells % value
4. These values were exported to GraphPad Prism for graphical representation

### 2.2.9 GraphPad Prism

All graphs were produced in GraphPad Prism version 6. Statistical differences between groups were evaluated by the Student's t-test or ANOVA when samples were normally distributed. Correlation analysis was carried out using Spearman test.

# Chapter 3: The mutational landscape of the LRF UK CLL4 trial cohort

The resolution at which we can analyse the cancer genome has been greatly enhanced by the application of NGS technologies (158), allowing the study of subclonal architecture and clonal evolution of cancer patients. Whole-genome/-exome sequencing, as well as targeted re-sequencing panels have been applied to CLL, identifying many recurrently mutated genes across institutional and clinical trial cohorts (31,32,75,47,96,159,160). These include genes already identified using classical sequencing technologies (*TP53* and *ATM*), as well as the identification of additional recurrently mutated genes, such as; *SF3B1, NOTCH1, BIRC3, POT1, EGR2, RPS15* and *SAMHD1* (**Figure 16**; COSMIC v79**)**. Many of these studies have been conducted in institutional cohorts which are not homogenously treated nor composed entirely of patients requiring therapy at study entry (features that are pre-requisites in clinical trials) (**Figure 17**; 39/48 studies included in PubMed literature search analysis). This is important, because the protracted natural history of CLL, as well as the "watch and wait" clinical management until treatment requirement, means that analysis of these features in the context of patients requiring therapy is required to understand the true impact of recurrently mutated genes on prognosis. Therefore, when assessing the potential of predictive or prognostic biomarkers, clinical trial cohorts are more appropriate.

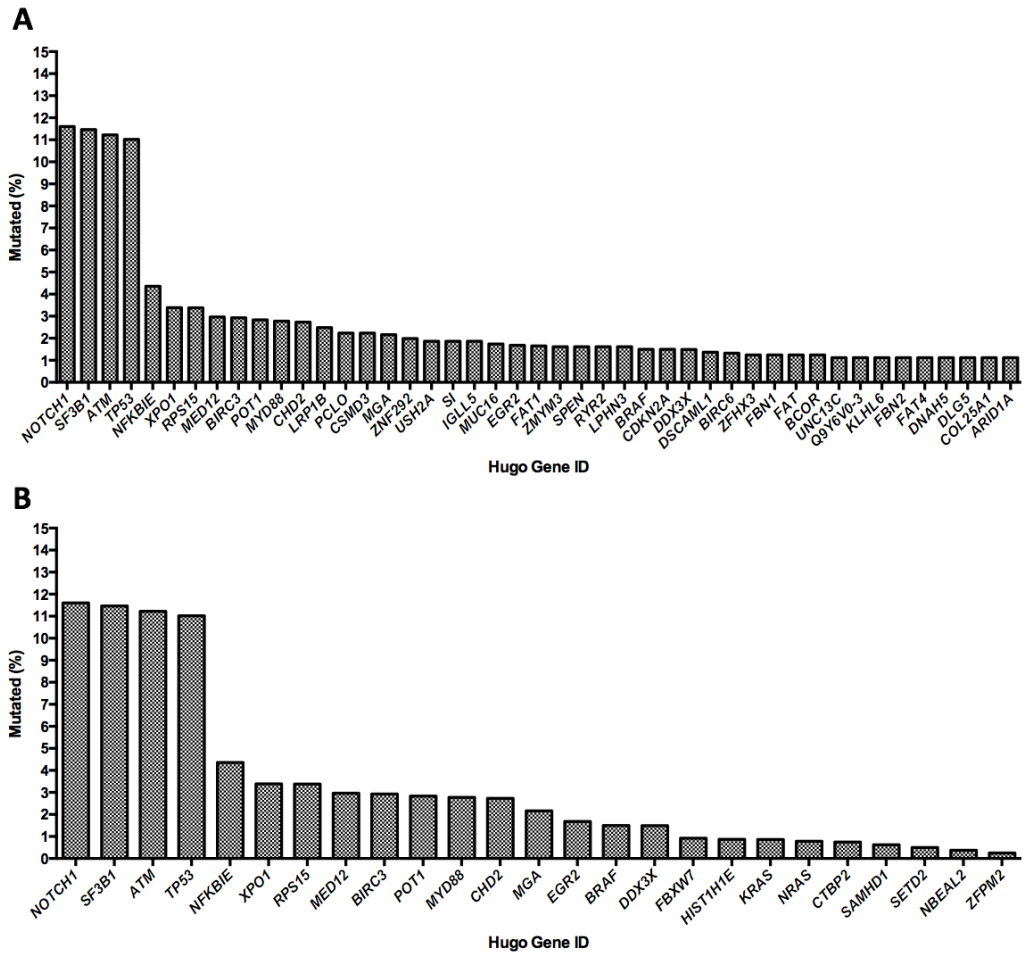



**Figure 16 COSMIC CLL recurrently mutated gene frequencies**

**A** Recurrently mutated genes reported in COSMIC v79 for CLL with a frequency of more than 1%. **B** The mutation frequency reported in COSMIC v79 of the 25 CLL mutated genes included in the CLL4 TruSeq study.

Of the NGS studies using clinical trial cohorts, the majority have been conducted using targeted re-sequencing (**Figure 17;** 8/9 clinical trial sample publications included in PubMed literature search analysis). However, except for the studies pertaining to CLL4 (data for these publications contributed by the PhD candidate), the clinical trials used have had relatively short follow up for OS (mean longest follow up time across five studies [WES n=1, targeted n=4] of 67.2 months), although the therapies used do consist largely of modern frontline regimens (WES; FC +/- Rituximab [CLL08](96), targeted; FC +/- Campath [HOVON 68](161), Chl + Gazya/Rituximab [CLL11](162), BOMP [ICLL01](163), FCM +/- R, F + Campath [UK NCRN CLL201 & CLL202](163), Binet Stage A observational [O-CLL1](164,165)). Despite these studies in clinical trials, further information concerning the clinical application of gene mutation data of patients requiring therapy is required. To assess the long-term impact of recurrently mutated genes in the clinical trial setting, the LRF UK CLL4 clinical trial (CLL4)(67), which compared the treatment of chlorambucil against fludarabine with or without cyclophosphamide, was selected for mutation screening by NGS as it has more than 150 months of survival follow up data. Furthermore, it has a large array of clinical and biological data either taken at trial entry, as well as from additional studies conducted on stored clinical samples.

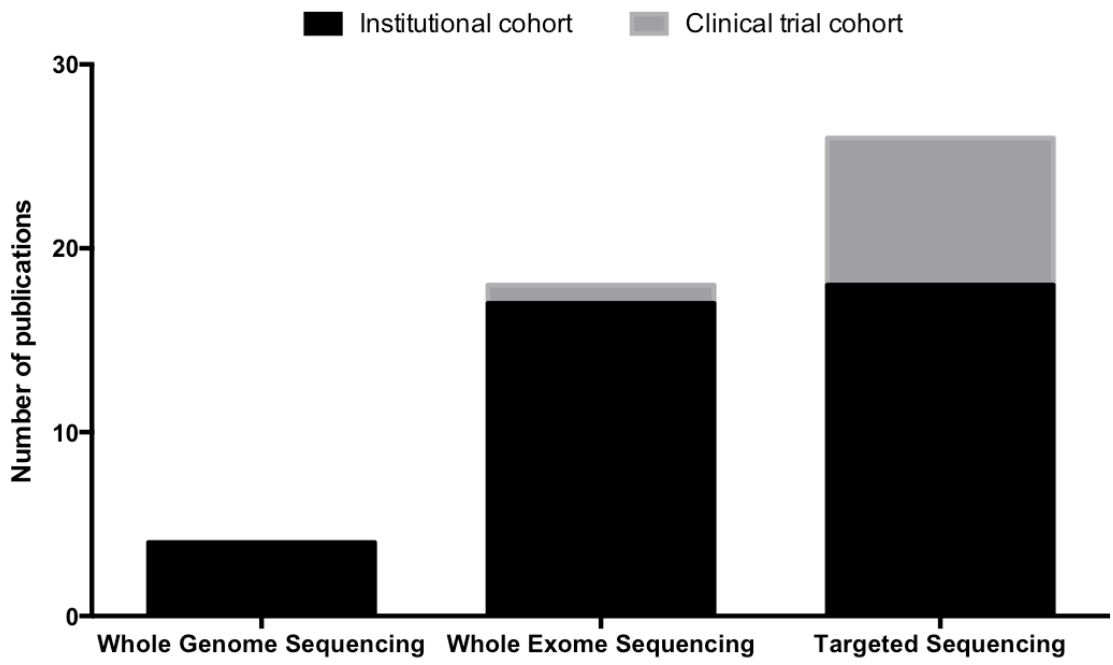



**Figure 17 Number of NGS CLL publications in PubMed separated by NGS technique**

Number of CLL genomics research articles in PubMed using each NGS technology as their primary sequencing output. Search terms used in PubMed: “CLL next generation sequencing”, “CLL targeted sequencing”, “CLL whole exome sequencing” and “CLL whole genome sequencing”. Black bars represent research articles using CLL patients not enrolled onto clinical trials, grey bars represent research articles using CLL patients enrolled onto clinical trials for at least part of the analysis. Articles included in the analysis n=48(15,31–33,57,75,47,96,254–261).

The targeted re-sequencing studies carried out in clinical trial cohorts range from targeting a single mutation (*NOTCH1* P2514fs) to a panel containing more than 80 genes(161–164). Furthermore, these study's research objectives are diverse, with:

1. Identification of cases mutated in multiple clinically important genes linked with survival (163)
2. Characterisation of the frequency of clonal and subclonal events in key genes (161)
3. Analysing gene mutations in the context of a complex karyotype and survival (162)
4. Assessing multiple NGS mutation techniques association with *IGFR1* expression in the *NOTCH1* P2514fs studies (164)

The intratumoral heterogeneity of CLL is a theme throughout these publications, however limited analysis regarding either the impact of clonal and subclonal mutations on prognosis, nor inference of clonal evolution patterns. Although well characterised work on subclonal mutations in a subset of the CLL8 trial has been presented using WES, there are only targeted re-sequencing studies that have investigated the clinical importance of subclonal mutations in CLL in non-randomised institutional cohorts(8,159). This highlights the need for targeted re-sequencing NGS studies in the context of clinical trials.

Additionally, many studies have reported potential genomic/epigenetic/clinical biomarkers in CLL. However, beside CNVs, diagnostic screening for mutations in *TP53* remains the only genomic biomarker from these studies to have been included in the iwCLL clinical guideline recommendations(166). This suggests that further studies are required to validate the importance of these additional prognostic or predictive biomarkers. Due to the extensive number of studies conducted on CLL4 (n=29 at the time of writing), and the large amount of clinical biomarker and long-term follow up data available, targeted sequencing of the CLL4 cohort is well suited to contribute to understanding the importance of other genes in CLL therapy response predication and disease prognostication.

This chapter describes the screening of 544 patient samples from CLL4, presenting mutation data on 500 cases for a panel of 25 genes using the Illumina TruSeq targeted re-sequencing platform, that have been previously shown to be mutated in CLL **(Figure 16B)**, with various levels of biological and clinical significance. The sequencing coverage achieved in this cohort was analysed, identifying genomic regions of good, poor, and dichotomous read depth. Next, assessment of the impact of missing data on clinical statistical analysis, using *TP53* as an example. The bioinformatic pipeline, as well as the *in silico* and molecular validation strategy to filter a list of 903 high confidence variants is described, along with comparative concordance analysis of the mutations observed between this study and previous CLL4 publications in recurrently mutated genes in CLL:

*TP53* (167)*, SF3B1* and *NOTCH1* (7). Following which, extended characterisation of the mutation landscape in CLL4 is presented, stratified by IGHV mutational status, as well as *SF3B1* mutational status owing to *SF3B1* mutations being the most recurrent in this cohort, occurring in just under 25% of all cases. Additionally, new recurrently mutated genes were identified, which previously were only found to be mutated in a handful of cases. Co-occurrence analysis was conducted at the gene and intracellular signalling pathway level, identifying established and novel associations. The subclonal architecture of the CLL4 cohort was explored, inferring the clonal evolution patterns and identifying established and novel interaction partners. Altogether, this chapter provides the framework for **Chapter 4:**, which aims to test the clinical importance of recurrently mutated genes in CLL at the clonal and subclonal level.

To summarise, this chapter aims to discuss the following:

1. Design and preparation of the CLL4 TruSeq study
2. Sequencing coverage and quality of the CLL4 TruSeq study
3. Bioinformatic and molecular strategies to identify high quality disease associated mutations
4. Comparative analysis of CLL4 TruSeq study observed mutations in: *TP53, SF3B1,* and *NOTCH1* with the relevant previous CLL4 clinical trial literature
5. Description of the observed mutation landscape across the 25-gene panel of the CLL4 TruSeq study
6. Statistical association analysis of the entire CLL4 TruSeq study, as well as analysis of statistical associations in clinically important CLL subsets
7. Characterisation of the subclonal architecture of the CLL4 TruSeq study

## 3.1 CLL4 TruSeq gene panel design

### 3.1.1 Initial design of the CLL4 TruSeq panel

The design of the CLL4 TruSeq gene panel was conducted using Illumina DesignStudio. Candidate genes were stratified into six categories based on their previously reported mutation frequency, predictive/prognostic biomarker evidence and supporting biological data in the literature. Decisions on gene inclusion were also made based on amplicon capture efficiency across the panel from *in silico* analysis. Priority 1 genes were selected as being functionally and prognostically important, with the clinical importance of sub-clonal mutations and correlation with biomarkers in a clinical trial setting withstanding. Priority 2 genes were those shown to be functionally important in diagnostic gene discovery cohorts, but have yet to be assigned prognostic utility. Priority 3 genes have been reported in the literature but their frequency and clinical importance is yet to be fully elucidated, with priority 4-6 genes being of shared research interest between Southampton and the collaborative sequencing site (Oxford National Institute for Health Research Biomedical Research Centre/Molecular Diagnostic Centre, University of Oxford). **Table 3** represents the final iteration of the 25-gene panel, including which exons were included in the design and the priority status of each gene (the manifest file for Design 1 as a representative example can be found in the appendix **A.1**).

### 3.1.2 Re-designs of the CLL4 TruSeq panel

During the sequencing stage of the study, it was identified that the coverage performance in certain regions was sub-optimal **(3.3)**. A review of the design with Illumina led to the identification of a technical error, with SNP sites excluded from the design process (post capture efficiency analysis), which lead to specific amplicon regions missing their desired target. Therefore, this error was corrected for future amplicon pools. Additionally, our collaborators made slight adjustments to the final combination of pooled amplicons (i.e. the exact number of primers targeting the genomic sequence), owing to striving for improved coverage performance for all amplicons. The nomenclature used below describes each design chronologically, with Design 1 pertaining to the initial amplicon pool with the SNP error, Design 2 the first re-design, and Design 3 the additional capture efficiency adjustments.

**Table 3 CLL4 TruSeq Panel Design, Coverage, and Priority Gene Status**

| Hugo Gene ID | Exons in Design | Priority Gene List Status |
|---|---|---|
| *ATM* | All exons | 1 |
| *BIRC3* | All exons | 1 |
| *NOTCH1* | 33 and 34 | 1 |
| *SF3B1* | 10 to 20 | 1 |
| *TP53* | All exons | 1 |
| *MYD88* | 2 to 5 | 1 |
| *EGR2* | All exons | 2 |
| *NFKBIE* | All exons | 2 |
| *POT1* | 5 to 19 | 2 |
| *SAMHD1* | All exons | 2 |
| *BRAF* | 2 to 18 | 2 |
| *FBXW7* | 5 to 12 | 3 |
| *XPO1* | 10 to 19 | 3 |
| *CHD2* | All exons | 3 |
| *DDX3X* | All exons | 3 |
| *MED12* | 1 to 4 | 3 |
| *SETD2* | All exons | 4 |
| *RPS15* | All exons | 4 |
| *CTBP2* | All exons | 4 |
| *MGA* | 16 to 18 | 4 |
| *NBEAL2* | All exons | 4 |
| *KRAS* | All exons | 4 |
| *NRAS* | 2 to 5 | 4 |
| *HIST1H1E* | All exons | 5 |
| *ZFPM2* | All exons | 6 |

## 3.2 Sample curation, library preparation and sequencing

An audit of available DNA, cell pellets, and viable cells was conducted in the Cancer Genomics group laboratory, followed by additional audits at the Royal Bournemouth Hospital and the primary CLL4 trial repository centre; the Institute of Cancer Research in Sutton. This led to the identification of 595 samples from the first randomisation of the trial. DNA quantification, library preparation and sequencing was conducted in Oxford, including an 8-week placement during which the PhD candidate contributed to all stages of the process, with 20% of the library preparations (n=119) achieved during this placement. DNA quantity was assessed post DNA extraction using the Qubit BR quantification system **(2.1.6.3)**, with samples taken forward to library preparation if they met the minimum quantification requirements for low (50ng) or high (250ng) input **(2.1.12)**.

Post-quantification, 36 cases were removed due to low concentration, leading to the preparation of libraries for 559 samples using either the low input or high input protocol **(Figure 18)**. Ten out of the 559 prepared libraries did not sufficiently amplify, 9/10 of which were prepared using the low input protocol. This led to 549 prepared libraries taken forward for MiSeq sequencing, in maximum batch sizes of 20 per MiSEQ run, to allow for sufficient coverage across all cases based on the average sequence output of a MiSEQ cartridge. Twenty was selected as the optimum number of cases based on the read depth and coverage required per sample, and the cost of sequencing. Post-sequencing, 5/549 were determined to have failed, with 100% of amplicons in these five cases below 100 reads. Therefore, 544 cases were taken forward for coverage analysis (full CLL4 trial n=777).

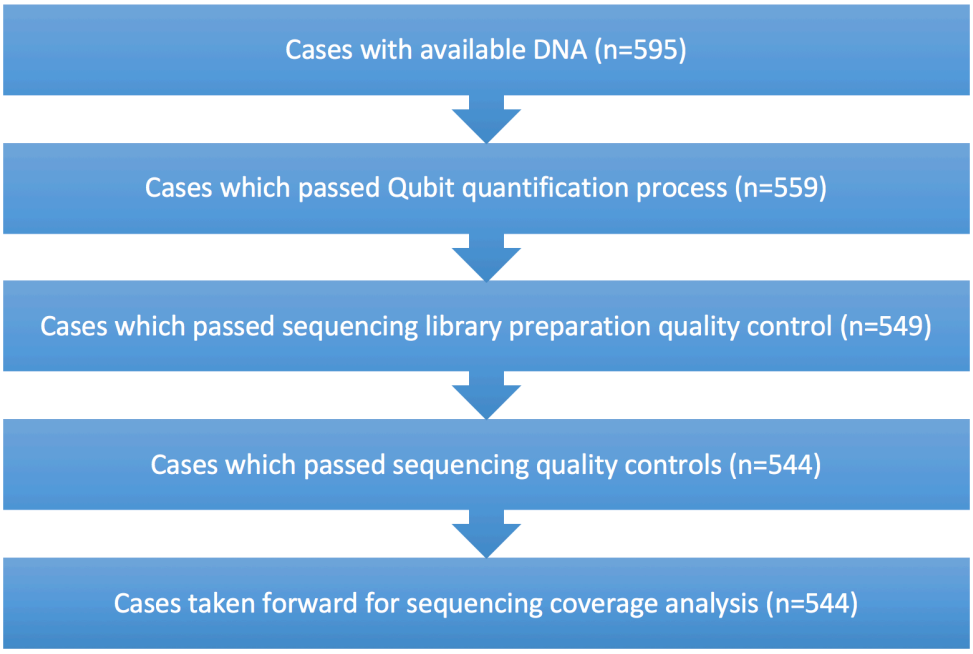



**Figure 18 Flow diagram of the sample sequencing process**

Flow diagram representing the number of cases taken forward from each step in the data collection process.

## 3.3 CLL4 TruSeq study sequencing coverage analysis

The identification of high confidence variants requires adequate and consistent sequence coverage across the gene panel. Therefore, the uniformity in sequence coverage was tested, by analysis of the raw sequencing coverage data throughout all samples sequenced across the three panel designs. Mean coverage statistics, amplicon correlation matrix analysis, sequencing coverage per gene, individual amplicon performance and inference of missing data was all assessed and presented below.

### 3.3.1 Identifying the distribution of sequencing coverage in the CLL4 TruSeq study per MiSEQ run

The first comparison was to assess the general performance of each MiSEQ run (n=28), as this is important to understand whether the data is impacted by batch effects. Batch effects are defined by inherent differences between sequencing runs (or batches), in terms of coverage, that may have a downstream impact on the results. Calculating the mean coverage per run is a broad way of identifying batch effects, since a lack of consistency in sequencing coverage within a sequencing run and across all runs will highlight discrepancies in data quality.

Mean run statistics were calculated, using the proportion of amplicons with a read depth above 100 (>100DP). A depth of one-hundred reads was set as the cut off for this analysis, since, except in specific circumstances, mutations were only considered if they had a read depth of more than 100 reads. The mean percentage of amplicons with more than 100 reads across all 544 cases was 94.38% (Design 1 = 91.38%, Design 2 = 96.06%, Design 3 = 95.96%). Most the runs had a mean of >90% (25/28), with all runs achieving more than >70% of amplicons having a read depth of more than 100 **(Figure 19)**. This means that at a broad level, the sequencing coverage per run was very good, but as can be observed in **Figure 19**, there was also specific runs either with consistent but much reduced sequencing coverage (**Figure 19,** 14MiSEQ099), or high variability within the run itself (**Figure 19**, 15MiSEQ066). Therefore, inference of the performance of the sequencing coverage at the amplicon level is required.

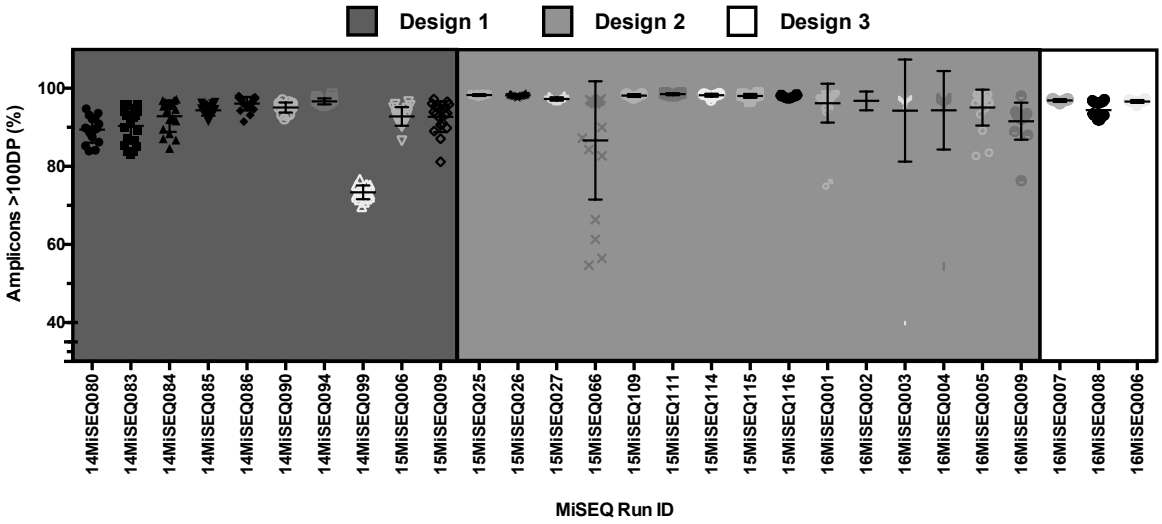



**Figure 19 Mean amplicon coverage distribution with more than 100 reads**

Percentage of amplicons with more than 100 reads, per case per run. Dark grey represents runs conducted using Design 1, light grey Design 2, and white for Design 3. Bars represent mean within each run and standard deviation.

### 3.3.2 Correlating the mean sequencing coverage between MiSEQ runs at the amplicon level

Assessing the sequencing coverage data distribution of each MiSEQ run is a good way to identify whether they are consistent with each other or not, thereby providing some general insights into data variability. However, this type of analysis does not offer any information on how each of the MiSEQ runs differ from each other. Therefore, assessment of the correlation between the average sequencing depth of each amplicon in each MiSEQ run and per CLL4 TruSeq study design will offer an insight into these potential differences.

To further analyse the sequencing coverage data, correlation matrix analysis using Spearman Rank **(2.2.6.4)** was conducted at the amplicon level. Prior to analysis, the mean coverage per amplicon was calculated across each MiSEQ run, followed by calculation of the GC content in percentage terms of each amplicon, as well as the length of the design primers and target regions. The GC content was calculated since it is known that Illumina sequencers can have sequencing bias in GC regions (168), whilst the target region length was included as read length has been reported to correlate with sequencing quality (169).

Designs 1 and 2 showed significant batch effects, as the MiSEQ runs clustered into two distinct groups based on average read depth per amplicon, per run (Design 1: Cluster 1 = 3 runs, Cluster 2 = 5 runs, Intermediate = 1 run. Design 2: Cluster 1 = 9 runs, Cluster 2 = 5 runs, Intermediate = 1 run. Design 3: Cluster 1 = 3 runs) **(**representative example **Figure 20** [Designs 2 and 3 in **A.2** and **A.3**], overview of all designs **Figure 21A)**. Cluster 1 MiSEQ runs were defined as those with high levels of similarity that had been hierarchically clustered to the top of the corr plot. Cluster 2 MiSEQ runs represent the next group of MiSEQ runs with high levels of similarity, but whose average amplicon coverage profile is distinctly different from Cluster 1 MiSEQ runs. Intermediate clusters define specific runs which share features of both Cluster 1 and Cluster 2, since they were placed in between each cluster in the corr plot hierarchical clustering procedure **(Figure 20A)**. Correlation matrix analysis and the identified correlated clusters do not represent a direct identifier of “good” vs. “poor” data quality, they are instead an indicator of the consistency of the library preparation and sequencing processes, and offer an insight into how comparable the data within each MiSEQ run is to every other MiSEQ run of the same design.

As previously stated, re-designs were conducted to improve data capture **(3.1.2)**. Design 2 was found to have significantly reduced correlation between amplicon read depth and GC content in relation to Design 1 ($P$<0.001), representing an increase in data quantity and a reduction in the impact of GC content on sequencing (**Figure 21B**; right). However, this re-design had a negative impact on the correlation between amplicon read depth and target length in comparison to

Design 1 ($P$<0.05), suggesting that multiple factors are at play regarding optimal targeted re-sequencing coverage (**Figure 21B**; left).

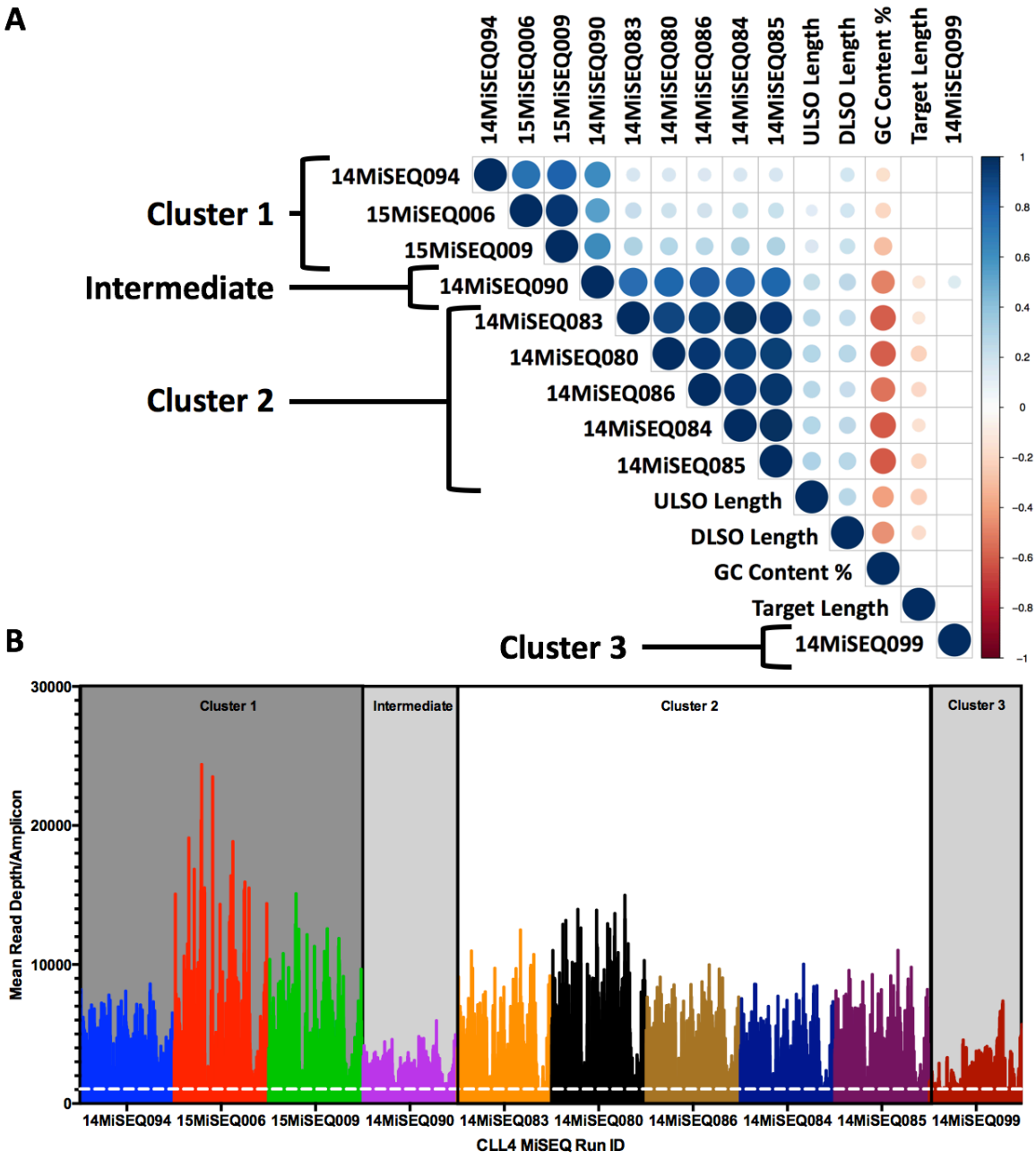



**Figure 20 CLL4 TruSeq Design 1 correlation coverage analysis**

**A** Correlation matrix with hierarchical clustering of CLL4 TruSeq Design 1 MiSEQ runs (n=10) read depth per amplicon (n=610), including ULSO, DLSO and target length, as well as GC content % per amplicon. Blue represents positive correlations (0 to 1), red represents negative correlations (0 to -1). Size of circle represents strength of correlation score. Non-significant correlations (p=>0.01) excluded from correlation matrix. Clustering of MiSEQ runs based on similarity denoted by open brackets. **B** Histogram denoting the read depth per amplicon, per MiSEQ run (n=10) for CLL4 TruSeq Design 1. Histogram ordered based on correlation matrix hierarchical clustering. Design 2 & 3 corr plots and read depth per amplicon in **A.2** and **A.3**

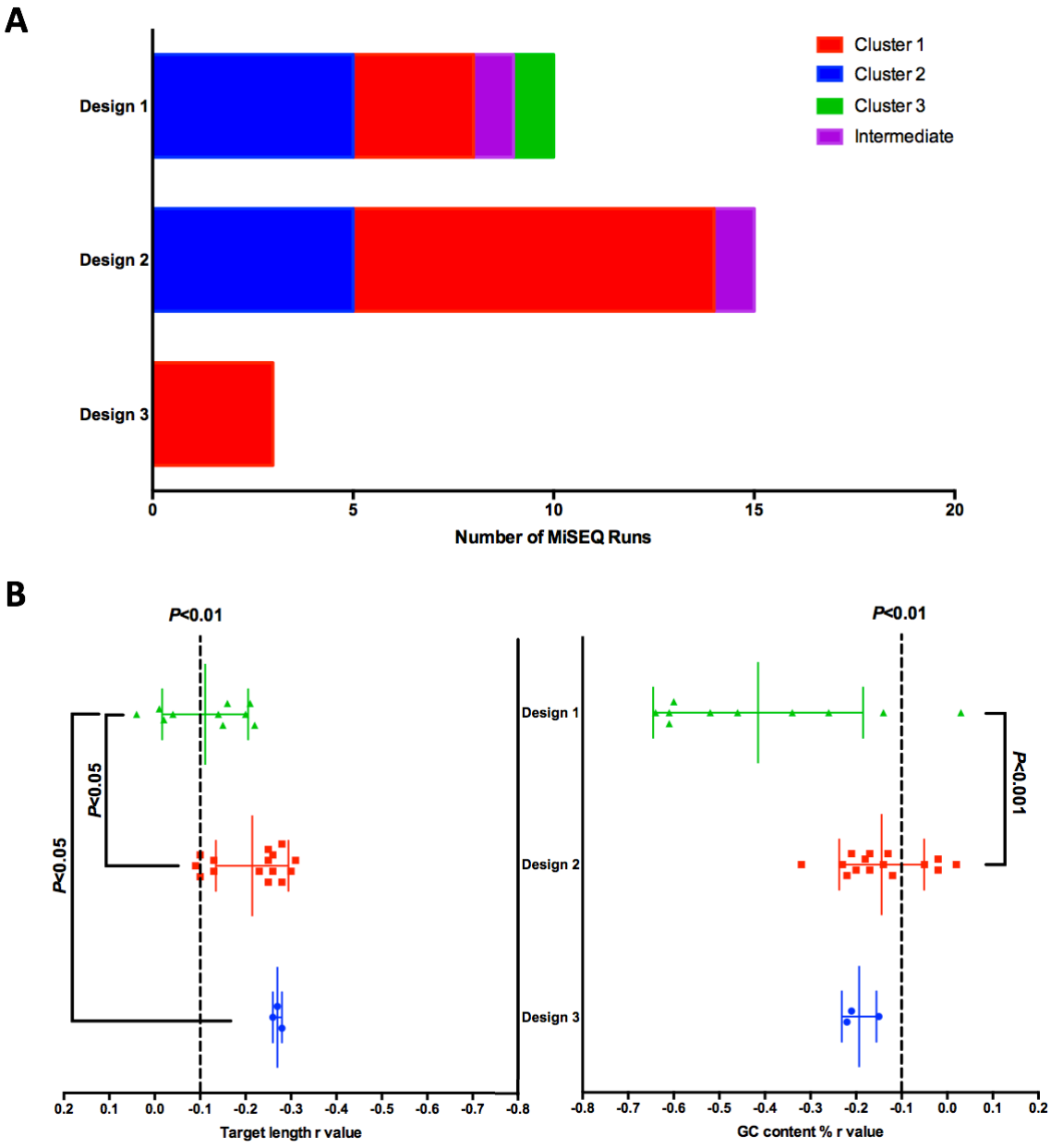



**Figure 21 Overview of sequencing coverage correlation matrix analysis**

**A** Number of MiSEQ runs per design which positively correlate with each other, with cluster number ordered by hierarchical clustering. Intermediate MiSEQ runs shared between clusters. See Appendices for corr plots of each design **(A.2, A.3)**. **B** Correlation between MiSEQ runs and target length (left) or GC content % (right) stratified by design. Raw r values from each MiSEQ run plotted, vertical dotted line denoting whether a correlation is significant ($P$<0.01). Two-way ANOVA with multiple comparisons used to compare between designs.

### 3.3.3 Identifying the distribution of sequencing coverage in the CLL4 TruSeq study per gene

To better understand why runs cluster into distinct groups within each design, mean coverage per gene analysis was undertaken. This is important, because the genomic sequence context (such as GC content) is gene specific, therefore identifying whether all genes across designs perform in a similar manner is required. To identify whether certain genes may be accounting for poor read depth performance, and to assess inter-design performance, the mean read depth per amplicon, per gene was calculated for each design.

The sequencing coverage per gene was on average found to be high, with most genes across designs having an average read depth of more than 2000 reads (Design 1 = 17/25 [**Figure 22A**], Design 2 = 21/25, Design 3 = 21/25 [**A.4**]). Across all three designs, *MED12, NFKBIE,* and *NOTCH1* were all shared as genes which had an average below 2000 reads, all genes with high GC content (>60%). In most cases, the change in design did not alter the average sequencing coverage of the gene **(Figure 22B)**. However, Design 2 represents the most stable performing design in terms of average coverage per gene, whereas Design 1 and 3 show more variability between genes. Therefore, to better understand the impact of poor performing amplicons, individual analysis at the per amplicon per patient resolution was undertaken.

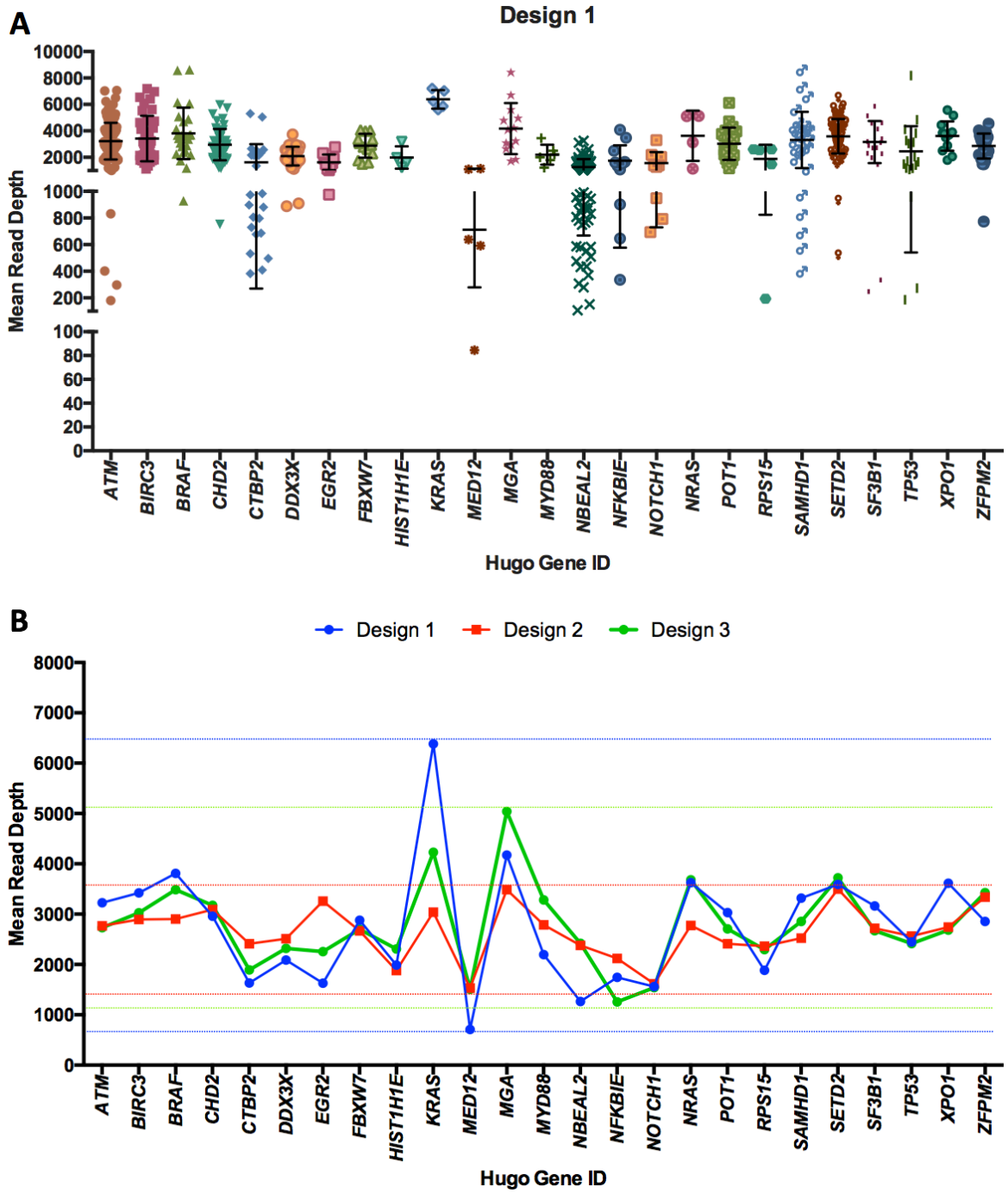



**Figure 22 Mean coverage per gene**

**A** Mean read depth per gene of CLL4 TruSeq Design 1. Bars represent mean ± SD. Designs 2 and 3 can be found in **A.4**. B Mean read depth per gene stratified by CLL4 TruSeq panel design. Blue represents Design 1, red Design 2 and green Design 3. Thin dotted lines represent the minimum and maximum average values across the design. See Appendices for distribution per design per amplicon **(A.4)**.

### 3.3.4 Individual amplicon performance analysis per patient

The above analysis highlights the run and design coverage variability, with high levels of mean coverage, but highly variable features based on per design and per gene analysis. However, these approaches do not give an insight into the individual performance of each amplicon. This is important, since identification of poor performing amplicons has potential downstream applications for mutation calling.

To achieve this, each case/amplicon value in every MiSEQ run was given a score based on multiple read depth (DP) cut offs (≤50DP = -2 [bright blue], ≤100DP = -1 [dull blue], ≤500DP = 0 [black], ≤1000DP = 1 [dull yellow], >1000DP = 2 [bright yellow]). These cut offs were set to represent amplicons where:

1. Mutation data was not discernible (≤50DP)
2. Only clonal mutation data was discernible, since the poor coverage was associated with high GC content, and the amplicon covers a hot spot CLL mutation (51-100DP)
3. Amplicon performance was sufficient to capture mutations down to 2.8% VAF (101-500DP),
4. Amplicon performance was sufficient to capture mutations down to 2.4% VAF (501-1000DP)
5. Amplicon performance was sufficient to capture mutations down to 1% VAF (>1000DP)

These dataframes were then visualised using hierarchically clustered heat maps, followed by the identification of amplicons shared within distinct clusters **(2.2.2)**. This was conducted to easily identify which amplicons are featured within each cluster.

The largest variability in amplicon performance was found within Design 1, and it also contained the highest proportion of amplicons with a read depth >1000DP **(Figure 23A)**. Data from Designs 2 and 3 show that the subsequent panels became more consistent, at the expense of the number of >1000DP performing amplicons **(Figure 23B&C)**. However, there was also a reduction in the number of poor performing amplicons (<50DP) in these designs in comparison to Design 1. Across all three designs there were amplicons which were consistently good (dull/bright yellow across all cases), consistently poor (dull/bright blue or black), and dichotomous in their sequencing coverage (yellow and blue). Therefore, the proportion of amplicons falling into each group per gene were calculated for each design and analysed by two-way ANOVA, with multiple comparisons between groups.

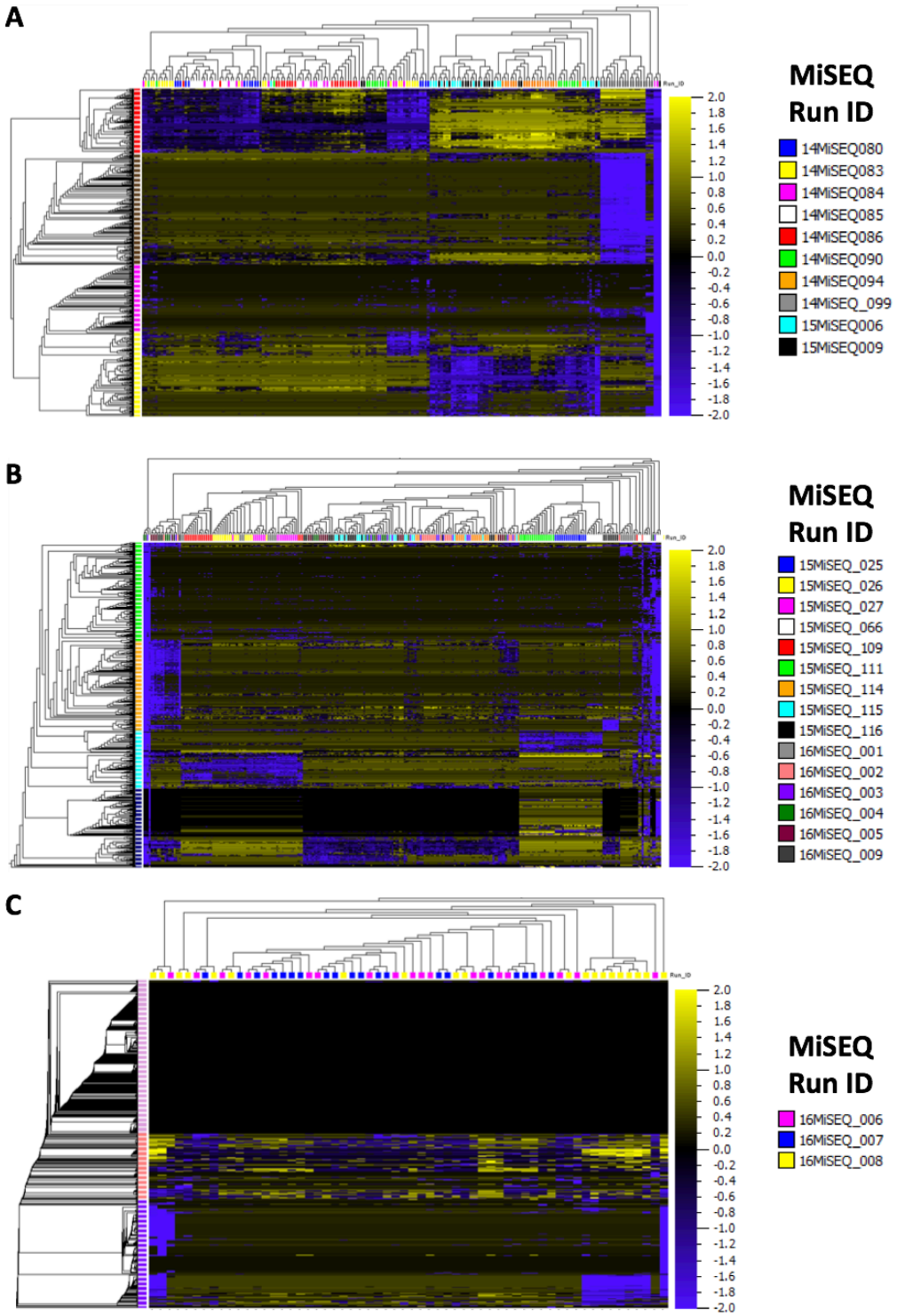



**Figure 23 Heat map analysis of coverage data**

Visual representation of sequencing coverage data per amplicon, per patient for Design 1 **(A)**, Design 2 **(B)**, and Design 3 **(C)**. ≤50DP = -2 [bright blue], ≤100DP = -1 [dull blue], ≤500DP = 0 [black], ≤1000DP = 1 [dull yellow], >1000DP = 2 [bright yellow]. Colours along the top of the heat map define which MiSEQ run each patient was sequenced in (colour legend to the right of each heat map). Colours along the left of the heat map denote the visually identified clusters of amplicons hierarchically clustered into similarly performing groups.

Only 5 genes were found to have significantly different distributions of amplicons in each respective cluster across all three designs, with *ATM* in designs 1 and 2, *NBEAL2* in designs 1 and 3, and *CHD2* as well as *SETD2* in design 3. For *ATM* in designs 1 and 2, cluster 2 had significantly more *ATM* amplicons in comparison to cluster 1 (Design 1: $P<0.0001$, Design 2: $P = 0.0029$), cluster 3 (Design 1: $P<0.0001$, and Design 2: $P<0.0001$), cluster 4 (Design 1: $P<0.0001$, Design 2: $P<0.0001$). This means that most *ATM* amplicons in designs 1 and 2 had >1000DP coverage. *NBEAL2* amplicons were significantly clustered in cluster 1, in comparison to both cluster 2 and cluster 3 in design 1 (Cluster 2: $P = 0.0275$, Cluster 3: $P = 0.0275$), and cluster 3 in design 3 (Cluster 3: $P = 0.0011$). *CHD2* and *SETD2* were only found to have significantly different distributions of amplicons in design 3, with *CHD2* having significantly more amplicons in cluster 1 vs. cluster 2 ($P = 0.015$), and *SETD2* having significantly more amplicons in cluster 1 vs. both cluster 2 ($P = 0.0005$), and cluster 3 ($P = 0.0211$).

### 3.3.5 Biological and clinical implications of missing data: Insight using *TP53*

As shown above, there are amplicons which have sequencing coverage that precludes assessment of whether the genomic region it covers contains mutations. This means that using a dataset with missing sequencing coverage data might have significant implications for downstream analysis of the data. For example, if only 90% of designed amplicons covering a gene have a DP of more than 50 reads in a single patient, but no mutations were identified in these regions, does that mean that this patient is wild type for that gene, and if not, does the missing mutation information from the 10% which could not be assessed matter? Therefore, assessing the impact of missing data in this dataset is important.

To understand if the differential coverage of amplicons might have an impact on the inference of clinical associations in this dataset, analysis of the coverage of *TP53* in all cases was conducted. *TP53* was selected since it is the only gene to be included in the clinical guidelines, and represents a group of patients with dismal clinical outcome. Therefore, if missing data does impact characterisation of mutated genes, it will be observable in *TP53.* First, amplicon coverage data across all three designs was aligned with the exon target **(Figure 24A)**. This identified that although most exons had exceptional median coverage (9/11 exons >1000DP median coverage), amplicons in exon 10 had <50DP in 62% of cases.

Next, the number of exon targeting amplicons with less than 100 reads was calculated, to evaluate the extent of missing *TP53* sequencing data. Fifty seven percent of cases (312/544) had one amplicon with less than 50 reads, increasing to 75% of cases that had at least one amplicon with a read depth of less than 100 **(Figure 24B)**. Therefore, it is important to identify whether this lack of complete coverage of *TP53* would affect capturing mutations associated with poor clinical outcome.

Therefore, Kaplan Meier survival analysis stratified by the number of missing amplicons was conducted **(Figure 24C&D)**. This type of analysis compares the survival between different groups in a pairwise manner. Furthermore, this analysis was carried out after excluding all cases with deletions in the short arm of chromosome 17, as well as excluding all cases which have previously been published to have a *TP53* mutation (167), since these cases would confound this analysis.

The cases were stratified into 9 groups based on the number of *TP53* exon targeting amplicons which had a read depth below 100. None of the 9 groups were found to be significantly different to the group with no missing amplicons, suggesting that there was not a significantly impact of missing data on the clinical data **(Figure 24D)**. However, cases with 3 missing amplicons were shown to have a significantly poorer survival than cases with 1 and 2 amplicons missing, respectively. When analysing this group further, and using mutation data presented later, 14/18 cases were found to carry a mutation in *TP53* (not previously published), *ATM, NOTCH1* or *SF3B1*, or a CNV with del(11q) or +(12). Therefore, this provides confidence that the gaps in *TP53* coverage are not having a statistical impact on the application of this sequencing data on survival analysis. Although this analysis was restricted to *TP53*, it has relevance to the rest of the genes on the TruSeq panel. This is because that after accounting for strong predictors of OS, no significant differences were found between cases with various amounts of missing data.

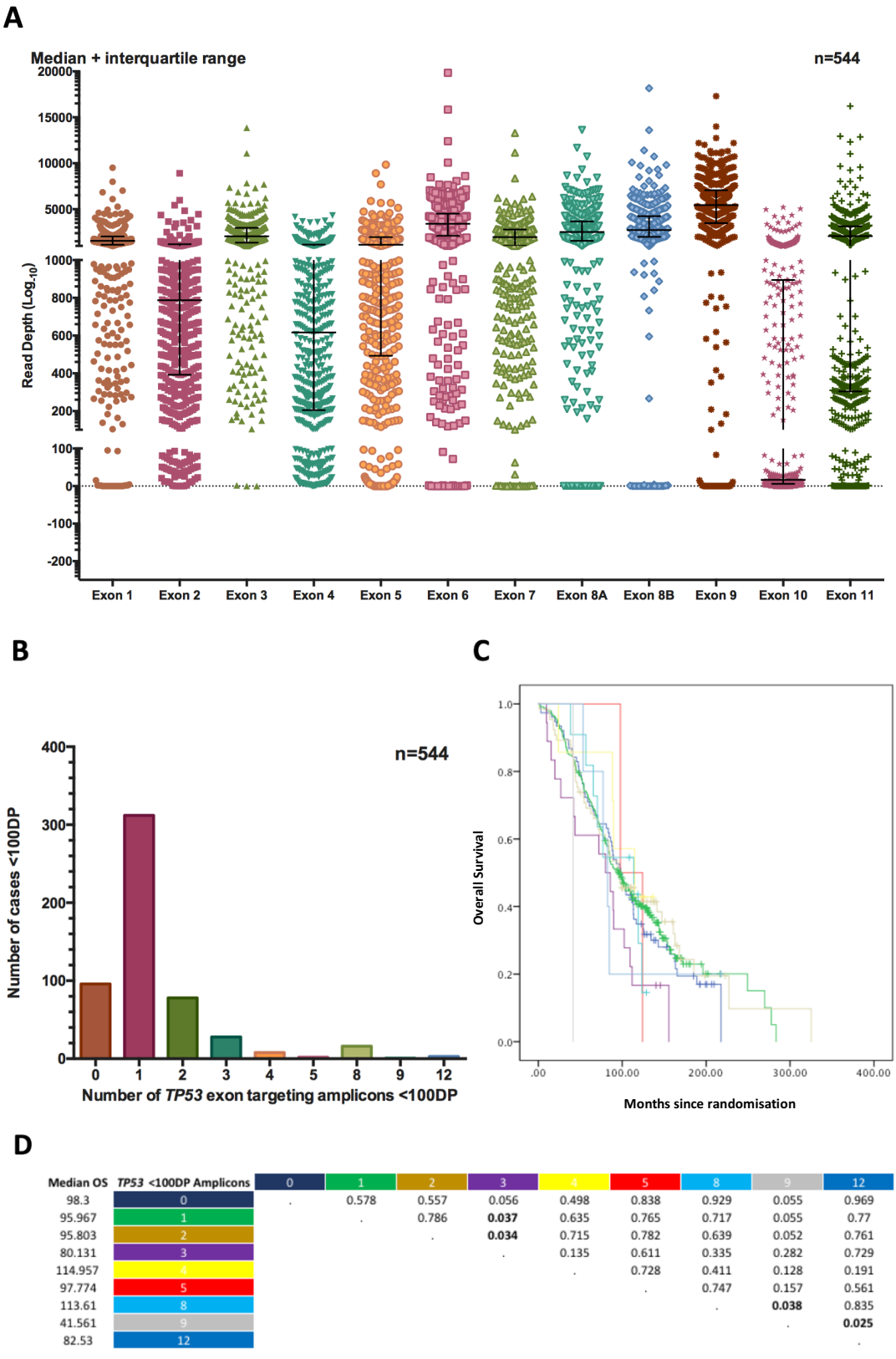



**Figure 24 Clinical analysis of missing data in *TP53***

**A** Coverage of *TP53* per exon across all cases (n=544), bars represent median read depth and the interquartile range. **B** Proportion of cases containing amplicons with <100 read depth. **C** Kaplan Meier for Overall Survival of number of poor performing amplicons, after accounting for *TP53* mutation and del(17p) status. **D** Table of Log Rank *P* values adjusted for multiple comparisons based on the data presented in **C,** with the colours of the box relevant to the colour of the line in the Kaplan Meier graph.

## 3.4 Bioinformatics and mutation data processing of the CLL4 TruSeq study

Having assessed data quality and explored the variability within the sequencing data, defining the bioinformatic steps applied to the raw sequencing data as well as the process of high quality variant selection is important. Furthermore, these steps provide the frame for variant calling, which is a critical feature when performing survival association analysis downstream, as incorrect mutation data processing and bioinformatics could confound these associations.

Variant calling and annotation was conducted as described **(2.2.3)**. As previously mentioned, mutation data was only presented on 500 cases, since the pipeline had only been completed on that many cases at the time of writing. This section reports the strategy employed to identify true positive and high confidence variants. This was achieved using additional variant filtering, extensive manual curation, molecular validation/orthogonal sequencing and clinical variant selection. An over view of the mutation data processing strategy, and the number of variants at each stage is presented below, and will be referred to throughout this section **(Figure 25)**.

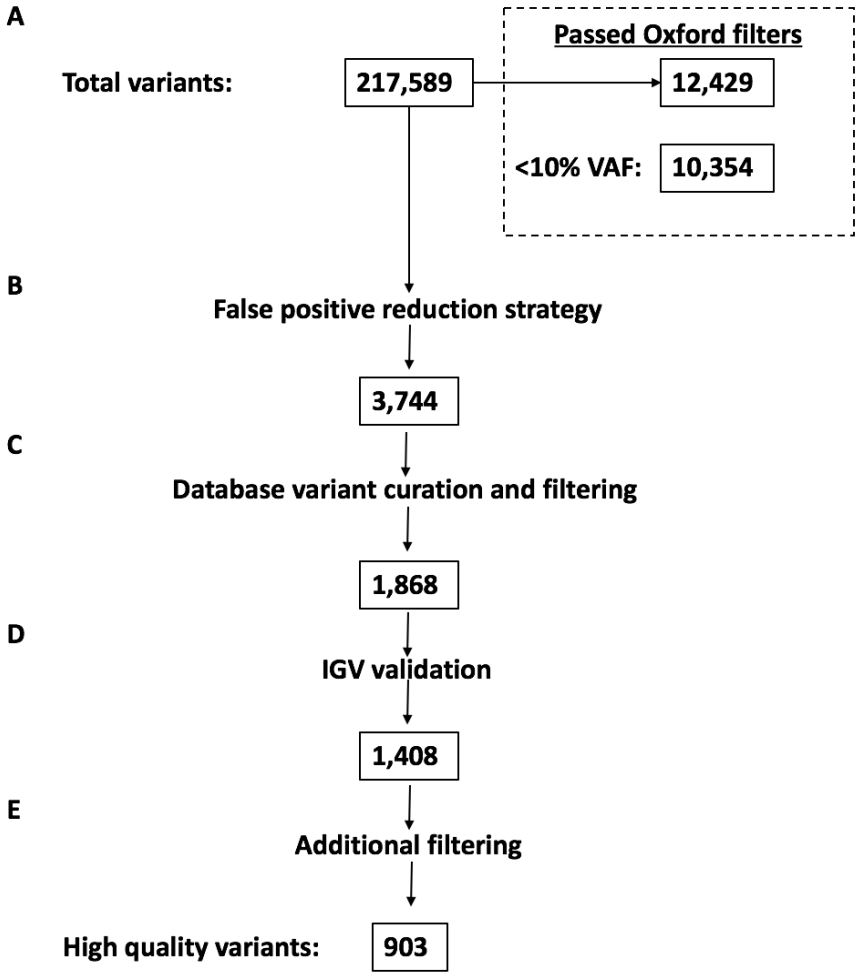



**Figure 25 Flow diagram output of variant filtering strategy**

**A** Total number of variants identified without bioinformatic filtering. Hashed box represents the filtering applied in the first instance. **B** Number of variants left after removal of recurrent false positives. **C** Number of variants left after curating for mutations with known impact in CLL or predictive evidence of pathogenicity. **D** Number of variants which were classified as real variants after visualisation of each variant in IGV. **E** Final number of variants included for downstream analysis of molecular validation, orthogonal sequencing and additional filtering.

### 3.4.1 Handing of variant call files (.vcf)

In the 500 cases, 217,589 annotated variants were identified before filtering. Of which, 12,429 variants passed the initial filtering strategy from the bioinformatics pipeline, with 10,354 variants having a VAF of less than 10% **(Figure 25A)**. This represents a mutation burden of 25 variants per case across the 25 genes covered by the panel. When considered in the context of the low mutation burden (0.5/Mb) previously observed in CLL from WES studies (25), it suggests that there is potential for a high number of false positives in this TruSeq data set, particularly at low VAFs. This observation highlights the importance of molecular validation and orthogonal sequencing of low VAF and clinically-relevant mutations.

### 3.4.2 Removing recurrent false positives from the mutation data

To filter the variant list to remove recurrent false positives, a training strategy using IGV was applied. Each variant was assessed for the number of times its genomic position occurred within a run, followed by the selection of the smallest number of recurrent variants, whilst still capturing either a) the IGV validated variants in cases presented in an interim analysis (n=366), or b) the IGV validated variants without additional filtering of *TP53, ATM, NOTCH1* and *SF3B1* in additional cases (n=144). This strategy reduces the number of candidate variants to 3,744 **(Figure 25B)**. This strategy succeeded in the removal of 70% of the variants passed at the first filtering step. Although it is not possible to confirm that all recurrent false positives have been removed, a two-way binomial distribution test found a significant difference between the number of variants found after this filtering process in comparison to the 12,429 variants found after the first set of filtering ($P$<0.0001), assuming that the distribution to support the null hypothesis was 0.5 (whether the variant was a recurrent false positive or not being random).

### 3.4.3 Filtering strategy to retain high quality, clinically and biologically relevant mutations

Using public databases of germline mutation variation (dbSNP, 1000g, ESP), somatic cancer mutation variation (COSMIC), and mutation effect prediction algorithms (SIFT, PolyPhen), as well as knowledge of previously observed somatically acquired mutations in CLL using the most recently described patient series (57,96), a filtering strategy to select for the removal of non-pathogenic mutations was conducted. Mutations were included if they were previously observed in COSMIC or CLL (regardless of germline mutation annotation or benign mutation effect prediction scores), with non-cancer associated variants only included if they were deleterious/damaging for both the SIFT and PolyPhen predication algorithms. This approach provides the best possible method between removing germline variation (which cannot be

achieved experimentally due to lack of available germline DNA), and keeping potentially pathogenic variants. After accounting for quality metrics (such as Phred score) and confidence in calling the mutation at its reported VAF, this produced a final list of 1,868 variants to take forward for IGV validation **(Figure 25C)**.

### 3.4.4 *In silico* mutation validation using Integrated Genomics Viewer

All 1868 variants were visualised in IGV, following a strict set of criteria to define a variant as a true positive **(2.2.5)**. Additionally, variants were assessed per gene to account for recurrent batch false positives which pass the initial criteria. From this analysis, 1,408 variants were identified to be true positives, with an *in silico* true positive percentage of 75.38%, representing the number of manually curated variants vs. the 1,868 variants described in section **3.4.3 (Figure 25D)**.

### 3.4.5 Assessment of minor subclones using orthogonal sequencing and method agreement analysis

Forty seven percent (659/1408) of IGV validated variants had a VAF <12%. To verify whether potential *in silico* true positives were sequencing artefacts, 60 variants from 4 genes (*TP53, ATM, NOTCH1* and *SF3B1*) with a VAF below 12% were taken forward for orthogonal sequencing validation **(2.1.10, 2.1.13, 2.1.14)**. The concordance between the TruSeq data and the orthogonal sequencing was low, with only 58.33% of variants being identified. Furthermore, the average TruSeq VAF of validated variants was 6.1%, with the average TruSeq VAF of the failed variants being 3.7%, suggesting that false positives are more likely to have a lower VAF. However, of these validated variants, 100% have been annotated in COSMIC, as well as 94.29% being previously observed to be somatically acquired in CLL. Meanwhile, only 24% of non-validated variants were annotated in COSMIC, with only 40% of variants having either COSMIC or previous CLL mutation annotation.

To assess the concordance between TruSeq and orthogonal sequencing validation, Bland-Altman plots were produced. Bland-Altman plots are used to assess the agreement between two techniques, with the bias output value representing the mean difference between the two techniques. The agreement between the TruSeq VAF and the validation VAF was assessed using a Bland-Altman plot, showing a bias of 0.0198 (mean difference in VAF across between the TruSeq VAF and the orthogonal sequencing VAF) across all 60 variants **(Figure 26A)**. However, when the variants which had 0 mutant reads (in the orthogonal sequencing validation experiment) were removed, the agreement between technologies was less than 1% VAF (bias = 0.0078) **(Figure 26B)**. This shows that using orthogonal sequencing to validate subclonal mutations found in the

TruSeq is required to validate high confidence variants. When considering only variants with a VAF of <5%, it has been previously observed that TruSeq does not perform well (170). Therefore, to improve confidence in variants <5% VAF being true positives, all variants below this threshold without either COSMIC or previous CLL annotation were filtered out (n = 271). This improves the final list of high confidence variants since it removes the variants most likely to not have a clinical impact in CLL.

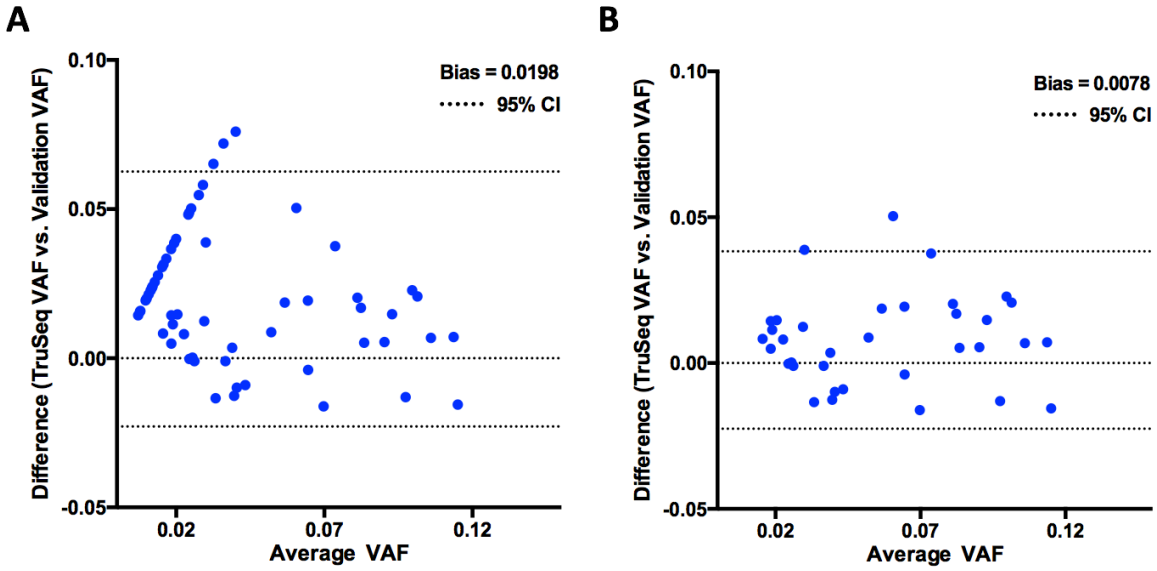



**Figure 26 Bland-Altman plots of minor subclone validation**

Bland-Altman plots of the agreement between the reported TruSeq VAFs and orthogonal sequencing validation VAFs of all 60 <12% VAF variants taken for validation **(A)**, and only variants which had at least 10 mutant reads in the orthogonal sequencing experiment **(B)**. Blue dots represent the XY scatter of the average VAF between TruSeq and orthogonal sequencing (*x* axis) and the raw VAF difference between TruSeq and the orthogonal sequencing (*y* axis).

### 3.4.6 Dealing with regions of high percentage GC content

Considering the impact of GC content on sequencing coverage **(Figure 21B)**, it is possible that true positives are being missed from the bioinformatics filtering steps due to not meeting the strict quality requirements of the filtering process **(Figure 25)**. To assess this, amplicons with more than 60% GC content were identified, and cross-referenced with whether hotspot mutations occurred within their target regions. Of the genes identified as having more than 60% GC content within their amplicon targets, only *NOTCH1* and *NFKBIE* also carried hotspot mutations: *NOTCH1* P2514fs and *NFKBIE* T253fs. Therefore, these two indels were therefore taken forward for manual screening across all 500 cases to identify potential low coverage mutations that could be re-introduced into the variant list. Indels were re-introduced if they had >20 reads and a VAF >5%, or if they had <20 reads, the variant was required to have a VAF of >12%. This led to the re-introduction of 15 *NOTCH1* variants and 1 *NFKBIE* variant. This increased the total number of *NOTCH1* P2514fs variants to 50 and *NFKBIE* T253fs variant to 40.

### 3.4.7 Additional annotation and filtering to ensure high quality variants

Additional annotation was conducted to take forward a list of high quality and clinically relevant variants for association and survival analysis **(2.2.4)**. In brief, the following annotations were created: exon inclusion criteria from a recurrent mutated gene study by the European Research Initiative in CLL (ERIC), removal of *ATM* variants shown not to be pathogenic based on the Leiden Open Variation Database (LOVD), removal of variants in *CTBP2* that may be affected by pseudogene usage **(A.6)**, exclusion of variants without COSMIC or CLL annotation that fall outside of previously described CLL mutated exons, and inclusion of mutations validated using an orthogonal sequencing technique. This additional filtering led to a final total of 903 high confidence variants **(A.7)**. A full description of the mutations found across the TruSeq panel is described below **(3.6)**.

## 3.5 Concordance between TruSeq and previous CLL4 studies

Before outlining the mutation landscape of the CLL4 TruSeq study, it is important to describe how TruSeq compares to previous studies to capture mutations in key genes. This is important to assess the sensitivity of TruSeq, not only to capture mutations below the detection threshold of traditional techniques such as Sanger sequencing, but to capture the mutations found in previous studies in an unbiased manner.

There have been three publications using the CLL4 clinical trial cohort to identify mutations in genes included on the TruSeq panel: Gonzalez *et al*. (2011) for *TP53* (167)*,* Skowronska *et al.* (2012) for *ATM* (9)*,* and Oscier *et al.* (2012) for *SF3B1* and *NOTCH1* (7). Unfortunately, at the time of writing the full list of CLL4 cases used in the Skowronska paper were not available to the PhD candidate, therefore an assessment of the concordance between the two studies cannot be conducted for *ATM*.

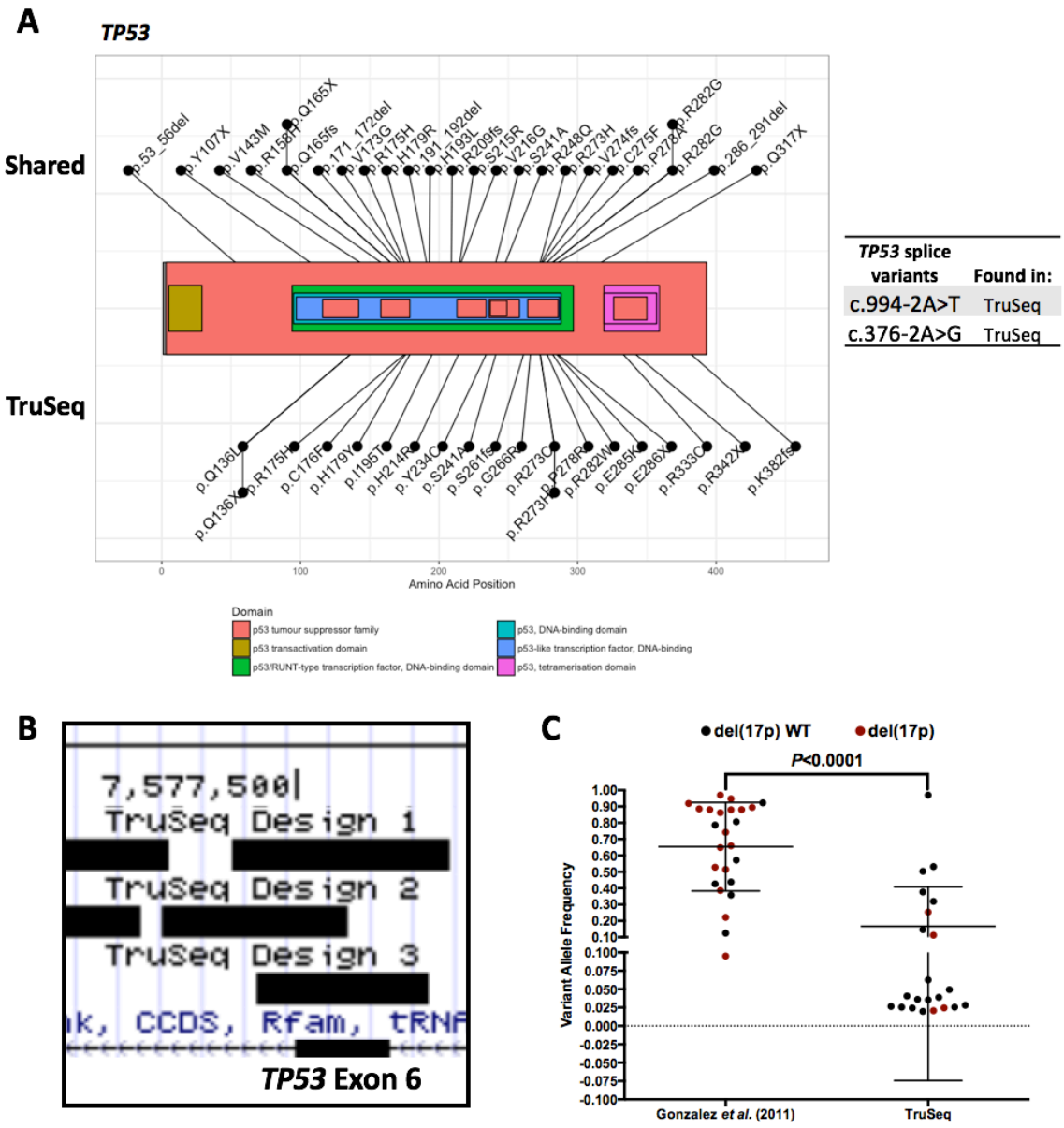
The approach undertaken for each gene was first to identify the cases shared between the CLL4 TruSeq study and the previous publication. Following which, the mutations that were shared and novel to each study were identified, and plotted using a Lolliplot. After which, the concordance between TruSeq and the previous publication technique was tested using Bland-Altman plots (described above **3.4.5**), with the detail of the variants discordant between studies explored.

### 3.5.1 CLL4 TruSeq study identified more *TP53* variants than Gonzalez *et al.* (2011), with increased sensitivity

Gonzalez *et al.* (2011) found *TP53* mutations in in 40/529 of the CLL4 cohort, a mutation frequency of 7.6%(167). From these 529 cases, 424 cases were shared between the former study and the TruSeq study. In these 424 cases, TruSeq identified 47 variants, 25 of which were previously identified in the Gonzalez study, as well as 22 variants novel to TruSeq **(Figure 27A)**. However, 6 variants in these 424 cases found by the Gonzalez study were not identified by TruSeq. Analysis of the coverage of the 3 designs in UCSC genome browser **(2.2.1),** identified a region of Exon 6 which was not covered in Design 2 **(Figure 27B)**, of which all 6 of these unidentified variants resided. Next, the sensitivity of each approach was assessed by comparing the TruSeq VAF distribution in the shared (previously identified by Gonzalez) and novel (TruSeq only) variants, stratified by del(17p) status **(Figure 27C).** This was conducted for two reasons. Firstly, to assess the distribution of identified variants between TruSeq and Sanger sequencing for *TP53* mutations. Secondly, to assess whether variants with a VAF of over 50% are associated with del(17p) events.

TruSeq only variants were found at a significantly reduced VAF (Gonzalez mean VAF = 0.6541, SEM $\pm$ 0.0541, TruSeq mean VAF = 0.1669 SEM $\pm$ 0.0514, $P$<0.0001). Furthermore, 14/20 variants with a VAF greater than 0.5 had concomitant deletions in the short arm of chromosome 17. Additionally, there were 7 >12% VAF variants in the 424 cases not captured by the technique applied in the original study, but captured by TruSeq **(Figure 27D)**. Since the original study used a bespoke technique (Capillary Electrophoresis Single-Stranded Conformation Analysis [CE-SSCA]) to filter for variants to screen by Sanger sequencing, the variant's IARC annotation was investigated to understand whether there was any *in silico* evidence to not include these mutations. Five out of seven were shown on IARC to be non-functional, therefore there is not a clear rationale as to why these cases would not have been picked up for Sanger screening in the original analysis by Gonzalez *et al.* (2011).

In total, TruSeq identified 53 variants after removing orthogonal sequencing negative variants **(Figure 28A)**. All variants <12% VAF, as well as a selection of clonal variants were taken forward for orthogonal sequencing validation (n=29). Post-sequencing, the variants were assessed in IGV **(2.2.5)**. The agreement between TruSeq and the validation technology (Ion Torrent or SureSelect) was assessed using a Bland-Altman plot, identifying a bias across all variants of 0.0165 **(Figure 28B).** Furthermore, additional Sanger sequencing was conducted, and combined with the confirmation data from the Gonzalez study, identifying 8/54 variants to be false positives **(Figure 28C&D)**.

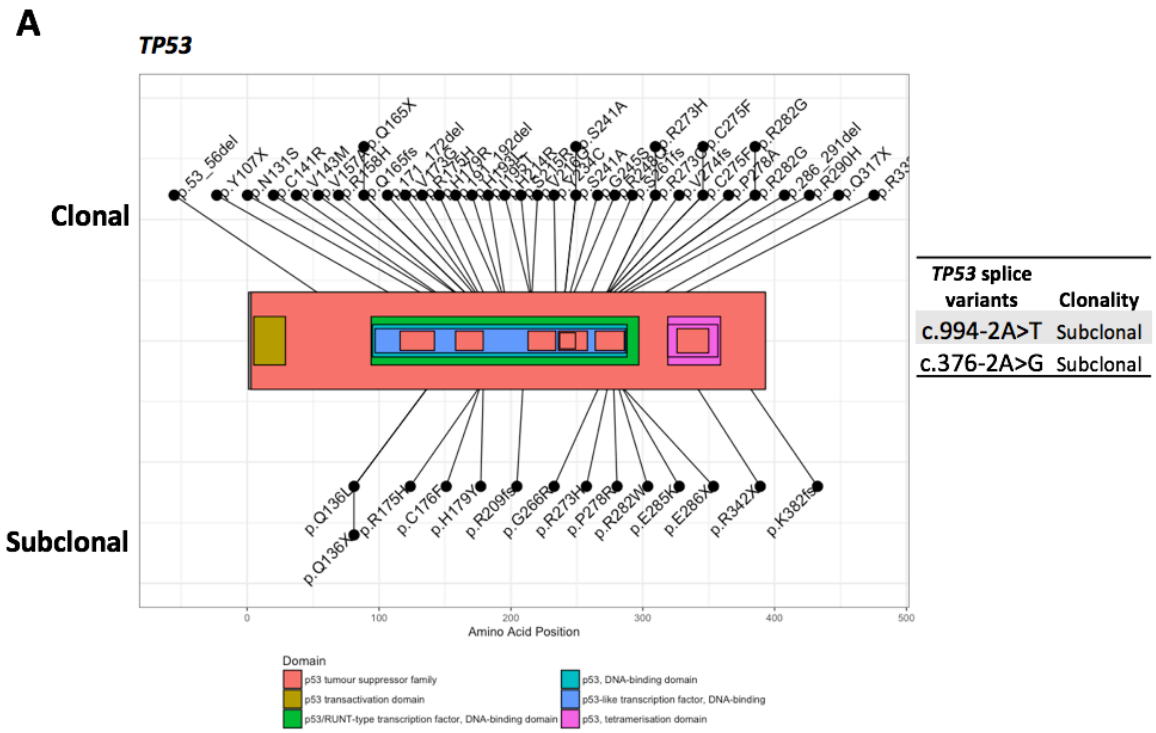


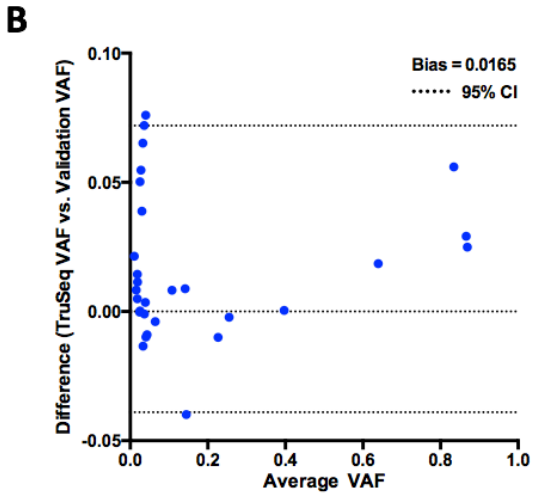

**D**

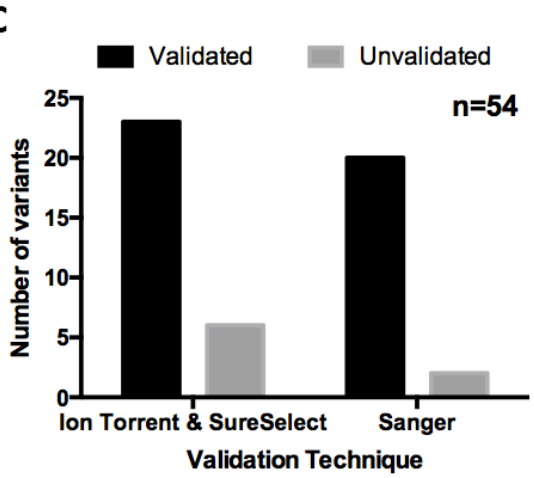
| CLL4 ID | Amino Acid Change | VAF | Exon | del(17p) | IARC Annotation | Screened in Gonzalez *et al.* (2011) |
|---|---|---|---|---|---|---|
| 4016 | p.R333C | 0.5319 | 10 | wt | Functional | No |
| 4070 | p.S241A | 0.9698 | 7 | wt | Non-Functional | Yes |
| 4105 | p.S261fs | 0.3193 | 8 | wt | Not Found | Yes |
| 4140 | p.R273C | 0.5036 | 8 | wt | Non-Functional | Yes |
| 4527 | p.I195T | 0.377 | 6 | wt | Non-Functional | Yes |
| 4662 | p.Y234C | 0.1459 | 7 | wt | Non-Functional | Yes |
| 4674 | p.H214R | 0.2541 | 6 | del | Non-Functional | Yes |

**Figure 27 Analysis of concordance between Gonzalez *et al.* (2011) and TruSeq for *TP53* mutations**

Concordance analysis between Gonzalez *et al.* (2011) and TruSeq, using only cases analysed in the former study. **A** Lolliplot showing the variants found by TruSeq and Gonzalez (top), and only found by TruSeq (bottom). Table to the right shows the spice variants only found by TruSeq. **B** Loss of capture of 50% of exon 6 in Design 2. **C** Distribution of VAF in shared and TruSeq only cases, stratified by del(17p) status. Bars represent mean and standard deviation. **D** Table showing the clonal (>12% VAF) variants found only by TruSeq, as well as their IARC annotation, and whether these cases were screened by Gonzalez.




D

| CLL4 ID | Amino Acid Change | TruSeq VAF | Exon | del(17p) | Clonal *TP53*mut | COSMIC/CLL/IARC Annotation | Orthogonal sequencing method |
|---|---|---|---|---|---|---|---|
| 4486 | p.S376T | 0.0502 | 11 | No | No | Yes/No/Functional | Ion Torrent & SureSelect |
| 4384 | p.S371fs | 0.055 | 11 | No | No | No/No/Not Found | Ion Torrent |
| 4486 | p.P142L | 0.0214 | 5 | No | No | Yes/Yes/Partially Functional | Ion Torrent & SureSelect |
| 4170 | p.C368A | 0.072 | 4 | Yes | Yes | No/Yes/Not Found | Ion Torrent |
| 4235 | p.G350A | 0.0781 | 4 | No | No | Yes/No/Not Found | Ion Torrent |
| 4486 | p.W91fs | 0.0652 | 4 | No | No | No/Yes/Not Found | Ion Torrent & SureSelect |
| 4412 | p.R114C | 0.9411 | 4 | No | NA | Yes/Yes/Not Found | Sanger |
| 4143 | p.Y236D | 0.3564 | 7 | No | NA | Yes/Yes/Non-Functional | Sanger |

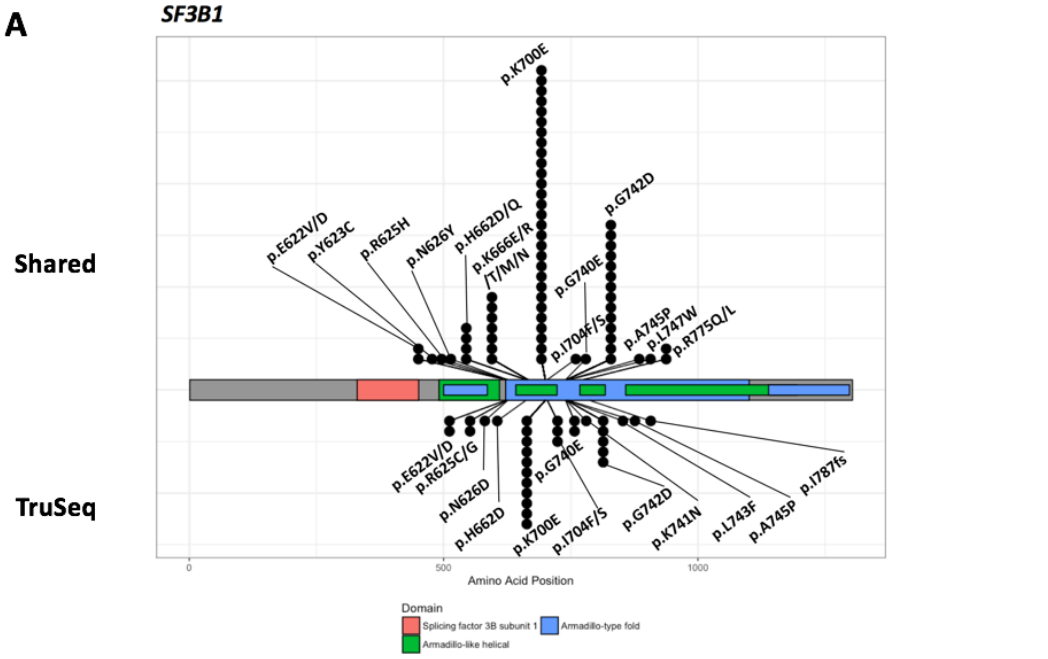
**Figure 28 Molecular validation of *TP53* mutations**

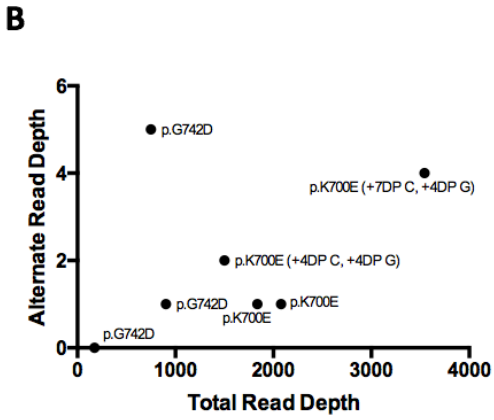
**A** Lolliplot of all TruSeq mutations included in final list of 903 variants. **B** Bland-Altman plot of the agreement between TruSeq variants and orthogonal sequencing. Blue dots represent the XY scatter of the average VAF between TruSeq and orthogonal sequencing (*x* axis) and the raw VAF difference between TruSeq and the orthogonal sequencing (*y* axis). **C** The number of variants taken for validation, stratified by success (Sanger includes Gonzalez *et al.* (2011) data). **D** Table of variants found to be false positives by orthogonal sequencing.

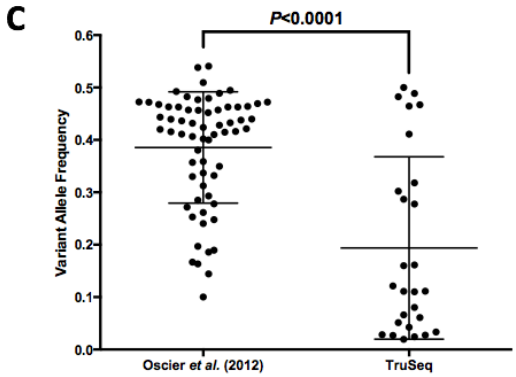
### 3.5.2 CLL4 TruSeq study identified more *SF3B1* variants than Oscier *et al.* (2012), with increased sensitivity

Oscier *et al.* (2012) observed *SF3B1* mutations in 73/437 cases, a mutation frequency of 16.7% using HRM followed by Sanger sequencing (7). Out of these 437 cases, 344 were shared between both studies, and taken forward for concordance analysis and molecular validation. In this sub-cohort, TruSeq identified 96 variants, of which 65 variants were previously observed in the Oscier study **(Figure 29A)**. However, there were 7 variants not identified by TruSeq, 4 were p.K700E and 3 were p.G742D. All 7 variants failed to be picked up by the bioinformatic pipeline, with all having >100 reads at the specific loci. However, upon IGV inspection these variants were not included because: they did not have >10 alternate reads, they contained multiple low level variants at the mutation locus, or they were found to be wild type with 0 alternate reads **(Figure 29B)**. There were 34 variants unique to TruSeq, 4 of which were additional *SF3B1* variants to those already identified as *SF3B1* mutants in the original Oscier study (3x p.K700E, 1x p.A745P) **(Figure 29A)**. Mutations found only by TruSeq also had a significantly reduced VAF in comparison to those found in the Oscier study (Oscier mean VAF = 0.3855, SEM $\pm$ 0.013, TruSeq mean VAF = 0.1939, SEM $\pm$ 0.033, $P$<0.0001) **(Figure 29C)**. Finally, there were 13 variants unique to TruSeq that were not observed in the Oscier study, but which all fell within the screening region (exons 14-16) and at a VAF expected to be captured by Sanger sequencing **(Figure 29D)**. However, in the previous study (Oscier *et al.*) cases were initially screened using WGA DNA and HRM analysis, with positive curves taken forward for Sanger validation. It could be that these cases failed the HRM process, and since only a selection of the HRM negative cases were selected for wild type Sanger sequencing, these variants could have been missed.

Overall, 134 variants were identified in *SF3B1* spanning 121 cases post molecular validation **(Figure 30A)**. Twenty-two variants were taken forward for orthogonal sequencing validation, spanning the range of observed VAFs (0.0333 – 0.5408), with a bias of 0.0166 by Bland-Altman agreement analysis **(Figure 30B)**. Additionally, 11 variants were screened and validated by Sanger sequencing, which led to 67 validated variants inclusive of the cases originally screened in the Oscier study. No variants taken for molecular validation were found to be false positives **(Figure 30C)**.




D

| CLL4 ID | Amino Acid Change | VAF | Exon | Exon screened in Oscier *et al.* (2012) |
|---|---|---|---|---|
| 4032 | p.K700E | 0.5 | 15 | Yes |
| 4037 | p.I704F | 0.4887 | 15 | Yes |
| 4141 | p.I704F | 0.467 | 15 | Yes |
| 4166 | p.K700E | 0.1599 | 15 | Yes |
| 4172 | p.I704S | 0.3022 | 15 | Yes |
| 4290 | p.K700E | 0.2868 | 15 | Yes |
| 4354 | p.G742D | 0.4646 | 16 | Yes |
| 4439 | p.K700E | 0.411 | 15 | Yes |
| 4485 | p.K700E | 0.4824 | 15 | Yes |
| 4578 | p.K700E | 0.1213 | 15 | Yes |
| 4670 | p.G740E | 0.1613 | 15 | Yes |
| 4704 | p.G742D | 0.2776 | 16 | Yes |
| 4748 | p.H662D | 0.3181 | 14 | Yes |

**Figure 29 Analysis of concordance between Oscier *et al.* (2012) and TruSeq *SF3B1* mutations**

Concordance analysis between Oscier *et al.* (2012) and TruSeq, only using cases analysed in the former study. **A** Lolliplot showing the variants found by TruSeq and Oscier (top), and only found by TruSeq (bottom). **B** XY scatter graph of total read depth and alternate read depth from variants observed in the Oscier study but not identified by TruSeq **C** Distribution of VAF in shared and TruSeq only cases. Bars represent mean and standard deviation. **D** Table showing the clonal (>12% VAF) variants found only by TruSeq, and whether these cases were screened by Oscier.

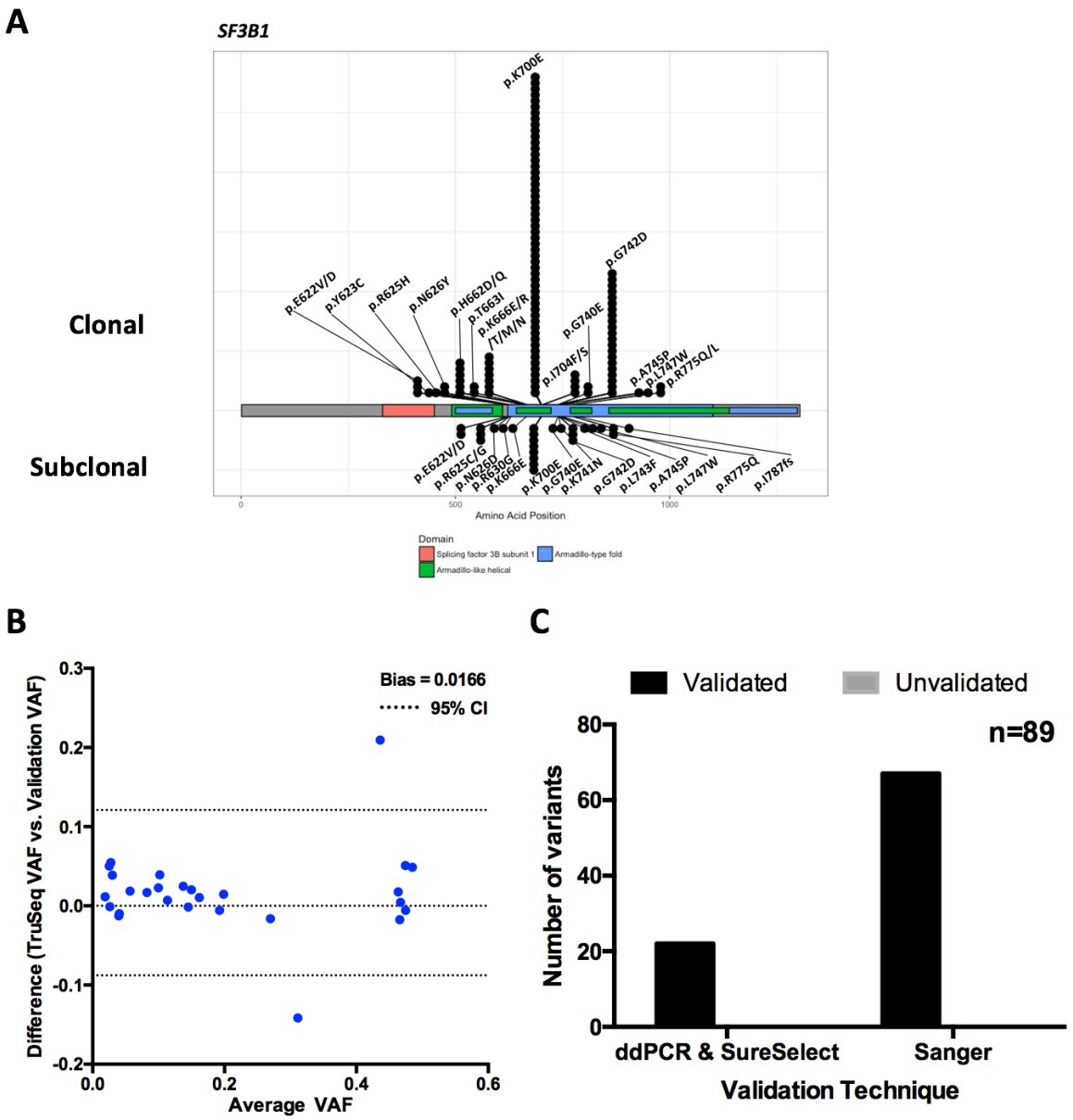


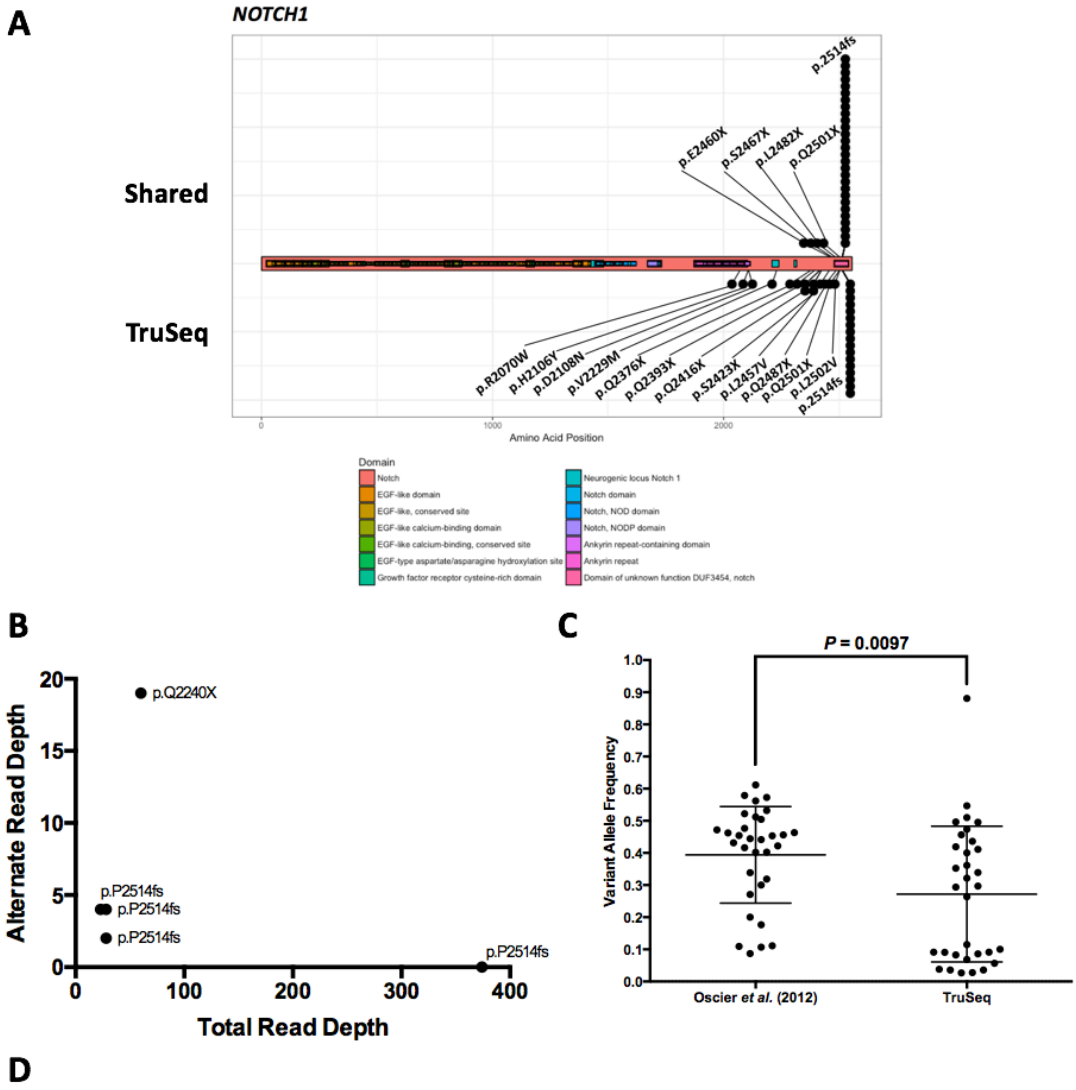

**Figure 30 Molecular validation of *SF3B1* mutations**

**A** Lolliplot of all TruSeq mutations included in final list of 903 variants. **B** Bland-Altman plot of the agreement between TruSeq variants and orthogonal sequencing. Blue dots represent the XY scatter of the average VAF between TruSeq and orthogonal sequencing (*x* axis) and the raw VAF difference between TruSeq and the orthogonal sequencing (*y* axis). **C** The number of variants taken for validation, stratified by success (Sanger includes Oscier *et al.* (2012) data).

### 3.5.3 CLL4 TruSeq study identified more *NOTCH1* variants than Oscier *et al.* (2012), with increased sensitivity

In addition to mutations in *SF3B1*, Oscier *et al.* (2012) also screened 466 cases for mutations in *NOTCH1,* identifying 46 variants and reporting a mutation frequency of 9.9%. Out of the 365 shared cases, 32 previously identified variants and 31 TruSeq novel variants were found in this sub-cohort **(Figure 31A)**. However, five variants were not observed by TruSeq, of which 4/5 were the hotspot p.P2514fs variants **(Figure 31B)**. All 4 of these variants had less than 10 mutant reads, and failed to meet the additional requirements for variants in high GC content regions described in **3.4.6**. The remaining variant (p.Q2240X), had less than 100 reads and more than 10 mutant reads, but didn't pass the filtering criteria because it was not a hotspot variant. TruSeq identified more low-VAF mutations than the Oscier study, with a significantly lower mean VAF across variants (Oscier mean VAF = 0.3942, SEM $\pm$ 0.0265, TruSeq mean VAF = 0.2720, SEM $\pm$ 0.0373, $P$ = 0.0097) **(Figure 31C)**. Finally, there were several >12% VAF variants unique to TruSeq (n=16), that weren't observed in the original Oscier study **(Figure 31D)**. These could have been missed in a similar fashion as described for the *SF3B1* mutations **(3.5.2)**, however considering the high GC content in this region of *NOTCH1,* it could also be attributed to poor sequencing in these cases.

In total 72 *NOTCH1* variants were identified in this cohort **(Figure 32A)**. Twenty variants were taken forward for molecular validation using either: Sanger, ddPCR, or SureSelect, from the range of VAFs reported (0.027 – 0.8806), and their agreement tested using a Bland-Altman plot **(Figure 32B)**. Although 19/20 variants were validated **(Figure 32C)**, the agreement between sequencing technologies was larger than for mutations in *TP53* or *SF3B1* (Bias = 0.046). When two variants with more than 20% VAF difference between each sequencing technology were removed, this brought the bias value in line with the validation in the other genes (Bias = 0.0177). This suggest that in these two cases the library preparation in either the TruSeq or validation sequencing under or over represented the mutant allele. This is important, especially for low VAF variants, since misrepresentation of the true VAF could lead to the inclusion of a false positive, which could potentially weaken the statistical survival associations performed in **Chapter 4:.**




D

| CLL4 ID | Amino Acid Change | VAF | Exon | Exon screened in Oscier *et al.* (2012) |
|---|---|---|---|---|
| 4097 | p.P2514fs | 0.4363 | 34 | Yes |
| 4221 | p.Q2487X | 0.4003 | 34 | Yes |
| 4366 | p.P2514fs | 0.8806 | 34 | Yes |
| 4436 | p.P2514fs | 0.2637 | 34 | Yes |
| 4436 | p.Q2376X | 0.3524 | 34 | Yes |
| 4437 | p.H2106Y | 0.297 | 34 | Yes |
| 4486 | p.P2514fs | 0.5465 | 34 | Yes |
| 4544 | p.S2423X | 0.4113 | 34 | Yes |
| 4544 | p.V2229M | 0.5104 | 34 | Yes |
| 4565 | p.D2108N | 0.2938 | 34 | Yes |
| 4595 | p.Q2416X | 0.4568 | 34 | Yes |
| 4602 | p.P2514fs | 0.4964 | 34 | Yes |
| 4658 | p.Q2416X | 0.4953 | 34 | Yes |
| 4672 | p.P2514fs | 0.3214 | 34 | Yes |
| 4784 | p.Q2501X | 0.3392 | 34 | Yes |
| 4785 | p.Q2393X | 0.3616 | 34 | Yes |

**Figure 31 Analysis of concordance between Oscier et al. (2012) and TruSeq *NOTCH1* mutations**

Concordance analysis between Oscier *et al.* (2012) and TruSeq, only using cases analysed in the former study. **A** Lolliplot showing the variants found by TruSeq and Oscier (top), and only found by TruSeq (bottom). **B** XY scatter graph of total read depth and alternate read depth from variants observed in the Oscier study but not identified by TruSeq **C** Distribution of VAF in shared and TruSeq only cases. Bars represent mean and standard deviation. **D** Table showing the clonal (>12% VAF) variants found only by TruSeq, and whether these cases were screened by Oscier.

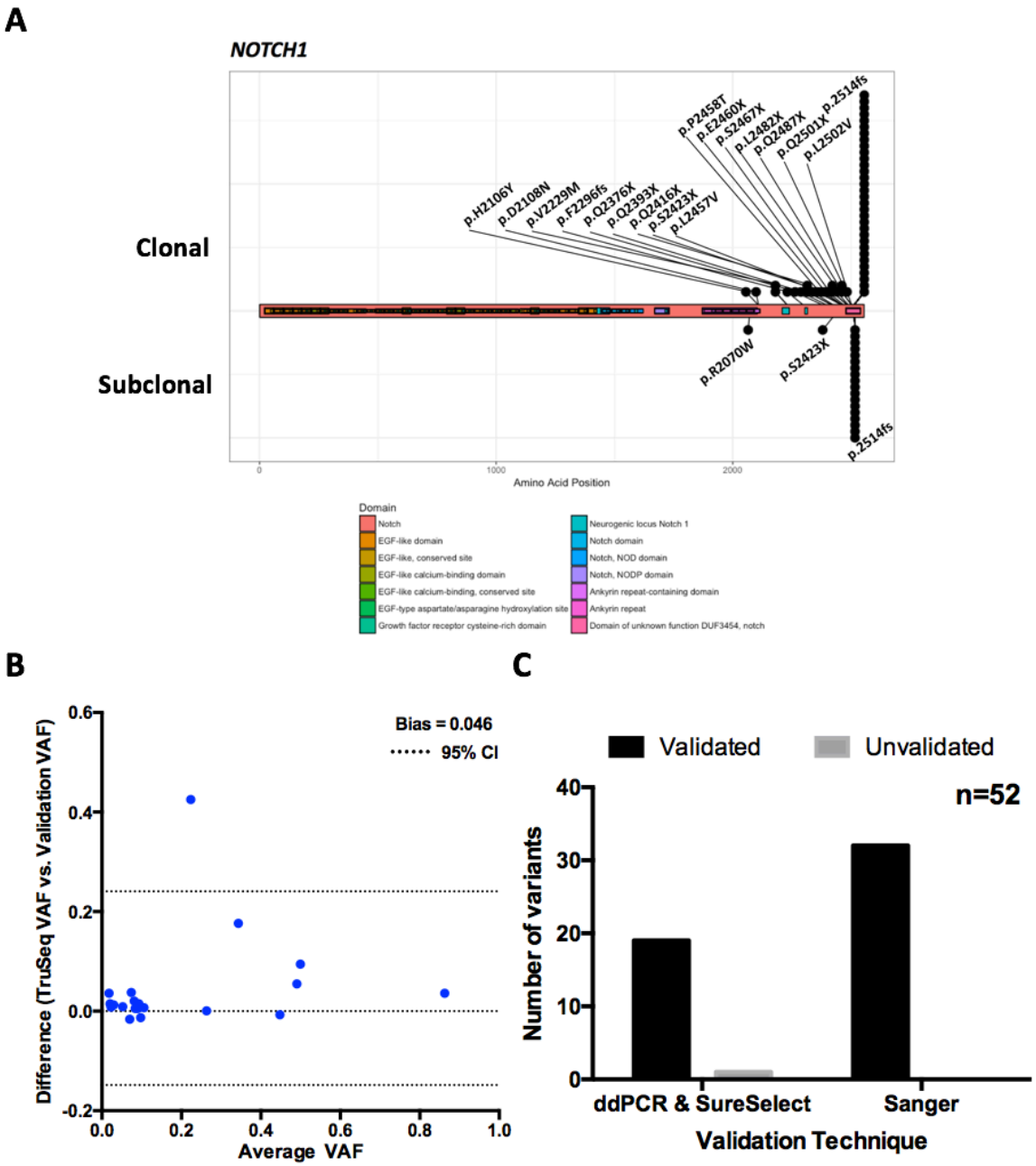



**Figure 32 Molecular validation of *NOTCH1* mutations**

**A** Lolliplot of all TruSeq mutations included in final list of 903 variants. **B** Bland-Altman plot of the agreement between TruSeq variants and orthogonal sequencing. Blue dots represent the XY scatter of the average VAF between TruSeq and orthogonal sequencing (*x* axis) and the raw VAF difference between TruSeq and the orthogonal sequencing (*y* axis). **C** The number of variants taken for validation, stratified by success (Sanger includes Oscier *et al.* (2012) data).

## 3.6 The mutation landscape in CLL4

After mutation data processing, a list of high quality variants was generated. To understand the impact that these genes are having on CLL pathogenesis, analysis of the frequency that each gene is mutated, the distribution of the mutation types found within each gene, and the clonal architecture of the mutations within each gene, is important. Furthermore, stratifying this analysis by key features of disease biology (such as IGHV mutational status), may also highlight genes which associate with good (M-CLL) or poor (U-CLL) risk disease.

Due to this study using targeted re-sequencing, and therefore not having matched copy number status for all genomic locations covered in the panel design (data inappropriate for analysis in MutSig, a mutation analysis software which assesses the frequency and local copy number status of each mutation in a dataset (171)), a power calculation using the cohort size (n=500) and the background mutation frequency in CLL (0.5) was conducted on www.tumorportal.org/power to identify significantly mutated genes. The result of this calculation was that this cohort was 100% powered to identify genes mutated in 5% of cases, whilst being; 96%, 69% and 12% powered to identify genes mutated in 3%, 2% and 1% of cases, respectively. This calculation was also calculated for the U-CLL (n=263) and M-CLL (n=157) sub-cohorts, identifying that both are only 100% powered to identify genes mutated in 10% of cases, with the U-CLL cohort 96% powered to identify genes mutated in 5% of cases, and the M-CLL cohort underpowered (i.e. <95% power) to identify genes mutated in <10% of cases **(A.8)**.

### 3.6.1 Overview of the mutation landscape in CLL4: *SF3B1* found to be the most recurrently mutated gene in CLL4

This study identified 18 recurrently mutated genes in the full cohort **(Figure 33)**, with 12 genes found to be recurrently mutated in the U-CLL subset, and 2 found to be recurrently mutated in M-CLL subset **(A.9)**. The CNV landscape, as reported in previous studies, is shown in **Figure 34**. *SF3B1* was found to be the most recurrently mutated gene in the cohort, accounting for 24.2% of all cases (135/903 variants). As observed in previous studies (7,10,31,82,172–174), mutations in *SF3B1* were centralised around exons 14-16, with the p.K700E variant representing most mutations (46.27%) **(Figure 35A)**. Excluding *MYD88*, all priority 1 genes were found to be mutated in >10% of cases (*TP53* = 10.6%, *NOTCH1* = 14.4%, *ATM* = 26.8%, *BIRC3* = 14.4%, *MYD88* = 4%), with the proportion of mutation types observed per gene representing what has been previously reported **(Figure 33)**. Mutations in *BRAF* in CLL have previously been observed not to contain the pan-cancer hotspot p.V600E variant, with previous studies suggesting this has implications for

*BRAF* specific therapy in these cases. However, in this cohort 7/20 p.V600E variants were observed, representing less than 50% of the observed mutations in this gene **(Figure 35B)**.

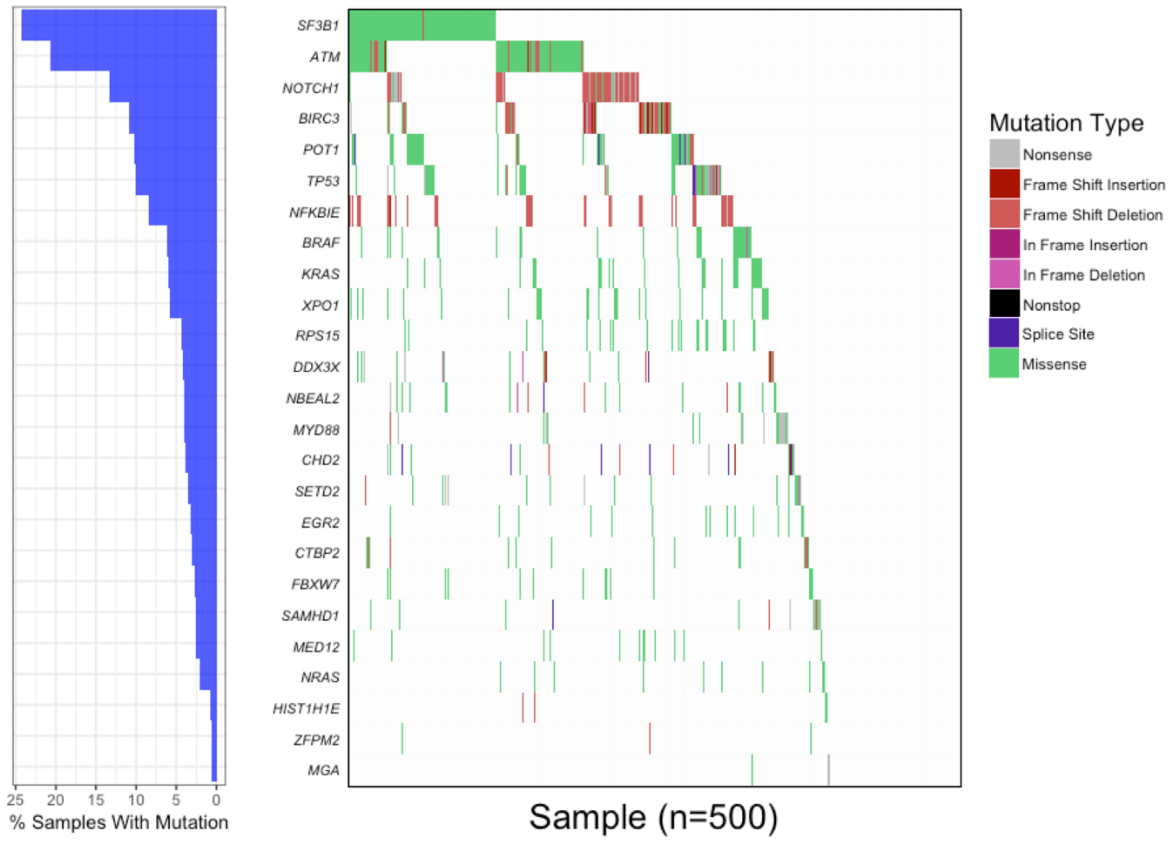



**Figure 33 Mutation Landscape of CLL4**

Waterfall plot showing the distribution of high confidence variants stratified by mutation type (key to the right) across the 25 genes for 500 cases, along with their mutation frequency (bar chart, left).

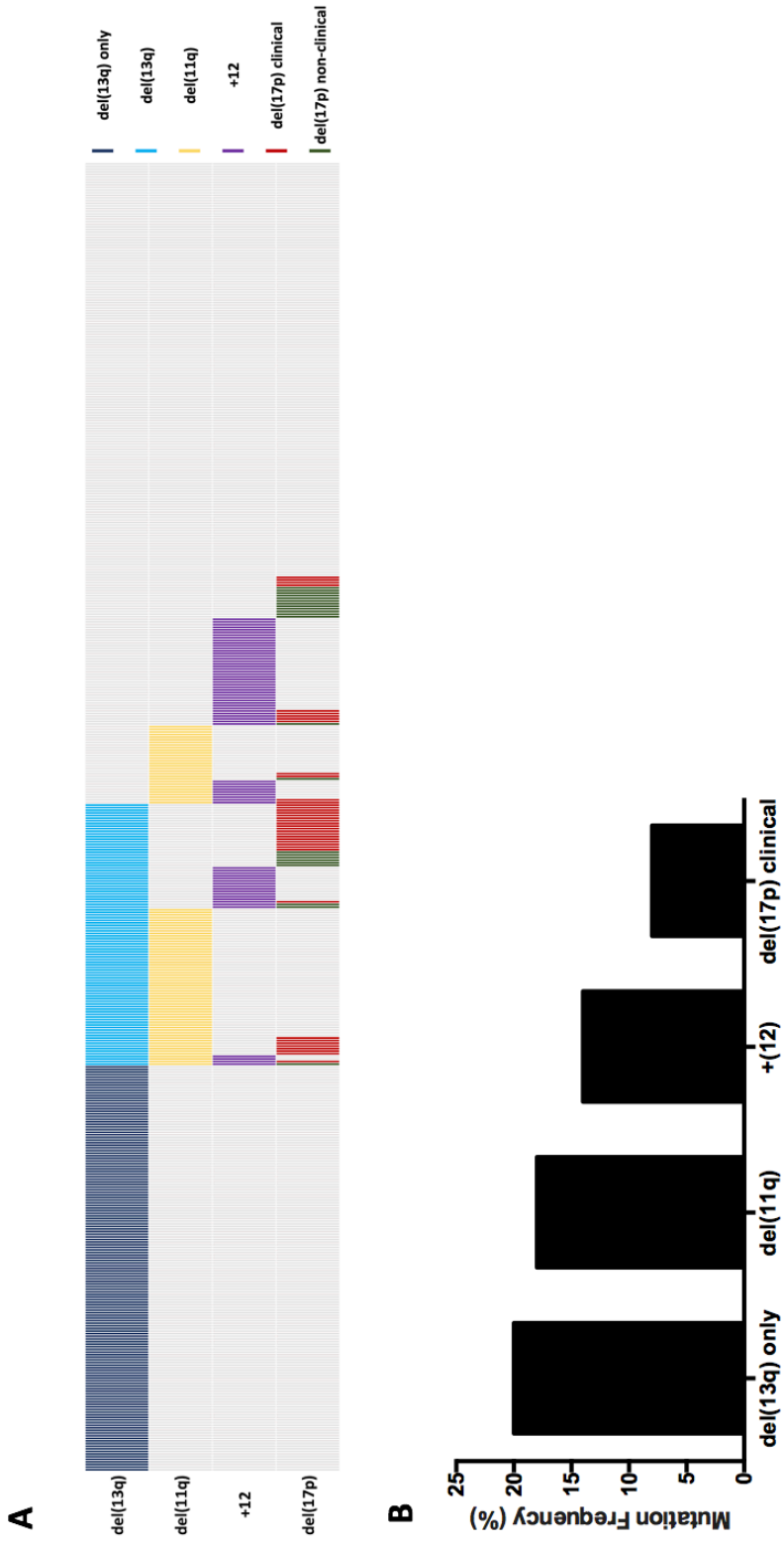



**Figure 34 Frequency of CLL4 clinical trial FISH CNVs**

**A** Heat map showing the distribution of CNVs in the CLL4 TruSeq cohort. **B** Frequency of CNVs in the cohort according to FISH analysis, with del(13q) cases with other CNVS, and specific del(17p) cases not meeting clinical criteria removed.

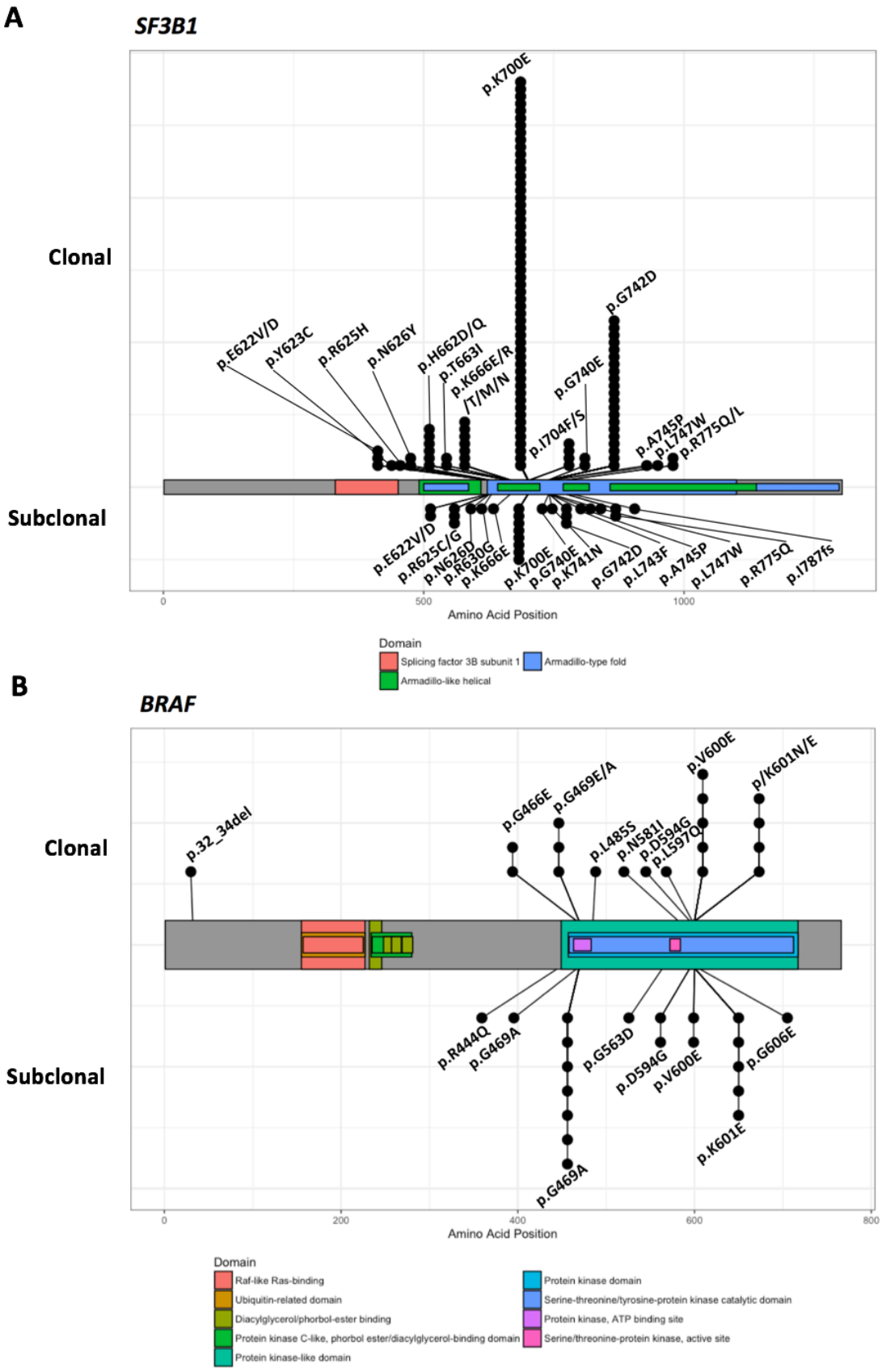



**Figure 35 Mutation Lolliplots of *SF3B1* and *BRAF* variants**

Lolliplots showing the location and amino acid change of mutations in *SF3B1* **(A)** and *BRAF* **(B)**. Variants plotted on the top of the protein map represent variants with a VAF >12% (clonal), whilst those plotted below the protein map represent variants with a VAF <12% (subclonal).

### 3.6.2 Distribution of mutation type in the full cohort and stratified by IGHV mutational status: Transitions events predominate in CLL4

Missense mutations dominate the variation seen in this CLL4 cohort, accounting for 71.65% of all variants **(Figure 36A; left)**. 22.59% of variants were indels, predominated by frameshift deletions (n = 172), with nonsense and splicing variants making up the residual 5.76% of variants **(Figure 36A; left)**. The distribution of VAF between mutation types was assessed by ANOVA, identifying indels as occurring at a significantly lower VAF than missense mutations (*P* = 0.0024) **(Figure 36A; right)**. Since targeted sequencing without copy number data is not suitable for ABSOLUTE analysis, assessment of the mutation signature of these CLL cases was not possible. This process identifies the sequence context of each mutation (i.e. the 3' and 5' sequence context either side of the mutation site), allowing the calculation of the mutation frequency of all 96 conceivable missense mutation combinations (25). In this rudimentary analysis, the 16 possible combinations for each transversion and transition event was calculated in combination for all CLL, as well as the U-CLL and M-CLL subsets, showing that transition events predominate in CLL4 **(Figure 36B)**. This "signature" is comparable to the 1B and 2 mutation signatures described by Alexandrov *et al.* (2013), where C>T transversions had a profile that was associated with ageing, as well as over-activity of APOBEC genes which were linked with viral infection (25).

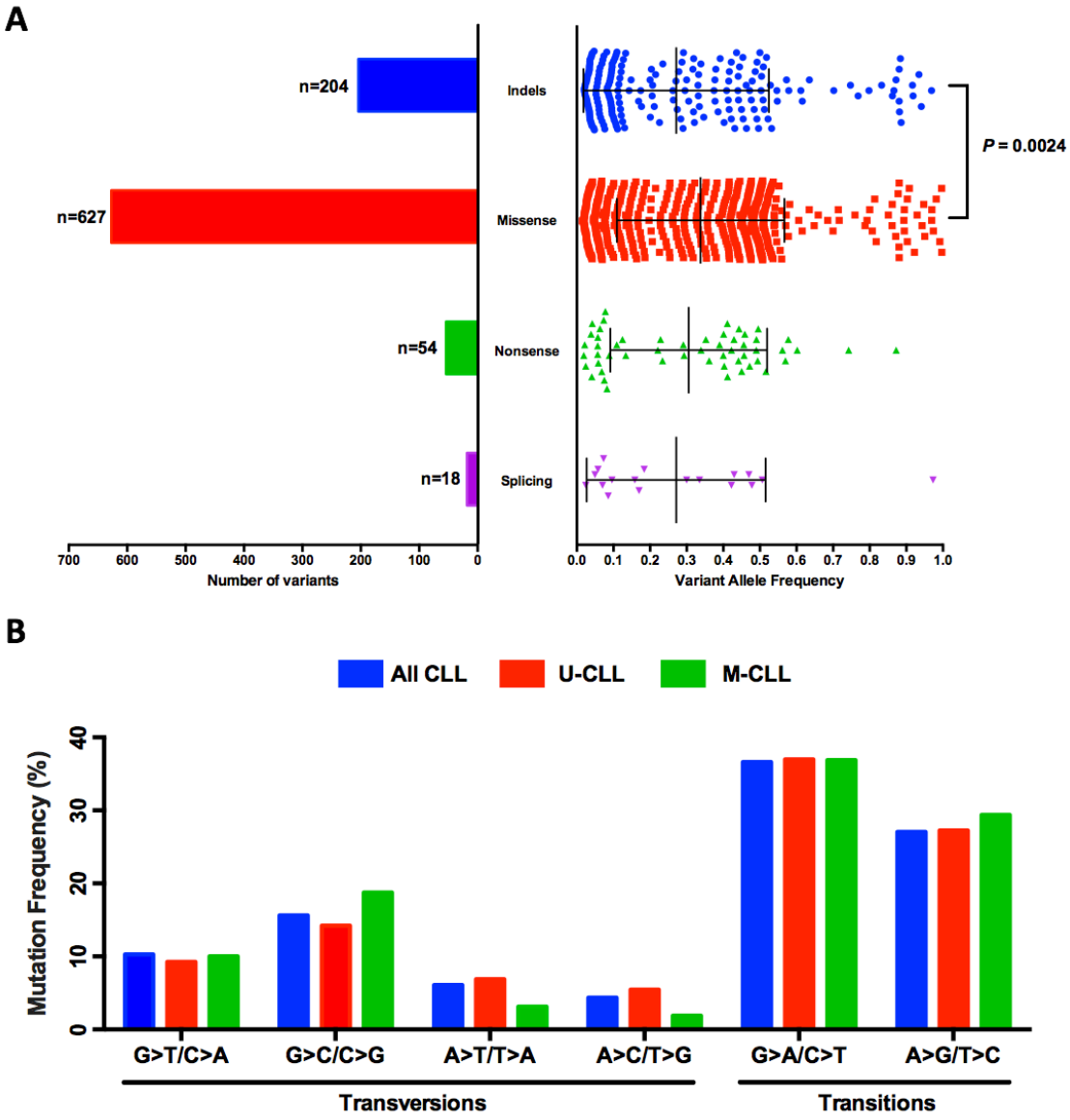



**Figure 36 Distribution of mutation type**

**A** Distribution of VAF of all mutations between mutation types. Bars represent mean and standard deviation. One-way ANOVA used to calculate differences between groups (right). Total number of variants per mutation type (left). **B** Distribution of missense mutations by transversion and transition mutation class.

### 3.6.3 Mutation landscape in U-CLL and M-CLL subsets of CLL4

*SF3B1* retained the highest mutation frequency in both the U-CLL and M-CLL subsets (U-CLL = 28.52% of cases, M-CLL=17.83% of cases), of which the latter was shared with *ATM* (**Figure 37A**; right). To identify genes which were significantly mutated in U-CLL vs. M-CLL or *vice versa*, a Fisher's exact test between each gene and the IGHV mutational status was conducted on genes which had >10 mutated cases, identifying *KRAS, RPS15* and *XPO1* to all be significantly associated with the U-CLL subset ($P$<0.01) (**Figure 37A**; left; **3.6.3.1**). Additionally, *SF3B1, NFKBIE* and *DDX3X* were shown to associate with U-CLL ($P$<0.05), with *MYD88* shown to associate with M-CLL ($P$<0.05) (**Figure 37A**; left). Since the general trend per gene was that the mutation frequency was higher in U-CLL than M-CLL (18/22 genes), the average number of mutations per case in each sub-group was calculated, showing no statistical difference between U-CLL and M-CLL (mean number of variants per subgroup; U-CLL = 2.379, M-CLL = 2.061, $P$=0.067). Furthermore, an analysis of the distribution of mutation VAFs between U-CLL and M-CLL showed no statistical difference (mean VAF per subgroup; U-CLL = 0.311, M-CLL = 0.335, $P$=0.204). This suggests that when comparing the mutation burden between U-CLL and M-CLL using a targeted panel pre-selected with known genes of pathogenicity in CLL, that the previously observed increased mutation burden in M-CLL from WGS data does not hold (57). However, this was an unbiased analysis which did not control for copy number or high risk variants in genes such as *ATM* or *TP53.* Furthermore, the CLL4 TruSeq panel is enriched for genes mutated in U-CLL, therefore the frequency of mutations in U-CLL was to some extent expected to be higher.

#### 3.6.3.1 Recurrently mutated *KRAS, RPS15,* and *XPO1* are significantly associated with U-CLL in CLL4

*KRAS, RPS15* and *XPO1* have all been previously reported to be associated with U-CLL(96). The mutation spectrum observed in this study for these genes is comparable to the previously reported studies, with the predominance of p.G13D in *KRAS,* C-terminal mutations in *RPS15* and the hotspot p.E571K variant in *XPO1* **(Figure 38)**. To establish whether there were clonal or subclonal mutation differences between the U-CLL and M-CLL subgroups within each gene, Pearson's Chi-squared 2x2 contingency calculations were carried out. *KRAS* showed a strong trend towards most clonal mutations being associated with U-CLL, while conversely subclonal mutations skewed towards being largely from the M-CLL subset ($P$=0.012), whilst *RPS15* and *XPO1* showed no differences (*RPS15* $P$=0.159, *XPO1* $P$=0.781). The subclonal architecture of the entire CLL4 cohort is described in more detail below **(3.9)**.

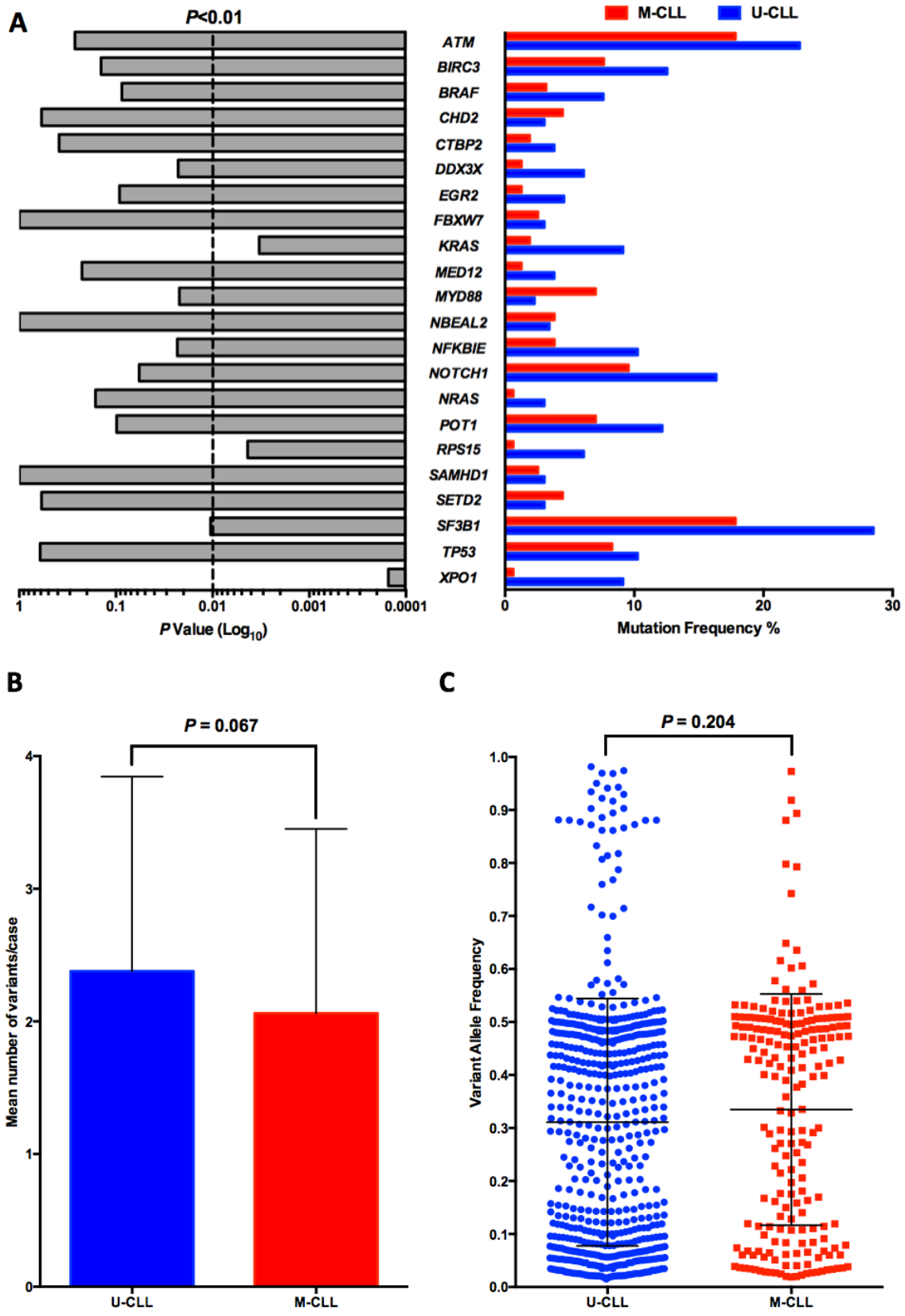
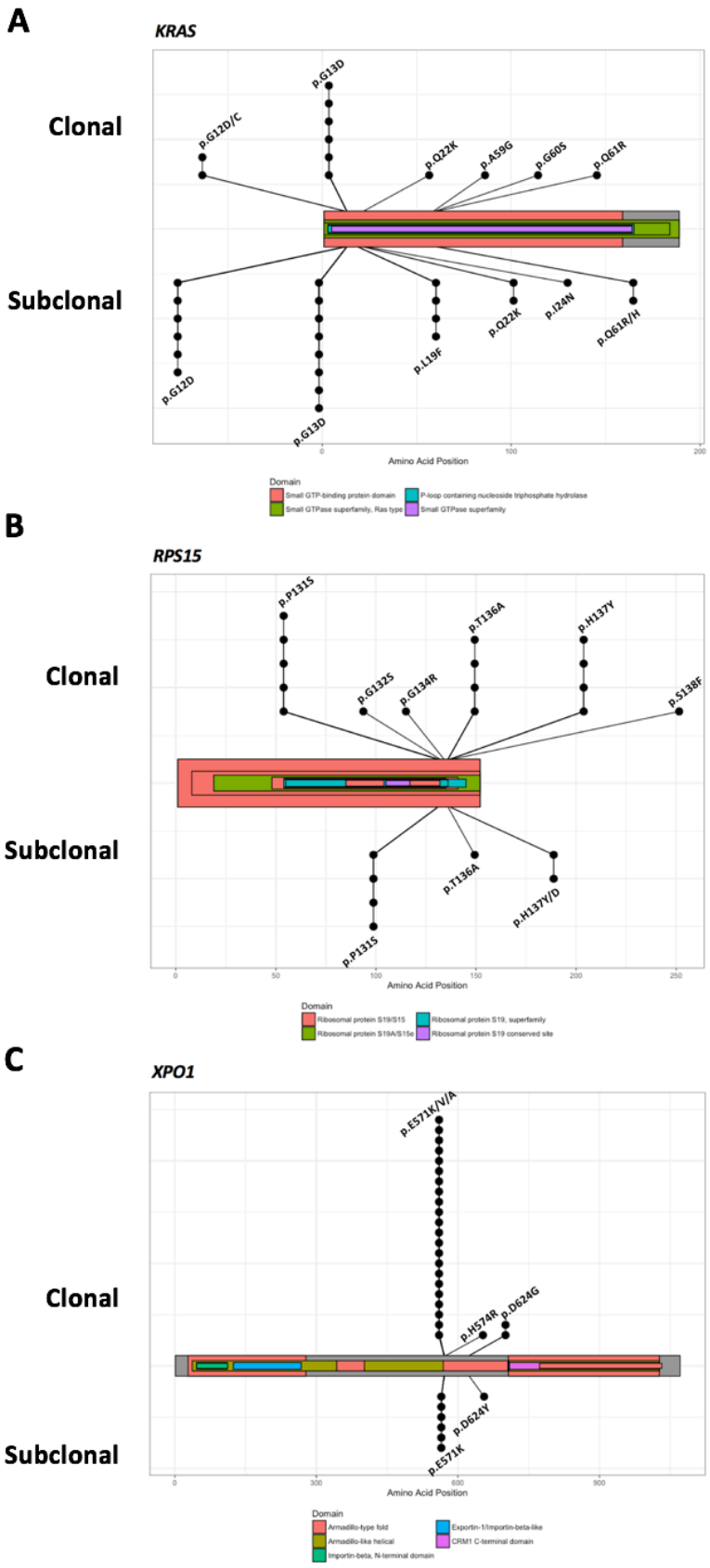



**Figure 37 *IGHV* mutational status co-occurrence analysis**

**A** Mutation frequency of all genes with >10 mutated cases, stratified by *IGHV* mutational status (right, U-CLL blue bars, M-CLL red bars). Association testing conducted for each gene vs. *IGHV* mutational status using Fisher's exact test (left, significance cut off *P*<0.01). **B** Mean number of variants per case between U-CLL and M-CLL. Bars represent standard deviation. **C** Distribution of VAF of all mutations between U-CLL and M-CLL. Bars represent mean and standard deviation. Unpaired student's T-test used to calculate differences between groups in **B** and **C**.




| *IGHV* mutational status | Clonal | Subclonal |
|---|---|---|
| U-CLL | 11 | 16 |
| M-CLL | 1 | 4 |
| Missing | 0 | 3 |

| *IGHV* mutational status | Clonal | Subclonal |
|---|---|---|
| U-CLL | 11 | 7 |
| M-CLL | 4 | 0 |
| Missing | 1 | 1 |

| *IGHV* mutational status | Clonal | Subclonal |
|---|---|---|
| U-CLL | 22 | 5 |
| M-CLL | 1 | 0 |
| Missing | 2 | 2 |

**Figure 38 Mutation Lolliplots of *KRAS, RPS15* and *XPO1***

Lolliplots showing the location and amino acid change of mutations in *KRAS* **(A)**, *RPS15* **(B)** and *XPO1* **(C)**. Variants plotted on the top of the protein map represent variants with a VAF >12% (clonal), whilst those plotted below the protein map represent variants with a VAF <12% (subclonal). Tables to the right of each Lolliplot show the number of variants found at a clonal or subclonal level, stratified by *IGHV* mutational status.

### 3.6.4 Overview of mutation landscape in *SF3B1* mutated cases of CLL4

Due to mutations in *SF3B1* being the most recurrent CLL mut driver gene in this cohort, independent of IGHV mutational status, co-occurrence analysis of other mutated genes within *SF3B1* mutated CLL was undertaken. Mutations in *ATM* were the most recurrent in both the *SF3B1* wild type and mutant sub groups (*SF3B1* WT = 18.73%, *SF3B1* mutants = 26.45%) **(Figure 39A; right)**. Association analysis via Fisher's exact testing found only *POT1* to be significantly associated with *SF3B1* mutants (**Figure 39A**; left; **3.6.4.1**). The mutation burden between *SF3B1* wild type and mutant subsets was assessed, identifying *SF3B1* mutants as having significantly reduced number of mutations in comparison to *SF3B1* wild type cases (mean number of variants per subgroup; *SF3B1* WT = 1.17, *SF3B1* mutants = 0.64, $P<0.0001$) **(Figure 39B)**. Finally, the distribution of the VAF across all mutations between the *SF3B1* wild type and mutant subgroups was assessed, identifying no significant difference ($P = 0.7382$) **(Figure 39C)**.

#### 3.6.4.1 Recurrently mutated *POT1* significantly co-occurs with *SF3B1* mutated cases in CLL4

*POT1* mutations have been observed previously to be associated with U-CLL, with mutations localised to the two oligonucleotide/oligosaccharide-binding (OB) domain, which impact telomere function and associate with increased frequency of chromosome abnormalities(175). Here, *POT1* mutations were shown to be significantly associated with *SF3B1* mutated CLL **(**$P = 0.00529$, **Figure 39A; left)**, therefore additional analysis to better understand the association between these two genes was undertaken. Most *POT1* variants were located within the OB domain (46/57) **(Figure 40A;** Lolliplot**)**. The distribution of hotspot *SF3B1* mutations in *POT1* mutated CLL was identified, and compared with the distribution of *SF3B1* mutations in the rest of the CLL4 cases using a Pearson's chi-squared 3x2 contingency table **(2.2.8.1)**. This analysis identified no differences between the two groups, suggesting no individual *SF3B1* variant to be associated with *POT1* mutations ($P=0.21$) **(Figure 40B)**.

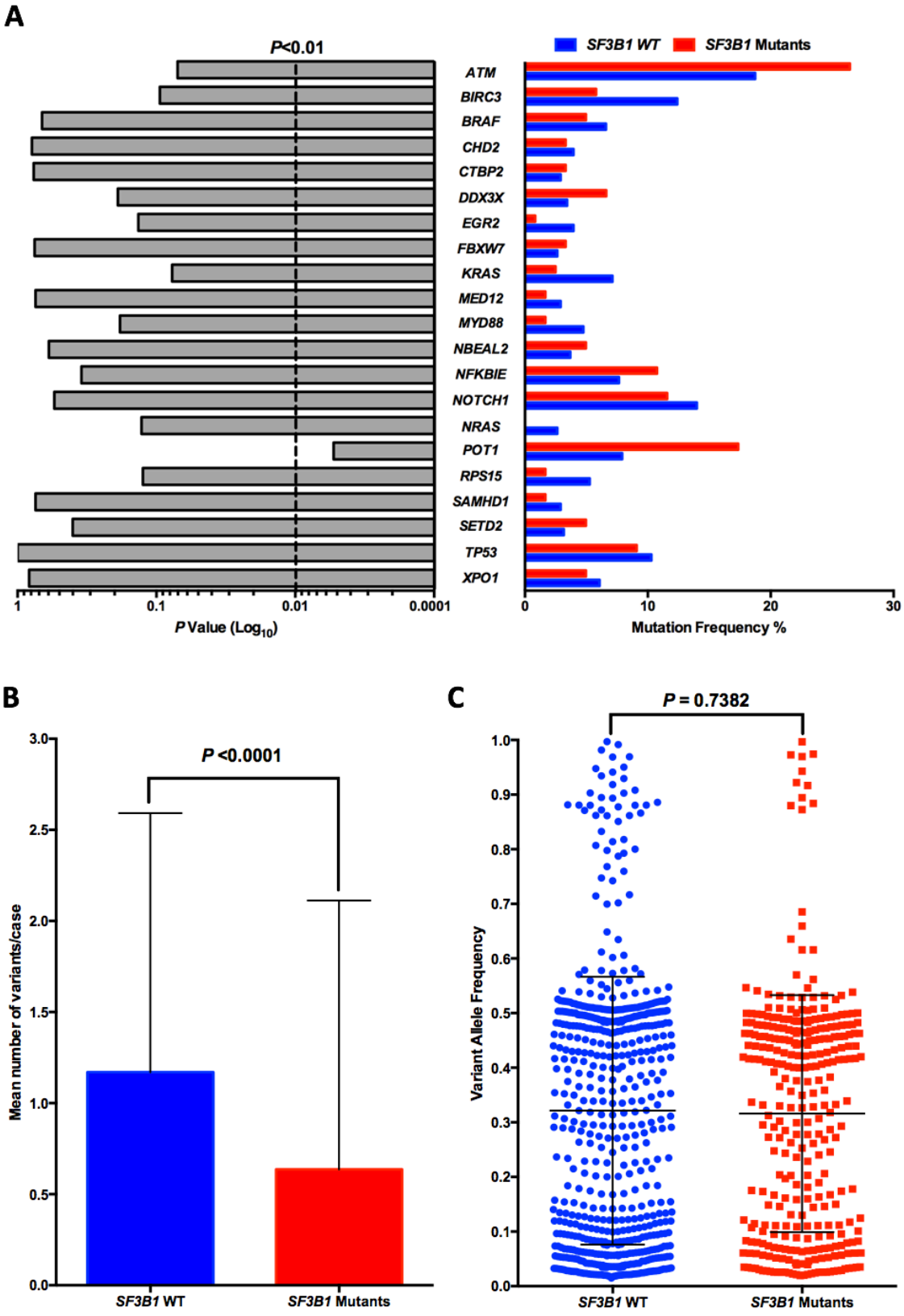
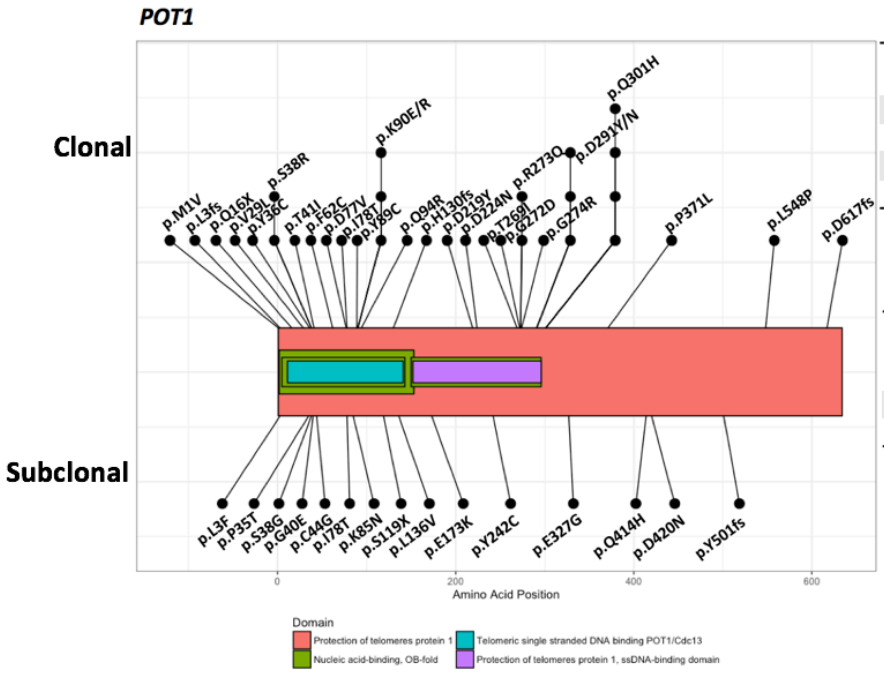



**Figure 39 Mutated *SF3B1* co-occurrence analysis**

**A** Mutation frequency of all genes with >10 mutated cases, stratified by *SF3B1* mutational status (right, *SF3B1* wild type blue bars, *SF3B1* mutants red bars). Association testing conducted for each gene vs. *SF3B1* mutational status using Fisher's exact test (left, significance cut off *P*<0.01). **B** Mean number of variants per case between *SF3B1* wild types and mutants. Bars represent standard deviation. **C** Distribution of VAF of all mutations between *SF3B1* wild types and mutants. Bars represent mean and standard deviation. Unpaired student's T-test used to calculate differences between groups in **B** and **C**.

**A**




| *POT1* splice variants | Clonality |
|---|---|
| c.125-2A>T | Clonal |
| c.124+2->AACTGG | Clonal |
| c.949+1G>A | Clonal |
| c.48+1G>T | Subclonal |

| *SF3B1* mutational status | Clonal | Subclonal |
|---|---|---|
| Wild type | 20 | 11 |
| Mutant | 17 | 5 |

**B**

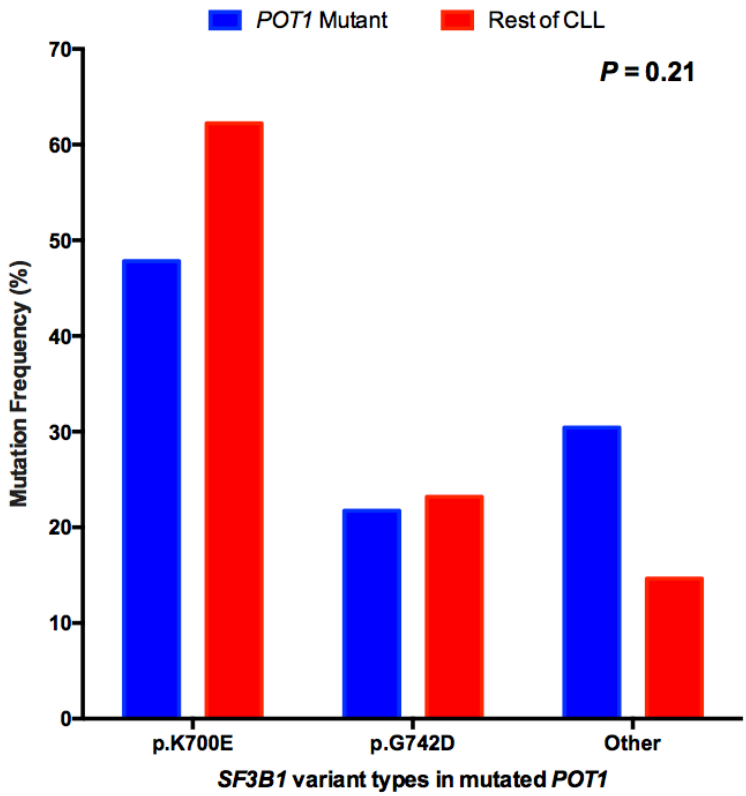



**Figure 40 Association analysis between *SF3B1* and *POT1* mutated CLL**

**A** Lolliplot showing the location and amino acid change of mutations in *POT1* (left), top table denoting splice variants which cannot be loaded in the GenVisR Lolliplot function, bottom table denoting the number of clonal and subclonal variants stratified by *SF3B1* mutational status. Variants plotted on the top of the protein map represent variants with a VAF >12% (clonal), whilst those plotted below the protein map represent variants with a VAF <12% (subclonal). **B** Mutation frequency of *SF3B1* mutations in *POT1* mutated and all CLL minus *POT1* mutated cases. Pearson's chi-squared 3x2 contingency tables used to calculate the difference between *SF3B1* variant types and *POT1* mutated vs. rest of CLL.

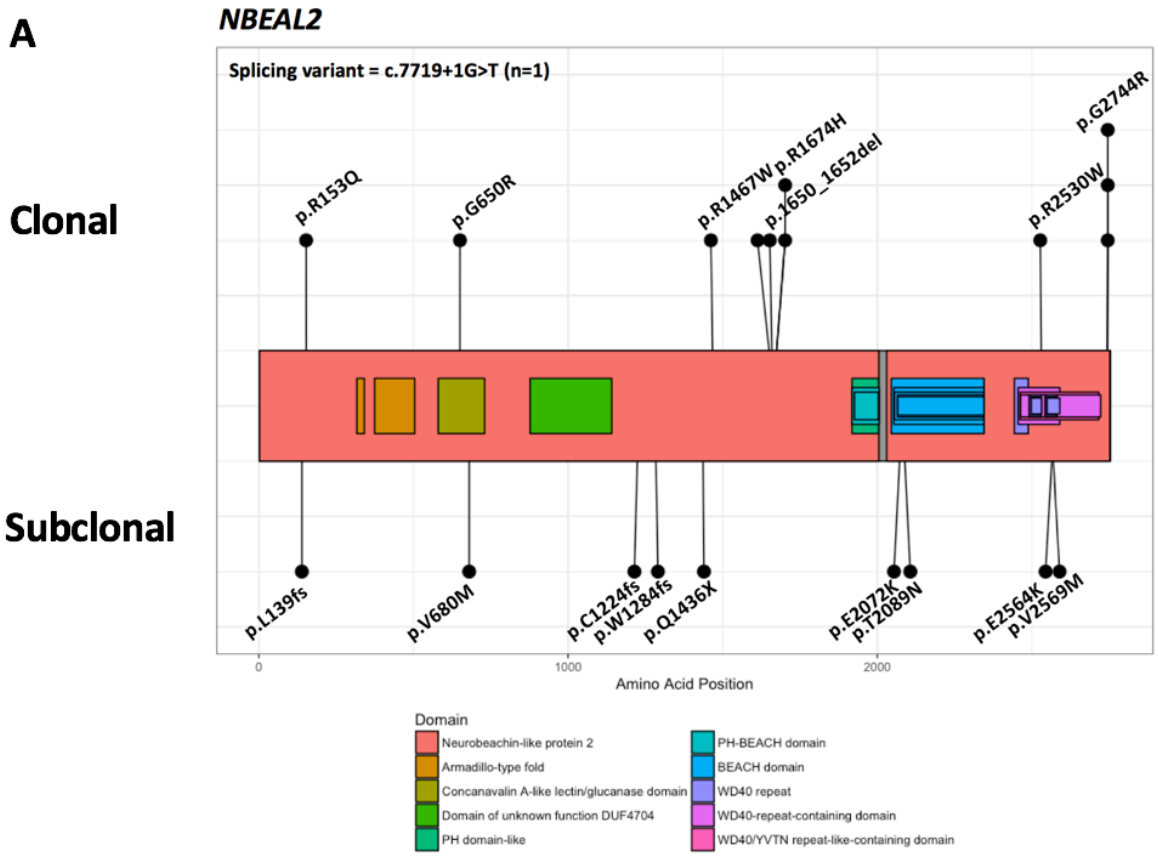
### 3.6.5 Mutated *NBEAL2* and *CTBP2* found to be recurrently mutated in CLL4: First reported case of recurrent incidence in CLL

Novel recurrently mutated genes *NBEAL2* (4%) and *CTBP2* (3%), previously reported to be mutated in CLL at extremely low frequencies were observed to be mutated at a significant frequency in this cohort. However, due to the lack of copy number data on these mutated cases at the genomic loci of *CTBP2* and *NBEAL2,* their true significance will need to be validated in external CLL cohorts with this data available. Regardless, since these genes have not been shown to be mutated at this frequency before, the observed mutations are described below.

#### 3.6.5.1 *NBEAL2* mutations in CLL4

*NBEAL2* is a member of the BEACH domain containing family, with a predicted role in megakaryocyte alpha-granule biogenesis (176). *NBEAL2* has been previously shown by the Cancer Genomics group to be part of the Minimally Deleted Region (MDR) of del(3p) in CLL, with studies showing mutations in *NBEAL2* are associated with Grey Platelet Syndrome (GPS) (176), and a role in autophagy via interactions with *LYST*(177). Mutations in *NBEAL2* occur across the entirety of the gene, pre-dominated by missense mutations **(Figure 41A)**. None of the mutations found in this study were observed in the publication which found a causal link between mutations in *NBEAL2* and GPS (176), although the mutations found in GPS also occur across the entirety of the gene. Only one of the 21 variants observed in CLL was annotated in COSMIC, in one patient with Acute Lymphoblastic Leukaemia (ALL). Although none of the variants appear in dbSNP, 7/21 are annotated in the ESP database and 3/21 in the 1000 genomes database, with all variants occurring in less than 1% of the normal populations studied **(Figure 41B)**. In combination with these variants all being observed at near 50% VAF, could suggest potential germline variants, however without matched germline material this cannot be confirmed. Therefore, the likelihood is that these mutations are not likely to have a role in CLL pathogenesis.

**A**




**B**

| CLL4 ID | *NBEAL2* amino acid change | Exon | VAF | COSMIC ID | 1000g | ESP |
|---|---|---|---|---|---|---|
| 4036 | p.G650R | 14 | 0.4832 | . | 0.0005 | 0.000893 |
| 4047 | p.E2564K | 50 | 0.0656 | . | . | . |
| 4048 | p.R153Q | 5 | 0.4839 | . | . | . |
| 4074 | p.V680M | 15 | 0.0848 | COSM1216726 | . | . |
| 4086 | p.G2744R | 54 | 0.4035 | . | . | 0.001886 |
| 4263 | p.V2569M | 50 | 0.0615 | . | . | . |
| 4265 | p.C1224fs | 25 | 0.0635 | . | . | . |
| 4272 | p.G2744R | 54 | 0.4607 | . | . | 0.001886 |
| 4347 | p.1650_1652del | 31 | 0.5096 | . | . | 0.001313 |
| 4384 | c.7719+1G>T | 50 | 0.0731 | . | . | . |
| 4473 | p.R1674H | 31 | 0.5074 | . | 0.0005 | 0.000646 |
| 4486 | p.Q1436X | 28 | 0.0778 | . | . | . |
| 4486 | p.E2072K | 38 | 0.0805 | . | . | . |
| 4494 | p.E1659K | 31 | 0.4452 | . | . | . |
| 4613 | p.G2744R | 54 | 0.467 | . | . | 0.001886 |
| 4638 | p.R1467W | 28 | 0.5215 | . | . | 0.000079 |
| 4704 | p.T2089N | 38 | 0.0606 | . | . | . |
| 4726 | p.R2530W | 49 | 0.4788 | . | . | . |
| 4737 | p.R1674H | 41 | 0.4757 | . | 0.0005 | 0.000646 |
| 4752 | p.L139fs | 5 | 0.0822 | . | . | . |
| 4772 | p.W1284fs | 27 | 0.0606 | . | . | . |

**Figure 41 *NBEAL2* mutations in CLL4**

**A** Lolliplot showing the location and amino acid change of mutations in *NBEAL2*. Variants plotted on the top of the protein map represent variants with a VAF >12% (clonal), whilst those plotted below the protein map represent variants with a VAF <12% (subclonal). **B** Table of disease and normal variation annotation for COSMIC, 1000g and ESP of all *NBEAL2* variants.

### 3.6.5.2 *CTBP2* mutations outside of pseudogene affect regions in CLL4

*CTBP2* is a transcriptional repressor, which has been shown to be implicated in Wnt signalling(178). However, the *CTBP2* gene can be transcribed into two distinct proteins; CTBP2 (the transcriptional repressor), and RIBEYE, a protein expressed in neuronal cells of the retina, and involved in the transport of vesicles to synaptic ribbons (179). Mutations in *CTBP2* were observed across all exons, but because of sequence homology between exons 2-10 to several known pseudogenes, only mutations in exon 1 were included in the final list of variants (15/903) **(Figure 42, A.5**, **A.6)**. The TruSeq design of this gene covered the *CTBP2* transcript (ENST00000337195), as well as the RIBEYE transcript (ENST00000309035). Mutations were observed to uniquely map to exon 1 of RIBEYE (mutually exclusive to exon 1 of *CTBP2*), raising the possibility that this protein could be expressed outside of the retina. However, caution must be taken, as exon 1 of RIBEYE is also an intronic region of *CTBP2*, therefore these mutations could just be intronic variants of unknown significance.

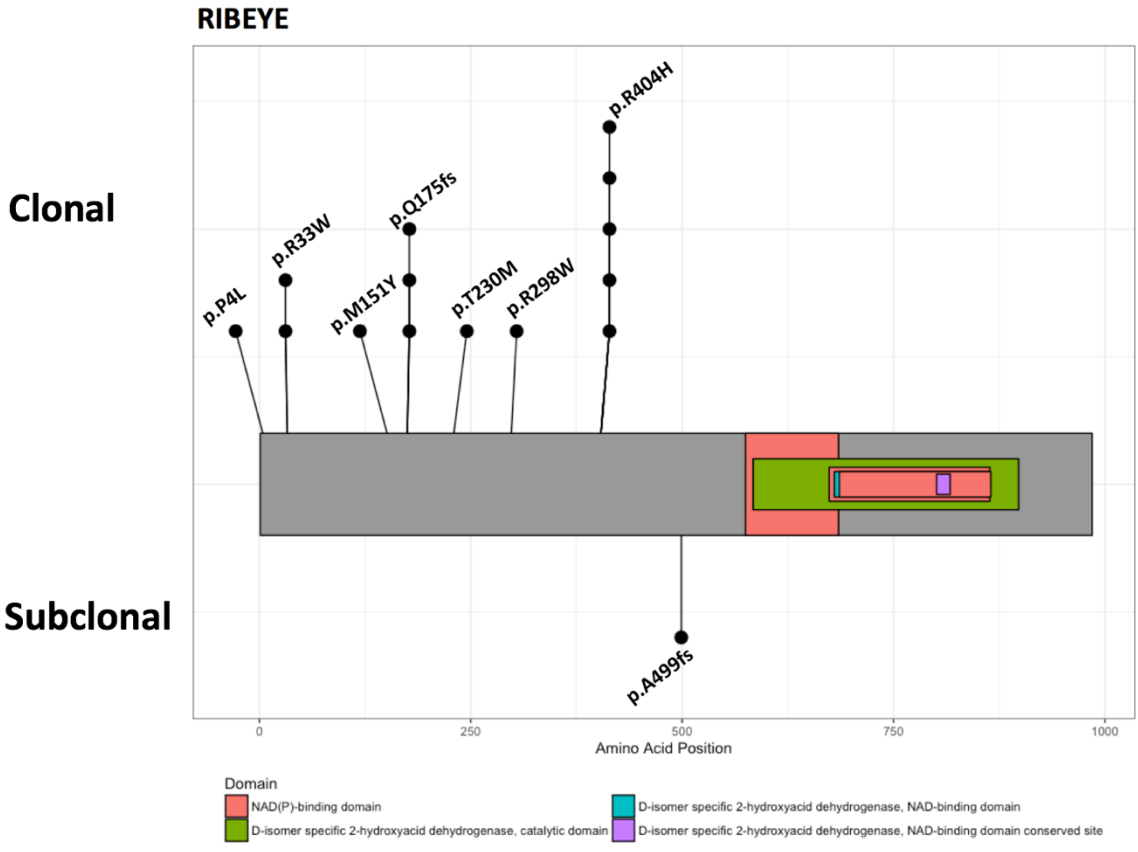



**Figure 42 Lolliplot of mutations found in *CTBP2* (RIBEYE) in CLL4**

**A** Lolliplot showing the location and amino acid change of mutations in *NBEAL2*. Variants plotted on the top of the protein map represent variants with a VAF >12% (clonal), whilst those plotted below the protein map represent variants with a VAF <12% (subclonal).

## 3.7 Identification of mutated pathways in CLL4

To study the potential impact of biological pathway dysregulation, mutated genes were positioned with their signalling pathways based on the classification published in Landau *et al.* (2016), with genes not described in that analysis prescribed a pathway based on a literature search **(Table 4)**.

**Table 4 Stratification of CLL4 panel genes into intracellular signalling pathways**

| Intracellular signalling pathway | Hugo Gene ID (bold = dual pathway members) |
|---|---|
| Inflammatory pathway/B-Cell Receptor signalling | *MYD88, BIRC3, BRAF, KRAS, NRAS, NFKBIE, EGR2, SAMHD1* |
| RNA processing | *SF3B1, RPS15, XPO1,* ***DDX3X*** |
| DNA Damage Response/Cell Cycle Control | *ATM, POT1, TP53* |
| NOTCH signalling | *NOTCH1,* ***FBXW7*** |
| Chromatin Modification | *SETD2, CHD2, HIST1H1E* |
| Wnt signalling | *CTBP2, MED12, ZFPM2* |
| Inflammatory pathway/ B-Cell Receptor signalling/RNA processing | ***DDX3X*** |
| MYC | *MGA* |
| Autophagy | *NBEAL2* |
| NOTCH signalling/MYC | ***FBXW7*** |

Dual pathway members defined as genes which have been shown to have a functional role in more than one intracellular signalling pathway.

### 3.7.1 One mutated intracellular signalling pathway

First, the distribution of the number of cases: not carrying a mutation, only carrying mutations within one intracellular signalling pathway, or carrying mutations in more than one intracellular signalling pathway, was evaluated **(Figure 43A)**. This showed that most cases had mutations in genes in more than one intracellular signalling pathway (n=213), followed by 181 cases having one mutated pathway, and 106 cases not carrying a mutation. For cases only carrying mutations within one intracellular signalling pathway, an assessment of the number of mutated genes per case was conducted. This analysis identified less than 8 cases per intracellular signalling pathway had mutations in more than one gene (**Figure 43B**, red bars).

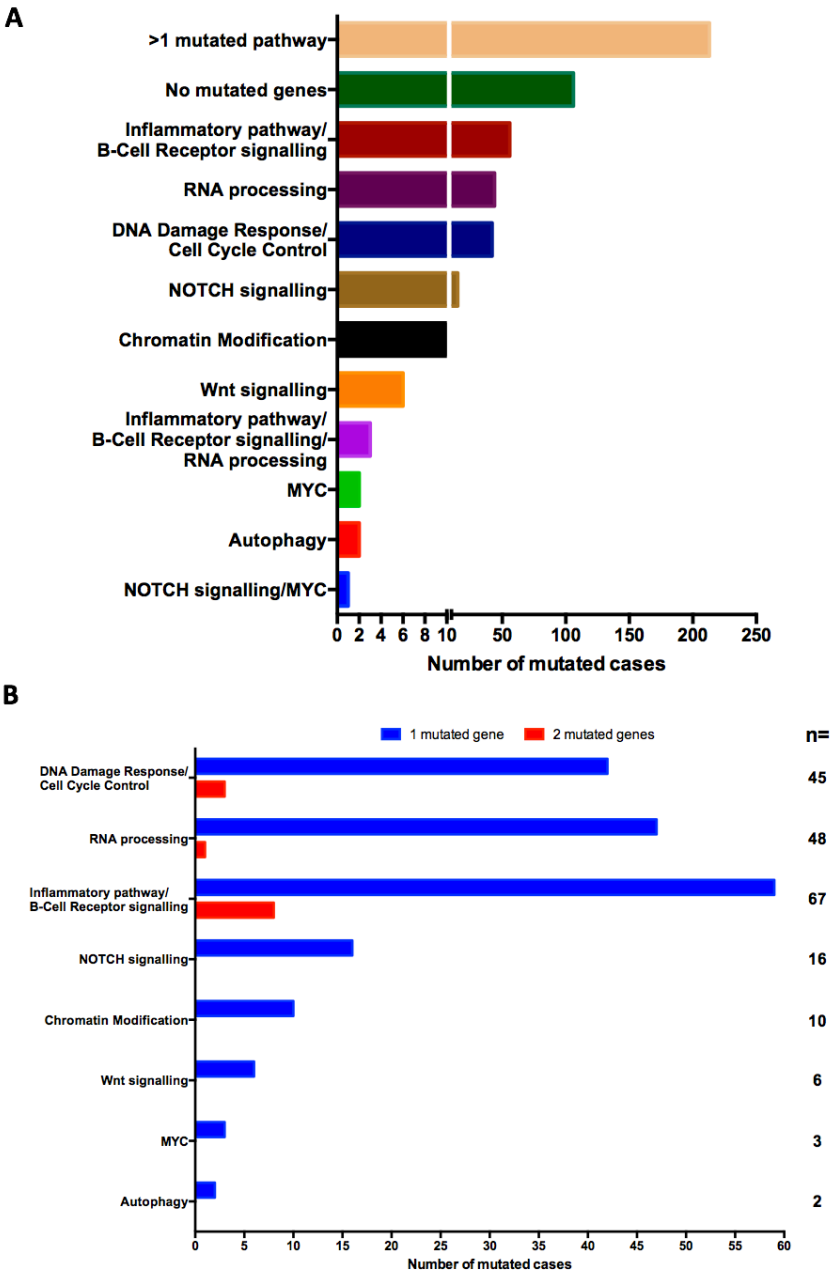



**Figure 43 Overview of mutated intracellular signalling pathways in CLL4**

**A** Number of mutated cases, stratified by number and type of intracellular signalling pathway. **B** Number of cases in each mutated pathway with 1 or 2 mutated genes.

### 3.7.2 More than one mutated intracellular signalling pathway

Next, the proportion of cases with two mutated intracellular signalling pathways was calculated, with each combination targeted in greater than >10 cases being taken forward for heat-map analysis, with the defined dual pathway genes (those found to have a functional role in more than one intracellular signalling pathway) excluded due to the confounding impact of assessing two pathways with the same mutation data within it **(Table 5, A.10, Figure 44)**. The shared mutated cases of the DNA Damage Response (DDR)/Cell Cycle Control (CC) and RNA processing pathways were dominated by *SF3B1* mutations (33/41) **(Figure 44A)**, with *ATM* accounting for most mutations in the DDR/CC and inflammatory pathway/BCR signalling group (17/27) **(Figure 44B)**. Within the RNA processing and inflammatory pathway/BCR signalling group the distribution was more even, with mutations in *SF3B1, RPS15* and *NFKBIE* contributing to 8, 7, and 7 out of the 17 cases, respectively **(Figure 44C)**. Finally, there was only one group of cases with three mutated intracellular signalling pathways in >10 cases (DDR/CC, RNA processing, inflammatory pathway/BCR signalling). Five genes within this group were found to be mutated in 5 or more of the 13 cases; *ATM* (n=5), *TP53* (n=5), *SF3B1* (n=7), *RPS15* (n=5) and *NFKBIE* (n=5) **(Figure 44D)**.

**Table 5 Breakdown of 2 mutated intracellular signalling pathways**

| Pathway 1 | Pathway 2 | Number of cases | Percentage |
|---|---|---|---|
| DNA Damage Response/Cell Cycle Control | RNA processing | 41 | 30.60 |
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | 27 | 20.15 |
| Inflammatory pathway/B-Cell Receptor signalling | RNA processing | 17 | 12.69 |
| DNA Damage Response/Cell Cycle Control | NOTCH signalling | 8 | 5.97 |
| Inflammatory pathway/B-Cell Receptor signalling | NOTCH signalling | 8 | 5.97 |
| Chromatin Modification | Inflammatory pathway/ B-Cell Receptor signalling/RNA processing | 6 | 4.48 |
| DNA Damage Response/Cell Cycle Control | Wnt signalling | 4 | 2.99 |
| NOTCH signalling | RNA processing | 4 | 2.99 |
| Chromatin Modification | DNA Damage Response/ Cell Cycle Control | 3 | 2.24 |
| Inflammatory pathway/B-Cell Receptor signalling | Wnt signalling | 3 | 2.24 |
| Autophagy | DNA Damage Response/Cell Cycle Control | 2 | 1.49 |
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | 2 | 1.49 |
| Inflammatory pathway/B-Cell Receptor signalling | Inflammatory pathway/ B-Cell Receptor signalling/RNA processing | 2 | 1.49 |
| Autophagy | Inflammatory pathway/ B-Cell Receptor signalling | 1 | 0.75 |
| Autophagy | NOTCH signalling | 1 | 0.75 |
| Autophagy | RNA processing | 1 | 0.75 |
| Inflammatory pathway/B-Cell Receptor signalling | NOTCH signalling/MYC | 1 | 0.75 |
| Inflammatory pathway/ B-Cell Receptor signalling/RNA processing | NOTCH signalling | 1 | 0.75 |
| Inflammatory pathway/ B-Cell Receptor signalling/RNA processing | RNA processing | 1 | 0.75 |
| NOTCH signalling/MYC | Wnt signalling | 1 | 0.75 |

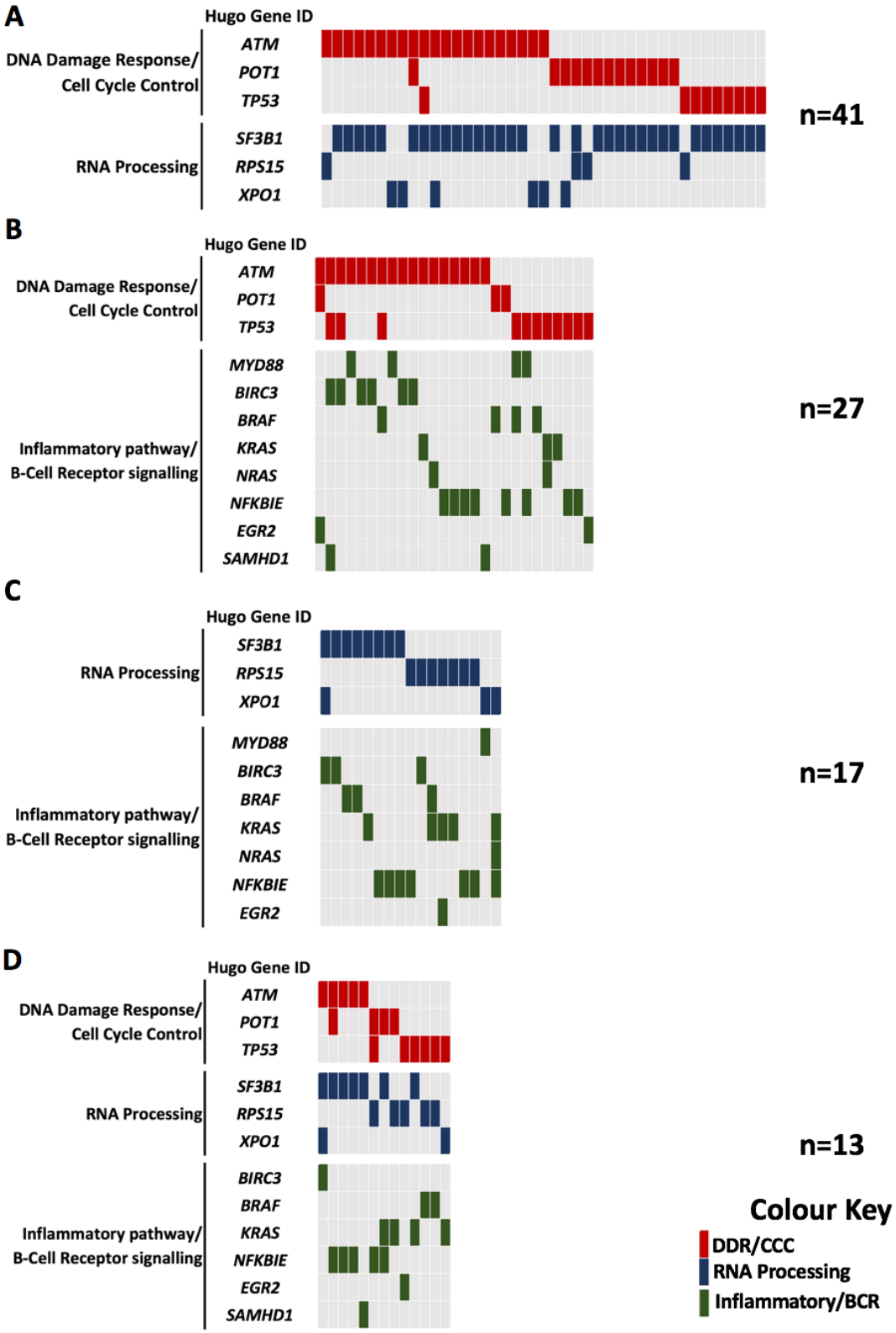



**Figure 44 Heat-maps of co-mutated intracellular signalling pathways**

Heat map analysis of cases co-mutated between 2 or 3 intracellular signalling pathways in >10 cases. **A** DNA Damage Response/Cell Cycle Control vs. RNA Processing. **B** DNA Damage Response/Cell Cycle Control vs. Inflammatory pathway/B-Cell Receptor signalling. **C** RNA Processing vs. Inflammatory pathway/B-Cell Receptor signalling. **D** DNA Damage Response/Cell Cycle Control vs. RNA Processing vs. Inflammatory pathway/B-Cell Receptor signalling. Red bars represent mutations in DDR/CC signalling genes, blue bars represent mutations in RNA processing genes, green bars represent mutations in BCR genes, and grey bars represent cases unmutated for the gene in that row.

## 3.8 Statistical association analysis of recurrently mutated genes in CLL4

As shown above, specific recurrently mutated genes in the CLL4 dataset have associated with one another, as is the case with *SF3B1* and *POT1* **(Figure 39A)**. However, this association was discovered through a targeted analysis, since the high mutation frequency of *SF3B1* in this cohort warranted further investigation. To be certain that all the potentially important statistical associations between recurrently mutated CLL genes are identified, an unbiased approach is required. Therefore, all mutation data was combined with the known CNV data from the clinical trial (using the raw clone size data from the original trial FISH analysis, without a cut off), with analysis conducted on the full dataset, as well as being stratified by; clonal events (>12% VAF/clone size), and subclonal events (<12% VAF/clone size). This data frame was then taken forward for correlation matrix analysis and Fisher's Exact testing **(2.2.6.4)**.

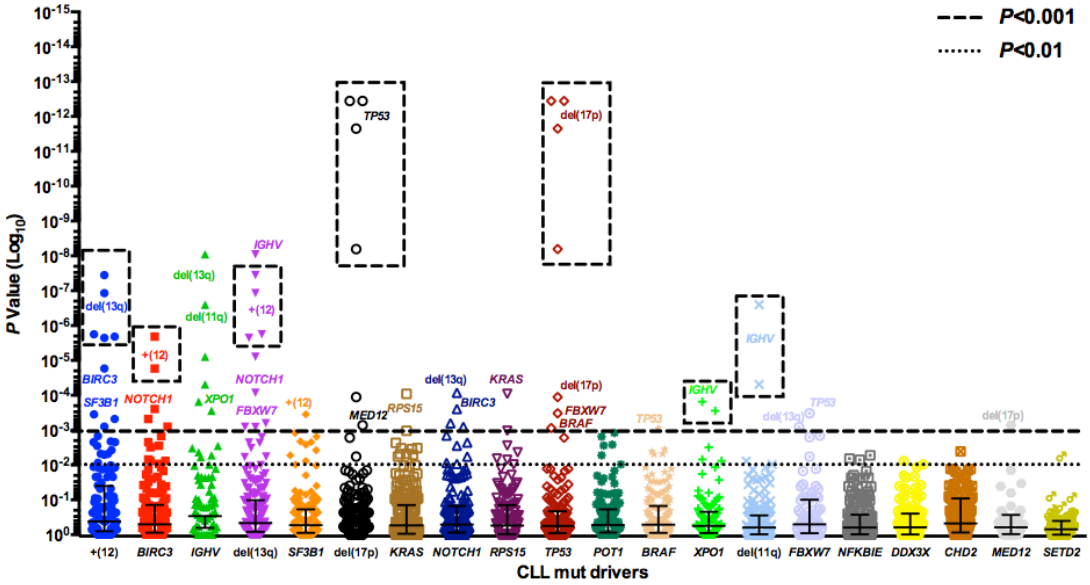
A total of 2800 Fisher Exact tests were conducted across all variables **(Figure 45A)**. Due to the high volume of multiple hypothesis tests, strict significance cut offs need to be applied to limit the effect of Type I errors. Therefore, the False Discovery Rate (FDR) *Q* value was ascertained based on a *P* value of <0.01, by ranking the tests by their *P* value **(2.2.8.2)**. Three levels of significance were set; *Q>P* (*Q* is greater than *P*, very high confidence), *P*<0.001 (high confidence), and *P*<0.01 (intermediate confidence). Very high confidence associations were observed for previously identified CLL interactors, such as; *TP53* and del(17p) ($P = 3.57x10^{-13}$), *BIRC3* and +(12) ($P = 1.70x10^{-5}$), as well as *RPS15* and *KRAS*, of which *RPS15* mutated CLL has been previously shown to correlate with overexpression of the Ras pathway genes *RASGRP4* and *RASGRF1* ($P = 9.03x10^{-5}$) **(Figure 45A, A.11)**.

Within the intermediate to high confidence significance levels, several novel associations were observed **(Figure 45B-D, A.11)**. The pan cancer genes *KRAS* and *BRAF* are known to be mutually exclusive in Colorectal Cancer (CRC) (180) and melanoma (181), however 5 cases from this cohort were co-mutated, including combinations of the hotspot mutations; *KRAS* p.G13D and *BRAF* p.V600E ($P = 3.49x10^{-3}$) **(Figure 45B)**.

Mutated *FBXW7* was found to be associated with subclonal mutations in *TP53* **(Figure 45C**; $P = 3.26x10^{-4}$**)**. An *in vivo* lymphoma study assessing the role of *Tp53* and *Fbxw7* conditional loss, found that mice that were *Tp53* null and heterozygous for *Fbxw7* had accelerated tumorigenesis post irradiation in comparison to mice which were *Tp53* null and *Fbxw7* wild type. Therefore, the abnormal status of *TP53* (*TP53* mutation as well as del(17p) CNV, termed *TP53*ab) was assessed in the context of *FBXW7* mutated cases, and the functional impact of subclonal *TP53* mutations ascertained **(Figure 45C)**. Of the four co-mutated cases, two were found to be *TP53*ab, whilst the other two cases had unknown or partial TP53 function.

Mutations in *FBXW7* and *NOTCH1* have previously been shown to be mutually exclusive in CLL, however in this cohort were found to be significantly associated **(Figure 45D**; $P = 5.72x10^{-3}$). In paediatric T-CLL Acute Lymphoblastic Leukaemia (ALL), *FBXW7* mutated cases which co-occur with *NOTCH1* mutations have been shown to have a superior OS than cases with mutations in either gene alone. One hundred percent of the co-mutated cases carried the recurrent *NOTCH1* p.2514fs variant, which accounts for 48% of all *FBXW7* mutated cases.

**A**

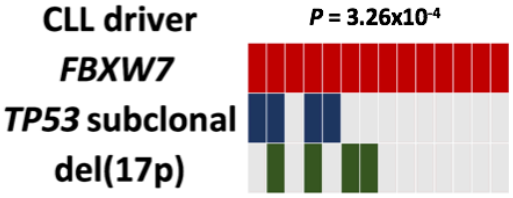



**B**

| CLL4 ID | *KRAS* variant | *KRAS* VAF | *BRAF* variant | *BRAF* VAF |
|---|---|---|---|---|
| 4365 | p.G13D | 0.0149 | p.V600E, p.L597Q, p.G469A | 0.1191, 0.2281, 0.0209 |
| 4432 | p.G60S | 0.3498 | p.D594G | 0.4585 |
| 4736 | p.Q22K | 0.0213 | p.K601E, p.V600E, p.G469A | 0.0181, 0.1431, 0.0456 |
| 4742 | p.I24N | 0.0217 | p.G466E | 0.5068 |
| 4745 | p.G13D | 0.0314 | p.V600E | 0.2729 |

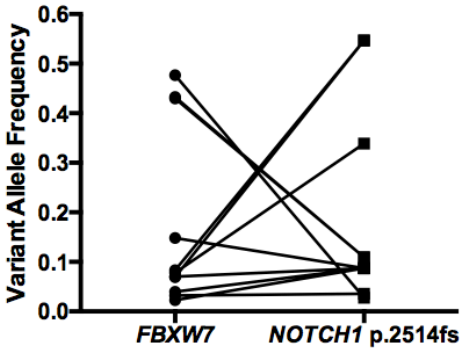
$P = 3.49 \times 10^{-3}$

**C**




| *FBXW7* mutant | *TP53* subclonal variant | VAF | Exon | del(17p) | IARC Annotation |
|---|---|---|---|---|---|
| 4263 | p.K382fs | 0.0355 | 11 | wt | Not found |
| 4267 | p.R342X | 0.0246 | 10 | deleted | Not found |
| 4419 | p.E286X, p.H179Y | 0.0388, 0.0256 | 8, 5 | wt | Not found, partially functional |
| 4432 | p.R273H | 0.0208 | 8 | deleted | Non-functional |

**D**




| *FBXW7* mutant | *NOTCH1* variant | *NOTCH1* VAF |
|---|---|---|
| 4216 | p.2514fs | 0.0278 |
| 4263 | p.2514fs | 0.036 |
| 4267 | p.2514fs | 0.1094 |
| 4419 | p.2514fs | 0.0871 |
| 4486 | p.2514fs | 0.5465 |
| 4488 | p.2514fs | 0.3387 |

$P = 5.72 \times 10^{-3}$

**Figure 45 CLL mut driver association analysis**

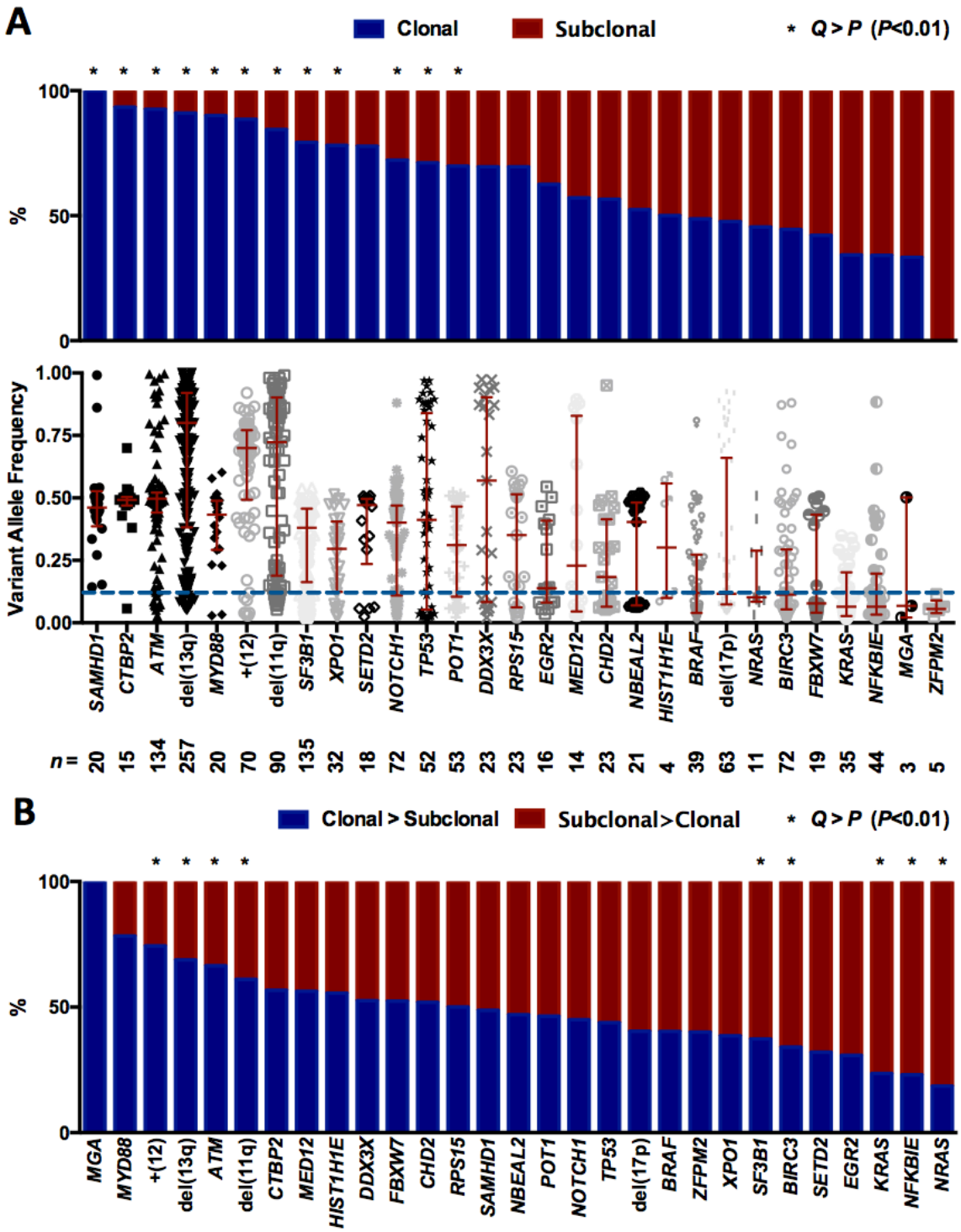
**A** Manhattan plot of all CLL drivers with at least one intermediate significant Fisher's Exact test association (*P*>0.01). **B** Table of *KRAS* and *BRAF* co-mutated cases, showing the amino acid change and VAF in each case. **C** Association between *FBXW7* and *TP53* subclonal mutations. Heat map of *FBXW7* mutated CLL (left), table of *FBXW7* and *TP53* subclonal co-mutated cases, showing the *TP53* amino acid change, VAF, exon, del(17p) status, and functional effect from IARC (right). **D** Association between *FBXW7* and *NOTCH1* p.2514fs co-mutated cases, showing the VAF relationship between both genes (left) and a table of the CLL4 IDs and *NOTCH1* p.2514fs VAF values (right).

## 3.9 Subclonal architecture of CLL4

Assessing the subclonal architecture of disease is important, as it allows the identification of whether recurrently mutated genes are predominantly found to be clonal or subclonal. This is important, since subclonal mutations have been shown to undergo clonal expansion in response to therapy, impacting patient survival (8,96,159). Furthermore, it provides an insight into whether specific genetic events might be evolutionarily linked.

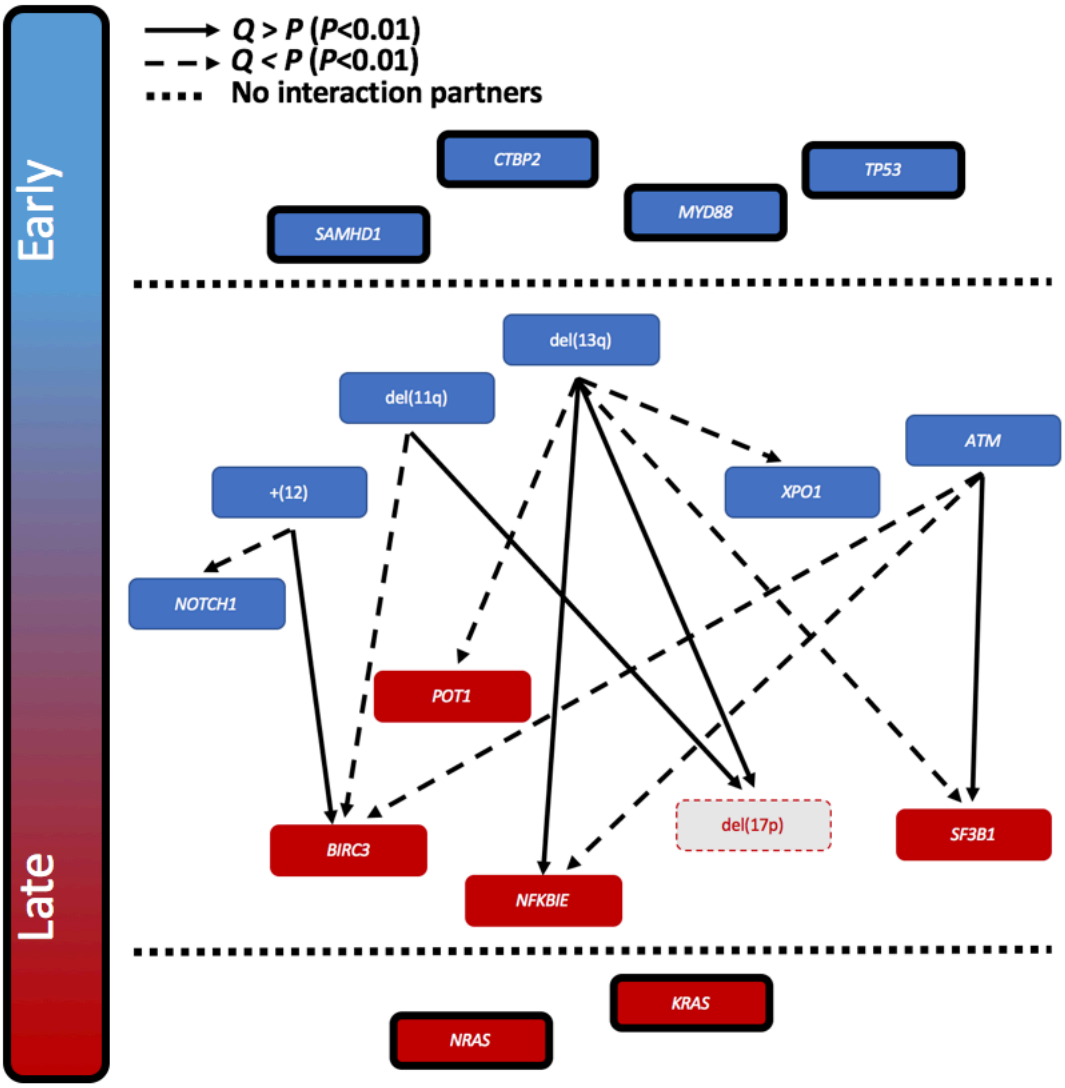
Using WES, mutated genes and CNVs have had their clonal evolution inferred in CLL, through the identification of events as "early" or "late", defined by the Cancer Cell Fraction (CCF) (96). Since the CLL4 dataset does not have matched copy number data for all loci, the CCF cannot be calculated. Furthermore, the CLL4 clinical trial did not capture tumour purity data (percentage of CD19 x CD5 positive cells within the PBMC sample), therefore this could not be applied to calculate a cancer associated VAF. However, there is evidence in the literature that to use the limit of Sanger detection (12% VAF) as a cut off to define clonal (>12% VAF) and subclonal (<12% VAF) variants (as used throughout this chapter) has clinical utility (159). These cut offs were applied, and the proportion of subclonal and clonal mutations per CLL driver was calculated, along with the distribution of VAF **(Figure 46A)**. This led to the identification of 12 recurrently mutated genes composed significantly of clonal events.

Next, all cases with more than one recurrently mutated gene were tested by Pearson's chi-squared contingency tables to identify those which contained mutations at significantly different VAFs, by using the raw total read and alternate read depth data. Three hundred and ninety cases passed these criteria at a significance cut off of $P<0.01$. After this, in-going (subclonal to clonal) and out-going (clonal to subclonal) edges were drawn between all variants within each case, and enumerated to calculate how often each driver occurs at an early (higher VAF) or late (lower VAF) event. This analysis identified 4 significant early drivers (del(13q), +(12), *ATM* and del(11q), and 5 significant late drivers (*NFKBIE, KRAS, BIRC3, SF3B1* and *NRAS*) **(Figure 46B).** Finally, the number of edges between specific interaction partners was calculated on a per case basis, and combined to provide a picture of the clonal evolution in this cohort **(Figure 46C&D)**. Five highly significant interactions were observed ($Q>P$), with an additional 7 interactions which are still highly significant, but caution should be taken in their interpretation because the level of multiple hypothesis testing conducted ($Q<P$, $P<0.01$) **(Figure 46D)**. Multiple hypothesis testing was controlled for by calculating the False Discovery Rate (FDR) and inferring the $Q$ value as previously described (96). This analysis identified interactions previously observed in Landau *et al.* (2015), namely del(13q) with *SF3B1* or *POT1,* and +(12) with *BIRC3* (96).




| Driver | *P*-value | *Q*-value | Clonal inference |
|---|---|---|---|
| del(13q) | 1.55E-14 | 3.45E-04 | Early |
| +(12) | 5.94E-11 | 6.90E-04 | Early |
| *ATM* | 2.14E-09 | 1.03E-03 | Early |
| *NFKBIE* | 2.53E-09 | 1.38E-03 | Late |
| *KRAS* | 2.70E-08 | 1.72E-03 | Late |
| *BIRC3* | 3.36E-07 | 2.07E-03 | Late |
| *SF3B1* | 9.09E-06 | 2.41E-03 | Late |
| del(11q) | 3.00E-04 | 2.76E-03 | Early |
| *NRAS* | 6.01E-04 | 3.10E-03 | Late |

C




D

| Driver interactions | *P*-value | *Q*-value | Statistical significance |
|---|---|---|---|
| +(12)>*BIRC3* | 6.87E-05 | 1.72E-04 | High |
| del(13q)>*NFKBIE* | 6.87E-05 | 3.45E-04 | High |
| *ATM>SF3B1* | 1.33E-04 | 5.17E-04 | High |
| del(13q)>del(17p) | 6.85E-04 | 6.90E-04 | High |
| del(11q)>del(17p) | 8.54E-04 | 8.62E-04 | High |
| +(12)>*NOTCH1* | 1.04E-03 | 1.03E-03 | Intermediate |
| del(13q)>*POT1* | 1.83E-03 | 1.21E-03 | Intermediate |
| *ATM>BIRC3* | 1.95E-03 | 1.38E-03 | Intermediate |
| del(13q)>*SF3B1* | 4.38E-03 | 1.55E-03 | Intermediate |
| *ATM>NFKBIE* | 5.37E-03 | 1.72E-03 | Intermediate |
| del(11q)>*BIRC3* | 5.37E-03 | 1.90E-03 | Intermediate |
| del(13q)>*XPO1* | 9.77E-03 | 2.07E-03 | Intermediate |

**Figure 46 Inferred clonal evolution of CLL4**

**A** The frequency that a CLL driver was found as clonal (>12% VAF) or subclonal (≤12% VAF) across 500 CLL4 cases is presented (top), along with the distribution of driver's VAF values (bars represent median and interquartile range, blue dashed line representing the 12% clonal/subclonal cut off) (bottom). **B** The frequency of out-going edges (clonal>subclonal) and in-going edges (subclonal to clonal) in 390 CLL4 cases with >1 CLL driver (after accounting for missing CNV data) is presented (top), along with a table showing the significant drivers, including whether they were found to be early or late evolutionary events. **C** Direct edges between two drivers within the sample were drawn when they occurred at a different VAF. These edges are summed and tested to infer the clonal evolution in the CLL4 cohort. Blue boxes denote early drivers, red late drivers, and grey boxes intermediate/not powered drivers. **D** Table denoting the significant interactions observed in **C**, along with their *P*- and *Q*-values as well as the strength of significance. Drivers were tested for being significantly composed of clonal or subclonal events/edges/interactions using a two-tailed binomial test, followed by calculation of the *Q*-value to account for multiple hypothesis testing. Drivers were considered significant if their *Q*-value was greater than their *P*-value **(A&B).** Driver interactions were considered high confidence if their *Q*-value was greater than their *P*-value, and intermediate confidence if their *P*-value was <0.01 **(C)**. All multiple hypothesis testing was conducted using a cut off *P*<0.01. See methods and Appendix for more information.

## 3.10 Discussion and Future Directions

Recurrently mutated genes in CLL can affect patient prognosis and response to treatment, underlying the importance of screening for and understanding the role of these mutations within clinical trial cohorts. Mutations in *TP53, ATM, SF3B1* and *NOTCH1* have previously been shown to be clinically important (7,61,182), however the evidence for genes such as *BIRC3, RPS15, EGR2* as well as others in CLL is not fully established, pertaining to the application of NGS targeted re-sequencing in a clinical trial cohort to address these questions.

In this chapter, curation of the LRF CLL4 trial cohort for available material was undertaken, identifying a total cohort of 500 cases from the first randomisation with available mutation data. This cohort represents one of the largest in the literature to assess the mutation landscape of many previously reported recurrently mutated genes and putative candidates in CLL together in the context of a clinical trial. After characterisation of the sequencing coverage and mutation data processing, 903 high quality variants were identified across the 25 genes. *SF3B1* was found to be the most recurrently mutated gene across the entire cohort. Co-occurrence analysis within the U-CLL, M-CLL, and *SF3B1* mutated subsets identified previously observed associations with *KRAS, RPS15,* and *XPO1* in U-CLL, and a novel association of *POT1* in *SF3B1* mutated CLL. Finally, pathway analysis, statistical associations between recurrently mutated genes and subclonal architecture of the CLL4 cohort was described, identifying groups of co-mutated cases, that provided an insight into the inferred clonal evolution of the CLL4 TruSeq cohort.

### 3.10.1 CLL4 biological material

When conducting mutation screening experiments, the starting biological material is an important feature which can limit the eventual downstream analysis of mutation data. In the CLL4 TruSeq study, the starting material included: high quality and concentration DNA samples, high quality but low concentration DNA samples, and cells frozen in liquid nitrogen. The DNA samples had been extracted from PBMCs, and therefore it is likely that they also contain a proportion of non-CLL cells. Therefore, all additional samples extracted from viable cells had to be processed in the same way to be consistent. However, when screening for mutations and aiming to use the VAF to classify mutated cases into subclonal (<12% VAF) and clonal (>12% VAF) groups, precise measurement of the VAF is vital, and could be affected by “contaminating” non-CLL cells. The best way to account for this would be to sort the PBMCs for positivity of CLL markers CD19 and CD5, providing a population of double positive cells to extract the DNA from. However, considering this approach was not suitable for this project, other options to assess tumour purity of the CLL4 DNA samples were explored.

Although there was flow cytometry data available for CD19 and CD5 at the start of the trial (67), dual CD19 x CD5 staining was not part of the original CLL4 trial protocol, and was not introduced until later into the trial. The percentage of lymphocytes in each sample was recorded after Ficoll extraction, but this value represents the total percentage of lymphocytes, not the total percentage of CLL cells. To assess whether lymphocyte percentage was comparable to CD19 CD5 tumour purity, screening on a small amount of cases by CD19 x CD5 flow cytometry was conducted, but unfortunately the cell viability of the samples assessed was not suitable for this analysis, therefore no comparative analyses could be conducted.

### 3.10.2 Sequencing data quality and coverage of CLL4 TruSeq panel

Confidence that the sequencing coverage obtained in NGS studies is of acceptable quality and depth is paramount. However, regardless of the design, there will be poor performing and good performing regions, as has been observed here. Three iterations of a CLL TruSeq custom panel were used for this chapter, with their main design difference being the exclusion of known SNP sites. However, that has led to differences in sequencing coverage. Furthermore, GC content was shown to significantly impact the obtained read depth, especially for genes such as *NFKBIE* and *NOTCH1.* This is not just a phenomenon of this CLL TruSeq custom panel, as the Landau *et al.* (2015)(96) WES study implemented a TruSeq targeted panel to capture *NOTCH1* mutations after observing poor coverage in their WES data, whilst Puente *et al.* (2015)(57) applied Sanger sequencing on its WGS cohort due to similar problems. It could be argued that if Landau *et al.* (2015) generated good *NOTCH1* coverage using TruSeq that it is a specific problem with this design, however the number of amplicons on their panel and the number of genes covered, as well as the number of cases sequenced per run was not published, therefore this comparative analysis to understand more about the relative performances of both kits is not possible. Due to the many studies that have come across this problem, this analysis identifies that it is the high GC content of the region which does not sequence well, and this is especially a feature of Illumina sequencing technologies. To try and account for these limitations, quantitation of the GC content of each amplicon was conducted, providing information on the amplicons and genes poorly effected by GC content. This allowed assessment of whether hotspot mutations known to be important in CLL were found in these amplicons, which could potentially impact capturing hotspot variants. Only two such variants were identified, *NOTCH1* P2514fs, and *NFKBIE* T253fs. Manually screening of these locations across all 500 patients identified variants missed by the bioinformatics pipeline, therefore justifying the inclusion of this troubleshooting process to screen for additional mutations.

The individual amplicon performance analysis identified amplicons that clustered into distinct performance groups. This potentially has implications on downstream data analysis, since if not every designed amplicon provides data of sufficient quality to identify a mutation, can that case be called wild type for the gene in question. To address this, missing data analysis of *TP53,* the gene with the strongest impact on survival in CLL within the 25 genes of the TruSeq panel, was conducted. This analysis showed that there were many cases which could not capture mutation data for every exon of *TP53,* but after accounting for previously observed *TP53* mutations, as well as del(17p) events, no differences between cases with missing data and those without were observed. Due to this observation, it was suggested that missing data does not significantly impact the identification of poor risk cases in *TP53,* and that this should be applied to the rest of the genes on the panel. This analysis was not conducted on the rest of the panel, because their known impact on survival, and association with clinical biomarkers of the CLL is not fully understood. Therefore, attempting such an analysis in genes other than *TP53* would not add benefit.

The only way to be truly certain about the downstream analysis of NGS data with missing data, is to use complete case analysis, defining any case which is wild type for mutations in a gene, but with missing data for any region, as missing. However, this is impractical for several reasons. Firstly, due to the number of missing amplicons across the panel, this would lead to a severe reduction in the total number of cases to take forward for survival analysis, reducing the power of the study to assess associations of mutated genes with survival outcome. Secondly, many NGS studies have reported mean coverage statistics comparable to those observed in this chapter, therefore it is likely that these studies have come across similar technique limitations. Finally, as will be presented in **Chapter 4:**, survival analysis ignoring the impact of missing data, still identifies associations between genes with OS and PFS that have previously been observed in the literature. This suggests that the impact of missing data in individual cases is minimal in comparison to the strength of associations between observed mutated genes and survival outcome.

Furthermore, the available concentration of DNA from the CLL4 samples may impact the downstream analysis of the mutation data too. This is due to the technical capacity of the Illumina TruSeq methodology, since it has two levels of starting concentration for library preparation **(2.1.12)**. Although high quality library preparations from both DNA inputs were created, it is possible that the sequencing depth could have been effected. Since there was a strict filtering criteria for the inclusion of subclonal mutations based on the total read depth relative to the mutant allele read depth, real subclonal mutations could have been missed due to insufficient sequencing coverage.

### 3.10.3 Bioinformatics and mutation data processing of CLL4 sequencing data

Identifying mutations from NGS data is a computationally intensive process, involving alignment of the raw sequencing data, pre-processing steps using publicly available tools to identify the mutation types, and variant annotation using publicly available tools and databases to provide: biological, clinical and *in silico* mutation predication information. However, once an initial list of variants is produced, it needs to be filtered to identify potential disease associated variants amongst the large number of variants initially identified. This full list of variants, dependent on the technique used, will contain many false positives, including sequencing artefacts. Therefore, manual curation via visualisation of mutations in IGV is required, to be able to separate out true positives from false positives.

In this chapter, 217,589 variants were identified using the standard methodology, which was reduced to 12,429 after applying the aforementioned filtering strategies. For 500 cases, this is a considerable number of variants to take forward for manual curation, especially considering the low mutation burden of CLL (0.5/Mb). Therefore, additional strategies to assess the variant lists and remove recurrent false positives were conducted. This was a labour-intensive process, involving the manual curation of many variants in a training phase, followed by the identification of recurrent false positives per run, using an iterative process of working out the frequency of recurrent variants to use as a cut off, whilst keeping all the manually curated variants identified in the training phase. However, this process allowed the reduction of the dataset to a number that was suitable for manual curation. Even so, over 400 of these variants were found to be false positive by IGV manual curation, emphasising the difficulty in processing mutation data from NGS studies. This shows the importance of implementing bioinformatic processes that are bespoke to the sequencing data generated, and how when dealing with large volumes of data, troubleshooting measures to reduce the size of the data are required.

### 3.10.4 Orthogonal sequencing and molecular validation of CLL4 mutations

Validation of mutation data from NGS studies is required to improve data confidence. This issue has been addressed in the literature in multiple ways(183). Classical validation of clonal variants by Sanger sequencing is common(33,57,96,159), however this technique is not applicable to subclonal mutations which fall below the sensitivity of this sequencing technology. Validation via sensitive PCR techniques, such as ddPCR, has been implemented in the validation of subclonal mutations, however these techniques are highly specific for each variant location and therefore represent a considerable financial burden. Serial dilution experiments of clonal mutations for multiple genes showed sensitivity to correctly identify the mutation at low levels, with the

expected and observed VAFs being highly concordant(8). NGS library replication (sequencing the same cases using the same NGS technology more than once) has also been presented(159), however this does not account for potential amplification bias that might occur during NGS library preparation and sequencing using a particular technology. Therefore, repeating the same cases using an orthogonal NGS chemistry has been suggested to overcome these concerns(183).

In this chapter, application of orthogonal NGS chemistries were used to validate subclonal *TP53* mutations (Ion Torrent, SureSelect), clonal and subclonal mutations in *ATM, SF3B1,* and *NOTCH1* (SureSelect), as well as *SF3B1* K700E and *NOTCH1* P2514fs variants using ddPCR. Furthermore, clonal *TP53, SF3B1,* and *NOTCH1* mutations were validated using Sanger sequencing. This validation highlighted the importance of orthogonal sequencing of subclonal variants, as many of the variants identified by TruSeq were not observed in the validation sequencing experiment. Due to the association found between a variant being identified by orthogonal sequencing and having cancer associated mutation annotation (i.e. from COSMIC), this allowed additional conservative filtering of low VAF variants in any gene on the panel that did not have published supportive evidence of being somatically acquired in cancer. Orthogonal sequencing of all subclonal mutations of all genes was beyond the scope of this thesis. The variants which were identified by TruSeq, *in silico* validated in IGV, but not found by orthogonal sequencing, offer an interesting insight into the limitations of NGS. Most likely, the TruSeq variants were the product of PCR bias, where a single DNA molecule, carrying a variant, was amplified preferentially. This could be overcome by using Unique Molecular Indexes (UMIs), which bind to each fragmented DNA molecule, allowing the removal of all reads which have the same UMI tag (184). Naturally this is a process that needs to be considered during the study design process, which for this project pre-dates the commercial availability of UMIs.

### 3.10.5 Available data from previous CLL4 publications

Along with the data generated from the clinical trial, there was available mutation data for *TP53, SF3B1,* and *NOTCH1*, which has proved invaluable as a comparative tool to validate the sensitivity of the Illumina TruSeq custom panel technology. Although some mutations were missed due to technical reasons (i.e. lack of capture of part of exon 6 in *TP53* due to design changes), there was strong concordance between the variants that were found at VAFs that surpass the sensitivity threshold of the previous studies sequencing technique. However, although FISH data on key CNVs was available from the original trial, as well as SNP6.0 data on a handful of cases, copy number status of the CLL4 TruSeq cohort is a considerable disadvantage to the study. This is because that without that additional genetic information, true assessment of the mutation frequency and significance of recurrently mutated genes using MutSig is not possible.

Furthermore, without complete copy number status, true assessment of the cancer cell fraction (CCF) using ABSOLUTE cannot be undertaken (171), and therefore whether a variant is truly clonal or subclonal is also not possible to ascertain.

### 3.10.6 Mutation landscape of CLL4

One of the largest studies published to date, Landau *et al.* (2015) (96), presented germline matched WES data on a sub-cohort of the CLL8 clinical trial (n=278, FC vs. FCR, Phase 3 clinical trial), with additional matched relapse associated time point WES data (n=59). This study, including further WES data from a diagnostic cohort boosting the total cohort size to 538 patients, also had rich RNA-Seq data, however this was not conducted on patients from the CLL8 cohort. Landau *et al.* (2015) (96) identified 44 recurrently mutated genes (n=20 on TruSeq CLL panel, genes not identified on TruSeq CLL panel: *CTBP2, NBEAL2, SETD2, NFKBIE* and *ZFPM2*), and 11 recurrently somatic copy number alterations. Clonal evolution was identified as a frequent event in the matched WES pre-treatment and post-relapse cohort, with CLL drivers that are stable in response to treatment (del(11q), del(13q), +(12), *XPO1*), have increasing cancer cell fraction (CCF) after relapse (*TP53,* del(17p)), or showing 'shifting' CCF after relapse (*SF3B1, ATM, DDX3X, POT1, NOTCH1, SAMHD1, KRAS, MGA* and *BIRC3*).

This chapter identified 18 genes to be recurrently mutated in this cohort, including genes that previously haven't been found at such mutation frequencies. Furthermore, it has described the associations between different mutations and pathways, confirming previously observed associations, as well as offering unique insight into subgroups of co-mutated cases. This is something that has enriched this study, since the Landau *et al.* study and others had focused on the known clinical features for their co-occurrence analyses. This is also a feature that has been conducted, but will be presented throughout the clinical and statistical survival analysis chapter, when appropriate. This chapter has attempted to infer the clonal evolution of the CLL4 cohort, but to do so, a clinically relevant cut off needed to be selected. The Nadeu *et al.* study offered such a cut off, showing that by basing calling clonal or subclonal mutations on the detection limit of Sanger sequencing (12% VAF), insight into the survival associations and clonal expansion of these subclonal mutations was possible. Therefore, this cut off was selected for this study, allowing the identification of 12 significantly clonal genes, with no statistically significant genes found to be mostly subclonal. Regardless, the true challenge for this dataset is in its clinical implementation, with the opportunity to study the clinical impact of these genes concurrently.

### 3.10.7 Future work

Additional work that would add value to the data presented here should be focused on describing the copy number landscape of all CLL4 cases, so that assessment of true mutation recurrence (via analysis with MutSig) and clonality of mutations and genes (via ABSOLUTE) could be undertaken. This could be achieved using shallow WGS, an NGS strategy that covers the whole genome at a limited depth, allowing inference of copy number status. Furthermore, since additional studies after the design of the CLL4 TruSeq custom panel found mutations in the 3'UTR of *NOTCH1,* screening of these variants across all cases would improve the analysis of *NOTCH1.* Data from this TruSeq study was used as part of a HRM screening project of CLL4 cases for the *NOTCH1* 3'UTR variants, therefore some data is already available. However, considering the number of *NOTCH1* subclonal mutations, NGS screening of this region is still required. Splice variants in Dishevelled 2 (*DVL2*), a gene involved in Wnt signalling which is known to limit NOTCH signalling through inhibition of CSL (185), have recently been shown to occur in *SF3B1* mutated CLL, and led to activation of NOTCH signalling, therefore screening this large *SF3B1* mutated cohort for this splice variant would be interesting, especially since the study which observed the splice variant did not provide any associated survival data between those cases which had the *DVL2* splice variant or not. However, the number of CLL4 cases with available RNA is low, which may impede this analysis from taking place, since it is not guaranteed that any of the CLL4 cases with available RNA are *SF3B1* mutants.

On the technical side, finding ways to improve the bioinformatics strategy to identify recurrent false positives would be advantageous, as the volume of variants required to be checked in IGV was extensive. The method applied here to reduce these false positives would be a good addition to the full process if it was automated, however that would still leave many variants to manually curate. Automating the assessment of the sequence context of variants would be one way to overcome this, but this would require in depth coding knowledge to be able to use available genomics tools to use the .bam and .bam.bai files to: search for the variant, count the number of reads assigned to each base at that location, scan the sequence around the variant to see whether it is near the end of the read, is surrounded by messy reads, or is directly next a large deletion event not identified in the original filtered analysis pipeline. Previous work within the cancer genomics group has assessed the quality of each read rather than each variant, creating a read Phred score, which removes poor quality reads, therefore reducing the calling of potentially false positive variants (186).

### 3.10.8 Conclusions

This chapter has described the: sample preparation, sequencing, mutation data analysis, and mutation landscape of CLL4 using an Illumina TruSeq custom panel. It has identified 18 recurrently mutated genes, with *SF3B1* found to be the most recurrent, including within both U-CLL and M-CLL disease subsets. The outcome of this analysis is a dataframe of 903 variants to take forward for clinical statistical analysis, to assess the survival associations of the observed mutated genes.

# Chapter 4: Clinical and statistical survival analysis of the LRF UK CLL4 TruSeq study

Identification of predictive and prognostic biomarkers is important for optimal cancer patient management, especially in a disease like CLL which has a heterogeneous clinical course. CLL has many clinical biomarkers, however most of these are prognostic, with *TP53* status being the standout predictive biomarker. A plethora of candidate biomarkers have been presented, however they have seldom been analysed together, with many multivariate Cox proportional hazard and disease stratification models excluding variables essential for CLL prognostication. Candidate biomarker exclusion reasons range from the experimental procedure being difficult or costly to perform, to them either not being popular within the CLL community, or falling out of the multivariate model due to lack of significance in that dataset.

Assessing variables within a patient population can be of great utility in improving patient care and disease management. By performing statistical analysis of whether a specific covariate (e.g. type of therapy, genetic or biological variable, clinical feature) significantly impacts the survival time of patients can greatly aid understanding of the disease and improve subsequent patient treatment. This can be achieved by performing survival analysis which uses two key features: the survival time at the point of data collection (usually date of randomisation of each patient to the trial until the prescribed date of data collection), and the survival event (usually a binary scoring system which stratifies patients into two groups based on whether the outcome measure has occurred (death, disease progression) or not (censored) (187). Classical survival analysis involves both univariate and multivariate stages, implementing Kaplan Meier curves and Cox proportional hazard ratio (HR) regression models. More recently however, implementation of machine learning tools to survival analysis has been included alongside classical analysis, providing a multivariate approach which does not have any underlying assumptions of a variable's hazard (188).

Of the targeted NGS studies using clinical trial cohorts **(described in Chapter 3 intro)**, only limited survival data was presented, without any stratified survival analysis of the impact of clonal and subclonal mutations. Herling *et al.* (2016) (162) showed an association with poor OS in *POT1* mutated cases, along with a clinical association between *KRAS* mutants and a complex karyotype. Guièze *et al.* (2015) (163) showed that when *TP53, SF3B1,* and pathogenic *ATM* are co-mutated (termed "multiple hit") it leads to reduced PFS. The only other targeted NGS study of a CLL clinical trial showed the impact of *NOTCH* P2514fs variants on PFS, however this was a comparison of different mutation capture techniques (165).

Survival analysis investigates time to event data, where a response variable (time) is measured in relation to a specific end point (event) (187). The two main techniques employed for survival analysis are Kaplan Meier curves, and Cox Proportional Hazard (PH) modelling. Kaplan Meier survival analysis estimates the probability of surviving until a specific point in time (survival function, "t") vs. the probability of dying at "t" (187). Cox PH modelling is a type of linear regression, where the dependent variable is the probability of dying (Hazard Ratio, HR). Both methods can analyse the differences between groups using the Log Rank test (187).

Multivariate survival analysis using Cox PH is the gold standard in the identification of clinically important variables which impact prognosis. However, the statistical modelling process accepts certain assumptions of the variables being tested (such as hazard proportionality). To overcome this restriction, machine learning tools have been developed to assess the impact of variables within a dataset in an unsupervised, unbiased fashion. Random Forests (RF) are a non-parametric statistical method which do not infer any assumption over the distribution of data points concerning the response variable (OS or PFS) (189). Random Survival Forests (RSF) are a right-censored model of RF, which function through the production of forests of decision trees based on the Log Rank score from the response variable and the experimental variables and cases randomly selected for each decision tree (189) .

This chapter, applying the data generated from **Chapter 3:**, as well as all available clinical and previous biomarker study data (from previous publications from the Cancer Genomics Group and others), investigated the predictive and prognostic biomarker status of the mutated genes in the CLL4 TruSeq study. Validation of the TruSeq technology to identify additional poor risk patients using genes with previous CLL4 study data (*TP53, SF3B1, NOTCH1*) with sensitivity specificity analysis was conducted. Time-dependent coefficient survival analysis was undertaken to better understand the importance of long-term survival data, since CLL is a disease of the elderly and age can be a competing risk. Classical univariate survival analysis of all covariates was conducted for OS and PFS, to identify candidate biomarkers for multivariate analysis. This was followed by construction of traditional multivariate Cox PH models. Finally, the application of unsupervised machine learning strategies was explored for survival analysis.

In summary, this chapter aims to discuss the following:

1. The composition of the CLL4 TruSeq cohort, and how it compares to the previous CLL4 gene mutation publications in terms of identification of survival events
2. The available time to event survival data, exploring the differences between time points and the functions of survival analysis, to better interpret key analytical read outs
3. The outcome of univariate survival analysis on the full CLL4 TruSeq cohort, including in depth analysis of genes found to significantly associate with reduced OS or PFS, exploring pairwise Kaplan Meier survival analysis
4. Multivariate Cox Proportional Hazards model of all mutated genes and clinical biomarkers significantly associated with OS and PFS in univariate survival analysis
5. Application of Random Survival Forests to identify the most important variables in the CLL4 TruSeq dataset, in an unbiased fashion

## 4.1 Comparison of the CLL4 TruSeq cohort to full the CLL4 clinical trial

When analysing sub-cohorts of clinical trials, it is important to assess whether the sub-cohort is comparable to and representative of the full trial cohort. This is because over-representation of specific clinical groups could bias survival analysis. Therefore, 2x2 and 3x2 Chi-squared contingency table calculations were conducted for all major clinical characteristics reported in the original trial **(Table 6)**. All the clinical characteristics were observed to be non-significantly different between the CLL4 TruSeq cohort and the full CLL4 trial cohort (*P*>0.05, see **Table 6** for details). This shows that the CLL4 TruSeq cohort is comparable to the full trial. In addition, the percentage of missing data was calculated, identifying the following: *IGHV* mutational status (15.83%), CD38 (15.43%), ZAP70 (23.85%), del(17p) (8.62%), del(11q) (7.01%), +(12) (7.01%), and del(13q) (7.41%). All other variables had a full complement of data. Handling of missing data can be conducted in multiple ways, with the most common ways being: data imputation, and complete case analysis. Since data imputation randomly assigns each missing data value to "become" mutated or wild type, complete case analysis will be used in this thesis so that all observations are based on real data. However, to maintain confidence in the survival associations observed, power calculations for all comparisons will be conducted.

**Table 6 CLL4 TruSeq vs. full CLL4 trial: cohort characteristics**

| | | CLL4 TruSeq cohort (n=499) | | | Full CLL4 cohort (n=777) | | | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | **Subgroup** | **No. Cases** | **%** | **Total no. of cases with data** | **No. Cases** | **%** | **Total no. of cases with data** | ***P* value** |
| **Gender** | Male | 361 | 72.34 | 499 | 573 | 73.75 | 777 | 0.582 |
| | Female | 138 | 27.66 | | 204 | 26.25 | | |
| **Age at diagnosis** | Mean | 64 | N/A | 499 | 64 | N/A | 777 | 0.428 |
| | Median | 65 | N/A | | 64 | N/A | | |
| | Range | 42-86 | N/A | | 35-86 | N/A | | |
| **Binet stage** | A | 121 | 24.25 | 499 | 191 | 24.58 | 777 | 0.909 |
| | B | 222 | 44.49 | | 352 | 45.30 | | |
| | C | 156 | 31.26 | | 234 | 30.12 | | |
| **Treatment** | Chl | 237 | 47.49 | 499 | 387 | 49.81 | 777 | 0.358 |
| | FC | 144 | 28.86 | | 196 | 25.23 | | |
| | FDR | 118 | 23.65 | | 194 | 24.97 | | |
| ***IGHV* Mutational Status** | Unmutated | 263 | 62.62 | 420 | 327 | 61.35 | 533 | 0.689 |
| | Mutated | 157 | 37.38 | | 206 | 38.65 | | |
| **CD38 positivity** | Positive | 268 | 63.51 | 422 | 334 | 62.55 | 534 | 0.760 |
| | Negative | 154 | 36.49 | | 200 | 37.45 | | |
| **ZAP70 positivity** | Positive | 189 | 49.74 | 380 | 236 | 49.37 | 478 | 0.916 |
| | Negative | 191 | 50.26 | | 242 | 50.63 | | |
| **FISH del(17p)** | Normal | 430 | 94.30 | 456 | 538 | 94.22 | 571 | 0.958 |
| | Deleted | 26 | 5.70 | | 33 | 5.78 | | |
| **FISH del(11q)** | Normal | 374 | 80.60 | 464 | 462 | 79.93 | 578 | 0.786 |
| | Deleted | 90 | 19.40 | | 116 | 20.07 | | |
| **FISH +(12)** | Normal | 394 | 84.91 | 464 | 487 | 84.26 | 578 | 0.770 |
| | Deleted | 70 | 15.09 | | 91 | 15.74 | | |
| **FISH del(13q) alone** | Normal | 295 | 63.85 | 462 | 361 | 62.46 | 578 | 0.643 |
| | Deleted | 167 | 36.15 | | 217 | 37.54 | | |

Variables compared using 2x2 or 3x2 Chi-squared contingency tables, as appropriate. Comparison of age at diagnosis between groups was calculated using an unpaired two-sided t-test.

## 4.2 The CLL4 TruSeq custom panel has greater sensitivity but reduced specificity in comparison to previous CLL4 mutation publications

Targeted re-sequencing platforms offer the capacity to identify mutations below the detection thresholds of standard techniques used in molecular diagnostic laboratories. However, increased sensitivity to mutation detection may not necessarily correlate with increased identification of patients at risk of an event (death/progression). Therefore, comparative analysis of the CLL4 TruSeq mutation data with previous CLL4 publications (*TP53*: n = 424, *SF3B1:* n = 344, *NOTCH1:* n = 365) was carried out for OS and PFS using sensitivity and specificity analysis. Sensitivity and specificity analysis represents the measure of performance of binary classification tests (i.e. lists of cases mutated for a gene vs. survival censor data). This type of analysis has been used by this laboratory previously, showing the added benefit of screening for *NOTCH1* 3'UTR non-coding mutations by HRM and Sanger sequencing in comparison to *NOTCH1* coding mutations alone(190). For ease of comparison between the CLL4 TruSeq study and previous publications, all OS analyses will be conducted using the OS 2012 and PFS 2011 datasets, the time to event survival data employed in the *SF3B1* and *NOTCH1* paper (7).

### 4.2.1 Specificity higher in previous CLL4 study data for capturing OS events in all genes

Across all genes, TruSeq had an increased sensitivity (True Positive Rate [TPR]) than the original publications (*TP53* 12.82% vs. 9.62%, *SF3B1* 30.16% vs. 21.83%, *NOTCH1* 17.39 vs. 11.96%). TPR is defined in this context as the proportion of mutated cases for a gene that received a survival event. However, TruSeq had reduced specificity (True Negative Rate[TNR]) than the original publications (*TP53* 97.32% vs. 99.11%, *SF3B1* 88.04% vs. 95.65%, *NOTCH1* 91.01 vs. 95.51%). TNR is defined in this context as the proportion of wild type cases for a gene that received a survival event. The TPR and TNR were then compared to the False Positive Rate (FPR) and False Negative Rate (FNR) (measures of the proportions of mutated cases to not receive a survival event [FPR], and the proportions of wild type cases to receive an event [FNR]), to generate accuracy values and Likelihood Ratios (LR) in the positive and negative directions. LR+ represents the probability of mutated cases receiving a survival event divided by the probability of wild type cases receiving a survival event, whilst the LR- represents the same, but for cases who do not receive a survival event. Finally, a combined LR +/- value, also known as the Diagnostic Odds Ratio (DOR), was calculated, allowing direct comparison of TruSeq vs. the original publication in terms of likelihood of gene mutations predicting a survival event. The DOR represents the likelihood of mutated cases that receive a survival event in relation to likelihood of cases that do not. In all genes, the original publication had higher LR+/- values (*TP53* 4.29 vs. 9.82, *SF3B1* 2.00 vs. 4.10, *NOTCH1* 1.76 vs.

2.45), suggesting that mutations outside the detection limits of the technologies used in these original publications do not contribute to an increased capturing of OS events, at least in these genes **(Figure 47)**.

**A**

| *TP53* Mutation - OS 2012 | Sensitivity% (TPR) | Specificity% (TNR) | False Negative Rate% (FNR) | False Positive Rate% (FPR) | Accuracy% | LR+ | LR- | LR+/LR- |
|---|---|---|---|---|---|---|---|---|
| *TP53* TruSeq | 12.82 (43/424) | 97.32 (3/112) | 87.18 (40/312) | 2.68 (3/43) | 8.73 (37/424) | 4.79 | 1.12 | 4.29 |
| *TP53* Gonzalez | 9.62 (31/424) | 99.11 (1/112) | 90.38 (30/312) | 0.89 (1/31) | 6.84 (29/424) | 10.77 | 1.10 | 9.82 |

**B**

| *SF3B1* Mutation - OS 2012 | Sensitivity% (TPR) | Specificity% (TNR) | False Negative Rate% (FNR) | False Positive Rate% (FPR) | Accuracy% | LR+ | LR- | LR+/LR- |
|---|---|---|---|---|---|---|---|---|
| *SF3B1* TruSeq | 30.16 (87/344) | 88.04 (11/92) | 69.84 (76/252) | 11.96 (76/87) | 18.90 (65/344) | 2.52 | 1.26 | 2.00 |
| *SF3B1* Oscier | 21.83 (59/344) | 95.65 (4/92) | 78.17 (55/252) | 4.35 (55/59) | 14.83 (51/344) | 5.02 | 1.22 | 4.10 |

**C**

| *NOTCH1* Mutation - OS 2012 | Sensitivity% (TPR) | Specificity% (TNR) | False Negative Rate% (FNR) | False Positive Rate% (FPR) | Accuracy% | LR+ | LR- | LR+/LR- |
|---|---|---|---|---|---|---|---|---|
| *NOTCH1* TruSeq | 17.39 (56/365) | 91.01 (8/89) | 82.61 (48/276) | 8.99 (8/56) | 10.96 (40/365) | 1.93 | 1.10 | 1.76 |
| *NOTCH1* Oscier | 11.96 (37/365) | 95.51 (4/89) | 88.04 (33/276) | 4.49 (4/37) | 7.95 (29/365) | 2.66 | 1.08 | 2.45 |

**Figure 47 Sensitivity Specificity analysis of *TP53, SF3B1,* and *NOTCH1* for OS**

Sensitivity specificity tables for OS for *TP53* **(A)**, *SF3B1* **(B)**, and *NOTCH1* **(C)**.

### 4.2.2 Comparable specificity between technologies for capturing *NOTCH1* PFS events

Sensitivity specificity tables were calculated for *TP53, SF3B1,* and *NOTCH1,* using the 2011 PFS time point data, since both *SF3B1* and *NOTCH1* were published using this dataset. Although the *TP53* Gonzalez paper was published in 2010, the 2011 dataset was used to be consistent across genes. As seen above for OS, TruSeq had an increased sensitivity for PFS than the original publications (*TP53* 10.80% vs. 7.97%, *SF3B1* 26.43% vs. 18.79%, *NOTCH1* 16.27 vs. 10.65%) **(Figure 48)**. Furthermore, TruSeq had reduced specificity for *TP53* and *SF3B1* than the original publications (*TP53* 97.14% vs. 100%, *SF3B1* 86.67% vs. 100%), whilst sharing comparable specificity for *NOTCH1* (*NOTCH1* 96.30 vs. 96.30%). *TP53* and *SF3B1* both had 100% specificity in their respective original publications, i.e. all mutated cases progressed. This means that LR+ values, and the LR+/- ratio are non-calculable, since numbers are not divisible by zero. Therefore, this shows that mutated cases from the original publications for *TP53* and *SF3B1* are better predictors of progression than TruSeq. However, for *NOTCH1* this is not the case, with TruSeq having an increased LR+/- value to its original publication (*NOTCH1* 3.82 vs. 2.67). This suggests that increased sensitivity of detection using NGS technologies may impact identification of cases who may progress, but that it is gene specific. Furthermore, this comparative analysis of OS and PFS suggests that a similar analysis for each gene in the TruSeq cohort between clonal and subclonal mutations (representing standard Sanger sequencing mutation detection sensitivity) would be beneficial, to assess the applicability of NGS as a tool to identify predictive biomarkers.

**A**

| ***TP53*** **Mutation - PFS 2011** | **Sensitivity% (TPR)** | **Specificity% (TNR)** | **False Negative Rate% (FNR)** | **False Positive Rate% (FPR)** | **Accuracy%** | **LR+** | **LR-** | **LR+/LR-** |
|---|---|---|---|---|---|---|---|---|
| *TP53* TruSeq | 10.80 (43/424) | 97.14 (1/35) | 89.20 (42/389) | 2.86 (1/43) | 9.67 (41/424) | 3.78 | 1.09 | 3.47 |
| *TP53* Gonzalez | 7.97 (31/424) | 100.00 (0/35) | 92.03 (31/389) | 0.00 (0/31) | 7.31 (31/424) | NA | 1.09 | NA |

**B**

| ***SF3B1*** **Mutation - PFS 2011** | **Sensitivity% (TPR)** | **Specificity% (TNR)** | **False Negative Rate% (FNR)** | **False Positive Rate% (FPR)** | **Accuracy%** | **LR+** | **LR-** | **LR+/LR-** |
|---|---|---|---|---|---|---|---|---|
| *SF3B1* TruSeq | 26.43 (87/344) | 86.67 (4/30) | 73.57 (83/314) | 13.33 (83/87) | 22.97 (79/344) | 1.98 | 1.18 | 1.68 |
| *SF3B1* Oscier | 18.79 (59/344) | 100.00 (0/30) | 81.21 (59/314) | 0.00 (59/59) | 17.15 (59/344) | NA | 1.23 | NA |

**C**

| ***NOTCH1*** **Mutation - PFS 2011** | **Sensitivity% (TPR)** | **Specificity% (TNR)** | **False Negative Rate% (FNR)** | **False Positive Rate% (FPR)** | **Accuracy%** | **LR+** | **LR-** | **LR+/LR-** |
|---|---|---|---|---|---|---|---|---|
| *NOTCH1* TruSeq | 16.27 (56/365) | 96.30 (1/27) | 83.73 (55/338) | 3.70 (1/56) | 14.79 (54/365) | 4.39 | 1.15 | 3.82 |
| *NOTCH1* Oscier | 10.65 (37/365) | 96.30 (1/27) | 89.35 (36/338) | 3.70 (1/37) | 9.59 (35/365) | 2.88 | 1.08 | 2.67 |

**Figure 48 Sensitivity Specificity analysis of TP53, SF3B1, and NOTCH1 for PFS**

Sensitivity specificity tables for PFS for *TP53* **(A)**, *SF3B1* **(B)**, and *NOTCH1* **(C)**.

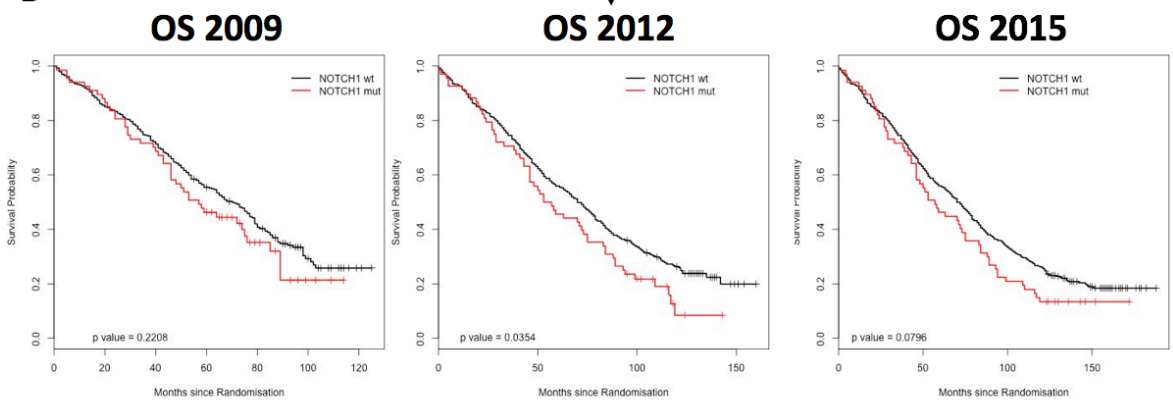
## 4.3 Analysis of the CLL4 time to event survival data

During clinical trials, repeat collection of survival data over time is taken to assess the longevity of patients under treatment. This provides multiple time points in which to test the effects of experimental variables on survival. However, as the number of patients at risk change at each data collection, and the age of patients increases over time, the selection of the optimal survival data time point, minimising the effect of confounding variables, is critical. To assess the impact of these factors on survival, a multifactorial approach was undertaken, an overview of which can be seen in **Figure 49**. For initial comparison, the median survival times were calculated for all time points **(Figure 49A, 4.3.1)**, followed by univariate Kaplan Meier plots of all variables at all time points **(Figure 49B, 4.3.2)**. Next, univariate Cox PH models were created, followed by assessment of their hazard proportionality via time-dependent covariate analysis **(Figure 49C, 4.3.3)**. Finally, the survival data was split at the median survival months, followed by Cox PH model generation of each split and comparison of all variables HRs **(Figure 49D, 4.3.4)**.

A

**Comparison of median OS & PFS survival data time points**

B

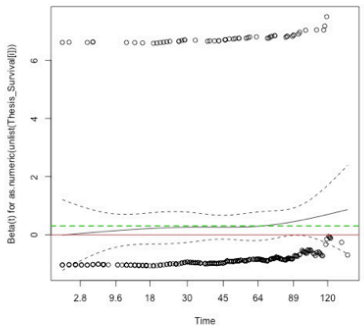



C

**Univariate Cox PH**

| Variable | Hazard Ratio | *P* value |
|---|---|---|
| *NOTCH1* | 1.354 | 0.036 |

→ **Log transform** →




D

**SurvSplit Univariate Cox PH**

| Variable | <70 months HR | <70 months P value | >70 months HR | >70 months *P* value | Overall model *P value* |
|---|---|---|---|---|---|
| *NOTCH1* | 1.1465 | 0.6274 | 2.15 | 0.0316 | 0.0786 |

**Figure 49 Overview of survival data time point analysis strategy**

**A** Comparison of all time points for OS and PFS survival data (only OS shown as a representative example), to identify inherent differences. **B** Representative Kaplan Meier plots from each OS time point for *NOTCH1*. Comparison of all variables at each time point undertaken to assess impact of time points on survival. **C** Representative workflow of assessment of time-dependent covariates, using *NOTCH1*. **D** Representative SurvSplit Cox PH model of *NOTCH1*, where the survival data was split at the median cohort survival point.

### 4.3.1 Median time to event in CLL4 remains consistent over time

To assess the impact of survival data time points on experimental variables in CLL4, survival data for OS was attained from the 2009, 2012, and 2015 updates, and for PFS from the 2009 and 2011 updates. At the time of writing, the original CLL4 trial result survival data from 2007 was unavailable, therefore analysis was only undertaken on the 2009, 2012, and 2015 survival updates for OS, with the survival updates from 2009 and 2011 analysed for PFS. Assessment of 2011 OS data was precluded due to it not being used in previous publications. PFS data was discarded as a continued measure after the 2011 update by the clinical trials team, since most of the patients in the trial had already reached their date of best response.

The median OS and PFS for each survival data time point was calculated, along with: 95% confidence intervals, minimum and maximum months since randomisation of patients receiving an event, and the number of events **(Figure 50)**. For OS, the median OS remained similar across all three time points (2009 = 68 months, 2012 = 70 months, 2015 = 70 months), whilst the number of cases recorded as dead at each time point was significantly different by Fisher's exact testing (2009 vs. 2012 vs. 2015 $P<0.0001$, 2009 vs. 2012 $P<0.0001$, 2012 vs. 2015 $P = 0.0135$) **(Figure 50A)**. For PFS, the median PFS was identical at 35 months, with a non-significant difference in the total number of cases defined as progressed (2009 vs. 2011 $P = 0.7184$) **(Figure 50B)**. These differences in the number of events at each time point could be important, since they could associate with experimental variables.

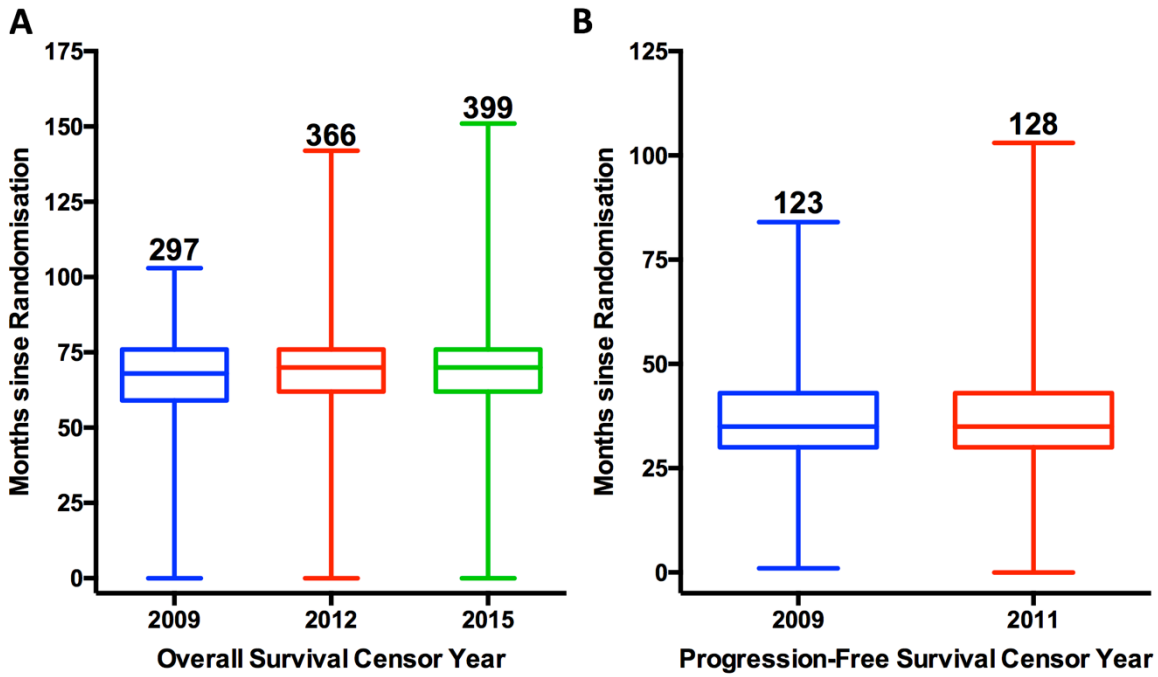



**Figure 50 Comparison of CLL4 survival data time points for OS and PFS**

Box and whisker plots of median survival of TruSeq CLL4 cohort (n=499) for OS **(A)** and PFS **(B)**. Box represents upper and lower 95% confidence intervals, with the horizontal line within the box representing the median survival across the entire cohort. Bars represent the minimum and maximum months since randomisation of patients succumbing to an event. Number above each box and whisker plot represents the number of events recorded at each time point.

### 4.3.2 OS time to event data from 2012 and PFS time to event data from 2011 identify the most variables associated with survival

To get a better understanding of the effect of survival data time points on experimental variables, univariate Kaplan Meier survival analysis was conducted for all variables across all survival data time points for OS and PFS, respectively **(Figure 51 & Figure 52)**. This identified 47 variables that had a significant impact ($P$<0.05) for at least one time point for OS **(Figure 51)** and 28 variables for PFS **(Figure 52)**. For OS, this analysis found variables significant at all time points, as well as variables with a dichotomous impact on survival depending on the time point in question. Interestingly, for OS there were 13 variables which showed no impact on survival in 2009, but by 2012 12/13 of these variables were found to have a significant impact on OS, with 6/13 returning to show no impact on survival by 2015 **(Figure 51)**. However, for PFS the impact was more straightforward, with only 3 affected variables, 2 found only to have a significant impact on progression at the later time point (2011), and only 1 variable found have a significant impact on progression in 2009 **(Figure 52)**. This suggests that these dichotomous variables may be time-dependent.

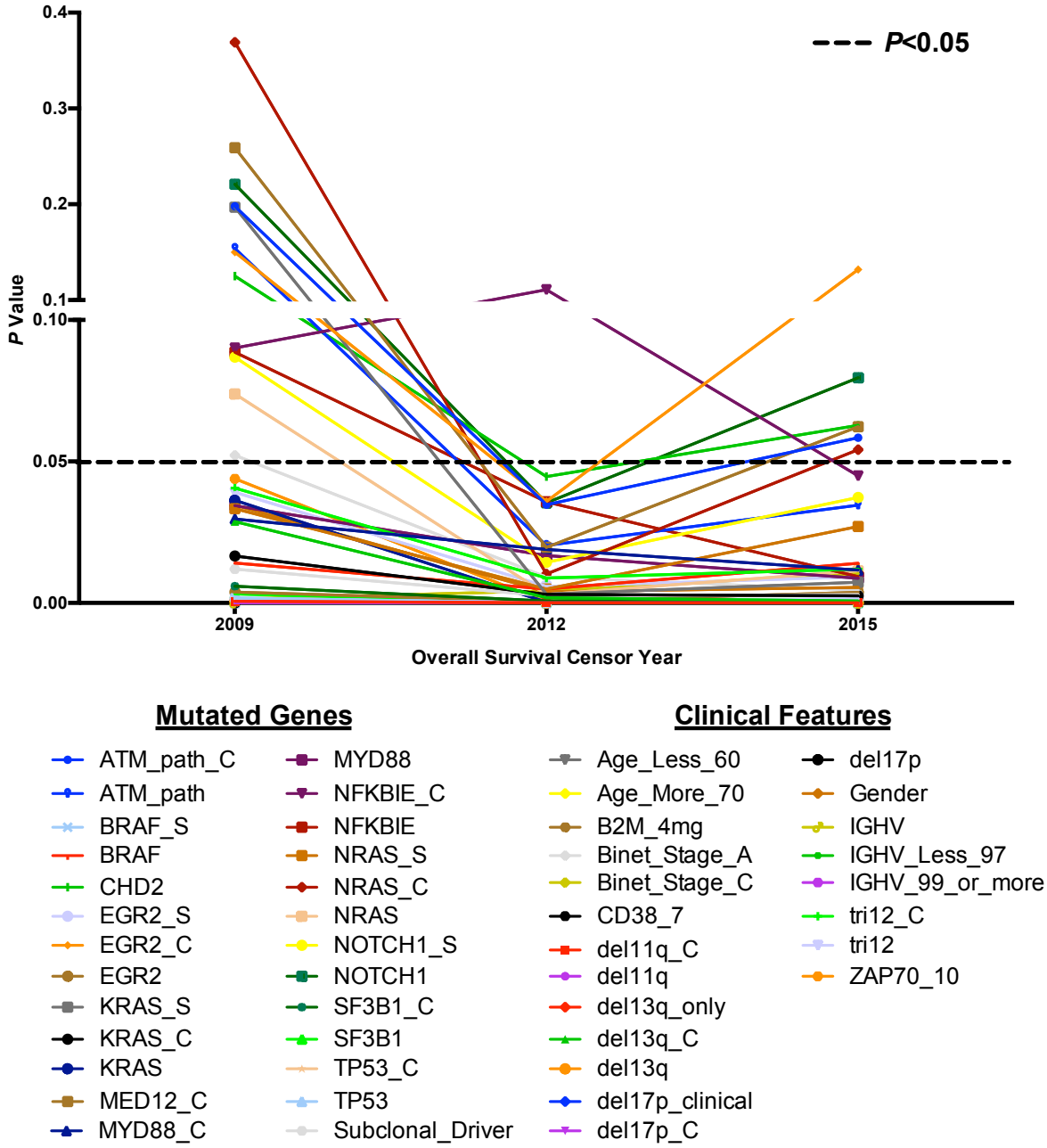



**Figure 51 Overview of univariate OS Kaplan Meier survival analysis over time**

Line plot showing the OS Kaplan Meier survival analysis Log Rank *P* values of significant (*P*<0.05) variables at 2009, 2012 and 2015 survival data time points (upper). List of variables separated by mutated genes (left) and clinical features (right) (lower)

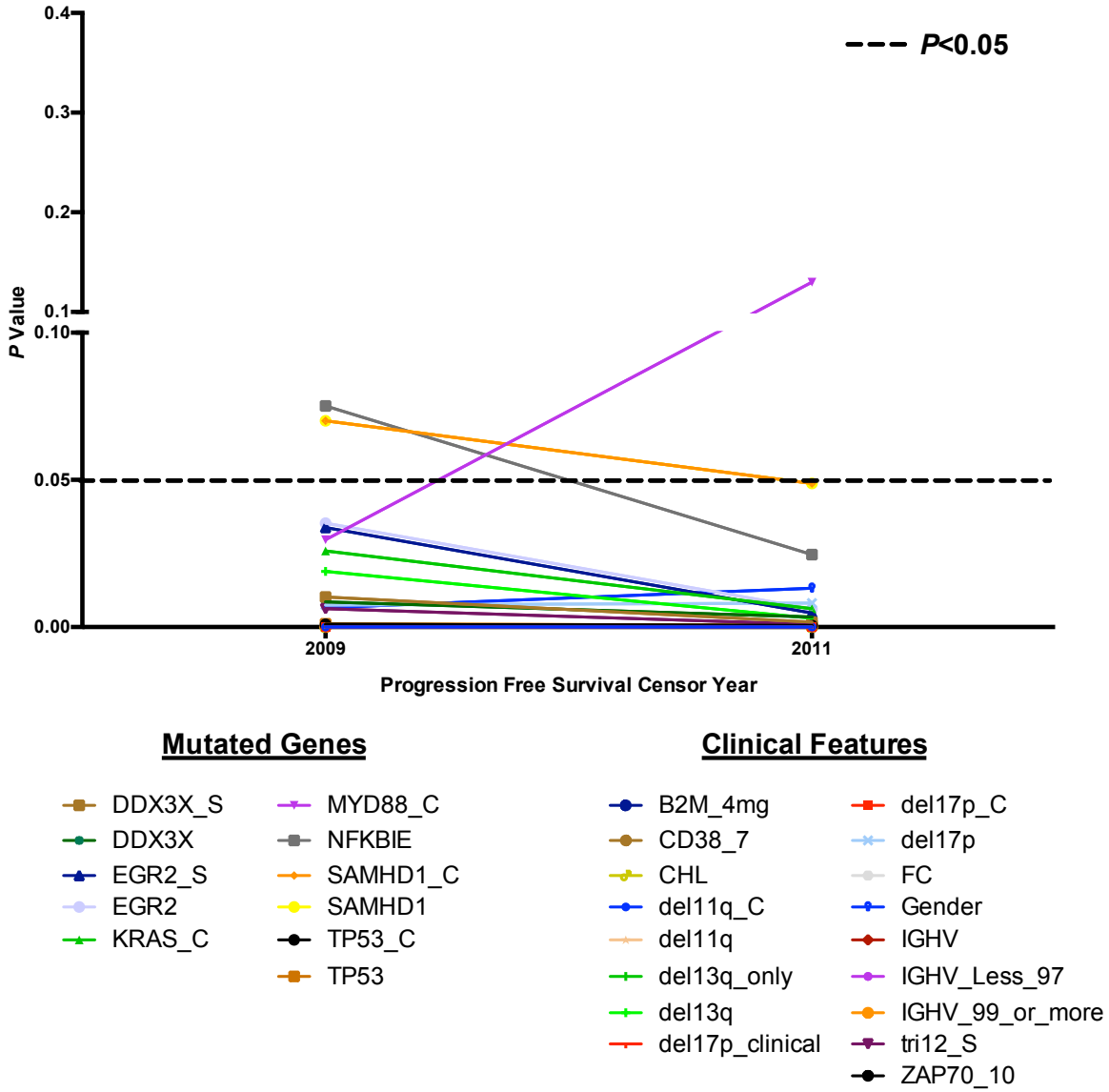



**Figure 52 Overview of univariate PFS Kaplan Meier survival analysis over time**

Line plot showing the PFS Kaplan Meier survival analysis Log Rank *P* values of significant (*P*<0.05) variables at 2009 and 2011 survival data time points (upper). List of variables separated by mutated genes (left) and clinical features (right) (lower)

### 4.3.3 Mutated genes in CLL4 TruSeq data have proportional and non-proportional hazards

When assessing the survival associations of a binary covariate with right censored survival data using Cox PH, the underlying linear regression model assumes that the associated hazard is proportional. What that means is that the relative risk of the variable under survival association testing is set to be linear over time. However, this trust in proportionality is not always warranted, especially when it comes to strong predictors of poor survival, which tend to be non-proportional (191). Non-proportional variables are skewed in relation to time and event, with most events either occurring at the beginning (early) or at the end (late) of the measured time frame. Therefore, to assess whether experimental variables in this dataset were affected by time, testing of non-proportionality using the cox.zph function was performed. In brief, a univariate Cox PH model was created for each of the variables, followed by log transformation of OS and PFS months since randomisation and plotting of the Schoenfeld residuals (191). Diagnosis of non-proportionality was also performed, identifying variables which significantly differed from its HR. For both OS and PFS, the survival data with the longest follow up was used (OS = 2015, PFS = 2011).

When the line of a variable on a Schoenfeld residuals plot, termed Beta(t) (log transformed HR), is a horizontal line not different to 0, it is defined as being proportional (i.e. the original Cox PH model is sufficient to predict the impact of a variable on survival). However, when the line of a variable significantly differs from 0, it suggests that the associated hazard differs over time. The red line on each plot represents 0, the dashed green line represents the HR for the variable, and the solid black line represents the variable as a function of Beta(t). For OS, 11 variables were classified as having significantly non-proportional hazards, and 16 variables for PFS (**Figure 53** and **Figure 54**).

#### *4.3.3.1* *TP53* and *SAMHD1* are non-proportional poor predicting hazards acting early in time in OS, and PFS, respectively

Representative examples of the types of non-proportional hazards observed are presented for OS in **Figure 53** and PFS in **Figure 54**. *TP53* was found to have a strong impact on OS and PFS, with time-dependent HR values much higher than the HR calculated assuming proportionality (dashed green line) early post randomisation (OS = 0-45 months until intersection, PFS 0-12 months). The effect of *TP53* then continues to diminish after this point, confirming the importance of screening for *TP53* mutations at diagnosis, but identifying that *TP53* mutated CLL had an ever-reducing impact on OS and PFS over time (**Figure 53** and **Figure 54**; top left). Gender (Male) and *SAMHD1* follow a similar trend as *TP53* for PFS, having significantly greater risk at earlier time points, supporting previously published data in CLL4 that men have a poorer response to therapy than

women, as well as data from other UK clinical trials that *SAMHD1* mutations are linked to having a poor PFS (manuscript under review at time of writing)**.** However, these hazards are lost after about 10 and 8 months, respectively (**Figure 54**; middle right & top right).

#### 4.3.3.2 *SF3B1* and *ATM* are non-proportional poor predicting hazards acting late in time in OS

Clonal *SF3B1* mutations, age <60 years and *ATM* mutations were found to have an initial good prognosis impact on OS, followed by progressive movement towards and past the PH, with *SF3B1* clonal mutations Beta(t) reducing again after 94 months, whilst age <60 and *ATM* mutations Beta(t) continued to increase 100-months post randomisation (**Figure 53**; *SF3B1* clonal middle right, Age <60 years bottom left, *ATM* top right). Similar trends were observed in PFS for *NFKBIE* mutations and cases treated with FC. This late increase in Beta(t) may also be indicative of the variable being confounded by age.

#### 4.3.3.3 Visual inspection of proportionality required to capture all non-proportional variables, as is the case for *NOTCH1* in PFS

In addition to the above described overt patterns of non-proportionality, more subtle effects were also observable using Schoenfeld residual plots. Namely, variables which remain similar to 0 (red line) for the most part, but with a small deviation in the middle of the line. These fall into two categories: those which deviate significantly, and those which do not. It was originally reported that visual inspection of all Schoenfeld residual plots is required, since the diagnostic test of significance does not function well under conditions of small deviations over a short time period. An example of this phenomenon includes *RPS15* subclonal mutations and *NOTCH1* mutations in OS. Whilst *RPS15* subclonal mutations were found to be significantly non-proportional, *NOTCH1* mutations were not, even though the lines share a similar trajectory **(Figure 54)**. There were 37 variables with non-significant but visually different curves from 0 for OS, and 38 variables for PFS, showing that using the test statistic alone does not capture all potential events of non-proportionality. Considering the impact of some variables at early and late time points in this analysis, splitting of the survival data to account for this should be undertaken.

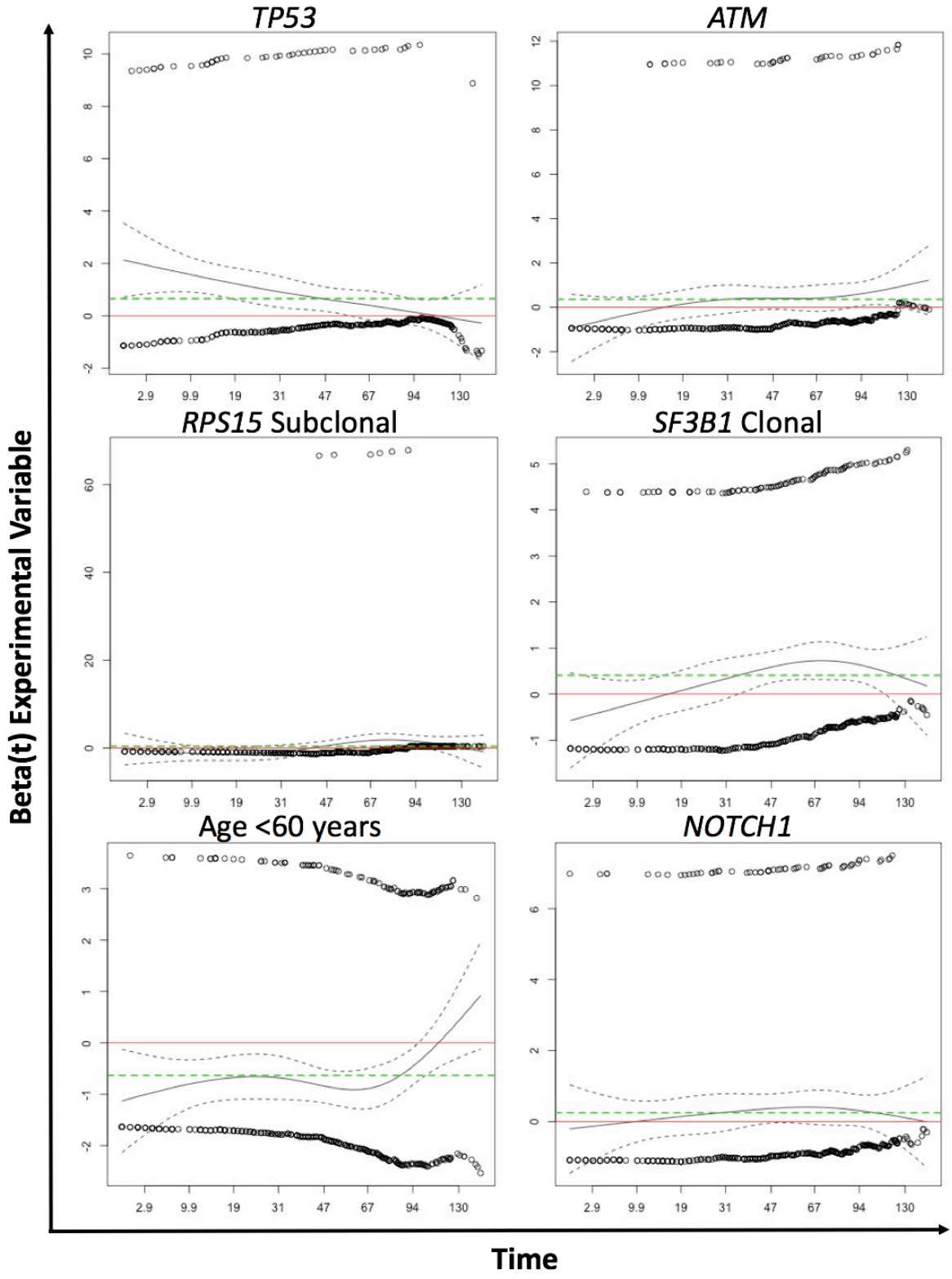



**Figure 53 Scaled Schoenfeld residuals for OS**

Testing proportionality of OS hazards of CLL4 TruSeq study variables. Beta(t) is the log transformed hazard (*y* axis), with time since randomisation on the *x* axis. Selected test variables for OS data presented, all significant test variables shown in **(A.12)**. The red line on each plot represents 0, the dashed green line represents the HR for the variable, and the solid black line represents the variable as a function of Beta(t)

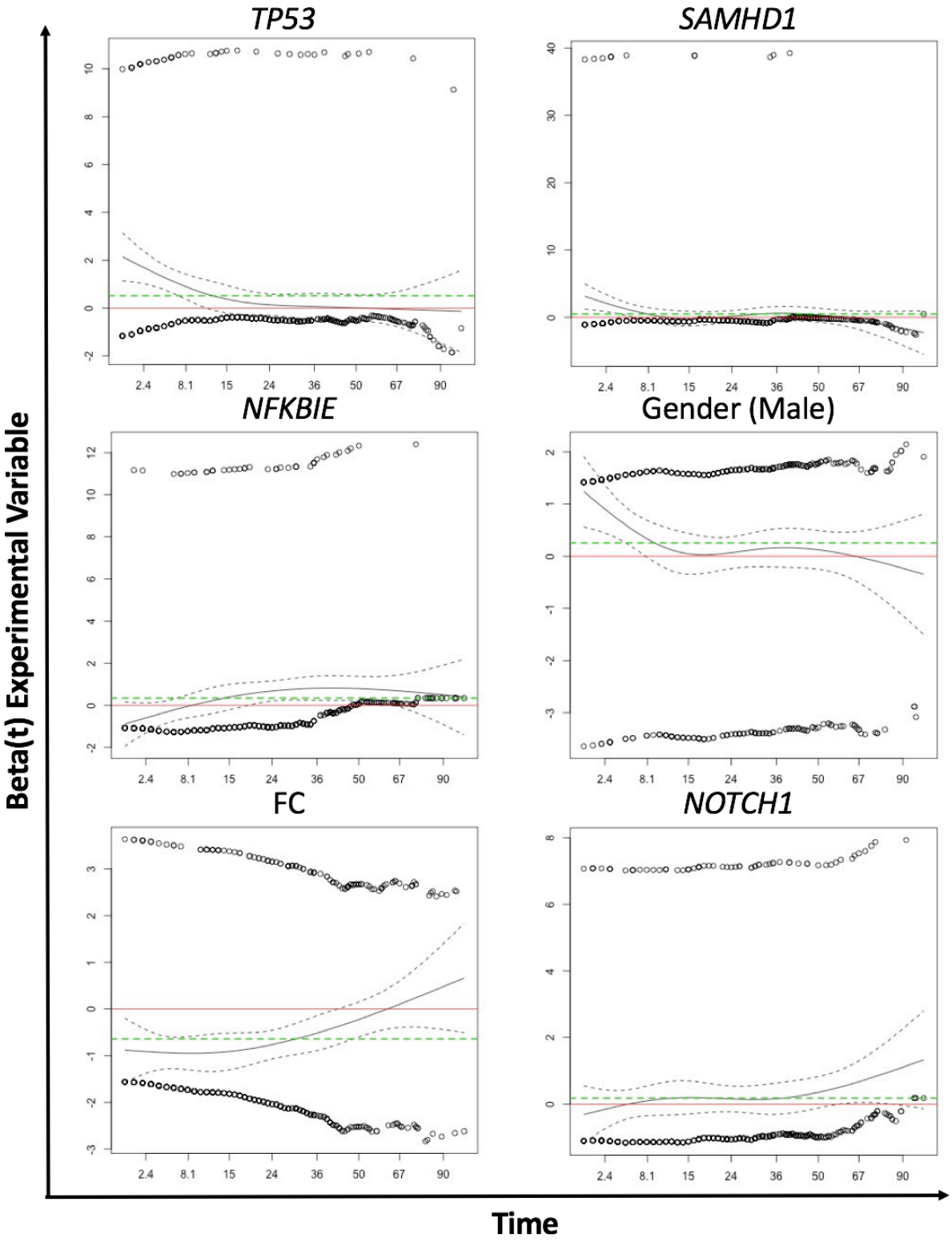



**Figure 54 Scaled Schoenfeld residuals for PFS**

Testing proportionality of PFS hazards of CLL4 TruSeq study variables. Beta(t) is the log transformed hazard (*y* axis), with time since randomisation on the *x* axis. Selected test variables for PFS data presented, all significant test variables shown in **(A.12).** The red line on each plot represents 0, the dashed green line represents the HR for the variable, and the solid black line represents the variable as a function of Beta(t)

### 4.3.4 Splitting of OS and PFS survival analysis by median event time identifies additional survival associations

As identified above, there were variables which were affected by time, either through univariate Kaplan Meier analysis at various time points **(4.3.2)**, or through the generation of Schoenfeld residuals from univariate Cox PH models in time-dependent covariate analysis **(4.3.3)**. This suggests that splitting the survival data to assess the impact of variables over time is required. Therefore, the SurvSplit function within the survival library was applied. The selection of the split point was defined by the median survival of the OS 2015 dataset (70 months) and PFS 2011 dataset (35 months). In brief, SurvSplit creates an additional line per case; if that case survives past the median survival time, leading to the generation of a Cox PH model before the split (all cases) and after the split (only cases still at risk).

#### *4.3.4.1* Mutated *NOTCH1* associated with poor OS after 70 months

The OS SurvSplit univariate Cox PH models identified two groups of variables: those which have an overall significant impact on OS ($P$<0.05), and those which do not ($P$<0.05). The significant overall model variables were further broken down into those which only had an early significant impact (<70 months), those which maintained significance before and after the median split, and those which only had a late significant impact on OS (>70 months) (**Figure 55**; first three groups on graph, left to right**)**. The non-significant overall models were only found to have variables with a significant late impact on OS (**Figure 55**; last group on graph, left to right). Out of the 11 variables identified as having significant non-proportional hazards for OS**,** 7 were found to have an early or late impact on OS in this analysis, reflecting the comparability of the Schoenfeld residuals line and the SurvSplit function. Interestingly, mutated *NOTCH1* was found to significantly impact OS after 70 months but not before, with the overall model found not to significantly impact OS. This is comparable to the Schoenfeld residuals line observed in the bottom right of **Figure 53**. Furthermore, 4 of these late variables were not found to be significant in the 2009 time point survival data, but became so at later time points (*NOTCH1,* subclonal *KRAS, NFKBIE*, Binet Stage A; **Figure 51**). No early events were found to be significant in 2009 and lose their significance at later time points.

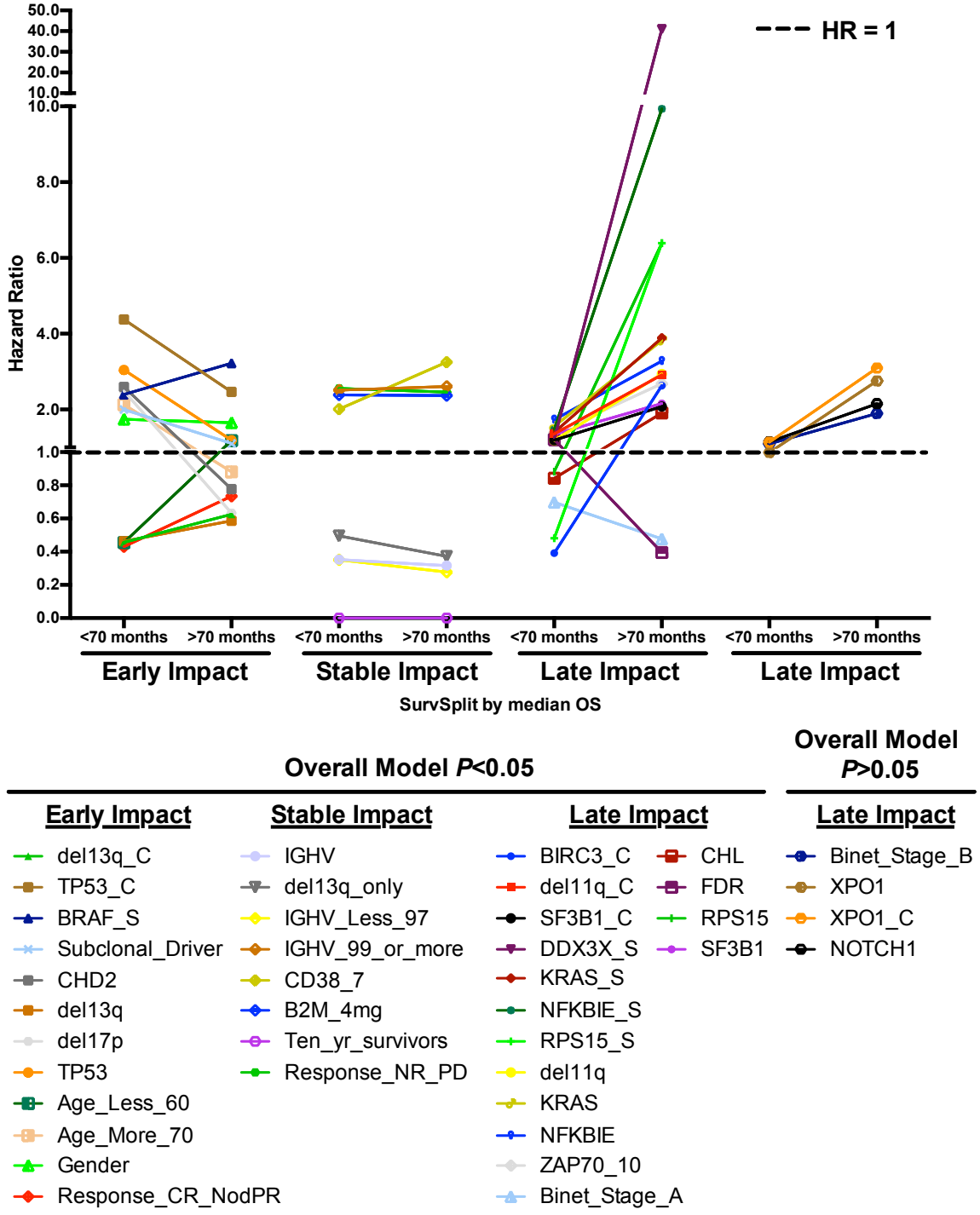



**Figure 55 SurvSplit of OS data**

Line plot showing the OS Hazard Ratio <70 months and >70 months of SurvSplit Cox PH models for all variables which were significant for the overall model, or at at least one time point within the model (*P*<0.05). **Early Impact** Variables that within the model were *P*<0.05 for <70 months, *P*>0.05 for >70 months, and *P*<0.05 for the overall model. **Stable Impact** Variables that were *P*<0.05 at both time points and in the overall model. **Late Impact (Overall Model *P*<0.05)** Variables that within the model were *P*<0.05 for >70 months, *P*>0.05 for <70 months, and *P*<0.05 for the overall model. **Late Impact (Overall Model *P*>0.05)** Variables that within the model were *P*<0.05 for >70 months, *P*>0.05 for <70 months, and *P*>0.05 for the overall model.

#### *4.3.4.2* *SAMHD1* increases its association with poor PFS after 35 months

The PFS SurvSplit univariate Cox PH models identified two groups of variables: those which have an overall significant impact on PFS ($P$<0.05), and those which do not ($P$<0.05). The significant overall model variables were further broken down into those which only had an early significant impact (<35 months), those which maintained significance before and after the median split, and those which only had a late significant impact on PFS (>35 months) **(Figure 56**; first three groups on graph, left to right**)**. The non-significant overall models were only found to have variables with a significant early impact on PFS **(Figure 55**; last group on graph, left to right**)**. *SAMHD1* was found to be the only late variable which was not significant in the 2009 data, with no early time point variables significant in 2009 losing impact on PFS at later time points **(Figure 52)**.

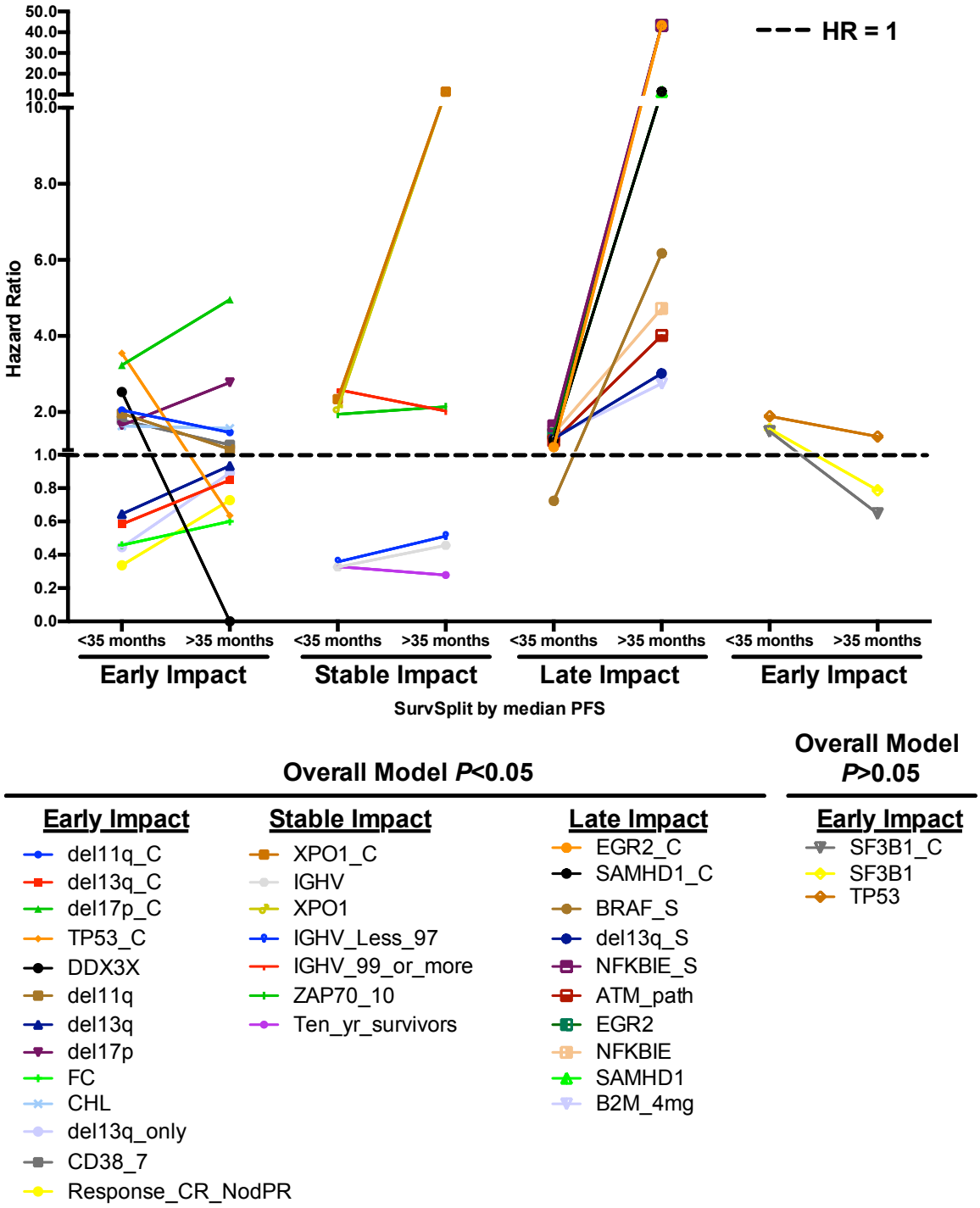



**Figure 56 SurvSplit of PFS data**

Line plot showing the PFS Hazard Ratio <35 months and >35 months of SurvSplit Cox PH models for all variables which were significant for the overall model, or for at least one time point within the model (*P*<0.05). **Early Impact (Overall Model *P*<0.05)** Variables that within the model were *P*<0.05 for <35 months, *P*>0.05 for >35 months, and *P*<0.05 for the overall model. **Stable Impact** Variables that were *P*<0.05 at both time points and in the overall model. **Late Impact (Overall Model *P*<0.05)** Variables that within the model were *P*<0.05 for >35 months, *P*>0.05 for <35 months, and *P*<0.05 for the overall model. **Early Impact (Overall Model *P*>0.05)** Variables that within the model were *P*<0.05 for <35 months, *P*>0.05 for >35 months, and *P*>0.05 for the overall model.

### 4.3.5 Final remarks of overview of the CLL4 time to event survival data

As described above, time does effect the impact of observed variables in time to event survival data. Furthermore, depending on when the observation takes place, significant differences can be found. Therefore, it is important to not only assess the univariate Kaplan Meier plots and univariate Cox PH models, but to also analyse and interpret the proportionality of the hazard in respect of time, and to split the Cox PH models at a sensible time point to assess the timing of the impact of observed variables. However, since the latter process leads to the reduction of cases to observe after the split time point, this analysis is constrained by the power of the cohort to support the observations. To maintain confidence in the survival associations made, sufficient power is required in each association test conducted. However, splitting each variable by the median survival value significantly reduces the number of cases at risk, and therefore the power to make survival associations. Therefore, the survival data time points for OS and PFS which showed the largest number of potentially predictive/prognostic variables were selected for the univariate survival analysis (OS = 2012, PFS = 2011), and the decision to not split the survival data was taken, to allow high confidence survival associations to be made.

## 4.4 Univariate Survival Analysis of CLL4 recurrently mutated genes and clinical biomarkers

To identify potential prognostic biomarkers, univariate Kaplan Meier and Cox PH survival analysis for OS and PFS is required. This was achieved using a standard process of analysis (overview in **Figure 57**). Firstly, univariate Kaplan Meier graphs and Cox PH models of individual variables were created for OS and PFS, identifying an initial list of variables for further interrogation. Next, if the significant variables were of clonal (>12% VAF) or subclonal (<12% VAF) origin, they were tested in pairwise Kaplan Meier curves (hierarchical in structure, with cases called as clonal if containing both a clonal and subclonal mutation in the same gene) and bivariate Cox PH models (comparing both groups regardless of clonality), with only those which retained significance in both tests being taken forward for further analysis. Additionally, after a defined list of significant variables was generated, each association's statistical power was calculated, using the relevant number of cases, survival functions, and Hazard Ratio, with a cut off for inclusion of 80%. This final list is presented in for OS in **Table 7**, and for PFS in **Table 8**, where the univariate Kaplan Meier and Cox PH statistics for OS and PFS associated variables are defined. For all genes found to be significantly associated with poor OS or PFS, the survival data in the context of relevant biological questions and CLL subsets was explored. Significant clinical biomarkers for OS and PFS were not considered for discussion within this section (unless stated), since they have all previously been published to significantly impact survival in CLL4. However, these variables will be re-introduced in the multivariate modelling section.

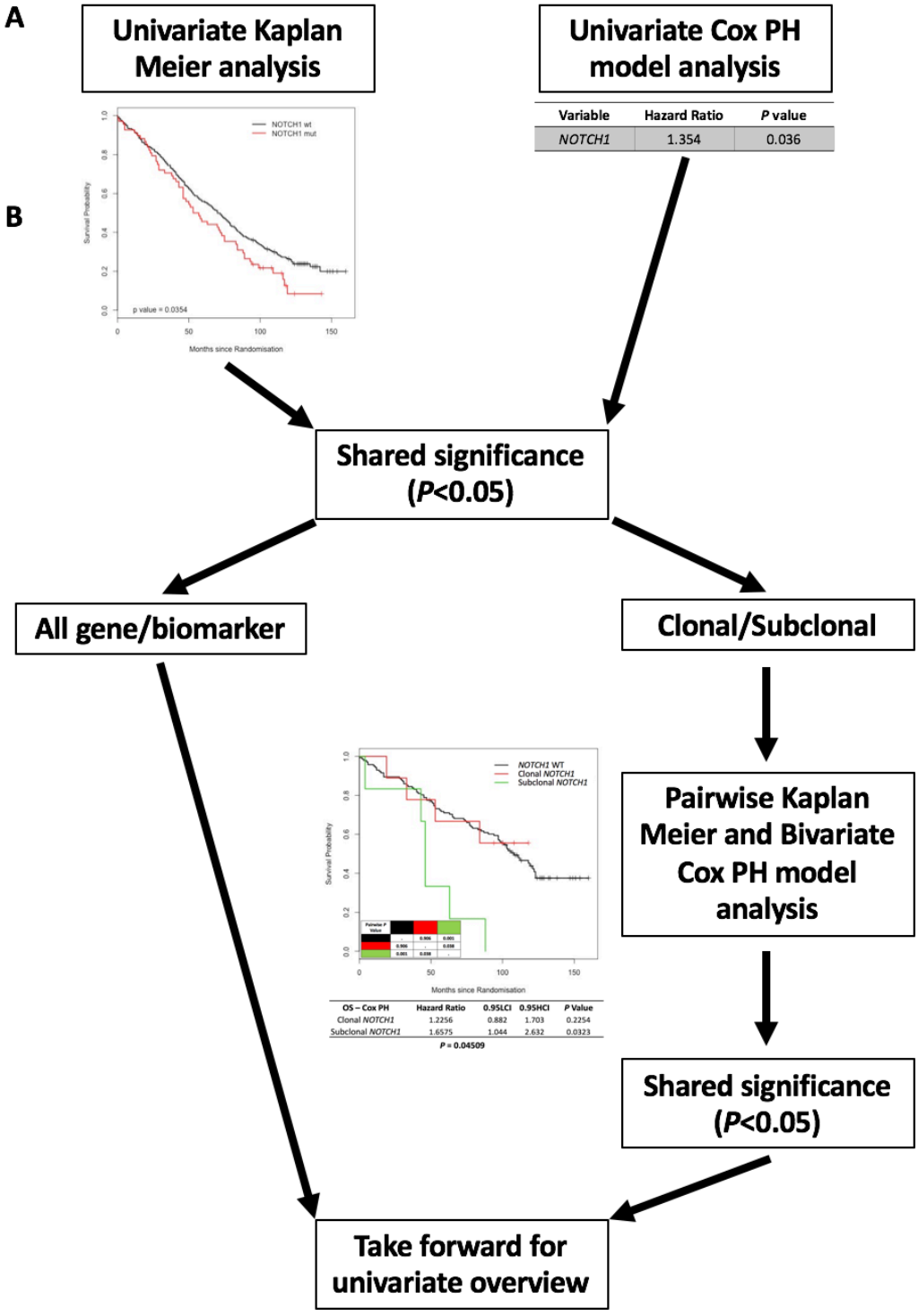



**Figure 57 Overview of Univariate Survival Analysis Process**

Flow diagram of the univariate survival analysis process. Initially, univariate analysis of all variables (subclonal and clonal variables considered as single entities) conducted, with those with shared impact from both Kaplan Meier analysis and Cox PH model analysis taken forward to the next stage ($P$<0.05). If a complete variable, it is taken forward for univariate over and description if a Priority 1 gene. If it is a clonal or subclonal variable, it is tested against its gene partner in pairwise Kaplan Meier and bivariate Cox PH analysis, with only those with concordant impact from both methods taken forward.

### 4.4.1 Overview of variables associated with significant impact on OS in all CLL4 cases

In an analysis of all CLL cases (n = 499), 36 variables were found to have a significant impact on OS in Kaplan Meier analysis (Log Rank *P*<0.05), of which 29 met the basic statistical cut offs used throughout this thesis for other types of multiple hypothesis testing (Log Rank *P*<0.01). Strong concordance was found between Kaplan Meier and univariate Cox PH model analysis, with 34/36 variables found to have a significant impact on OS by both methods (Log Rank *P*<0.05), with clonal mutations in *EGR2* (Log Rank *P* = 0.0619) and all mutations in *CHD2* (Log Rank *P* = 0.0641) being the two variables identified in Kaplan Meier analysis but not Cox PH model analysis.

Clinical CLL biomarkers previously reported to significantly impact OS accounted for 17 out of the 36 variables. Age at randomisation (<60 years), Binet stage A, IGHV mutational status (M-CLL), IGHV homology (<97%), complete response (CR) or nodular partial response (NodPR) (CR/NodPR), del(13q) only, and the group of ten year survivors, all contributed to increased survival in comparison to all other cases. Age at randomisation (>70 years), gender (male), Binet Stage C, *IGHV* homology (>99%), ZAP70 expression, CD38 expression, $\beta$2M expression, del(17p), +(12), and no response/progressive disease (NR/PD) all predicted a poor prognosis in relation to all other cases.

In addition, the following genes were found to have a significant impact on OS (*P*<0.05): pathogenic *ATM* (all), *BRAF* (subclonal, all), *EGR2* (subclonal, all), *KRAS* (subclonal, clonal, all), *MYD88* (all), *NFKBIE* (all), *NOTCH1* (all), *NRAS* (all), *SF3B1* (clonal, all), *TP53* (clonal, all), and any case with a subclonal mutation in any of the genes in the TruSeq panel (denoted as subclonal driver). Pairwise Kaplan Meier analysis and bivariate Cox PH models validated (*P*<0.05) that: subclonal *BRAF*, subclonal *EGR2,* clonal and subclonal *KRAS,* clonal *MYD88,* clonal *SF3B1,* and clonal *TP53* mutations*,* with subclonal mutations in *NOTCH1* and *NRAS* only retaining significance in bivariate Cox PH model analysis.

Power calculations were conducted for all 36 variables for OS, identifying 7 variables which fell below the 80% cut off: Subclonal driver, *CHD2, MYD88, NFKBIE, NOTCH1,* gender, and Binet stage A. Therefore, these variables were excluded from downstream analysis.

**Table 7 All significant and powered CLL univariate OS survival analysis variables**

| Genetic/clinical biomarkers | OS Survival Analysis – All CLL4 cases | | Univariate Kaplan Meier graphs | | Univariate Cox PH models | | |
|---|---|---|---|---|---|---|---|
| | Variable | Number of cases | Median Months OS | Log Rank *P* Value | Hazard Ratio | Log Rank *P* Value | Power |
| Clonal/Subclonal mutations in recurrently mutated genes | Clonal *KRAS* | 12/499 | 32 | 0.00303 | 2.339 | 0.0108 | 0.96 |
| | Clonal *SF3B1* | 104/499 | 54 | 0.00069 | 1.507 | 0.00112 | 0.94 |
| | Clonal *TP53* | 34/499 | 26 | 3.18E-06 | 2.323 | 5.03E-05 | 0.99 |
| | Subclonal *BRAF* | 15/499 | 37 | 2.46E-05 | 2.843 | 0.000458 | 0.99 |
| | Subclonal *EGR2* | 6/499 | 41 | 0.000109 | 4.01 | 0.00306 | 0.99 |
| | Subclonal *KRAS* | 20/499 | 46 | 0.00309 | 1.961 | 0.00811 | 0.93 |
| | Subclonal Driver | 214/499 | 58 | 0.0025 | 1.372 | 0.00276 | 0.85 |
| All mutations in recurrently mutated genes | Pathogenic *ATM* | 39/499 | 52 | 0.0203 | 1.495 | 0.028 | 0.81 |
| | *BRAF* | 31/499 | 46.5 | 0.00494 | 1.732 | 0.00979 | 0.89 |
| | *EGR2* | 16/499 | 41 | 0.000621 | 2.354 | 0.0032 | 0.99 |
| | *KRAS* | 30/499 | 44 | 0.000263 | 2.011 | 0.0011 | 0.99 |
| | *NRAS* | 10/499 | 38 | 0.00303 | 2.428 | 0.0113 | 0.96 |
| | *SF3B1* | 122/499 | 53 | 0.000373 | 1.508 | 0.000599 | 0.96 |
| | *TP53* | 48/499 | 31 | 4.59E-06 | 2.064 | 4.10E-05 | 0.99 |
| Recurrent CNVs | del(17p) | 26/456 | 15 | 0 | 4.983 | 1.30E-10 | 1 |
| | del(11q) | 90/464 | 51 | 5.66E-05 | 1.672 | 0.000146 | 0.99 |
| | +(12) | 70/464 | 44 | 0.0058 | 1.481 | 0.00867 | 0.84 |
| | del(13q) only | 167/462 | 91 | 9.87E-08 | 0.539 | 5.98E-08 | 0.99 |
| *IGHV* | *IGHV* (M-CLL) | 157/420 | 104 | 3.20E-12 | 0.42 | 1.13E-12 | 0.99 |
| | IGHV Homology (<97%) | 122/413 | 118 | 4.17E-12 | 0.376 | 3.78E-13 | 0.99 |
| | IGHV Homology (>99%) | 224/413 | 53 | 3.07E-09 | 2.006 | 3.13E-09 | 0.99 |
| Clinical and biological biomarkers | Age at randomisation (<60 years) | 152/499 | 101 | 5.85E-09 | 0.497 | 1.86E-09 | 0.99 |
| | Age at randomisation (>70 years) | 153/499 | 46.5 | 6.26E-12 | 2.094 | 8.69E-11 | 0.99 |
| | Binet Stage C | 156/499 | 51 | 0.00424 | 1.37 | 0.00527 | 0.84 |
| | ZAP70 | 189/380 | 66 | 0.000281 | 1.553 | 0.000289 | 0.94 |
| | CD38 | 268/422 | 53 | 1.83E-11 | 2.324 | 6.36E-12 | 0.99 |
| | β-2-M | 164/357 | 49 | 1.81E-12 | 2.409 | 6.22E-12 | 0.99 |
| | Response CR/NodPR | 184/470 | 88 | 2.53E-05 | 0.625 | 2.18E-05 | 0.97 |
| | Response NR/PD | 98/470 | 37 | 3.98E-07 | 1.877 | 2.51E-06 | 0.99 |
| | Ten year survivors | 119/494 | NA | 0 | 1.15E-09 | 0 | 1 |

Median OS for all wild type groups was 70 months

### 4.4.2 Overview of variables associated with significant impact on PFS in all CLL4 cases

In all CLL cases studied 23 variables were found to have a significant impact on PFS in Kaplan Meier analysis (Log Rank $P<0.05$), with 18 of these significant variables surpassing the $P<0.01$ level. Complete concordance between Kaplan Meier analysis and Cox PH models was observed, with all variables in both survival methods having a significant impact on PFS ($P<0.05$).

Previously identified clinical CLL biomarkers with a statistical impact on PFS were observed in 14/23 variables **(Table 8)**. IGHV mutational status (M-CLL), del(13q) only, IGHV Homology (<97%), CR/NodPR, ten year survivors, and FC treatment all showed increased PFS, with: del(11q), del(17p), gender (male), IGHV Homology (>99%), ZAP70 expression, CD38 expression, β2M expression, and NR/PD showing reduced PFS.

From the TruSeq data, the following variables had a significant impact on PFS ($P<0.05$): *DDX3X* (subclonal, all), *EGR2* (subclonal, all), *KRAS* (clonal), *NFKBIE* (all), *TP53* (clonal, all). Pairwise Kaplan Meier analysis and bivariate Cox PH models validated ($P<0.05$) that: subclonal *EGR2,* clonal *KRAS,* and clonal *TP53* mutations, with subclonal *DDX3X* mutations only retaining significance in pairwise Kaplan Meier analysis.

Power calculations were conducted for all 23 variables for PFS, identifying 2 variables which fell below the 80% cut off: *NFKBIE,* and gender. Therefore, these variables were excluded from downstream analysis.

**Table 8 All significant and powered CLL univariate PFS survival analysis variables**

| Genetic/clinical biomarkers | PFS Survival Analysis – All CLL4 cases | | Univariate Kaplan Meier graphs | | Univariate Cox PH models | | |
|---|---|---|---|---|---|---|---|
| | Variable | Number of cases | Median Months PFS | Log Rank *P* Value | Hazard Ratio | Log Rank *P* Value | Power |
| Clonal/Subclonal mutations in recurrently mutated genes | Clonal *KRAS* | 12/499 | 19 | 0.0181 | 2.0008 | 0.0163 | 0.89 |
| | Clonal *TP53* | 34/499 | 6 | 0.00112 | 1.8137 | 0.001051 | 0.98 |
| | Subclonal *DDX3X* | 7/499 | 5 | 0.00955 | 2.6271 | 0.008631 | 0.98 |
| | Subclonal *EGR2* | 6/499 | 8.5 | 0.025 | 2.4651 | 0.02361 | 0.92 |
| All mutations in recurrently mutated genes | *DDX3X* | 21/499 | 6 | 0.00796 | 1.8262 | 0.007632 | 0.93 |
| | *EGR2* | 16/499 | 13.5 | 0.015 | 1.851 | 0.01465 | 0.87 |
| | *TP53* | 48/499 | 7 | 0.00104 | 1.6597 | 0.001017 | 0.97 |
| Recurrent CNVs | del(11q) | 90/464 | 15 | 9.68E-06 | 1.7114 | 8.88E-06 | 0.99 |
| | del(17p) | 26/456 | 3 | 0 | 5.244 | 0 | 1 |
| | del(13q) only | 167/462 | 33 | 1.68E-05 | 0.6445 | 1.61E-05 | 0.98 |
| *IGHV* | *IGHV* (M-CLL) | 157/420 | 35 | 9.70E-12 | 0.4778 | 1.01E-11 | 0.99 |
| | *IGHV* Homology (<97%) | 122/413 | 40 | 9.39E-12 | 0.4485 | 1.00E-11 | 0.99 |
| | *IGHV* Homology (>99%) | 224/413 | 18.5 | 5.05E-11 | 1.9793 | 5.18E-11 | 0.99 |
| Clinical and biological biomarkers | FC Treatment | 144/499 | 41 | 9.04E-10 | 0.527 | 9.53E-10 | 0.99 |
| | CHL Treatment | 237/499 | 19 | 1.35E-06 | 1.57259 | 1.34E-06 | 0.99 |
| | ZAP70 | 189/380 | 19 | 1.37E-06 | 1.682 | 1.38E-06 | 0.99 |
| | CD38 | 268/422 | 19 | 9.76E-06 | 1.6055 | 9.58E-06 | 0.98 |
| | β-2-M | 164/357 | 22 | 4.41E-06 | 1.6707 | 4.32E-06 | 0.99 |
| | Response CR/NodPR | 184/470 | 40.5 | 0 | 0.4262 | 0 | 0.99 |
| | Response NR/PD | 98/470 | 3 | 0 | 39.8494 | 0 | 1 |
| | Ten year survivors | 119/494 | 51 | 0 | 0.3395 | 0 | 0.99 |

Median PFS for all wild type groups was 24 months

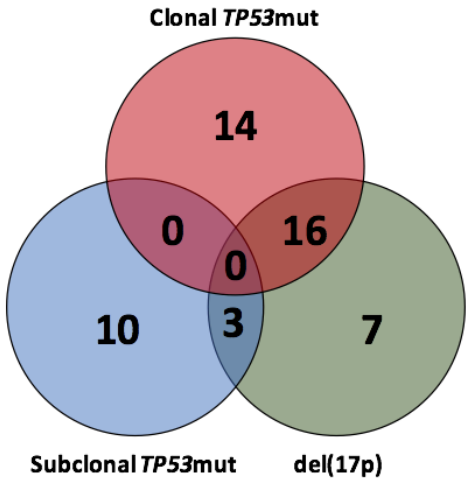
### 4.4.3 Impact of the DNA Damage Response and inflammatory/BCR signalling pathway genes on survival in CLL4

#### 4.4.3.1 Subclonal *TP53* mutations on their own do not infer poor prognosis in CLL4

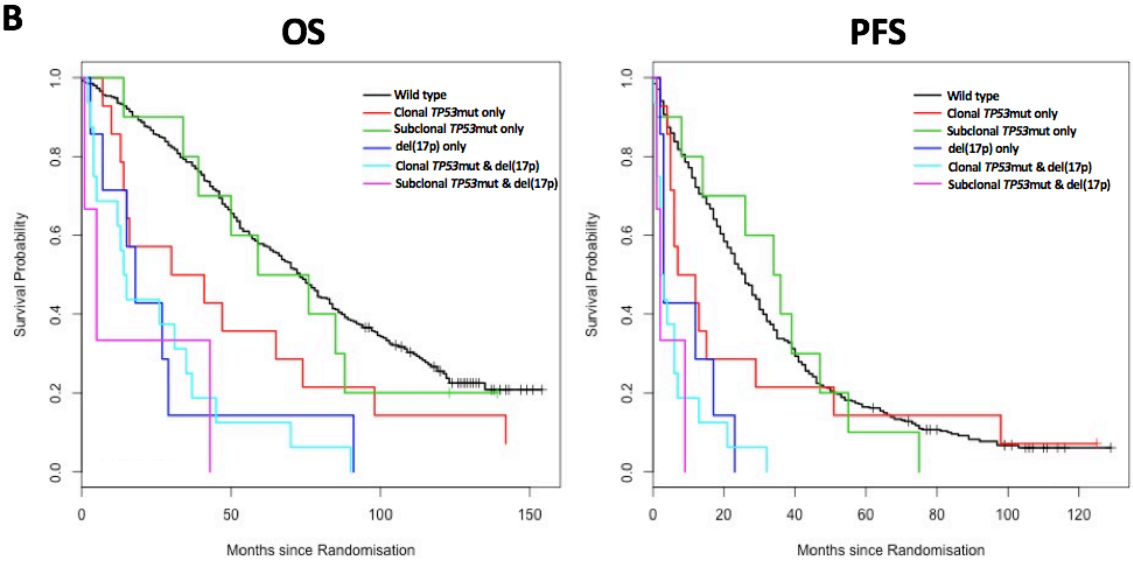
Subclonal *TP53* mutations have been proposed to infer a poor prognosis in CLL, comparable to that of clonal *TP53* mutations(8,159). However, this inference has only been studied in non-clinical trial cohorts. Furthermore, the pairwise assessment between all subgroups including del(17p) has not been presented in the literature, with comparisons only drawn to either clonal *TP53* mutations (8,159), or "clonal *TP53* lesions", which is a combined variable of clonal *TP53* mutations and del(17p) (8). Previously, *TP53* mutations were found to define a subgroup of patients with poor outcome, and supported the notion that screening of *TP53* mutations added clinical value (167). Therefore, since this TruSeq study had the sensitivity to capture mutations down to 1% VAF, analysis of clonal (>12% VAF) and subclonal (<12% VAF) mutations was assessed in the context of del(17p) to verify the prognostic impact of subclonal *TP53* mutations.

In univariate survival analysis, clonal *TP53* mutations and del(17p) events were found to associate with poor OS and PFS (clonal *TP53* = OS: Median = 26 vs. 72, $P<0.001$, PFS: Median = 6 vs. 26, $P = 0.001$, del(17p) = OS: Median = 15 vs. 72, $P<0.001$, PFS: Median = 3 vs. 25.5, $P<0.001$), with subclonal *TP53* mutations showing no impact on either OS or PFS (OS: Median = 43 vs. 70, $P = 0.095$, PFS: Median = 20 vs. 24, $P = 0.29$). Clonal and subclonal *TP53* mutations were mutually exclusive (**Figure 58A)**, although there were 3 cases with more than 1 *TP53* variant (2x clonal variants n = 2, 2x subclonal variants n = 1). Furthermore, 73% of del(17p) events had either a clonal (n = 16) or subclonal mutation (n = 3), with 53% of clonal mutations having a del(17p) event, and 23% of subclonal mutations having a del(17p) **(Figure 58A)**. In pairwise survival analysis, subclonal *TP53* mutations without del(17p) events were found to have no impact on OS nor PFS (OS: Median = 67.5 vs. 73, $P = 0.743$, PFS: Median = 35 vs. 26, $P = 0.972$) (**Figure 58B**, left and right). Intriguingly, clonal *TP53* mutations alone did not predict for poor PFS (Median = 9.5 vs. 26, $P = 0.324$), although until 20 months the separation from the wild type curve was indicative of poor outcome, and since the median PFS is 9.5 months, it suggests that not all clonal *TP53* variants have the same impact on PFS. Furthermore, in co-occurrence analysis clonal *TP53* mutated ($P<0.0001$), as well as all *TP53* mutated cases ($P<0.0001$) associated with non-response/progressive disease, further suggesting that clonal *TP53* mutations associate with poor PFS. From the four clonal *TP53* mutation only cases which had a PFS higher than the median for wild type cases, two were del(13q) only cases, one was Binet stage A, and all four were treated with FC, all features that could mitigate the impact of *TP53* aberration.

**A**




**B**




| Survival | Pairwise OS - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI | Pairwise *P* Value (to WT) |
|---|---|---|---|---|---|---|---|
| OS | Wild type | 406 | 289 | 73 | 67 | 79 | - |
| | Clonal *TP53*mut only | 14 | 13 | 35.5 | 15 | 142 | 0.023 |
| | Subclonal *TP53*mut only | 10 | 8 | 67.5 | 39 | NA | 0.743 |
| | del(17p) only | 7 | 7 | 18 | 7 | NA | <0.001 |
| | Clonal *TP53*mut & del(17p) | 16 | 16 | 14.5 | 5 | 45 | <0.001 |
| | Subclonal *TP53*mut & del(17p) | 3 | 3 | 5 | 1 | NA | <0.001 |
| PFS | Wild type | 406 | 371 | 26 | 23 | 28 | - |
| | Clonal *TP53*mut only | 14 | 13 | 9.5 | 6 | 98 | 0.324 |
| | Subclonal *TP53*mut only | 10 | 10 | 35 | 14 | NA | 0.972 |
| | del(17p) only | 7 | 7 | 3 | 3 | NA | <0.001 |
| | Clonal *TP53*mut & del(17p) | 16 | 16 | 2.5 | 2 | 13 | <0.001 |
| | Subclonal *TP53*mut & del(17p) | 3 | 3 | 2 | 1 | NA | <0.001 |

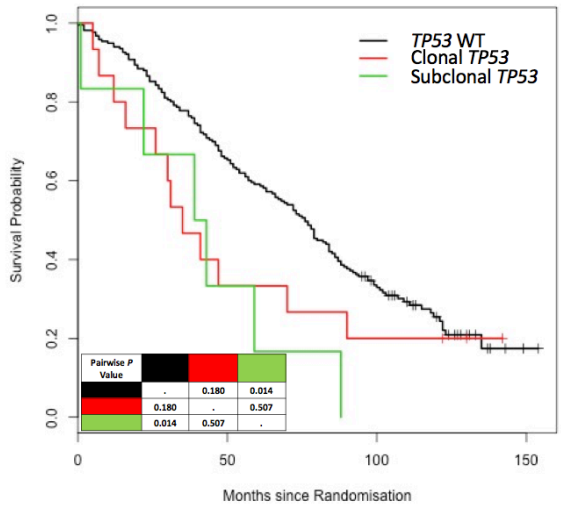
**Figure 58 *TP53* survival analysis in the context of del(17p)**

**A** Venn diagram of clonal *TP53* mutation subclonal *TP53* mutation and del(17p) status. **B** Pairwise Kaplan Meier analysis of stratified clonal and subclonal *TP53* mutations in the context of del(17p) (black = wild type, red = clonal *TP53*mut only, green = subclonal *TP53*mut only, blue = del(17p) only, light blue = clonal *TP53*mut and del(17p), magenta = subclonal *TP53*mut and del(17p) in OS (left graph) and PFS (right graph) (upper = Kaplan Meier graphs, lower = Kaplan Meier associated survival tables).

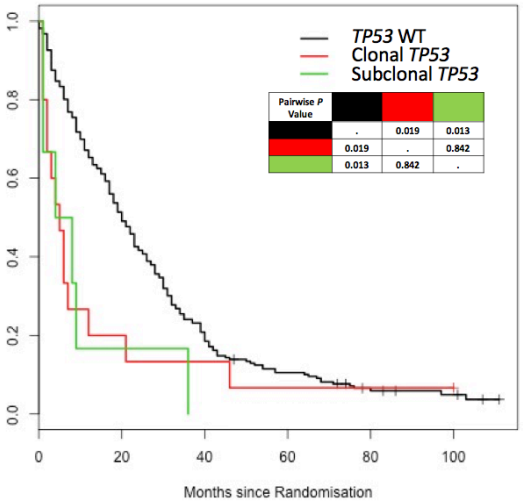
#### 4.4.3.1.1 Stratified survival analysis of CLL4 treatment groups unearths reduced survival for subclonal *TP53* mutated cases treated with chlorambucil

*TP53* mutated cases respond poorly to chemotherapy, therefore stratified analysis based on treatment type was conducted to assess whether specific treatment regimens account for the impact of subclonal *TP53* mutations in CLL4. The FC treatment group showed no impact of subclonal *TP53* mutations on either OS or PFS (OS: Median = 76 vs. 67, *P* = 0.827, PFS: Median = 39 vs. 41, *P* = 0.425), with the fludarabine treatment group being underpowered to assess the role of subclonal *TP53* mutations (n = 1). Unfortunately, analysis of subclonal *TP53* mutations in the context of del(17p) events in the CHL treatment group was also underpowered, however analysis of only clonal and subclonal *TP53* mutations was possible. Subclonal *TP53* mutated cases predicted for poor OS and PFS (OS: Median = 41 vs. 76, *P* = 0.014, PFS: Median = 6 vs. 20, *P* = 0.013), however clonal *TP53* mutated cases only predicted for poor PFS (OS: Median = 35 vs. 76, *P* = 0.180, PFS: Median = 5 vs. 20, *P* = 0.019) **(Figure 59)**. There was only 1 clonal *TP53* mutated case which survived past the median wild type survival time, but this patient was a long-term survivor, censored at last follow up 142 months since randomisation. These results together suggest that subclonal *TP53* mutations might have an impact on OS and PFS, however they need to be interpreted carefully in the context of not being able to conduct powered analysis of the same event in other treatment groups, nor being able to analyse subclonal *TP53* mutated cases in the context of del(17p) events. Regardless, since an association with poor OS and PFS was found for subclonal *TP53* mutations in CHL treated cases, assessment of clonal expansion in CHL treated *TP53* mutated cases (clonal or subclonal) of available 2$^{nd}$ randomisation material was undertaken, to identify the occurrence of chemotherapy induced selection.

**A**




**B**




**C**

| Survival | Pairwise - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|---|
| **OS** | *TP53* WT | 216 | 156 | 76 | 66 | 84 |
| | Clonal *TP53* | 15 | 12 | 35 | 26 | NA |
| | Subclonal *TP53* | 6 | 6 | 41 | 22 | NA |
| **PFS** | *TP53* WT | 216 | 204 | 20 | 17 | 23 |
| | Clonal *TP53* | 15 | 14 | 5 | 2 | 21 |
| | Subclonal *TP53* | 6 | 6 | 6 | 1 | NA |

**D**

| Survival | Cox PH | Hazard Ratio | 0.95LCI | 0.95UCI | *P* Value | Model *P* Value |
|---|---|---|---|---|---|---|
| **OS** | Clonal *TP53* | 1.4942 | 0.8285 | 2.695 | 0.182 | 0.02613 |
| | Subclonal *TP53* | 2.6361 | 1.1611 | 5.985 | 0.0205 | |
| **PFS** | Clonal *TP53* | 1.8925 | 1.098 | 3.263 | 0.0217 | 0.004693 |
| | Subclonal *TP53* | 2.6147 | 1.156 | 5.914 | 0.021 | |

**Figure 59 Subclonal *TP53* mutations with impact on OS and PFS in CHL treated CLL**

Pairwise Kaplan Meier analysis of clonal and subclonal *TP53* mutated cases in CHL treated CLL for OS **(A)** and PFS **(B)** (black = wild type, red = clonal *TP53*, green = subclonal *TP53*). Cases with both clonal and subclonal mutations were placed in the clonal group. **C** Pairwise Kaplan Meier table of data from **(A)** and **(B)**. **D** Bivariate Cox PH model for OS and PFS, clonal *TP53* mutated cases vs. subclonal *TP53* mutated cases.

#### 4.4.3.1.2 Clonal expansion in response to CHL by a *TP53* R282W subclonal mutation

Previous studies have assessed the link between subclonal *TP53* mutations and clonal evolution during patient therapy. There were three *TP53* mutated cases with available DNA from the second randomisation of the trial treated with CHL, therefore these 2nd randomisation samples were screened for mutations in *TP53* using the Ion Torrent NGS platform **(2.1.13)**. Only 1/3 cases showed clonal expansion post therapy (U-CLL 4572, **Figure 60A**). One case had two clonal mutations in *TP53* at diagnosis, neither of which changed significantly in VAF (M-CLL 4491, **Figure 60B**). The final case (M-CLL 4267) had a low subclonal mutation (2% VAF), which increased marginally at 2nd randomisation (8% VAF), suggesting that a mutation in *TP53* at a subclonal level is not sufficient to be selected for by chemotherapy in this patient, and that specific types of mutations may be more important than others (**Figure 60C**). However, this analysis was specific to *TP53,* therefore it was unable to screen the other genomic aberrations and mutations in these cases. This means that tracking of the other events was not possible, and would be beneficial to complete the analysis of clonal evolution in these patients.

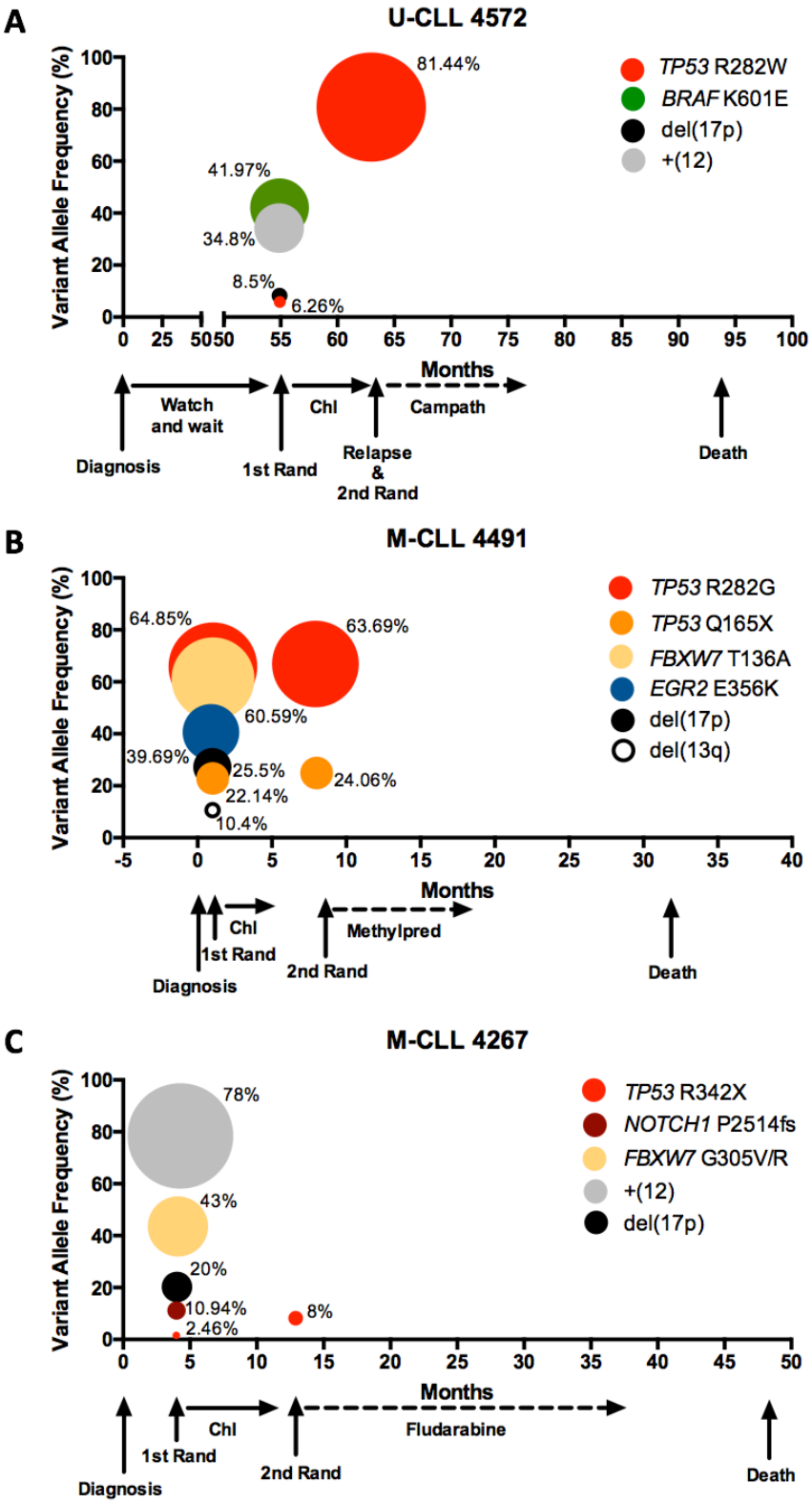



**Figure 60 Longitudinal analysis of clonal evolution in *TP53* mutated cases**

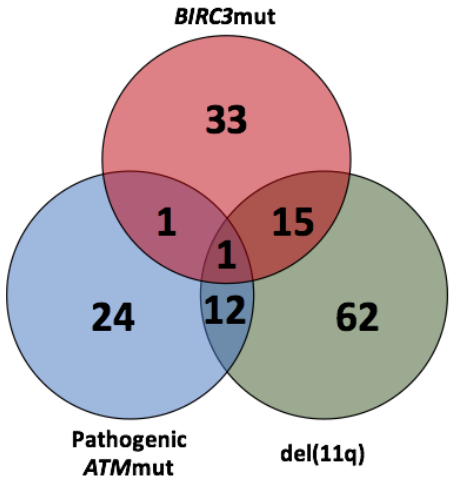
Graphical representation of mutation profile of three *TP53* mutated cases, including *TP53* mutation status at 2nd randomisation. Graph shows VAF (*y* axis) and months since diagnosis (*x* axis), with the clinical timeline for each patient below each graph. Rand = Randomisation

#### 4.4.3.2 Co-mutated/deleted del(11q) and *BIRC3,* and biallelic *ATM* predict poor PFS in CLL4

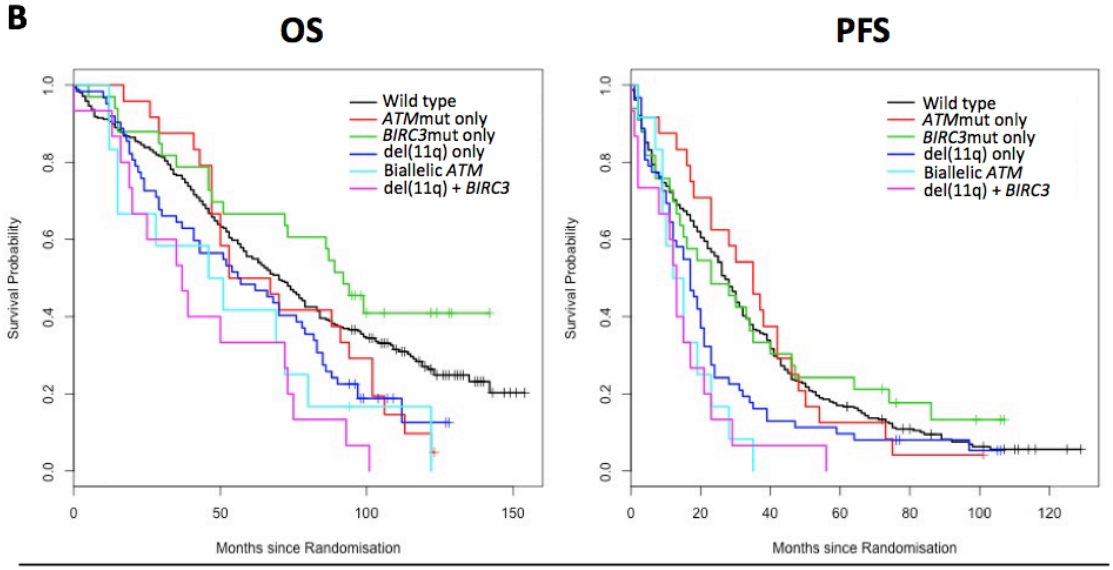
The question of the impact of biallelic *ATM* and *BIRC3* on prognosis in the context of del(11q) cases in CLL has been controversial, since there are contrasting studies in the literature (61,49). In a study of multiple non-clinical trial cohorts, biallelic *BIRC3* (*BIRC3* mutation and deletion) was found to be mutually exclusive to *TP53* aberrations (mutation and/or deletion), and predicted for a poor OS (49). However, this study did not provide univariate biallelic *ATM* survival analysis, only stating that biallelic *BIRC3* was independent of biallelic *ATM* in a multivariate Cox PH model. In comparison, previous work from this laboratory showed in an analysis of 133 CLL4 cases that biallelic *ATM* (mutation and deletion) predicted a subgroup with poor outcome, but not *BIRC3*(61). Therefore, since this study had an increased number of del(11q) cases to the previous CLL4 study, as well as more sensitive mutation screening of *ATM* and *BIRC3*, an analysis of the impact of biallelic *BIRC3* and *ATM* was undertaken.

In univariate Kaplan Meier analysis of OS and PFS in the entire CLL4 cohort, *BIRC3* was not found to significantly reduce survival (OS: Median = 73 vs. 68, $P = 0.95$, PFS: Median =20 vs. 25, $P = 0.61$). However, Pathogenic *ATM* mutations did significantly impact OS (OS: Median = 53 vs. 70, $P = 0.03$, PFS: Median = 23 vs. 24.5, $P = 0.37$). When evaluated in the context of del(11q), it was observed that all but two cases (*ATM* and *BIRC3* mutated n=1, biallelic *ATM* and del(11q) + *BIRC3* n=1), were mutually exclusive (*ATM* n=12, *BIRC3* n=15, **Figure 61A**). These subgroups were then taken forward for pairwise Kaplan Meier survival analysis, identifying that both del(11q) + *BIRC3* (PFS: Median = 13 vs. 27, $P = 0.001$) and biallelic *ATM* (PFS: Median = 13.5 vs. 27, $P =0.002$) predicted for poor PFS in comparison to wild type cases **(Figure 61B, right)**. Interestingly, only del(11q) + *BIRC3* was found to significantly impact OS (OS: Median = 37 vs. 70, $P = 0.001$) in comparison to wild type cases, with a strong trend towards a difference in OS to del(11q) only cases ($P = 0.061$) **(Figure 61B, left)**. Biallelic *ATM* showed a strong trend towards an impact on OS (OS: Median = 46 vs. 70, $P = 0.051$), but was not significantly different to del(11q) only cases ($P = 0.396$). However, caution must be taken from the findings of this analysis, since copy number data on all cases was not available, therefore the true deletion status of *BIRC3* is not known.

**A**




**B**




| Survival | Pairwise OS - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI | Pairwise *P* value (to WT) | Pairwise *P* value (to del(11q)) |
|---|---|---|---|---|---|---|---|---|
| OS | Wild type | 316 | 223 | 70 | 61 | 78 | - | 0.023 |
| | *ATM*mut only | 24 | 22 | 60 | 50 | 102 | 0.218 | 0.535 |
| | *BIRC3*mut only | 33 | 19 | 92 | 72 | NA | 0.162 | 0.008 |
| | del(11q) only | 62 | 51 | 56.5 | 41 | 79 | 0.023 | - |
| | Biallelic *ATM* | 12 | 11 | 48.5 | 15 | NA | 0.051 | 0.396 |
| | del(11q) + *BIRC3* | 15 | 15 | 37 | 20 | 75 | 0.001 | 0.061 |
| PFS | Wild type | 316 | 289 | 27 | 25 | 30 | - | 0.009 |
| | *ATM*mut only | 24 | 23 | 35 | 23 | 46 | 0.834 | 0.06 |
| | *BIRC3*mut only | 33 | 28 | 23 | 14 | 46 | 0.523 | 0.074 |
| | del(11q) only | 62 | 58 | 17 | 12 | 21 | 0.009 | - |
| | Biallelic *ATM* | 12 | 12 | 13.5 | 9 | NA | 0.002 | 0.218 |
| | del(11q) + *BIRC3* | 15 | 15 | 13 | 8 | 23 | 0.001 | 0.114 |

**Figure 61 del(11q) + *BIRC3* and biallelic *ATM* survival analysis**

**A** Venn diagram of *BIRC3, ATM,* and del(11q) mutation/deletion status. **B** Pairwise Kaplan Meier analysis of stratified biallelic *BIRC3* and *ATM* in the context of del(11q) (black = wild type, red = *ATM*mut only, green = *BIRC3*mut only, blue = del(11q) only, light blue = biallelic *ATM,* magenta = biallelic *BIRC3*) in OS (left graph) and PFS (right graph) (upper = Kaplan Meier graphs, lower = Kaplan Meier associated survival tables).

#### 4.4.3.3 Inflammatory pathway/BCR signalling genes *KRAS, NRAS,* and *BRAF* all significantly associate with poor OS in CLL4

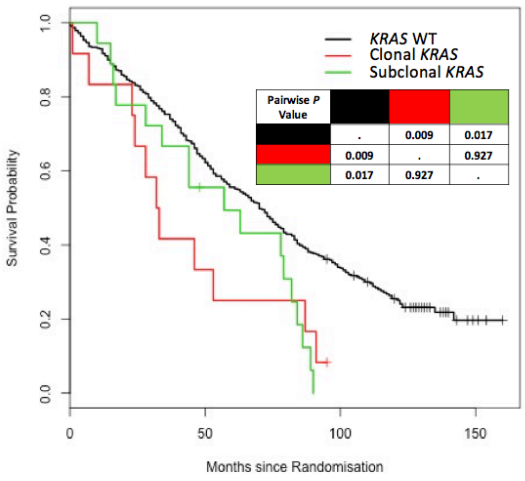
Essential intracellular signalling pathway transducers, Kristen Rat Sarcoma (*KRAS*) and Neuroblastoma Rat Sarcoma (*NRAS*) are GTPases which after Guanine Exchange Factor (GEF) activation facilitate downstream pathway transduction (192). KRAS and NRAS binding and subsequent activation of other core signalling components such as Grb2 and SOS perpetuates signal specific cellular programs, from apoptosis to proliferation. *KRAS* and *NRAS* are oncogenes, with mutations or overexpression of the gene found in 25% of all cancers. A tyrosine kinase implicated in the Ras intracellular signalling pathways, v-Raf murine sarcoma viral oncogene homologue B1 (*BRAF*) is essential for MAP Kinase signalling, leading to the transcription of genes associated with growth and survival (193). *BRAF* mutations in cancer have been well documented, with activating mutations defining over 50% of advanced melanoma patients, of which the V600E variant is the most common (194).

*KRAS, NRAS,* and *BRAF* mutations have recently been described in the CLL literature, with all genes observed in less than 5% of cases in the Landau and Puente studies (57,96). Since, a targeted re-sequencing study identified mutations in each gene at higher frequencies (*KRAS* = 8%, *NRAS* = 2%, and *BRAF* = 3.5%), however except for a co-occurrence between *KRAS* mutation and non-responsiveness, no association with survival outcome was reported for any of the genes(162). Therefore, considering the cohort size of this CLL4 TruSeq study, although the frequency of mutations in these genes were found to be comparable to the aforementioned targeted re-sequencing study (*KRAS* = 6.5%, *NRAS* = 2%, and *BRAF* = 7%, **Figure 33**), it has the appropriate sample size and statistical power (*KRAS* = 99%, *NRAS* = 96%, *BRAF* = 89%) to test their association with outcome. In a pan-European study, *BRAF* mutation data from this cohort (n = 247), was included in a study describing the resistance of genes to fludarabine *in vitro* after transposase mutagenesis, identifying *BRAF* as a target (195). No survival analysis of the CLL4 data was presented, only genetic characterisation of clonal *BRAF* mutations (<12% VAF) found in the cohort, since the filtering of high quality variants at this stage of the project was rudimentary.

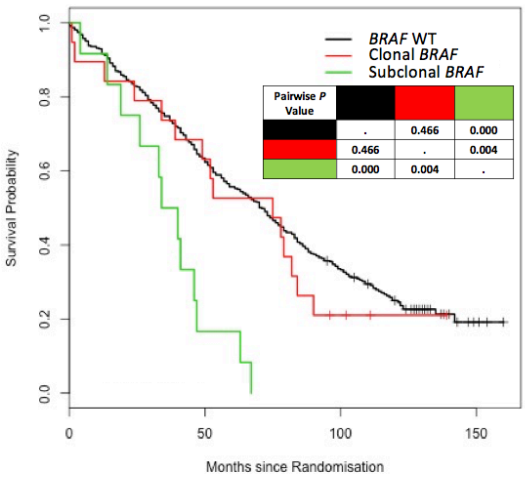
*KRAS* mutations were found to associate with *IGHV* with homology to germline >99% ($P$<0.001) and CD38 positive cases ($P$ = 0.006), whilst *NRAS* and *BRAF* showed no associations with clinical biomarkers after multiple hypothesis testing ($Q$>$P$, $P$<0.01). In univariate Kaplan Meier survival analysis, all three genes predicted for poor OS (*KRAS*: Median = 44 vs. 70, $P$<0.0001, *NRAS:* Median = 38 vs. 70, $P$ = 0.0117 **[Figure 62D&E]**, *BRAF:* Median = 46.5 vs. 71, $P$ = 0.005). In pairwise analysis of clonal and subclonal mutations, both clonal and subclonal *KRAS* mutations predicted poor OS (Clonal: Median = 32.5 vs. 70, $P$ = 0.009, Subclonal: Median = 57 vs. 70, $P$ = 0.017) **(Figure**

**62A&C)**, whilst neither subgroup predicted reduced survival in *NRAS,* and only subclonal mutations in *BRAF* associated with reduced OS (Clonal: Median = 75 vs. 71, *P* = 0.466, Subclonal: Median = 37 vs. 71, *P*<0.0001) **(Figure 62B&C)**. Altogether, this suggests that abrogation of the Ras intracellular signalling pathway in CLL associates with reduced survival. However, since mutations in these genes are generally mutually exclusive in cancer, assessing the survival impact of these genes after removal of co-mutated cases is required.

**A**



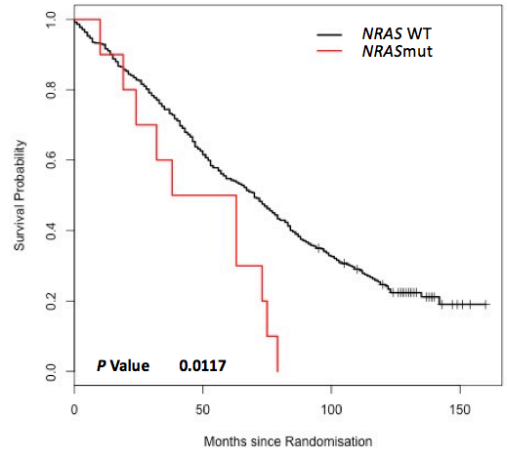

**B**




**C**

| Hugo Gene ID | Pairwise OS - Kaplan Meier | Number of cases | Number of events | Median OS | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|---|
| ***KRAS* (A)** | Wild type | 469 | 338 | 70 | 63 | 77 |
| | Clonal | 12 | 11 | 32.5 | 24 | NA |
| | Subclonal | 18 | 17 | 57 | 34 | 86 |
| ***BRAF* (B)** | Wild type | 468 | 339 | 71 | 64 | 78 |
| | Clonal | 19 | 15 | 75 | 49 | NA |
| | Subclonal | 12 | 12 | 37 | 26 | NA |

**D**




**E**

| Univariate - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|
| *NRAS* WT | 489 | 356 | 70 | 62 | 77 |
| *NRAS*mut | 11 | 11 | 38 | 24 | NA |

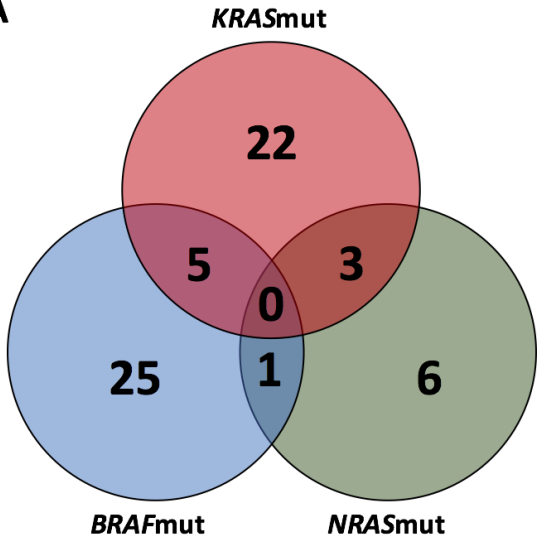
**Figure 62 Overview of impact of *KRAS, NRAS,* and *BRAF* mutations on OS in CLL4**

Pairwise Kaplan Meier analysis of clonal and subclonal mutations for OS in *KRAS* **(A)** and *BRAF* **(B)**, with associated Kaplan Meier survival table in **(C)**. **D** Univariate Kaplan Meier analysis for OS in *NRAS.* **E** Univariate Kaplan Meier table for *NRAS,* data from **(D)**.

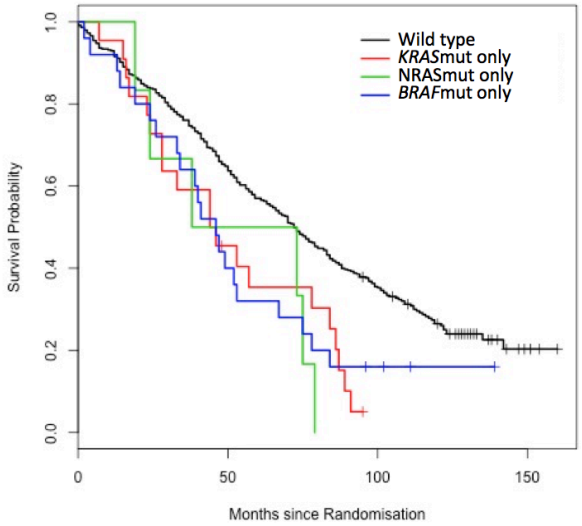
#### 4.4.3.3.1 Mutually exclusive *KRAS* and *BRAF* mutations have a negative impact on OS in CLL4

Chemotherapy-refractory metastatic colorectal cancer patients treated with cetuximab, cetuximab plus chemotherapy, or panitumumab, were found to have added benefit from screening of *KRAS, NRAS, BRAF,* and *PIK3CA* mutations, by identifying mutually exclusive cases that were *KRAS* wild type (196). Although anti-EGFR based therapies are not common in CLL, the anti-EGFR compound gefitinib has been tested in ZAP70 positive CLL cells *in vitro*, showing pre-clinical efficacy (197). Therefore, the mutual exclusivity of *KRAS, NRAS,* and *BRAF* was assessed, followed by pairwise Kaplan Meier survival analysis of mutually exclusive cases from each gene. 27/35 *KRAS,* 7/11 *NRAS,* and 25/31 *BRAF* mutated cases were mutually exclusive of the other two genes in this analysis, with 5 cases sharing *KRAS* and *BRAF* mutations, 3 cases sharing *KRAS* and *NRAS* mutations, and 1 case having dual *NRAS* and *BRAF* mutations **(Figure 63A)**. In pairwise Kaplan Meier survival analysis, mutually exclusive *KRAS* (OS: Median = 44 vs. 72, $P = 0.003$) and *BRAF* (OS: Median = 46 vs. 72, $P = 0.014$) mutations were found to significantly impact OS in comparison to wild type, with mutations in *NRAS* showing a strong trend for reduced OS (OS: Median = 38 vs. 72, $P = 0.063$) **(Figure 63B&C)**. Although no significant co-occurrence (after accounting for multiple hypothesis testing: $Q>P$ or $P<0.01$) was observed in all mutations of *KRAS* ($P = 0.015$), *NRAS* ($P = 0.501$), or *BRAF* ($P = 0.407$), high proportions of the mutually exclusive cases for each gene were ZAP70 positive (*KRAS* = 14/19, *NRAS* = 4/5, *BRAF* = 10/18, reduced numbers due to missing ZAP70 expression data), suggesting that if anti-EGFR therapy is going to be pursued for ZAP70 positive CLL, screening for mutations in these genes is advised based on the findings in colorectal cancer. Furthermore, *KRAS* and *NRAS* mRNA has been found to be upregulated in cases not responding to CHL combination therapy with Rituximab in CLL (198), again suggesting the potential importance of screening for these mutations in CLL.

A




B




C

| Pairwise OS - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI | Pairwise *P* value (to WT) |
|---|---|---|---|---|---|---|
| Wild type | 438 | 311 | 72 | 66 | 80 | - |
| *KRAS*mut only | 22 | 20 | 45 | 28 | 86 | 0.003 |
| *NRAS*mut only | 6 | 6 | 55.5 | 24 | NA | 0.063 |
| *BRAF*mut only | 25 | 21 | 46 | 34 | 75 | 0.014 |

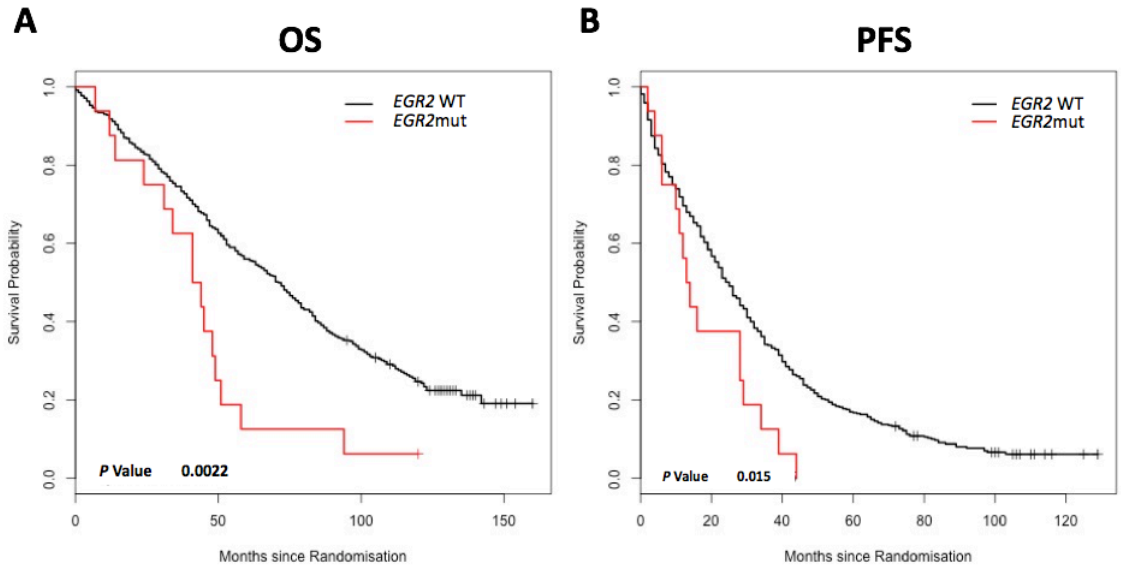
**Figure 63 OS survival analysis of mutually exclusive *KRAS, NRAS,* and *BRAF* mutations**

**A** Venn diagram of *KRAS, NRAS,* and *BRAF* mutations. **B** Pairwise Kaplan Meier analysis of mutually exclusive mutations from each gene (black = wild type, red = *KRAS*mut only, green = *NRAS*mut only, blue = *BRAF*mut only) in OS **B** Pairwise Kaplan Meier analysis table associated with the data from **(B)**.

#### 4.4.3.4 Subclonal *EGR2* mutations predict for a poor OS and PFS in CLL4

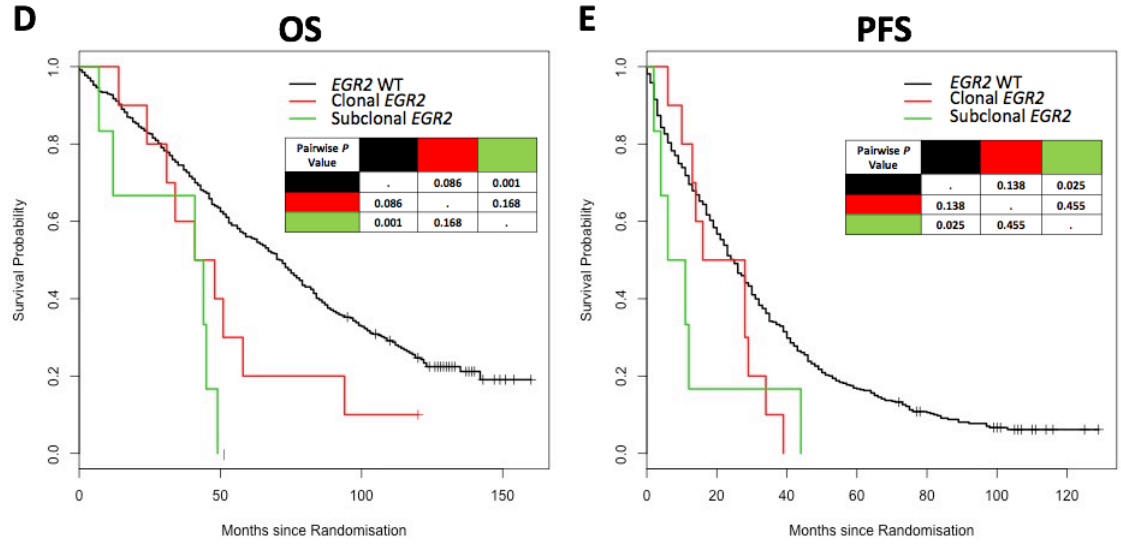
An integral contributor to haematopoietic cell development pathways, Early Growth Response gene 2 (*EGR2*) is a zinc finger transcription factor which supports the differentiation of B-cells(199), T-cells(199) and myeloid cells (200). In cancer, overexpression of EGR2 in cell lines has been associated with increased apoptosis (201), whilst overexpression of PTEN in endometrial cancer cell lines transcriptionally transactivated EGR2 (202). *EGR2* mutations in CLL have been found to occur in haematopoietic progenitors, almost exclusively mutated in the zinc finger domains of the gene (12). The suggested functional role of these mutations in this study are that they modify transcription factor function. In a pan-European study, including mutation data from this CLL4 study (n = 366) as a validation cohort, *EGR2* mutations in the screening cohort were found to associate with clinically aggressive subgroups of CLL (CD38 expression, U-CLL), were co-mutated *ATM* lesions and *TP53* aberrations, and predicted shorter time-to-first-treatment (TTFT) and OS, with the CLL4 data and other confirmation cohorts also finding a survival association between *EGR2* mutations and OS (11). However, the mutations provided for this publication were capped at 5% VAF, as the training strategy for *in silico* validation of minor subclones (<5% VAF) was in its infancy. Furthermore, the aim of this study was not to assess the importance of clonal or subclonal mutations in *EGR2* on survival, but to confirm across multiple cohorts that screening for mutations in *EGR2* was beneficial for identifying poor risk patients. Therefore, in this expanded cohort, including mutations at a higher resolution (down to 1% VAF), the impact of clonal and subclonal mutations in *EGR2* was explored.

*EGR2* mutations were found to co-occur with cause of death by Richter's syndrome ($P = 0.004$), and were found to be mutually exclusive from IGHV cases with homology to germline <97%. In an analysis of clonal and subclonal *EGR2* mutations, and after accounting for multiple hypothesis testing, only the association with cause of death by Richter's syndrome was retained ($P = 0.009$). When considered, together, *EGR2* mutations were found to associate with poor OS and PFS in this cohort (OS: Median = 41 vs. 71, $P = 0.0022$, PFS: Median = 13.5 vs. 25, $P = 0.015$) **(Figure 64A-C)**. Subclonal *EGR2* mutations were found to associate with poor OS and PFS in stratified pairwise Kaplan Meier survival analysis (OS: Median = 42.5 vs. 71, $P = 0.001$, PFS: Median = 8.5 vs. 25, $P = 0.025$), whilst clonal *EGR2* mutations were not found to significantly associate with PFS, and despite an infer median OS similar to subclonal *EGR2* mutated cases, only a trend towards an association with reduced OS (OS: Median = 44.5 vs. 71, $P = 0.086$, PFS: Median = 22 vs. 25, $P = 0.138$) **(Figure 64D-F)**. This suggests that it is the subclonal mutations which carry the risk of poor outcome in *EGR2,* and supports the idea that increasing the sensitivity of mutation capture is important to improve CLL patient prognostication.




C

| Survival | Univariate - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|---|
| OS | *EGR2* WT | 483 | 351 | 71 | 64 | 78 |
| | *EGR2*mut | 16 | 15 | 41 | 31 | 51 |
| PFS | *EGR2* WT | 483 | 443 | 25 | 22 | 28 |
| | *EGR2*mut | 16 | 16 | 13.5 | 10 | 34 |




F

| Survival | Pairwise - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|---|
| OS | *EGR2* WT | 483 | 351 | 71 | 64 | 78 |
| | Clonal *EGR2* | 10 | 9 | 44.5 | 31 | NA |
| | Subclonal *EGR2* | 6 | 6 | 42.5 | 12 | NA |
| PFS | *EGR2* WT | 483 | 443 | 25 | 22 | 28 |
| | Clonal *EGR2* | 10 | 10 | 22 | 13 | NA |
| | Subclonal *EGR2* | 6 | 6 | 8.5 | 4 | NA |

**Figure 64 Subclonal *EGR2* mutations significantly associate with poor OS and PFS in CLL4**

Univariate Kaplan Meier analysis in *EGR2* for OS **(A)** and PFS **(B),** with associated Kaplan Meier survival table in **(C)**. Pairwise Kaplan Meier analysis of clonal and subclonal mutations in *EGR2* for OS **(D)** and PFS **(E)**, with associated Kaplan Meier survival table in **(F)**.

### 4.4.4 The RNA splicing and metabolism genes *SF3B1* and *DDX3X* are associated with poor OS and PFS, respectively

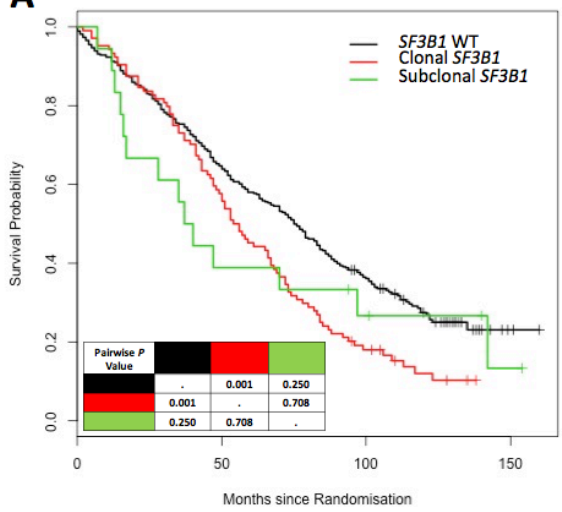
#### 4.4.4.1 Dichotomous impact on OS of clonal and subclonal mutations in *SF3B1* in U-CLL and M-CLL

Previously in CLL4, *SF3B1* mutations were identified as a marker of poor prognosis independent of IGHV mutational status (7). This was achieved by multivariate Cox PH modelling, where *SF3B1* mutated cases in the publication cohort were tested against: U-CLL, mutated *NOTCH1,* age, gender, Binet Stage (A vs. B/C), del(11q), *TP53*ab*,* and CLL4 treatment (CHL vs. FC/F). Recently, assessment of clonal and subclonal mutations for their impact on OS was conducted, finding neither clonal or subclonal events to associate with poor OS(159). However, stratified assessment of clonal and subclonal *SF3B1* mutations has not been reported in the context of a clinical trial, although the clonal evolution of *SF3B1* has been inferred(96). Furthermore, identifying the survival impact of both clonal and subclonal mutations within the U-CLL and M-CLL subgroups has not been conducted.

*SF3B1* mutated CLL and clonal *SF3B1* mutated CLL were found to have a significant impact on OS in all cases, with no observed impact on PFS (clonal *SF3B1*: OS: Median =54 vs. 75, $P = 0.0007$, PFS: Median = 23 vs. 25, $P = 0.179$. *SF3B1:* OS: Median = 53 vs. 75, $P = 0.0003$, PFS: Median = 22.5 vs. 26, $P = 0.09$). Subclonal mutations in *SF3B1* did not associate with poor OS or PFS in all cases (clonal *SF3B1*: OS: Median = 51.5 vs. 70, $P = 0.483$, PFS: Median = 20 vs. 24.5, $P = 0.892$). In pairwise Kaplan Meier survival analysis of all cases, only clonal *SF3B1* mutations were associated with poor OS (Clonal: Median = 55 vs. 75, $P = 0.001$. Subclonal: Median = 38.5 vs. 75, $P = 0.250$), with subclonal mutations having a reduced median OS, but potentially due to the subclonal curve crossing with wild type curve, it was found to not associate with poor OS **(Figure 59A&B)**.

Interestingly, dichotomous impact of *SF3B1* clonal and subclonal mutations on OS was observed in a stratified analysis of U-CLL and M-CLL. In this context, clonal *SF3B1* mutations predicted for poor OS in M-CLL (Median = 53 vs. 112, $P<0.0001$), whilst subclonal mutations did not (Median = 97 vs. 112, $P = 0.624$) **(Figure 65C&E)**. However, in U-CLL, the reverse was true, with subclonal mutations predicting for poor OS (Median = 16 vs. 52, $P = 0.041$), and clonal mutations no survival association (Median = 66 vs. 52, $P = 0.875$) **(Figure 65D&E)**.

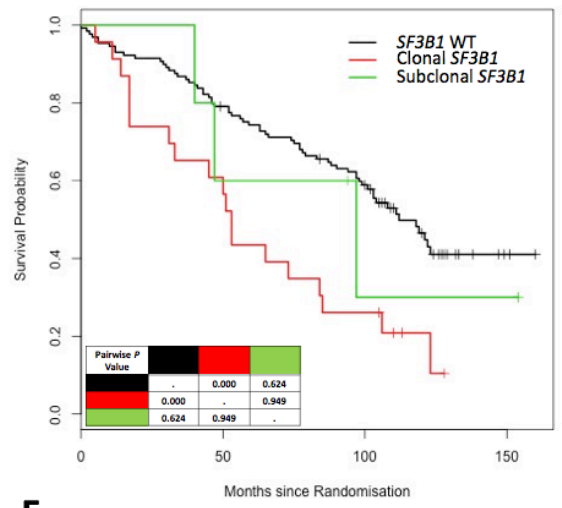
A



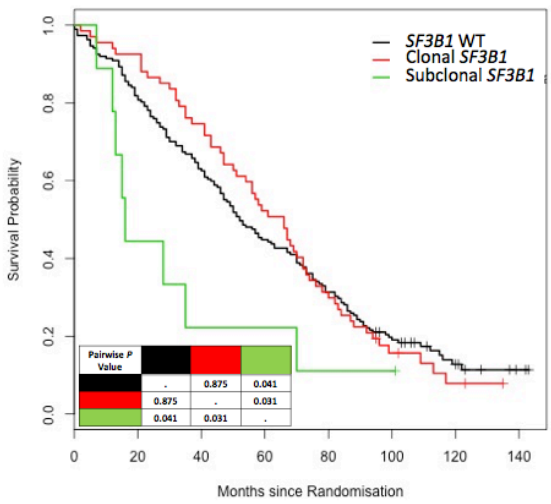

B

| Pairwise - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|
| *SF3B1* WT | 377 | 262 | 75 | 67 | 83 |
| Clonal *SF3B1* | 104 | 90 | 55 | 48 | 68 |
| Subclonal *SF3B1* | 18 | 14 | 38.5 | 17 | NA |

C




D




E

| *IGHV* Mutational Status | Pairwise OS - Kaplan Meier | Number of cases | Number of events | Median OS | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|---|
| M-CLL | *SF3B1* WT | 129 | 64 | 112 | 101 | NA |
| | Clonal *SF3B1* | 23 | 19 | 53 | 33 | 106 |
| | Subclonal *SF3B1* | 5 | 3 | 97 | 47 | NA |
| U-CLL | *SF3B1* WT | 187 | 157 | 52 | 47 | 67 |
| | Clonal *SF3B1* | 67 | 59 | 66 | 54 | 73 |
| | Subclonal *SF3B1* | 9 | 8 | 16 | 13 | NA |

F

| *IGHV* Mutational Status | OS Cox PH | Hazard Ratio | 0.95LCI | 0.95UCI | *P* Value | Model *P* Value |
|---|---|---|---|---|---|---|
| M-CLL | Clonal *SF3B1* | 2.4886 | 1.4869 | 4.165 | <0.001 | 0.001 |
| | Subclonal *SF3B1* | 1.7098 | 0.6231 | 4.692 | 0.297 | |
| U-CLL | Clonal *SF3B1* | 0.9537 | 0.7026 | 1.295 | 0.761 | 0.947 |
| | Subclonal *SF3B1* | 0.98513 | 0.5418 | 1.791 | 0.961 | |

**Figure 65 Distinct impact of clonal and subclonal *SF3B1* mutations on OS by *IGHV* mutational status**

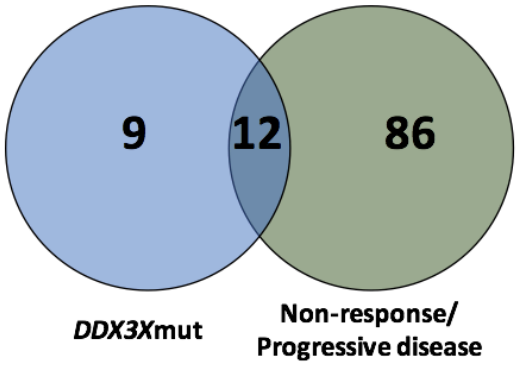
Pairwise Kaplan Meier analysis of clonal and subclonal *SF3B1* mutated cases **(A)**, with Kaplan Meier table data **(B)**. Pairwise Kaplan Meier analysis of clonal and subclonal *SF3B1* mutated cases in M-CLL **(C)** and U-CLL **(D)** for OS (black = wild type, red = clonal *SF3B1*, green = subclonal *SF3B1*). Cases with both clonal and subclonal mutations were placed in the clonal group. **E** Pairwise Kaplan Meier table of data from **(C)** and **(D)**. **F** Bivariate Cox PH model for OS, clonal *SF3B1* mutated cases vs. subclonal *SF3B1* mutated cases.

#### 4.4.4.2 *DDX3X* mutations are associated with non-response to CLL4 chemotherapy and progressive disease in CLL4, with reduced PFS

A member of the DEAD-box RNA helicase family, DEAD-box polypeptide 3 X-linked (*DDX3X*) has important roles in various levels of protein synthesis, including splicing, transcription, translation and RNA transport (203). *DDX3X* mutations in CLL are rare at a frequency of around 3%(31), however it has been proposed along with *SF3B1* and *NOTCH1* that *DDX3X* has a role in convergent evolution of CLL in response to therapy (204). However, substantial analysis on the impact of *DDX3X* mutations on survival, including stratification by clonal and subclonal variants, is withstanding.

*DDX3X* mutations were found to associate with non-response/progressive disease ($P<0.0001$) **(Figure 66A)**, with 57% of cases carrying a *DDX3X* mutation being in that subgroup (data not shown). This association was also found for clonal *DDX3X* mutations ($P = 0.008$), but did not associate with subclonal mutations ($P = 0.037$) after accounting for multiple hypothesis testing ($Q>P$, or $P<0.01$) **(Figure 66B)**. In univariate Kaplan Meier analysis, *DDX3X* mutations were found to associate with poor PFS (Median = 6 vs. 25.5, $P = 0.008$), but not for OS (Median = 68 vs. 70, $P = 0.551$, data not shown) **(Figure 66C&D)**. When stratified by clonal and subclonal mutations, subclonal mutations were found to associate with poor PFS in pairwise Kaplan Meier analysis (Median = 4 vs. 25.5, $P<0.0001$), with clonal mutations showing strong trend towards poor PFS (Median = 9.5 vs. 25.5, $P = 0.064$) **(Figure 66E&F)**.

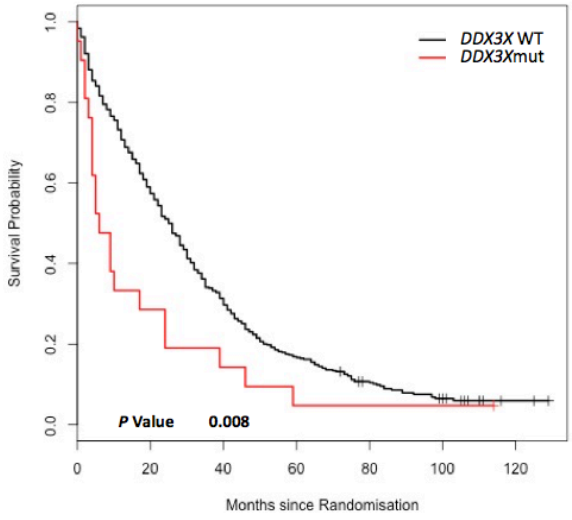
**A**




**B**

| Clonality of *DDX3X* mutated case | Response CR/NodPR | Response PR | Response NR/PD |
|---|---|---|---|
| Clonal | 4 | 2 | 8 |
| Subclonal | 0 | 1 | 4 |
| Both | 0 | 2 | 0 |

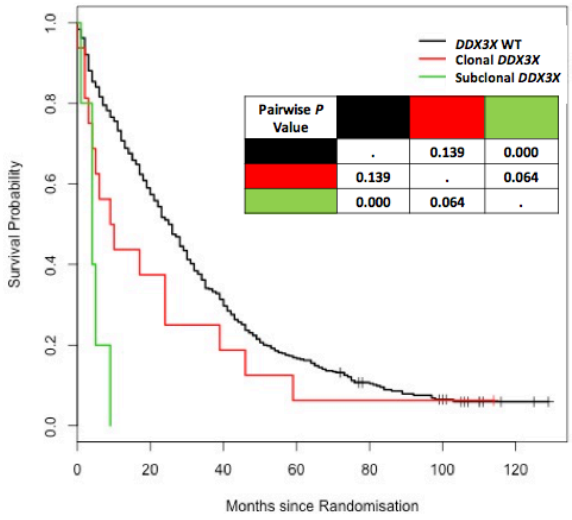
**C**




**D**

| Univariate - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|
| *DDX3X* WT | 478 | 439 | 25.5 | 23 | 28 |
| *DDX3X*mut | 21 | 20 | 6 | 4 | 24 |

**E**




**F**

| Pairwise - Kaplan Meier | Number of cases | Number of events | Median Survival | 0.95LCI | 0.95UCI |
|---|---|---|---|---|---|
| *DDX3X* WT | 478 | 439 | 25.5 | 23 | 28 |
| Clonal *DDX3X* | 16 | 15 | 9.5 | 4 | 46 |
| Subclonal *DDX3X* | 5 | 5 | 4 | 4 | NA |

**Figure 66 Impact of *DDX3X* mutations on PFS**

**A** Venn diagram of *DDX3X* and non-response/progressive disease cases. **B** Table showing the distribution of clonal and subclonal *DDX3X* mutations across all three response groups. **C** Univariate Kaplan Meier analysis of *DDX3X* on PFS. **D** Kaplan Meier analysis table associated with data from **(C)**. **E** Pairwise Kaplan Meier analysis of clonal and subclonal *DDX3X* mutated cases (black = wild type, red = clonal *DDX3X*mut, green = subclonal *DDX3X*mut) in PFS. **F** Pairwise Kaplan Meier analysis table associated with the data from **(E)**.

## 4.5 Multivariate Survival Analysis in CLL4

Univariate survival analysis provides the basis of whether cases with a variable may describe cases at risk from an event in comparison to all other cases without that variable. However, it fails to consider additional factors which may impact the relative importance of any one variable. Therefore, multivariate survival analysis accounting for these additional factors is required. Traditional multivariate survival analysis using Cox PH models was conducted on all variables significant for OS and PFS within the dataset, where backwards selection was used until all variables in the model met the significance criteria of $P<0.05$. This started with a model including all significant univariate variables, with stepwise removal of the least significant variable (highest *P* value), followed by modelling of the remaining variables. In addition, Random Survival Forests were implemented to assess all variables without any underlying assumption of the hazard of each variable.

### 4.5.1 *SF3B1* and *NRAS* found to independently predict poor OS in CLL4

In the multivariate analysis for OS, two models were created by backwards selection: significant univariate variables for all mutations in each gene, with all significant univariate clinical features, and significant univariate clonal and subclonal variables with all significant univariate clinical features. When all mutations per gene were considered together, *SF3B1* (HR = 1.43, 95%CI = 1.00-2,04, *P* = 0.047) and *NRAS* (HR = 3.80, 95%CI = 1.59-9.09, *P* = 0.003) were found to independently predict poor OS in a model with: M-CLL, *TP53*ab, age at diagnosis, Binet stage C, del(13q) only, CD38 positivity, β-2-M positivity, and non-response/progressive disease (see HRs, 95%CIs and *P* values in table) **(Table 9)**. Interestingly, Binet stage C in this model predicts for an improved prognosis, which suggests that after identifying poor risk patients with other factors, Binet stage C cases without these factors have a better outcome. In the multivariate model stratified by clonal and subclonal mutations, no TruSeq variable was retained in the final model, with all significant clinical features from the first model found in the second model.

**Table 9 Multivariate Cox PH model for OS in CLL4**

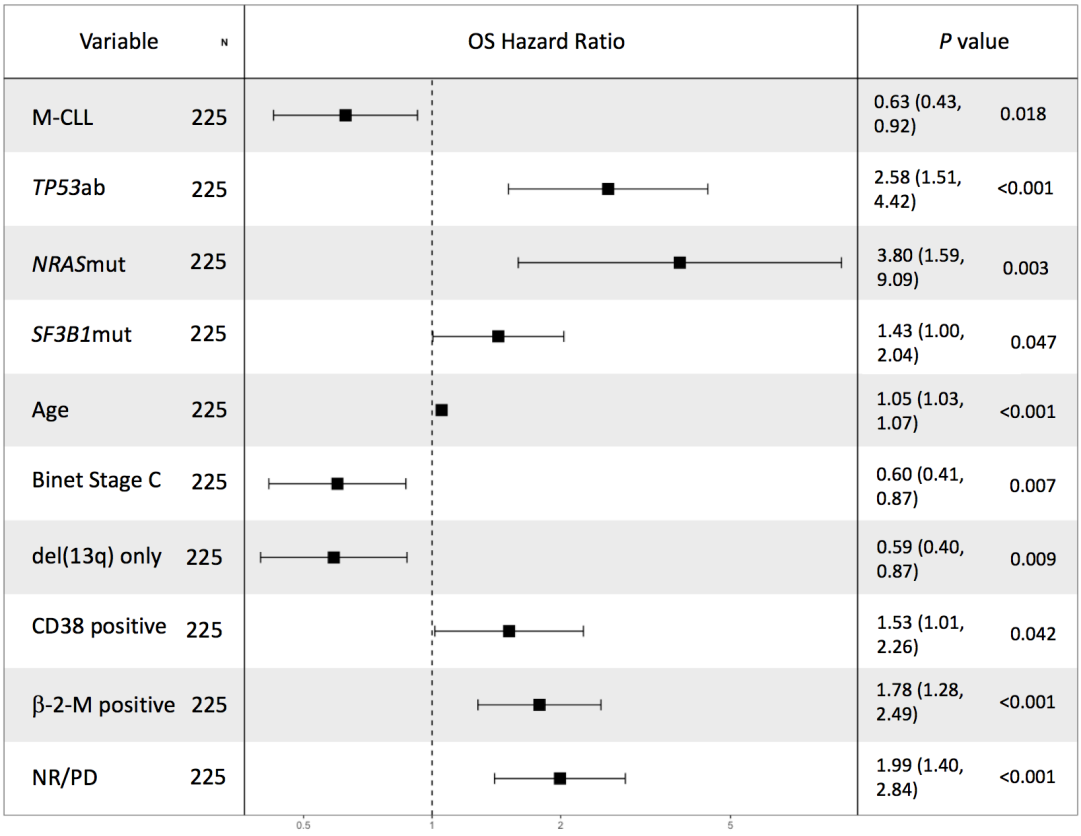



Table of multivariate Cox PH model for OS in CLL4 generated using backwards selection ($P$<0.05). Each row denotes a variable, with the columns (left to right) representing: variable name, number of cases in final analysis after accounting for missing data, forest plot show the HR with 95% CI (vertical dashed line represents HR = 1, *x* axis of HR values at the bottom of the column), HR and 95%CI values, followed by the *P* value of each variable.

### 4.5.2 Co-mutated/deleted del(11q) and *BIRC3* independently predicts poor PFS in CLL4

As described for multivariate Cox PH model analysis for OS, two models were created for PFS, one including all mutations in each gene, and another where clonal and subclonal mutations in univariate significant genes were assessed. In both models, co-mutated/deleted del(11q) and *BIRC3* (del(11q) + *BIRC3*) was found to significantly associate with reduced PFS (HR = 1.82, 95%CI = 1.01-3.29, *P* = 0.046), in a model with: M-CLL, *TP53*ab, FC treatment, complete response/partial nodular response, and non-response/progressive disease (see HRs, 95%CIs and *P* values in table) **(Table 10)**. This supports the pairwise analysis presented **4.4.3.2**, where del(11q) + *BIRC3* was found to associate with poor PFS in the context of del(11q) only cases and biallelic *ATM* cases.

**Table 10 Multivariate Cox PH model for PFS in CLL4**

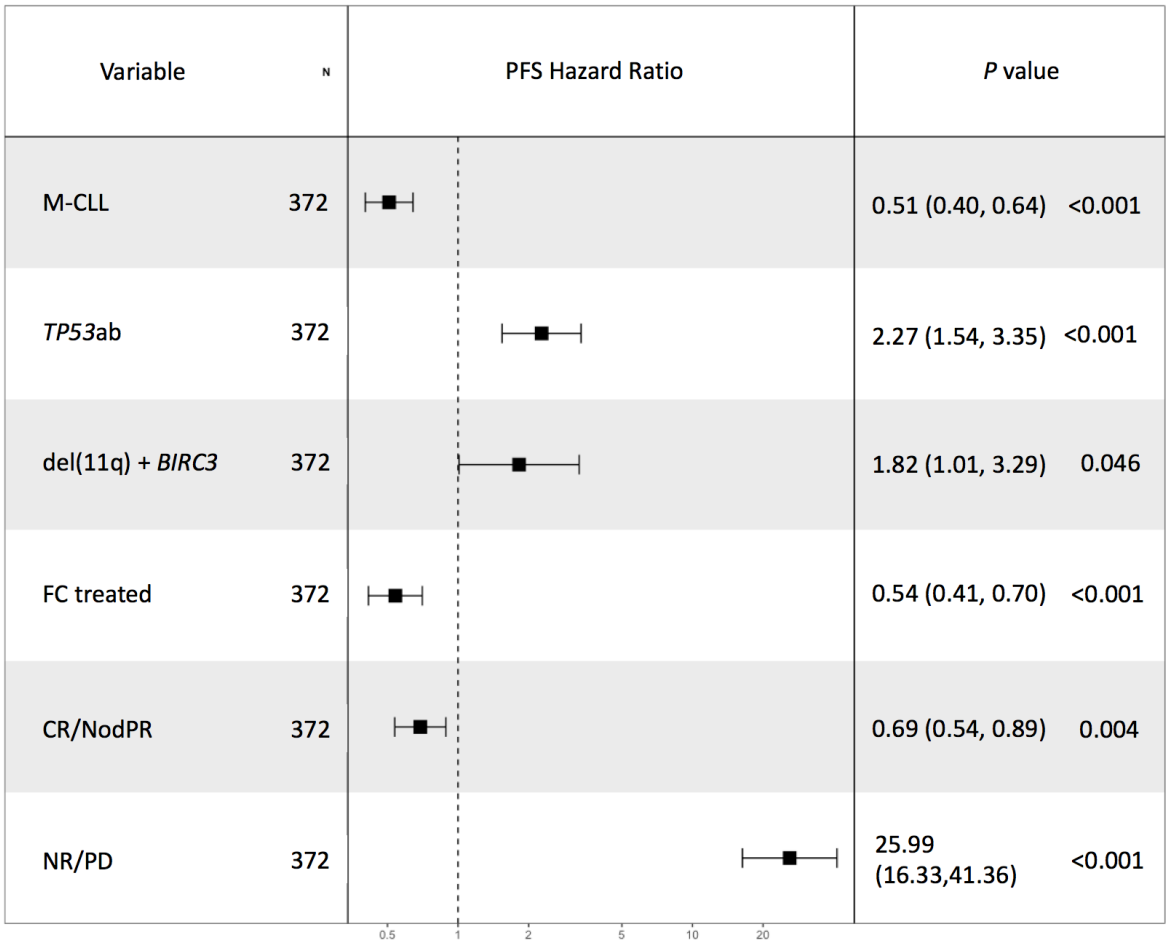



Table of multivariate Cox PH model for PFS in CLL4 generated using backwards selection (*P*<0.05). Each row denotes a variable, with the columns (left to right) representing: variable name, number of cases in final analysis after accounting for missing data, forest plot show the HR with 95% CI (vertical dashed line represents HR = 1, *x* axis of HR values at the bottom of the column), HR and 95%CI values, followed by the *P* value of each variable.

### 4.5.3 Random Survival Forests analysis in CLL4

To make this analysis comparable to the multivariate Cox PH analysis presented above, two RSF models were generated. The first was considering all mutations per gene, if each gene had more than 5 events. These were then placed in a dataframe with all clinical features. For the second, each gene was stratified by clonality (i.e. whether clonal or subclonal), removing variables with less than 5 events, and placed in a dataframe with all clinical features. Following which, both dataframes were used to generate RSF models for OS and PFS. Once a RSF has been built, data can be extracted in several ways to understand more about the impact the variables have on the response variable. Ranking variables based on Variable Importance (VIMP) identifies variables which could be strong predictors. VIMP functions by calculating the prediction error (or: Out of Bag rate [OOB]) for the observed variable, then randomly permutes the variable to test whether that impacts its effect on the response variable (189) . Positive VIMP scores are indicative of a variable being resistant to permutation, whilst those with a negative VIMP score are not. Finally, the minimal depth of each variable can be calculated, to work out which variables partition the largest amount of the dataset. Minimal depth is calculated by working out the average distance from the root node per variable across the entire forest. The VIMP score and the minimal depth can then be interpreted together to identify variables which predict risk, are relatively resistant to permutation, and account for the partitioning of a high number of cases within the dataset.

#### 4.5.3.1 *SF3B1* mutated CLL identified as an important feature in OS from RSF model in CLL4

A RSF model was created (depth = 1000 trees), using the Log Rank score and the OS survival censor as the split rule. From this OS RSF model, 12 features were found to have high VIMP and low minimal depth scores **(Figure 67A)**. From this, *SF3B1* (VIMP: Rank = 53/54, Score = -0.1067, Minimal Depth: Rank = 6/54, Score = 6.998) was the only gene to feature in the top variable list, with top variables according to the minimal depth vs. VIMP model being: Ten year survivors, age at diagnosis, β-2-M positivity, del(13q) events, *TP53*ab, CD38 positivity, M-CLL, complete response/partial nodular response, ZAP70 expression, non-response/progressive disease, and gender (male) **(Figure 67B)**. Interestingly, the VIMP rank for *SF3B1* was extremely low, but since it had a low minimal depth, therefore was found near the root node of decision trees during forest building, it was found to have a strong impact on OS.

**A**

| Variable | Minimal Depth | VIMP |
|---|---|---|
| Ten year survivors | 2.742 | 0.044 |
| Age at diagnosis | 3.075 | 0.016 |
| β-2-M positive | 5.877 | 0.004 |
| del(13q) | 6.645 | 0.009 |
| *TP53*ab | 6.727 | 0.011 |
| *SF3B1* | 6.998 | -0.005 |
| CD38 positive | 7.262 | -0.005 |
| M-CLL | 7.428 | 0 |
| CR/NodPR | 7.45 | 0.002 |
| ZAP70 positive | 7.474 | -0.003 |
| NR/PD | 7.507 | 0.004 |

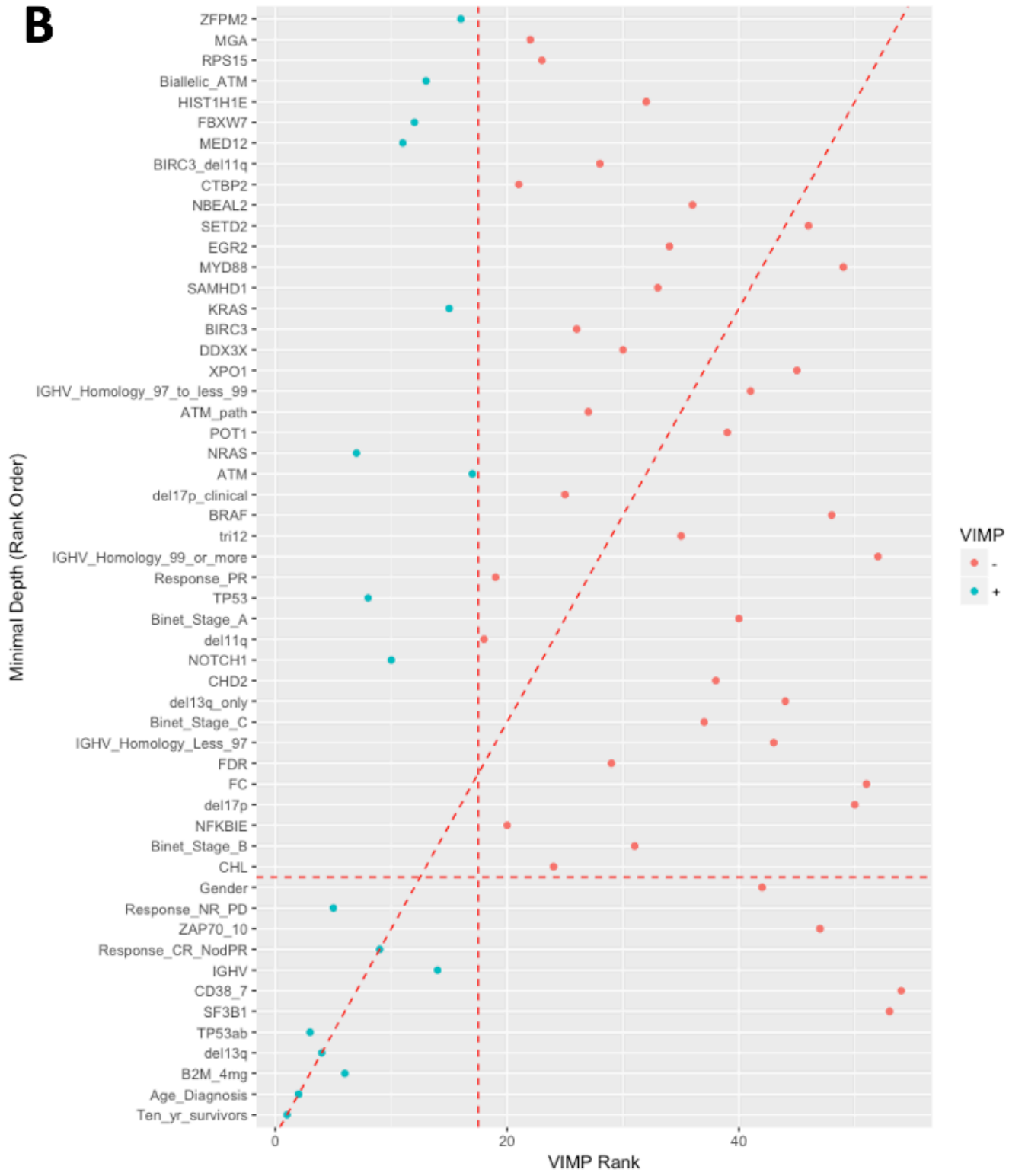



**Figure 67 RSF model of OS in CLL4 - Minimal depth vs. VIMP**

**A** Table of top variables based on minimal depth vs. VIMP analysis, ranked in order of smallest minimal depth. **B** XY scatter plot representing the minimal depth score vs. VIMP rank from the RSF model. Variables are ordered from bottom to top on the *y* axis by minimal depth ranking, and from left to right on the *x* axis for VIMP rank. Points on the diagonal red dashed line represent variables where minimal depth and VIMP agree, with points below the horizontal red dashed line representing variables listed in the top variables by minimal depth, and points to the left of the vertical red line representing variables with a positive VIMP score

#### 4.5.3.2 No putative recurrently mutated genes from TruSeq study identified as an important feature in PFS from RSF model in CLL4

A RSF model was created (depth = 1000 trees), using the Log Rank score and the PFS survival censor as the split rule. From this PFS RSF model, 8 features were found to have high VIMP and low minimal depth scores **(Figure 68)**. No recurrently mutated gene, nor clonal or subclonal mutated subgroup, were found in the top variable list from the minimal depth vs. VIMP model. The features found to impact PFS from this RSF model were: non-response/progressive disease, age at diagnosis, ten year survivors, ZAP70 expression, FC treatment, CHL treatment, complete response/partial nodular response, and M-CLL.

**A**

| Variable | Minimal Depth | VIMP |
|---|---|---|
| NR/PD | 2.934 | 0.038 |
| Age at Diagnosis | 3.548 | 0.003 |
| Ten year survivors | 4.746 | 0.005 |
| ZAP70 positive | 6.017 | 0 |
| FC treated | 6.089 | 0.001 |
| CHL treated | 6.229 | 0.001 |
| CR/NodPR | 6.239 | 0.01 |
| M-CLL | 6.267 | 0 |

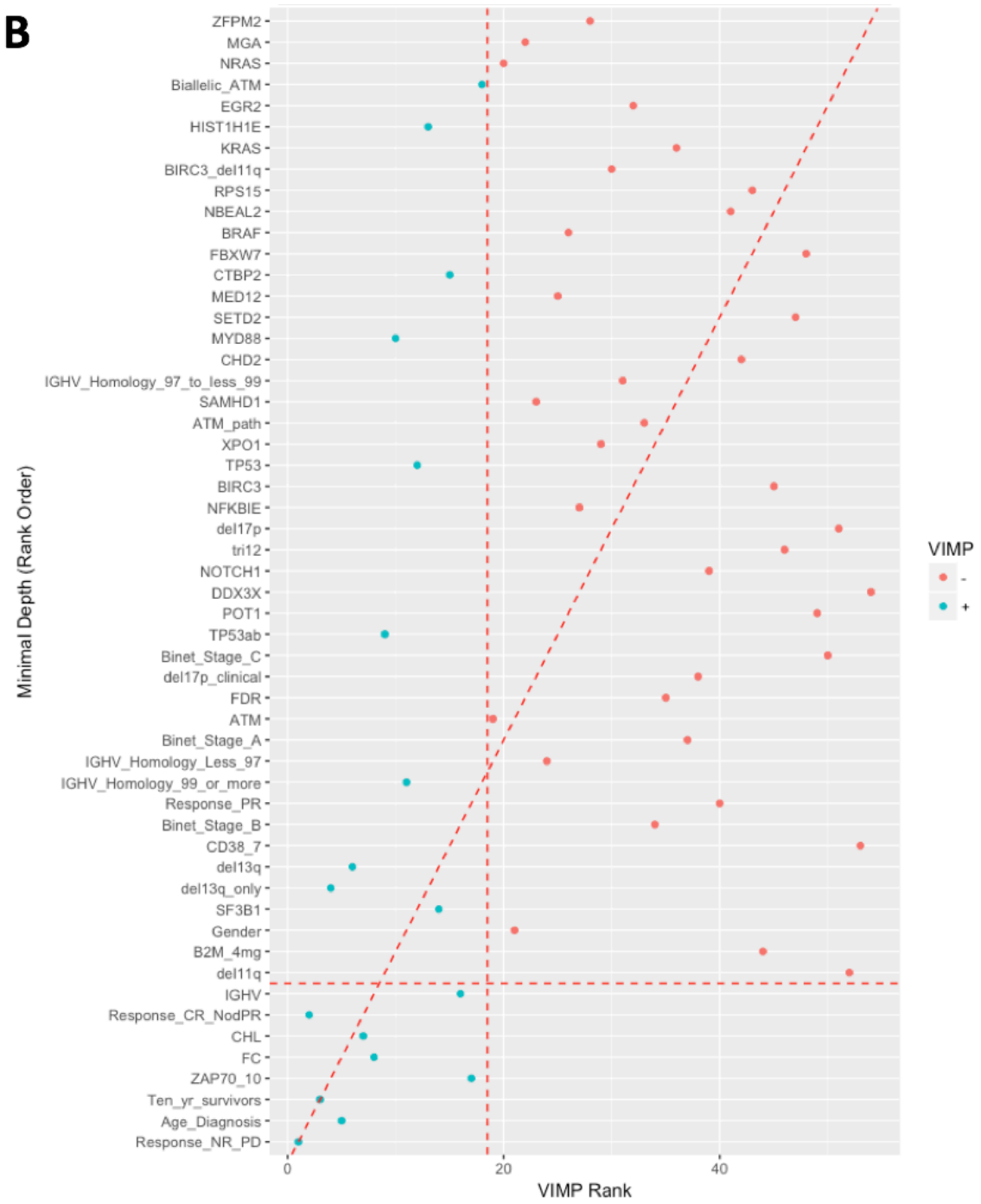



**Figure 68 RSF model of PFS in CLL4 - Minimal depth vs. VIMP**

**A** Table of top variables based on minimal depth vs. VIMP analysis, ranked in order of smallest minimal depth. **B** XY scatter plot representing the minimal depth score vs. VIMP rank from the RSF model. Variables are ordered from bottom to top on the *y* axis by minimal depth ranking, and from left to right on the *x* axis for VIMP rank. Points on the diagonal red dashed line represent variables where minimal depth and VIMP agree, with points below the horizontal red dashed line representing variables listed in the top variables by minimal depth, and points to the left of the vertical red line representing variables with a positive VIMP score

## 4.6 Discussion and Future Directions

Survival association analysis of genetic, biological, and clinical features of disease are important to identify subgroups of patients that may require specific clinical management. In CLL, this has already been realised for biological and clinical features, with: *IGHV* mutational status, ZAP70 expression, CD38 expression, β-2-M expression, Binet stage, as well as others tested at diagnosis to stratify patients (30,34–36,87). However, except for the recurrent CNVs in CLL (del(13q) only, del(17p), del(11q), +(12)), the only gene to be screened for is *TP53* (166). Furthermore, international authorities and committees only recommend the analysis of deletions and mutations of *TP53* when therapy is required. Building evidence suggests that the inclusion of other genes should be explored, but many have not been assessed together in the context of a clinical trial at the higher resolution permitted by NGS approaches.

In this chapter, analysis of the clinical importance of gene mutations in the context of the LRF CLL4 clinical trial was undertaken, identifying the importance of: a) comparative analysis to previous publications, b) assessing survival at multiple time points, and c) understanding how survival analysis functions, handle right-censored survival data. Following which, comprehensive univariate survival analysis of all 25 genes, stratified by clonal and subclonal mutations, as well as all available CLL4 clinical features, was conducted. This analysis identified that 7 genes were associated with poor OS (Pathogenic *ATM, BRAF, EGR2, KRAS, NRAS, SF3B1,* and *TP53*), and 4 genes associated with poor PFS (*DDX3X, EGR2, KRAS, TP53*), either within clonal or subclonal subsets, or for all mutations in each gene.

In addition, the impact of IGHV and *TP53* status on the prognostic important of these new mutations was also ascertained, to provide initial insight into the potential clinical utility of these mutations. These analyses were extended further with multivariate Cox PH modelling using backwards selection, that identified *SF3B1* and *NRAS* to independently associate with poor OS, in a model with significant univariate genetic and clinical features. In a multivariate Cox PH model for PFS using backwards selection, del(11q) with concomitant *BIRC3* mutation was found to associate with poor PFS, in a similarly comprehensive model including established confounders. Application of supervised machine learning to OS and PFS in the context of all available variables identified *SF3B1* as the only gene to have low minimal depth (splits decision trees near the root node, therefore strong predictor of survival event), but was ranked low for VIMP (high rate of permutation, i.e. randomly change *SF3B1* mutants to wild type for model testing reduces its impact on OS) for OS, whilst no gene from the TruSeq data was identified to be in the top-ranking variables in terms of minimal depth or VIMP for PFS.

### 4.6.1 Sensitivity Specificity analysis

When assessing the impact of recurrently mutated genes on survival, especially in the context of using a new and more sensitive technique, comparative analysis to previous data (if available) produced with more traditional technology, offers a chance to test technique performance. Sensitivity and specificity analysis can be applied to assess these differences, comparing the ability of each test to capture survival events. Although the TruSeq technology overall had higher sensitivity, and captured more survival events than the previous studies for *TP53, SF3B1,* and *NOTCH1* (7,167)*,* it did not exhibit increased specificity compared to the previous study techniques, techniques that included a molecular screening technique (e.g. HRM) followed by confirmatory Sanger sequencing. This suggests that whilst capturing mutations below the technical sensitivity of Sanger sequencing may identify additional patients destined to have adverse survival events, there appear to be cases with subclonal mutations that exhibit protracted survival. This may reflect the fact that those more clonal mutations detectable by more traditional technologies have a more significant impact on the biology of the tumour cells, and therefore a more aggressive disease. Furthermore, mutations at low frequencies within a tumour cell population may have specific impacts on survival depending on the gene in question, the clonal cell fraction of the mutation as well as features of the micro-environment where cell proliferation occurs.

It is important to note that sensitivity specificity analysis only accounts for whether each technique captures a survival event, with no assessment of the timing of this event. In the context of identifying cases which may be treatment resistant, through the clonal evolution of a subclonal mutation, sensitivity specificity analysis cannot account for these features. Therefore, although targeted re-sequencing may be less specific than standard technologies of mutation capture in terms of calling survival events, it allows the identification of mutated cases that might benefit from being treated in a way that is based on their biology.

### 4.6.2 Time to event survival data time points, time-dependent covariates and SurvSplit functions

Understanding the distribution and impact of clinical variables over time is important in the context of survival analysis. This is particularly relevant in a disease of the elderly population, such as CLL, as assessing associations with survival is likely to be partially confounded by events (e.g. deaths) not causatively linked with the disease. Furthermore, addressing the question of variable proportionality is important to enable appropriate interpretation of significant survival associations, and permits the identification of the most informative timescale for the assessment

of disease-related survival. Assessment of Schoenfeld residuals has been reported in survival association studies in CLL(205), but the PhD candidate's knowledge at the time of writing SurvSplit functions have not been used in CLL. Schoenfeld residuals allow the assessment of proportionality of HRs over time, rather than assuming the HR is constant. This means that for non-proportional variables, the proportional HR considered in a multivariate model might confound the significance of it or other variables, limiting the validity of the model. SurvSplit represents analysing time to event data in a different manner to the traditional analysis. Significant associations in the secondary split (data after the median survival time) were observed for variables found to not impact survival in the traditional analysis, therefore caution should be taken with its implementation. This may be because they are impacted by reduced power, since after the split the Cox model only assesses cases yet to receive the event (i.e. die or progress). To simplify the analysis in this chapter, the SurvSplit function was not used in the final univariate and multivariate analysis. However, implementation of this type of analysis in future work could enrich the available data, if it is statistically powered to do so.

### 4.6.3 Univariate survival analysis

Univariate survival analysis using Kaplan Meier curves and Cox PH models is the first step used by many publications interested in the survival impact of genetic, biological, and clinical features in disease. Both allow easy interpretation of the importance of variables on survival. In CLL, these approaches are a standard way of presenting data on putative biomarkers of disease. Furthermore, they are often used to show pairwise associations, plotting the survival curve of various subgroups of the disease to identify poor and good risk groups of patients in the context of known biomarkers.

In this study, subclonal *TP53* mutations were not found to associate with a similar survival outcome to clonal *TP53* mutations, a result which is not concordant with the published literature on subclonal *TP53* mutations in CLL (8,159). Given the discordance between the data presented in this thesis, and several published studies, it was important to be confident that our *TP53* mutations were real. Therefore, orthogonal sequencing of all subclonal *TP53* mutations was conducted to guarantee only the selection of high quality variants, an analysis that suggests that the *TP53* mutations identified in CLL4 were real. Consequently, the associations observed here are likely to be associated with the fact that sample was performed at treatment requirement, the general composition of the cohort or the nature of the follow-up data. However, none of the published studies conducted analysis in the context of del(17p) events stratified by clonal and subclonal *TP53* mutations. However, in **Chapter 3** it was observed that a portion of exon 6 was not captured in cases sequenced using design 2 of the CLL4 TruSeq panel, therefore this could

represent a loss of subclonal *TP53* mutations that could impact outcome. Target screening of this region in cases currently without mutation data would contribute to the resolution of this question.

The presence of a *BIRC3* mutation was not found to associate with reduced PFS or OS in an analysis of all cases employing univariate survival analysis. However, in pairwise analysis of *BIRC3* mutations and pathogenic *ATM* mutations stratified by del(11q), co-mutated/deleted del(11q) and *BIRC3* was found to associate with poor OS and PFS. This analysis offers significant insight into this question of whether *BIRC3* or *ATM* is the most important gene in co-occurrence with del(11q) events, since the published work either had small numbers of *BIRC3* mutated cases (61), or excluded either *ATM* or *BIRC3* from a combined analysis altogether in the context of del(11q) (49,159). Key questions regarding the applicability of these findings are associated with the decision-making process regarding inclusion of *ATM* variants into the "pathogenic *ATM*" variable. The most pertinent of these is the definition of what makes an *ATM* mutation pathogenic or not. Using a strategy based on the Nadeu *et al.* (2016) study(159), variants were included if they had previously appeared in the LOVD *ATM* database, if they had been tested *in vitro* as having ATM dysfunction, as well as the exclusion of potential germline SNPs which have shown not to impact *ATM* function. This analysis has been supported by Professor Tanja Stankovic, who provided annotation to the pathogenicity of the CLL4 TruSeq *ATM* variant list (personal communication). Reviewing the biallelic *ATM* variants, and discussing these with Professor Stankovic will be important in realising the impact of survival on biallelic *ATM* cases. Regarding "biallelic" *BIRC3* cases, they were not classified as such, due to not having the appropriate copy number data to demonstrate that *BIRC3* was deleted. Screening all del(11q) cases, as well as all *BIRC3* mutated cases, for deletions of *BIRC3* using shallow WGS, for example, would clarify the survival association of these cases with OS and PFS.

*KRAS, NRAS,* and *BRAF* mutations in CLL have been shown to be mutated at low frequency, and associated with poor risk CLL features, but limited survival analysis has been conducted (57,96,162,195). Here, survival impact in all three genes was observed for OS, with an analysis of mutually exclusive mutated cases unearthing that both *KRAS* and *BRAF* only mutated CLL have a poor outcome. However, this analysis did not consider mutations in other genes on the panel. Across all mutually exclusive *KRAS, NRAS,* and *BRAF* cases, only 1/53 did not have a mutation in another gene (*BRAF*), with all other cases having a range of 2-5 mutated genes per case. It could be an important additional feature to remove cases if they had mutations in other genes found to impact prognosis in this cohort, improving the potential applications of this observation, however, this could impact the power to statistical support this observation.

Mutated *EGR2* has been observed to lead to poor outcome in CLL, with data from a subset of this chapter contributing to this work (11). This chapter expands on the impact of these mutation in OS and PFS, providing data that subclonal mutations predict for poor OS and PFS. However, *EGR2* did not infer a poor outcome in the context of multivariate Cox PH model for either survival category, but an analysis in the publication for OS found *EGR2* to be independent of: age, gender, Binet stage, IGHV mutational status, del(11q), and *TP53*ab. This is likely to be because of one of two factors. Firstly, in the publication the variables going into the model for stepwise selection processing were only the *EGR2* mutated variable, followed by the clinical variables available for CLL4. As described in the multivariate section below, the interpretation of multivariate models should be done so with caution, since the selection criteria used has an impact on the final model. Secondly, this chapter presents multivariate models where many other interacting variables were included, which could reduce the impact of *EGR2* in this context.

*SF3B1* mutations have has been shown to be an independent poor prognostic marker across multiple institutional and clinical trial cohorts (7,10,69,96). Here, mutated *SF3B1* associated with poor OS, as well as having a dichotomous inferior impact on survival for both clonal and subclonal mutations in U-CLL and M-CLL. This chapter confirms the observation found in Oscier *et al.* (2012), that *SF3B1* has no association with reduced PFS in the CLL4 trial cohort, even when subclonal mutations were included. However, the impact of clonal and subclonal *SF3B1* mutations on survival is not comparable to the only study to have presented such data, Nadeu *et al.* (2016). In that study, no association for either clonal or subclonal *SF3B1* mutations was found for OS, the only survival comparison that can be made with the survival data presented in this chapter. This could be due to cohort characteristic differences, since the Nadeu *et al*. study is not based on a clinical trial of patients requiring therapy. This observation can only be confirmed with further studies in other large clinical trial cohorts, where over time the impact of clonal and subclonal *SF3B1* mutations can be assessed. *SF3B1* was found alongside *NRAS* to retain significance in multivariate Cox PH modelling for OS, as well as ranked as an important variable in RSF model analysis, the only gene to do so. This further supports the continuing evidence of screening for *SF3B1* mutations at diagnosis in CLL, as in the context of all the available CLL4 clinical biomarkers, it still predicted for poor OS.

A significant survival association in the context of *DDX3X* mutated cases has not been reported, although their prevalence and impact in relation to clonal evolution in response to therapy has been assessed (96,204). In this chapter, *DDX3X* mutations were found to associate with non-response/progressive disease, and significantly associated with poor PFS in all mutations, as well as in subclonal mutations. Although these survival associations were not observed in multivariate Cox PH modelling, it offers a potential new gene to screen for to predict poor PFS in response to

chemotherapy, and supports the use of sensitive NGS techniques to identify subclonal mutations. Molecular validation of these variants would be particularly pertinent. Furthermore, requesting *DDX3X* mutation data from other cohorts from collaborators in Europe to strengthen the hypothesis that *DDX3X* mutated cases have poor response to therapy would be a useful approach.

Although identified as significant predictors of poor OS in univariate survival analysis, *NOTCH1* and *MYD88* were excluded from further analysis due to being statistically underpowered to make these associations. This was not expected, since both *NOTCH1* and *MYD88* have been shown to be strong predictors of poor, and good OS, respectively (7,33,206). Although in the case of *MYD88,* it is important to note that analysis in the context of M-CLL, of which this subgroup is where most mutations occur, confounds its positive impact on outcome (207). As there is no missing clinical data, it is not the case that the study does not have a large enough sample size, therefore the problem likely pertains to the effect size of each variable. For *NOTCH1,* there were 57 mutated cases, with a HR of 1.35, suggesting an increased risk of death of 35% ($P$ = 0.0425). In comparison, Binet stage C cases had a univariate HR of 1.37 ($P$ = 0.005), and was powered to make that survival association (84% powered). However, the Binet stage C variable had nearly 3 times more cases (n = 156). *MYD88* was also found to be underpowered (HR = 0.47, $P$ = 0.008, n = 10), whilst other variables with a similar HR were appropriately powered (M-CLL: HR = .42, $P$<0.0001, n = 157). It is not the case that power is dominated by the number of cases in the group to be tested, as it is relative to the effect size observed. This is further explained by variables with small numbers, but large effect sizes, as is seen for *NRAS* (HR = 2.43, $P$ = 0.01, n = 10). This shows the importance of power calculations in survival analysis, as well as understanding the required numbers to be statistically confident in the associations observed.

Interestingly, no survival association with poor OS or PFS was observed for *RPS15* in this cohort, a gene which a sub-cohort of this chapter provided data to a pan-European study (n = 329), finding that *RPS15* mutated cases that were *TP53* wild type conferred a similar poor prognosis (14). However, for CLL4 in this publication only co-occurrence analysis was presented, showing a weak association of *RPS15* cases co-occurring with *TP53*ab (3 *RPS15* mutated cases out of 26 *TP53*ab cases, "$P$<0.05" reported). In this chapter, co-occurrence analysis did find *RPS15* mutated cases to be mutually exclusive of *IGHV* homology to germline (<97%) ($P$ = 0.005), with most cases being in the IGHV homology to germline (>99%) group (15/16, $P$ = 0.005). However, this was not discussed in the main chapter because no survival associations were observed. It is important to note that for this publication, the PhD candidate screened the cohort for *RPS15* mutations and created the dataframe with associated clinical data, but did not conduct the survival of co-occurrence analysis for publication.

### 4.6.4 Multivariate survival analysis

#### 4.6.4.1 Cox PH models

Multivariate survival analysis using Cox PH is a major feature of all publications addressing the impact of genetic, biological, and clinical features on the survival of patients with disease, especially in publications where the patients studied were enrolled onto a clinical trial. In CLL, they are readily used to show the importance of the putative biomarker being presented, and contribute evidence to support its screening at diagnosis, to either predict poor prognosis, or to identify an optimum clinical management strategy. However, the construction process of multivariate Cox PH models is complicated, with forward selection and backward selection methods available. Backwards selection was chosen as the method to be used in this chapter, due to its ease of application in comparison to the other methods. Forward selection requires exploration of every variable combination, including complicated decisions such as which order to include variables into the model. Whereas by comparison, including all variables significant at the univariate level into a model, then removing the variable with the highest *P* value each time is considerably simpler. In the iterative process of backward selection, the *P* value of each variable in each model changes after each stepwise removal of the last least significant variable, with some variables always remaining below the $P<0.05$ cut off, whilst others "become" and "lose" significance during backwards selection. This suggests that this process may be error prone, or at the very least could mask the true value of the *P* value presented in the final model. Therefore, caution must be taken when interpreting the outcome of multivariate Cox PH models.

#### 4.6.4.2 RSF models

Supervised machine learning in the context of survival analysis offers the potential of assessing the impact of experimental variables without any biased assumptions on the data presented. Although not used extensively, various types of RSF models have been presented in the cancer genomics literature(188,196), but always alongside a multivariate Cox PH model. In CLL, two papers have applied different aspects of RSF analysis methodology, both coming from the Rossi CLL group. The first was a prognostication model building on the work of Döhner *et al.* (2000)(30), where they integrated genetic data from: *TP53, SF3B1, NOTCH1,* and *BIRC3* with the classical cytogenetic features del(11q), del(13q) only, +(12), and del(17p)(208). To do this, the implemented recursive partitioning, an approach which aims to identify the variable which describes most of the data, following which it recursively tests the subsequent data subgroups to identify whether there are other significant groups within each partition. This stratified patients into the following groups: *TP53* disrupted/*BIRC3* disrupted, *SF3B1* mutated and/or *NOTCH1* mutated and/or del(11q), and a group not defined by either of the previous groups, but were

further stratified into +(12) cases and del(13q) only cases. They then conducted pairwise Kaplan Meier analysis of these groups, showing that they offered distinct prognostic utility. This approach was attempted here, however due to the sheer number of variables, a clear biologically and clinically sensible stratification was not apparent. The second was an analysis of the minimal depth of a RSF model in the study of CD49d expression in CLL (209). Univariate and multivariate survival analysis identified expression of CD49d as a predictive biomarker, and this was confirmed using the minimal depth scores extracted from a RSF model. However, the data included only the variables presented in the multivariate Cox PH model, therefore it is unknown whether all available variables were considered in the RSF model.

In this chapter, a feature identified within an extensive publication of how to analyse RSF data was selected, minimal depth vs. VIMP. This is an important feature to assess, as it considers both the strength of the variable to withstand permutation (i.e. strong variables being likely true positives), whilst also residing close to the root node across the forest (i.e. splits the decision tree early in the process, therefore a strong predictor). The Ehrlinger *et al.* (2016) (210) publication goes on to explore the various methods that can be used to gain insight into RSF models, however this was outside of the remit of this project, but would be a key feature of future work in understanding the importance of using machine learning tools in survival, as well as exploring the additional data held within each forest.

### 4.6.5 Future directions

1. Molecular validation of all subclonal mutations in genes linked with survival from univariate survival analysis
2. Additional screening of second randomisation cases for all subclonal mutations in all genes found to have an impact in univariate survival analysis, to assess whether clonal expansion occurred in these cases
3. Further assessment of time dependent covariates and SurvSplit, especially in the context of multivariate Cox PH modelling
4. Assessment of the possibility to generate a 5-year and 10-year time to event survival dataset, rather than using time to event data from a specific time point
5. Further analysis of the importance of RSF models, including features such as response date, time to first treatment, and the assessment of continuous variables, such as age
6. Optimisation of R code and plot generation, to streamline the process of data analysis

### 4.6.6 Conclusions

This chapter, and the one preceding it, have identified that *SF3B1* mutations are an important feature of CLL, in terms of high recurrence in the disease (26%), and in terms of being associated with poor OS in univariate survival, multivariate survival, and machine learning survival models. Therefore, *SF3B1* presents itself as an interesting target for *in vitro* analysis using Small Molecule Inhibitors (SMIs), to be explored in **Chapter 5**:.

# Chapter 5: Elucidating the clinical utility and mechanism of action of splicing inhibition in CLL

Mutated *SF3B1* in CLL has been shown by this work **(Chapter 4:),** and previously by others (7) to be an independent marker of poor prognosis, with others showing its negative impact on time-to-first treatment (TTFT) (159), OS (10,47), splicing (143,211) and clonal evolution (75,96,159,173) in a range of institutional and clinical trial cohorts. Since the identification of mutated *SF3B1* in CLL by Wang *et al.* (2011) (31), targeted inhibition of the spliceosome using small-molecule inhibitors (SMIs) has been hypothesised for CLL therapy(212). This has led to the implementation of multiple spliceosome inhibitors **(Figure 69)**, which have been derived from three major sources. Derivatives of the natural products pladienolide B and FR901464 namely Spliceostatin A (213), Meayamycin B (214), Sudemycins (215), and E7107 (216), target the SF3b complex, and ablate SF3b function, as well as inducing nuclear retention of pre-mRNA consequently promoting alternative splicing. Additionally, a benzothiazole compound (TG003 (217)) led to the inhibition of splicing via abrogation of phosphorylation activities. These phosphorylation sites are controlled by Cdc2-like kinase (Clk) family proteins, including serine/arginine (SR) rich proteins, which are important sites for cross-exon interactions that locate on exonic splicing enhancers (ESE) (217). Finally, GEX1A derived from *Streptomyces sp* (218,219) directly binds to SF3B1 and like FR901464 and Pladienolide compounds, induces cell cycle arrest via p27 alternative splicing in HeLa cells (218).

Spliceosome inhibition in CLL was first described using Sudemycins, where they showed sensitivity *in vitro* in samples with an *SF3B1* mutation*,* and importantly synergised with Ibrutinib in an *in vivo* CLL xenograft model (220). The spliceosome inhibitor E7107 was the first to reach Phase I clinical trials in a range of solid tumours, however unforeseeable toxicities inducing bilateral optic neuritis led to the discontinuation of the study (221). This underlines the importance of validating the disease specific mechanistic action of spliceosome inhibitors using different compounds. Therefore, Spliceostatin A (SSA) was selected to assess whether treating *SF3B1* mutated CLL with spliceosome inhibitors offers clinical utility.

SSA is a methylated product, isolated from the FR901464 natural compound which is produced by *Pseudomonas sp* (222). FR901464 was originally shown to increase SV40 DNA virus transcriptional activity as well as inducing cell death in a number of human and murine cancer cell lines (222). SSA and its structurally similar synthetic analogue MMB are both postulated to target the SF3b complex via SF3B1 (213,223). In neuroblastoma and HeLa cell lines, SSA has been shown to induce $G_2$ arrest *in vitro* (213), possibly via deregulation of MCL-1 (224). It has also been demonstrated that MMB treatment can modulate *MCL-1* pre-mRNA in a non-small cell lung cancer cell line, with a combination of MMB and the pan BCL-2 family member (BCL-2, BCL-w, BCL-$X_L$) inhibitor ABT-737 re-sensitising ABT-737 resistant cell lines to the drug, eliciting comparable levels of apoptosis (223).

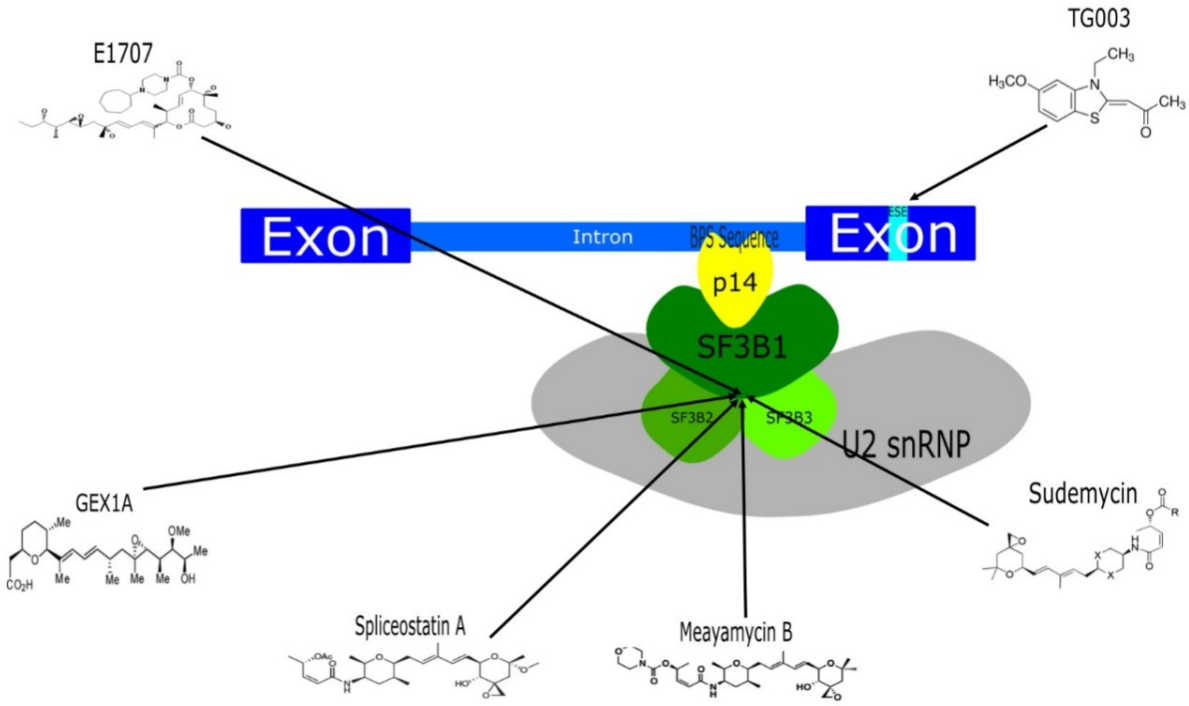



**Figure 69 Chemical Structures and targets of Splicing Inhibitors**

FR901464 derivatives Spliceostatin A, Meayamycin B and Sudemycin, the Pladienolide derivative E1707 and the Herboxidiene compound GEX1A all inhibit the SF3b complex, with all three components (SF3B1, SF3B2 and SF3B3) shown to be bound to all compounds through immunoprecipitation, cross-linking and mass spectrometry analysis. These compounds inhibit SF3b from co-operating with p14 and binding the Branch Point Sequence (BPS Sequence). TG003, a Clk family inhibitor disrupts cross-exon interactions at Exon Splicing Enhancer (ESE) sites. SSA(213), MMB(223), TG003(217), E7107(225), Sudemycin(215)  Figure design adapted from Wahl *et al.* (2009)(129) and Dehm (2013)(262)

Data generated by this PhD candidate as well as where appropriate data from others within the group (Dr Marta Larrayoz), explores the efficacy and mechanism of action of SSA and MMB in CLL cells and B-cell malignancy models *in vitro* (Ramos cell line, Eμ-TCL1 mouse cells). The effect of microenvironmental support on SSA induced apoptosis, as well as combination treatment with BCL-2 inhibitors was assessed (data from the PhD candidate supported a publication as second author; Larrayoz *et al.* (2015) (145)). Additionally, due to the effect of SSA on MCL-$1_L$ expression levels at the protein and RNA level in CLL and Ramos cells, direct targeting of MCL-1 using the selective small-molecule MCL-1 inhibitor UMI-77 was also conducted. This was to explore whether direct targeting of Mcl-1 was important of the observed affects with the drug or whether other off target effects may also be involved. Consequently, the efficacy of UMI-77 in CLL cells and the effect of microenvironmental support on UMI-77 induced apoptosis was analysed, with and without co-treatment with BCL-2 inhibitors.

## 5.1 SSA induces apoptosis in CLL cells in a dose-, time-, and caspase-dependent manner

### 5.1.1 SSA induces apoptosis in CLL cells in a dose- and time-dependent manner

To assess the efficacy of SSA to elicit cell death in primary CLL cells *in vitro*, PBMCs of local patients from Southampton General Hospital (SGH) (n=9) and the Royal Bournemouth Hospital (RBH) (n=2) were treated with SSA (2.5-20nM) for 24 hours (efficacious range identified using the MTT assay in n=5 cases**, A.14**) (cohort characteristics in **A.13**). SSA-induced cell death was analysed by flow cytometry using Annexin-V FITC/Propidium Iodide (Annexin V/PI), a technique which measures surface expression of phosphatidylserine (as a marker of caspase activation) and cell membrane permeability respectively, or by immunoblotting for PARP cleavage, which is cleaved by activated caspases **(Figure 70)**. SSA treatment induced a dose-dependent reduction in CLL cell viability at 24h, with a significant dose dependent increase in Annexin V/PI binding of cells between 2.5-20nM (**Figure 70A,** lower flow cytometry plots). SSA (5nM) induced >50% cell death in 6/11 cases, but all samples reduced viability by >50% at 10nM SSA or above. This was observed across all cases after 24-hour treatment, with 10nM SSA inducing more than fifty percent cell death following normalisation to the vehicle control in all CLL cases ($P$ = 0.003) **(Figure 70B)**. Furthermore, work by Dr Larrayoz showed that SSA induced apoptosis also occurred in time-dependent manner in CLL cells, with decreasing cell viability after 6, 12 and 24 hours treatment at 5nM, 10nM and 20nM SSA(145) (Attached publication Figure 1D).

### 5.1.2 SSA induces apoptosis in CLL cells in a caspase-dependent manner

To ascertain whether SSA-induced cell death of CLL cells was caspase-dependent, the cells were pre-incubated for 30 minutes with the pan-caspase inhibitor QVD at 25μM, followed by SSA treatment for 24 hours. Apoptosis was determined as in 5.1.1. by flow cytometry (**Figure 70C**, n=5) and immunoblotting (**Figure 70D**, n=3, provided by Dr Larrayoz). QVD treatment significantly inhibited SSA-induced apoptosis by flow cytometry ($P$<0.0001) and western blotting. Therefore, SSA induced cell death requires activation of caspases which is indicative of apoptosis.

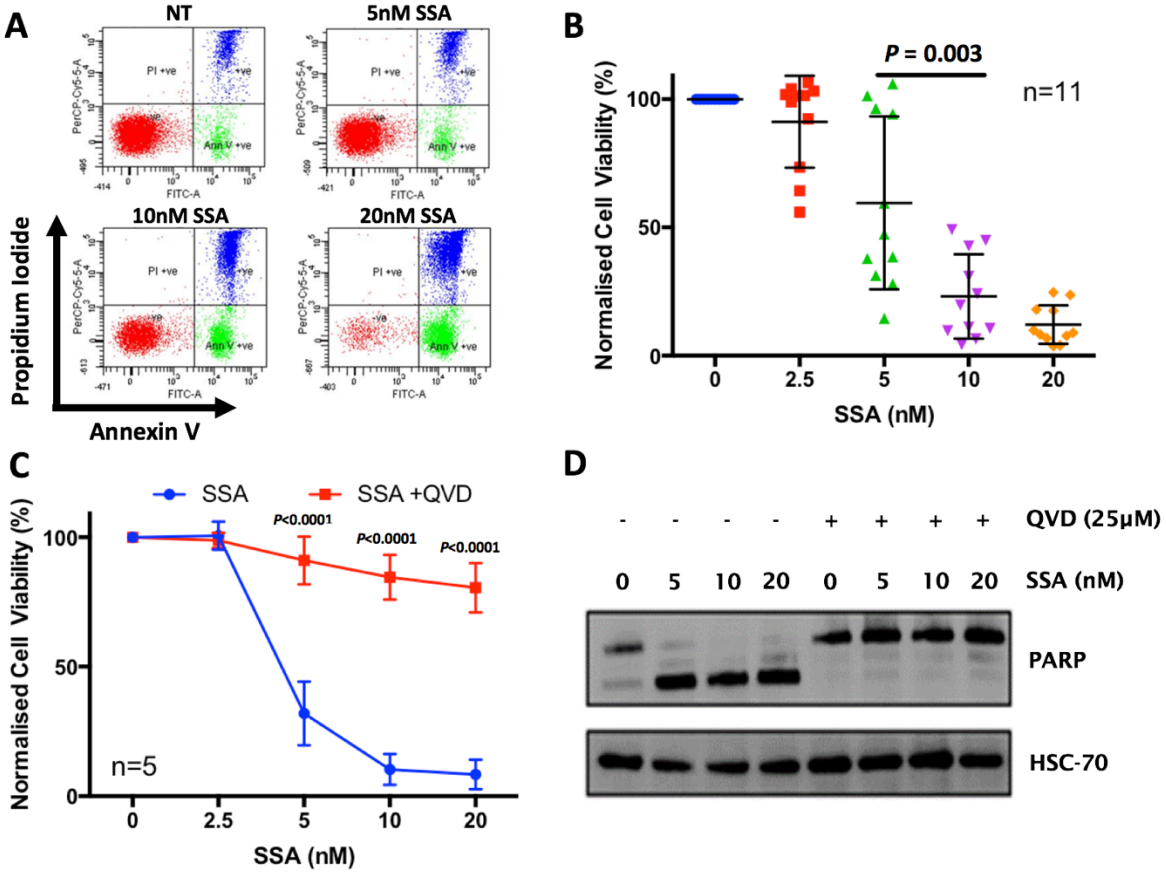



**Figure 70 SSA induces apoptosis in CLL cells in a dose- and caspase-dependent manner**

CLL cells were incubated with vehicle control or 2.5-20nM SSA for 24 hours prior to quantification of apoptosis by Annexin V/PI staining using flow cytometry (n=11). **A** Representative flow cytometry plots of SSA induced apoptosis at various concentrations. **B** Overview of SSA induced cell death in the CLL cohort studied **(**n=11). **C** CLL cell death was analysed by flow cytometry using Annexin V/PI, after 24-hour treatment of SSA in the presence or absence of QVD. The percentage of viable cells was normalised to the vehicle control (n=5). **D** CLL cell death was analysed by immunoblotting using PARP cleavage, after 24-hour treatment of SSA in the presence or absence of QVD. A representative blot of 3 patient samples is shown, and HSC-70 was used as a loading control. Bars represent mean ± standard deviation.

## 5.2 Sensitivity of CLL cells to SSA induced apoptosis

Assessment of the differential sensitivity of CLL cells to SSA-induced apoptosis was carried out, stratifying patient samples by clinico-biological characteristics known to be important for the disease. To do this, the 50% inhibitory concentration values ($IC_{50}$) were calculated using CalcuSyn software (Cambridge, UK). A Mann-Whitney test was used to assess the differences between each defined subgroup. Cases in green were performed by this PhD candidate and combined with the cases included in the SSA publication by Dr Larrayoz (black data points) (145) **(Figure 71)**.

No significant differences were observed across all clinico-biological characteristics, with SSA being similarly efficacious at inducing apoptosis independently of *SF3B1* mutational status (*P* = 0.74) **(Figure 71A)**, IGHV (p=0.81) **(Figure 71B)**, ZAP70 (*P* = 0.96) **(Figure 71C)**, CD38 (*P* = 0.39) **(Figure 71D),** CD49d (*P* = 0.19) **(Figure 71E),** surface IgM expression (*P* = 0.43) **(Figure 71F)** and anti-IgM induced calcium flux (*P* = 0.85) **(Figure 71G)**. SSA treatment was equally effective and independent of chromosomal aberrations (145) (FISH data for green data points unavailable at the time of writing, therefore data was not included here). SSA was less effective at inducing apoptosis of B and T cells from control healthy donors compared to CLL cells and was even apparent in non-CLL B cells (CD19+CD5-) and T cells (CD3+) from CLL patients (Attached publication Figure 2F & 2G).

Next, *SF3B1* mutant and wildtype CLL cells (n=5) were treated with titrations of SSA and CHL for 24 hours, with cell viability assessed using the MTT mitochondrial activity assay **(A.14)**. No significant differences between *SF3B1* mutant and wildtype cases were observed for either drug in this analysis, although there was a trend towards greater levels of cell death in *SF3B1* mutants, numbers were too small to perform statistical analysis. Furthermore, no difference between the $IC_{50}$ values of SSA were observed between wild type and *SF3B1* mutated samples **(Figure 71A)**. In a separate analysis of the impact of CHL stratified by *SF3B1* mutation status, Dr Larrayoz demonstrated that *SF3B1* mutated cases showed greater resistance to chlorambucil in comparison to wildtype cases (*P* = 0.0016)(145).

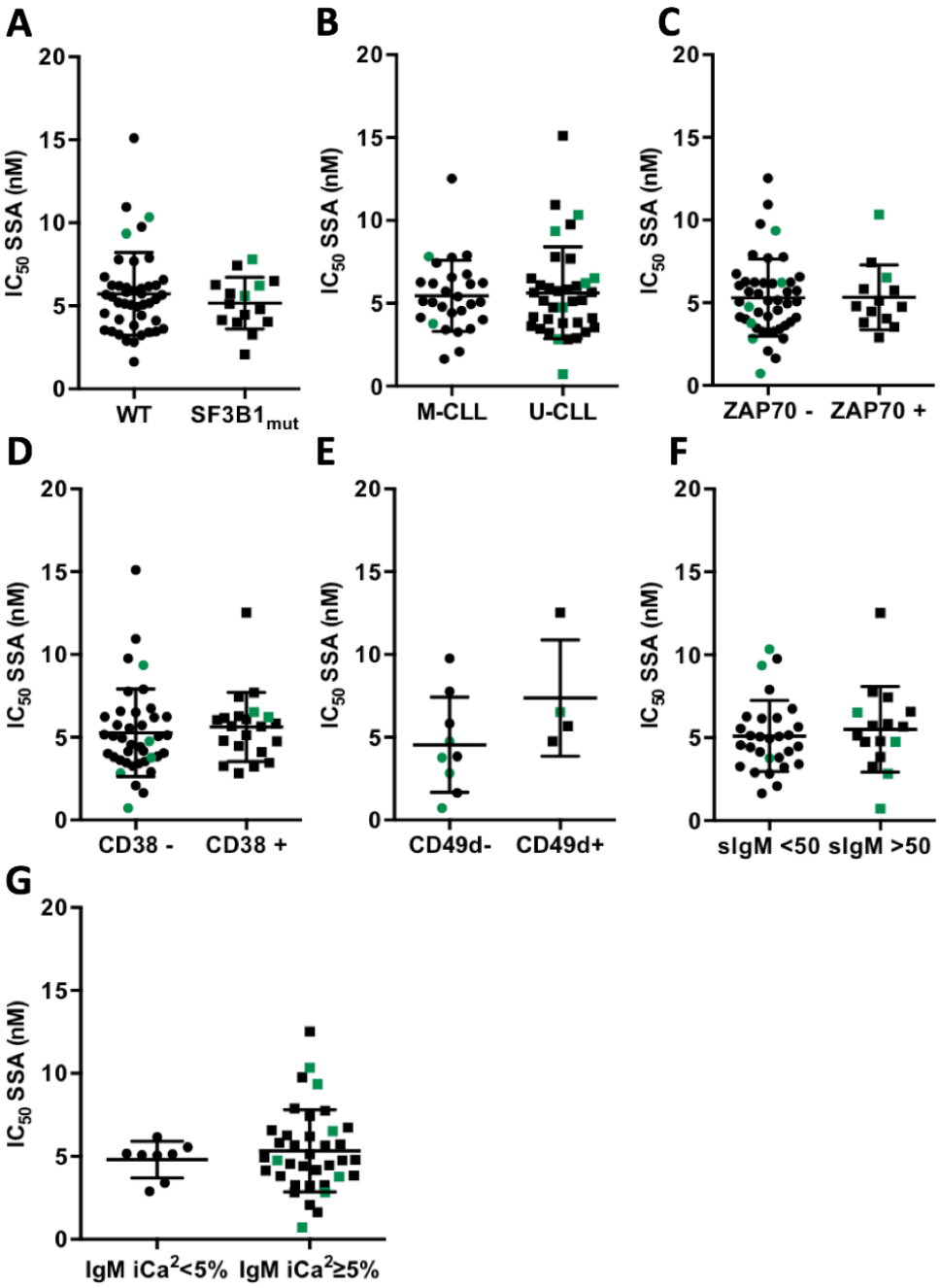



**Figure 71 Sensitivity of CLL cells to SSA induced apoptosis**

Cell viability was assessed by Annexin V/PI flow cytometry, with $IC_{50}$ values calculated using CalcuSyn. The $IC_{50}$ values were plotted for **(A)** *SF3B1* mutational status (wildtype n=44, mutant n=16), **(B)** IGHV mutational status (M-CLL n=28, U-CLL n=36), **(C)** ZAP70 positivity (30% cut off) (ZAP70- n=45, ZAP70+ n=13), **(D)** CD38 positivity (30% cut off) (CD38- n=40, CD38+ n=21), **(E)** CD49d positivity (30% cut off) (CD49d- n=9, CD49d+ n=4), **(F)** Surface IgM (sIgM) expression (cut off 50 MFI) (sIgM <50 n=29, sIgM >50 n=16), **(G)** IgM Calcium signalling (IgM $iCa^2$) capacity (cut off 5%) (IgM $iCa^2$ <5% n=8, IgM $iCa^2$ >5% n=37). Green data points represent samples contributed by the PhD candidate. Black points represent data provided by Dr Larrayoz. There are different numbers of green points per plot due to the availability of the various CLL biomarkers for each CLL case at the time of writing. No variables found to be significantly different by student t-test, *P* values can be found in text. Additional important variables such as CNVs was not available at the time of writing for all cases.

## 5.3 SSA treatment causes intron retention and modifies *MCL-1* and *SF3B1* RNA splicing in Ramos cells

SSA was previously reported to induce intron retention as part of its mechanism of action (213). Therefore, to assess whether SSA functioned analogously in the B cell setting, RT-PCR of *RIOK3* and *DNAJB1* was performed in Ramos cells following treatment with 40nM SSA for 24h (**Figure 72A**). The Burkitt's Lymphoma cell line Ramos was selected to test splicing inhibition, since they are of B-cell origin and have been used previously as a model for CLL publications. *RIOK3* and *DNAJB1* have previously been identified to accumulate after treatment with pladienolides (225), and were therefore selected to assess SSA splicing inhibition efficacy. For this initial experiment, 40nM was used as this was the initial top dose of the treatment dilution, however this was reduced for later experiments due to finding 20nM SSA was sufficient to induce substantial apoptosis. SSA induced intron retention in *RIOK3* and *DNAJB1* in Ramos cells, as shown by the additional larger RT-PCR fragments (650 and 591bp respectively), confirm its function as a splicing inhibitor (**Figure 72A**).

Since siRNA-mediated knockdown of *SF3B1* in solid tumours led to modulation of *MCL-1* splicing (226), the role of SSA in *MCL-1* and *SF3B1* splicing in Ramos cells was evaluated. Ramos cells were treated with 10nM SSA for 6, 12, 18 and 24 hours, in the presence or absence of the pan-caspase inhibitor Z-VAD.fmk (ZVAD) at 25μM, following which samples were prepared for PCR **(2.1.9.1)**. SSA induced a time-dependent reduction of *MCL-1*$_L$ RNA expression, with a concurrent increase in *MCL-1*$_S$ RNA levels **(Figure 72B)**. Importantly, this effect occurred independently of ZVAD treatment, demonstrating that the increase in the *MCL-1*$_S$ isoform was not a consequence of apoptosis. This confirmed our data with primary CLL cells in which SSA induced intron retention (Figure 3A attached manuscript) and *Mcl-1* splicing in a dose- and time-dependent manner (Attached publication Figure 3B &D-F). However, *SF3B1* RNA splicing and expression was unaffected (**Figure 72B**).

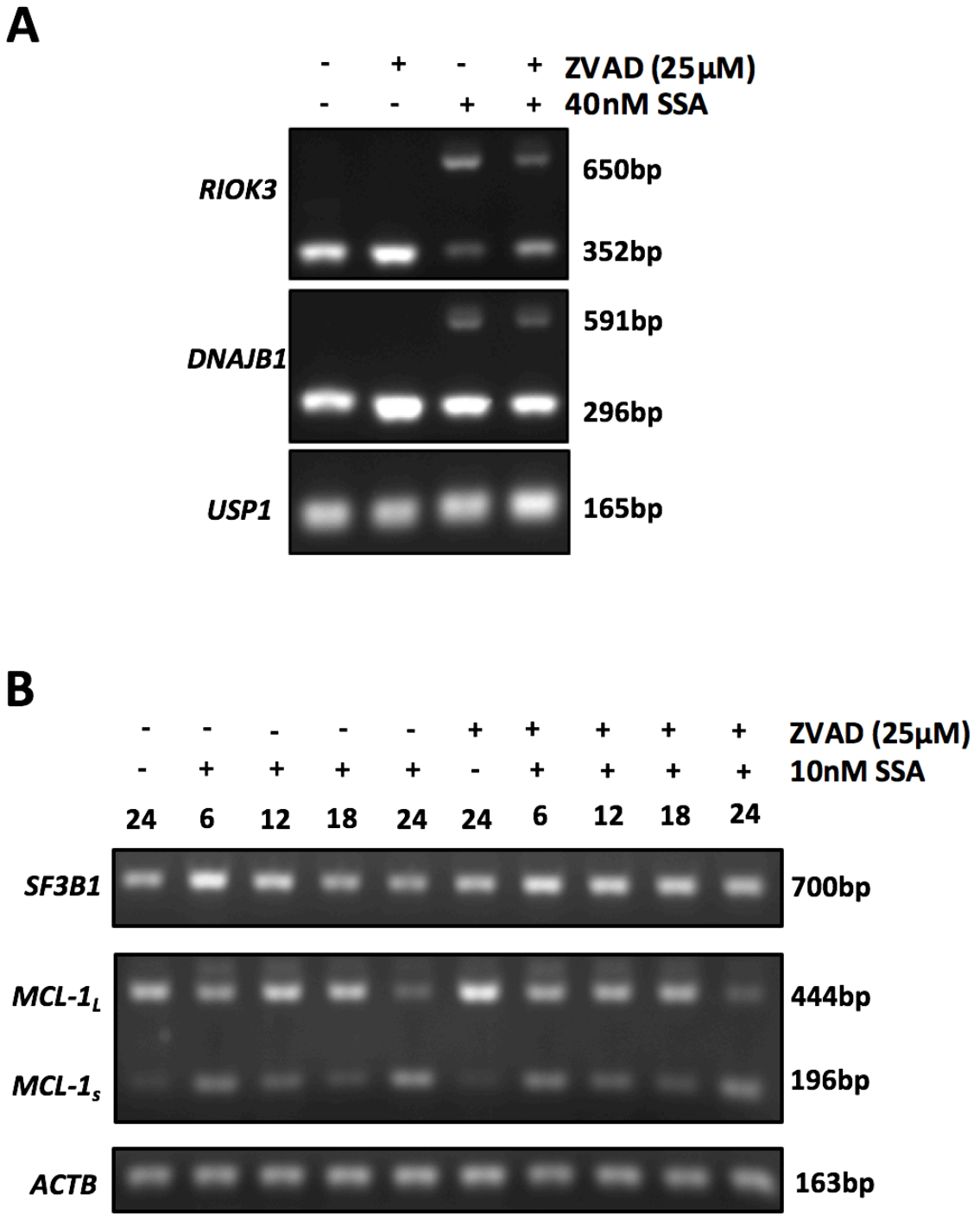



**Figure 72 SSA treatment causes intron retention and modifies *MCL-1*$_{L/S}$ and *SF3B1* RNA splicing in Ramos cells**

**A** Ramos cells were incubated with 40nM SSA for 24h and assessed for intron retention by PCR of *DNAJB1* and *RIOK3*. **B** Ramos cells were incubated with 10nM SSA for 24h and assessed for *MCL-1 and SF3B1* splicing in a time-dependent manner in the presence or absence of Z-VAD.fmk. Representative RT-PCR samples are shown. Actin (*ACTB*) and Ubiquitin (*USP1*) were used as house-keeping genes.

## 5.4 SSA treatment reduces MCL-$1_L$ protein expression in Ramos cells

Since SSA treatment reduced RNA expression of the *MCL-$1_L$* isoform, the consequence of SSA treatment on MCL-$1_L$ protein expression in the presence and absence of ZVAD was assessed. SSA (10nM) treatment induced a time-dependent reduction of MCL-$1_L$ expression in Ramos cells **(Figure 73A&B)**. However, MCL-$1_S$ levels remained unchanged in the presence of ZVAD, suggesting that although accumulation of this isoform occurs at the RNA level, it was not translated into protein **(Figure 73A)**. Furthermore, SF3B1 protein expression in response to SSA was assessed. In the absence of ZVAD, SF3B1 was exposed to a unique cleavage event, which was likely caused by apoptosis and not SSA induced alternative splicing **(Figure 73A)**. However, we did see a small reduction in SF3B1 expression following SSA treatment even in the ZVAD treated samples indicating that SF3B1 expression may also be affected but further work is required to test this hypothesis.

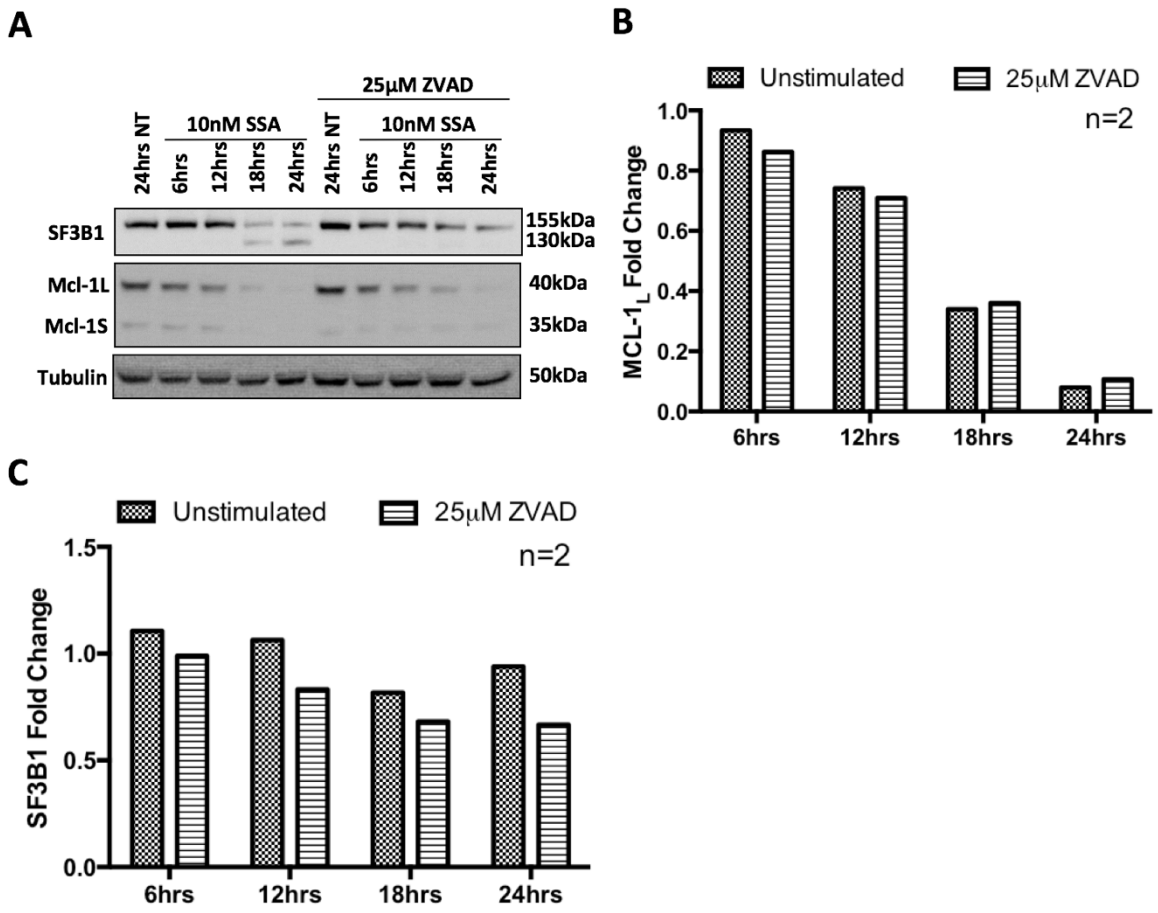



**Figure 73 SSA treatment reduces Mcl-$1_L$ protein expression in Ramos cells**

**A** Ramos cells were incubated with 10nM SSA for 6, 12, 18 or 24 hours in the presence or absence of 25μM ZVAD, and assessed for SF3B1, MCL-1 and Tubulin (loading control) protein expression (n=2). **B** MCL-$1_L$ fold change in response to SSA treatment with and without caspase inhibition over time (6, 12, 18, 24 hours), normalised to Tubulin loading control (n=2). **C** SF3B1 fold change in response to SSA treatment with and without caspase inhibition over time (6, 12, 18, 24 hours), normalised to Tubulin loading control (n=2).

## 5.5 MCL-$1_L$ overexpression in Ramos cells induces resistance to SSA

Analogously to CLL, SSA induced a dose-, time- and caspase-dependent apoptosis in Ramos cells at low nanomolar concentrations **(Figure 74A&B)**, using Annexin V/PI flow cytometry. To determine if the SSA induced reduction in MCL-$1_L$ protein and RNA expression was implicated in SSA-induced cell death, MCL-$1_L$ was overexpressed in Ramos cells using a Myc-tagged construct that was unable to be alternatively spliced by SSA. Next, Ramos cells with and without the Myc-tagged MCL-$1_L$ construct were treated with 2.5-5nM SSA for 24 hours and apoptosis assessed for Annexin V/PI positivity by flow cytometry (data provided by Dr Larrayoz). Compared with both wildtype and mock transfected Ramos cells, the MCL-$1_L$ overexpressing Ramos cells showed a statistically significant protection from SSA-induced apoptosis ($P$<0.01) at 5nM **(Figure 74C)**.

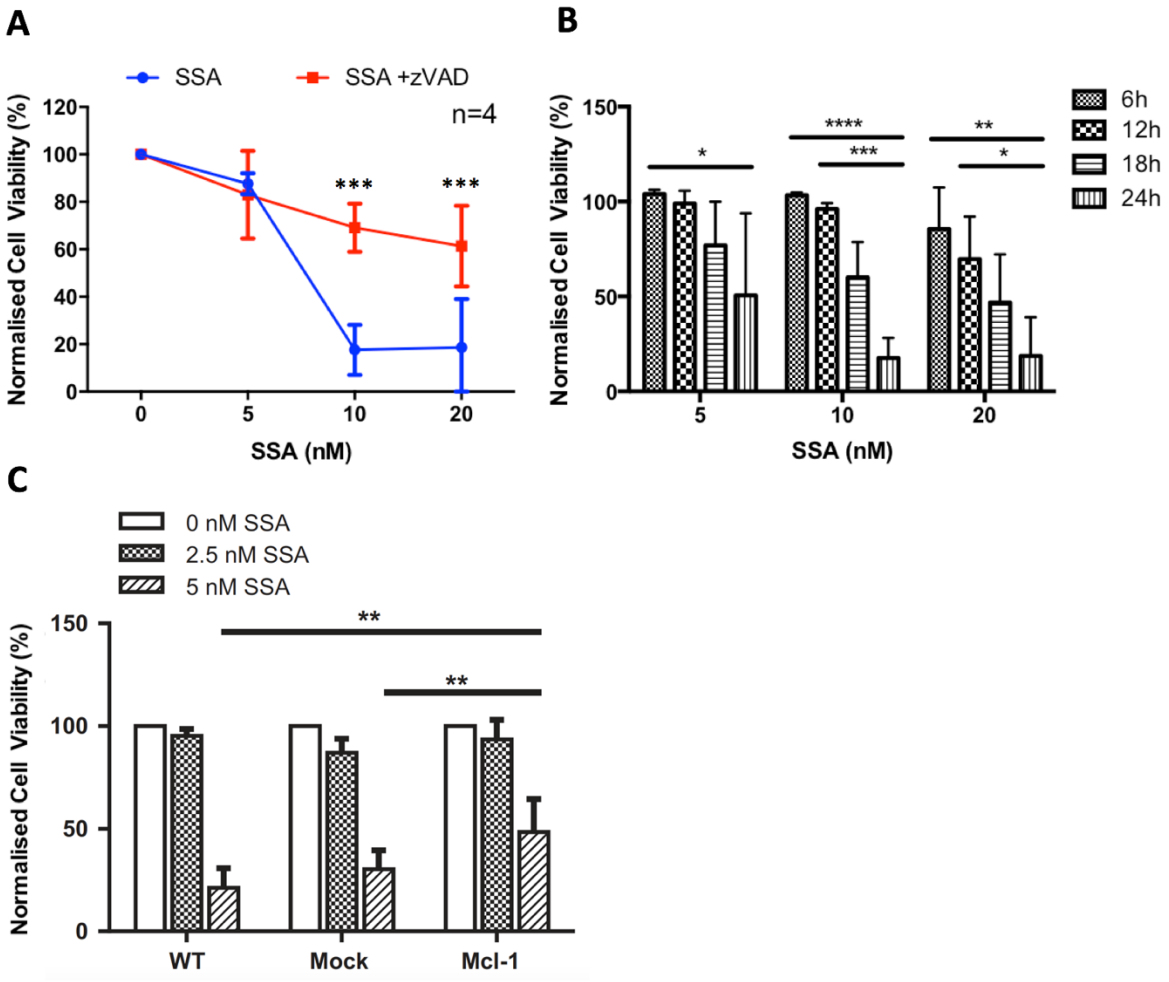



**Figure 74 MCL-1$_L$ overexpression in Ramos cells induces resistance to SSA induced apoptosis**

**A** Ramos cells (n=4) were treated with 5-20nM SSA for 24 hours in the presence or absence of ZVAD. Apoptosis was quantified by Annexin V/PI and the percentage of viable cells normalised to the vehicle control. **B** Ramos cells (n=4) were incubated with (5, 10 or 20 nM) or without SSA for 6, 12, 18 or 24 hours. Cell viability was assessed by Annexin V/PI and normalised to vehicle control at each time point. Bars represent mean ± s.d. **C** Ramos MCL-1$_L$ overexpressing cells and control cell lines were treated with 2.5-5nM SSA for 24 hours then assessed by Annexin V/PI flow cytometry (plot taken from (145)). Bars represent mean ± s.d. *$P$<0.05; **$P$<0.01; ***$P$<0.001; ****$P$<0.0001.

## 5.6 SSA induces intrinsic apoptosis in Ramos cells in a dose- and time-dependent manner

To understand the SSA-induced apoptotic mechanism further, Ramos cells which ectopically expressed Bcl-2 (an anti-apoptotic protein for intrinsic cell death)(83) and CrmA (an anti-apoptotic protein for extrinsic cell death)(83) were treated with 5-20nM SSA for 6, 12, 18 and 24 hours. Bcl-2 and CrmA were overexpressed in Ramos cell lines and confirmed to have increased expression of either protein in comparison to wild type cells by flow cytometry and western blotting **(Figure 75A&B)**. In comparison to wild type Ramos cells, Bcl-2 overexpressing cells were significantly protected against SSA-induced apoptosis ($P<0.001$), whilst overexpression of CrmA had a partial protective effect ($P = 0.28$) **(Figure 75C)**. This suggested that SSA induces cell death via intrinsic apoptosis. A similar trend was observed in a time-dependent manner, however statistical comparison between groups was not possible due to low experimental duplicates **(Figure 75D)**.

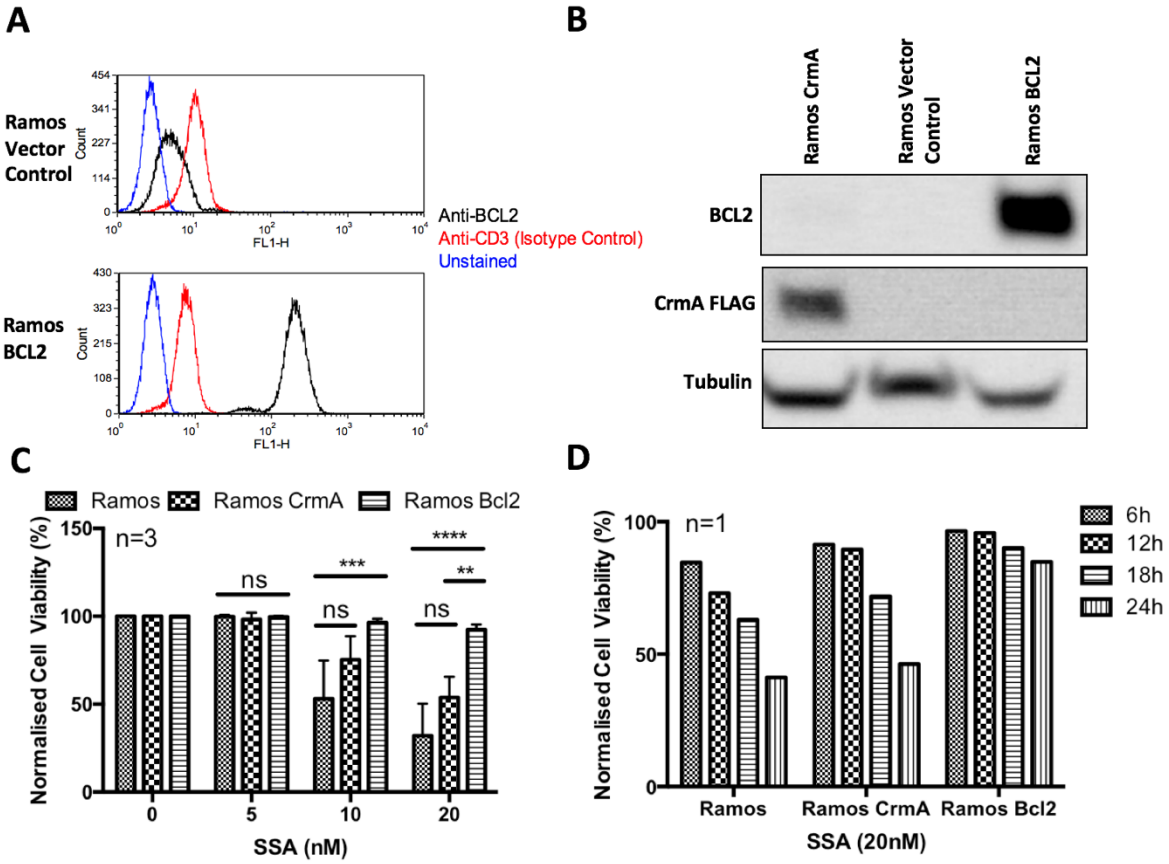


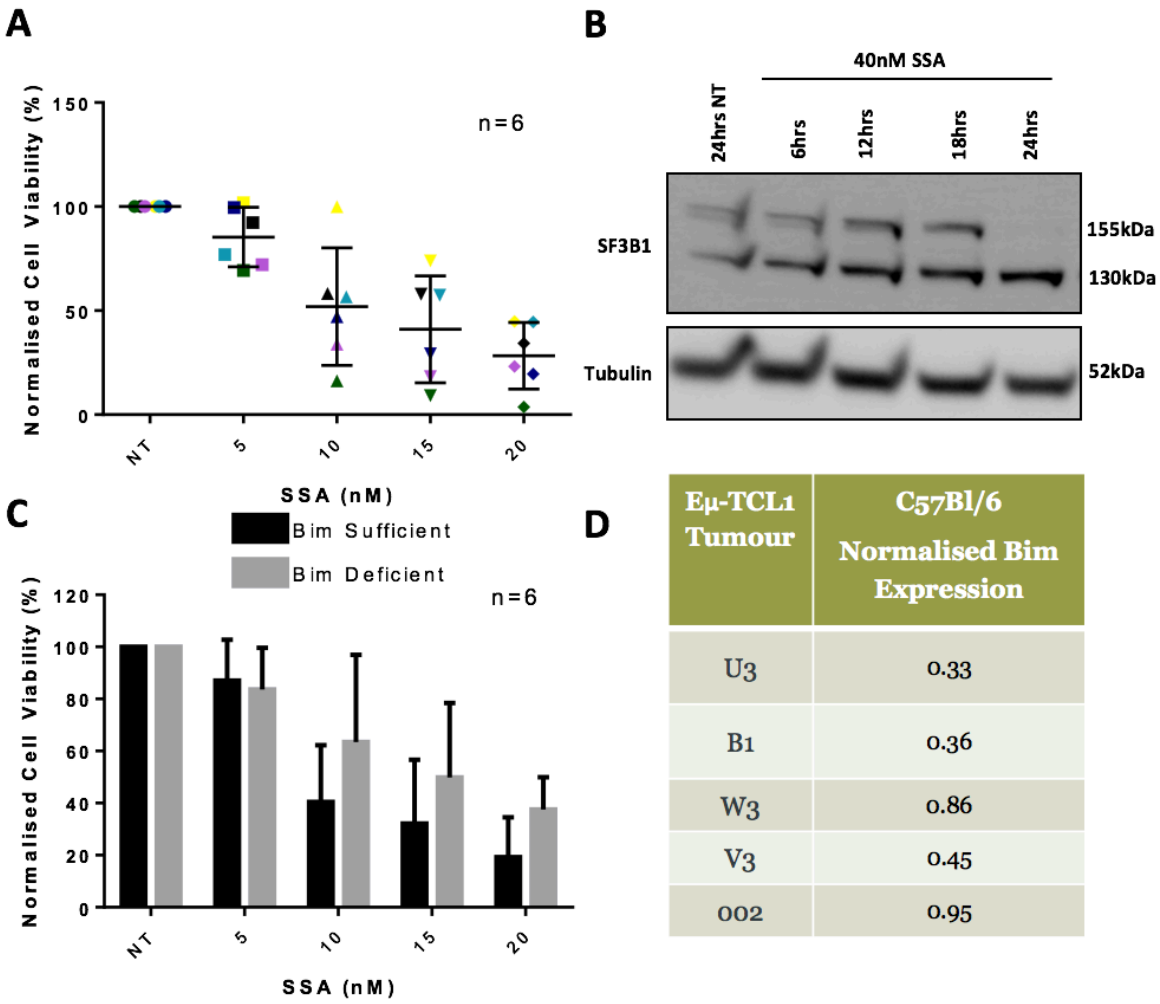

**Figure 75 SSA induces intrinsic apoptosis in Ramos cells in a dose- and time-dependent manner**

**A** Ramos vector control and Ramos Bcl2 overexpression cells assessed for Bcl2 expression by intracellular flow cytometry. **B** Ramos vector control, Bcl2 and CrmA overexpression cells assessed for Bcl2 and CrmA overexpression, respectively, by immunoblotting. **C** Ramos cells (wildtype, CrmA overexpression, Bcl2 overexpression) (n=3) were treated with 5-20nM SSA for 24 hours. Apoptosis was quantified by Annexin V/PI and the percentage of viable cells normalised to the vehicle control. **D** Ramos cells (wildtype, CrmA overexpression, Bcl2 overexpression) (n=1) were treated with 20nM SSA for 6, 12, 18 or 24 hours. Apoptosis was quantified by Annexin V/PI and the percentage of viable cells normalised to the vehicle control. Bars represent mean ± s.d. **P*<0.05; ***P*<0.01; ****P*<0.001; *****P*<0.0001.

## 5.7 SSA induces cell death in Eμ-TCL1 cells *in vitro* and Bim may play a role in SSA sensitivity in Eμ-TCL1 cells

Pre-clinical studies of novel therapeutics require *in vivo* testing before being taken forward for first in man clinical trials. However, SSA requires DMSO as a solvent, a compound which is known to have local toxic effects *in vivo*(227). Therefore, SSA was tested in a panel of the CLL-like sporadic Eμ-TCL1 mouse model tumours *in vitro* (n=6). SSA induced cell death in a dose-dependent manner as assessed by flow cytometry **(Figure 76A)**. However, SSA-induced cell death could not be assessed by immunoblotting, since multiple bands were observed for PARP, preventing accurate calculation of PARP cleavage. As previously shown in this model it was not possible to use ZVAD or QVD to confirm SSA induces caspase-dependent apoptosis in this model, as these chemicals cannot inhibit caspase activation in these cells. However, cleavage of Sf3b1 was shown in a dose-dependent manner, showing an analogous function of SSA between murine and human B-cells **(Figure 76B)**.

Due to the important role of the BH3 only protein Bim in intrinsic apoptosis, as well as the availability of basal Bim protein expression data immediately post mouse culling and tumour harvesting from the spleen of these mice (*pers comms* Dr Matt Carter), SSA sensitivity based on normalised Bim expression was assessed. Tumours with <0.5 normalised expression were considered “Bim deficient” and those with >0.5 “Bim sufficient” **(Figure 76D)**. A trend was observed that Eμ-TCL1 cells that had sufficient Bim expression levels were more sensitive to SSA than Bim deficient cells **(Figure 76C)**.




| Eμ-TCL1 Tumour | C57Bl/6 Normalised Bim Expression |
|---|---|
| U3 | 0.33 |
| B1 | 0.36 |
| W3 | 0.86 |
| V3 | 0.45 |
| 002 | 0.95 |

**Figure 76 SSA induces cell death in Eμ-TCL1 cells *in vitro* that may be dependent on Bim expression**

**A** Eμ-TCL1 cells (n=6) treated with SSA (5-20nM) for 24 hours and analysed by Annexin V/PI flow cytometry, values normalised to untreated sample. **B** Immunoblot of Sf3b1 and loading control Tubulin, demonstrating cleavage of Sf3b1. **C** Annexin V/PI flow cytometry data from sporadic Eμ-TCL1 tumours treated with a range of doses of SSA (5nM, 10nM, 15nM, 20nM) for 24 hours and separated into Bim Sufficient (black bars, <0.5 normalised Bim expression) and Bim Deficient (grey bars, >0.5 normalised Bim expression) groups, showing a protection from death in response to SSA in the Bim deficient group (n=6). **D** C57Bl/6 normalised Bim expression of the five Eμ-TCL1 tumours generated from immunoblotting (courtesy of Dr Matt Carter). The sixth tumour could not be assessed for Bim expression since it was $Bim^{-/-}$.

## 5.8 MMB acts analogously to SSA in CLL

Due to the impracticalities of SSA production, its synthetic analogue MMB was assessed to compare whether it functions analogously to SSA in CLL. MMB induced apoptosis in CLL at low nanomolar concentrations **(Figure 77A&B)**. Furthermore, MMB-induced apoptosis was prevented using the pan-caspase inhibitor QVD shown by inhibition of PARP cleavage by immunoblotting **(Figure 77C)**. Furthermore, similarly to SSA, a dose-dependent decrease in MCL-$1_L$ protein expression was observed (n=1) in the presence and absence of QVD, confirming that it is the targeted action of MMB and not a consequence of protein degradation during cell death.

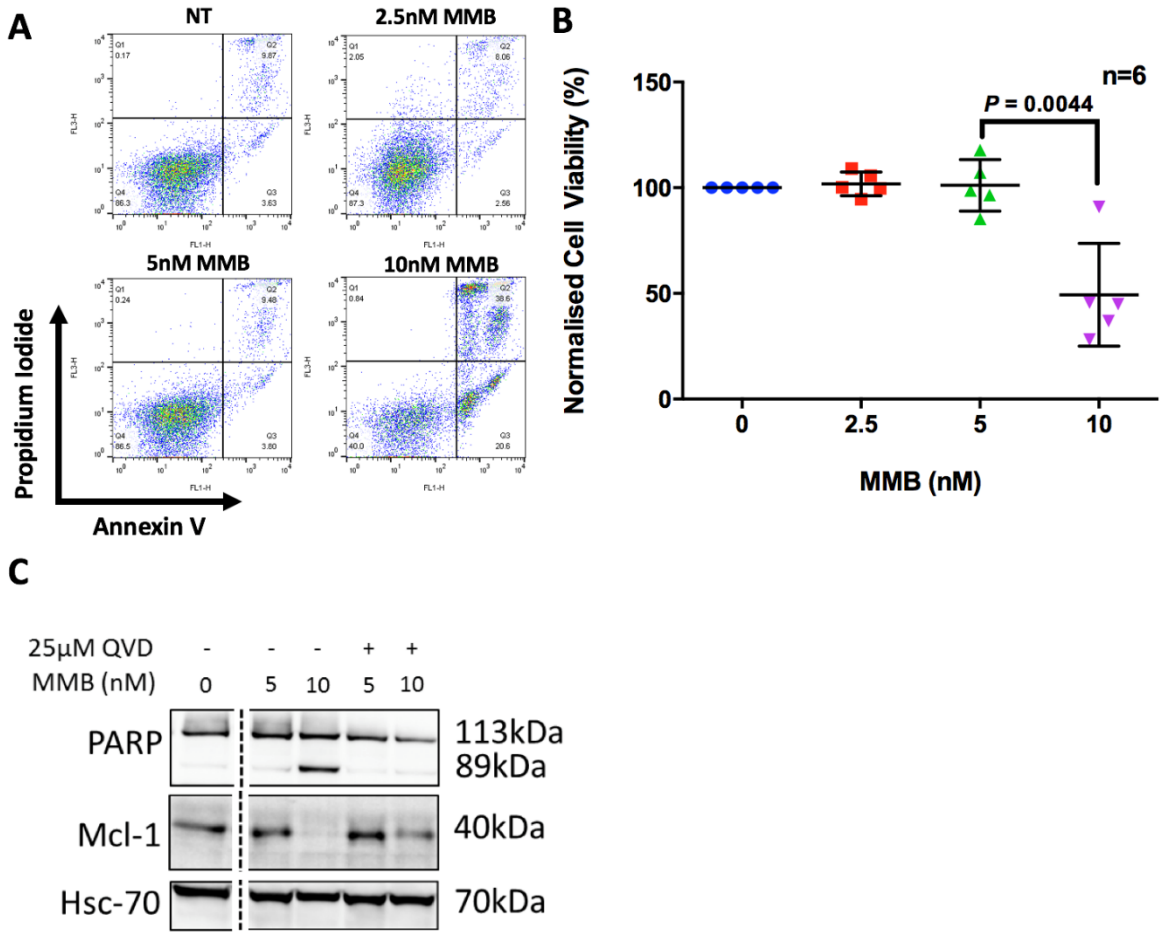



**Figure 77 MMB induces apoptosis in CLL cells in a dose-dependent manner**

**A** CLL cells were incubated with vehicle control or 2.5-10nM MMB for 24 hours. Apoptosis was quantified by Annexin V/PI staining and flow cytometry (n=6). **B** Induction of apoptosis by MMB across all cases by flow cytometry (n=6). **C** CLL cells were treated as above but in the presence or absence of QVD. Immunoblotting was performed on PARP and MCL-1, with Hsc-70 as a loading control (n=1).

## 5.9 The effect of microenvironmental support on SSA induced apoptosis

To determine whether splicing inhibition could be suggested as a therapeutic option for primary CLL cells, partial replication of signals found within the lymph node to determine the effect of microenvironmental support on SSA-induced signalling and cell death was conducted. Importantly, CLL cells are supported by a plethora of cell types and signals within the lymph nodes, however it is not possible to fully replicate all of these signals together. Since previous work by this group assessed the impact of CD40L/IL4 support in the context of SSA (assessed by Dr Larrayoz) (145), assessment of the effect of BCR engagement using bead immobilised αIgM was undertaken **(Figure 78)**. Following sIgM engagement with αIgM, CLL cells were subsequently treated with SSA to assess its efficacy as a monotherapy on tumour cell death.

BCR stimulation alone showed a trend (due to low number of replicates) for increased cell viability in comparison to unstimulated CLL cells **(Figure 78),** confirming our previous studies which demonstrated significant increases in CLL cell viability following treatment with bead immobilised anti-IgM (228). BCR stimulation demonstrated protection from SSA-induced apoptosis (5-20nM), however due to the low number of replicates this protection was not significant. This data supports our findings following treatment with CD40L/IL4 (145) , which also showed a protection from SSA-induced apoptosis. As previously demonstrated both anti-IgM and IL-4/CD40L treatment result in an increase in anti-apoptotic proteins Mcl-1 and Bcl-XL therefore we sought to determine their impact on SSA-induced killing.

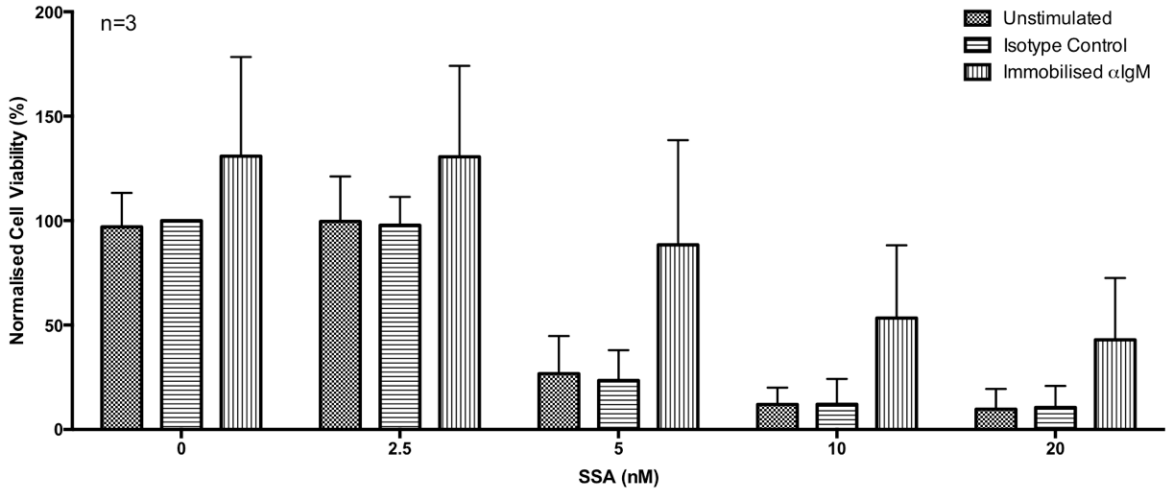



**Figure 78 The effect of immobilised αIgM stimulation on SSA induced apoptosis**

CLL cells were incubated with vehicle control or 2.5-20nM SSA for 24 hours, after 6-hour pre-stimulation with immobilised αIgM, isotype control or no stimulation. Apoptosis was quantified by Annexin V APC staining and flow cytometry (n=3). Bars represent mean ± s.d.

## 5.10 SSA combination treatment with BCL-2 inhibitors in the context of microenvironmental stimulation

Bcl-2 inhibitors such as venetoclax have been shown to be highly efficacious in CLL(229,74), however there is emerging *in vitro* evidence that they are ineffective at targeting CLL cells in secondary lymphoid organs receiving microenvironmental tumour support (230). Since SSA targets *MCL-1* RNA splicing and thus Mcl-1 protein levels, but is ineffective as a monotherapy to target CLL cells receiving microenvironmental support *in vitro,* it was hypothesised that venetoclax in combination with SSA may be able to overcome protective signals within the CLL lymph nodes.

No significant differences were observed between either venetoclax- or SSA-induced cell death as monotherapies. In combination however, differential effects were observed. Following anti-IgM treatment, 10nM venetoclax and SSA augmented cell death in both cases, with dual 20nM treatment leading to increased cell death in one case, whilst protection from apoptosis in another **(Figure 79A)**. This was mirrored in the fractional two-drug analysis, which showed strong synergistic interactions at 10nM, with synergistic and additive interactions observed at 20nM **(Figure 79B)**. However, these results should be interpreted with caution, since it was only conducted on two CLL patients. The 10nM BCR stimulated combination therapy data is comparable to that observed in the context of CD40L/IL4 microenvironmental support (Attached publication, Figure 7), which suggests that the combination of venetoclax with SSA might be effective at targeting CLL cells in secondary lymphoid organs receiving microenvironmental support.

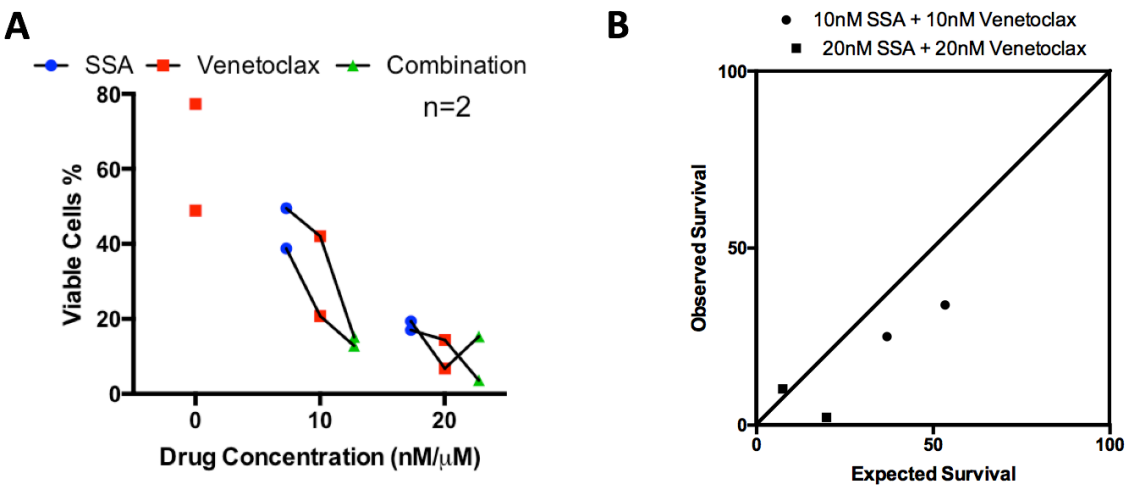



**Figure 79 SSA in combination with the Bcl-2 family inhibitor venetoclax augments apoptosis in CLL cells following immobilised αIgM stimulation**

CLL cells were incubated with vehicle control, 10nM and 20nM SSA, venetoclax or both drugs combined for 24 hours, after 6-hour pre-stimulation with immobilised αIgM, isotype control or no stimulation. Apoptosis was quantified by Annexin V APC staining and flow cytometry (n=2). **A** The effect of both drugs alone or combined (SSA = Blue, venetoclax = Red, combination = Green), cell viability not normalised to non-treatment control. **B** Fractional two-drug analysis for the combination at both doses (circle = 10nM, square = 20nM).

## 5.11 Direct targeting of MCL-1 using the selective small-molecule MCL-1 inhibitor UMI-77 in CLL cells

Splicing inhibitors SSA and MMB both modulate *MCL-1* RNA splicing and subsequent protein expression in CLL cells(145) (MMB; **Figure 77C&D**), and the Ramos cell line *in vitro* **(Figure 72B, Figure 73A&B)**. To better understand the mechanism of action of these splicing inhibitor compounds on MCL-1 in CLL, a comparative analysis using a selective small-molecule inhibitor of MCL-1 was performed. Achieving MCL-1 inhibition is suggested to be important for overcoming microenvironment induced drug resistance in CLL (231), as well as for malignancies reliant on MCL-1 for survival (232). The role of MCL-1 in this context has been recently highlighted in CLL, with low expression of Noxa and corresponding high expression of MCL-1 in the lymph nodes of CLL patients(233). However, until recently there has been a lack of compounds available that selectively target MCL-1 over other BCL-2 family member proteins (234).

An array of MCL-1 inhibitors have been reported, with varying degrees of selectivity for MCL-1 (90,232,234,235). These inhibitors are stratified by their ability to induce Bax/Bak dependent apoptosis, as well as their binding specificity to the BH3-domain of MCL-1 (mimicking Noxa activity) (231,234). Although there has been a number of reviews on the topic of direct targeting of MCL-1 using selective small-molecule inhibitors in CLL (231,235), evidence to support this rationale has been limited. A recent pan-cancer study using the S63845 compound in MCL-1 addicted tumours, showed anti-tumour activity at nanomolar concentrations (236). However, the assessment of targeting MCL-1 directly in non-MCL-1-addicted tumours (such as CLL) is currently withstanding. Therefore, the Noxa-like BH3-mimetic UMI-77 (237) was chosen to assess whether direct targeting of MCL-1 is efficacious in CLL *in vitro*.

UMI-77 is a synthetic small molecule analogue of UMI-59, which was the lead compound identified from a 53,000 synthetic small molecule library using a High-Throughput-Screening (HTS) approach (237). Abulwerdi *et al.* (2014) showed that UMI-77; has high affinity for MCL-1 over other BCL-2 family members (MCL-1=0.49$\pm$0.06, A1/BFL-1=5.33$\pm$1.0, BCL-2=23.83$\pm$1.81, BCL-$X_L$=32.99$\pm$4.33, BCL-w=8.19$\pm$1.91. $K_i \pm$ SD, μmol/L); binds the BH3-binding pocket of MCL-1 (assessed via computational and structural binding studies); and induces Bax/Bak-dependent intrinsic apoptosis(237). All of these characteristics have been previously identified as important for a selective small-molecule MCL-1 inhibitor (234). The following results sections report the efficacy of UMI-77 in CLL cells *in vitro*, as well as its effect on CLL cells with microenvironmental support alone, and in combination with the BCL-2 inhibitor venetoclax.

## 5.12 UMI-77 induces apoptosis in a dose-, caspase- and time-dependent manner

### 5.12.1 UMI-77 induces apoptosis in a dose-dependent manner in CLL cells

To assess whether UMI-77 induced cell death in CLL cells *in vitro* in a similar way to SSA, a cohort of local cases from SGH were identified for the study (n=28). Cohort characteristics can be found in **(A.13)**. The efficacious range of UMI-77 was evaluated between 0.1μM-1000μM using Annexin V/PI by flow cytometry. UMI-77 reduced cell viability at concentrations greater or equal to 10μM, with more than 75% cell death at 100μM. Consequently, all future experiments were performed at 10, 30, 50, or 100μM UMI-77, since 100μM induced >50% cell viability in all cases tested. UMI-77 induced cell death in a dose-dependent manner **(Figure 80A&B)**, with a mean $IC_{50}$ of 48.74μM (s.d=20.32μM) **(Figure 80B)**.

### 5.12.2 UMI-77 induces apoptosis in a caspase-dependent manner in CLL cells

To evaluate whether UMI-77 induced caspase-dependent apoptosis, CLL cells were pre-treated with and without 25μM QVD (30 minutes pre-incubation) followed by addition of UMI-77 for 24 hours. QVD treatment significantly protected CLL cells from UMI-77-induced apoptosis at 30μM ($P = 0.04$) and 50μM ($P = 0.002$), however this protection was lost at 100μM, likely due to off target effects of using such high doses **(Figure 80C)**. This trend was also observed at the protein level, with PARP cleavage blocked at 10μM and 50μM **(Figure 80D)**, and a non-significant reduction in the PARP cleavage ratio, however this could be attributed to low sample numbers **(Figure 80E)**. Therefore, UMI-77 induced caspase-dependent apoptosis in a dose-dependent manner in CLL cells.

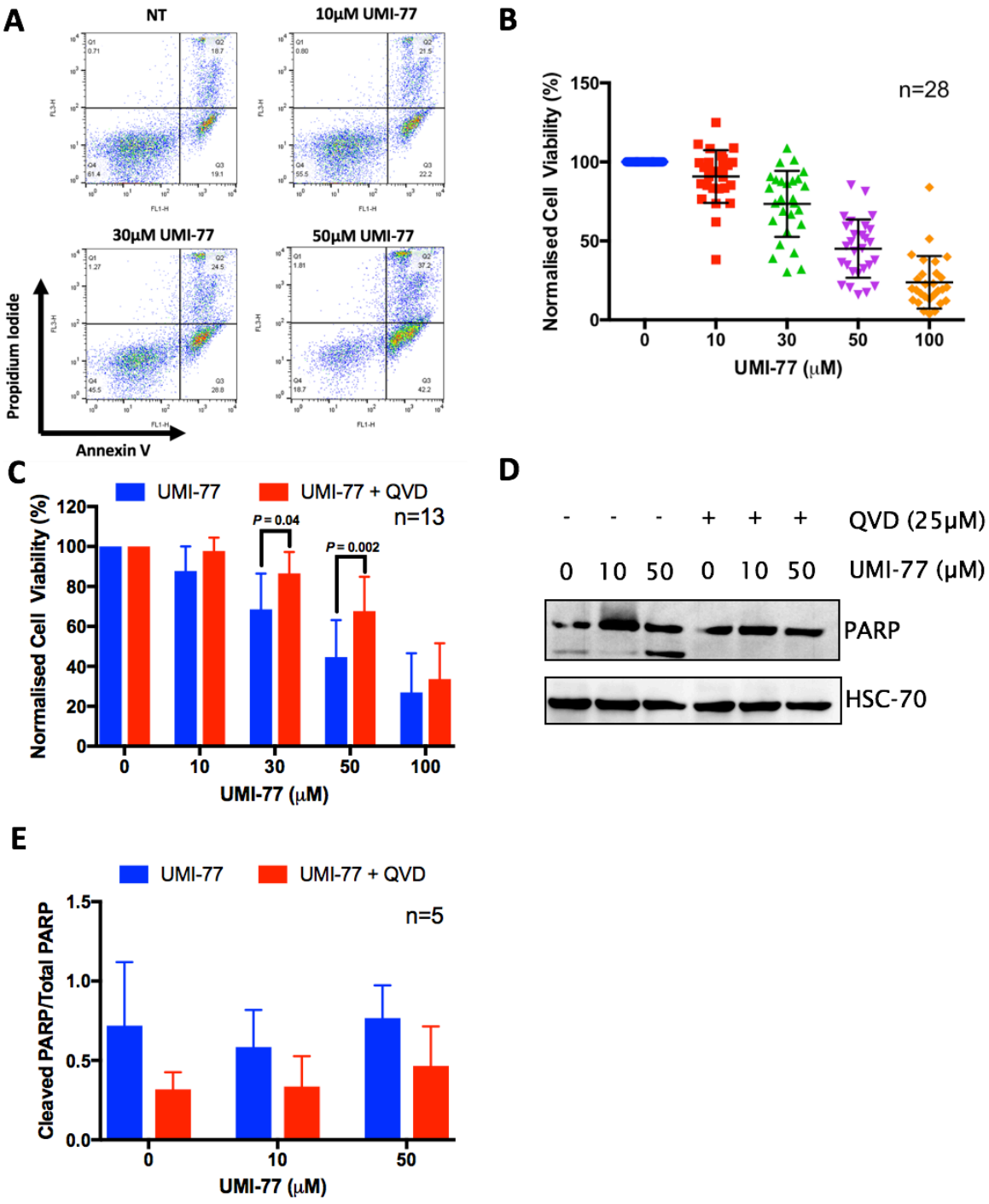



**Figure 80 UMI-77 induces apoptosis in a dose- and caspase-dependent manner in CLL cells**

**A** CLL cells were incubated with vehicle control or 10µM-100µM UMI-77 for 24 hours. Apoptosis was quantified by Annexin V/PI staining and flow cytometry. Representative flow cytometry plots are shown. **B** Summary of UMI-77 *in vitro* treated CLL cohort, assessed by Annexin V/PI staining and flow cytometry (n=28). (**C)** CLL cells were treated as **(A)** but in the presence or absence of QVD and apoptosis assessed by flow cytometry (n=13). **D** Immunoblotting was performed for PARP. HSC-70 was used as the loading control. A representative blot of 5 patient samples is shown. **E** Quantitation of PARP cleavage in CLL samples treated with UMI-77 with and without QVD (n=5).

### 5.12.3 UMI-77 induces apoptosis in a time-dependent manner in CLL cells

Considering the variable sensitivity of CLL cells to UMI-77 **(**4.69μM-93.88μM, **Figure 80B)**, the assessment of cases more resistant to UMI-77 over a 24, 48, and 72-hour time course was undertaken (n=4) **(Figure 81)**. Levels of UMI-77-induced apoptosis at 48 and 72 hours increased over time at 50μM in comparison to 24 hours (48 hours [*P* = 0.0006] and 72 hours [*P*<0.0001]). There was also a trend towards increased cell death at 48 hours in comparison to 24 hours at 30μM (*P* = 0.0781), which became significant at 72 hours (*P*<0.0001) **(Figure 81A)**. Cell viability was completely ablated after 48-hour treatment with 100μM UMI-77 (*P* = 0.0027) **(Figure 81A)**. This loss of resistance to UMI-induced apoptosis at later time points was confirmed by comparing the $IC_{50}$ values for each case at each time point **(Figure 81B)**, with significant reductions in the $IC_{50}$ values between 24-hour and 48-hour (*P* = 0.042), and 72-hour (*P* = 0.038) UMI-77 treatment. At the protein level, there was a trend of increased PARP cleavage in a time-dependent manner, but low sample numbers prohibited statistical assessment **(Figure 81C&D)**. Therefore, UMI-77 induced cell death in a concentration and time dependent manner.

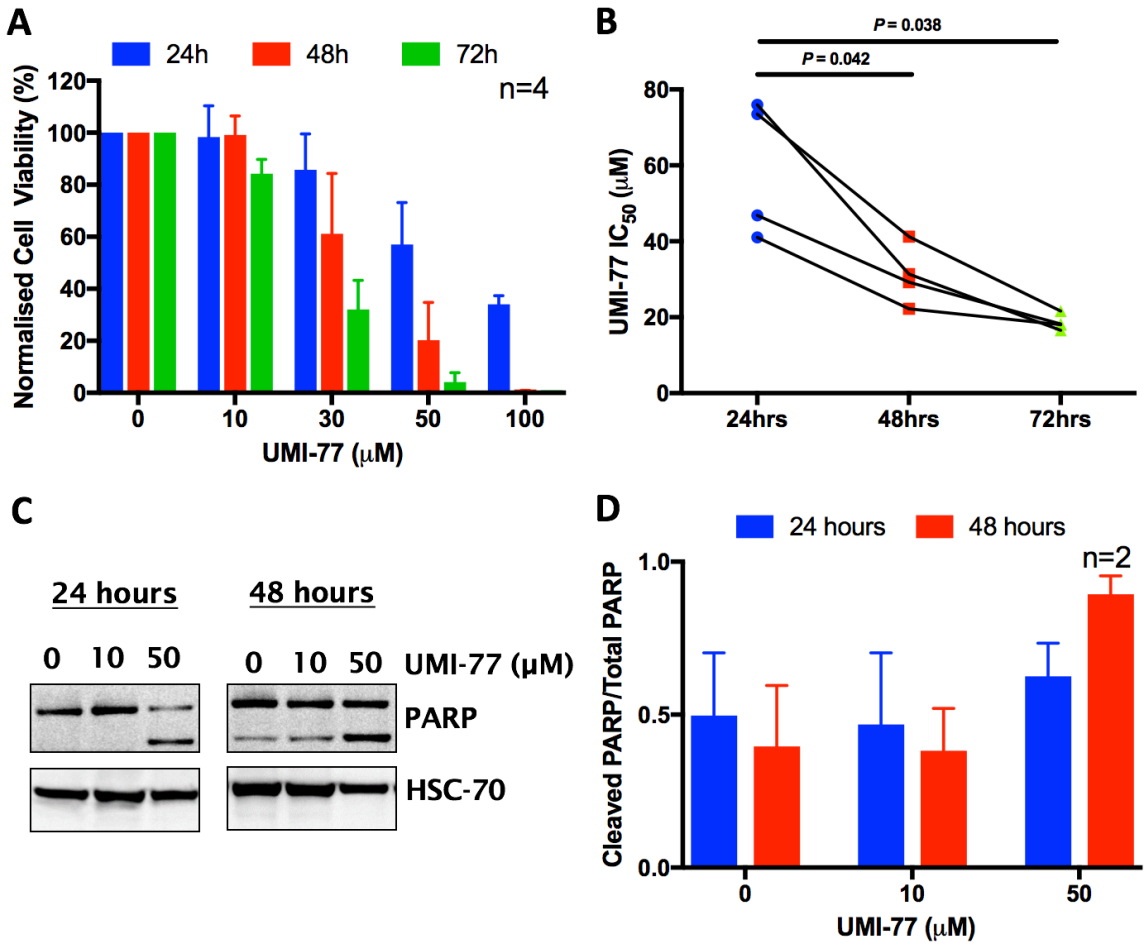



**Figure 81 UMI-77 induces apoptosis in a time-dependent manner in CLL cells**

**A** CLL cells were incubated with vehicle control or 10μM-100μM UMI-77 for 24 hours, 48 and 72 hours (n=4), with apoptosis quantified by Annexin V/PI staining and flow cytometry. **B** The $IC_{50}$ values of UMI-77 was calculated for each patient at 24, 48 and 72 hours. **C** CLL cells were incubated with vehicle control or 10μM and 50μM for 24 hours and 48 hours (n=2). Immunoblotting was performed for PARP. HSC-70 was used as a loading control. A representative blot of n=2 patient samples is shown. **D** Quantitation of PARP cleavage over time in CLL samples treated with UMI-77 (n=2).

## 5.13 Sensitivity of CLL cells to UMI-77 induced apoptosis

Considering the variability in sensitivity to UMI-77 that has been observed in CLL cases, assessment of whether this sensitivity could be explained by differences in clinico-biological factors of the disease was undertaken **(Figure 82)**. This was achieved using the same strategy as in **5.2**. Across all the clinico-biological factors available for evaluation, none were found to show significant differences in UMI-77 $IC_{50}$ (*IGHV* mutational status $P = 0.3595$; ZAP70 $P = 0.9260$; CD38 $P = 0.7662$; CD49d $P = 0.6194$; sIgM $P = 0.2588$; IgM $iCa^2$ $P = 0.8123$) **(Figure 82A-F)**.

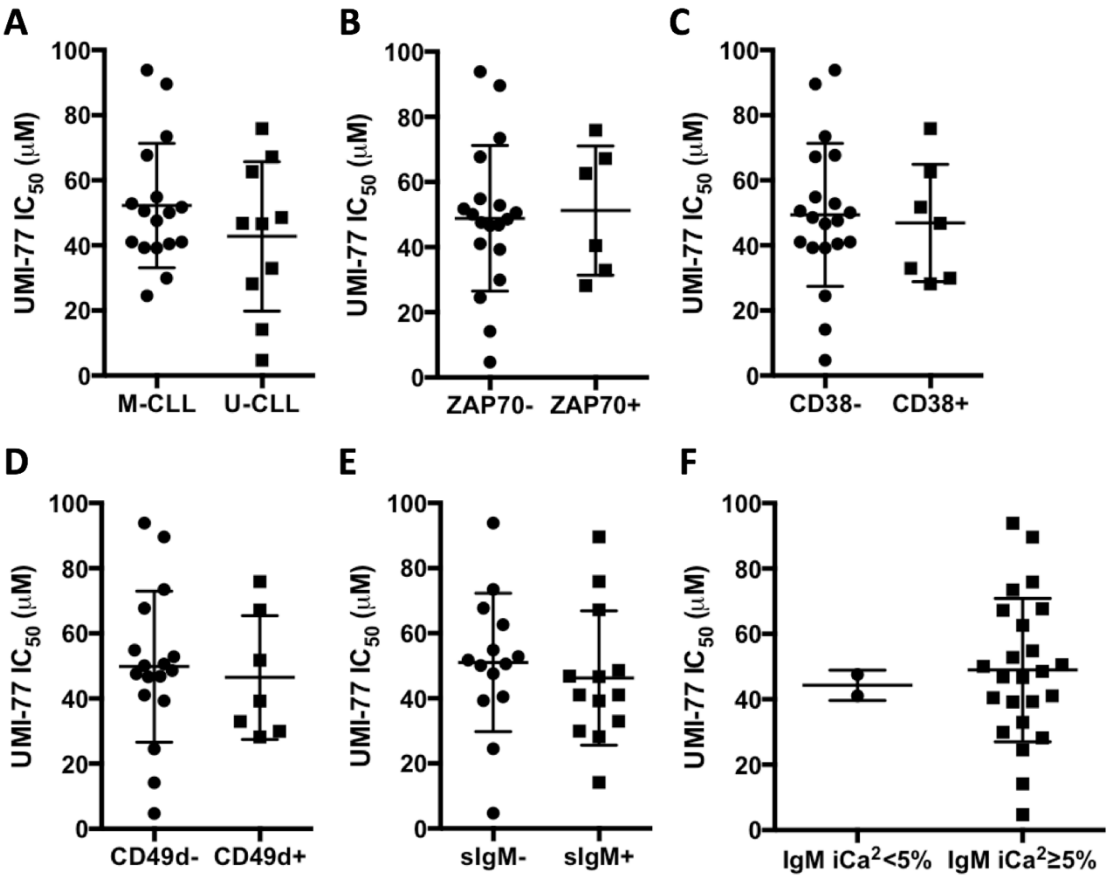



**Figure 82 Sensitivity of CLL cells to UMI-77 induced apoptosis**

Cell viability was assessed by Annexin V/PI flow cytometry, with $IC_{50}$ values calculated using CalcuSyn. The $IC_{50}$ values were plotted for **(A)** IGHV mutational status (M-CLL n=17, U-CLL n=10), **(B)** ZAP70 positivity (30% cut off) (ZAP70- n=19, ZAP70+ n=6), **(C)** CD38 positivity (30% cut off) (CD38- n=20, CD38+ n=7), **(D)** CD49d positivity (30% cut off) (CD49d- n=17, CD49d+ n=7), **(E)** Surface IgM (sIgM) expression (cut off 50 MFI) (sIgM <50 n=14, sIgM >50 n=13), (**F)** IgM Calcium signalling (IgM $iCa^2$) capacity (5% cut off) (IgM $iCa^2$<5% n=2, IgM $iCa^2$>5% n=24). CNV data was not available for analysis at the time of writing.

## 5.14 The effect of microenvironmental stimulation on UMI-77 induced apoptosis

Next, the efficacy of UMI-77 to induce apoptosis in the context of microenvironmental support was investigated. The two model systems described above were employed. No statistically significant differences between treatment groups were observed (largely due to the heterogeneity between samples and the low sample number, n=3) **(Figure 83)**. However, there was greater cell viability in cells stimulated using CD40L/IL4 than following BCR stimulation, compared to vehicle/isotype control treated cells.

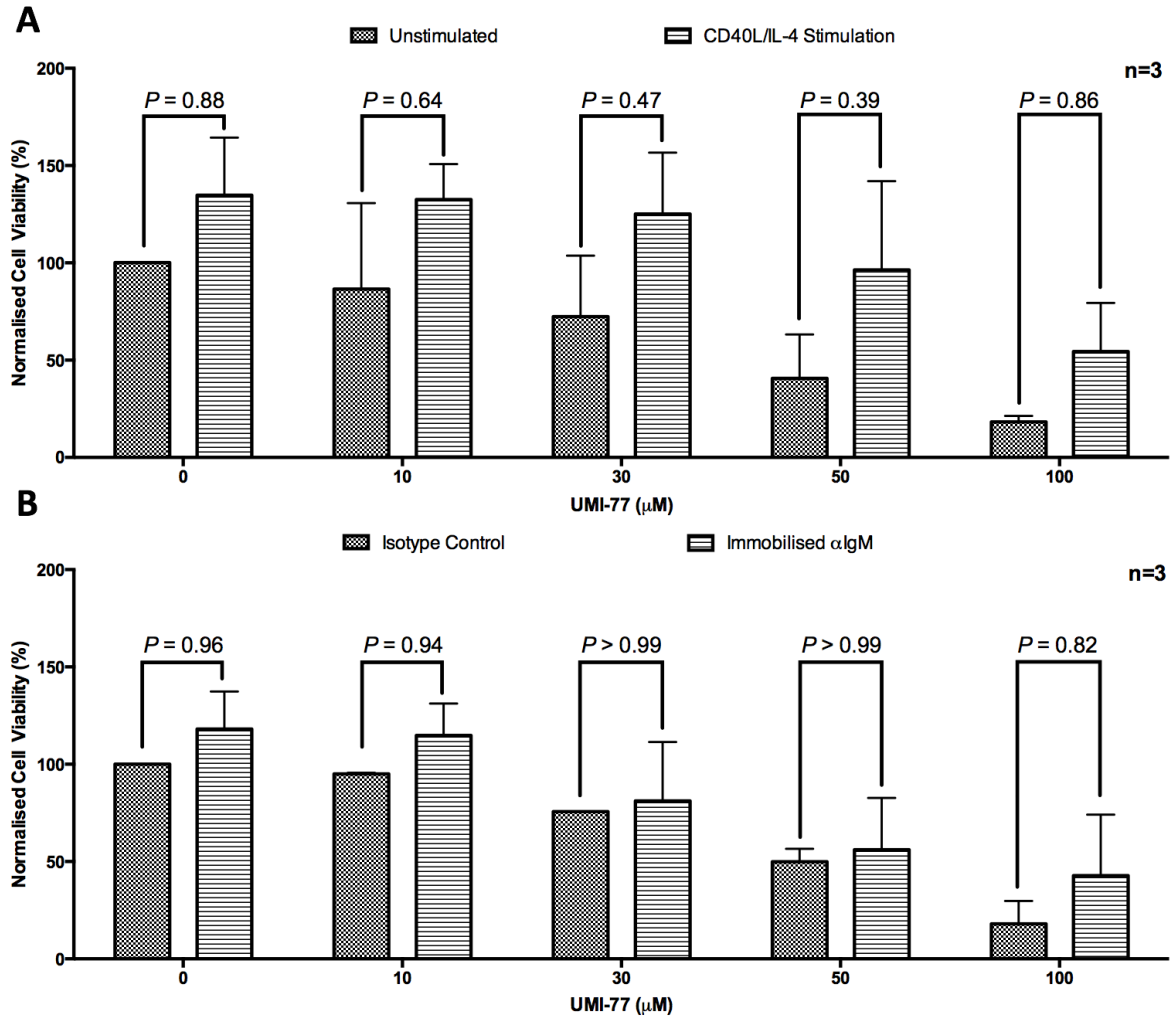



**Figure 83 The effect of CD40L/IL4 and αIgM stimulation on UMI-77 induced apoptosis**

**A** CLL cases were pre-treated with 300ng/mL CD40L and 10ng/mL IL-4 for 6 hours prior to 24-hour treatment with UMI-77 (10-100μM) and analysed for apoptosis by flow cytometry (n=3). CD40L/IL-4 stimulated data was normalised to unstimulated vehicle control. **B** CLL cells were pre-treated with 2:1 bead:cell immobilised αIgM or IgG isotype control for 6 hours prior to 24-hour treatment with UMI-77 (10-100μM) and analysed for apoptosis by flow cytometry (n=3). Immobilised αIgM data was normalised to isotype control vehicle control.

## 5.15 UMI-77 combination treatment with BCL-2 inhibitors in the context of microenvironmental stimulation

In tumours with high expression of BCL-2 family members, therapeutic resistance to BCL-2 family member inhibitors can occur when delivered as a monotherapy. Therefore, combination therapies targeting multiple members of the same pathway can be used to overcome this resistance. As shown above, SSA induced apoptosis and perturbed *MCL-1* RNA splicing in combination with venetoclax led to substantial re-sensitisation of CLL cells in contrast to either compound alone. Therefore, it was hypothesised that the high UMI-77 $IC_{50}$ values may be a result of overexpression of other BCL-2 anti-apoptotic family members such as BCL-2 or Bcl-$X_L$. Furthermore, that direct targeting of MCL-1 in combination with venetoclax may replicate what was observed with SSA and sensitise cells to UMI-77.

### 5.15.1 UMI-77 combination treatment with venetoclax in the context of CD40L/IL-4 stimulation

Initially, a dose-titration of venetoclax (3-100nM) alone, or in combination with either 10μM or 50μM UMI-77 was conducted **(Figure 84A)**. This showed that 10μM UMI-77 in combination with 10nM venetoclax was the lowest dose combination that led to a significant reduction in cell viability in comparison to 10nM venetoclax alone (*P* = 0.05, n=3, **Figure 84A**). Therefore, 10μM UMI-77 and 10nM venetoclax was taken forward in a larger number of cases (n=8), with analysis extended to 5μM UMI-77 and 5nM venetoclax, since venetoclax alone can induce substantial apoptosis at less than 10nM. This identified that certain patients were sensitive to the combination therapy, whilst others showed no difference in viability to venetoclax monotherapy **(Figure 84B)**. At the protein level, CD40L/IL4 pre-treatment led to an increase in MCL-1 and BCL-$X_L$, and to a lesser extent BCL-2 **(Figure 84C)**. When considered using the fractional 2-drug analysis method, UMI-77 in combination with venetoclax was not sufficient to induce synergy in most cases, suggesting other pro-survival proteins such as BCL- $X_L$ may play a role in treatment resistance **(Figure 84D)**. If time had permitted, this analysis would have been repeated with navitoclax which could inhibit both Bcl-2 and Bcl-XL at nanomolar concentrations, to examine this hypothesis.

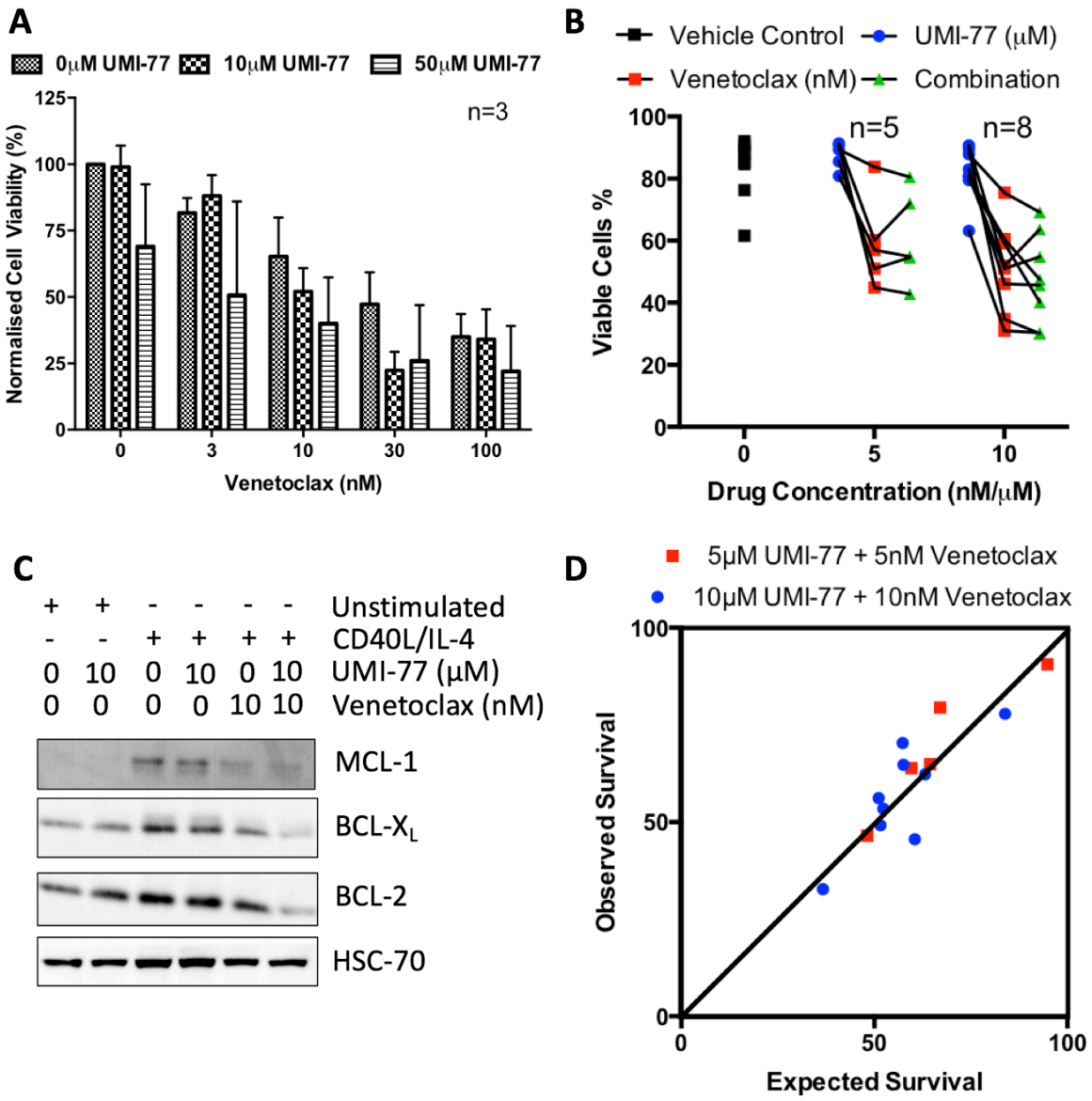



**Figure 84 UMI-77 in combination with venetoclax differentially augments apoptosis in CLL cells following CD40L/IL-4 stimulation**

**A** CLL cells were incubated with vehicle control, UMI-77, venetoclax or a combination of both drugs for 24 hours, after 6-hour pre-stimulation with and without CD40L/IL4. Apoptosis was quantified by Annexin V/PI staining and flow cytometry. **B** The effect of both drugs alone or combined on CLL viability was evaluated as indicated in **(A)** (UMI-77 = Blue, venetoclax = Red, combination = Green). Cell viability was not normalised to the untreated control. **C** Representative example of immunoblotting of BCL-2 family member proteins in response to mono- and combination treatment (n=4) **D** Fractional two-drug analysis for the combination at both doses (circle = 10μM/10nM, square = 5μM/5nM).

### 5.15.2 UMI-77 combination treatment with venetoclax in the context of BCR stimulation

The initial dose titration experiment in the context of CD40L/IL4 stimulation was used here to guide concentration selection **(Figure 84A, 5.15.1).** Single and combined doses of 5nM/5μM and 10nM/10μM venetoclax/UMI-77 post BCR stimulation with anti-IgM in CLL cells were selected. As demonstrated following CD40L/IL4 treatment, 5nM venetoclax alone induced similar levels of apoptosis as the combination of 5μM UMI-77 and 5nM venetoclax **(Figure 85A)**. CLL cells treated with the combination of 10μM UMI-77 and 10nM venetoclax, showed variable responses, with some cases sensitive to the combination whilst others were less sensitive **(Figure 85A)**. BCR stimulation increased the expression of MCL-1, BCL-2 and BCL-$X_L$ **(Figure 85B)**. In fractional 2-drug analysis there were 5 cases in the 10nM/10μM treatment group below the line, and therefore defined as synergistic **(Figure 85C)**. However, their proximity to the line and the centre of the graph suggests that this synergistic interaction is weak.

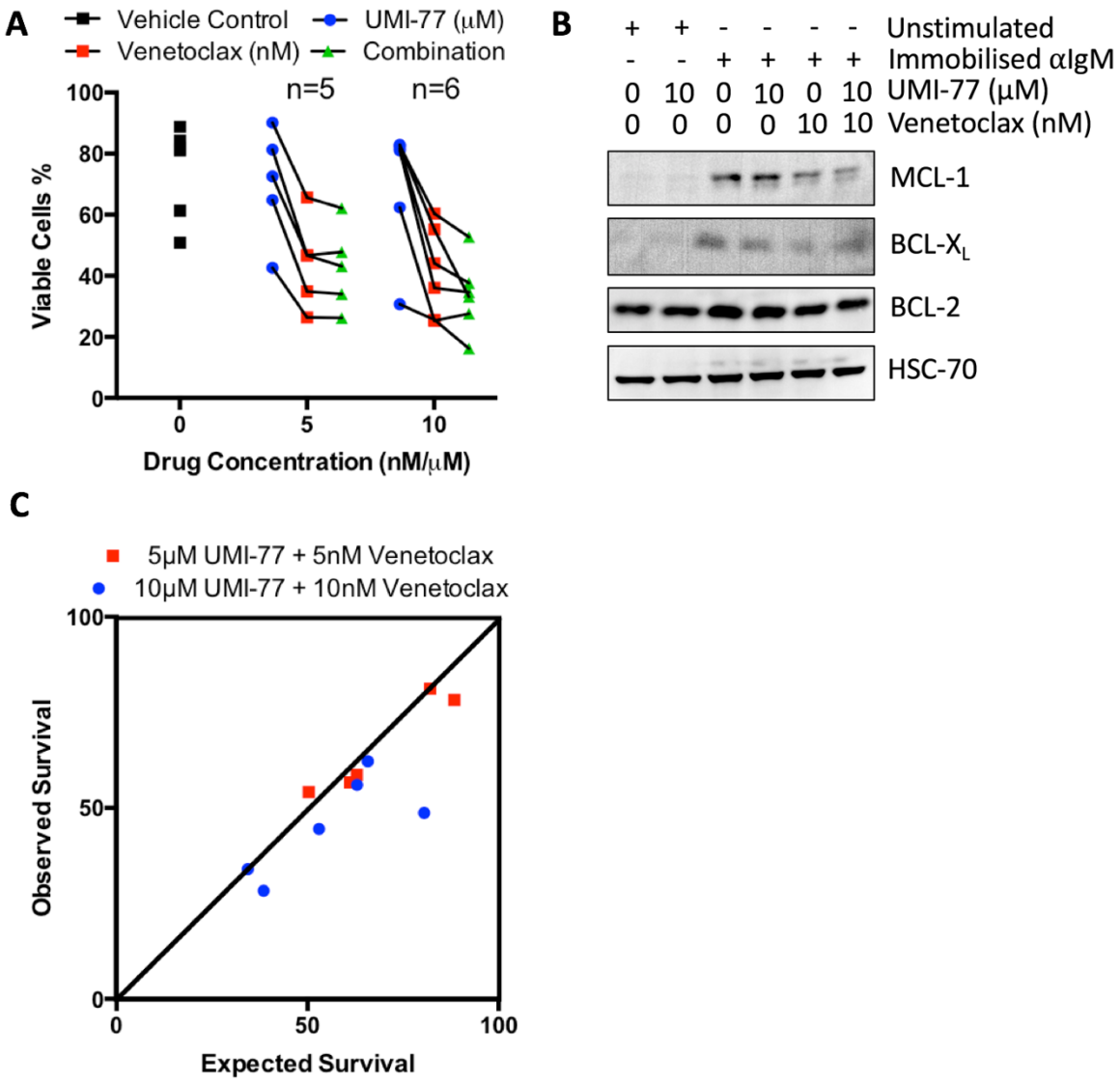



**Figure 85 UMI-77 in combination with venetoclax differentially augments apoptosis in CLL cells following immobilised αIgM stimulation**

CLL cells were incubated with vehicle control, UMI-77, venetoclax or both drugs in combination for 24 hours, with and without 6-hour pre-stimulation with 20μg/ml anti-IgM. Apoptosis was quantified by Annexin V APC staining and flow cytometry. **A** The effect of both drugs alone or in combination (UMI-77 = Blue, venetoclax = Red, combination = Green), on cell viability not normalised to non-treatment control. **C** Fractional two-drug analysis for the combination at both doses (circle = 10μM/10nM, square = 5μM/5nM).

## 5.16 Discussion & Future Directions

The application of omics-led technologies to the analysis of drug sensitivity is of use in identifying novel therapeutic agents or strategies, that may improve response to therapy as well as providing insight into treatment resistance. As the presence of mutations in *SF3B1* in CLL patients has an impact on OS independent of known disease biomarkers (shown in this thesis and multiple other cohorts), it was selected as a candidate for therapeutic targeting using inhibitors, such as SSA and MMB. These studies revealed that treatment with SSA induced apoptosis in most CLL cases evaluated at low nanomolar concentrations; although no difference was observed between *SF3B1* WT and mutant samples. *MCL-1* RNA splicing was seen to be dysregulated after *SF3B1* inhibition which prompted further investigation of the role of MCL-1 in the response to *SF3B1* inhibition and whether targeting MCL-1 directly was an efficacious strategy in *in vitro* assays. Therefore, the MCL-1 specific inhibitor UMI-77 was selected. These studies showed that treatment with UMI-77 induced apoptosis at micro molar concentrations. Combination *in vitro* treatment with the BCL-2 inhibitor venetoclax did not lead to re-sensitisation of CLL cells after pre-stimulation with micro-environmental support, suggesting that other proteins other than MCL-1 and BCL-2 are involved in treatment resistance in the secondary lymphoid organs of CLL patients.

### 5.16.1 Splicing inhibition

This work demonstrates that the splicing inhibitors SSA and MMB, which specifically target SF3B1, promote caspase-dependent apoptosis in CLL and Ramos cells. Furthermore, these compounds were found to elicit apoptosis in Eμ-TCL1 cells at nanomolar concentrations in Bim deficient and sufficient cells. SSA treatment in combination with venetoclax post anti-IgM treatment was undertaken in CLL cells, showing increased cell death than either drug alone, albeit in a small number of cases (n=2) **(Figure 79)**.

SSA treatment in this CLL cohort elicited apoptosis independent of *SF3B1* mutation status, an observation which is contrary to previously reported studies in CLL using the splicing inhibitor Sudemycin (220). Xargay-Torrent *et al.* (2015)(220) found that *SF3B1* and other RNA splicing gene mutated CLL patients were more sensitive to Sudemycin C and D1 than wild type patients, in their cohort of 41 CLL cases *in vitro*. Furthermore, *MCL-1$_S$* (*MCL-1* short isoform) RNA levels were shown to increase over time in Ramos and CLL cells treated with SSA, which has been observed previously with MMB in head and neck cancer cell lines (238), and with Sudemycin D1 in CLL cells (220). However, in our own studies this increase in RNA did not result in an increase at the protein level which contrasts with the work by Gao *et al.* (2014) (238). Indeed, MCL-1$_S$ has been purported as being pro-apoptotic, and shown to be increased at protein and RNA levels in HeLa cells after

SF3B1 knockdown with siRNA (226). Consequently in the studies by Gao et al (238) the apoptotic mechanism of action of splicing inhibitors was attributed to the increase of this pro-apoptotic $MCL_S$ isoform. However, it is important to mention that the role of other pro-apoptotic members in the context of splicing inhibition has yet to be fully explored. In contrast to these data our own work suggests that SSA induced apoptosis may be driven by a reduction in expression of $MCL_L$ rather than an increase in $MCL_S$.

When stratified by basal pro-apoptotic protein Bim expression, Eμ-TCL1 cells responded differently to SSA; samples with lower Bim expression were less sensitive to SSA than Bim deficient tumours. This correlates with work by others where Bim expression levels associated with fludarabine resistance, with cell lines with low Bim expression having higher survival *in vitro* (239). Bim expression in Eμ-TCL1 tumours as eluded to in the results, was calculated by immunoblotting from the founder animals of each sporadic tumour, whereas *in vitro* SSA treatment was conducted from homogenised spleens of different animal generations. Therefore, caution needs to be applied to these results as the Bim expression of the exact tumours treated with SSA were not assessed, however the potential trend identified warrants further investigation. Interestingly, the death of the Eμ-TCL1 tumour cells appeared to be independent of caspase activity. The lack of caspase inhibition observed in Eμ-TCL1 cells is intriguing and may be contributed to by a novel regulatory mechanism, where *in vivo* caspase inhibition showed elevated Interferon-β levels (IFN-β), followed by compromised HSC function, culminating in loss of Bak and Bax(240). Therefore, the basis behind a lack of caspase inhibition in the *in vitro* experiments could be that without functioning Bak and Bax in the Eμ-TCL1 cells, classical inhibition of caspases using ZVAD cannot be achieved. However further work needs to be conducted to elucidate this. Experiments to decipher whether the working concentration or the particular pan-caspase inhibitor used could have an effect on the lack of protection was carried out using a range of apoptosis inducing compounds (Dr Matt Carter, data not shown), showing that up to 75μM using ZVAD or QVD could not protect the cells from death.

The splicing inhibitor E7107 had reached Phase 1 clinical trials in a range of advanced solid tumours, however the trial was prematurely halted due to bilateral optic neuritis induced in two patients, with a further individual developing optic neuritis after cessation of treatment(221). This presents serious questions as to whether splicing inhibitors are suitable for use in humans, or whether the effect seen by Eskens *et al.* (2013)(221) is compound specific. Furthermore, the study could only interpret whether inhibition of splicing was occurring in peripheral blood, as biopsies of the tumour site were not taken. Since this clinical trial was for solid tumours not leukaemias, it does not preclude the potential use of splicing inhibitors in future clinical trials, however it does highlight the importance of pre-clinical and safety testing.

SF3B1 has been proposed to have a transcriptional regulation function in cancer, further to its established role as a splicing factor. Through interactions with the Far Upstream Element (FUSE)-binding protein-Interacting Repressor (FIR), regulatory features of c-myc expression, SF3B1 has been shown in co-immunoprecipitation studies using siRNA and SSA to bind to FIR and FIRΔexon2 (a splice variant of FIR which cannot repress c-myc activity(241)) in HeLa cells and primary colorectal cancer samples(242,243). This means that SF3B1 may have a role in c-myc regulation, a pan cancer oncogene. c-*myc* activation, *TP53* disruption and *CDKN2A* deletion occurs in approximately fifty percent of Richter transformation cases, with *c-myc* and *TP53* disruption occurring concurrently in the majority of patients(244,245). Experimental assessment of the potential role of the SF3B1/c-myc axis in Richter's Syndrome (RS) is outside the remit of this project, largely due to the complexity of gaining access to patient material, however future work in this area by others should not be overlooked.

Recently, a bulk and single RNA-seq study of *SF3B1* mutated CLL identified that *SF3B1* mutations: caused alternative splicing, associated with splice variants at the 3' splice site; led to changes in RNA expression in multiple intracellular signalling pathways; and modulated NOTCH signalling via a splice variant in *DVL2*(174). This highlights the unique biological differences between *SF3B1* mutated and wild type CLL, and offers other potential avenues of treating this CLL subset. Inhibition of NOTCH signalling via gamma-secretase inhibitors has shown promise in *in vitro* studies of CLL(119), therefore targeting NOTCH signalling in *SF3B1* mutated CLL might be an interesting lead to follow.

### 5.16.2 Mcl-1 inhibition

CLL cells are known to overexpress anti-apoptotic members of the BCL2 family such as Bcl-2 and Mcl-1 which protect cells from apoptosis(246) and are associated with progressive disease such as del(17p) CLL patients(247). Therefore, compounds which modulate these molecules are keenly sought. Venetoclax is effectively able to inhibit BCL2 and so other BH3 mimetics or inhibitors which target other anti-apoptotic proteins such as MCL-1, BCL2, BCL-$X_L$ and BCL-W are now required(246).

Mcl-1 inhibition using UMI-77 in CLL cells was undertaken, showing that UMI-77 elicits caspase-dependent apoptosis at micromolar concentrations in a dose- and time-dependent manner. Additionally, combination experiments with UMI-77 and the Bcl-2 inhibitor venetoclax show proof of principle, that increased cell death can be induced by blocking multiple pro-survival molecules; also in the context of overcoming micro-environmental support (CD40L/IL-4 and immobilised $\alpha$IgM stimulation) albeit in a patient specific manner.

The search for Mcl-1 specific inhibitors has been identified in the literature as a key bottleneck to overcoming resistance in cancer treatment. Putative Mcl-1 inhibitors have been described by Varadarajan *et al.* (2013) (234) as needing to meet the following requirements to be considered specific; affinity for Mcl-1 over other Bcl-2 family members; inability to induce apoptosis in BAX/BAK double knockout cells; and to induce cell death in Mcl-1 dependent tumour cells.

The compound UMI-77 has been reported to be an Mcl-1 inhibitor(237). Although being validated for Mcl-1 affinity and BAX/BAK double knockout cells, it has not been tested against Mcl-1 addicted tumour cells. Therefore, this compound was taken forward for *in vitro* testing in CLL, as assessing direct inhibition of Mcl-1 in comparison to reduction of Mcl-1 protein levels by modulated splicing would provide an insight into the importance of Mcl-1 in CLL drug resistance.

UMI-77 was shown to kill CLL cells at micromolar concentrations in a caspase-dependent manner. Furthermore, time-dependent apoptosis was observed at all doses in cases more resistant to UMI-77, suggesting that the potential mechanism of resistance at 24-hours post treatment can be overcome with longer periods of drug incubation *in vitro*. Mimicking of lymph node micro-environmental support via stimulation with CD40L/IL-4 or $\alpha$IgM protected against cell death by UMI-77 and venetoclax monotherapy in all cases. However, the UMI-77 venetoclax combination had limited increased impact on cell viability, acting synergistically in only a small number of cases. As SSA has been shown in our laboratory to be synergistic in combination with venetoclax(145), this suggests that SF3B1 inhibition targets more than just Mcl-1, and that other BCL-2 family member proteins are important in treatment resistance in the context of micro-environmental support. Testing of UMI-77 and other specific Mcl-1 inhibitors in combination with the pan Bcl-2 family inhibitor navitoclax would further elucidate the role of proteins such as BCL-$X_L$ and BCL-W in this process.

Recently, the specific Mcl-1 inhibitor S63845 was found to be highly efficacious in multiple Mcl-1 dependent cancer models *in vitro* and *in vivo* at nanomolar concentrations, as well as showing synergy with other inhibitor compounds(236). In comparison to UMI-77, S63845 could be a superior compound for Mcl-1 inhibition, however the Kotschy *et al*. (2016) study did not assess it in primary CLL cells, nor tumours which are not Mcl-1 dependent. Therefore, assessment of S63845 in CLL, especially in the context of microenvironmental stimulation, would be beneficial to understanding the true efficacy of Mcl-1 inhibition in CLL. In addition, alternative approaches to study the impact of Mcl-1, such as short interfering RNA (siRNA) or short hairpin RNA (shRNA) could be of use. In this context, Ramos cells with Mcl-1 knockdown could be tested for reduced sensitivity to venetoclax in comparison to Mcl-1 WT cells, which would provide additional evidence for the clinical relevance of targeting both BCL2 and Mcl-1.

### 5.16.3 Future work

1. Treatment of micro-environmental support stimulated CLL cells with UMI-77 in combination with Navitoclax, to assess whether treatment resistance is dependent on BCL-$X_L$
2. Assessment of the efficacious S63845 Mcl-1 specific inhibitor
3. Identification of additional inhibitor compounds which may target the biology of mutated *SF3B1,* for example gamma-secretase inhibitors due to the activation of NOTCH signalling in *SF3B1* mutated CLL cases

### 5.16.4 Conclusions

This chapter has studied the efficacy of the splicing inhibitors SSA and MMB in primary CLL cells and the Ramos cell line *in vitro,* identifying that they induce apoptosis in a caspase dependent manner, and downregulate Mcl-1 at the protein and RNA level, and SSA is synergistic with the BCL-2 inhibitor venetoclax (in the context of micro-environmental support. Owing to the splicing inhibitor's impact on Mcl-1, comparative analysis of the Mcl-1 specific inhibitor UMI-77 was undertaken, showing efficacy at micro molar concentrations, however not providing the levels of synergy in combination observed with SSA in the context of micro-environmental support.

# Chapter 6: Discussion and future directions

## 6.1 Discussion

Clinical management of cancer patients has improved, through the introduction of targeted therapies with a better understanding of disease biology. This is the case with the treatment of CLL, with the introduction of the BCR inhibitor Ibrutinib (248) and the BCL-2 inhibitor venetoclax (229), which are targeted therapies based on disease biology. Furthermore, the knowledge that *TP53*ab CLL patients respond poorly to chemotherapy, therefore should be treated with non-chemotherapeutic regimens in the first line. However, currently *TP53* is the only gene screened for mutations at diagnosis, and targeted therapies of gene mutated subsets in CLL has not surpassed the pre-clinical stage. Furthermore, there is building evidence that many of the putative recurrently mutated genes in CLL define specific patient subgroups, therefore linking these to CLL biology and designing relevant treatment strategies could improve patient outcome. These putative genes gain greater credence with every study that shows a survival association, or finds them to co-occur with a poor risk subgroup, something this dataset in the context of the literature will be contributing to. Furthermore, outsourcing this data to other smaller scale studies on individual genes has already supported this goal (11,13,14,190,195), and besides the forthcoming publication on the final cohort of this dataset (in preparation), contribution of mutation data to ERIC for their next CLL genomics publication for many of the genes linked with poor outcome in this thesis will help to build momentum for the clinical importance of these putative recurrently mutated genes.

This thesis presented targeted re-sequencing data on a panel of 25 genes in the LRF CLL4 UK clinical trial cohort. Most of these genes were included in the panel design because they had previously been described in the literature as putative recurrently mutated genes in CLL, but had not been tested concurrently in the context of a clinical trial with appropriate long term follow up in a cohort of this size using targeted re-sequencing. It found 18 genes to be recurrently mutated in the full cohort, with *SF3B1* identified as the most recurrent. Furthermore, it identified genes to be recurrently mutated in CLL that previously have not been unearthed, however careful interpretation and further experimental work is required before they can be considered new genes to be recurrently mutation in CLL.

Following which, analysis of the impact of these genes on prognosis was explored, identifying that after multivariate Cox PH analysis with all significant genes and clinical biomarkers, *SF3B1* and *NRAS* were independently associated with reduced OS, with clonal *SF3B1* mutations and subclonal *NOTCH1* mutations independently associated with reduced OS in an analysis where mutations of

each gene were stratified into clonal (>12% VAF) and subclonal (<12% VAF) groups. However, no gene or clonal or subclonal gene subgroup was found to independently predict poor PFS in the context of clinical biomarkers, even after removal of treatment groups from the model. The application of supervised machine learning tools to survival analysis confirmed the importance of *SF3B1* mutated CLL in OS, whilst identifying no gene to define a group of poor PFS cases.

Since mutated *SF3B1* not only represented the most common mutated gene in the CLL4 cohort, and the most important mutated gene after *TP53* in terms of OS, it was selected as an ideal target for *in vitro* studies using splicing inhibitors. SSA and MMB both induced caspase-dependent apoptosis in CLL cells at nanomolar concentrations, with SSA found to downregulate Mcl-1 at the protein and RNA level. Therefore, SSA was selected to test the impact mono and combination therapy with venetoclax in the context of *in vitro* CLL microenvironment conditions, as Mcl-1 is a known to have increased expression in the lymph nodes of CLL patients, and may be contributing to BCL-2 inhibitor treatment resistance. This showed synergistic interactions between the two compounds, suggesting that this could be a novel combination to overcome microenvironmental support in the lymph nodes of CLL patients. To assess whether the SSA induced downregulation of Mcl-1 alone was responsible for this synergy, comparative analysis using the Mcl-1 specific inhibitor UMI-77 was undertaken. UMI-77 was found to induce apoptosis at micromolar concentrations, and did not work as effectively as SSA in combination with venetoclax in the context of microenvironmental support.

In the most comparable study using targeted re-sequencing in terms of cohort size, Nadeu *et al.* (2016) (159) screened 406 untreated CLL cases for mutations in the following genes at the following mutation frequencies: *TP53* (10.6%)*, SF3B1* (12.6%)*, BIRC3* (4.2%)*, NOTCH1* (21.8%)*,* and *ATM* (11.1%). In a study of 452 cases using WGS, all genes except *NOTCH1* and *ATM* were found at a frequency of less than 10%(57). Except for *TP53,* these frequencies vary considerably to those observed in this thesis, showing that cohort characteristics as well as sequencing techniques used and mutation data processing applied makes comparisons between studies without the relevant raw mutation data challenging. However, in the WES CLL8 clinical trial cohort of the Landau *et al.* (2015) (96) study, *SF3B1* and *ATM* were found to highly recurrent as found in this thesis, at 31% and 32%, respectively. Which although higher than observed in this study, is more comparable. Since both CLL4 and CLL8 trial patients were enrolled if they were treatment naïve and requiring therapy, it further supports the idea that similar cohort characteristics are a key feature when drawing conclusions from them, something that is largely reserved to the clinical trial setting.

The Nadeu study was the only one where direct comparison of clonal and subclonal mutation survival analysis could be undertaken for OS with the data from this thesis, however out of the five genes shared between both studies, only clonal *TP53* mutation's impact on OS was a common feature. The Landau study provided supplementary data on univariate survival analysis on 8 genes for PFS and OS, 7 of which had enough mutations to conduct survival analysis in this thesis. For OS, both studies found *BRAF* mutations to associate with poor OS, whilst both finding no impact on OS for: *ATM, POT1,* and *BIRC3,* with the Landau study not finding an association for *NOTCH1* or *EGR2* when univariate analysis here did find an association, whilst showing an association with OS for XPO1, which was not found in the CLL4 cohort. For PFS, all associations between both studies were comparable, except for *EGR2* which was found to predict for PFS in this thesis, but not in the Landau study. The Puente *et al.* (2015) (57) did not provide any OS or PFS data for comparison. From the original CLL8 study, multivariate analysis was conducted including mutation data for *TP53, SF3B1,* and *NOTCH1,* finding that as well as *TP53, SF3B1* was an independent marker of poor prognosis (69). However, in that study the association was with PFS, whereas in this thesis it was only for OS.

*In vitro* analysis of splicing inhibitors in CLL has only been published in three publications, of which data from this thesis has contributed to one. Sudemycins were found to significantly elicit apoptosis in CLL cells, with increased efficacy in cases with mutations in splicing machinery genes (220). The macrolide products FD-895 and pladienolide-B were also found to induce apoptosis in CLL cells (249), with both studies identifying downregulation of MCL-1 as a mechanism of splicing inhibitor function. Both studies support the observations found in this thesis (and the associated publication), that splicing inhibition elicits apoptosis in CLL cells *in vitro*, and that splicing inhibition leads to downregulation of MCL-1. Although all studies listed here also tested splicing inhibitors in combination with other compounds, and in the context of microenvironmental support, comparative analysis of direct MCL-1 inhibition was not considered in any of the publications. However, recently a large pan-cancer study has presented the impact of a MCL-1 inhibitor, showing potent efficacy in MCL-1 dependent tumours. Comparatively, the efficacy of UMI-77 is reduced to the S3685 compound, however CLL is not a MCL-1 dependent disease.

## 6.2 Future Directions

Further to the technical and chapter specific future work ideas discussed within each chapter, there are other concepts that have been highlighted through the production of this thesis that should be addressed. These pertain to the design and structure of work to be undertaken during a thesis, with the aim of describing the clinical and biological implications of genomic lesions in CLL.

### 6.2.1 Targeted drug compound and Clustered Regularly Interspaced Short Palindromic Repeats (CRIPSR) screening libraries of all mutations in recurrently mutated genes in CLL, and their impact *in vitro* and *in vivo*

Not all mutation's impact on disease within a given gene have the same biological impact. Whilst *in silico* prediction tools using online databases infer whether an individual variant's impact on the structure of the protein is damaging or deleterious, true understanding of the effect of each mutation can only be studied in a cellular context. In some cases, this data is readily available through the published work of others (for example, the LOVD database for pathogenic *ATM* mutations, and the IARC database for *TP53* mutations), however for putative genes, especially outside of hotspot regions, functional mutation data is lacking. Furthermore, *in vitro* targeting the proteins of genes found to be recurrently mutated in CLL have been conducted in isolation, however large scale screening of CLL relevant genes in this manner is currently withstanding.

To achieve this, CRISPR libraries based on observed mutations in the CLL literature to create mutated B-cell lines would offer the possibility of testing whether each mutation changes the biology of the cell line, and whether the mutation sensitises the cells to specific inhibitor compounds that target the protein of the recurrently mutated genes in CLL. Following which, combination therapies (with and without microenvironmental support) with current front and second line targeted therapies would elucidate which drugs and at which doses in the context of specific mutations. Successful inhibitor/mutation combinations could then be taken forward to *in vivo* assessment using the Eμ-TCL1 mouse model, a lymphoproliferative disease model with hallmarks of CLL.

This is an approach that was attempted in a targeted manner during the early phase of this project for *SF3B1*. Generation of *SF3B1* K700E CRISPR constructs were attempted, with the view of evaluating the impact of this mutation on Eμ-TCL1 disease pathogenesis, and to identify the optimal targeted therapy for *SF3B1* K700E mutated cases (data presented in appendices: ). Unfortunately, generation of the construct at that time was not possible, due to complications with the design of the construct, the repetitive nature of the region targeted, and that the technology was in the early stages of development when the experiments were performed.

Furthermore, it was later discovered through a CRISPR forum that the construct had a mutation in its Puromycin cassette, making it incapable of being Puromycin resistant. If it was possible to attempt the project again, identifying a potential collaborator to go and work in their lab to learn the skills would be a key feature, as well as selecting multiple targets rather than just *SF3B1* K700E.

### 6.2.2 *In vitro* and *in vivo* testing of novel mono and combination therapies in specific CLL mutated subsets based on the compound drug and CRISPR library screening results

Using the results of the targeted compound drug screen and CRISPR libraries to select the optimal drug combinations for mutated genes, *in vitro* testing of these compounds using CLL samples with mutations in the relevant genes would allow the identification of gene mutated subgroups in CLL that respond to the novel therapy combinations. This could also be attempted *in vivo* using patient derived xenograft models (PDX), however the aggressiveness of the disease is a key feature, since not all CLL patient samples engraft during this process (personal communication: Professor Mark Cragg).

The *in vitro* component of this suggested approach was tested in a small scale in this thesis, with an analysis of *SF3B1* mutated CLL cases treated with SSA, finding no association with increased or reduced sensitivity to the compound in comparison to wild type CLL cases. In the light of the recent single cell and bulk RNA-seq study into *SF3B1* mutated CLL, which found that some *SF3B1* mutated CLL cases had a splice variant in *DVL2* which led to constitutive activation of notch signalling, consideration of the biological impact of mutated genes in intracellular signalling pathways is important. This is because inhibitors of other recurrently mutated genes or members of other pathways may have an impact on drug sensitivity in this context.

## 6.3 Final Conclusion

This PhD project, and the accumulation of the data generated forming this thesis, has attempted to implement targeted re-sequencing to characterise the mutation landscape of CLL4, and to apply that to identify novel survival associations. These aims have been realised, with confirmation known associations maintaining significance in the context of multivariate modelling (*SF3B1*), and the identification of novel genes to associate with survival, which previously have not been observed (*DDX3X*). Furthermore, application of this data to investigation the importance of splicing inhibitors in CLL was realised, with further analysis of a specific MCL-1 inhibitor, in the context of the tumour micro-environment. This has ultimately led to the conclusion that inhibition of BCL-2 and MCL-1 alone is not sufficient overcome these survival signals, and further work to assess the role of other anti-apoptotic proteins in this process is required.

# Appendices

## A.1 Table of Design 1 CLL4 TruSeq panel manifest file

| Target | Build ID | Chromosome | Start Position | End Position | Target strand | ULSO Sequence | DLSO Sequence | Species | GC Content of target (%) | ULSO_Length | DLSO_Length |
|---|---|---|---|---|---|---|---|---|---|---|---|
| TP53_UserDefined (32452970)_48838278 | hg19 | chr17 | 7576525 | 7577155 | + | TTAGTTAGCTACAACCAGGAGCCATTG | GCAAGGAAAGGTGATAAAAGTGAATCTG | Homo sapiens | 44.44444444 | 27 | 28 |
| TP53_UserDefined (32452970)_48838279 | hg19 | chr17 | 7576525 | 7577155 | + | AGGACAGGCACAAACACGCACCTCAAA | CCCAAAGCCAGAGAAAAGAAAACTGAG | Homo sapiens | 51.85185185 | 27 | 27 |
| TP53_UserDefined (32452970)_48838280 | hg19 | chr17 | 7576525 | 7577155 | - | ATCAGGGGGCAGTGATGCCTCAAAG | AGTTGAAAATGGATTTAATACACCTAATCT | Homo sapiens | 56 | 25 | 30 |
| TP53_UserDefined (32452970)_48838281 | hg19 | chr17 | 7576525 | 7577155 | - | TAGTGGTAATCTACTGGGACGGAACAG | TCTTTCCTAGCACTGCCCAACAACACC | Homo sapiens | 48.14814815 | 27 | 27 |
| TP53_UserDefined (32418540)_48838282 | hg19 | chr17 | 7578177 | 7578811 | + | CACCATCGCTATCTGAGCAGCGCTCAT | GCACAAGTGAACAGATAAAGCAACTGG | Homo sapiens | 55.55555556 | 27 | 27 |
| TP53_UserDefined (32418540)_48838283 | hg19 | chr17 | 7578177 | 7578811 | - | AGGTTGTGAGGCGCTGCCCCCAC | TCTGGTTTGCAACTGGGGTCTCTGGG | Homo sapiens | 69.56521739 | 23 | 26 |
| TP53_UserDefined (32452969)_48838284 | hg19 | chr17 | 7571720 | 7573008 | + | CTCGAGCCGGGGAGGCAAAGGCTGCAGTAA | GATGGGGTGAGATTTCCTTTTAGGTAC | Homo sapiens | 63.33333333 | 30 | 27 |
| TP53_UserDefined (32452969)_48838285 | hg19 | chr17 | 7571720 | 7573008 | + | TACCTAACCAGCTGCCCAACTGTAGAA | GCCTTGAAGTTAGAGAAAATTCAACAGTG | Homo sapiens | 48.14814815 | 27 | 29 |
| TP53_UserDefined (32452969)_48838286 | hg19 | chr17 | 7571720 | 7573008 | + | ATCTAAGCTGGTATGTCCTACTCCCCA | TCAGTCTGAGTCAGGCCCTTCTGTCTT | Homo sapiens | 48.14814815 | 27 | 27 |
| TP53_UserDefined (32452969)_48838287 | hg19 | chr17 | 7571720 | 7573008 | - | TTTTCACCCCACCCTTCCCCTCCTTCTCCC | CCAGTGCCCGGGTTCCGGGAGG | Homo sapiens | 60 | 30 | 22 |
| TP53_UserDefined (32452969)_48838288 | hg19 | chr17 | 7571720 | 7573008 | - | TTGACCCCCTTGAGGGTGCTTGTTCC | CTTGTATATGATGATCTGGATCCACCAAG | Homo sapiens | 57.69230769 | 26 | 29 |
| TP53_UserDefined (32452969)_48838289 | hg19 | chr17 | 7571720 | 7573008 | - | ACTGGTGTTTTGTTGTGGGGAGGAGG | AAATGCTGGCATTTGCACCTACCTCAC | Homo sapiens | 53.84615385 | 26 | 27 |
| TP53_UserDefined (32452969)_48838290 | hg19 | chr17 | 7571720 | 7573008 | - | GCTCAGTAAACATATTTGCATGGGGTG | ATTCTCCACTTCTTGTTCCCCACTGAC | Homo sapiens | 44.44444444 | 27 | 27 |
| TP53_UserDefined (32408345)_48838291 | hg19 | chr17 | 7577449 | 7577658 | - | AAAAAGGCCTCCCCTGCTTGCCACAG | TTGCCTCTGACCCCTGGGCCCACCTC | Homo sapiens | 57.69230769 | 26 | 26 |
| TP53_UserDefined (32453006)_48838292 | hg19 | chr17 | 7590695 | 7590878 | - | TACTTGCCCTTACTTGTCATGGCGACT | TAAGCTCCTGACTGAACTTGATGAGTCC | Homo sapiens | 48.14814815 | 27 | 28 |
| ATM_UserDefined (32418544)_48838293 | hg19 | chr11 | 108164040 | 108164204 | - | GATGAAATTTTTATTACCTCCAAGAGTGA | AAAGTAAGCCAACAAAACTTACTGTAAT | Homo sapiens | 31.03448276 | 29 | 28 |
| ATM_UserDefined (32442456)_48838294 | hg19 | chr11 | 108098322 | 108098615 | + | AGTGTGATTAGTAACCCATTATTATTTCCT | TTTCTGTTGTGATATTACTTTTGTGTGT | Homo sapiens | 30 | 30 | 28 |
| ATM_UserDefined (32442456)_48838295 | hg19 | chr11 | 108098322 | 108098615 | - | GCTTAAATTTCTCAACTTCTTTCTGAAAAT | TGCATATATAGAGAGAAAGAGGTGTGA | Homo sapiens | 26.66666667 | 30 | 27 |
| ATM_UserDefined (32442457)_48838296 | hg19 | chr11 | 108141791 | 108142133 | + | GTTCTGTTAAGCTTATAAAGTTGAACTTTT | TGGTAGGTACAGTCTATTTTGTGGTCC | Homo sapiens | 26.66666667 | 30 | 27 |
| ATM_UserDefined (32442457)_48838297 | hg19 | chr11 | 108141791 | 108142133 | - | CAAAGAACACACATTGCTGTGGTAAAAA | AAAAATCAGGTTTATCCACAAATCATTTAC | Homo sapiens | 35.71428571 | 28 | 30 |
| ATM_UserDefined (32442457)_48838298 | hg19 | chr11 | 108141791 | 108142133 | - | GAGTTAAACGGTTTAGAAAGCCAAGCC | AATGCTCCAATTACTGTAAGAAACTGT | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_UserDefined (32442458)_48838299 | hg19 | chr11 | 108143259 | 108143579 | + | TTCATATTTAACCACAGTTCTTTTCCCG | CTTACATAGTAACAGCTCACAGTTGCA | Homo sapiens | 35.71428571 | 28 | 27 |
| ATM_UserDefined (32442458)_48838300 | hg19 | chr11 | 108143259 | 108143579 | - | TGGCCCATTTTGAATAAGGATCAGCC | AATAATCATTCTGATAAAAGCACTCAGAA | Homo sapiens | 46.15384615 | 26 | 29 |
| ATM_UserDefined (32442459)_48838301 | hg19 | chr11 | 108186550 | 108186840 | + | TAGACTACGAACATATGAACACGAAGC | CAGAGGATTAGGCTAAACATTCAGGGA | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_UserDefined (32442459)_48838302 | hg19 | chr11 | 108186550 | 108186840 | - | TATGTTACTAGGGCTTTGCCCCACAT | GAAAACATGGAAACAACAAAAATTGAAGA | Homo sapiens | 46.15384615 | 26 | 29 |
| ATM_Exon (32218271)_48838303 | hg19 | chr11 | 108121428 | 108121799 | + | GAAGTTGCATTGTGTCAAGACAAGAGG | CAGCATGCATTATGTCTGACTTACAGA | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32218271)_48838304 | hg19 | chr11 | 108121428 | 108121799 | - | CTGACTTTTGTGAGCTTTCTAGGTTTG | AGGATAATTTGAAAACTATTCCATCACAT | Homo sapiens | 40.74074074 | 27 | 29 |
| ATM_Exon (32442432)_48838305 | hg19 | chr11 | 108124541 | 108124766 | + | ACCAGAATCTCAAGGAATCACTGGATC | GCAGTAAAGGGCTCTAAATTGGACAAC | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32442432)_48838306 | hg19 | chr11 | 108124541 | 108124766 | - | AGCTGTTCTGATAATCCCAGAAGACAG | ATATATAAACCAAGTACCTAATGCTTTGC | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_Exon (32083136)_48838307 | hg19 | chr11 | 108137898 | 108138069 | - | TGAGGCCTCTTATACTGCCAAATCAAT | CAAGGAAGAGAAACAAAGATGTTCAAGA | Homo sapiens | 40.74074074 | 27 | 28 |
| ATM_Exon (32452996)_48838308 | hg19 | chr11 | 108151722 | 108151895 | + | TGATAGCTGTGGTTTTATCCTGTAGCC | TTTAGGTTGGGAATATTGGAAAGCAGT | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32452996)_48838309 | hg19 | chr11 | 108151722 | 108151895 | - | GCAAACAAAGCCTGTTTTTCGCAGATA | ACACATTTCAGAAACCTCCTACTCACT | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32444329)_48838310 | hg19 | chr11 | 108160329 | 108160528 | + | GGCTTAGGAGGAGCTTGGGCCTTTGTT | CTGGGTTCAGATTATCCTTCCTCCTCA | Homo sapiens | 55.55555556 | 27 | 27 |
| ATM_Exon (32444329)_48838311 | hg19 | chr11 | 108160329 | 108160528 | - | GTGAATCAAAGTATAAATAACGTCTCGAA | GTAAAAACTTGTTTAATTAGATACTCGGC | Homo sapiens | 31.03448276 | 29 | 29 |
| ATM_Exon (32442435)_48838312 | hg19 | chr11 | 108175402 | 108175579 | - | CCATCTTTCTCTAGAACTGAGTTTACA | AGAATATACAAACAAATCCACCCAATAGT | Homo sapiens | 37.03703704 | 27 | 29 |
| ATM_Exon (32442426)_48838313 | hg19 | chr11 | 108188100 | 108188248 | + | CATGGAGGAATATGCAGTGGGACCATT | ACATACCATATATATAGTGTGTATGTGTGT | Homo sapiens | 48.14814815 | 27 | 30 |
| ATM_Exon (32442426)_48838314 | hg19 | chr11 | 108188100 | 108188248 | - | AGTCAAATTTCTTACCTGACGGAAGTG | TCAAAGATATCAACAGAAAACCAGAGAA | Homo sapiens | 40.74074074 | 27 | 28 |
| ATM_Exon (32442427)_48838315 | hg19 | chr11 | 108196037 | 108196271 | + | TCTCACCAAACACCTTGTAGAACTCTC | TTAAGCCTTGGGCAAGGGTACTTAATC | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32442427)_48838316 | hg19 | chr11 | 108196037 | 108196271 | - | GAGTGTTCTTGAAAGTTCTGGCCAGTAT | GAATGCCACGACATAGAGTTCATGTAA | Homo sapiens | 42.85714286 | 28 | 27 |
| ATM_Exon (32422463)_48838317 | hg19 | chr11 | 108196785 | 108196952 | + | GAGTCTTGCCCTGAGTATTCTCAAGCA | CAATTCTATCAGTCCATCATGTGGTCG | Homo sapiens | 48.14814815 | 27 | 27 |
| ATM_Exon (32422463)_48838318 | hg19 | chr11 | 108196785 | 108196952 | - | CAGCTGGCATCCAACTTCTTGATCAT | GAAAGCAGCAACTACCATTCATTGAGG | Homo sapiens | 46.15384615 | 26 | 27 |
| ATM_Exon (32422470)_48838319 | hg19 | chr11 | 108200941 | 108201148 | - | GGGTAGAATATTGGGCTGAGTAACACT | GAATAATTAAGGTAAAACACAAAAGCCAT | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_Exon (32442429)_48838320 | hg19 | chr11 | 108213949 | 108214098 | - | CTGACAGCTGTCAGCTTTAATAAGCCA | ACCTTTTAATTTTAAAAGGCCATTAAAGAG | Homo sapiens | 44.44444444 | 27 | 30 |
| ATM_Exon (32422468)_48838321 | hg19 | chr11 | 108115515 | 108115753 | + | CAACTGCAAATTTATATCCATCATCCG | GTGGTAGCAGATTCTGTTATGCCTTTC | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32422468)_48838322 | hg19 | chr11 | 108115515 | 108115753 | - | CCTTTTTCTTGGGTTTTGGCTCCTT | GTGAAGAAACAAACAAGCTCAACTGGG | Homo sapiens | 44 | 25 | 27 |
| ATM_Exon (32442436)_48838323 | hg19 | chr11 | 108123544 | 108123639 | - | AACCTCTTCCCTGGCTAACAGTTGCAA | ACAAAGAAAAACAAAAAGCTTCACTTAGA | Homo sapiens | 48.14814815 | 27 | 29 |
| ATM_Exon (32444314)_48838324 | hg19 | chr11 | 108139137 | 108139336 | - | GAGCTTCTTCATTTAACGTAACTCACC | ATAAAAAGCACTCATTAACACTAAGAGAA | Homo sapiens | 40.74074074 | 27 | 29 |
| ATM_Exon (32083101)_48838325 | hg19 | chr11 | 108159704 | 108159830 | + | GCAATTGTCATAAAACCAAGTTAAAAAGC | GAACAGAATCTTGTGACACTTTAGTGA | Homo sapiens | 31.03448276 | 29 | 27 |
| ATM_Exon (32083101)_48838326 | hg19 | chr11 | 108159704 | 108159830 | - | CAGGGCTTTTGGAAAGAATTTCTAAAA | AACCCCCAAATTCAAAAACCATTTTTC | Homo sapiens | 33.33333333 | 27 | 27 |
| ATM_Exon (32218248)_48838327 | hg19 | chr11 | 108165654 | 108165786 | - | GAGTCTGAACTAAGTCTGTGACAATGA | GAGAAAATGTAACTTGCTTTTAACACG | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32422459)_48838328 | hg19 | chr11 | 108170441 | 108170612 | - | GTGTGAAGTATCATTCTCCATGAATGT | AGAAATTCTATTTCAGTAGAGTCATAGGT | Homo sapiens | 37.03703704 | 27 | 29 |
| ATM_Exon (32218251)_48838329 | hg19 | chr11 | 108172375 | 108172516 | - | CATGACCCACAGCAAACAGAACTGTTT | AACAATTCAAATGCCTATCAAGTACAC | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32083117)_48838330 | hg19 | chr11 | 108180887 | 108181042 | + | GCAGAAATCTATGCAGATAAGAAAAGT | TCCTGTTCTGAAATAATAGGAATGCTT | Homo sapiens | 33.33333333 | 27 | 27 |
| ATM_Exon (32083117)_48838331 | hg19 | chr11 | 108180887 | 108181042 | - | TCCATTACCTTTTCTCTTGATCATCCA | ACAATGGTAGTGACTGCTTAACAGAAA | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32083128)_48838332 | hg19 | chr11 | 108192028 | 108192147 | - | CCGTGCCCGGCCTATAGTTTTTATTTT | AAGAAAAGACACTCTGAAAATAAAGAAGA | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_Exon (32452995)_48838333 | hg19 | chr11 | 108216470 | 108216635 | + | GTTTTCCGTTACTTCTGCATGGAAAAA | AGTGATGATGTGGTTAGTAACCAACCC | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32452995)_48838334 | hg19 | chr11 | 108216470 | 108216635 | - | TCTCAAACCAAATAGCTGGATCCAAGA | GGGGCAATAAAGATGAACAGGATACAT | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_FiveUtrExon (32422466)_48838335 | hg19 | chr11 | 108093559 | 108093913 | + | CTTGGCGTTGCTTCTTCCTCCAGAAGT | GAAATTCAGTCGTGTGCGACCCAGTCT | Homo sapiens | 51.85185185 | 27 | 27 |
| ATM_FiveUtrExon (32422466)_48838336 | hg19 | chr11 | 108093559 | 108093913 | - | TTCAAACCCTGCGTGACTGCCCAGCGCC | GCTCTCACCCACCCTCTTCGCC | Homo sapiens | 64.28571429 | 28 | 22 |
| ATM_Exon (32422474)_48838337 | hg19 | chr11 | 108106397 | 108106561 | + | TACTGGTGTGAAATATCTCAGCAACAG | TGATATCAGAGCCGGAATTACAGTTGA | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32422474)_48838338 | hg19 | chr11 | 108106397 | 108106561 | - | ACAACAACCTTCAAAACATACCTAACC | GAACAAAAATAAGACACTTGGAATGGC | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32442428)_48838339 | hg19 | chr11 | 108117691 | 108117854 | + | CAGATATCTGTCACCAGGTACAGTAAG | GATTGTCTCTAGGAAATAAGGTGACCC | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32442428)_48838340 | hg19 | chr11 | 108117691 | 108117854 | - | CAGGAAATTTCTAAATGTGACATGACCT | TTTTTATGTATGTACTCTGTTTACACTGC | Homo sapiens | 35.71428571 | 28 | 29 |
| ATM_Exon (32422473)_48838341 | hg19 | chr11 | 108122564 | 108122758 | + | AGGGTGACTTAGAAAATAGCACAGAAG | TGCTTCTGAAATAAAGGGTTGTCTCAC | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32422473)_48838342 | hg19 | chr11 | 108122564 | 108122758 | - | CTTAAATTACCTGTGAAGAATTGGAGGC | AACAAAACAAAGGACAATTGTGAAAAG | Homo sapiens | 39.28571429 | 28 | 27 |
| ATM_Exon (32442433)_48838343 | hg19 | chr11 | 108128208 | 108128333 | - | CTACAATGAGTTGTGACAATCCCACTG | ATTGTGAATAGAAGGAAAAATGCAAATTC | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_Exon (32422469)_48838344 | hg19 | chr11 | 108153437 | 108153606 | - | GAGGGGACTTGCTAAGTATTGTTAACT | GCACGAACCAAAAAGTCAATGTGTTAA | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32422462)_48838345 | hg19 | chr11 | 108154954 | 108155200 | + | GCTAATCAGATTCAAGAGGACTGGAAA | TACTTTTCATTCCCTGAATGATATGAGA | Homo sapiens | 40.74074074 | 27 | 28 |
| ATM_Exon (32422462)_48838346 | hg19 | chr11 | 108154954 | 108155200 | - | TCTTTGGAAAGCAGTCTGTTAGAAGAC | AACTGTTTTAATACCATCAGCATTAAAGT | Homo sapiens | 40.74074074 | 27 | 29 |
| ATM_Exon (32422460)_48838347 | hg19 | chr11 | 108163346 | 108163520 | + | TATACCCCTTGTGTATGAGCAGGTGGA | TACTGTGCTGAAGTAGAGAGGTAGTCA | Homo sapiens | 48.14814815 | 27 | 27 |
| ATM_Exon (32422460)_48838348 | hg19 | chr11 | 108163346 | 108163520 | - | GAGTCAGAAAATTACCTGTTTCTGAAC | AAAACAACAACAACCAGTAAGTTCTGG | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32442434)_48838349 | hg19 | chr11 | 108168014 | 108168109 | - | ACTGAATACTACAGGCAACAGAAAACA | AAAATTAGTTTAAACTAGCATGAAGACAAC | Homo sapiens | 37.03703704 | 27 | 30 |
| ATM_Exon (32442425)_48838350 | hg19 | chr11 | 108178624 | 108178711 | - | ACACTTTGCAACACCTTCACCTAAAAT | CAAACAAAAATATCTCCCCTCAGTCCC | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_Exon (32083097)_48838351 | hg19 | chr11 | 108202171 | 108202284 | - | GTATTTCCATTTCTTAGAGGGAATGGT | GACACAAAATTAAGAAACAAGGTATTTCA | Homo sapiens | 37.03703704 | 27 | 29 |
| ATM_Exon (32442430)_48838352 | hg19 | chr11 | 108203489 | 108203627 | - | GAGAGTAACACAGCAAGAAAGTAACGT | AAACATTACATAAGGAACTACCCAACA | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_Exon (32452994)_48838353 | hg19 | chr11 | 108204613 | 108204695 | - | AGGTTGAAACATATGAAATTTGCCAAT | AAAAATGAAGAGATTTTAGTGACAAACTG | Homo sapiens | 29.62962963 | 27 | 29 |
| ATM_Exon (32442431)_48838354 | hg19 | chr11 | 108206572 | 108206688 | - | CACTTCACCCAACCAAATGGCATCTTT | ACACAGAGTCATCTGAGAATAGCAAAG | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_Exon (32422457)_48838355 | hg19 | chr11 | 108225538 | 108225601 | - | GTGTGAAAAGGCTAAGTGCACAATCTC | AGAGTCAAACACATTATAAATACGGATTT | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_UserDefined (32452988)_48838356 | hg19 | chr11 | 108235759 | 108239876 | + | TGGAGCATTATGCTAGGATAGTTTCCTC | TTTGACTGGACCATGAATCCTTTGAAAG | Homo sapiens | 42.85714286 | 28 | 28 |
| ATM_UserDefined (32452988)_48838357 | hg19 | chr11 | 108235759 | 108239876 | + | TTTTAAGAAGGTCCTGTTGTCAGTTTTTC | ACAAGTGAATTTGCTCATACAGCAGGC | Homo sapiens | 34.48275862 | 29 | 27 |
| ATM_UserDefined (32452988)_48838358 | hg19 | chr11 | 108235759 | 108239876 | + | GATCTTCAGTATATGAATTACCCTTTCATT | TTTCACTCTTTAGAAATAATGGTCATTCGG | Homo sapiens | 30 | 30 | 30 |
| ATM_UserDefined (32452988)_48838359 | hg19 | chr11 | 108235759 | 108239876 | + | TTCCTGCTGTTTATATCTGTTGATTTTTG | CTGAAGTGTAGCATAAATACTGATAGGAG | Homo sapiens | 31.03448276 | 29 | 29 |
| ATM_UserDefined (32452988)_48838360 | hg19 | chr11 | 108235759 | 108239876 | + | ATTTGATTGATGCCTTTTTCACTGAGAG | GCTCACATATTCACCTCTCTGGTTTTTC | Homo sapiens | 35.71428571 | 28 | 28 |
| ATM_UserDefined (32452988)_48838361 | hg19 | chr11 | 108235759 | 108239876 | + | TTCTACATTTTATCCAGCATTTCTCTGTG | GGAAAATGTTCATCCCAGCTGCGGAGATTA | Homo sapiens | 34.48275862 | 29 | 30 |
| ATM_UserDefined (32452988)_48838362 | hg19 | chr11 | 108235759 | 108239876 | + | TTTCTTATGAAGAGTTGGCATTTCTTTT | TCTTACAAGCTGCCTAAATGAATATTTGG | Homo sapiens | 28.57142857 | 28 | 29 |
| ATM_UserDefined (32452988)_48838363 | hg19 | chr11 | 108235759 | 108239876 | + | TCATGGGTGAAATTAGAAATTATCAACTAG | GGAGAAATAACTAATTCACAGATGACAG | Homo sapiens | 30 | 30 | 28 |
| ATM_UserDefined (32452988)_48838364 | hg19 | chr11 | 108235759 | 108239876 | + | TGGAACTCTTTCTGCAAATGACTAAGA | AGTTGTCCAAGGCAAGAAGATAGTAAA | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_UserDefined (32452988)_48838365 | hg19 | chr11 | 108235759 | 108239876 | + | TAAGGGAGATAATAGCTTTCCCACCCT | ATATAGAGATGTGTCCCAATTTCAAGT | Homo sapiens | 44.44444444 | 27 | 27 |
| ATM_UserDefined (32452988)_48838366 | hg19 | chr11 | 108235759 | 108239876 | + | AAATATTTGTGTGTTCTGTGTGCCAGG | TTGGGCCTAGAAAGTGATCTAGTTGAC | Homo sapiens | 40.74074074 | 27 | 27 |
| ATM_UserDefined (32452988)_48838367 | hg19 | chr11 | 108235759 | 108239876 | - | GAACAGTTTAAAGGCCTTGGGAATAAGAA | CATATAGAAGGACCTAAACAGAGAGAAC | Homo sapiens | 37.93103448 | 29 | 28 |
| ATM_UserDefined (32452988)_48838368 | hg19 | chr11 | 108235759 | 108239876 | - | CCTACTTAAAGTATGTTGGCAGGTTAAAAA | CCACTCCTTTCAGTTTCTCTTGTAGTCT | Homo sapiens | 33.33333333 | 30 | 28 |
| ATM_UserDefined (32452988)_48838369 | hg19 | chr11 | 108235759 | 108239876 | - | GACAAAACTGGATGAACAGCCTATCAAAT | AACAATTTCTGAAAGAAGAAGCCCAAT | Homo sapiens | 37.93103448 | 29 | 27 |
| ATM_UserDefined (32452988)_48838370 | hg19 | chr11 | 108235759 | 108239876 | - | TAAATGTACCTCCTTCCACCCCTGCCATAA | ATAAGAATGGGCTTGTGTGCAGCAAC | Homo sapiens | 46.66666667 | 30 | 26 |
| ATM_UserDefined (32452988)_48838371 | hg19 | chr11 | 108235759 | 108239876 | - | TTTTAGGGGAGGGGAAGGTGAGGAGATCA | CTCACTAATACACGAGTCAGTAGTTTTA | Homo sapiens | 51.72413793 | 29 | 28 |
| ATM_UserDefined (32452988)_48838372 | hg19 | chr11 | 108235759 | 108239876 | - | CTGGATGCAACAACTAGCAAGAATTAG | AAAGATACACCAAGGTAAGTTTGTAATTT | Homo sapiens | 40.74074074 | 27 | 29 |
| ATM_UserDefined (32452988)_48838373 | hg19 | chr11 | 108235759 | 108239876 | - | CTCCTCATTTGTCCTTGGCAGTTTTCT | AGGCTTGTGATCTGTATGTAAATTTGATA | Homo sapiens | 44.44444444 | 27 | 29 |
| ATM_UserDefined (32452988)_48838374 | hg19 | chr11 | 108235759 | 108239876 | - | TTTGGGGAGGTATGACCTGCACAAAG | ATTTCTCCAAATTCTGATTTCTCCAGC | Homo sapiens | 50 | 26 | 27 |
| ATM_UserDefined (32452988)_48838375 | hg19 | chr11 | 108235759 | 108239876 | - | TGCTTTCTCCTTAGAACGAGTCCCAT | TCAAACCAATATACTGGCTTTCTTTTTAA | Homo sapiens | 46.15384615 | 26 | 29 |
| ATM_UserDefined (32408355)_48838376 | hg19 | chr11 | 108199698 | 108200015 | + | AGCTAAGAAATGGAAAAATGAAGGCAT | TGCATAGAAAGAGTGACTTGGTCTTTT | Homo sapiens | 33.33333333 | 27 | 27 |
| ATM_UserDefined (32408355)_48838377 | hg19 | chr11 | 108199698 | 108200015 | - | GTATCTGAAAACCGGGCTAATGAGAGA | GAACTATTGCAAACTTGTCCAACCCA | Homo sapiens | 44.44444444 | 27 | 26 |
| ATM_UserDefined (32452989)_48838378 | hg19 | chr11 | 108224443 | 108224657 | + | TGTCTTCAGAAGGTAAGTGATATGAAG | ACACACACGGGTTAGAGATTTGTATAG | Homo sapiens | 37.03703704 | 27 | 27 |
| ATM_UserDefined (32452989)_48838379 | hg19 | chr11 | 108224443 | 108224657 | - | TTGACATCAAAAATTATTTCCCTCCTTT | CTTTCCAGTTAGAATGTGAACTTTACAA | Homo sapiens | 28.57142857 | 28 | 28 |
| SF3B1_UserDefined (32442466)_48838380 | hg19 | chr2 | 198266124 | 198266854 | + | TAAATCAACTGACCTGAAATGAAGAGA | TCTCGAATAAGGATTAACATCACTTCTC | Homo sapiens | 33.33333333 | 27 | 28 |
| SF3B1_UserDefined (32442466)_48838381 | hg19 | chr2 | 198266124 | 198266854 | + | TGGTGGATTTACCTTTCCTCTGTGTTG | TTCAGCCAAACTGCAGAATATGTTCAC | Homo sapiens | 44.44444444 | 27 | 27 |
| SF3B1_UserDefined (32442466)_48838382 | hg19 | chr2 | 198266124 | 198266854 | - | TTTTAAAATCACGTAATCAGCAATGAGTA | AAAGAATATTCTCAAAAGTCCTTTCAACT | Homo sapiens | 27.5862069 | 29 | 29 |
| SF3B1_UserDefined (32442466)_48838383 | hg19 | chr2 | 198266124 | 198266854 | - | GTTAAAGCCTTTATGGAAGGGTATCCG | TTCCAGTCTCCTGATGAGGAAATGAAA | Homo sapiens | 44.44444444 | 27 | 27 |
| SF3B1_Exon (32442437)_48838384 | hg19 | chr2 | 198265439 | 198265660 | + | CTGGATATAATTTCTGCTGCACCTACT | GTTGAAGAGAAAAGTGACCAAACATCG | Homo sapiens | 40.74074074 | 27 | 27 |
| SF3B1_Exon (32442437)_48838385 | hg19 | chr2 | 198265439 | 198265660 | - | TGATACTACTGTGGAGTTGGCAAACA | TGAAAACTATTATGGCACATTGCTACT | Homo sapiens | 42.30769231 | 26 | 27 |
| XPO1_Exon (32442439)_48838386 | hg19 | chr2 | 61715300 | 61715406 | - | TCCCAATAAATTTTGTGAACTTTTGGC | GTGAGAGATGCTTTGCTTTTGATTTCT | Homo sapiens | 33.33333333 | 27 | 27 |
| XPO1_Exon (32218516)_48838387 | hg19 | chr2 | 61715723 | 61715906 | - | CTGTTCCCAGCCGTGGATAATTCTTAT | GGTGATTCTGATTTCTTACACAATGCAG | Homo sapiens | 44.44444444 | 27 | 28 |
| XPO1_Exon (32442440)_48838388 | hg19 | chr2 | 61717777 | 61717911 | - | GCTTGTGTGTGGTTTTCCTTATTAGGG | TCTACCTTTTCTTAAAAGTTTTCATGGG | Homo sapiens | 44.44444444 | 27 | 28 |
| XPO1_Exon (32218512)_48838389 | hg19 | chr2 | 61721029 | 61721226 | + | GTTTCTTCTACTTCAGATACCAACAAC | TGCCCTAAAGCACTTCAAAAATCTCTC | Homo sapiens | 37.03703704 | 27 | 27 |
| XPO1_Exon (32218512)_48838390 | hg19 | chr2 | 61721029 | 61721226 | - | GGTCTTCTGTGTAGGCCCTTCATTATA | TAACAGAGCGGTTGGTTGAGTGTTCT | Homo sapiens | 44.44444444 | 27 | 26 |
| XPO1_Exon (32442438)_48838391 | hg19 | chr2 | 61722590 | 61722748 | - | AAATGCGAACTGATCGTCCCTTTTAAA | GATATGTGGAGCTTTTGCTCCTAAAAT | Homo sapiens | 37.03703704 | 27 | 27 |
| XPO1_Exon (32218514)_48838392 | hg19 | chr2 | 61724014 | 61724142 | - | TTTTAAGTAAAGTGTGTATCAGGCCAA | TCTTGAAAATTGGCTGAATAATTTTAGTG | Homo sapiens | 33.33333333 | 27 | 29 |
| DDX3X_UserDefined (32442453)_48838393 | hg19 | chrX | 41204433 | 41204801 | + | GATGTTTAGTGCTACTTTTCCTAAGGA | TGACCTCCTAAATGCAACAGGTAACA | Homo sapiens | 37.03703704 | 27 | 26 |
| DDX3X_UserDefined (32442453)_48838394 | hg19 | chrX | 41204433 | 41204801 | - | ATTTGCAACATCAAACAGTACCTGTAT | TCAACATGGTTTTAGTATGGCTGTTAG | Homo sapiens | 33.33333333 | 27 | 27 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| DDX3X_UserDefined_(32442453)_48838395 | hg19 | chrX | 41204433 | 41204801 | - | CACACTAAGTTTGGCATCAACATGGTA | GCAGAAATGACCGTTTGTCTGATTCTT | Homo sapiens | 40.74074074 |
| DDX3X_UserDefined_(32442507)_48838396 | hg19 | chrX | 41202990 | 41203652 | + | AGGGCTTTCTAAATGATGCTAAGACTT | CCAACGAGAGAGTTGGCAGTACAGAT | Homo sapiens | 37.03703704 |
| DDX3X_UserDefined_(32442507)_48838397 | hg19 | chrX | 41202990 | 41203652 | + | GTTTTCATACCGATCTAGAGTTCGTCC | TGTTCTTTTGTTTTTGGCAGAGTTGCG | Homo sapiens | 44.44444444 |
| DDX3X_UserDefined_(32442507)_48838398 | hg19 | chrX | 41202990 | 41203652 | - | CACTTGAAAACTGGTTTTCATTCTACT | GTTCAAAAACAAACAAACAAATCCTGCTC | Homo sapiens | 33.33333333 |
| DDX3X_UserDefined_(32442507)_48838399 | hg19 | chrX | 41202990 | 41203652 | - | AATATCGGCACCACCATAAACCACGCA | GCTAATACCAAGGAGATTGGGTATTGT | Homo sapiens | 48.14814815 |
| DDX3X_UserDefined_(32442509)_48838400 | hg19 | chrX | 41206565 | 41207022 | + | AACTTCATAGCTTTGGGAAGGGTTTTT | ATGGAGGAAATTATAACTCCCAGGGGG | Homo sapiens | 37.03703704 |
| DDX3X_UserDefined_(32442509)_48838401 | hg19 | chrX | 41206565 | 41207022 | - | AGGCTTTGAAAAAGATGACTAAAAGGA | AAATTATGTTGGTTCACATGAAGATGAA | Homo sapiens | 33.33333333 |
| DDX3X_UserDefined_(32442509)_48838402 | hg19 | chrX | 41206565 | 41207022 | - | GGGCTGCTGATTGTTCCTTTCATTTCT | TCCATCACTGTTGTAAAAGCCTCCATA | Homo sapiens | 44.44444444 |
| EGR2_UserDefined_(32442464)_48838403 | hg19 | chr10 | 64575621 | 64575839 | + | ATTTGCTCCTCGCACAACCTGGAGA | CTATTTGCCACTGACTCTCTCCTGTCT | Homo sapiens | 52 |
| EGR2_UserDefined_(32442464)_48838404 | hg19 | chr10 | 64575621 | 64575839 | - | TGACACTCCAGGTAGCGAGGGAGTTGG | TAAGCCGCGTGTCTTCGCAGGCCCGG | Homo sapiens | 59.25925926 |
| EGR2_FiveUtrExon_(32453018)_48838405 | hg19 | chr10 | 64578301 | 64578407 | - | TTGCTGCCGGTGGGTGCCTCACCTAAG | AAGAACGCAGCTATCTTTTCTTGCTTT | Homo sapiens | 62.96296296 |
| SETD2_Exon_(32442490)_48838406 | hg19 | chr3 | 47084051 | 47084190 | - | GGGCTATTTTGTCCTATTCAGCCAGAT | TAAGAGGAGCTGTATGGGCGTTCCCA | Homo sapiens | 44.44444444 |
| SETD2_Exon_(32442492)_48838407 | hg19 | chr3 | 47108560 | 47108608 | - | TTTTTCTTGTTTTTAAGGGGCCAGGAT | GGTATTCAGAGATACTCTGTGTTGATGG | Homo sapiens | 37.03703704 |
| SETD2_Exon_(32442484)_48838408 | hg19 | chr3 | 47129603 | 47129737 | - | GGTTGAGCAACTTTTTGGGTTTTGGT | TTTTTATTTCTGGTTTTGTGGTGGGAA | Homo sapiens | 40.74074074 |
| SETD2_Exon_(32442488)_48838409 | hg19 | chr3 | 47147487 | 47147610 | - | TGCCTGAGCATATATGTTTTCCCTACT | GCAGCTTGGATATGTTTGCCTTGAAAC | Homo sapiens | 40.74074074 |
| SETD2_Exon_(32442488)_48838410 | hg19 | chr3 | 47161672 | 47166038 | + | TAACATGGCATTTTCCCTTGCTTGGC | TTTTTAAGAGGCCCTCTATCTTTGATATC | Homo sapiens | 46.15384615 |
| SETD2_Exon_(32442488)_48838411 | hg19 | chr3 | 47161672 | 47166038 | + | GTCCTTAATGGAATTGCTGCTTATTTCAGG | AGGTCTACCTGATCTTGGATCCCAGTA | Homo sapiens | 40 |
| SETD2_Exon_(32442488)_48838412 | hg19 | chr3 | 47161672 | 47166038 | + | GAAGGCTTTTCTTGAGAGAAGTCCCA | GCAGCTCTTCTGCTGCCTAGAAGGTATT | Homo sapiens | 46.15384615 |
| SETD2_Exon_(32442488)_48838413 | hg19 | chr3 | 47161672 | 47166038 | + | TCACATCTGTATGACTTGTACTATCAACCC | GTCTCCTTGACTCCAATCTCTCATCTTC | Homo sapiens | 40 |
| SETD2_Exon_(32442488)_48838414 | hg19 | chr3 | 47161672 | 47166038 | + | TTTCTTCCATGGGCAAAGTAGAATTCT | ATAAGTTACACCATCACTGTCTTCCATGG | Homo sapiens | 37.03703704 |
| SETD2_Exon_(32442488)_48838415 | hg19 | chr3 | 47161672 | 47166038 | + | TCTAAAGAGATTTCTGGTCTTCCTCTTC | TTTTTACTCTTTAGCACTGCATCCAGA | Homo sapiens | 39.28571429 |
| SETD2_Exon_(32442488)_48838416 | hg19 | chr3 | 47161672 | 47166038 | + | GATGAAGCAATACCTGAACTCCCAATAGG | TAACTGAAGGCTCAATTATTTTCAGCTTCAG | Homo sapiens | 44.82758621 |
| SETD2_Exon_(32442488)_48838417 | hg19 | chr3 | 47161672 | 47166038 | + | TTTGGTTTTGCAGCAAGAAACCCTCG | ATATTCTGAAATCCATTTGATGAAAGCATG | Homo sapiens | 46.15384615 |
| SETD2_Exon_(32442488)_48838418 | hg19 | chr3 | 47161672 | 47166038 | + | TCAGATTCTGCCCCAGGAGATCCATTTA | GCTGGAGACCCAGCCTTTTCTCTTTCA | Homo sapiens | 46.42857143 |
| SETD2_Exon_(32442488)_48838419 | hg19 | chr3 | 47161672 | 47166038 | + | TTCTTCATTTAATTCTGTACAACAGAAAGA | TTCTAGTTTTGAAGAATACTTGCCTCTTC | Homo sapiens | 26.66666667 |
| SETD2_Exon_(32442488)_48838420 | hg19 | chr3 | 47161672 | 47166038 | + | CGCCTTGAGTATGTCTTCTTATACTCTTC | AGTGAGATCTGCTCCGCCGTCGCTCTCTTT | Homo sapiens | 41.37931034 |
| SETD2_Exon_(32442488)_48838421 | hg19 | chr3 | 47161672 | 47166038 | + | CTTTTTAAAGGTGCTGAGCTCTTTTTA | TGAGGAAATATCTGCTTGCTCATTCAA | Homo sapiens | 33.33333333 |
| SETD2_Exon_(32442488)_48838422 | hg19 | chr3 | 47161672 | 47166038 | + | AACTGTTATTGGTGTATGTGGCAAGGC | TAAAATGAATTTTGCCCAATTCCACCC | Homo sapiens | 44.44444444 |
| SETD2_Exon_(32442488)_48838423 | hg19 | chr3 | 47161672 | 47166038 | + | TTGCCAAGTGCAGTGAGAAACCTATTC | AGTTACTTTCAAATGGACTGCACAGTT | Homo sapiens | 44.44444444 |
| SETD2_Exon_(32442488)_48838424 | hg19 | chr3 | 47161672 | 47166038 | - | ACTTCAGGGACCCACAGCGATGGAAG | AGAGTTGTGGGATAAAGTAAATTTAAAAGT | Homo sapiens | 57.69230769 |
| SETD2_Exon_(32442488)_48838425 | hg19 | chr3 | 47161672 | 47166038 | - | CAGACCAGTCCGATAAATTTCTTCTATCC | AAGTGATGGTGAGCTTCAGGACAGAA | Homo sapiens | 41.37931034 |
| SETD2_Exon_(32442488)_48838426 | hg19 | chr3 | 47161672 | 47166038 | - | GCACTCATCAGAAGAGCTCAGAAACTTAG | ACTGGGGTTGTGTATGATCGAACTCA | Homo sapiens | 44.82758621 |
| SETD2_Exon_(32442488)_48838427 | hg19 | chr3 | 47161672 | 47166038 | - | GTCCAGTAGAAAACAAATAGATAATCGCC | CAAATAAGTCTTGGCAACAGACCACTTT | Homo sapiens | 37.93103448 |
| SETD2_Exon_(32442488)_48838428 | hg19 | chr3 | 47161672 | 47166038 | - | GATTCGGATGATAGCAGTATTCCAAGAAA | GCAAGTAAAGCCTGTCCTCAAACTGAT | Homo sapiens | 37.93103448 |
| SETD2_Exon_(32442488)_48838429 | hg19 | chr3 | 47161672 | 47166038 | - | CAAGAAGAAGGGAATTCCATATTGCCT | CCAGAAATTGTGTCTACAGTTCATGAAGAT | Homo sapiens | 40.74074074 |
| SETD2_Exon_(32442488)_48838430 | hg19 | chr3 | 47161672 | 47166038 | - | TCAGCTTCGTTAATCATTTTGATGATTT | AAGAAACAATAGTAGAAGTAGGTAGTGAC | Homo sapiens | 27.5862069 |
| SETD2_Exon_(32442488)_48838431 | hg19 | chr3 | 47161672 | 47166038 | - | GTCTATGCCAGTTATGACTGTGGATTATT | GTGCATTTGGAATCAAAACCAGTTATATG | Homo sapiens | 37.93103448 |
| SETD2_Exon_(32442488)_48838432 | hg19 | chr3 | 47161672 | 47166038 | - | GATCAATTAAGAAGTTTTTGTCCCATAGA | GGTGCAAAGAAAAAGACTTGGATGATACC | Homo sapiens | 31.03448276 |
| SETD2_Exon_(32442488)_48838433 | hg19 | chr3 | 47161672 | 47166038 | - | GCTTCCCGTTATAAATCTACCCTTTCAAAA | CATCAAATCGATTAAATGATTCACCTACTT | Homo sapiens | 36.66666667 |
| SETD2_Exon_(32442488)_48838434 | hg19 | chr3 | 47161672 | 47166038 | - | TAACAGAGCACGAGAGAGTTCTGACTCA | CTAAAAGGACTTCAGAAAATGAAGCAATT | Homo sapiens | 46.42857143 |
| SETD2_Exon_(32442488)_48838435 | hg19 | chr3 | 47161672 | 47166038 | - | CAGCAAGCATTGAAAAGGAAAGAGATT | GGTCTGAGAGAGGCTCTAGAACTAATT | Homo sapiens | 37.03703704 |
| SETD2_Exon_(32442488)_48838436 | hg19 | chr3 | 47161672 | 47166038 | - | ATCACCAGCACCAGTAACAGAGCCAGT | AAAAAGAAGATTCCCATATTGGGAAGG | Homo sapiens | 51.85185185 |
| SETD2_Exon_(32442488)_48838437 | hg19 | chr3 | 47161672 | 47166038 | - | GCTTCAGCCTTACAAAGAAAACTTTGC | GAAACATCTGCTTCATGTAACATCCAG | Homo sapiens | 40.74074074 |
| BIRC3_Exon_(32105204)_48838438 | hg19 | chr11 | 102198783 | 102198861 | - | CTCTTGTCTTTTCACTCAAATAGCACT | TTGAGGGATTATTTATACAAGTGAAATGAA | Homo sapiens | 37.03703704 |
| BIRC3_Exon_(32442502)_48838439 | hg19 | chr11 | 102199628 | 102199676 | - | ACCAACACAGTATATTTGAAAATGCATT | CTTACCCATTGCAATGAAAATGACCCA | Homo sapiens | 28.57142857 |
| BIRC3_Exon_(32442503)_48838440 | hg19 | chr11 | 102201730 | 102201972 | + | ATGCAGAAGATGAAATAAGGGAAGAGG | TTTGAGTTGCTCATAAGTCAAAAATGT | Homo sapiens | 40.74074074 |
| BIRC3_Exon_(32442503)_48838441 | hg19 | chr11 | 102201730 | 102201972 | - | TCTTTTTCCTCAGTTGCTCTTTCTCTC | TGGAACAAAGAATTTAACACATATTAGGAT | Homo sapiens | 40.74074074 |
| BIRC3_FiveUtrExon_(32442501)_48838442 | hg19 | chr11 | 102188194 | 102188302 | - | ACATCTCTTACACCAGGGGAGGCGCG | TTTTAAATGCGTCACCCAAATCCCCACCCC | Homo sapiens | 61.53846154 |
| BIRC3_UserDefined_(32442463)_48838443 | hg19 | chr11 | 102192568 | 102196296 | + | TATAAATCAAAACATCTCAAAATGGAGACC | GGAAATAAAGGAAAAGTGATTCTAGCTGGG | Homo sapiens | 30 |
| BIRC3_UserDefined_(32442463)_48838444 | hg19 | chr11 | 102192568 | 102196296 | + | ATAGTACCTGTGTTTGGGAAAGATTTT | AAATGTGATGATTGGTCAAGAAATTATCC | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838445 | hg19 | chr11 | 102192568 | 102196296 | + | TGAGAATGATGGTTGAAGGTTACATTT | CACCTGGAAACAAAGCATTGAAGTCTG | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838446 | hg19 | chr11 | 102192568 | 102196296 | + | AAGCCACAGTGACTTGCTTATTGGTC | AACTAAATTGATGCAATTTGATTATCCATC | Homo sapiens | 46.15384615 |
| BIRC3_UserDefined_(32442463)_48838447 | hg19 | chr11 | 102192568 | 102196296 | + | TTCAAAAATGTTTTCAGGAGGTGATAAG | AAGTTCAGTTTTGTTATTGGTAGTTTGGGC | Homo sapiens | 32.14285714 |
| BIRC3_UserDefined_(32442463)_48838448 | hg19 | chr11 | 102192568 | 102196296 | + | AACTTAACATTGAGTTGCTTCAACAGC | TAAGATGTGTTAGCCAGTCCTGTTACT | Homo sapiens | 37.03703704 |
| BIRC3_UserDefined_(32442463)_48838449 | hg19 | chr11 | 102192568 | 102196296 | + | GACTAATTGGAGAAAAATTGGGGATATATC | AATGTATTATTTTCCTCCTTTGAGTTAGG | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838450 | hg19 | chr11 | 102192568 | 102196296 | + | GCTGTGAAACACTTGACAGCAGAGCTT | TCAACACAGCTTACTCTGTAGCATCATG | Homo sapiens | 48.14814815 |
| BIRC3_UserDefined_(32442463)_48838451 | hg19 | chr11 | 102192568 | 102196296 | + | TTAGAATCTCCATGTTGAAACTCTAAATGC | ATTTGATGAAAAGCGCCAACACGTTTG | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838452 | hg19 | chr11 | 102192568 | 102196296 | + | TTGCTCGTGCTGGTTTCTATTACACTG | TTCAGTAACAAATTCCACACACTCATTAC | Homo sapiens | 44.44444444 |
| BIRC3_UserDefined_(32442463)_48838453 | hg19 | chr11 | 102192568 | 102196296 | + | TTTTTCTGCCTTGATGAGAAGTTCCTACC | TAATGCTATGTCAGAACACCTGAGACA | Homo sapiens | 41.37931034 |
| BIRC3_UserDefined_(32442463)_48838454 | hg19 | chr11 | 102192568 | 102196296 | + | TAGTGTTCTAGTTAATCCTGAGCAGCT | CATGCCAAGTGGTTTCCAAGGTAATTG | Homo sapiens | 40.74074074 |
| BIRC3_UserDefined_(32442463)_48838455 | hg19 | chr11 | 102192568 | 102196296 | - | CGAGATTAGACTAAGTCCCTTTAAGGATTT | AGAGTACTAGTAATTACTGATCGTTTACC | Homo sapiens | 36.66666667 |
| BIRC3_UserDefined_(32442463)_48838456 | hg19 | chr11 | 102192568 | 102196296 | - | GCTCTTTGTTCCAAATAACTGTCACTCAT | TGGTATACCTTATATAACTCGATCATGTT | Homo sapiens | 37.93103448 |
| BIRC3_UserDefined_(32442463)_48838457 | hg19 | chr11 | 102192568 | 102196296 | - | AATCTTCTTTGTATTCATCACTGTCTTTAT | CTTTTTACTTCCTCTAAATTTGTTGCACAT | Homo sapiens | 26.66666667 |
| BIRC3_UserDefined_(32442463)_48838458 | hg19 | chr11 | 102192568 | 102196296 | - | GTAAAGAGGTCTGAGTCGATAATACTAG | ACAATTATTTCCCTCAAAGGTAAGTAATTG | Homo sapiens | 37.93103448 |
| BIRC3_UserDefined_(32442463)_48838459 | hg19 | chr11 | 102192568 | 102196296 | - | CCTCAAAGAACTAACATTGTAGAGTTATTC | GGTCAGAAATTATTTCATGGATAGCATACC | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838460 | hg19 | chr11 | 102192568 | 102196296 | - | AATCAGAGATTAATGTGTTTGTTCATTTG | TGCTATATTCAGGTTACTTTTGGAAAAAC | Homo sapiens | 27.5862069 |
| BIRC3_UserDefined_(32442463)_48838461 | hg19 | chr11 | 102192568 | 102196296 | - | CAACTGAAGACATTTTGAATTCAGTGAAAT | GTTATTTCCCATAGATCATCAGGAATATGA | Homo sapiens | 30 |
| BIRC3_UserDefined_(32442463)_48838462 | hg19 | chr11 | 102192568 | 102196296 | - | GCATTCAATCAACCCAAAGATGTAATGTG | TTTCACTGATTCTAAGATGGCTATTTTAAA | Homo sapiens | 37.93103448 |
| BIRC3_UserDefined_(32442463)_48838463 | hg19 | chr11 | 102192568 | 102196296 | - | CAGTTTTAATACTAGGCTGAAAAGCCAAAA | AACCCTTCACAGAAGAGTGATGTGTAT | Homo sapiens | 33.33333333 |
| BIRC3_UserDefined_(32442463)_48838464 | hg19 | chr11 | 102192568 | 102196296 | - | CAGAAGCATTTGACCTTGTCATTCACA | GTTCATAATGAAATGAATGGGGAAGAAA | Homo sapiens | 40.74074074 |
| BIRC3_UserDefined_(32442463)_48838465 | hg19 | chr11 | 102192568 | 102196296 | - | CTGGCATTTTCGTTATTCATTGCACAG | GCTTCCAAGTTGTTAACGGAATTTAGACTC | Homo sapiens | 40.74074074 |
| BIRC3_UserDefined_(32442463)_48838466 | hg19 | chr11 | 102192568 | 102196296 | - | CCCACATAATAAAAACCCGCACTTGC | CCTTCGGTTCCCAATTGCTCAATTTTC | Homo sapiens | 46.15384615 |
| BIRC3_UserDefined_(32442463)_48838467 | hg19 | chr11 | 102192568 | 102196296 | - | CTTCGTTGTCTTAAGTCCATCGAAACT | TGAACCCATGGATCATCTCCAGATTCC | Homo sapiens | 40.74074074 |
| BIRC3_UserDefined_(32408352)_48838468 | hg19 | chr11 | 102206647 | 102207001 | + | TGTTATTAAACAGAAGACACAGACGTC | AATAATTCAGGCTGAATGTGGTGACTC | Homo sapiens | 37.03703704 |
| BIRC3_UserDefined_(32408352)_48838469 | hg19 | chr11 | 102206647 | 102207001 | - | CGTATCAATCAGTTCTCTTGCTTGTAA | CCTTCCAGTCCTATATATATAGCAGGAT | Homo sapiens | 37.03703704 |
| NFKBIE_ThreeUtrExon_(32442505)_48838470 | hg19 | chr6 | 44226908 | 44227021 | - | GTTCCCTGATGGTCTTCAACTGTCATT | GAGCCATAACTGCTGCAGTTTGG | Homo sapiens | 44.44444444 |
| NFKBIE_Exon_(32129086)_48838471 | hg19 | chr6 | 44229363 | 44229585 | + | TGGTCCAGATGTACAGCCAGATGGAGT | GGTGAAGAAAGTAGCTCTCTGCTTTTC | Homo sapiens | 51.85185185 |
| NFKBIE_Exon_(32129086)_48838472 | hg19 | chr6 | 44229363 | 44229585 | - | ATGACCCTCTATTTGTGCCTCCAGACAG | TGTGGACAGTCTTTAAGACTGCCCTT | Homo sapiens | 50 |
| NFKBIE_Exon_(32442506)_48838473 | hg19 | chr6 | 44230297 | 44230399 | - | CCTACCCTATGGGTAGCTAGTTAAAGT | TTGGTGGCGAGAGAATGTTATGCCCA | Homo sapiens | 44.44444444 |
| NFKBIE_UserDefined_(32442504)_48838474 | hg19 | chr6 | 44232719 | 44233525 | + | TAGGTGGAATCAGCCCGCTCCCCAT | CGCGCCTCCGACATGCCCGCGGCTCT | Homo sapiens | 60 |
| NFKBIE_UserDefined_(32442504)_48838475 | hg19 | chr6 | 44232719 | 44233525 | + | CAGGTCCACCCAGCGGTTACTGT | TTCATTCATTCACTCTTTCCCCCTTCC | Homo sapiens | 60.89656522 |
| NFKBIE_UserDefined_(32442504)_48838476 | hg19 | chr6 | 44232719 | 44233525 | - | AAGGAACCACAGGAGAAGGAAGACGCG | TGAGTATGGGGGCTAGGCTGAATCGG | Homo sapiens | 55.55555556 |
| NFKBIE_UserDefined_(32442504)_48838477 | hg19 | chr6 | 44232719 | 44233525 | - | AGGACCCGCTTGGGCGGTTCGGCTGCC | AAGGGGCCGGACGAGGCGGAGGAGA | Homo sapiens | 74.07407407 |
| MED12_UserDefined_(32442514)_48838478 | hg19 | chrX | 70339223 | 70340020 | + | CCTAAGTGGCTGAGTTTGCCTTCAT | CCAGAGTGCCATTAACACTTGGTTCA | Homo sapiens | 48 |
| MED12_UserDefined_(32442514)_48838479 | hg19 | chrX | 70339223 | 70340020 | + | GAGACCTCACTATTTGCAATGTCCATC | TGGAACCATCCTCCTTCTTAATCTAGA | Homo sapiens | 44.44444444 |
| MED12_UserDefined_(32442514)_48838480 | hg19 | chrX | 70339223 | 70340020 | - | GGGCAAATGCCAATGAGATAGTATTAGGT | TGAGGCAGGAAAGCGGCGTTTAGTTG | Homo sapiens | 41.37931034 |
| MED12_UserDefined_(32442514)_48838481 | hg19 | chrX | 70339223 | 70340020 | - | CTTCCTTCTTACTGAAAATGGGGACCT | ATCGTCAGTCACCAGCCAGAAGTTAT | Homo sapiens | 44.44444444 |
| MED12_FiveUtrExon_(32442511)_48838482 | hg19 | chrX | 70338555 | 70338703 | - | ATTCCCAGCCGTCAGTGCCTCCTCCTA | GAACGGGAAGGGGGAGAGGGAG | Homo sapiens | 59.25925926 |
| NBEAL2_Exon_(32442513)_48838483 | hg19 | chr3 | 47035427 | 47035510 | - | AGTGAGCTTTAGGGCAGGGCAGTCAC | TTGGGGTGGGGAGGGGCCATCAGGAG | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined_(32442467)_48838484 | hg19 | chr3 | 47030159 | 47031083 | + | GTGCAGAGGTCCCGCTGCTACCACT | AGGCAGCCGGGGCCAAGTGCTAGT | Homo sapiens | 64 |
| NBEAL2_UserDefined_(32442467)_48838485 | hg19 | chr3 | 47030159 | 47031083 | + | GTTGGAGAATGTGGCCCTACATGCTCT | CAAATACAAGTTCCCTCCTGCTGCTT | Homo sapiens | 51.85185185 |
| NBEAL2_UserDefined_(32442467)_48838486 | hg19 | chr3 | 47030159 | 47031083 | - | TGTTCGGCCAGCACATGCAGCTCATC | TAGGCAGAACAGGGGAACAGTGGTGG | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined_(32442467)_48838487 | hg19 | chr3 | 47030159 | 47031083 | - | TCAAAGAGGCCCTCGCAGAGAAG | TATGTTCTCCAGGTTCCTGGGGGAGTC | Homo sapiens | 56.52173913 |
| NBEAL2_UserDefined_(32442467)_48838488 | hg19 | chr3 | 47030159 | 47031083 | - | CCTAAACTCTTGCTTGGTCTTTGGTTC | TCTTCTCCTTGGAGCTGATGACTTCCC | Homo sapiens | 44.44444444 |
| NBEAL2_UserDefined_(32453015)_48838489 | hg19 | chr3 | 47039557 | 47041894 | + | GAGGGCAGCGGGCCCCTCCTGTACCTA | CATGGGGCGGCTGAGGGGACTATGA | Homo sapiens | 70.37037037 |
| NBEAL2_UserDefined_(32453015)_48838490 | hg19 | chr3 | 47039557 | 47041894 | + | CAGCATAGTTCGGGAGCACAGACAGAA | TTTGGGACCTGGCTGGGTGTGCCTACT | Homo sapiens | 51.85185185 |
| NBEAL2_UserDefined_(32453015)_48838491 | hg19 | chr3 | 47039557 | 47041894 | + | GGTCACGCAGACCATGCTGAGCTTTTT | TTGGAGCCAGGGAACCTCGAAGTGCTA | Homo sapiens | 51.85185185 |
| NBEAL2_UserDefined_(32453015)_48838492 | hg19 | chr3 | 47039557 | 47041894 | + | ACAGCAGGCCTTGGGCCAGCGGATGAA | GTCTCCAGGGTCTGGTTGCCTGCTT | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined_(32453015)_48838493 | hg19 | chr3 | 47039557 | 47041894 | + | AGAAGGAAGATAGTGGCATAGGGCC | GAGGGGGAGGGTGGCTTCTACCCTCT | Homo sapiens | 52 |
| NBEAL2_UserDefined_(32453015)_48838494 | hg19 | chr3 | 47039557 | 47041894 | + | CGGACAGCCAGATGTAGTGCGGCTTCT | CTCTCTCCCCATTCTGCACGCCCTTT | Homo sapiens | 59.25925926 |
| NBEAL2_UserDefined_(32453015)_48838495 | hg19 | chr3 | 47039557 | 47041894 | + | AATGTGCTGGAGGACGGCAGCCT | CTCATAGCTCTCATCTGCAGTTGTGT | Homo sapiens | 60.89656522 |
| NBEAL2_UserDefined_(32453015)_48838496 | hg19 | chr3 | 47039557 | 47041894 | - | GAAAGTTGAAGAGCAAATGCTGGTAGA | AAACATGGCCACCAGCCTGCTCAGAA | Homo sapiens | 40.74074074 |
| NBEAL2_UserDefined_(32453015)_48838497 | hg19 | chr3 | 47039557 | 47041894 | - | ATAAACTGGACGCCGTACTTCTTCCGCAG | CATTCTGGATTCCTAACAGGTGAGTCC | Homo sapiens | 51.72413793 |
| NBEAL2_UserDefined_(32453015)_48838498 | hg19 | chr3 | 47039557 | 47041894 | - | TACCTGACCGTCATCGCCTGTGGCCGC | AAAGGCGGGGAGAGAGGGTGGGCAGGT | Homo sapiens | 66.66666667 |

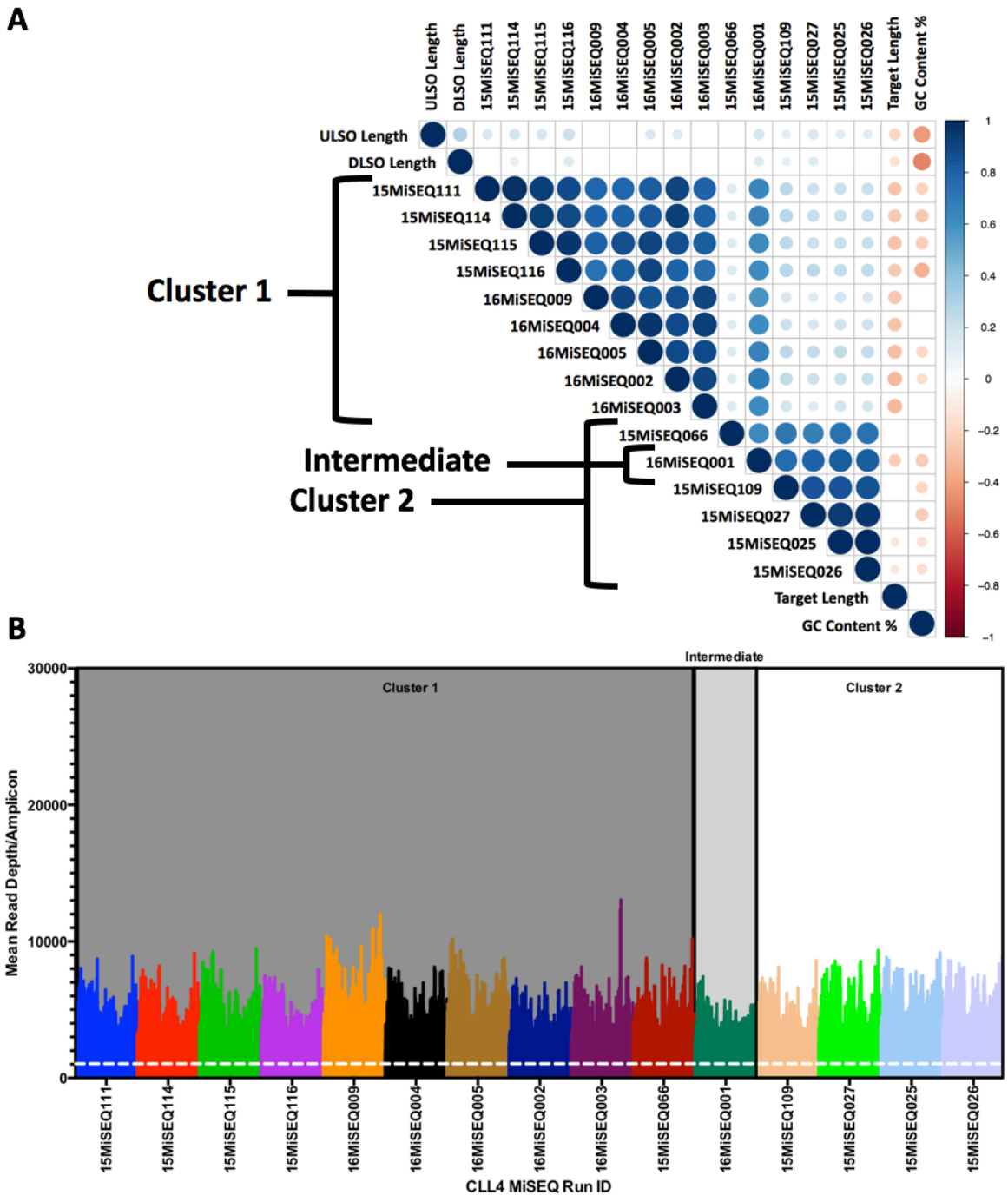
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| POT1_Exon (32129095)_48838513 | hg19 | chr7 | 124481027 | 124481232 | + | CTGGGGTTTTAGTTGCACCATCCTGAA | TGGAGATAATTCACTTTGCAGAACAAA | Homo sapiens | 48.14814815 |
| POT1_Exon (32129095)_48838514 | hg19 | chr7 | 124481027 | 124481232 | - | CCACATGAGGGCGATTTGGATATAATT | TTTAAGTCACTGTTTTGTTAGAATACTCC | Homo sapiens | 40.74074074 |
| POT1_Exon (32442521)_48838515 | hg19 | chr7 | 124491926 | 124492005 | - | GCAATAGATCCTTTACTGAAGAATGCA | CTTTAACATTGACAGTTTAGTTCTTGCT | Homo sapiens | 37.03703704 |
| POT1_ThreeUtrExon (32453038)_48838516 | hg19 | chr7 | 124462532 | 124462655 | - | GCTTGTTTGAGACAATGAGAAAGACTG | ACAAAGGAAAACTCTTCCTATGAATCTT | Homo sapiens | 40.74074074 |
| POT1_Exon (32442529)_48838517 | hg19 | chr7 | 124482861 | 124483017 | - | TGGTTCAGGAGGATGCATGTCATAAAT | AAACAAACGTTTCATATTATTTGTGTGTA | Homo sapiens | 40.74074074 |
| POT1_ThreeUtrExon (32442519)_48838518 | hg19 | chr7 | 124463966 | 124464128 | - | GTGGTAAGCTTAGGTACCTATTTCAGAGT | TGAGCTTCTTTTCCTGACCTGAGTTTT | Homo sapiens | 41.37931034 |
| POT1_Exon (32442528)_48838519 | hg19 | chr7 | 124477210 | 124477270 | - | CCTCTGAAACATTTTGTATTGTTTCCA | TGGTGAAGTACTCAAGGTTAAAGACTT | Homo sapiens | 33.33333333 |
| POT1_UserDefined (32442530)_48838520 | hg19 | chr7 | 124475283 | 124475518 | + | ACTTGTTCGAGAGTTGCAAATTTCAC | CAATGCACAAAAATGCTGAGACTTTGG | Homo sapiens | 37.03703704 |
| POT1_UserDefined (32442530)_48838521 | hg19 | chr7 | 124475283 | 124475518 | - | TTTATTCTCTTAACAGGAGGTACACTC | GGTCATGTGCTGAACATGGAGGTTT | Homo sapiens | 37.03703704 |
| CTBP2_ThreeUtrExon (32442533)_48838522 | hg19 | chr10 | 126678037 | 126678267 | + | GGGATGCCTCCAGGGATGATCCCTT | TCAAGTGATCCCACCTCAGACACCCAA | Homo sapiens | 60 |
| CTBP2_ThreeUtrExon (32442533)_48838523 | hg19 | chr10 | 126678037 | 126678267 | - | AGGAGGACTTCCTGCAGCCATG | AATAGATGAACTAAGAGAAAAAGAATCGG | Homo sapiens | 59.09090909 |
| CTBP2_Exon (32442531)_48838524 | hg19 | chr10 | 126681771 | 126681898 | - | TGTAGCTACTTTTCTACATGGCCAAGA | TGAGCCCCTCCGTGCCTCTGGCCTGC | Homo sapiens | 40.74074074 |
| CTBP2_Exon (32442534)_48838525 | hg19 | chr10 | 126682424 | 126682554 | - | CAACAAAACTAGCTTGCAGCTTTCTTC | TGTTGCCCCAGCCGCCTCACCAGGAGC | Homo sapiens | 40.74074074 |
| CTBP2_Exon (32453041)_48838526 | hg19 | chr10 | 126686533 | 126686739 | + | ACGGCTGCAGACGGGATGTTGCACA | CATGGGAAGCTCTAGCCCCAACATCTA | Homo sapiens | 60 |
| CTBP2_Exon (32453041)_48838527 | hg19 | chr10 | 126686533 | 126686739 | - | TTGGCCTCATCCCTTTCAGGAATTGCC | TACGTGCCACCCTCTGCGCTGACCCC | Homo sapiens | 51.85185185 |
| CTBP2_FiveUtrExon (32442535)_48838528 | hg19 | chr10 | 126714651 | 126716378 | + | CCGGGCCACCTTCTGGTACGGTGAGT | TCTGGGCCTGGGTGAAGGGGTTTGA | Homo sapiens | 65.38461538 |
| CTBP2_FiveUtrExon (32442535)_48838529 | hg19 | chr10 | 126714651 | 126716378 | + | CTGTCAGGGCGCAAGGGGTGGGAGA | ACCATGGAGCAGTTCGCTTCGGT | Homo sapiens | 68 |
| CTBP2_FiveUtrExon (32442535)_48838530 | hg19 | chr10 | 126714651 | 126716378 | + | ACAGACTTGATATCCGCGCTCCTCA | ATGCTTTCCCGGGCAGGCCCGTA | Homo sapiens | 52 |
| CTBP2_FiveUtrExon (32442535)_48838531 | hg19 | chr10 | 126714651 | 126716378 | + | ATGTCGCTGAAGACCTGGCTGA | GTTCTGCTGCCAAGCACTCCGTA | Homo sapiens | 54.54545455 |
| CTBP2_FiveUtrExon (32442535)_48838532 | hg19 | chr10 | 126714651 | 126716378 | + | GTCGTAGAAGGTGAAGTCAGGAGTAGA | AACGACCAATATTTATATGCCTGCTGGG | Homo sapiens | 48.14814815 |
| CTBP2_FiveUtrExon (32442535)_48838533 | hg19 | chr10 | 126714651 | 126716378 | - | CGCCGGCCAGCCAGAGCCTCCA | ACTCACCAGGCCACACACACACACACC | Homo sapiens | 77.27272727 |
| CTBP2_FiveUtrExon (32442535)_48838534 | hg19 | chr10 | 126714651 | 126716378 | - | ATTCTGCCCCTGGCGTGGCACCCATGCCC | ATGGAGCTGTCAAAGAGGGAGCGCAA | Homo sapiens | 70 |
| CTBP2_FiveUtrExon (32442535)_48838535 | hg19 | chr10 | 126714651 | 126716378 | - | TTTGACTCCCGGCCGTCCTGGCTACC | ACTTGGCTTCTCTGGGCGCTGCGGCTC | Homo sapiens | 66.66666667 |
| CTBP2_FiveUtrExon (32442535)_48838536 | hg19 | chr10 | 126714651 | 126716378 | - | TATGGGGCGGAGGTGCCTGCCTACCC | CCATCCAAGATGGCTTACGAGACTTAC | Homo sapiens | 69.23076923 |
| CTBP2_FiveUtrExon (32442535)_48838537 | hg19 | chr10 | 126714651 | 126716378 | - | TATAACTCAGTGGCTGCAAGAAAGGG | TCATGGGATCCAATGCAAGGCCGGTCA | Homo sapiens | 46.15384615 |
| CTBP2_FiveUtrExon (32442535)_48838538 | hg19 | chr10 | 126714651 | 126716378 | - | CGGAGGATTAATAATTCAGGCCATGCT | TCAGAGCTGGGATGCTGCTGGGTGGTA | Homo sapiens | 44.44444444 |
| CTBP2_FiveUtrExon (32453040)_48838539 | hg19 | chr10 | 126727586 | 126727673 | - | AGAGGGTCTCCAGAGAGTAAGGATTCT | TAAGAGCTGTTTGGCGGTTTTTACC | Homo sapiens | 48.14814815 |
| CHD2_Exon (31361024)_48838540 | hg19 | chr15 | 93499689 | 93499879 | + | TCTGGTCCTTGCTGCACTTTATTATGC | CCTGTTAAGGAGGAAAGCAGTATTGTC | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361024)_48838541 | hg19 | chr15 | 93499689 | 93499879 | - | AATACACAGGAAGGAGGAGCTTACTTCTCC | AAAGTAAAAACAAAACAGAACGAGCCA | Homo sapiens | 44.44444444 |
| CHD2_Exon (32442543)_48838542 | hg19 | chr15 | 93510555 | 93510743 | - | GTTTCATGGTACTGATCTCAACCCCTT | GGGAAGAAACAGTATAGTGTTATTTAGC | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361014)_48838543 | hg19 | chr15 | 93514995 | 93515157 | + | TTAAACCCCCTGAAGAAAATGAAAGGG | GGAAAAAGAACATTTCGGGAATTGCTG | Homo sapiens | 40.74074074 |
| CHD2_Exon (31361014)_48838544 | hg19 | chr15 | 93514995 | 93515157 | - | CCTACCAGAAGAATCTCCTGTCCATTT | AACAGAAAGTTGTAGGTACCAGAAGAG | Homo sapiens | 44.44444444 |
| CHD2_Exon (32453042)_48838545 | hg19 | chr15 | 93515495 | 93515647 | - | AGTGTAGGAAGGTTGGGGAGGGGAGGC | AAAGACATTCCAATAATTATGTTGAAACAC | Homo sapiens | 62.96296296 |
| CHD2_Exon (31361015)_48838546 | hg19 | chr15 | 93518109 | 93518180 | - | TTCTTCTCAAGGCAAAAGAATTGCACT | TGTAAAAATAAACCACAAGAAAACTGTAAG | Homo sapiens | 37.03703704 |
| CHD2_Exon (31361031)_48838547 | hg19 | chr15 | 93521464 | 93521613 | - | TAACATCCTGAAAGGGCTGAACAAGCC | GTACACAATGCAATATACTACTTTATGCA | Homo sapiens | 48.14814815 |
| CHD2_Exon (32442538)_48838548 | hg19 | chr15 | 93522365 | 93522513 | - | GAATCAATTTGATTCTGCCTACCTCCC | AGGAAAAGGAACTGTTATGCTTATGCC | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361011)_48838549 | hg19 | chr15 | 93524045 | 93524141 | - | TCCTACGAAAGAAAAACACACTGTCCT | GTGGAAGTAAAATGTAGTTGTTAGAAACA | Homo sapiens | 40.74074074 |
| CHD2_Exon (32442544)_48838550 | hg19 | chr15 | 93524595 | 93524687 | - | TAAAGAGAAAAGCAGAGCGCCTAGCAC | AAGAACAAAAGCCAACACGAATAGAAT | Homo sapiens | 48.14814815 |
| CHD2_Exon (31361029)_48838551 | hg19 | chr15 | 93527560 | 93527730 | - | ATCAACACTTCTGTAGCATTCACAACA | AAACAGACAACCCATTGTCTGACATTT | Homo sapiens | 37.03703704 |
| CHD2_Exon (32442545)_48838552 | hg19 | chr15 | 93528728 | 93528903 | - | GATCCAGTCAACCCAAGACTTACAGTA | TGAGAAAGATCGAGGCAAAGTCACAC | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361004)_48838553 | hg19 | chr15 | 93534706 | 93534747 | - | ATGTTCCCAGCATTAGTATTAAATGGA | GTTGATGCTAGCATATCCTGTGATGAC | Homo sapiens | 33.33333333 |
| CHD2_Exon (32442546)_48838554 | hg19 | chr15 | 93536089 | 93536228 | - | CCCTTTAGCAACAGAAATCCAGTCCAA | TGAGGAGGAACAAATATGGGGACACT | Homo sapiens | 44.44444444 |
| CHD2_UserDefined (32453005)_48838555 | hg19 | chr15 | 93540187 | 93540633 | + | TAGATGGTCGTACCGTAAGAAAATAGA | GGGAGTTAATTAAAACAGACCCAGAGC | Homo sapiens | 37.03703704 |
| CHD2_UserDefined (32453005)_48838556 | hg19 | chr15 | 93540187 | 93540633 | - | CTATGGAACTACACAAAGTCTAACTTCAA | ACCAGTTTCACATTAGTTAAATGAAAATCT | Homo sapiens | 34.48275862 |
| CHD2_UserDefined (32453005)_48838557 | hg19 | chr15 | 93540187 | 93540633 | - | CCCTTTATCTTCATCCGTGAAACAGTG | GTTTCCATAGCCATGTTCATAAATCCC | Homo sapiens | 44.44444444 |
| CHD2_UserDefined (32453004)_48838558 | hg19 | chr15 | 93489051 | 93489446 | + | CCACATGATGTTTTATTGATTATGCATAAA | CAGTAGGAACAACTCAAAAACCATCCC | Homo sapiens | 26.66666667 |
| CHD2_UserDefined (32453004)_48838559 | hg19 | chr15 | 93489051 | 93489446 | - | CTGTCAATTAAAAGTGAAAAATGACATTCT | AGACAGAAAGTTATAAACACAATGAAGA | Homo sapiens | 26.66666667 |
| CHD2_UserDefined (32453004)_48838560 | hg19 | chr15 | 93489051 | 93489446 | - | AGATTAAACAAGTCAGCAAAGCCATGA | GGAAGCTGTCAATGCAATTCTGGAATT | Homo sapiens | 37.03703704 |
| CHD2_Exon (32442537)_48838561 | hg19 | chr15 | 93467551 | 93467782 | + | CTCTGAATCTTCTGAGAGTCAGTCGGA | TCTTGATGATGAAACCTACAGCCATTT | Homo sapiens | 48.14814815 |
| CHD2_Exon (32442537)_48838562 | hg19 | chr15 | 93467551 | 93467782 | - | TTTGGAACCTGCTGATTCGCTCTCAGA | TTCTCATGGTAAAACATTTGCAGATAG | Homo sapiens | 48.14814815 |
| CHD2_Exon (31361019)_48838563 | hg19 | chr15 | 93470474 | 93470560 | - | CCCTCATGCCTGAAGTTCAGTCTTTTT | GTAAAATTAAGAGACCACGCATTAATCTT | Homo sapiens | 44.44444444 |
| CHD2_Exon (32442550)_48838564 | hg19 | chr15 | 93472260 | 93472321 | - | AGTAAGAGAGACAGAAGCAGACTAAGG | GGGGGAAAAATGATTACAGACTTTAATTC | Homo sapiens | 44.44444444 |
| CHD2_Exon (32442540)_48838565 | hg19 | chr15 | 93480748 | 93480855 | - | CATTAAAGCTCTCATTTCAGCACGACA | GTGATCAAAAGCAAAACAAAGTAATTTTT | Homo sapiens | 40.74074074 |
| CHD2_Exon (31361037)_48838566 | hg19 | chr15 | 93482808 | 93482948 | - | CTATTAAGTGCACAACCCCGGACTTAT | AATGAGAAGAGAAACAAAAGGCTAAGT | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361027)_48838567 | hg19 | chr15 | 93485052 | 93485185 | - | TACAGCAGTAGACTAAAATTCATCACA | GCAAGAGGAATAATGCTGAATAAATGAGA | Homo sapiens | 33.33333333 |
| CHD2_Exon (32442551)_48838568 | hg19 | chr15 | 93486073 | 93486298 | + | GCACATGGGAGAGTGAAGAATCCTTA | CTTTGATCATCTAAACATGGAGCCATT | Homo sapiens | 46.15384615 |
| CHD2_Exon (32442551)_48838569 | hg19 | chr15 | 93486073 | 93486298 | - | TTTTAGGCCCTTCACTTTCTGTTGC | AGTACACTTGAAAGAGAGTATATTACACC | Homo sapiens | 44 |
| CHD2_Exon (32442541)_48838570 | hg19 | chr15 | 93487645 | 93487745 | - | CACGTGTCTAGTCTCCAAAGTGCATTT | GAATGGCAGACAGACAGAAAAAGTACA | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361006)_48838571 | hg19 | chr15 | 93496587 | 93496803 | + | ATCATTCCTCTCCTACCTGTTCCACCA | ACTTAAAGCCTTCAGTAGATCATCAGT | Homo sapiens | 48.14814815 |
| CHD2_Exon (31361006)_48838572 | hg19 | chr15 | 93496587 | 93496803 | - | TATAAGAAAGGGGCCATACAGCTGGTG | CAATAACAGCCAAGCAAGGTGTCATTC | Homo sapiens | 48.14814815 |
| CHD2_Exon (32442542)_48838573 | hg19 | chr15 | 93498653 | 93498742 | - | GGACAAGTCGCACTGGAAAAGATTTAG | TGTGATTGATGTAATAGAAACATGCCA | Homo sapiens | 44.44444444 |
| CHD2_Exon (32442547)_48838574 | hg19 | chr15 | 93541729 | 93541851 | - | CTCTTCCTACCTAATTCTGCAGGGTAG | AACAAGAAAGGAAGAGAAGCACCCGCC | Homo sapiens | 48.14814815 |
| CHD2_Exon (31361018)_48838575 | hg19 | chr15 | 93543742 | 93543870 | - | TCCACTCGTAAACCAAACACATTTCAG | TGGAAAAGGAGAGAAAATGATGATTAAAAT | Homo sapiens | 40.74074074 |
| CHD2_Exon (32442548)_48838576 | hg19 | chr15 | 93545407 | 93545547 | - | GGGCCCTCAAATCTCATTCTACTTCTA | GGAAACAAAACAGGGCTATGAATTTTTG | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361017)_48838577 | hg19 | chr15 | 93547847 | 93547981 | - | CTATTGCTTAGCTCAACTGAAGCTCTG | AAAAAGAGTTACAAGGAGTCATTGTCA | Homo sapiens | 44.44444444 |
| CHD2_Exon (31361020)_48838578 | hg19 | chr15 | 93552375 | 93552553 | - | AACAGAATCCACTCCCACACTCAGTCA | GCAACAGGACAACAGAGTAAAGACATA | Homo sapiens | 48.14814815 |
| CHD2_Exon (32442539)_48838579 | hg19 | chr15 | 93557926 | 93558139 | + | ATGGACACCCAAGAGATAACTACAATC | AGAAAGCATGAAATAAAGATCCCCCAA | Homo sapiens | 40.74074074 |
| CHD2_Exon (32442539)_48838580 | hg19 | chr15 | 93557926 | 93558139 | - | GCATTACTGAAGTGTCTCTTGTTGGGG | AGAGTAAAAGGAAACCAAAGAAGCTCA | Homo sapiens | 48.14814815 |
| CHD2_Exon (32442549)_48838581 | hg19 | chr15 | 93555575 | 93555674 | - | CCAGCAATAACCCATGATTCTGTGATC | AAAGAAAAATGAACAGTTCTGGTGAAT | Homo sapiens | 44.44444444 |
| CHD2_UserDefined (32442449)_48838582 | hg19 | chr15 | 93492182 | 93492360 | - | CTTGGTTGAGATACAATAAGGGTGTGA | GAAAAACAAAAGCCCAAGTCTGAGAAA | Homo sapiens | 40.74074074 |
| CHD2_Exon (31361010)_48838583 | hg19 | chr15 | 93563242 | 93563488 | + | CCGTTCCGGAAGCTATCGACCCAACAA | ACAGCGAATCCCATGTCTTGGTTATAG | Homo sapiens | 55.55555556 |
| CHD2_Exon (31361010)_48838584 | hg19 | chr15 | 93563242 | 93563488 | - | TGGTCATAAGGCCTCTTTCTGGACAT | GCAGGAAAAGAAAGAATGATCTGATGACT | Homo sapiens | 46.15384615 |
| CHD2_FiveUtrExon (32129237)_48838585 | hg19 | chr15 | 93444418 | 93444529 | - | CAGCAATCCATGTCTGTCAAGGTAAAC | TCCTGTGCCCAGGTATTTACTGTCA | Homo sapiens | 44.44444444 |
| CHD2_ThreeUtrExon (32105245)_48838586 | hg19 | chr15 | 93567602 | 93567985 | + | AGATAAACTGGGGGAATATAAACAGCC | CTGATCCTACAGTAGCCGGTTATCTAG | Homo sapiens | 40.74074074 |
| CHD2_ThreeUtrExon (32105245)_48838587 | hg19 | chr15 | 93567602 | 93567985 | - | TGAGACTGCAGGGTGCAATGGGGGTAG | GAAATAGAGATGCAAGGATTAGTCGAAAG | Homo sapiens | 59.25925926 |
| CHD2_ThreeUtrExon (32105245)_48838588 | hg19 | chr15 | 93567602 | 93567985 | - | AACACATCCCCAGTGAAGCTGGTACACA | CCAAAAATATCCACGTGTTTGGTGACT | Homo sapiens | 48.14814815 |
| HIST1H1E_ThreeUtrExon (32442552)_48838589 | hg19 | chr6 | 26156568 | 26157328 | + | TGTCCGAGCTCATTACTAAAGCTGTTG | TGGGGAAGCCAAGCCTAAGGCTAAAAA | Homo sapiens | 44.44444444 |
| HIST1H1E_ThreeUtrExon (32442552)_48838590 | hg19 | chr6 | 26156568 | 26157328 | + | AAGGCGAAAGCAGCCAAGCCAAAAA | CAGTAAAAGAGCTGTTGCACTATTAGGG | Homo sapiens | 48 |
| HIST1H1E_ThreeUtrExon (32442552)_48838591 | hg19 | chr6 | 26156568 | 26157328 | - | AAAGATACCCCGCTGCCGCTCCTTGGAGGCG | TCGAGCCGCGGCGCTGCGCCCGCCCGTTT | Homo sapiens | 63.33333333 |
| HIST1H1E_ThreeUtrExon (32442552)_48838592 | hg19 | chr6 | 26156568 | 26157328 | - | TTTGGCCTTCGCTGGGCTCTTGGGCGCC | AGGCCGCCTTCTTGTTGAGTTTGAA | Homo sapiens | 67.85714286 |
| ZFPM2_FiveUtrExon (32129111)_48838593 | hg19 | chr8 | 106331147 | 106331209 | - | TTCGACTGCTCAGGCAAAGTTTGAGAC | TCGCGGCTGCCGCTCCCTCGGCTC | Homo sapiens | 48.14814815 |
| ZFPM2_Exon (32129112)_48838594 | hg19 | chr8 | 106431372 | 106431530 | - | GTGGACTCACATATTACATCCAACCAT | GAAAAAGAATCAGGGAAACTGTACCAAA | Homo sapiens | 40.74074074 |
| ZFPM2_Exon (32129113)_48838595 | hg19 | chr8 | 106456508 | 106456609 | - | CTCTCATGTAACCAGCACACACACAAAAA | GGAAAACACCAACACAAGGAGAAGAAG | Homo sapiens | 40.74074074 |
| ZFPM2_Exon (32129114)_48838596 | hg19 | chr8 | 106573591 | 106573709 | - | GGAAAGCAAGCAGAAGAGAATTTCTGAG | GCAGACAAAGGAACAGAAGTGTTTAGA | Homo sapiens | 44.44444444 |
| ZFPM2_Exon (32442555)_48838597 | hg19 | chr8 | 106648474 | 106646585 | - | CTTTGTACTTCCATTCCAGCTCTTCTG | GTAGAAAGAATGATAACAAACAGATGCCA | Homo sapiens | 44.44444444 |
| ZFPM2_Exon (32284888)_48838598 | hg19 | chr8 | 106800946 | 106801152 | - | GCTACACAGCACTGAGCACTTACTA | AGAAAAAGATAAAATTGACATAGGATGCT | Homo sapiens | 48 |
| ZFPM2_Exon (32442557)_48838599 | hg19 | chr8 | 106810952 | 106811176 | + | AGTGCACCAAGAGCTTTTCAAATGCTC | ACTCAAAGGACTTGGGTCTGAATTCTT | Homo sapiens | 44.44444444 |
| ZFPM2_Exon (32442557)_48838600 | hg19 | chr8 | 106810952 | 106811176 | - | GTGTGAATTCAGGTGCATTTCTAGAGCT | GCAGGTAAAAACACAAGAGACACATTT | Homo sapiens | 42.85714286 |
| ZFPM2_ThreeUtrExon (32284890)_48838601 | hg19 | chr8 | 106813275 | 106815816 | + | AGAATTCCTGCCCCCTGGTGCTAGTCTAAA | CAGAGAAAGTGACATGGAACACTCTCC | Homo sapiens | 50 |
| ZFPM2_ThreeUtrExon (32284890)_48838602 | hg19 | chr8 | 106813275 | 106815816 | + | ACAGAGCTGGACAAGTGTGAGAAAAAG | TTTCCCCCAAGATATCACCATGGTC | Homo sapiens | 44.44444444 |
| ZFPM2_ThreeUtrExon (32284890)_48838603 | hg19 | chr8 | 106813275 | 106815816 | + | TGGCAGTAGTAGCTACCCTCCCGTCATTTA | AGTCCCAACACTGGCCAAACCTCCATAAA | Homo sapiens | 50 |
| ZFPM2_ThreeUtrExon (32284890)_48838604 | hg19 | chr8 | 106813275 | 106815816 | + | ATCAATTCTTCCACTGTCTTAGATTTAATT | TGCAACATTACCTTCAGCCGGCACGAAA | Homo sapiens | 26.66666667 |
| ZFPM2_ThreeUtrExon (32284890)_48838605 | hg19 | chr8 | 106813275 | 106815816 | + | CTGCTTCCAACAAAGTGCCTGCCAT | AAGATGTGATATCTTTCCAGGAATTGTCTC | Homo sapiens | 52 |
| ZFPM2_ThreeUtrExon (32284890)_48838612 | hg19 | chr8 | 106813275 | 106815816 | + | TGATACTACTCATTCCAGTGTTTCCTGC | CACATCAGCGTAATGACCTGGGTCAA | Homo sapiens | 42.85714286 |
| ZFPM2_ThreeUtrExon (32284890)_48838613 | hg19 | chr8 | 106813275 | 106815816 | + | TGAGAAAAATGGGAATTTGAAGCAGCC | TCCATATTATGGAATCAAGCCAAGTGA | Homo sapiens | 40.74074074 |
| ZFPM2_ThreeUtrExon (32284890)_48838614 | hg19 | chr8 | 106813275 | 106815816 | + | TTCTCTGCCATTGTTCCCAAAAATCG | CAGAGGAACAGTTGTCTAGTATAGCAA | Homo sapiens | 44.44444444 |
| ZFPM2_ThreeUtrExon (32284890)_48838615 | hg19 | chr8 | 106813275 | 106815816 | + | GCAGCAGAACATGTCAAATGAACTAAC | CCAGCAGTATTCATAGCTGTGGTTATG | Homo sapiens | 40.74074074 |
| ZFPM2_ThreeUtrExon (32284890)_48838616 | hg19 | chr8 | 106813275 | 106815816 | - | GGTGTTGGTGAAAGTTGATCACGGAAT | ATGAAATCTTTCAAAATTTCATTCGTGAT | Homo sapiens | 44.44444444 |
| ZFPM2_ThreeUtrExon (32284890)_48838617 | hg19 | chr8 | 106813275 | 106815816 | - | TGTATCTCTGGTCTCTGGTTCTGTGAGA | AAGTTTGCCGCTAGGGACATGGAGCTCC | Homo sapiens | 48.14814815 |
| ZFPM2_ThreeUtrExon (32284890)_48838618 | hg19 | chr8 | 106813275 | 106815816 | - | ATTGAATGTTATGTTACACTCAAAACAAG | GAAAGAAGGCCCAATATTAGGCTGAAC | Homo sapiens | 27.5862069 |
| ZFPM2_ThreeUtrExon (32284890)_48838619 | hg19 | chr8 | 106813275 | 106815816 | - | CTTCTTAGTTTGAGTGGAAAAGTCCTTGT | TAGGGAACTCTGGGGACTTAGCCATCT | Homo sapiens | 37.93103448 |
| ZFPM2_ThreeUtrExon (32284890)_48838620 | hg19 | chr8 | 106813275 | 106815816 | - | CAGGTAGGCACATCTCATACATCTTTC | TTTCCCCATCCACTAAGGGGACACTGG | Homo sapiens | 44.44444444 |
| ZFPM2_ThreeUtrExon (32284890)_48838621 | hg19 | chr8 | 106813275 | 106815816 | - | GCTCAGACTGAGATAAACACTTTTGCTG | GGTTCTTGAGTGGAGTACAAGGATTAT | Homo sapiens | 41.37931034 |
| ZFPM2_ThreeUtrExon (32284890)_48838622 | hg19 | chr8 | 106813275 | 106815816 | - | GGTGGGATGAAAATAAGTTTCCATTGG | CCAGATAGTTTTCTACCTTATTGAAACTGA | Homo sapiens | 40.74074074 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| NOTCH1_ThreeUtrExon (32452971)_48838631 | hg19 | chr9 | 139390473 | 139392010 | + | GGTGCTGGGCAGGCCCTGGTAGCTCAT | TGGTGGACCCGCCCACAGTGAAATTCA | Homo sapiens | 66.66666667 |
| NOTCH1_ThreeUtrExon (32452971)_48838632 | hg19 | chr9 | 139390473 | 139392010 | + | CACGTTCAGGTGCCCGATGCCCA | GTCCTTGGCCTCCTTGCTTCCACA | Homo sapiens | 65.2173913 |
| NOTCH1_ThreeUtrExon (32452971)_48838633 | hg19 | chr9 | 139390473 | 139392010 | + | TGTACTCGTCCAGCAGCCTCACGATGT | CCATCACCAGAGGAAGCAGCAGTACAA | Homo sapiens | 55.55555556 |
| NOTCH1_ThreeUtrExon (32452971)_48838634 | hg19 | chr9 | 139390473 | 139392010 | - | TGTGGACAACACCCCCAGCCACCAG | GCTCTGTGGATGCCAGGGCCGA | Homo sapiens | 64 |
| NOTCH1_ThreeUtrExon (32452971)_48838635 | hg19 | chr9 | 139390473 | 139392010 | - | TTGCTGCCAGCGCCCTGTCCCAGAT | TTCCTGAGTGGAGAGCCGAGCCAGGCA | Homo sapiens | 64 |
| NOTCH1_ThreeUtrExon (32452971)_48838636 | hg19 | chr9 | 139390473 | 139392010 | - | ACCTGCCTGGGATGCCCGACACCCAC | TTTGAATGGTCAATGCGAGTGGCTGTC | Homo sapiens | 69.23076923 |
| NOTCH1_ThreeUtrExon (32452971)_48838637 | hg19 | chr9 | 139390473 | 139392010 | - | ACATCGCACAGGAGCGCATGCATCAC | TCAAGGCACGGAGGAAGAAGTCCCAGG | Homo sapiens | 57.69230769 |
| BRAF_Exon (32452983)_48838638 | hg19 | chr7 | 140453075 | 140453193 | - | CTTACCTAAACTCTTCATAATGCTTGCTC | AGAATTGAGGCTATTTTTCCACTGATTA | Homo sapiens | 37.93103448 |
| BRAF_Exon (32452985)_48838639 | hg19 | chr7 | 140481376 | 140481493 | - | TCCCTCTCAGGCATAAGGTAATGTACT | GTATGTAATGTGGTGACATTGTGACAA | Homo sapiens | 44.44444444 |
| BRAF_Exon (32400620)_48838640 | hg19 | chr7 | 140477791 | 140477875 | - | AATAATGGTATGGAGTTAGGGCTATGA | GAGCTTGTGTGAATTACTCTTTTCCAG | Homo sapiens | 37.03703704 |
| BRAF_Exon (32452982)_48838641 | hg19 | chr7 | 140439612 | 140439746 | - | CAGCTTATTGTGTTTTCAATGAGTGTG | GTAAAAGCTTCATGCTATCCAAAAGAA | Homo sapiens | 37.03703704 |
| BRAF_Exon (32400624)_48838642 | hg19 | chr7 | 140449087 | 140449218 | - | AGCATTGCTCTAGGAATTATAGTAGGT | AAATATTTACCACGTCTTGGTGTTTATTT | Homo sapiens | 37.03703704 |
| BRAF_Exon (32400623)_48838643 | hg19 | chr7 | 140453987 | 140454033 | - | CAGTCTATTTTAATATTCTGTGAAGGGTT | CCTTCCTGAAATTTGTCTGCGAAGTTT | Homo sapiens | 31.03448276 |
| BRAF_Exon (32452984)_48838644 | hg19 | chr7 | 140476712 | 140476888 | + | CACATGTCGTGTTTTCCTGTACAAAGA | CAGGCAACTAGAGGTTGCTTAAATACC | Homo sapiens | 40.74074074 |
| BRAF_Exon (32452984)_48838645 | hg19 | chr7 | 140476712 | 140476888 | - | GTCGACATTTAATGTTTACTGTCACAT | AGTTCCATTCGTTAAATGTCTTGTAAAT | Homo sapiens | 33.33333333 |
| BRAF_Exon (32452973)_48838646 | hg19 | chr7 | 140482821 | 140482957 | - | GTCTAAATATCCCCAACCTTCTACCC | TTTCTCTGAATTCTTTCTTCTAGAAGTC | Homo sapiens | 48.14814815 |
| BRAF_Exon (32452974)_48838647 | hg19 | chr7 | 140487348 | 140487384 | - | GTTCACTAGATAATTCCAAATTGTTTTGTG | TTTTAAAAACTCAAGGAAACTGCAATTG | Homo sapiens | 30 |
| BRAF_Exon (32452975)_48838648 | hg19 | chr7 | 140494108 | 140494267 | + | GGAAGGCTGTGGAATTGGAATGGATT | AAATCCAATACTGCCAAAGCTGCTTTG | Homo sapiens | 48.14814815 |
| BRAF_Exon (32452975)_48838649 | hg19 | chr7 | 140494108 | 140494267 | - | AAATTCTCACCAGTCCGTCTCCTTCA | AGTGCCATTTTTCATTTATCACAGTTA | Homo sapiens | 46.15384615 |
| BRAF_Exon (32452976)_48838650 | hg19 | chr7 | 140500162 | 140500281 | - | GTTTCTGAGAATGGAATTTGATCTCAG | TATGGTTTGACTTCTGCTCTTGGGCG | Homo sapiens | 37.03703704 |
| BRAF_Exon (32452977)_48838651 | hg19 | chr7 | 140501212 | 140501360 | - | AAACAAGAGAGTAGATACGTCAGTTTC | TCCAAAAATATCTCTTTTCTACCTACCAT | Homo sapiens | 37.03703704 |
| BRAF_Exon (32452978)_48838652 | hg19 | chr7 | 140507760 | 140507862 | + | TATCAGTGTCCCAACCAATTGGTTTCT | AAACAGCAGGGTAAAGAGATGTACTCT | Homo sapiens | 40.74074074 |
| BRAF_Exon (32452978)_48838653 | hg19 | chr7 | 140507760 | 140507862 | - | AGGCTTGTGTCTACATTTTTCAGAGAG | TTGGAACTGACATTTTACCTGAGTTGA | Homo sapiens | 40.74074074 |
| BRAF_Exon (32452979)_48838654 | hg19 | chr7 | 140508692 | 140508795 | - | GGTGTTGTATCTGACCTAGTAACCCAT | GGTTTGTATGTGACGTGAAATTTTGTT | Homo sapiens | 44.44444444 |
| BRAF_Exon (32452980)_48838655 | hg19 | chr7 | 140534409 | 140534672 | + | TAACGGTATCCATTGATGCAGAGCTAG | CACCAGTTGATAGGGTTTTAAGTTTCA | Homo sapiens | 44.44444444 |
| BRAF_Exon (32452980)_48838656 | hg19 | chr7 | 140534409 | 140534672 | - | CTGGGGAACGGAACTGATTTTTCTGTT | GAGTCAGTTTTAATATCACCATTTTGCTG | Homo sapiens | 44.44444444 |
| BRAF_Exon (32452981)_48838657 | hg19 | chr7 | 140549911 | 140550012 | + | ATAGGGCCTCTATATGTTCCTGTGTCA | TACACTAGTAATACAACGTATTCTCAACT | Homo sapiens | 44.44444444 |
| BRAF_Exon (32452981)_48838658 | hg19 | chr7 | 140549911 | 140550012 | - | AGGTGTGGAATATCAAACAAATGATTAAG | GGGATTAGGAATAGATTTGGGAAGAGAG | Homo sapiens | 31.03448276 |
| BRAF_ThreeUtrExon (32452972)_48838659 | hg19 | chr7 | 140434347 | 140434570 | + | TAAAATCCTCTGTTTGGAAACCAGCCC | CCATCACACTTTACAAGAAAGCACCTG | Homo sapiens | 44.44444444 |
| BRAF_ThreeUtrExon (32452972)_48838660 | hg19 | chr7 | 140434347 | 140434570 | - | CAGTGCATCAGAACCCTCCTTGAAT | CCAAAGAACACTTGTGTGGTTAAAGAC | Homo sapiens | 48 |
| BRAF_UserDefined (32453043)_48838661 | hg19 | chr7 | 140624315 | 140624555 | + | CGCCGCCACCACCGCCACCGCTCA | TGAGGGGATTGGGGGAAGGACGCTA | Homo sapiens | 79.16666667 |
| BRAF_UserDefined (32453043)_48838662 | hg19 | chr7 | 140624315 | 140624555 | - | GCTCTCGGTTATAAGATGGCGGC | CTGGTGTTTATTTTGGAAAGAGGCGG | Homo sapiens | 56.52173913 |
| MYD88_Exon (32452986)_48838663 | hg19 | chr3 | 38181355 | 38181489 | - | CTGCTTCAAACACCCATGCTCTTAACA | TGTGGGGAAGAGACAGAGTAAAGCAG | Homo sapiens | 44.44444444 |
| MYD88_UserDefined (32442465)_48838664 | hg19 | chr3 | 38181879 | 38182339 | + | CTCATCGAAAAGAGGTTGGCTAGAAGG | TGCAGAGCAAGGAATGTGACTTCCA | Homo sapiens | 48.14814815 |
| MYD88_UserDefined (32442465)_48838665 | hg19 | chr3 | 38181879 | 38182339 | - | TTCATGCATCCACGCACCCACCCCGTG | TGGGTGCAGAGCACAAGGGTGG | Homo sapiens | 62.96296296 |
| MYD88_UserDefined (32442465)_48838666 | hg19 | chr3 | 38181879 | 38182339 | - | CCTGGCACACATTCACTCACACAAAAA | GTAATCATCAGAGACAACCACCACCAT | Homo sapiens | 44.44444444 |
| MYD88_ThreeUtrExon (32452987)_48838667 | hg19 | chr3 | 38182623 | 38182827 | + | GTGTGTATCTGTCTGCCTGTCCATGTA | TATAGAACAGGACCAGCTGAGACTAAG | Homo sapiens | 48.14814815 |
| MYD88_ThreeUtrExon (32452987)_48838668 | hg19 | chr3 | 38182623 | 38182827 | - | TACAACGAAAGGAGGAGGCAGGGCAGAA | TGCAAGCCAAGGGGTGCCATGC | Homo sapiens | 53.57142857 |
| ATM_Exon (32442422)_48838669 | hg19 | chr11 | 108150218 | 108150335 | - | ACCAAACTTGGTGAAGTAATTTATGGG | CAAGCAAGCAAACAAACAAACAAAAAGT | Homo sapiens | 37.03703704 |
| ATM_Exon (32452991)_48838670 | hg19 | chr11 | 108173580 | 108173756 | + | GGAGGCACAAAATGTGAAATTCTTCAA | TTCAGCTTCGGGATAGCAACATACTAG | Homo sapiens | 37.03703704 |
| ATM_Exon (32452991)_48838671 | hg19 | chr11 | 108173580 | 108173756 | - | CTTCTTACTTCACACATTGGCTTTAATA | GGGATTAGAAAATGTAAATAAGAGAAGTGG | Homo sapiens | 32.14285714 |
| ATM_Exon (32083144)_48838672 | hg19 | chr11 | 108190681 | 108190785 | - | CAAGTTTTCAGAAAAGAAGCCATGACA | CATGTATAAATAGCTTCACCAAAGGAC | Homo sapiens | 37.03703704 |
| ATM_Exon (32218262)_48838673 | hg19 | chr11 | 108202606 | 108202764 | + | TACTGGCCTTAGCAAATGCAAACAGAG | ATCTAAAATCGGTTCAAGGCTGGCAC | Homo sapiens | 44.44444444 |
| ATM_Exon (32218262)_48838674 | hg19 | chr11 | 108202606 | 108202764 | - | GGCTACCTCTGGTTTAGTCAGAAATTCA | TAGATACTGCAGTGGGTAGAGCGTGTA | Homo sapiens | 42.85714286 |
| ATM_Exon (32083111)_48838675 | hg19 | chr11 | 108126942 | 108127067 | + | TAGCTGAAGAGGAAGCATATAAGTCAG | TAAGTTCTGGGGTACATGATGTAGAGC | Homo sapiens | 40.74074074 |
| ATM_Exon (32083111)_48838676 | hg19 | chr11 | 108126942 | 108127067 | - | TCTCCTACCTTGGCTTTCTGGAATAAT | AGTTCAGTTAACTGAGTTAAAGGTGGC | Homo sapiens | 40.74074074 |
| ATM_Exon (32083119)_48838677 | hg19 | chr11 | 108129713 | 108129802 | - | GCACCTGGCCTTAATTTCCACATTTGT | GCAATTAAGACAAACAAAGAACCAAGC | Homo sapiens | 44.44444444 |
| ATM_Exon (32083126)_48838678 | hg19 | chr11 | 108158327 | 108158442 | - | TGAAATAGACATTGAAGGTGTCAACCA | TGCAAGGGCACAAAATTAAAATATATTTAA | Homo sapiens | 37.03703704 |
| ATM_Exon (32452992)_48838679 | hg19 | chr11 | 108183138 | 108183225 | - | TGATGCTTAGTCCAGTAAGTAAATTCA | GTTTGTCACAAGAACACCTCTTTTAATGA | Homo sapiens | 33.33333333 |
| ATM_Exon (32083148)_48838680 | hg19 | chr11 | 108198372 | 108198485 | - | TGATCTTGATGAAAAGATGAAGCATATTC | AGGCAAAATCTTAAAATATAATGCATTACC | Homo sapiens | 31.03448276 |
| ATM_Exon (32218265)_48838681 | hg19 | chr11 | 108205696 | 108205836 | - | TCACTAAAACTCTAAGGGCTAAGCCAG | AACATGATTGATAATAGGTCAGATACACT | Homo sapiens | 44.44444444 |
| ATM_Exon (32083140)_48838682 | hg19 | chr11 | 108218006 | 108218092 | - | ACCCAATATTTACTGCTTCAATTTATCTG | AAGAAACAAACAATCTGAAATCACTGA | Homo sapiens | 31.03448276 |
| ATM_Exon (32083092)_48838683 | hg19 | chr11 | 108099905 | 108100050 | - | AAAAACTCACGCGACAGTAATCTGTTA | CAAGAAAACAAAATGCAATTTCAGAAAC | Homo sapiens | 37.03703704 |
| ATM_Exon (32452990)_48838684 | hg19 | chr11 | 108114680 | 108114845 | + | AATTCATGCTGTTACCAAAGGATGCTG | GGATTCCTAAAGAGACAACCAAGTCCA | Homo sapiens | 40.74074074 |
| ATM_Exon (32452990)_48838685 | hg19 | chr11 | 108114680 | 108114845 | - | TTTGGAATTTAATCCGTCAGTCTGAGA | TTCCATAATCCCATACAGAAAAAGGAT | Homo sapiens | 37.03703704 |
| ATM_Exon (32452993)_48838686 | hg19 | chr11 | 108119660 | 108119829 | + | AATCTAGTGATTACAGTGTCCCTTGCA | ACCAGAACTAAGTCATTTGTCTACCCC | Homo sapiens | 40.74074074 |
| ATM_Exon (32452993)_48838687 | hg19 | chr11 | 108119660 | 108119829 | - | TCCCAGCCTAGTTCTATTTTCTTCCTT | ACGATCTCGTATCACCACACAACCTCTAC | Homo sapiens | 40.74074074 |
| SF3B1_UserDefined (32442455)_48838688 | hg19 | chr2 | 198267280 | 198267759 | + | GCCCATAAGAATAGCTATCTGTTGTAC | TGTAAACAAAGAAAGGACAGTCATGAG | Homo sapiens | 40.74074074 |
| SF3B1_UserDefined (32442455)_48838689 | hg19 | chr2 | 198267280 | 198267759 | + | AATGACCACGAGGATCTGAAAAAGAGA | ATACACCCTATGAACATACTGCACCAT | Homo sapiens | 40.74074074 |
| SF3B1_UserDefined (32442455)_48838690 | hg19 | chr2 | 198267280 | 198267759 | - | GCAAGCGAGACACACTGGTATTAAGAT | ATCCTCCTGCTCAGCCTCCTTAGTAG | Homo sapiens | 44.44444444 |
| SF3B1_UserDefined (32442455)_48838691 | hg19 | chr2 | 198267280 | 198267759 | - | TGATGTGAAAGTGTAGCTTCTTCTCTT | TTTAGGCTGCTGGTCTGGCTACTATGA | Homo sapiens | 37.03703704 |
| SF3B1_UserDefined (32442454)_48838692 | hg19 | chr2 | 198264779 | 198265158 | + | TGATGCAAAAGTTTACCTCTTGACAAG | CTGCACAGTTGAAATACACTAAGAGTC | Homo sapiens | 37.03703704 |
| SF3B1_UserDefined (32442454)_48838693 | hg19 | chr2 | 198264779 | 198265158 | - | GATTTCTCGAACTGCTGTTGTCATGAAG | GGAATTTGCTGGTTACATAAGACTTTT | Homo sapiens | 42.85714286 |
| SF3B1_Exon (32218495)_48838694 | hg19 | chr2 | 198268309 | 198268488 | + | TCACGAGCTTTATCAGTAATCTGACGC | AATAGAACGAACCACAGAGGTATGTAC | Homo sapiens | 44.44444444 |
| SF3B1_Exon (32218495)_48838695 | hg19 | chr2 | 198268309 | 198268488 | - | TACTCTGCTCTTTTTCCCAGGCTGCAT | TTGTTCTCCTTAATAGTGTTTCTGATAGT | Homo sapiens | 48.14814815 |
| SF3B1_UserDefined (32453045)_48838696 | hg19 | chr2 | 198269680 | 198270250 | + | TTTTCTCTCTTTTTGCTCTTCTGGACT | TCTTCAGTTTGCATGTGGAAACCAGTC | Homo sapiens | 37.03703704 |
| SF3B1_UserDefined (32453045)_48838697 | hg19 | chr2 | 198269680 | 198270250 | - | AGGTTGATGTTGATGAATCAACACT | TGCTTTGGCCTACCTCCCAAAGTG | Homo sapiens | 36 |
| SF3B1_UserDefined (32453045)_48838698 | hg19 | chr2 | 198269680 | 198270250 | - | AGCAAGATAGATGTTGTCTCCAAGCAA | CGAACTATGAAAAGTGTTAATGACCAG | Homo sapiens | 40.74074074 |
| XPO1_UserDefined (32453021)_48838699 | hg19 | chr2 | 61719170 | 61720188 | + | ATCATGGGTCTCTAACAAGACAAAAAT | AACGTGGGTATTGACCTACTATGTACA | Homo sapiens | 33.33333333 |
| XPO1_UserDefined (32453021)_48838700 | hg19 | chr2 | 61719170 | 61720188 | + | ACAAGAAATCGTTTTTCGTCCTCTTCA | ATGAGCAGAATTTTATTGATGAGTAACAA | Homo sapiens | 37.03703704 |
| XPO1_UserDefined (32453021)_48838701 | hg19 | chr2 | 61719170 | 61720188 | + | AATTCTCTCACAACTTCTCCTTGATCA | TCATGCCTATAATTCCAACTATTCCGG | Homo sapiens | 37.03703704 |
| XPO1_UserDefined (32453021)_48838702 | hg19 | chr2 | 61719170 | 61720188 | - | GATGGAGTCCAGGACAAAATAAATGAA | TGTTAAGCACTGAGCATATGTTATTTTT | Homo sapiens | 37.03703704 |
| XPO1_UserDefined (32453021)_48838703 | hg19 | chr2 | 61719170 | 61720188 | - | AATAGGCTCCATTAGTGGAGCAATGC | TTTGAGAGCTCACTGGAAATTTCTGAA | Homo sapiens | 46.15384615 |
| XPO1_UserDefined (32453021)_48838704 | hg19 | chr2 | 61719170 | 61720188 | - | AACCAGAGGAAGTATTGGTTGTAGAGA | ATTTGCATTTGAAACCATGTTTTTCTT | Homo sapiens | 40.74074074 |
| FBXW7_Exon (32316629)_48838705 | hg19 | chr4 | 153245336 | 153245546 | + | CACATCCCAAACACGGATTGATGTATC | CTCACACACTTCTATCAGAAAGGCTCA | Homo sapiens | 44.44444444 |
| FBXW7_Exon (32316629)_48838706 | hg19 | chr4 | 153245336 | 153245546 | - | TGGTATCCATGTGGTGAGTGGATCTCT | GTCTTACTCACCTTTCGTTTTGTTCAT | Homo sapiens | 48.14814815 |
| FBXW7_Exon (32452998)_48838707 | hg19 | chr4 | 153247158 | 153247383 | + | CAAAACATGTAAACACTGGCCTGTCTC | AACATAATAAGGGAATGGGAAGTTGGG | Homo sapiens | 44.44444444 |
| FBXW7_Exon (32452998)_48838708 | hg19 | chr4 | 153247158 | 153247383 | - | GATGCCACCTCTTAGGGTTTGGGATAT | GATCTCTTATCTCTCCCTTAAATGCTC | Homo sapiens | 46.15384615 |
| FBXW7_Exon (32316624)_48838709 | hg19 | chr4 | 153249360 | 153249541 | + | CCATACTCCACCTGTATGTCCCACTAA | AGCACAAGCAATTGTGTATGTAGTATG | Homo sapiens | 48.14814815 |
| FBXW7_Exon (32316624)_48838710 | hg19 | chr4 | 153249360 | 153249541 | - | CTTAAGTGTTTTTCCAGTGTCTGAGAAC | AGGGAAAATCTTGTCATGGTTGGGAC | Homo sapiens | 39.28571429 |
| FBXW7_Exon (32452999)_48838711 | hg19 | chr4 | 153250824 | 153250937 | - | ATGCCTTCATTTTTCTCTTCACCAGTA | GTTTACTCTTCCTACACTTCACAAAGT | Homo sapiens | 37.03703704 |
| FBXW7_Exon (32316628)_48838712 | hg19 | chr4 | 153251884 | 153252020 | - | GGTTTTGAGCAGAGAGATGGAATACTC | AACTGAACTCTGCACAATGTAACAGAA | Homo sapiens | 44.44444444 |
| FBXW7_Exon (32453000)_48838713 | hg19 | chr4 | 153253748 | 153253871 | - | ATCAAGTATCTCATCCTGTGGAGAATA | AAGGTTCAAAATTGCATAGAGAATATGTT | Homo sapiens | 37.03703704 |
| FBXW7_Exon (32452997)_48838714 | hg19 | chr4 | 153258954 | 153259088 | - | ACCCTGATCAACTATGGTTGTGTGTAG | AGATTATCAATTACATTCTTCATTGGGTG | Homo sapiens | 44.44444444 |
| FBXW7_ThreeUtrExon (32453001)_48838715 | hg19 | chr4 | 153243983 | 153244301 | + | TCATGTCCACATCAAAGTCCAGCAC | TCTTGTTGAACTGTAAACAGGTCACAGC | Homo sapiens | 48 |
| FBXW7_ThreeUtrExon (32453001)_48838716 | hg19 | chr4 | 153243983 | 153244301 | - | AATGGGACTGAAGAAACCAAGCTGCT | AGACCTACAGACTAAGAAGGAAAAGAA | Homo sapiens | 46.15384615 |
| FBXW7_ThreeUtrExon (32453001)_48838717 | hg19 | chr4 | 153243983 | 153244301 | - | GCTGATCCTAAAATACTGAGGACATGG | CTTTGTAATTACCAGCTCAGATGATGG | Homo sapiens | 44.44444444 |
| DDX3X_Exon (32218356)_48838718 | hg19 | chrX | 41196661 | 41196718 | - | AAAGGCAACGATGCTTTAGATTGATCA | TAGAAGAGAGTGAGATGGTTAGCACCCC | Homo sapiens | 37.03703704 |
| DDX3X_Exon (32453002)_48838719 | hg19 | chrX | 41198289 | 41198336 | - | CAAATGCCTGGTTTTCCTTAGGCTTTA | CCAGCCTTCCAATCTGATTACTTACCT | Homo sapiens | 40.74074074 |
| DDX3X_Exon (32218358)_48838720 | hg19 | chrX | 41200737 | 41200869 | + | TGCGTATAGCAGTTTTGGATCTCGTAG | TAGCCGTAAGAGGCTTTACTAAATTG | Homo sapiens | 44.44444444 |
| DDX3X_Exon (32218358)_48838721 | hg19 | chrX | 41200737 | 41200869 | - | GAAGAAGCTAGACTTCCCTCTTGAATC | CTATCCCTACCATGGCACTGTAAGAGA | Homo sapiens | 44.44444444 |
| DDX3X_Exon (32453003)_48838722 | hg19 | chrX | 41202469 | 41202604 | - | CAGCTGCAAAACTGCATGACCCATAAA | GTTAAGAAATTCAATGGGGTTAGGAATGC | Homo sapiens | 44.44444444 |
| DDX3X_UserDefined (32453007)_48838723 | hg19 | chrX | 41201748 | 41202089 | + | GTGAACGCTTGGAACAGTAAGTTTTTG | TGCTTGACTTTTTAAGACACAGAGAGG | Homo sapiens | 40.74074074 |
| DDX3X_UserDefined (32453007)_48838724 | hg19 | chrX | 41201748 | 41202089 | - | AAGGCTTCATCACAAGTAACATACACT | GAACAGCACAAGCAATTAATCAGAATCAT | Homo sapiens | 37.03703704 |
| DDX3X_UserDefined (32453024)_48838725 | hg19 | chrX | 41205482 | 41206265 | + | CCAGTTCCGCTCAGGAAAAAGCCCAATTT | TATTGAAGAATATGTACATCGTATTGGTCG | Homo sapiens | 48.27586207 |
| DDX3X_UserDefined (32453024)_48838726 | hg19 | chrX | 41205482 | 41206265 | + | ACAATGAAAGTGACAAAGATTTTGCTC | TTAGAAAACATGGCTTATGAACACCAC | Homo sapiens | 33.33333333 |
| DDX3X_UserDefined (32453024)_48838727 | hg19 | chrX | 41205482 | 41206265 | - | GATGTTTATACATACTGCTGTAGCCACT | GTCGTTGAAAAATTATAGGTTTTTGTCAC | Homo sapiens | 39.28571429 |
| DDX3X_UserDefined (32453024)_48838728 | hg19 | chrX | 41205482 | 41206265 | - | TTACCACGTAATATTTAAACAAGTGCTT | CACTTGGCAAGTCAAAATTGATAACATGT | Homo sapiens | 28.57142857 |
| DDX3X_UserDefined (32453024)_48838729 | hg19 | chrX | 41205482 | 41206265 | - | CCCAACAGAGTGTCACCCATTCTAATA | CAAGACGGCACTTCTTGTTTAGCTTCA | Homo sapiens | 44.44444444 |
| DDX3X_FiveUtrExon (32129131)_48838730 | hg19 | chrX | 41193456 | 41193550 | - | GCACCCAGAAAATGAAAAGTAAGGGTG | TTCCACCGCGAAGGCCCTCTCACGGGAG | Homo sapiens | 44.44444444 |
| KRAS_Exon (32453009)_48838731 | hg19 | chr12 | 25378548 | 25378707 | - | ATGACAAAAGTTGTGGACAGGTTTTGA | AGTAACACTGAAATAAATACAGATCTGTTT | Homo sapiens | 37.03703704 |
| KRAS_Exon (32333277)_48838732 | hg19 | chr12 | 25380168 | 25380346 | - | GGTGCACTGTAATAATCCAGACTGTGT | TGGGTTTAAATTGAATATAATAAGCTGACA | Homo sapiens | 44.44444444 |
| KRAS_ThreeUtrExon (32453008)_48838733 | hg19 | chr12 | 25362678 | 25362845 | - | CCTGTACACATGAAGCCATCGTATATA | AGCATTTGTTTTAGCATTCCTAATTTTT | Homo sapiens | 40.74074074 |
| KRAS_ThreeUtrExon (32401272)_48838734 | hg19 | chr12 | 25368371 | 25368494 | - | GCACATGGCTTTCCCAGTAAATTACTC | GTTTAAGTTCAGCACATTAATTTTGGC | Homo sapiens | 44.44444444 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| MGA_UserDefined (32442571)_48838749 | hg19 | chr15 | 42040693 | 42042813 | + | GTCATGAATCCTGTAATTCAAGCTGTTGG | GACAGCCTGTTGGTACAGCCAGTCTTATT | Homo sapiens | 41.37931034 |
| MGA_UserDefined (32442571)_48838750 | hg19 | chr15 | 42040693 | 42042813 | + | CTCAGAATCCAGACCAGAAAGATGAAAC | CAGAAGAAGGTTGTGCAACTGTCAAAC | Homo sapiens | 42.85714286 |
| MGA_UserDefined (32442571)_48838751 | hg19 | chr15 | 42040693 | 42042813 | + | GTGGCTGTTCTGGAAGTTAGGACCATT | AAGTTATGGAGCAGCAATCTAATCTACAGC | Homo sapiens | 48.14814815 |
| MGA_UserDefined (32442571)_48838752 | hg19 | chr15 | 42040693 | 42042813 | + | TAACCAGGAAATGTGTTGGAGCTTCAC | TGTTCCTAGGAGAGCTGCAAAAAGCA | Homo sapiens | 44.44444444 |
| MGA_UserDefined (32442571)_48838753 | hg19 | chr15 | 42040693 | 42042813 | + | ATTGTGGATGTTGTTTCTGACTACCAG | CTTTAGGCTGTGTCAATCAAGCATCAT | Homo sapiens | 40.74074074 |
| MGA_UserDefined (32442571)_48838754 | hg19 | chr15 | 42040693 | 42042813 | - | GGTTCCGAAACATGACAGGCTGTAA | CACTGGAATCAACAACAATCGAGGTC | Homo sapiens | 48 |
| MGA_UserDefined (32442571)_48838755 | hg19 | chr15 | 42040693 | 42042813 | - | CCTGAGTGATAACATTCACTGCTGAAGAA | GACTTCCAATGTTTGCTTTTCTTTTTGGAA | Homo sapiens | 41.37931034 |
| MGA_UserDefined (32442571)_48838756 | hg19 | chr15 | 42040693 | 42042813 | - | CTTCGTTTCTTGCTCTCTTTTATTGAGT | GAGCTATAAGTTATTTTGGTTTCAGAGCC | Homo sapiens | 34.48275862 |
| MGA_UserDefined (32442571)_48838757 | hg19 | chr15 | 42040693 | 42042813 | - | CTTAAATTTTGGACTGTCTTATTACTGGC | GCATTCTTCATCCAAATGATCACCCCTA | Homo sapiens | 34.48275862 |
| MGA_UserDefined (32442571)_48838758 | hg19 | chr15 | 42040693 | 42042813 | - | TCCTGACATGTCTTTGTTTCTTTAATTAA | CTTCCTGACAATTCAAATTTACTTTCTTCC | Homo sapiens | 27.5862069 |
| MGA_UserDefined (32442571)_48838759 | hg19 | chr15 | 42040693 | 42042813 | - | CCTTTTCTACATCATCAACCTCCTCAC | TGGAGAAGGCTGAAGGATTAGAAATCC | Homo sapiens | 44.44444444 |
| SETD2_Exon (32129037)_48838760 | hg19 | chr3 | 47059128 | 47059229 | - | GTTAACCTGAGGGAGGAAAAGGAGAAG | TACTCCCCTTGCTTTTGGTAGCCCAC | Homo sapiens | 48.14814815 |
| SETD2_Exon (32105530)_48838761 | hg19 | chr3 | 47061250 | 47061330 | - | GGGCCTCTAAACGGATTCTAATACCAA | AAGGTTTGGTGGATATAAGCACTAAGC | Homo sapiens | 44.44444444 |
| SETD2_Exon (32129039)_48838762 | hg19 | chr3 | 47079156 | 47079267 | - | ACTCATGGAGCTGACTTATTCTATGTC | GAGTGTGGCTTACACATGTATTTCCTG | Homo sapiens | 40.74074074 |
| SETD2_Exon (32442496)_48838763 | hg19 | chr3 | 47087977 | 47088111 | - | TCTTTAAGATTGTGCCAACTTGGAAGT | GAAATGTTTCAGTTGTTAAAACTCATTGG | Homo sapiens | 37.03703704 |
| SETD2_Exon (32105534)_48838764 | hg19 | chr3 | 47098311 | 47098980 | + | CACTATAATTCTGTCCCTGAACTGGGC | CATGATCATAAGGAGCAGGAAACACAC | Homo sapiens | 48.14814815 |
| SETD2_Exon (32105534)_48838765 | hg19 | chr3 | 47098311 | 47098980 | + | TAACCAAGAGAGTCATAGGGCAGTGGT | GTTGCTCCCTAAATCATAGACACAATC | Homo sapiens | 48.14814815 |
| SETD2_Exon (32105534)_48838766 | hg19 | chr3 | 47098311 | 47098980 | - | TTGCTGTCTTGCCAGTCCGGCCCCC | ACTAATTCTGAAACTTTCCTTTTCTGCT | Homo sapiens | 69.23076923 |
| SETD2_Exon (32105534)_48838767 | hg19 | chr3 | 47098311 | 47098980 | - | ACAGATGCAGAACCTGGGAATGACAT | TCAGCCCTTGGTGGGACATTCTACAGA | Homo sapiens | 46.15384615 |
| SETD2_Exon (32129043)_48838768 | hg19 | chr3 | 47103653 | 47103836 | + | AGTCTTCGGGGCAGGTGTTTGATCTCT | ACAAGGTCTTATGTCTGTGTTCACTCT | Homo sapiens | 51.85185185 |
| SETD2_Exon (32129043)_48838769 | hg19 | chr3 | 47103653 | 47103836 | - | TGAACGAGGAAGGGATGCTGTTGGCTT | TGTTTTTGTTTTTGTTTTGAGCAAGTTAA | Homo sapiens | 51.85185185 |
| SETD2_Exon (32442499)_48838770 | hg19 | chr3 | 47127685 | 47127804 | - | CTCTCCTCGGGCTCTTTGTAATATTCT | CAGTTCACATTTAGCTTGCTTGTTTCT | Homo sapiens | 44.44444444 |
| SETD2_Exon (32105540)_48838771 | hg19 | chr3 | 47139445 | 47139571 | - | CTGGTATGACTTTTGTGTGTATTGCCA | GTTTGCATTTTTCTTTCAAGTTTCACA | Homo sapiens | 40.74074074 |
| SETD2_Exon (32129049)_48838772 | hg19 | chr3 | 47142948 | 47143045 | - | CATTGCCTAAAGACCCACTTGCATAAT | AAGTATTGTTTCTACTAATTTTGCTGCTT | Homo sapiens | 40.74074074 |
| SETD2_Exon (32442500)_48838773 | hg19 | chr3 | 47144836 | 47144913 | - | CTGGTTAGGTTTCGCAGCACAATCTTT | AAGTTGAGGTGGATTTAGAGTTTGAGG | Homo sapiens | 44.44444444 |
| SETD2_Exon (32129053)_48838774 | hg19 | chr3 | 47158113 | 47158244 | - | ACCCAAAAGAATCTAATGAGATTTGGC | GTAAATTAGAATACTCGGCTTTTGCTC | Homo sapiens | 37.03703704 |
| SETD2_Exon (32442494)_48838775 | hg19 | chr3 | 47168138 | 47168153 | - | GCCCTTAGCACAGTATGTGACACTATC | TTAACTAGTCATTCAATGATTATGAGATGG | Homo sapiens | 48.14814815 |
| SETD2_ThreeUtrExon (32442498)_48838776 | hg19 | chr3 | 47058533 | 47058744 | + | TCCAGGTCCTCAGGATTCTTACAGTAC | TATGATCACCAGAAAGGACAAGCTGAG | Homo sapiens | 48.14814815 |
| SETD2_ThreeUtrExon (32442498)_48838777 | hg19 | chr3 | 47058533 | 47058744 | - | CTCACGGTGTTATGAATAAGGAGCTGA | AGAGGAAATCCTGTGGGCCTTTCT | Homo sapiens | 44.44444444 |
| SETD2_FiveUtrExon (32105548)_48838778 | hg19 | chr3 | 47205344 | 47205464 | - | AGGGAGCGACGAGCGAGGTAGCGA | CCCCGCTGCTGTCGGCCGCCTT | Homo sapiens | 66.66666667 |
| SETD2_UserDefined (32442574)_48838779 | hg19 | chr3 | 47155316 | 47155544 | + | CGTCTATTGGAACAATAATCCCCATTTGG | ACAGATCATACCCACAATACAGCCTTG | Homo sapiens | 41.37931034 |
| SETD2_UserDefined (32453044)_48838780 | hg19 | chr3 | 47125210 | 47126200 | + | ATAGGTTCTTCTAGTTTTGTGCCGTTG | CTCACTGGTTGCATCAGAGATTGCA | Homo sapiens | 40.74074074 |
| SETD2_UserDefined (32453044)_48838781 | hg19 | chr3 | 47125210 | 47126200 | - | TGACGCTGAAATAGAGCCCAAAGAGA | GGAAGCATCTTTTGTTCATTTCAACTTGA | Homo sapiens | 46.15384615 |
| SETD2_UserDefined (32453044)_48838782 | hg19 | chr3 | 47125210 | 47126200 | - | AGACTAAGACTGCTGTCCCTCCGTTG | TAGAAGGCAAGGATGGCAAAGAGGATC | Homo sapiens | 53.84615385 |
| BIRC3_UserDefined (32408353)_48838783 | hg19 | chr11 | 102207441 | 102208515 | + | GAAGCAAACTGCCTTTTATTAAAAATTCT | TCATGAAGAAGAACCAAAACATCGTCT | Homo sapiens | 27.5862069 |
| BIRC3_UserDefined (32408353)_48838784 | hg19 | chr11 | 102207441 | 102208515 | + | AAGAGAATGATAGGCTTTTGTTCTTATGA | TCCCAGGCTGAGGCAAGAGAATTACTT | Homo sapiens | 31.03448276 |
| BIRC3_UserDefined (32408353)_48838785 | hg19 | chr11 | 102207441 | 102208515 | + | CGAAGGTGTGCATATATGTTGAATGAC | AACACGGGGTCAACACATCATAAAATC | Homo sapiens | 40.74074074 |
| BIRC3_UserDefined (32408353)_48838786 | hg19 | chr11 | 102207441 | 102208515 | - | CTTCCACTGGTAGATCTAGAAAGAGGA | TTTAAAAATGAACACTTTAGACTATAGCCA | Homo sapiens | 44.44444444 |
| BIRC3_UserDefined (32408353)_48838787 | hg19 | chr11 | 102207441 | 102208515 | - | TGTGTTTGTAGTGCTACCTCTTTTTCG | AGAAATGTACGAACTGTACCCTTGATT | Homo sapiens | 40.74074074 |
| NFKBIE_UserDefined (32453030)_48838788 | hg19 | chr6 | 44227780 | 44228276 | + | ACTGTACCAGGCCCCGCTCTTGGGTTT | TCAGCTCCATTCCGAAGCAGCAATT | Homo sapiens | 59.25925926 |
| NFKBIE_UserDefined (32453030)_48838789 | hg19 | chr6 | 44227780 | 44228276 | - | TAAGACAGCGCTGCACCTGGCTGTG | TAAGTGGCATATGCTTTGCTCTCCAG | Homo sapiens | 60 |
| NFKBIE_UserDefined (32453030)_48838790 | hg19 | chr6 | 44227780 | 44228276 | - | CTTGTGATTAGAGACATTGAGTGGGCT | ATTGATGTGCAGGTGAGACAGCCAGCA | Homo sapiens | 44.44444444 |
| NBEAL2_Exon (32453027)_48838791 | hg19 | chr3 | 47033965 | 47034045 | - | ATTTCTAGACCACACCAGGCCTTACAC | TGTGGGGAAGTGGGGAGGCACAGGGT | Homo sapiens | 48.14814815 |
| NBEAL2_Exon (32453026)_48838792 | hg19 | chr3 | 47035945 | 47036043 | - | CAGGCTACAGGTAACAGTATGTACTGG | GTGTAAGGTAGTATGGGGAAGAAGTGA | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32453011)_48838793 | hg19 | chr3 | 47032709 | 47033436 | + | GAGAACGGGCAGATGGCTGTAAGTGAT | CTCAGGCCCCGAAGAGCCCTTGTCA | Homo sapiens | 51.85185185 |
| NBEAL2_UserDefined (32453011)_48838794 | hg19 | chr3 | 47032709 | 47033436 | + | GGGCCTGACCCTTGACCCTTCCACAGA | AAACATGGGCTGGACAGAGTCAGTGTT | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined (32453011)_48838795 | hg19 | chr3 | 47032709 | 47033436 | - | ACACTCAGCAGGCCCTTCACAGAGC | GGGGTGTGGCAGGGTAATGTC | Homo sapiens | 60 |
| NBEAL2_UserDefined (32453011)_48838796 | hg19 | chr3 | 47032709 | 47033436 | - | TTCACATGCCAGCATCATGGGGATGGC | TCAGACCAGCTGGCCAGTCAGCATTAA | Homo sapiens | 55.55555556 |
| NBEAL2_UserDefined (32453013)_48838797 | hg19 | chr3 | 47036522 | 47039184 | + | CGAGCAGCCGGTACTGGTGCTG | AGCTGCTGGCACTGCTACAAGCACT | Homo sapiens | 68.18181818 |
| NBEAL2_UserDefined (32453013)_48838798 | hg19 | chr3 | 47036522 | 47039184 | + | GCACATTATCAGGCATGGCCAG | GCTGTACAGGTGGGTGACTGCCA | Homo sapiens | 54.54545455 |
| NBEAL2_UserDefined (32453013)_48838799 | hg19 | chr3 | 47036522 | 47039184 | + | GCGGCCGGGACCCTGGTGGTGGCTGTGT | GCCTGGGCGCCGGCCCTTCAGCCAGAA | Homo sapiens | 75 |
| NBEAL2_UserDefined (32453013)_48838800 | hg19 | chr3 | 47036522 | 47039184 | + | CACAGTGGGCAGGTGTAGCCTG | CGCACAACGACCACCACCACAG | Homo sapiens | 63.63636364 |
| NBEAL2_UserDefined (32453013)_48838801 | hg19 | chr3 | 47036522 | 47039184 | + | GCAGCATCGACTCTACCCTCGCA | AGGCATCTCTGGGGGTCTCTCCTGT | Homo sapiens | 60.86956522 |
| NBEAL2_UserDefined (32453013)_48838802 | hg19 | chr3 | 47036522 | 47039184 | + | AGCACCAGGCTGCTCCTCCATTACTCA | CTTGATGGGCGCCTGACGGGCCACAGA | Homo sapiens | 55.55555556 |
| NBEAL2_UserDefined (32453013)_48838803 | hg19 | chr3 | 47036522 | 47039184 | + | GGAGGATTTGGACAGTCCAGCTCAAGA | GCATGACCTCGTGGGTCCTGAACTGA | Homo sapiens | 51.85185185 |
| NBEAL2_UserDefined (32453013)_48838804 | hg19 | chr3 | 47036522 | 47039184 | + | TTGACACACATCTGTCCCCTTCCTGAT | GTCAGTCATGAAGAACCTCCTGGGAGA | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32453013)_48838805 | hg19 | chr3 | 47036522 | 47039184 | - | TCAGCGGTGGGCAATGACGGCAGCCA | ATGATCTACTCCATCCTGCCCACAGGT | Homo sapiens | 65.38461538 |
| NBEAL2_UserDefined (32453013)_48838806 | hg19 | chr3 | 47036522 | 47039184 | - | AAAGTAGCGCAGAGCACGTCGCAGGACCCTG | TCTTGGCAGCGGGCCTCCAGGGCTAG | Homo sapiens | 60 |
| NBEAL2_UserDefined (32453013)_48838807 | hg19 | chr3 | 47036522 | 47039184 | - | GCAAACTCATGGTCAAATACTCCTTCC | TTTCGCTGGAGTGGTCGGGTGGGGGCAGG | Homo sapiens | 44.44444444 |
| NBEAL2_UserDefined (32453013)_48838808 | hg19 | chr3 | 47036522 | 47039184 | - | TGTGGGGCAGCACGGGACACAAGGGC | ATGGACGATAGCCACGCAGTGCTAGA | Homo sapiens | 69.23076923 |
| NBEAL2_UserDefined (32453013)_48838809 | hg19 | chr3 | 47036522 | 47039184 | - | TGCCCCACCGAGTGCTCCTGGGTGC | TATCCAGCGGAGCCGATACAGCAGGAGGA | Homo sapiens | 70.83333333 |
| NBEAL2_UserDefined (32453013)_48838810 | hg19 | chr3 | 47036522 | 47039184 | - | ACTCCCCTGGGGCCCCCAAATACCTG | CCCCCGGACCCCTAGCTTCTGG | Homo sapiens | 65.38461538 |
| NBEAL2_UserDefined (32453013)_48838811 | hg19 | chr3 | 47036522 | 47039184 | - | CAGACCACCCAGAAAGGAAGTTTTG | ATGACTGGGGGACAGGTCCAGGCAGAT | Homo sapiens | 48 |
| NBEAL2_UserDefined (32453013)_48838812 | hg19 | chr3 | 47036522 | 47039184 | - | TTCCTCTCCATCCGTTCCTCTGCCAGC | TTTCAGCTGGACCTGCTTCAGCCTCTT | Homo sapiens | 59.25925926 |
| NBEAL2_UserDefined (32453012)_48838813 | hg19 | chr3 | 47045306 | 47046080 | + | ACCGAGGAGGGTGCGTCCTGGTGGTGT | ACTACTTCCTCAACTTCCCATGCAAG | Homo sapiens | 66.66666667 |
| NBEAL2_UserDefined (32453012)_48838814 | hg19 | chr3 | 47045306 | 47046080 | + | CGTGCCTCAGGCCTTACCCAGGTGAGA | GTGAGGGTCCCACTCTGCACCCCT | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined (32453012)_48838815 | hg19 | chr3 | 47045306 | 47046080 | - | CCATTCCCTGCCATCCTCCTAAAC | TAAAGACCAGCAACCCACCTCACGCT | Homo sapiens | 54.16666667 |
| NBEAL2_UserDefined (32453012)_48838816 | hg19 | chr3 | 47045306 | 47046080 | - | ACTTCACATCCCTCACACCCAGGG | GGCCTGATCGATAAAGAAGAGCTCAAG | Homo sapiens | 58.33333333 |
| NBEAL2_UserDefined (32453012)_48838817 | hg19 | chr3 | 47045306 | 47046080 | - | ACAGGATTCCTTTTGAGGAATAGAGGC | ACAGGGTACTGAGACAGGTCATTGTAG | Homo sapiens | 44.44444444 |
| NBEAL2_UserDefined (32442572)_48838818 | hg19 | chr3 | 47046462 | 47047939 | + | CCAGCCGTCTTCCGGGACCTGTCTAA | TTCCACTATGGCACCCACTACTCCAAT | Homo sapiens | 61.53846154 |
| NBEAL2_UserDefined (32442572)_48838819 | hg19 | chr3 | 47046462 | 47047939 | + | TGGTGGGCAGACTAGGGGGCAGATAA | CAGGCACGCCTGGAGAGCCCTGCCGAT | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined (32442572)_48838820 | hg19 | chr3 | 47046462 | 47047939 | + | GTGCTCACACTGGAGCGGGCAGGTGTA | GCAGCACCGCCAGGCTCTGGTGAGGAA | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined (32442572)_48838821 | hg19 | chr3 | 47046462 | 47047939 | + | TATGTGTCTGCACACCTACACGAGT | CGGGAACGGAAGGCTCTGGAGGGCATTAT | Homo sapiens | 48 |
| NBEAL2_UserDefined (32442572)_48838822 | hg19 | chr3 | 47046462 | 47047939 | + | GGACACTAACTCACCTAGCATCTTCCA | ATAAAGTGTCGACTCAACAAGCTCTGA | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32442572)_48838823 | hg19 | chr3 | 47046462 | 47047939 | - | ATGCTTGGGGTTCACCACACCGATGG | CCACAGGTTGGGGGAGCACCAC | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined (32442572)_48838824 | hg19 | chr3 | 47046462 | 47047939 | - | TCTCCCTGTCCATCCCGACTGCCCCC | AACTGCCGGTCGGAGCAGTCAAAGCT | Homo sapiens | 69.23076923 |
| NBEAL2_UserDefined (32442572)_48838825 | hg19 | chr3 | 47046462 | 47047939 | - | TGCTTGTAGCCAAAGATGAGGTCGATC | TGGCCCACGGGGGTAGCACCACAT | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32442572)_48838826 | hg19 | chr3 | 47046462 | 47047939 | - | AAGAATGCCTTGAGTTCGTCCAGGTG | TCATCTGTCACATGGTCCAGGTCTA | Homo sapiens | 50 |
| NBEAL2_UserDefined (32453016)_48838827 | hg19 | chr3 | 47048556 | 47050158 | + | CTGGTGCTAGCCCTGGTCCCCCA | CCTATGACCGCAACATAAGCAACTAC | Homo sapiens | 69.56521739 |
| NBEAL2_UserDefined (32453016)_48838828 | hg19 | chr3 | 47048556 | 47050158 | + | CCGGAATTATTCCCCTCCCTAATCAGA | GGATGGAAAGCTGCTATTCAGCGGT | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32453016)_48838829 | hg19 | chr3 | 47048556 | 47050158 | + | CTTGGAGCTGGGGAGAATGAGGCACA | TCAGGTGTGCCTCGTGGGCAGCTCTGT | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined (32453016)_48838830 | hg19 | chr3 | 47048556 | 47050158 | + | CTCTAGTCTTGGACAAGGCCTTCATT | GTATGCATGTGCCCTGGGGAGGGTGAGTTT | Homo sapiens | 46.15384615 |
| NBEAL2_UserDefined (32453016)_48838831 | hg19 | chr3 | 47048556 | 47050158 | + | CACATTCCCTGGACCTATTTTCCACCT | TTCACTGCCCCTGGCAGAGCAGCCTA | Homo sapiens | 48.14814815 |
| NBEAL2_UserDefined (32453016)_48838832 | hg19 | chr3 | 47048556 | 47050158 | - | AATGCACTTACCAACAGGTCTGGGGAAC | GTCATAGAGTAGGTGTCACAGATGCAG | Homo sapiens | 50 |
| NBEAL2_UserDefined (32453016)_48838833 | hg19 | chr3 | 47048556 | 47050158 | - | AAGAGACCCAGGCCCTCAGGGAGCCTGGC | AACCAGCTGTGGGTGCCCAGCAGCCCAC | Homo sapiens | 68.96551724 |
| NBEAL2_UserDefined (32453016)_48838834 | hg19 | chr3 | 47048556 | 47050158 | - | AGAACGGAGGAGTCACATCCGGCCTGG | CCAGTGCTTGTCCACTCACACC | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined (32453016)_48838835 | hg19 | chr3 | 47048556 | 47050158 | - | ACAGACCACCCTGTGGAGGACCAGGGT | TGTCCCGGAGCCTGAGATGAGGTAG | Homo sapiens | 62.96296296 |
| NBEAL2_UserDefined (32453016)_48838836 | hg19 | chr3 | 47048556 | 47050158 | - | ACAATCTGGCCTTCGGACCCCAATGC | ACAGCCATGTCAAGTTCAGTGCTGA | Homo sapiens | 57.69230769 |
| NBEAL2_UserDefined (32453016)_48838837 | hg19 | chr3 | 47048556 | 47050158 | - | GAAGATGAAGGGACAGGTAACAAACCT | CCCGCAACTTCCCATTGACTGAATA | Homo sapiens | 44.44444444 |
| NBEAL2_UserDefined (32453014)_48838838 | hg19 | chr3 | 47042491 | 47044799 | + | CGTGGCGTGGCGTGGAAGGCAGCGATGA | AGTTCCCCACTCACCCGCCAGCCT | Homo sapiens | 67.85714286 |
| NBEAL2_UserDefined (32453014)_48838839 | hg19 | chr3 | 47042491 | 47044799 | + | GTGCTGAAGGCCATGGTAACCAGGAA | TGGCCCCCAGCGCAAACTTTACTTT | Homo sapiens | 53.84615385 |
| NBEAL2_UserDefined (32453014)_48838840 | hg19 | chr3 | 47042491 | 47044799 | + | AACGGCACAGCTGATCTCCGTGAGAT | ACATGCTGCTACAGACTGCAGTG | Homo sapiens | 53.84615385 |
| NBEAL2_UserDefined (32453014)_48838841 | hg19 | chr3 | 47042491 | 47044799 | + | GCTGCAGCAGCTGCAGCAGCTGCAGA | GGTGCCTGGAGGTTGGGGCCCAGGAAGA | Homo sapiens | 65.38461538 |
| NBEAL2_UserDefined (32453014)_48838842 | hg19 | chr3 | 47042491 | 47044799 | + | GGTTTCTGGAATGCCTGCTATGACAT | CCCCACCCCAGGAGCTGGTGCTGGAA | Homo sapiens | 46.15384615 |
| NBEAL2_UserDefined (32453014)_48838843 | hg19 | chr3 | 47042491 | 47044799 | + | GCCCATGTGGGGCCTGGGCGCTGAGGT | CGTCTGAAGCTGGTGCCCAACCATCA | Homo sapiens | 74.07407407 |
| NBEAL2_UserDefined (32453014)_48838844 | hg19 | chr3 | 47042491 | 47044799 | + | CCTGACACCCACCGAGGAGGCCTCACT | GGATCAACTCCCTGAGTGTGTCCTCTT | Homo sapiens | 66.66666667 |
| NBEAL2_UserDefined (32453014)_48838845 | hg19 | chr3 | 47042491 | 47044799 | - | AAAACCTGGCCACGCTCCCGCCAGGCAG | TAGCCACCCAGGAGTCAGAGCATTTG | Homo sapiens | 67.85714286 |
| NBEAL2_UserDefined (32453014)_48838846 | hg19 | chr3 | 47042491 | 47044799 | - | TCTGAGGGGTCGCACCTTCTCACTCC | GGGTCTACCTGACTATCTTCTCTCCTT | Homo sapiens | 61.53846154 |
| NBEAL2_UserDefined (32453014)_48838847 | hg19 | chr3 | 47042491 | 47044799 | - | AAGTACAAGCCGTAGGCCAATCTGCG | TAGGGTTTGTGCACACAGGCAAGTCC | Homo sapiens | 53.84615385 |
| NBEAL2_UserDefined (32453014)_48838848 | hg19 | chr3 | 47042491 | 47044799 | - | ACCAGTGGCACCAGCCAGGAGCA | AATCTCACGTAGGCCTGGGCATGCAGCTGT | Homo sapiens | 65.2173913 |
| NBEAL2_UserDefined (32453014)_48838849 | hg19 | chr3 | 47042491 | 47044799 | - | CGCCGCTGCCCACTGCTCATAA | TTTGTCGATGAAGTGGCGCCATTCGGGT | Homo sapiens | 63.63636364 |
| NBEAL2_UserDefined (32453014)_48838850 | hg19 | chr3 | 47042491 | 47044799 | - | ACAGCGACCCTGCCCCAAGCCCGGCC | TAATAGGTCAGTCCCCTGGCTGCTGCC | Homo sapiens | 76.92307692 |
| NBEAL2_UserDefined (32453014)_48838851 | hg19 | chr3 | 47042491 | 47044799 | - | TTTGGCCTCTTTGGTCACTGCCAGAGG | ATGCGTGAATATGTCTCGGCGGCTGG | Homo sapiens | 55.55555556 |
| NBEAL2_FiveUtrExon (32453028)_48838852 | hg19 | chr3 | 47021302 | 47021402 | - | AAGCTGGGGGCACCGCGGGTGCGGGT | TACTACTGGGGAAGTCTGCGGCTGTG | Homo sapiens | 76.92307692 |

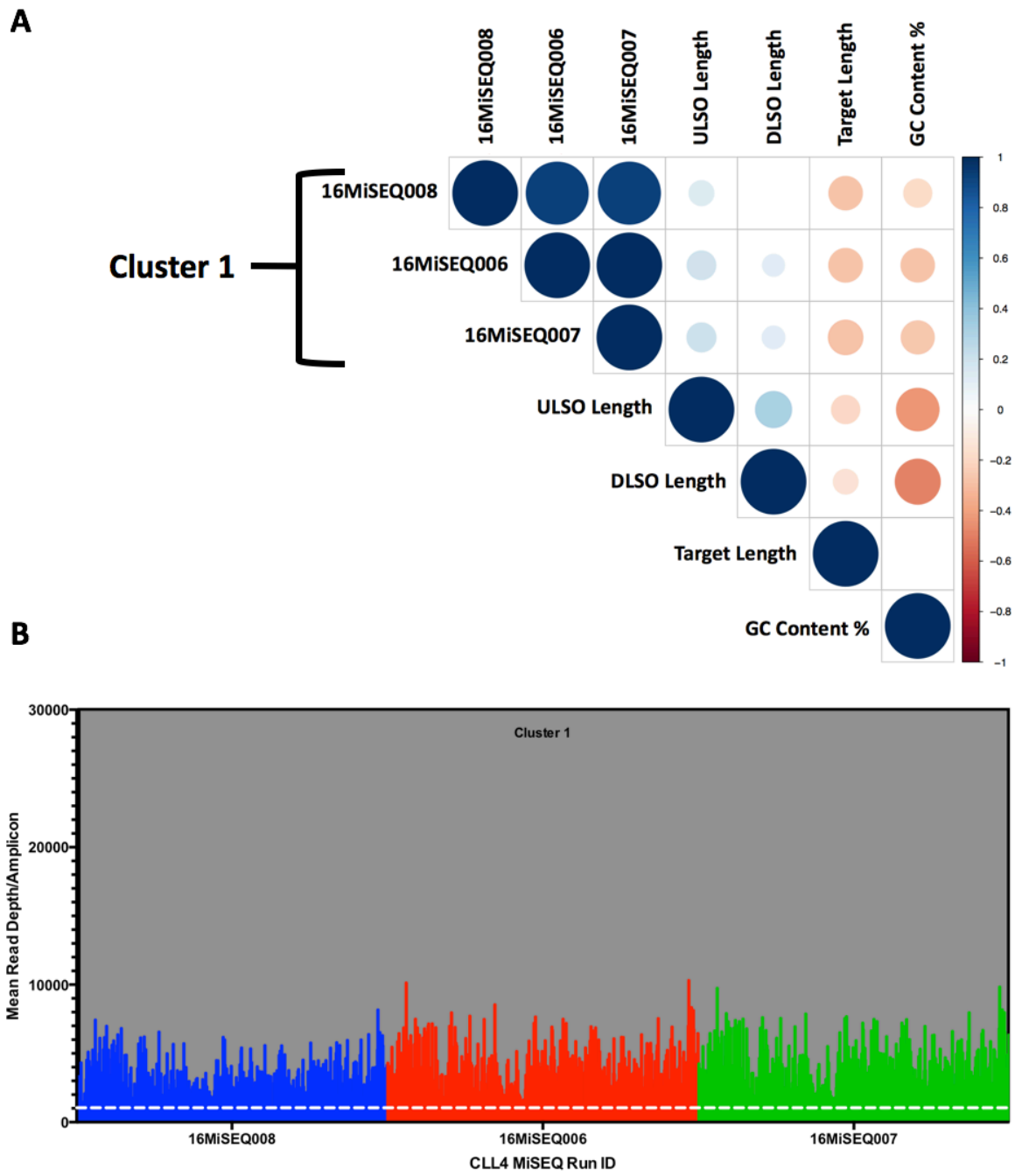
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CTBP2_Exon (32453039)_48838867 | hg19 | chr10 | 126681264 | 126681381 | - | CTTATTTCCCCACATAAGCAGCTACAG | GAGAAAAGGAACTGACTTTGTCTCAGC | Homo sapiens | 44.44444444 |
| CTBP2_Exon (32129228)_48838868 | hg19 | chr10 | 126683038 | 126683252 | - | TCACCTGGACGCTCTGTTTCTGTACAG | TTTGAAAATAGATGCCATGGAAGCTTC | Homo sapiens | 51.85185185 |
| SAMHD1_Exon (32442560)_48838869 | hg19 | chr20 | 35526225 | 35526362 | - | AAAGGAATAGTTAGGAGCCTAGGGACC | GTGGCCTGTTACTCACAAACTTTGC | Homo sapiens | 48.14814815 |
| SAMHD1_Exon (32442563)_48838870 | hg19 | chr20 | 35526843 | 35526947 | - | CCTCTGTTCCATTTCTTGTTATGCTCC | ACTTAGGCAAAACAAATATACTATAGTGG | Homo sapiens | 44.44444444 |
| SAMHD1_Exon (32442565)_48838871 | hg19 | chr20 | 35533767 | 35533906 | - | GCGAACTGCCTGTTAAGATCATTTCAA | GAGTTATCAGGCATCTTTCCATTAGTGAA | Homo sapiens | 40.74074074 |
| SAMHD1_Exon (32442567)_48838872 | hg19 | chr20 | 35540864 | 35540955 | - | AGTGGGGTCAAAGGTGTTAATGAAGAA | AAGAAACTTGATTGTCATTTCCCTATAAA | Homo sapiens | 40.74074074 |
| SAMHD1_Exon (32442568)_48838873 | hg19 | chr20 | 35547767 | 35547922 | - | GGCTCCCAATGGGCTAGAATCTTTAAC | AGTTTATGTACAGCTTAATTTTAACTGACT | Homo sapiens | 48.14814815 |
| SAMHD1_Exon (32105522)_48838874 | hg19 | chr20 | 35555585 | 35555655 | - | GGAGATTGTGTTGAGCCAAGATTGTAC | TCAATAAAAAGAAAGACTATTTAAGCAGGT | Homo sapiens | 44.44444444 |
| SAMHD1_Exon (32442561)_48838875 | hg19 | chr20 | 35559163 | 35559278 | - | GGAAACATACTGAAGTTCTTGAGGACC | GCTGTACAAAGAGACAAAGTTGTTATG | Homo sapiens | 44.44444444 |
| SAMHD1_Exon (32105524)_48838876 | hg19 | chr20 | 35563432 | 35563592 | - | GCACTATAATTAGGAATGGCTGCACAC | TAAGAAGGGAATGGGGTGGGGAACTT | Homo sapiens | 44.44444444 |
| SAMHD1_Exon (32442562)_48838877 | hg19 | chr20 | 35569442 | 35569514 | - | AAGCAAATAATGATAGACTTGCTTATCC | CTGCTTCTCAGTGATAGAATTGTGAGT | Homo sapiens | 32.14285714 |
| SAMHD1_Exon (32105526)_48838878 | hg19 | chr20 | 35575141 | 35575207 | - | AAGTGGTACTTGGCACAATCCATTGCC | TGAAGAAATAAATTCACGTTATAACCAAAC | Homo sapiens | 48.14814815 |
| SAMHD1_UserDefined (32442559)_48838879 | hg19 | chr20 | 35520177 | 35521569 | + | CTGTCACTGGGGTAGGATCTCCCAACAA | TTCACTTATGCTTAGGTCACTCTTAGAGG | Homo sapiens | 53.57142857 |
| SAMHD1_UserDefined (32442559)_48838880 | hg19 | chr20 | 35520177 | 35521569 | + | TTTGTTTTTAGCCATGTTTGTCAATGAC | GAGTCCAAGCTGGGCAACAGAGCAAAA | Homo sapiens | 32.14285714 |
| SAMHD1_UserDefined (32442559)_48838881 | hg19 | chr20 | 35520177 | 35521569 | + | TTGCCTAATACCATCTTATTTCTTTGATT | TTTTGGATGCTTCTCGGAGGCGAGTT | Homo sapiens | 27.5862069 |
| SAMHD1_UserDefined (32442559)_48838882 | hg19 | chr20 | 35520177 | 35521569 | + | GAAAGTCTCAAGTGCGCACATCCTCAT | CTCTGTCCTGTTTTCCCAGGAGAAAGA | Homo sapiens | 48.14814815 |
| SAMHD1_UserDefined (32442559)_48838883 | hg19 | chr20 | 35520177 | 35521569 | - | GTGTATTTAAGAGCAAACACTAGCCAG | TGCTGAAATTAGTGACCACCCAGTG | Homo sapiens | 40.74074074 |
| SAMHD1_UserDefined (32442559)_48838884 | hg19 | chr20 | 35520177 | 35521569 | - | TCTTCGTGAGTGGAACATAAAGCCGCC | TCTCAGTATCCTGGAGAGCTGGCCAG | Homo sapiens | 51.85185185 |
| SAMHD1_UserDefined (32442559)_48838885 | hg19 | chr20 | 35520177 | 35521569 | - | TTTTGTTTTTAATGTAGAAATGAGACAAGG | TACAAGGTGCCAGGTTGGGGTAAAGAG | Homo sapiens | 26.66666667 |
| SAMHD1_UserDefined (32442559)_48838886 | hg19 | chr20 | 35520177 | 35521569 | - | CACGCATGCTGAAGCTAAGTAACTTT | AACTTTCCTGGGCTCAAGCGACCCTCC | Homo sapiens | 42.30769231 |
| SAMHD1_UserDefined (32442559)_48838887 | hg19 | chr20 | 35520177 | 35521569 | - | GGACAGCTAGAAGCAATCTATATGCAA | CAGAGTCCAGCTTTTTAAAGATGACCC | Homo sapiens | 40.74074074 |

## A.2 Correlation analysis of the sequencing coverage in Design 2



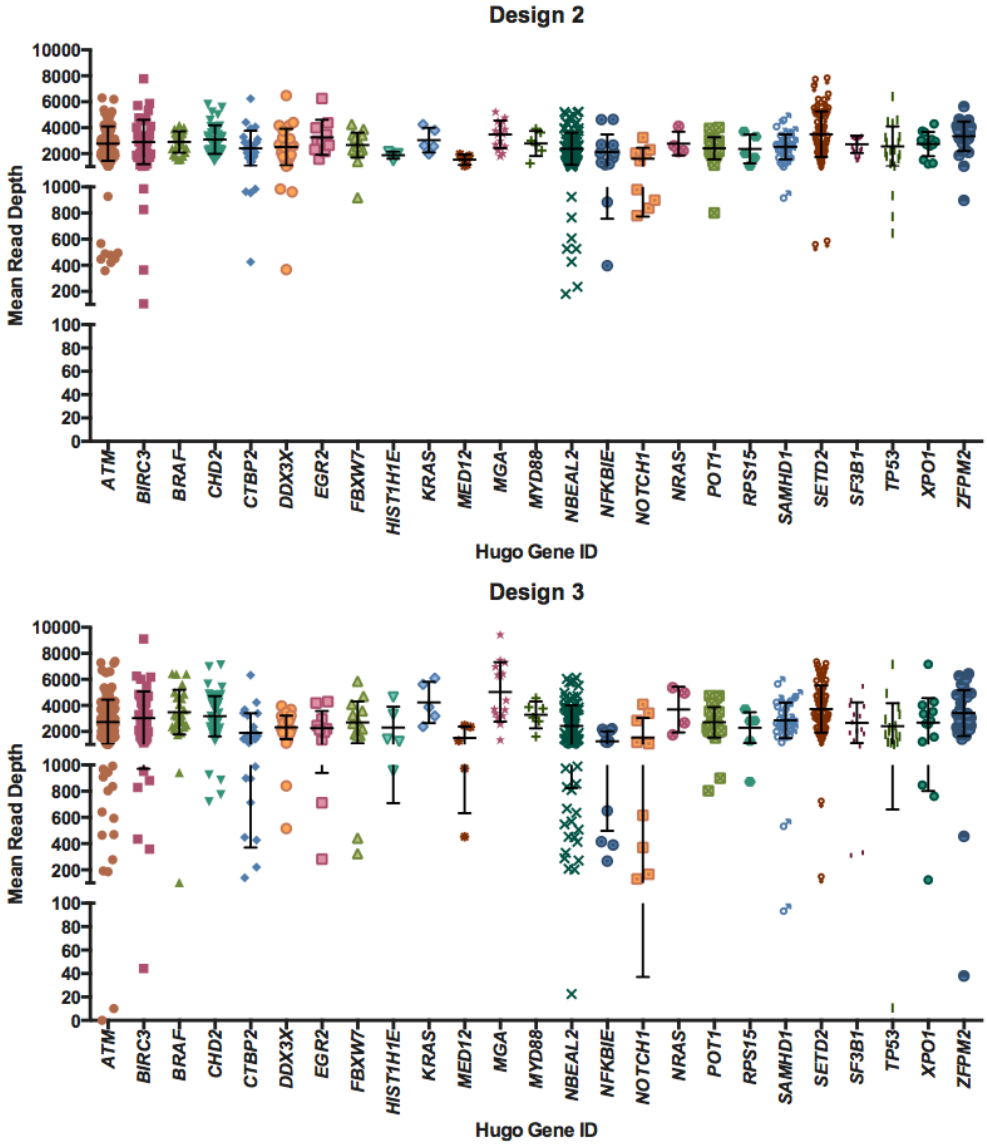

**A** Correlation matrix with hierarchical clustering of CLL4 TruSeq Design 2 MiSEQ runs (n=15) read depth per amplicon (n=618), including ULSO, DLSO and target length, as well as GC content % per amplicon. Blue represents positive correlations (0 to 1), red represents negative correlations (0 to -1). Size of circle represents strength of correlation score. Non-significant correlations (p=>0.01) excluded from correlation matrix. Clustering of MiSEQ runs based on similarity denoted by open brackets. **B** Histogram denoting the read depth per amplicon, per MiSEQ run (n=15) for CLL4 TruSeq Design 1. Histogram ordered based on correlation matrix hierarchical clustering.

## A.3 Correlation analysis of the sequencing coverage in Design 3




**A** Correlation matrix with hierarchical clustering of CLL4 TruSeq Design 3 MiSEQ runs (n=3) read depth per amplicon (n=610), including ULSO, DLSO and target length, as well as GC content % per amplicon. Blue represents positive correlations (0 to 1), red represents negative correlations (0 to -1). Size of circle represents strength of correlation score. Non-significant correlations (p=>0.01) excluded from correlation matrix. Clustering of MiSEQ runs based on similarity denoted by open brackets. **B** Histogram denoting the read depth per amplicon, per MiSEQ run (n=3) for CLL4 TruSeq Design 1. Histogram ordered based on correlation matrix hierarchical clustering.

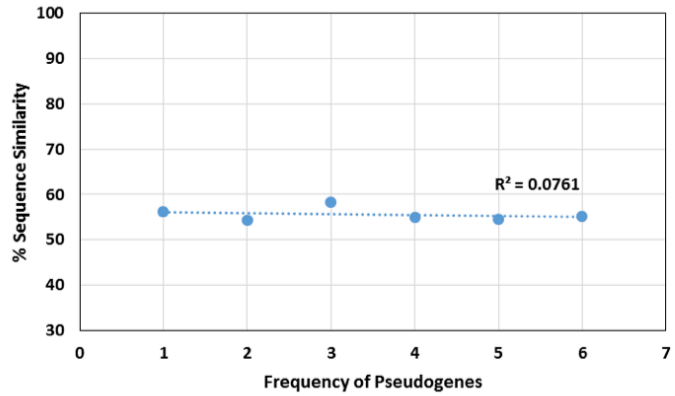
## A.4 Mean sequencing coverage per amplicon, per gene for the CLL4 TruSeq panel Design 2 and 3



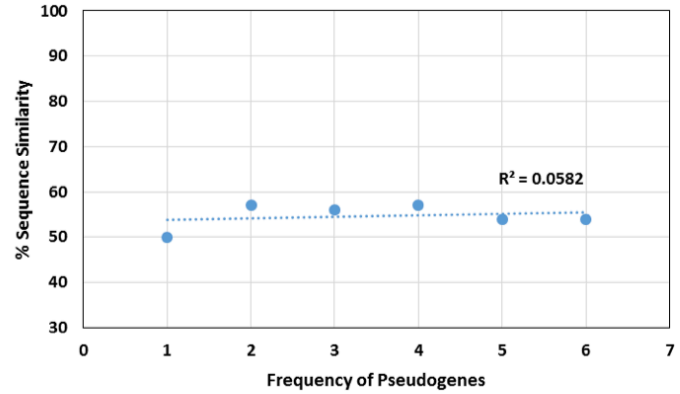

Mean sequencing coverage per gene. Dots represent mean coverage per amplicon. Bars represent mean coverage per gene and standard deviation.

## A.5 Sequence identity between *CTBP2* and *CTBP2* pseudogenes using the Pei *et al.* (2012)(250) dataset

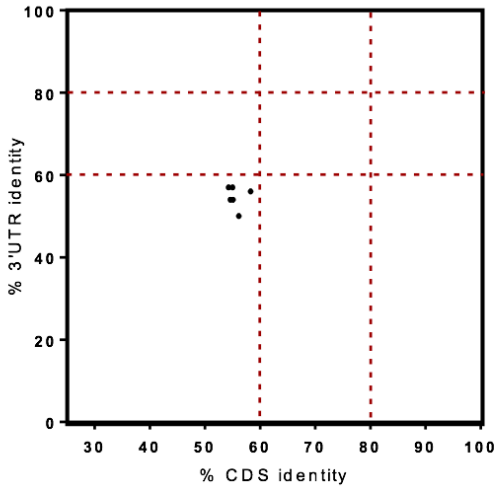
**A**




**B**



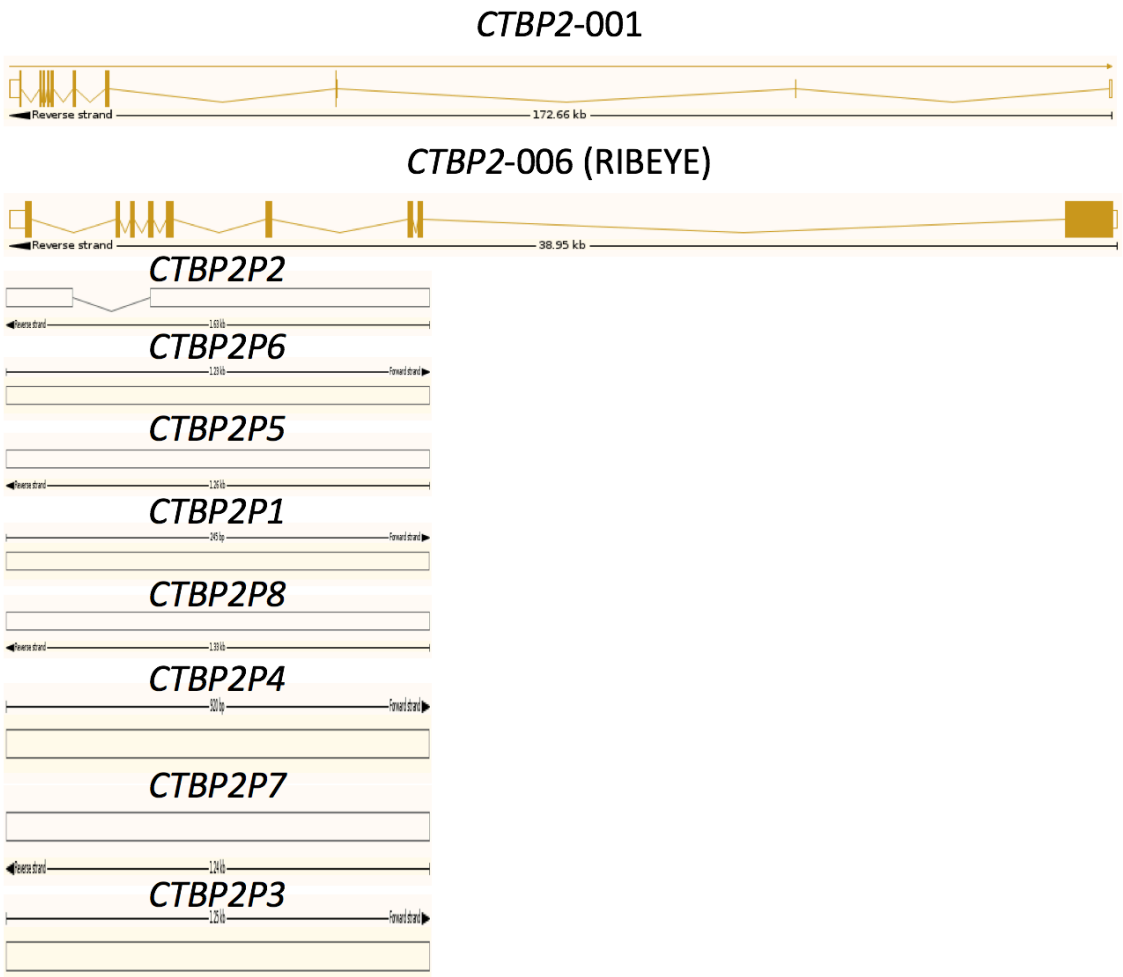

**C**




| Pseudogene Name | % CDS identity | 3'UTR identity (%) |
|---|---|---|
| *CTBP2P4* | 56.14 | 50 |
| *CTBP2P8* | 54.29 | 57 |
| *CTBP2P1* | 58.28 | 56 |
| *CTBP2P5* | 54.98 | 57 |
| *CTBP2P6* | 54.60 | 54 |
| *CTBP2P2* | 55.10 | 54 |

**A** Distribution of pseudogene sequence identity to coding exons (CDS) of *CTBP2*. **B** Distribution of pseudogene sequence identity to the 3'UTR of *CTBP2*. **C** Scatter plot of sequence identity of *CTBP2* pseudogenes to the CDS and 3'UTR regions of *CTBP2*. This analysis identifies that the pseudogenes of *CTBP2* fall into the lowest identity category identified by Pei *et al.* (2012)(250).

## A.6 Mapping of *CTBP2* pseudogenes




*MGC70870 does not have Ensembl annotation

Mapped using a local ClustalV pairwise alignment, a graphic representation of the mapping of each pseudogene to *CTBP2*.

# A.7 MAF file of high confidence variants (n=903)

| Hugo_Symbol | Entrez_Gene_Id | Center | NCBI_Build | Chromosome | Start_Position | End_Position | Strand | Variant_Classification | Variant_Type | Referece_Allele | Tumor_Seq_Allele1 | dbSNP_RS | Tumor_Sample_Barcode | Sequencing_Source | Sequencer | amino_acid_change | Pathway |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108180981 | 108180981 | + | Missense_Mutation | SNP | A | G | novel | 4039 | AMPLICON | Illumina MiSeq | p.T1953A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143528 | 108143528 | + | Missense_Mutation | SNP | T | G | novel | 4096 | AMPLICON | Illumina MiSeq | p.L1078R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108199926 | 108199926 | + | Missense_Mutation | SNP | A | G | novel | 4096 | AMPLICON | Illumina MiSeq | p.E2423G | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108205781 | 108205781 | + | Missense_Mutation | SNP | C | G | novel | 4133 | AMPLICON | Illumina MiSeq | p.P2699R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108170460 | 108170460 | + | Missense_Mutation | SNP | G | T | novel | 4384 | AMPLICON | Illumina MiSeq | p.L1675F | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236086 | 108236086 | + | Missense_Mutation | SNP | C | T | novel | 4393 | AMPLICON | Illumina MiSeq | p.R3008C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108186757 | 108186757 | + | Missense_Mutation | SNP | G | A | novel | 4426 | AMPLICON | Illumina MiSeq | p.E2039K | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108205780 | 108205780 | + | Missense_Mutation | SNP | C | T | novel | 4495 | AMPLICON | Illumina MiSeq | p.P2699S | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236086 | 108236086 | + | Missense_Mutation | SNP | C | T | novel | 4495 | AMPLICON | Illumina MiSeq | p.R3008C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108170479 | 108170479 | + | Missense_Mutation | SNP | G | C | novel | 4519 | AMPLICON | Illumina MiSeq | p.D1682H | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4002 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4006 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4006 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108180981 | 108180981 | + | Missense_Mutation | SNP | A | G | novel | 4006 | AMPLICON | Illumina MiSeq | p.T1953A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236086 | 108236086 | + | Missense_Mutation | SNP | C | T | novel | 4011 | AMPLICON | Illumina MiSeq | p.R3008C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4016 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4023 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108206581 | 108206581 | + | Missense_Mutation | SNP | G | A | novel | 4030 | AMPLICON | Illumina MiSeq | p.D2721N | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4036 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108201071 | 108201071 | + | Missense_Mutation | SNP | C | T | novel | 4042 | AMPLICON | Illumina MiSeq | p.H2480Y | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108205780 | 108205780 | + | Missense_Mutation | SNP | C | T | novel | 4060 | AMPLICON | Illumina MiSeq | p.P2699S | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4061 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4061 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108106541 | 108106541 | + | Frame_Shift_Del | DEL | TATCTC | T | novel | 4062 | AMPLICON | Illumina MiSeq | p.I159fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236096 | 108236096 | + | Missense_Mutation | SNP | T | A | novel | 4065 | AMPLICON | Illumina MiSeq | p.M3011K | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108119823 | 108119823 | + | Missense_Mutation | SNP | T | C | rs56128736 | 4071 | AMPLICON | Illumina MiSeq | p.V410A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4080 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108175463 | 108175463 | + | Missense_Mutation | SNP | A | T | rs1801673 | 4081 | AMPLICON | Illumina MiSeq | p.D1853V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108218089 | 108218089 | + | Missense_Mutation | SNP | C | G | novel | 4083 | AMPLICON | Illumina MiSeq | p.L2890V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4089 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4102 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4102 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4105 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4105 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108206581 | 108206581 | + | Missense_Mutation | SNP | G | A | novel | 4106 | AMPLICON | Illumina MiSeq | p.D2721N | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108121593 | 108121593 | + | Frame_Shift_Del | DEL | CAA | C | novel | 4120 | AMPLICON | Illumina MiSeq | p.K468fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108158425 | 108158425 | + | Frame_Shift_Del | DEL | CCTCTGTGACTTTTCAGG | C | novel | 4133 | AMPLICON | Illumina MiSeq | p.L1365fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4136 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4140 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236074 | 108236074 | + | Missense_Mutation | SNP | A | G | novel | 4142 | AMPLICON | Illumina MiSeq | p.K3004E | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4174 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108202647 | 108202647 | + | Frame_Shift_Del | DEL | GT | G | novel | 4175 | AMPLICON | Illumina MiSeq | p.F2558fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108186757 | 108186757 | + | Missense_Mutation | SNP | G | A | novel | 4201 | AMPLICON | Illumina MiSeq | p.E2039K | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108119823 | 108119823 | + | Missense_Mutation | SNP | T | C | rs56128736 | 4205 | AMPLICON | Illumina MiSeq | p.V410A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108119823 | 108119823 | + | Missense_Mutation | SNP | T | C | rs56128736 | 4211 | AMPLICON | Illumina MiSeq | p.V410A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4256 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4256 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108121593 | 108121593 | + | Frame_Shift_Del | DEL | CAA | C | novel | 4257 | AMPLICON | Illumina MiSeq | p.K468fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108199929 | 108199929 | + | Missense_Mutation | SNP | T | G | rs28904921 | 4259 | AMPLICON | Illumina MiSeq | p.V2424G | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4282 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4282 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108216635 | 108216635 | + | Missense_Mutation | SNP | G | A | novel | 4283 | AMPLICON | Illumina MiSeq | p.V2862I | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108117837 | 108117837 | + | Missense_Mutation | SNP | G | A | novel | 4285 | AMPLICON | Illumina MiSeq | p.A350T | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108139214 | 108139214 | + | Frame_Shift_Del | DEL | TTGTG | T | novel | 4289 | AMPLICON | Illumina MiSeq | p.L906fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4294 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4295 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4295 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108175463 | 108175463 | + | Missense_Mutation | SNP | A | T | rs1801673 | 4299 | AMPLICON | Illumina MiSeq | p.D1853V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4300 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108121593 | 108121593 | + | Frame_Shift_Del | DEL | CAA | C | novel | 4319 | AMPLICON | Illumina MiSeq | p.K468fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4330 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4334 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4334 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108175463 | 108175463 | + | Missense_Mutation | SNP | A | T | rs1801673 | 4344 | AMPLICON | Illumina MiSeq | p.D1853V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4347 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108114797 | 108114797 | + | Frame_Shift_Del | DEL | TA | T | novel | 4351 | AMPLICON | Illumina MiSeq | p.L205fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108216545 | 108216545 | + | Missense_Mutation | SNP | C | T | novel | 4351 | AMPLICON | Illumina MiSeq | p.R2832C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4358 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4364 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4364 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4366 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4366 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108186610 | 108186610 | + | Missense_Mutation | SNP | G | A | rs11212587 | 4367 | AMPLICON | Illumina MiSeq | p.G2023R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4368 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4368 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108218052 | 108218052 | + | Missense_Mutation | SNP | G | C | novel | 4369 | AMPLICON | Illumina MiSeq | p.L2877F | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4396 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4414 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4414 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4432 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108188115 | 108188115 | + | Missense_Mutation | SNP | G | A | novel | 4433 | AMPLICON | Illumina MiSeq | p.G2072R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108198384 | 108198384 | + | Frame_Shift_Del | DEL | CTAAAACT | C | novel | 4433 | AMPLICON | Illumina MiSeq | p.L2330fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108206616 | 108206616 | + | Missense_Mutation | SNP | C | G | novel | 4464 | AMPLICON | Illumina MiSeq | p.F2732L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108180945 | 108180945 | + | Missense_Mutation | SNP | G | C | novel | 4470 | AMPLICON | Illumina MiSeq | p.V1941L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4474 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4483 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098352 | 108098352 | + | Missense_Mutation | SNP | A | G | novel | 4490 | AMPLICON | Illumina MiSeq | p.M1V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4499 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4499 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108114797 | 108114797 | + | Frame_Shift_Del | DEL | TA | T | novel | 4510 | AMPLICON | Illumina MiSeq | p.L205fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108216545 | 108216545 | + | Missense_Mutation | SNP | C | T | novel | 4510 | AMPLICON | Illumina MiSeq | p.R2832C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108123551 | 108123551 | + | Missense_Mutation | SNP | C | T | rs2227922 | 4521 | AMPLICON | Illumina MiSeq | p.P604S | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108137925 | 108137925 | + | Missense_Mutation | SNP | C | T | novel | 4521 | AMPLICON | Illumina MiSeq | p.R832C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108117794 | 108117794 | + | In_Frame_Del | DEL | ATTTCGTAATATTGCC | A | novel | 4524 | AMPLICON | Illumina MiSeq | p.336_340del | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108202225 | 108202225 | + | Missense_Mutation | SNP | G | A | novel | 4524 | AMPLICON | Illumina MiSeq | p.A2524T | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4528 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4533 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4533 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4550 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4550 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4554 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4588 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4598 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4598 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4607 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4616 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236086 | 108236086 | + | Missense_Mutation | SNP | C | T | novel | 4626 | AMPLICON | Illumina MiSeq | p.R3008C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4629 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4633 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4633 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108190707 | 108190707 | + | Frame_Shift_Ins | INS | A | AT | novel | 4634 | AMPLICON | Illumina MiSeq | p.H2125fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4648 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4655 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4655 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4657 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4660 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4668 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4678 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4678 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4687 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108236087 | 108236087 | + | Missense_Mutation | SNP | G | A | novel | 4688 | AMPLICON | Illumina MiSeq | p.R3008H | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108138003 | 108138003 | + | Missense_Mutation | SNP | T | C | rs1800056 | 4696 | AMPLICON | Illumina MiSeq | p.F858L | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4696 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108124761 | 108124761 | + | Missense_Mutation | SNP | T | C | rs4986761 | 4711 | AMPLICON | Illumina MiSeq | p.S707P | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108121593 | 108121593 | + | Frame_Shift_Del | DEL | CAA | C | novel | 4715 | AMPLICON | Illumina MiSeq | p.K468fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108098576 | 108098576 | + | Missense_Mutation | SNP | C | G | rs1800054 | 4725 | AMPLICON | Illumina MiSeq | p.S49C | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4725 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108175463 | 108175463 | + | Missense_Mutation | SNP | A | T | rs1801673 | 4727 | AMPLICON | Illumina MiSeq | p.D1853V | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108143456 | 108143456 | + | Missense_Mutation | SNP | C | G | rs1800057 | 4749 | AMPLICON | Illumina MiSeq | p.P1054R | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108121593 | 108121593 | + | Frame_Shift_Del | DEL | CAA | C | novel | 4751 | AMPLICON | Illumina MiSeq | p.K468fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108119823 | 108119823 | + | Missense_Mutation | SNP | T | C | rs56128736 | 4755 | AMPLICON | Illumina MiSeq | p.V410A | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108192065 | 108192065 | + | Missense_Mutation | SNP | G | A | novel | 4772 | AMPLICON | Illumina MiSeq | p.E2164K | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108199899 | 108199899 | + | Frame_Shift_Ins | INS | A | AG | novel | 4772 | AMPLICON | Illumina MiSeq | p.Q2414fs | DDR/CC |
| ATM | 472 | Southampton_Cancer_Genomics | hg19 | 11 | 108199929 | 108199929 | + | Missense_Mutation | SNP | T | G | rs28904921 | 4775 | AMPLICON | Illumina MiSeq | p.V2424G | DDR/CC |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207704 | 102207704 | + | Frame_Shift_Del | DEL | CAAAG | C | novel | 4036 | AMPLICON | Illumina MiSeq | p.K563fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201765 | 102201765 | + | Missense_Mutation | SNP | G | A | novel | 4047 | AMPLICON | Illumina MiSeq | p.D373N | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206864 | 102206864 | + | Frame_Shift_Del | DEL | AC | A | novel | 4062 | AMPLICON | Illumina MiSeq | p.T498fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4101 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4104 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201962 | 102201962 | + | Frame_Shift_Del | DEL | AAAAGAATC | A | novel | 4104 | AMPLICON | Illumina MiSeq | p.K439fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201966 | 102201966 | + | Nonsense_Mutation | SNP | G | T | novel | 4145 | AMPLICON | Illumina MiSeq | p.E440X | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207821 | 102207821 | + | Frame_Shift_Del | DEL | AT | A | novel | 4196 | AMPLICON | Illumina MiSeq | p.F602fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201790 | 102201790 | + | Frame_Shift_Del | DEL | TGA | T | novel | 4264 | AMPLICON | Illumina MiSeq | p.V381fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4280 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207530 | 102207530 | + | Nonsense_Mutation | SNP | C | G | novel | 4282 | AMPLICON | Illumina MiSeq | p.S540X | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201943 | 102201943 | + | Frame_Shift_Ins | INS | G | GTTTA | novel | 4322 | AMPLICON | Illumina MiSeq | p.R432fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207651 | 102207651 | + | Frame_Shift_Ins | INS | G | GA | novel | 4327 | AMPLICON | Illumina MiSeq | p.E545fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201814 | 102201814 | + | Missense_Mutation | SNP | G | T | novel | 4335 | AMPLICON | Illumina MiSeq | p.G389V | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207508 | 102207508 | + | Nonsense_Mutation | SNP | T | TA | novel | 4357 | AMPLICON | Illumina MiSeq | p.Y533_I534delinsX | NFkB |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4359 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201928 | 102201928 | + | Frame_Shift_Ins | INS | T | TA | novel | 4362 | AMPLICON | Illumina MiSeq | p.I427fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4362 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206700 | 102206700 | + | Frame_Shift_Del | DEL | AT | A | novel | 4368 | AMPLICON | Illumina MiSeq | p.D443fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4374 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207659 | 102207659 | + | Frame_Shift_Ins | INS | A | AT | novel | 4374 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4389 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201919 | 102201919 | + | Frame_Shift_Ins | INS | A | AG | novel | 4399 | AMPLICON | Illumina MiSeq | p.E424fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201921 | 102201921 | + | Frame_Shift_Ins | INS | G | GA | novel | 4399 | AMPLICON | Illumina MiSeq | p.D425fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4425 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207736 | 102207736 | + | Missense_Mutation | SNP | G | T | novel | 4479 | AMPLICON | Illumina MiSeq | p.G573V | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201928 | 102201928 | + | Frame_Shift_Del | DEL | TAAGGG | T | novel | 4537 | AMPLICON | Illumina MiSeq | p.I427fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201970 | 102201970 | + | Nonsense_Mutation | SNP | C | A | novel | 4566 | AMPLICON | Illumina MiSeq | p.S441X | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201907 | 102201907 | + | Frame_Shift_Ins | INS | T | TA | novel | 4602 | AMPLICON | Illumina MiSeq | p.L420fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207489 | 102207489 | + | Splice_Site | INS | A | AGT | novel | 4602 | AMPLICON | Illumina MiSeq | NA | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4602 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207692 | 102207692 | + | Frame_Shift_Ins | INS | A | AG | novel | 4614 | AMPLICON | Illumina MiSeq | p.K558fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4678 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4687 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201930 | 102201930 | + | Frame_Shift_Del | DEL | AG | A | novel | 4696 | AMPLICON | Illumina MiSeq | p.R428fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4696 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207804 | 102207804 | + | Nonsense_Mutation | SNP | A | T | novel | 4696 | AMPLICON | Illumina MiSeq | p.K596X | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207684 | 102207684 | + | Frame_Shift_Del | DEL | AC | A | novel | 4733 | AMPLICON | Illumina MiSeq | p.T556fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207713 | 102207713 | + | Frame_Shift_Del | DEL | GTC | G | novel | 4767 | AMPLICON | Illumina MiSeq | p.S566fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207675 | 102207675 | + | Frame_Shift_Del | DEL | GAAGA | G | novel | 4769 | AMPLICON | Illumina MiSeq | p.E553fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201964 | 102201964 | + | Frame_Shift_Del | DEL | AAG | A | novel | 4051 | AMPLICON | Illumina MiSeq | p.K439fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206699 | 102206699 | + | Frame_Shift_Del | DEL | GATTT | G | novel | 4104 | AMPLICON | Illumina MiSeq | p.D443fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4105 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201928 | 102201928 | + | Frame_Shift_Del | DEL | TAAGGG | T | novel | 4204 | AMPLICON | Illumina MiSeq | p.I427fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4211 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4278 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207497 | 102207497 | + | Frame_Shift_Del | DEL | AG | A | novel | 4304 | AMPLICON | Illumina MiSeq | p.Q529fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4310 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201932 | 102201932 | + | Frame_Shift_Del | DEL | GGAAGA | G | novel | 4364 | AMPLICON | Illumina MiSeq | p.E429fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4378 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201962 | 102201962 | + | Missense_Mutation | SNP | A | C | novel | 4419 | AMPLICON | Illumina MiSeq | p.E438D | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201928 | 102201928 | + | Frame_Shift_Del | DEL | TAAGGG | T | novel | 4435 | AMPLICON | Illumina MiSeq | p.I427fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4453 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206699 | 102206699 | + | Frame_Shift_Del | DEL | GATTT | G | novel | 4488 | AMPLICON | Illumina MiSeq | p.D443fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207508 | 102207508 | + | Nonsense_Mutation | SNP | T | TA | novel | 4488 | AMPLICON | Illumina MiSeq | p.Y533_I534delinsX | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207656 | 102207656 | + | Frame_Shift_Del | DEL | AC | A | novel | 4488 | AMPLICON | Illumina MiSeq | p.Q547fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206907 | 102206907 | + | Frame_Shift_Del | DEL | ACT | A | novel | 4522 | AMPLICON | Illumina MiSeq | p.N512fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4532 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4554 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4566 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207723 | 102207723 | + | In_Frame_Del | DEL | TTTA | T | novel | 4595 | AMPLICON | Illumina MiSeq | p.569_570del | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201945 | 102201945 | + | Frame_Shift_Del | DEL | GAA | G | novel | 4602 | AMPLICON | Illumina MiSeq | p.E433fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207675 | 102207675 | + | Frame_Shift_Ins | INS | G | GA | novel | 4625 | AMPLICON | Illumina MiSeq | p.E553fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207503 | 102207503 | + | Frame_Shift_Del | DEL | TA | T | novel | 4651 | AMPLICON | Illumina MiSeq | p.I531fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207509 | 102207509 | + | Missense_Mutation | SNP | A | G | novel | 4651 | AMPLICON | Illumina MiSeq | p.Y533C | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4654 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201880 | 102201880 | + | Frame_Shift_Ins | INS | G | GACTA | novel | 4696 | AMPLICON | Illumina MiSeq | p.R411fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206790 | 102206790 | + | Frame_Shift_Ins | INS | T | TTA | novel | 4752 | AMPLICON | Illumina MiSeq | p.I473fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207648 | 102207648 | + | Frame_Shift_Ins | INS | G | GT | novel | 4752 | AMPLICON | Illumina MiSeq | p.V544fs | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102206695 | 102206695 | + | Splice_Site | SNP | A | T | novel | 4769 | AMPLICON | Illumina MiSeq | NA | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102207759 | 102207759 | + | Missense_Mutation | SNP | T | A | novel | 4769 | AMPLICON | Illumina MiSeq | p.C581S | NFkB |
| BIRC3 | 330 | Southampton_Cancer_Genomics | hg19 | 11 | 102201850 | 102201850 | + | Missense_Mutation | SNP | G | A | rs17881197 | 4774 | AMPLICON | Illumina MiSeq | p.R401K | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481477 | 140481477 | + | Missense_Mutation | SNP | C | T | novel | 4047 | AMPLICON | Illumina MiSeq | p.R444Q | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4078 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453154 | 140453154 | + | Missense_Mutation | SNP | T | C | novel | 4112 | AMPLICON | Illumina MiSeq | p.D594G | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481412 | 140481412 | + | Missense_Mutation | SNP | C | T | novel | 4169 | AMPLICON | Illumina MiSeq | p.G466R | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4345 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4345 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4365 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4365 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4389 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453154 | 140453154 | + | Missense_Mutation | SNP | T | C | novel | 4419 | AMPLICON | Illumina MiSeq | p.D594G | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453118 | 140453118 | + | Missense_Mutation | SNP | C | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.G606E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4559 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4559 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4570 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4583 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140476718 | 140476718 | + | Missense_Mutation | SNP | C | T | novel | 4661 | AMPLICON | Illumina MiSeq | p.G563D | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4674 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4674 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4736 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4736 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | T | novel | 4086 | AMPLICON | Illumina MiSeq | p.G469E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4150 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4179 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4211 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140477854 | 140477854 | + | Missense_Mutation | SNP | A | G | novel | 4314 | AMPLICON | Illumina MiSeq | p.L485S | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | G | novel | 4330 | AMPLICON | Illumina MiSeq | p.G469A | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453193 | 140453193 | + | Missense_Mutation | SNP | T | A | novel | 4345 | AMPLICON | Illumina MiSeq | p.N581I | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453145 | 140453145 | + | Missense_Mutation | SNP | A | T | novel | 4365 | AMPLICON | Illumina MiSeq | p.L597Q | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453154 | 140453154 | + | Missense_Mutation | SNP | T | C | novel | 4432 | AMPLICON | Illumina MiSeq | p.D594G | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140624403 | 140624403 | + | In_Frame_Del | DEL | GCGGCGC | G | novel | 4457 | AMPLICON | Illumina MiSeq | p.32_34del | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453132 | 140453132 | + | Missense_Mutation | SNP | T | A | novel | 4458 | AMPLICON | Illumina MiSeq | p.K601N | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481412 | 140481412 | + | Missense_Mutation | SNP | C | T | novel | 4469 | AMPLICON | Illumina MiSeq | p.G466R | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453134 | 140453134 | + | Missense_Mutation | SNP | T | C | novel | 4572 | AMPLICON | Illumina MiSeq | p.K601E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481402 | 140481402 | + | Missense_Mutation | SNP | C | T | novel | 4613 | AMPLICON | Illumina MiSeq | p.G469E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453132 | 140453132 | + | Missense_Mutation | SNP | T | A | novel | 4727 | AMPLICON | Illumina MiSeq | p.K601N | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4736 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4741 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140481411 | 140481411 | + | Missense_Mutation | SNP | C | T | novel | 4742 | AMPLICON | Illumina MiSeq | p.G466E | NFkB |
| BRAF | 673 | Southampton_Cancer_Genomics | hg19 | 7 | 140453136 | 140453136 | + | Missense_Mutation | SNP | A | T | novel | 4745 | AMPLICON | Illumina MiSeq | p.V600E | NFkB |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93536112 | 93536112 | + | Missense_Mutation | SNP | C | T | novel | 4047 | AMPLICON | Illumina MiSeq | p.A1160V | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93489418 | 93489418 | + | Frame_Shift_Ins | INS | C | CA | novel | 4075 | AMPLICON | Illumina MiSeq | p.S450fs | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93527671 | 93527671 | + | Missense_Mutation | SNP | C | T | novel | 4080 | AMPLICON | Illumina MiSeq | p.R1060W | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93558013 | 93558013 | + | Frame_Shift_Del | DEL | CA | C | novel | 4080 | AMPLICON | Illumina MiSeq | p.Q1594fs | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93521463 | 93521463 | + | Splice_Site | SNP | G | A | novel | 4271 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93527559 | 93527559 | + | Splice_Site | SNP | G | A | novel | 4302 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93527587 | 93527587 | + | Nonsense_Mutation | SNP | G | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.E1032X | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93515095 | 93515095 | + | Missense_Mutation | SNP | C | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.H764Y | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93527635 | 93527635 | + | Missense_Mutation | SNP | C | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.P1048S | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93541759 | 93541759 | + | Nonsense_Mutation | SNP | C | T | novel | 4699 | AMPLICON | Illumina MiSeq | p.Q1306X | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93498743 | 93498743 | + | Splice_Site | SNP | G | A | novel | 4047 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93496631 | 93496631 | + | Missense_Mutation | SNP | C | T | novel | 4111 | AMPLICON | Illumina MiSeq | p.T516I | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93515514 | 93515514 | + | Missense_Mutation | SNP | G | A | novel | 4302 | AMPLICON | Illumina MiSeq | p.G791E | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93498743 | 93498743 | + | Splice_Site | SNP | G | A | novel | 4335 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93486192 | 93486192 | + | Frame_Shift_Del | DEL | TAC | T | novel | 4355 | AMPLICON | Illumina MiSeq | p.Y316fs | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93487746 | 93487746 | + | Splice_Site | SNP | G | C | novel | 4364 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93498738 | 93498738 | + | Missense_Mutation | SNP | A | G | novel | 4432 | AMPLICON | Illumina MiSeq | p.D602G | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93521507 | 93521507 | + | Missense_Mutation | SNP | T | G | novel | 4468 | AMPLICON | Illumina MiSeq | p.I874S | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93547982 | 93547982 | + | Splice_Site | SNP | G | T | novel | 4701 | AMPLICON | Illumina MiSeq | NA | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93480801 | 93480801 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4705 | AMPLICON | Illumina MiSeq | p.A166fs | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93522464 | 93522464 | + | Missense_Mutation | SNP | A | G | novel | 4713 | AMPLICON | Illumina MiSeq | p.M943V | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93543783 | 93543783 | + | Frame_Shift_Ins | INS | C | CA | novel | 4742 | AMPLICON | Illumina MiSeq | p.N1350fs | Chromatin Modification |
| CHD2 | 1106 | Southampton_Cancer_Genomics | hg19 | 15 | 93470478 | 93470478 | + | Nonsense_Mutation | SNP | G | A | novel | 4753 | AMPLICON | Illumina MiSeq | p.W100X | Chromatin Modification |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126714833 | 126714833 | + | Frame_Shift_Del | DEL | GC | G | novel | 4486 | AMPLICON | Illumina MiSeq | p.A499fs | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715118 | 126715118 | + | Missense_Mutation | SNP | C | T | novel | 4062 | AMPLICON | Illumina MiSeq | p.R404H | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126716232 | 126716232 | + | Missense_Mutation | SNP | G | A | novel | 4086 | AMPLICON | Illumina MiSeq | p.R33W | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715878 | 126715878 | + | Missense_Mutation | SNP | T | C | novel | 4092 | AMPLICON | Illumina MiSeq | p.M151V | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715118 | 126715118 | + | Missense_Mutation | SNP | C | T | novel | 4140 | AMPLICON | Illumina MiSeq | p.R404H | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126716318 | 126716318 | + | Missense_Mutation | SNP | G | A | novel | 4281 | AMPLICON | Illumina MiSeq | p.P4L | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715118 | 126715118 | + | Missense_Mutation | SNP | C | T | novel | 4334 | AMPLICON | Illumina MiSeq | p.R404H | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715640 | 126715640 | + | Missense_Mutation | SNP | G | A | novel | 4357 | AMPLICON | Illumina MiSeq | p.T230M | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715803 | 126715803 | + | Frame_Shift_Del | DEL | CCT | C | novel | 4394 | AMPLICON | Illumina MiSeq | p.Q175fs | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715803 | 126715803 | + | Frame_Shift_Del | DEL | CCT | C | novel | 4470 | AMPLICON | Illumina MiSeq | p.Q175fs | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715437 | 126715437 | + | Missense_Mutation | SNP | G | A | novel | 4490 | AMPLICON | Illumina MiSeq | p.R298W | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715803 | 126715803 | + | Frame_Shift_Del | DEL | CCT | C | novel | 4577 | AMPLICON | Illumina MiSeq | p.Q175fs | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715118 | 126715118 | + | Missense_Mutation | SNP | C | T | novel | 4611 | AMPLICON | Illumina MiSeq | p.R404H | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126716232 | 126716232 | + | Missense_Mutation | SNP | G | A | novel | 4613 | AMPLICON | Illumina MiSeq | p.R33W | Wnt |
| CTBP2 | 1488 | Southampton_Cancer_Genomics | hg19 | 10 | 126715118 | 126715118 | + | Missense_Mutation | SNP | C | T | novel | 4725 | AMPLICON | Illumina MiSeq | p.R404H | Wnt |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204715 | 41204715 | + | Missense_Mutation | SNP | C | G | novel | 4036 | AMPLICON | Illumina MiSeq | p.S410C | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41206604 | 41206604 | + | Frame_Shift_Del | DEL | ACAAAGTAG | A | novel | 4136 | AMPLICON | Illumina MiSeq | p.Q604fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205589 | 41205589 | + | Missense_Mutation | SNP | C | T | novel | 4367 | AMPLICON | Illumina MiSeq | p.R475C | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41206128 | 41206128 | + | Missense_Mutation | SNP | C | A | novel | 4384 | AMPLICON | Illumina MiSeq | p.F544L | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204531 | 41204531 | + | Frame_Shift_Del | DEL | TCCGC | T | novel | 4566 | AMPLICON | Illumina MiSeq | p.V375fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204715 | 41204715 | + | Missense_Mutation | SNP | C | T | novel | 4595 | AMPLICON | Illumina MiSeq | p.S410F | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204495 | 41204495 | + | In_Frame_Del | DEL | GAAT | G | novel | 4648 | AMPLICON | Illumina MiSeq | p.363_364del | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204436 | 41204436 | + | Frame_Shift_Del | DEL | CT | C | novel | 4136 | AMPLICON | Illumina MiSeq | p.L344fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41203585 | 41203585 | + | In_Frame_Del | DEL | TTAG | T | novel | 4172 | AMPLICON | Illumina MiSeq | p.320_321del | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41203002 | 41203002 | + | Frame_Shift_Ins | INS | C | CTGCA | novel | 4260 | AMPLICON | Illumina MiSeq | p.T231fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41202530 | 41202530 | + | Frame_Shift_Ins | INS | G | GC | novel | 4299 | AMPLICON | Illumina MiSeq | p.R202fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41203491 | 41203491 | + | Splice_Site | SNP | G | A | novel | 4327 | AMPLICON | Illumina MiSeq | NA | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205595 | 41205595 | + | Missense_Mutation | SNP | C | G | novel | 4330 | AMPLICON | Illumina MiSeq | p.Q477E | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205493 | 41205493 | + | Missense_Mutation | SNP | C | G | novel | 4472 | AMPLICON | Illumina MiSeq | p.L443V | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41203282 | 41203282 | + | Splice_Site | SNP | G | C | novel | 4524 | AMPLICON | Illumina MiSeq | NA | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41202590 | 41202590 | + | Missense_Mutation | SNP | C | T | novel | 4582 | AMPLICON | Illumina MiSeq | p.A222V | NFkB/RNA processing |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205627 | 41205627 | + | Missense_Mutation | SNP | C | G | novel | 4595 | AMPLICON | Illumina MiSeq | p.F487L | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41203512 | 41203512 | + | Frame_Shift_Del | DEL | TC | T | novel | 4604 | AMPLICON | Illumina MiSeq | p.R296fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205830 | 41205830 | + | Nonsense_Mutation | SNP | G | T | novel | 4629 | AMPLICON | Illumina MiSeq | p.E524X | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205627 | 41205627 | + | Missense_Mutation | SNP | C | G | novel | 4658 | AMPLICON | Illumina MiSeq | p.F487L | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41204735 | 41204735 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4664 | AMPLICON | Illumina MiSeq | p.Q417fs | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41205613 | 41205613 | + | Missense_Mutation | SNP | G | A | novel | 4737 | AMPLICON | Illumina MiSeq | p.A483T | NFkB/RNA processing |
| DDX3X | 1654 | Southampton_Cancer_Genomics | hg19 | X | 41206192 | 41206192 | + | Nonsense_Mutation | SNP | G | T | novel | 4769 | AMPLICON | Illumina MiSeq | p.E566X | NFkB/RNA processing |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573769 | 64573769 | + | Missense_Mutation | SNP | G | T | novel | 4170 | AMPLICON | Illumina MiSeq | p.S210Y | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4188 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4265 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573332 | 64573332 | + | Missense_Mutation | SNP | C | T | novel | 4352 | AMPLICON | Illumina MiSeq | p.E356K | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573619 | 64573619 | + | Missense_Mutation | SNP | G | A | novel | 4486 | AMPLICON | Illumina MiSeq | p.P260L | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573332 | 64573332 | + | Missense_Mutation | SNP | C | T | novel | 4499 | AMPLICON | Illumina MiSeq | p.E356K | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573164 | 64573164 | + | Missense_Mutation | SNP | C | G | novel | 4023 | AMPLICON | Illumina MiSeq | p.E412Q | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573332 | 64573332 | + | Missense_Mutation | SNP | C | T | novel | 4111 | AMPLICON | Illumina MiSeq | p.E356K | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573332 | 64573332 | + | Missense_Mutation | SNP | C | T | novel | 4168 | AMPLICON | Illumina MiSeq | p.E356K | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4304 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573121 | 64573121 | + | Missense_Mutation | SNP | C | T | novel | 4379 | AMPLICON | Illumina MiSeq | p.R426Q | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573332 | 64573332 | + | Missense_Mutation | SNP | C | T | novel | 4491 | AMPLICON | Illumina MiSeq | p.E356K | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4614 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4742 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573248 | 64573248 | + | Missense_Mutation | SNP | G | T | novel | 4763 | AMPLICON | Illumina MiSeq | p.H384N | NFkB |
| EGR2 | 1959 | Southampton_Cancer_Genomics | hg19 | 10 | 64573167 | 64573167 | + | Missense_Mutation | SNP | C | A | novel | 4785 | AMPLICON | Illumina MiSeq | p.D411Y | NFkB |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153247246 | 153247246 | + | Missense_Mutation | SNP | T | C | novel | 4235 | AMPLICON | Illumina MiSeq | p.Y519C | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153250882 | 153250882 | + | Missense_Mutation | SNP | C | T | novel | 4263 | AMPLICON | Illumina MiSeq | p.R393Q | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249400 | 153249400 | + | Missense_Mutation | SNP | G | C | novel | 4419 | AMPLICON | Illumina MiSeq | p.H460D | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153250882 | 153250882 | + | Missense_Mutation | SNP | C | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.R393Q | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153250904 | 153250904 | + | Missense_Mutation | SNP | A | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.C386S | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153247288 | 153247288 | + | Missense_Mutation | SNP | C | T | novel | 4432 | AMPLICON | Illumina MiSeq | p.R505H | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249387 | 153249387 | + | Missense_Mutation | SNP | A | G | novel | 4486 | AMPLICON | Illumina MiSeq | p.V464A | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249450 | 153249450 | + | Missense_Mutation | SNP | A | G | novel | 4486 | AMPLICON | Illumina MiSeq | p.L443P | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249384 | 153249384 | + | Missense_Mutation | SNP | C | A | novel | 4488 | AMPLICON | Illumina MiSeq | p.R465L | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153247366 | 153247366 | + | Missense_Mutation | SNP | C | T | novel | 4493 | AMPLICON | Illumina MiSeq | p.R479Q | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153247246 | 153247246 | + | Missense_Mutation | SNP | T | C | novel | 4584 | AMPLICON | Illumina MiSeq | p.Y519C | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249510 | 153249510 | + | Missense_Mutation | SNP | C | A | novel | 4216 | AMPLICON | Illumina MiSeq | p.G423V | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249510 | 153249510 | + | Missense_Mutation | SNP | C | A | novel | 4267 | AMPLICON | Illumina MiSeq | p.G423V | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249511 | 153249511 | + | Missense_Mutation | SNP | C | T | novel | 4267 | AMPLICON | Illumina MiSeq | p.G423R | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153247366 | 153247366 | + | Missense_Mutation | SNP | C | T | novel | 4283 | AMPLICON | Illumina MiSeq | p.R479Q | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249384 | 153249384 | + | Missense_Mutation | SNP | C | T | novel | 4357 | AMPLICON | Illumina MiSeq | p.R465H | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249523 | 153249523 | + | Missense_Mutation | SNP | C | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.G419R | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249503 | 153249503 | + | Missense_Mutation | SNP | C | G | novel | 4704 | AMPLICON | Illumina MiSeq | p.W425C | NOTCH/MYC |
| FBXW7 | 55294 | Southampton_Cancer_Genomics | hg19 | 4 | 153249503 | 153249503 | + | Missense_Mutation | SNP | C | G | novel | 4723 | AMPLICON | Illumina MiSeq | p.W425C | NOTCH/MYC |
| HIST1H1E | 3008 | Southampton_Cancer_Genomics | hg19 | 6 | 26156672 | 26156672 | + | Frame_Shift_Del | DEL | TC | T | novel | 4648 | AMPLICON | Illumina MiSeq | p.P19fs | Chromatin Modification |
| HIST1H1E | 3008 | Southampton_Cancer_Genomics | hg19 | 6 | 26156648 | 26156648 | + | Frame_Shift_Del | DEL | TG | T | novel | 4655 | AMPLICON | Illumina MiSeq | p.A11fs | Chromatin Modification |
| HIST1H1E | 3008 | Southampton_Cancer_Genomics | hg19 | 6 | 26156811 | 26156811 | + | Missense_Mutation | SNP | G | C | novel | 4287 | AMPLICON | Illumina MiSeq | p.A65P | Chromatin Modification |
| HIST1H1E | 3008 | Southampton_Cancer_Genomics | hg19 | 6 | 26156758 | 26156758 | + | Missense_Mutation | SNP | C | T | novel | 4329 | AMPLICON | Illumina MiSeq | p.A47V | Chromatin Modification |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398255 | 25398255 | + | Missense_Mutation | SNP | G | T | novel | 4146 | AMPLICON | Illumina MiSeq | p.Q22K | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25380276 | 25380276 | + | Missense_Mutation | SNP | T | C | novel | 4161 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4197 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4352 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25380275 | 25380275 | + | Missense_Mutation | SNP | T | G | rs17851045 | 4356 | AMPLICON | Illumina MiSeq | p.Q61H | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4365 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398262 | 25398262 | + | Missense_Mutation | SNP | C | G | novel | 4377 | AMPLICON | Illumina MiSeq | p.L19F | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4378 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4378 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4454 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398262 | 25398262 | + | Missense_Mutation | SNP | C | A | novel | 4461 | AMPLICON | Illumina MiSeq | p.L19F | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4527 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | A | novel | 4548 | AMPLICON | Illumina MiSeq | p.G12V | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398262 | 25398262 | + | Missense_Mutation | SNP | C | A | novel | 4565 | AMPLICON | Illumina MiSeq | p.L19F | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4587 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398262 | 25398262 | + | Missense_Mutation | SNP | C | G | novel | 4641 | AMPLICON | Illumina MiSeq | p.L19F | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4655 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4655 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4701 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4701 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398255 | 25398255 | + | Missense_Mutation | SNP | G | T | novel | 4736 | AMPLICON | Illumina MiSeq | p.Q22K | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398248 | 25398248 | + | Missense_Mutation | SNP | A | T | novel | 4742 | AMPLICON | Illumina MiSeq | p.I24N | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4745 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4121 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4190 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398284 | 25398284 | + | Missense_Mutation | SNP | C | T | novel | 4257 | AMPLICON | Illumina MiSeq | p.G12D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25380276 | 25380276 | + | Missense_Mutation | SNP | T | C | novel | 4283 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25380280 | 25380280 | + | Missense_Mutation | SNP | C | T | novel | 4432 | AMPLICON | Illumina MiSeq | p.G60S | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4489 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398255 | 25398255 | + | Missense_Mutation | SNP | G | T | novel | 4546 | AMPLICON | Illumina MiSeq | p.Q22K | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4565 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25380282 | 25380282 | + | Missense_Mutation | SNP | G | C | novel | 4665 | AMPLICON | Illumina MiSeq | p.A59G | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398285 | 25398285 | + | Missense_Mutation | SNP | C | A | novel | 4701 | AMPLICON | Illumina MiSeq | p.G12C | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4719 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| KRAS | 3845 | Southampton_Cancer_Genomics | hg19 | 12 | 25398281 | 25398281 | + | Missense_Mutation | SNP | C | T | novel | 4760 | AMPLICON | Illumina MiSeq | p.G13D | NFkB |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339245 | 70339245 | + | Missense_Mutation | SNP | T | G | novel | 4061 | AMPLICON | Illumina MiSeq | p.V41G | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339251 | 70339251 | + | Missense_Mutation | SNP | A | C | novel | 4061 | AMPLICON | Illumina MiSeq | p.Q43P | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339253 | 70339253 | + | Missense_Mutation | SNP | G | A | novel | 4221 | AMPLICON | Illumina MiSeq | p.G44S | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339323 | 70339323 | + | Missense_Mutation | SNP | C | T | novel | 4384 | AMPLICON | Illumina MiSeq | p.A67V | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339254 | 70339254 | + | Missense_Mutation | SNP | G | A | novel | 4389 | AMPLICON | Illumina MiSeq | p.G44D | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339254 | 70339254 | + | Missense_Mutation | SNP | G | C | novel | 4693 | AMPLICON | Illumina MiSeq | p.G44A | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339253 | 70339253 | + | Missense_Mutation | SNP | G | A | novel | 4092 | AMPLICON | Illumina MiSeq | p.G44S | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339254 | 70339254 | + | Missense_Mutation | SNP | G | A | novel | 4310 | AMPLICON | Illumina MiSeq | p.G44D | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339230 | 70339230 | + | Missense_Mutation | SNP | T | C | novel | 4378 | AMPLICON | Illumina MiSeq | p.L36P | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339233 | 70339233 | + | Missense_Mutation | SNP | C | G | novel | 4425 | AMPLICON | Illumina MiSeq | p.T37R | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339254 | 70339254 | + | Missense_Mutation | SNP | G | T | novel | 4683 | AMPLICON | Illumina MiSeq | p.G44V | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339254 | 70339254 | + | Missense_Mutation | SNP | G | C | novel | 4705 | AMPLICON | Illumina MiSeq | p.G44A | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339230 | 70339230 | + | Missense_Mutation | SNP | T | G | novel | 4715 | AMPLICON | Illumina MiSeq | p.L36R | Wnt |
| MED12 | 9968 | Southampton_Cancer_Genomics | hg19 | X | 70339253 | 70339253 | + | Missense_Mutation | SNP | G | A | novel | 4758 | AMPLICON | Illumina MiSeq | p.G44S | Wnt |
| MGA | 23269 | Southampton_Cancer_Genomics | hg19 | 15 | 42041074 | 42041074 | + | Nonsense_Mutation | SNP | C | T | novel | 4346 | AMPLICON | Illumina MiSeq | p.R1818X | MYC |
| MGA | 23269 | Southampton_Cancer_Genomics | hg19 | 15 | 42040924 | 42040924 | + | Nonsense_Mutation | SNP | C | T | novel | 4677 | AMPLICON | Illumina MiSeq | p.Q1768X | MYC |
| MGA | 23269 | Southampton_Cancer_Genomics | hg19 | 15 | 42040838 | 42040838 | + | Missense_Mutation | SNP | A | C | novel | 4760 | AMPLICON | Illumina MiSeq | p.N1739T | MYC |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182689 | 38182689 | + | Frame_Shift_Del | DEL | TG | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.L289fs | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182032 | 38182032 | + | Missense_Mutation | SNP | C | G | novel | 4571 | AMPLICON | Illumina MiSeq | p.S219C | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4066 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4124 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182337 | 38182337 | + | Nonsense_Mutation | SNP | C | T | novel | 4158 | AMPLICON | Illumina MiSeq | p.Q153X,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182032 | 38182032 | + | Missense_Mutation | SNP | C | G | novel | 4168 | AMPLICON | Illumina MiSeq | p.S174C,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182025 | 38182025 | + | Missense_Mutation | SNP | G | T | novel | 4169 | AMPLICON | Illumina MiSeq | p.V172F,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182337 | 38182337 | + | Nonsense_Mutation | SNP | C | T | novel | 4178 | AMPLICON | Illumina MiSeq | p.Q153X,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182025 | 38182025 | + | Missense_Mutation | SNP | G | T | novel | 4306 | AMPLICON | Illumina MiSeq | p.V172F,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182025 | 38182025 | + | Missense_Mutation | SNP | G | T | novel | 4312 | AMPLICON | Illumina MiSeq | p.V172F,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4353 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4370 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182032 | 38182032 | + | Missense_Mutation | SNP | C | G | novel | 4384 | AMPLICON | Illumina MiSeq | p.S174C,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182025 | 38182025 | + | Missense_Mutation | SNP | G | T | novel | 4469 | AMPLICON | Illumina MiSeq | p.V172F,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4574 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4598 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182025 | 38182025 | + | Missense_Mutation | SNP | G | T | novel | 4607 | AMPLICON | Illumina MiSeq | p.V172F,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182032 | 38182032 | + | Missense_Mutation | SNP | C | G | novel | 4612 | AMPLICON | Illumina MiSeq | p.S174C,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4637 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| MYD88 | 17874 | Southampton_Cancer_Genomics | hg19 | 9 | 38182641 | 38182641 | + | Nonsense_Mutation | SNP | T | C | novel | 4661 | AMPLICON | Illumina MiSeq | p.X160R,MYD88 | NFkB |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47049647 | 47049647 | + | Missense_Mutation | SNP | G | A | novel | 4047 | AMPLICON | Illumina MiSeq | p.E2564K | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47037428 | 47037428 | + | Missense_Mutation | SNP | G | A | novel | 4074 | AMPLICON | Illumina MiSeq | p.V680M | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47049662 | 47049662 | + | Missense_Mutation | SNP | G | A | novel | 4263 | AMPLICON | Illumina MiSeq | p.V2569M | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47040930 | 47040930 | + | Frame_Shift_Del | DEL | CT | C | novel | 4265 | AMPLICON | Illumina MiSeq | p.C1224fs | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47049677 | 47049677 | + | Splice_Site | SNP | G | T | novel | 4384 | AMPLICON | Illumina MiSeq | NA | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47042491 | 47042491 | + | Nonsense_Mutation | SNP | C | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.Q1436X | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47046000 | 47046000 | + | Missense_Mutation | SNP | G | A | novel | 4486 | AMPLICON | Illumina MiSeq | p.E2072K | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47046052 | 47046052 | + | Missense_Mutation | SNP | C | A | novel | 4704 | AMPLICON | Illumina MiSeq | p.T2089N | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47030814 | 47030814 | + | Frame_Shift_Del | DEL | TG | T | novel | 4752 | AMPLICON | Illumina MiSeq | p.L139fs | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47041439 | 47041439 | + | Frame_Shift_Del | DEL | TG | T | novel | 4772 | AMPLICON | Illumina MiSeq | p.W1284fs | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47037253 | 47037253 | + | Missense_Mutation | SNP | G | A | novel | 4036 | AMPLICON | Illumina MiSeq | p.G650R | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47030856 | 47030856 | + | Missense_Mutation | SNP | G | A | novel | 4048 | AMPLICON | Illumina MiSeq | p.R153Q | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47050775 | 47050775 | + | Missense_Mutation | SNP | G | C | novel | 4086 | AMPLICON | Illumina MiSeq | p.G2744R | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47050775 | 47050775 | + | Missense_Mutation | SNP | G | C | novel | 4272 | AMPLICON | Illumina MiSeq | p.G2744R | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47043574 | 47043574 | + | In_Frame_Del | DEL | TGCAGCTGCA | T | novel | 4347 | AMPLICON | Illumina MiSeq | p.1650_1652del | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47043648 | 47043648 | + | Missense_Mutation | SNP | G | A | novel | 4473 | AMPLICON | Illumina MiSeq | p.R1674H | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47043602 | 47043602 | + | Missense_Mutation | SNP | G | A | novel | 4494 | AMPLICON | Illumina MiSeq | p.E1659K | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47050775 | 47050775 | + | Missense_Mutation | SNP | G | C | novel | 4613 | AMPLICON | Illumina MiSeq | p.G2744R | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47042584 | 47042584 | + | Missense_Mutation | SNP | C | T | novel | 4638 | AMPLICON | Illumina MiSeq | p.R1467W | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47049386 | 47049386 | + | Missense_Mutation | SNP | C | T | novel | 4726 | AMPLICON | Illumina MiSeq | p.R2530W | Autophagy |
| NBEAL2 | 23218 | Southampton_Cancer_Genomics | hg19 | 3 | 47043648 | 47043648 | + | Missense_Mutation | SNP | G | A | novel | 4737 | AMPLICON | Illumina MiSeq | p.R1674H | Autophagy |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4012 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4024 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4039 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232740 | 44232740 | + | Frame_Shift_Del | DEL | TAA | T | novel | 4046 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4121 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4174 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4180 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4205 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4216 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4271 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4285 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4310 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4312 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4316 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4367 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4453 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4474 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44233103 | 44233103 | + | Frame_Shift_Ins | INS | C | CG | novel | 4486 | AMPLICON | Illumina MiSeq | p.R133fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4489 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4532 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4588 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4617 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4653 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4684 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4694 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4719 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4733 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4752 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4046 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4127 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4259 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232739 | 44232739 | + | Nonsense_Mutation | SNP | G | GT | novel | 4265 | AMPLICON | Illumina MiSeq | p.Y254_I255delinsX | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4288 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232741 | 44232741 | + | Frame_Shift_Del | DEL | AAGTG | A | novel | 4289 | AMPLICON | Illumina MiSeq | p.L252fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4323 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4399 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4426 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4485 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4501 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4739 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4772 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232738 | 44232738 | + | Frame_Shift_Del | DEL | TGTAA | T | novel | 4785 | AMPLICON | Illumina MiSeq | p.T253fs | NFkB |
| NFKBIE | 4794 | Southampton_Cancer_Genomics | hg19 | 6 | 44232948 | 44232948 | + | Nonsense_Mutation | SNP | G | A | novel | 4785 | AMPLICON | Illumina MiSeq | p.Q185X | NFkB |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4023 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4146 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4158 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4211 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4216 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390923 | 139390923 | + | Nonsense_Mutation | SNP | G | T | novel | 4254 | AMPLICON | Illumina MiSeq | p.S2423X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4263 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4267 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4307 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4333 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4419 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391983 | 139391983 | + | Missense_Mutation | SNP | G | A | novel | 4437 | AMPLICON | Illumina MiSeq | p.R2070W | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4489 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4497 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4501 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4532 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4544 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4615 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4701 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4731 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390649 | 139390649 | + | Frame_Shift_Del | DEL | AGG | A | novel | 4006 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4014 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4089 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4097 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4099 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4101 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4130 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4145 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4157 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390732 | 139390732 | + | Nonsense_Mutation | SNP | G | A | novel | 4221 | AMPLICON | Illumina MiSeq | p.Q2487X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390791 | 139390791 | + | Nonsense_Mutation | SNP | G | T | novel | 4268 | AMPLICON | Illumina MiSeq | p.S2467X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4269 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4304 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4308 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4314 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4366 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4374 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4396 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391303 | 139391303 | + | Frame_Shift_Del | DEL | GA | G | novel | 4409 | AMPLICON | Illumina MiSeq | p.F2296fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4435 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4436 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391065 | 139391065 | + | Nonsense_Mutation | SNP | G | A | novel | 4436 | AMPLICON | Illumina MiSeq | p.Q2376X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391875 | 139391875 | + | Missense_Mutation | SNP | G | A | novel | 4437 | AMPLICON | Illumina MiSeq | p.H2106Y | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4447 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4453 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4454 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390822 | 139390822 | + | Missense_Mutation | SNP | G | C | rs61755043 | 4474 | AMPLICON | Illumina MiSeq | p.L2457V | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4475 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4486 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4488 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390813 | 139390813 | + | Nonsense_Mutation | SNP | C | A | novel | 4501 | AMPLICON | Illumina MiSeq | p.E2460X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4536 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390923 | 139390923 | + | Nonsense_Mutation | SNP | G | C | novel | 4544 | AMPLICON | Illumina MiSeq | p.S2423X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391506 | 139391506 | + | Missense_Mutation | SNP | C | T | novel | 4544 | AMPLICON | Illumina MiSeq | p.V2229M | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391869 | 139391869 | + | Missense_Mutation | SNP | C | T | novel | 4565 | AMPLICON | Illumina MiSeq | p.D2108N | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390945 | 139390945 | + | Nonsense_Mutation | SNP | G | A | novel | 4595 | AMPLICON | Illumina MiSeq | p.Q2416X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4602 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390746 | 139390746 | + | Nonsense_Mutation | SNP | AG | A | novel | 4616 | AMPLICON | Illumina MiSeq | p.L2482X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4621 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390687 | 139390687 | + | Missense_Mutation | SNP | G | C | novel | 4654 | AMPLICON | Illumina MiSeq | p.L2502V | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391506 | 139391506 | + | Missense_Mutation | SNP | C | T | novel | 4656 | AMPLICON | Illumina MiSeq | p.V2229M | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390945 | 139390945 | + | Nonsense_Mutation | SNP | G | A | novel | 4658 | AMPLICON | Illumina MiSeq | p.Q2416X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4672 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4705 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4711 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4726 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390690 | 139390690 | + | Nonsense_Mutation | SNP | G | A | novel | 4737 | AMPLICON | Illumina MiSeq | p.Q2501X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390648 | 139390648 | + | Frame_Shift_Del | DEL | CAG | C | novel | 4752 | AMPLICON | Illumina MiSeq | p.P2514fs | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390746 | 139390746 | + | Nonsense_Mutation | SNP | AG | A | novel | 4762 | AMPLICON | Illumina MiSeq | p.L2482X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390819 | 139390819 | + | Missense_Mutation | SNP | G | T | novel | 4773 | AMPLICON | Illumina MiSeq | p.P2458T | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139390690 | 139390690 | + | Nonsense_Mutation | SNP | G | A | novel | 4784 | AMPLICON | Illumina MiSeq | p.Q2501X | NOTCH |
| NOTCH1 | 4851 | Southampton_Cancer_Genomics | hg19 | 9 | 139391014 | 139391014 | + | Nonsense_Mutation | SNP | G | A | novel | 4785 | AMPLICON | Illumina MiSeq | p.Q2393X | NOTCH |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4121 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4389 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4396 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256530 | 115256530 | + | Missense_Mutation | SNP | G | T | novel | 4527 | AMPLICON | Illumina MiSeq | p.Q61K | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256530 | 115256530 | + | Missense_Mutation | SNP | G | T | novel | 4638 | AMPLICON | Illumina MiSeq | p.Q61K | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4655 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4187 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4493 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4527 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4738 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| NRAS | 4893 | Southampton_Cancer_Genomics | hg19 | 1 | 115256529 | 115256529 | + | Missense_Mutation | SNP | T | C | rs11554290 | 4749 | AMPLICON | Illumina MiSeq | p.Q61R | NFkB |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124510987 | 124510987 | + | Missense_Mutation | SNP | A | G | novel | 4078 | AMPLICON | Illumina MiSeq | p.I78T | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532325 | 124532325 | + | Missense_Mutation | SNP | C | T | novel | 4146 | AMPLICON | Illumina MiSeq | p.G40E | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503594 | 124503594 | + | Nonsense_Mutation | SNP | G | C | novel | 4176 | AMPLICON | Illumina MiSeq | p.S119X | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124481138 | 124481138 | + | Missense_Mutation | SNP | C | T | novel | 4206 | AMPLICON | Illumina MiSeq | p.D420N | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124510965 | 124510965 | + | Missense_Mutation | SNP | C | A | novel | 4221 | AMPLICON | Illumina MiSeq | p.K85N | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124487022 | 124487022 | + | Missense_Mutation | SNP | T | C | novel | 4342 | AMPLICON | Illumina MiSeq | p.E327G | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503433 | 124503433 | + | Missense_Mutation | SNP | C | T | novel | 4354 | AMPLICON | Illumina MiSeq | p.E173K | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532341 | 124532341 | + | Missense_Mutation | SNP | G | T | novel | 4356 | AMPLICON | Illumina MiSeq | p.P35T | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124481154 | 124481154 | + | Missense_Mutation | SNP | T | A | novel | 4486 | AMPLICON | Illumina MiSeq | p.Q414H | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124511090 | 124511090 | + | Missense_Mutation | SNP | A | C | novel | 4499 | AMPLICON | Illumina MiSeq | p.C44G | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493170 | 124493170 | + | Missense_Mutation | SNP | T | C | novel | 4635 | AMPLICON | Illumina MiSeq | p.Y242C | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532332 | 124532332 | + | Missense_Mutation | SNP | T | C | novel | 4694 | AMPLICON | Illumina MiSeq | p.S38G | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503544 | 124503544 | + | Missense_Mutation | SNP | A | C | novel | 4739 | AMPLICON | Illumina MiSeq | p.L136V | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124537219 | 124537219 | + | Missense_Mutation | SNP | C | A | novel | 4761 | AMPLICON | Illumina MiSeq | p.L3F | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124537267 | 124537267 | + | Splice_Site | SNP | C | A | novel | 4761 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124475336 | 124475336 | + | Frame_Shift_Del | DEL | TA | T | novel | 4780 | AMPLICON | Illumina MiSeq | p.Y501fs | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124491972 | 124491972 | + | Missense_Mutation | SNP | C | A | novel | 4033 | AMPLICON | Illumina MiSeq | p.Q301H | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493075 | 124493075 | + | Missense_Mutation | SNP | C | T | novel | 4061 | AMPLICON | Illumina MiSeq | p.G274R | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493080 | 124493080 | + | Missense_Mutation | SNP | C | T | novel | 4074 | AMPLICON | Illumina MiSeq | p.G272D | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503681 | 124503681 | + | Missense_Mutation | SNP | T | C | novel | 4085 | AMPLICON | Illumina MiSeq | p.K90R | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124492004 | 124492004 | + | Missense_Mutation | SNP | C | A | novel | 4089 | AMPLICON | Illumina MiSeq | p.D291Y | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124510987 | 124510987 | + | Missense_Mutation | SNP | A | G | novel | 4092 | AMPLICON | Illumina MiSeq | p.I78T | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124511097 | 124511097 | + | Splice_Site | SNP | T | A | novel | 4110 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493089 | 124493089 | + | Missense_Mutation | SNP | G | A | novel | 4135 | AMPLICON | Illumina MiSeq | p.T269I | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124499043 | 124499043 | + | Missense_Mutation | SNP | C | T | novel | 4140 | AMPLICON | Illumina MiSeq | p.D224N | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124537227 | 124537227 | + | Missense_Mutation | SNP | T | C | novel | 4157 | AMPLICON | Illumina MiSeq | p.M1V | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124492004 | 124492004 | + | Missense_Mutation | SNP | C | T | novel | 4209 | AMPLICON | Illumina MiSeq | p.D291N | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124491972 | 124491972 | + | Missense_Mutation | SNP | C | A | novel | 4223 | AMPLICON | Illumina MiSeq | p.Q301H | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503561 | 124503561 | + | Frame_Shift_Del | DEL | TG | T | novel | 4247 | AMPLICON | Illumina MiSeq | p.H130fs | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124491972 | 124491972 | + | Missense_Mutation | SNP | C | A | novel | 4253 | AMPLICON | Illumina MiSeq | p.Q301H | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532337 | 124532337 | + | Missense_Mutation | SNP | T | C | novel | 4285 | AMPLICON | Illumina MiSeq | p.Y36C | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532359 | 124532359 | + | Missense_Mutation | SNP | C | A | novel | 4314 | AMPLICON | Illumina MiSeq | p.V29L | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124464068 | 124464068 | + | Frame_Shift_Del | DEL | TTA | T | novel | 4347 | AMPLICON | Illumina MiSeq | p.D617fs | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124482912 | 124482912 | + | Missense_Mutation | SNP | G | A | novel | 4355 | AMPLICON | Illumina MiSeq | p.P371L | DDR/CC |

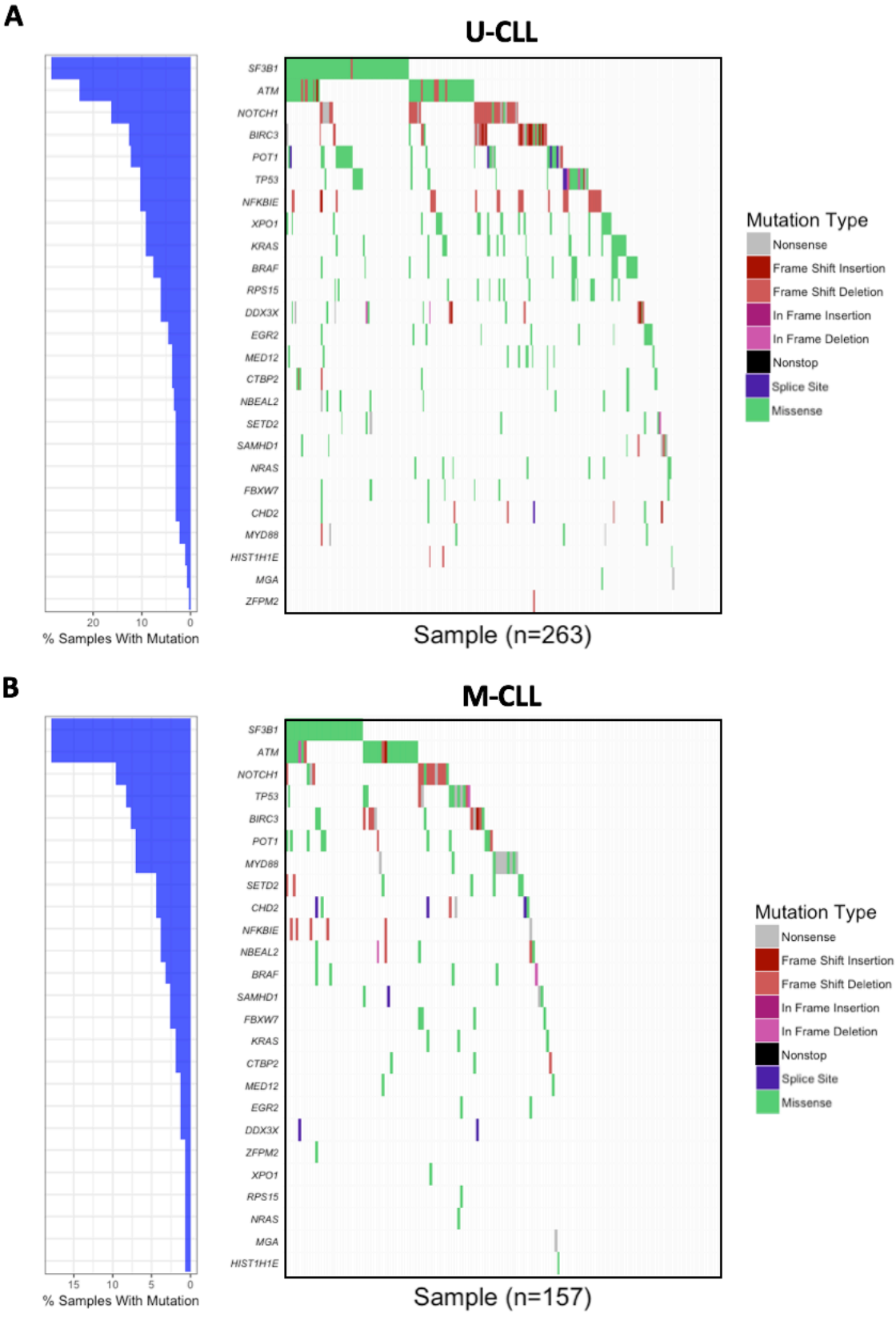
| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532322 | 124532322 | + | Missense_Mutation | SNP | G | A | novel | 4366 | AMPLICON | Illumina MiSeq | p.T41I | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532398 | 124532398 | + | Nonsense_Mutation | SNP | G | A | novel | 4437 | AMPLICON | Illumina MiSeq | p.Q16X | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532318 | 124532318 | + | Splice_Site | INS | A | ACCAGTT | novel | 4454 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124491972 | 124491972 | + | Missense_Mutation | SNP | C | A | novel | 4473 | AMPLICON | Illumina MiSeq | p.Q301H | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493077 | 124493077 | + | Missense_Mutation | SNP | C | T | novel | 4488 | AMPLICON | Illumina MiSeq | p.R273Q | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124493077 | 124493077 | + | Missense_Mutation | SNP | C | T | novel | 4545 | AMPLICON | Illumina MiSeq | p.R273Q | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124511035 | 124511035 | + | Missense_Mutation | SNP | A | C | novel | 4563 | AMPLICON | Illumina MiSeq | p.F62C | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124499058 | 124499058 | + | Missense_Mutation | SNP | C | A | novel | 4628 | AMPLICON | Illumina MiSeq | p.D219Y | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503682 | 124503682 | + | Missense_Mutation | SNP | T | C | novel | 4635 | AMPLICON | Illumina MiSeq | p.K90E | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503682 | 124503682 | + | Missense_Mutation | SNP | T | C | novel | 4683 | AMPLICON | Illumina MiSeq | p.K90E | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124467311 | 124467311 | + | Missense_Mutation | SNP | A | G | novel | 4701 | AMPLICON | Illumina MiSeq | p.L548P | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124510990 | 124510990 | + | Missense_Mutation | SNP | T | A | novel | 4713 | AMPLICON | Illumina MiSeq | p.D77V | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503684 | 124503684 | + | Missense_Mutation | SNP | T | C | novel | 4719 | AMPLICON | Illumina MiSeq | p.Y89C | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124503669 | 124503669 | + | Missense_Mutation | SNP | T | C | novel | 4731 | AMPLICON | Illumina MiSeq | p.Q94R | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532332 | 124532332 | + | Missense_Mutation | SNP | T | G | novel | 4747 | AMPLICON | Illumina MiSeq | p.S38R | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124532332 | 124532332 | + | Missense_Mutation | SNP | T | G | novel | 4748 | AMPLICON | Illumina MiSeq | p.S38R | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124491925 | 124491925 | + | Splice_Site | SNP | C | T | novel | 4755 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124537219 | 124537219 | + | Frame_Shift_Del | DEL | CA | C | novel | 4768 | AMPLICON | Illumina MiSeq | p.L3fs | DDR/CC |
| POT1 | 25913 | Southampton_Cancer_Genomics | hg19 | 7 | 124492004 | 124492004 | + | Missense_Mutation | SNP | C | A | novel | 4773 | AMPLICON | Illumina MiSeq | p.D291Y | DDR/CC |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4145 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4146 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4180 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440429 | 1440429 | + | Missense_Mutation | SNP | A | G | novel | 4180 | AMPLICON | Illumina MiSeq | p.T136A | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4352 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | G | novel | 4745 | AMPLICON | Illumina MiSeq | p.H137D | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | T | novel | 4769 | AMPLICON | Illumina MiSeq | p.H137Y | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440429 | 1440429 | + | Missense_Mutation | SNP | A | G | novel | 4024 | AMPLICON | Illumina MiSeq | p.T136A | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440415 | 1440415 | + | Missense_Mutation | SNP | C | G | novel | 4135 | AMPLICON | Illumina MiSeq | p.P131R | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | A | novel | 4150 | AMPLICON | Illumina MiSeq | p.P131T | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440417 | 1440417 | + | Missense_Mutation | SNP | G | A | novel | 4216 | AMPLICON | Illumina MiSeq | p.G132S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4295 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | T | novel | 4322 | AMPLICON | Illumina MiSeq | p.H137Y | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440429 | 1440429 | + | Missense_Mutation | SNP | A | G | novel | 4323 | AMPLICON | Illumina MiSeq | p.T136A | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | A | novel | 4342 | AMPLICON | Illumina MiSeq | p.P131T | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440429 | 1440429 | + | Missense_Mutation | SNP | A | G | novel | 4356 | AMPLICON | Illumina MiSeq | p.T136A | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440436 | 1440436 | + | Missense_Mutation | SNP | C | T | novel | 4426 | AMPLICON | Illumina MiSeq | p.S138F | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | T | novel | 4461 | AMPLICON | Illumina MiSeq | p.H137Y | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440429 | 1440429 | + | Missense_Mutation | SNP | A | G | novel | 4491 | AMPLICON | Illumina MiSeq | p.T136A | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | T | novel | 4565 | AMPLICON | Illumina MiSeq | p.H137Y | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440423 | 1440423 | + | Missense_Mutation | SNP | G | A | novel | 4643 | AMPLICON | Illumina MiSeq | p.G134R | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440414 | 1440414 | + | Missense_Mutation | SNP | C | T | novel | 4674 | AMPLICON | Illumina MiSeq | p.P131S | RNA processing |
| RPS15 | 6209 | Southampton_Cancer_Genomics | hg19 | 19 | 1440432 | 1440432 | + | Missense_Mutation | SNP | C | T | novel | 4694 | AMPLICON | Illumina MiSeq | p.H137Y | RNA processing |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35540904 | 35540904 | + | Nonsense_Mutation | SNP | T | A | novel | 4075 | AMPLICON | Illumina MiSeq | p.R372X | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563529 | 35563529 | + | Missense_Mutation | SNP | T | C | novel | 4075 | AMPLICON | Illumina MiSeq | p.T138A | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35545382 | 35545382 | + | Missense_Mutation | SNP | A | G | novel | 4175 | AMPLICON | Illumina MiSeq | p.I308T | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35559171 | 35559171 | + | Missense_Mutation | SNP | T | C | novel | 4175 | AMPLICON | Illumina MiSeq | p.H206R | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35540897 | 35540897 | + | Missense_Mutation | SNP | T | C | novel | 4201 | AMPLICON | Illumina MiSeq | p.Y374C | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35545234 | 35545234 | + | Splice_Site | SNP | C | T | novel | 4201 | AMPLICON | Illumina MiSeq | NA | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35559186 | 35559186 | + | Missense_Mutation | SNP | A | T | novel | 4268 | AMPLICON | Illumina MiSeq | p.I201N | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563576 | 35563576 | + | Missense_Mutation | SNP | A | G | novel | 4268 | AMPLICON | Illumina MiSeq | p.I122T | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35532564 | 35532564 | + | Missense_Mutation | SNP | A | C | novel | 4291 | AMPLICON | Illumina MiSeq | p.V500G | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563451 | 35563451 | + | Nonsense_Mutation | SNP | G | A | novel | 4293 | AMPLICON | Illumina MiSeq | p.R164X | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563447 | 35563447 | + | Missense_Mutation | SNP | A | G | novel | 4311 | AMPLICON | Illumina MiSeq | p.F165S | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563494 | 35563494 | + | Frame_Shift_Del | DEL | CT | C | novel | 4311 | AMPLICON | Illumina MiSeq | p.Q149fs | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35533826 | 35533826 | + | Missense_Mutation | SNP | G | A | novel | 4315 | AMPLICON | Illumina MiSeq | p.R451C | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563450 | 35563450 | + | Missense_Mutation | SNP | C | T | novel | 4368 | AMPLICON | Illumina MiSeq | p.R164Q | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563451 | 35563451 | + | Nonsense_Mutation | SNP | G | A | novel | 4480 | AMPLICON | Illumina MiSeq | p.R164X | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35563581 | 35563581 | + | Missense_Mutation | SNP | A | T | novel | 4480 | AMPLICON | Illumina MiSeq | p.D120E | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35545417 | 35545417 | + | Frame_Shift_Del | DEL | GA | G | novel | 4604 | AMPLICON | Illumina MiSeq | p.F296fs | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35559277 | 35559277 | + | Missense_Mutation | SNP | C | G | novel | 4604 | AMPLICON | Illumina MiSeq | p.V171L | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35559186 | 35559186 | + | Missense_Mutation | SNP | A | T | novel | 4613 | AMPLICON | Illumina MiSeq | p.I201N | NFkB |
| SAMHD1 | 25939 | Southampton_Cancer_Genomics | hg19 | 20 | 35532564 | 35532564 | + | Missense_Mutation | SNP | A | C | novel | 4781 | AMPLICON | Illumina MiSeq | p.V500G | NFkB |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47164594 | 47164594 | + | Missense_Mutation | SNP | G | T | novel | 4426 | AMPLICON | Illumina MiSeq | p.S511Y | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47108569 | 47108569 | + | Nonsense_Mutation | SNP | C | A | novel | 4704 | AMPLICON | Illumina MiSeq | p.E2034X | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47108569 | 47108569 | + | Nonsense_Mutation | SNP | C | A | novel | 4723 | AMPLICON | Illumina MiSeq | p.E2034X | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47163543 | 47163543 | + | Nonsense_Mutation | SNP | TG | T | novel | 4752 | AMPLICON | Illumina MiSeq | p.S861X | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47165830 | 47165830 | + | Missense_Mutation | SNP | T | C | novel | 4051 | AMPLICON | Illumina MiSeq | p.D99G | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47061309 | 47061309 | + | Frame_Shift_Del | DEL | CTGTCT | C | novel | 4089 | AMPLICON | Illumina MiSeq | p.K2456fs | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47164793 | 47164793 | + | Missense_Mutation | SNP | G | A | novel | 4111 | AMPLICON | Illumina MiSeq | p.R445C | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47125604 | 47125604 | + | Missense_Mutation | SNP | A | G | novel | 4172 | AMPLICON | Illumina MiSeq | p.M1889T | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47164711 | 47164711 | + | Missense_Mutation | SNP | C | T | novel | 4250 | AMPLICON | Illumina MiSeq | p.R472H | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47165978 | 47165978 | + | Missense_Mutation | SNP | C | T | novel | 4273 | AMPLICON | Illumina MiSeq | p.A50T | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47162897 | 47162897 | + | Missense_Mutation | SNP | T | C | novel | 4306 | AMPLICON | Illumina MiSeq | p.T1077A | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47079245 | 47079245 | + | Missense_Mutation | SNP | T | C | rs76132393 | 4354 | AMPLICON | Illumina MiSeq | p.T2421A | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47098389 | 47098389 | + | Missense_Mutation | SNP | T | C | novel | 4530 | AMPLICON | Illumina MiSeq | p.I2295M | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47165475 | 47165475 | + | Frame_Shift_Del | DEL | TG | T | novel | 4533 | AMPLICON | Illumina MiSeq | p.T217fs | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47165538 | 47165538 | + | In_Frame_Del | DEL | AGGTGGAGGC | A | novel | 4606 | AMPLICON | Illumina MiSeq | p.193_196del | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47125604 | 47125604 | + | Missense_Mutation | SNP | A | G | novel | 4715 | AMPLICON | Illumina MiSeq | p.M1889T | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47162528 | 47162528 | + | Nonsense_Mutation | SNP | C | A | novel | 4730 | AMPLICON | Illumina MiSeq | p.E1200X | Chromatin Modification |
| SETD2 | 29072 | Southampton_Cancer_Genomics | hg19 | 3 | 47165237 | 47165237 | + | Missense_Mutation | SNP | T | G | novel | 4762 | AMPLICON | Illumina MiSeq | p.K297Q | Chromatin Modification |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4028 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266512 | 198266512 | + | Missense_Mutation | SNP | C | T | novel | 4047 | AMPLICON | Illumina MiSeq | p.R775Q | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267484 | 198267484 | + | Missense_Mutation | SNP | G | A | novel | 4070 | AMPLICON | Illumina MiSeq | p.R625C | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266512 | 198266512 | + | Missense_Mutation | SNP | C | T | novel | 4104 | AMPLICON | Illumina MiSeq | p.R775Q | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266709 | 198266709 | + | Missense_Mutation | SNP | C | A | novel | 4108 | AMPLICON | Illumina MiSeq | p.K741N | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266596 | 198266596 | + | Missense_Mutation | SNP | A | C | novel | 4133 | AMPLICON | Illumina MiSeq | p.L747W | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267492 | 198267492 | + | Missense_Mutation | SNP | T | A | novel | 4137 | AMPLICON | Illumina MiSeq | p.E622V | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4152 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266603 | 198266603 | + | Missense_Mutation | SNP | C | G | novel | 4158 | AMPLICON | Illumina MiSeq | p.A745P | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266713 | 198266713 | + | Missense_Mutation | SNP | C | T | novel | 4166 | AMPLICON | Illumina MiSeq | p.G740E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4490 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267491 | 198267491 | + | Missense_Mutation | SNP | C | A | novel | 4490 | AMPLICON | Illumina MiSeq | p.E622D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267361 | 198267361 | + | Missense_Mutation | SNP | T | C | novel | 4508 | AMPLICON | Illumina MiSeq | p.K666E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4532 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4548 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4553 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267484 | 198267484 | + | Missense_Mutation | SNP | G | C | novel | 4583 | AMPLICON | Illumina MiSeq | p.R625G | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4628 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267484 | 198267484 | + | Missense_Mutation | SNP | G | A | novel | 4629 | AMPLICON | Illumina MiSeq | p.R625C | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4660 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266607 | 198266607 | + | Missense_Mutation | SNP | C | A | novel | 4670 | AMPLICON | Illumina MiSeq | p.L743F | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4670 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267481 | 198267481 | + | Missense_Mutation | SNP | T | C | novel | 4670 | AMPLICON | Illumina MiSeq | p.N626D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267469 | 198267469 | + | Missense_Mutation | SNP | T | C | novel | 4698 | AMPLICON | Illumina MiSeq | p.R630G | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266476 | 198266476 | + | Frame_Shift_Del | DEL | AT | A | novel | 4748 | AMPLICON | Illumina MiSeq | p.I787fs | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4773 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4784 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4008 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4012 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4016 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4030 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4032 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4033 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266822 | 198266822 | + | Missense_Mutation | SNP | T | A | novel | 4037 | AMPLICON | Illumina MiSeq | p.I704F | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4038 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4039 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4046 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267361 | 198267361 | + | Missense_Mutation | SNP | T | C | novel | 4049 | AMPLICON | Illumina MiSeq | p.K666E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267492 | 198267492 | + | Missense_Mutation | SNP | T | A | novel | 4059 | AMPLICON | Illumina MiSeq | p.E622V | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267361 | 198267361 | + | Missense_Mutation | SNP | T | C | novel | 4061 | AMPLICON | Illumina MiSeq | p.K666E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4071 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4071 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4084 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4084 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4085 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267361 | 198267361 | + | Missense_Mutation | SNP | T | C | novel | 4087 | AMPLICON | Illumina MiSeq | p.K666E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4089 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4091 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4100 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266603 | 198266603 | + | Missense_Mutation | SNP | C | G | novel | 4102 | AMPLICON | Illumina MiSeq | p.A745P | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4112 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4120 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4130 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4133 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266822 | 198266822 | + | Missense_Mutation | SNP | T | A | novel | 4141 | AMPLICON | Illumina MiSeq | p.I704F | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267492 | 198267492 | + | Missense_Mutation | SNP | T | A | novel | 4147 | AMPLICON | Illumina MiSeq | p.E622V | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267489 | 198267489 | + | Missense_Mutation | SNP | T | C | novel | 4158 | AMPLICON | Illumina MiSeq | p.Y623C | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4160 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4166 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266821 | 198266821 | + | Missense_Mutation | SNP | A | C | novel | 4172 | AMPLICON | Illumina MiSeq | p.I704S | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4173 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266713 | 198266713 | + | Missense_Mutation | SNP | C | T | novel | 4175 | AMPLICON | Illumina MiSeq | p.G740E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4190 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4205 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267360 | 198267360 | + | Missense_Mutation | SNP | T | G | novel | 4209 | AMPLICON | Illumina MiSeq | p.K666T | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4221 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4266 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267371 | 198267371 | + | Missense_Mutation | SNP | G | C | novel | 4268 | AMPLICON | Illumina MiSeq | p.H662Q | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4272 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266512 | 198266512 | + | Missense_Mutation | SNP | C | A | novel | 4280 | AMPLICON | Illumina MiSeq | p.R775L | RNA processing |

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267481 | 198267481 | + | Missense_Mutation | SNP | T | A | novel | 4282 | AMPLICON | Illumina MiSeq | p.N626Y | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4285 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4290 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4319 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4324 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4330 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4334 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4338 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267369 | 198267369 | + | Missense_Mutation | SNP | G | A | novel | 4342 | AMPLICON | Illumina MiSeq | p.T663I | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4354 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267371 | 198267371 | + | Missense_Mutation | SNP | G | T | novel | 4367 | AMPLICON | Illumina MiSeq | p.H662Q | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4369 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267491 | 198267491 | + | Missense_Mutation | SNP | C | G | novel | 4382 | AMPLICON | Illumina MiSeq | p.E622D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4393 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4414 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4419 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267373 | 198267373 | + | Missense_Mutation | SNP | G | C | novel | 4433 | AMPLICON | Illumina MiSeq | p.H662D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4439 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4440 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4449 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4451 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4456 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267369 | 198267369 | + | Missense_Mutation | SNP | G | A | novel | 4470 | AMPLICON | Illumina MiSeq | p.T663I | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4472 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4473 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267483 | 198267483 | + | Missense_Mutation | SNP | C | T | novel | 4474 | AMPLICON | Illumina MiSeq | p.R625H | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4485 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4486 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4495 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4501 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267480 | 198267480 | + | Missense_Mutation | SNP | T | C | novel | 4508 | AMPLICON | Illumina MiSeq | p.N626S | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267360 | 198267360 | + | Missense_Mutation | SNP | T | A | novel | 4524 | AMPLICON | Illumina MiSeq | p.K666M | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4533 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4535 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4545 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4558 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4563 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4576 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4578 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4586 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4603 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267373 | 198267373 | + | Missense_Mutation | SNP | G | C | novel | 4603 | AMPLICON | Illumina MiSeq | p.H662D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266822 | 198266822 | + | Missense_Mutation | SNP | T | A | novel | 4621 | AMPLICON | Illumina MiSeq | p.I704F | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4623 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4626 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266596 | 198266596 | + | Missense_Mutation | SNP | A | C | novel | 4628 | AMPLICON | Illumina MiSeq | p.L747W | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4634 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4635 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267359 | 198267359 | + | Missense_Mutation | SNP | C | G | novel | 4650 | AMPLICON | Illumina MiSeq | p.K666N | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4662 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266713 | 198266713 | + | Missense_Mutation | SNP | C | T | novel | 4670 | AMPLICON | Illumina MiSeq | p.G740E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4704 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4713 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4719 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4723 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4732 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4734 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266512 | 198266512 | + | Missense_Mutation | SNP | C | T | novel | 4737 | AMPLICON | Illumina MiSeq | p.R775Q | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267373 | 198267373 | + | Missense_Mutation | SNP | G | C | novel | 4747 | AMPLICON | Illumina MiSeq | p.H662D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267373 | 198267373 | + | Missense_Mutation | SNP | G | C | novel | 4748 | AMPLICON | Illumina MiSeq | p.H662D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198267360 | 198267360 | + | Missense_Mutation | SNP | T | C | novel | 4750 | AMPLICON | Illumina MiSeq | p.K666R | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4755 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266611 | 198266611 | + | Missense_Mutation | SNP | C | T | novel | 4769 | AMPLICON | Illumina MiSeq | p.G742D | RNA processing |
| SF3B1 | 23451 | Southampton_Cancer_Genomics | hg19 | 2 | 198266834 | 198266834 | + | Missense_Mutation | SNP | T | C | novel | 4777 | AMPLICON | Illumina MiSeq | p.K700E | RNA processing |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7574035 | 7574035 | + | Splice_Site | SNP | T | A | novel | 4127 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578406 | 7578406 | + | Missense_Mutation | SNP | C | T | rs28934578 | 4150 | AMPLICON | Illumina MiSeq | p.R175H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577085 | 7577085 | + | Missense_Mutation | SNP | C | T | novel | 4169 | AMPLICON | Illumina MiSeq | p.E285K | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7572962 | 7572962 | + | Frame_Shift_Del | DEL | GT | G | novel | 4263 | AMPLICON | Illumina MiSeq | p.K382fs | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7574003 | 7574003 | + | Nonsense_Mutation | SNP | G | A | novel | 4267 | AMPLICON | Illumina MiSeq | p.R342X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578556 | 7578556 | + | Splice_Site | SNP | T | C | novel | 4312 | AMPLICON | Illumina MiSeq | NA | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578523 | 7578523 | + | Missense_Mutation | SNP | T | A | novel | 4355 | AMPLICON | Illumina MiSeq | p.Q136L | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577082 | 7577082 | + | Nonsense_Mutation | SNP | C | A | novel | 4419 | AMPLICON | Illumina MiSeq | p.E286X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578395 | 7578395 | + | Missense_Mutation | SNP | G | A | novel | 4419 | AMPLICON | Illumina MiSeq | p.H179Y | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577120 | 7577120 | + | Missense_Mutation | SNP | C | T | rs28934576 | 4432 | AMPLICON | Illumina MiSeq | p.R273H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578403 | 7578403 | + | Missense_Mutation | SNP | C | A | novel | 4451 | AMPLICON | Illumina MiSeq | p.C176F | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578221 | 7578221 | + | Frame_Shift_Del | DEL | TTC | T | novel | 4546 | AMPLICON | Illumina MiSeq | p.R209fs | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577094 | 7577094 | + | Missense_Mutation | SNP | G | A | rs28934574 | 4572 | AMPLICON | Illumina MiSeq | p.R282W | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577142 | 7577142 | + | Missense_Mutation | SNP | C | T | novel | 4580 | AMPLICON | Illumina MiSeq | p.G266R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578524 | 7578524 | + | Nonsense_Mutation | SNP | G | A | novel | 4597 | AMPLICON | Illumina MiSeq | p.Q136X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577105 | 7577105 | + | Missense_Mutation | SNP | G | C | novel | 4694 | AMPLICON | Illumina MiSeq | p.P278R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7574030 | 7574030 | + | Missense_Mutation | SNP | G | A | novel | 4016 | AMPLICON | Illumina MiSeq | p.R333C | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577538 | 7577538 | + | Missense_Mutation | SNP | C | T | rs11540652 | 4049 | AMPLICON | Illumina MiSeq | p.R248Q | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577560 | 7577560 | + | Missense_Mutation | SNP | A | C | novel | 4070 | AMPLICON | Illumina MiSeq | p.S241A | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577114 | 7577114 | + | Missense_Mutation | SNP | C | A | novel | 4081 | AMPLICON | Illumina MiSeq | p.C275F | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578503 | 7578503 | + | Missense_Mutation | SNP | C | T | novel | 4091 | AMPLICON | Illumina MiSeq | p.V143M | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578204 | 7578204 | + | Missense_Mutation | SNP | A | C | novel | 4092 | AMPLICON | Illumina MiSeq | p.S215R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577154 | 7577154 | + | Frame_Shift_Del | DEL | CACTACTCAGG | C | novel | 4105 | AMPLICON | Illumina MiSeq | p.S261fs | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577106 | 7577106 | + | Missense_Mutation | SNP | G | C | rs17849781 | 4107 | AMPLICON | Illumina MiSeq | p.P278A | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577121 | 7577121 | + | Missense_Mutation | SNP | G | A | novel | 4140 | AMPLICON | Illumina MiSeq | p.R273C | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578457 | 7578457 | + | Missense_Mutation | SNP | C | T | novel | 4140 | AMPLICON | Illumina MiSeq | p.R158H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577064 | 7577064 | + | In_Frame_Del | DEL | TCTTGCGGAGATTCTCTTC | T | novel | 4170 | AMPLICON | Illumina MiSeq | p.286_291del | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578406 | 7578406 | + | Missense_Mutation | SNP | C | T | rs28934578 | 4171 | AMPLICON | Illumina MiSeq | p.R175H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577114 | 7577114 | + | Missense_Mutation | SNP | C | A | novel | 4197 | AMPLICON | Illumina MiSeq | p.C275F | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578509 | 7578509 | + | Missense_Mutation | SNP | A | G | novel | 4253 | AMPLICON | Illumina MiSeq | p.C141R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578460 | 7578460 | + | Missense_Mutation | SNP | A | G | novel | 4338 | AMPLICON | Illumina MiSeq | p.V157A | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577548 | 7577548 | + | Missense_Mutation | SNP | C | T | rs28934575 | 4366 | AMPLICON | Illumina MiSeq | p.G245S | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577069 | 7577069 | + | Missense_Mutation | SNP | C | T | rs55819519 | 4368 | AMPLICON | Illumina MiSeq | p.R290H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578538 | 7578538 | + | Missense_Mutation | SNP | T | C | novel | 4464 | AMPLICON | Illumina MiSeq | p.N131S | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578274 | 7578274 | + | In_Frame_Del | DEL | TGAG | T | novel | 4478 | AMPLICON | Illumina MiSeq | p.191_192del | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578436 | 7578436 | + | Frame_Shift_Del | DEL | TG | T | novel | 4481 | AMPLICON | Illumina MiSeq | p.Q165fs | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577094 | 7577094 | + | Missense_Mutation | SNP | G | C | rs28934574 | 4491 | AMPLICON | Illumina MiSeq | p.R282G | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578437 | 7578437 | + | Nonsense_Mutation | SNP | G | A | novel | 4491 | AMPLICON | Illumina MiSeq | p.Q165X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577094 | 7577094 | + | Missense_Mutation | SNP | G | C | rs28934574 | 4497 | AMPLICON | Illumina MiSeq | p.R282G | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578265 | 7578265 | + | Missense_Mutation | SNP | A | G | novel | 4527 | AMPLICON | Illumina MiSeq | p.I195T | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7579366 | 7579366 | + | Nonsense_Mutation | SNP | G | T | novel | 4548 | AMPLICON | Illumina MiSeq | p.Y107X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577118 | 7577118 | + | Frame_Shift_Ins | INS | C | CA | novel | 4594 | AMPLICON | Illumina MiSeq | p.V274fs | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578415 | 7578415 | + | In_Frame_Del | DEL | ACCT | A | novel | 4599 | AMPLICON | Illumina MiSeq | p.171_172del | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577120 | 7577120 | + | Missense_Mutation | SNP | C | T | rs28934576 | 4623 | AMPLICON | Illumina MiSeq | p.R273H | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577560 | 7577560 | + | Missense_Mutation | SNP | A | C | novel | 4643 | AMPLICON | Illumina MiSeq | p.S241A | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578394 | 7578394 | + | Missense_Mutation | SNP | T | C | novel | 4648 | AMPLICON | Illumina MiSeq | p.H179R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7577580 | 7577580 | + | Missense_Mutation | SNP | T | C | novel | 4662 | AMPLICON | Illumina MiSeq | p.Y234C | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578208 | 7578208 | + | Missense_Mutation | SNP | T | C | novel | 4674 | AMPLICON | Illumina MiSeq | p.H214R | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7579519 | 7579519 | + | In_Frame_Del | DEL | TTCAGTGAAC | T | novel | 4684 | AMPLICON | Illumina MiSeq | p.53_56del | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578412 | 7578412 | + | Missense_Mutation | SNP | A | C | novel | 4721 | AMPLICON | Illumina MiSeq | p.V173G | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578271 | 7578271 | + | Missense_Mutation | SNP | T | A | novel | 4727 | AMPLICON | Illumina MiSeq | p.H193L | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7576897 | 7576897 | + | Nonsense_Mutation | SNP | G | A | novel | 4753 | AMPLICON | Illumina MiSeq | p.Q317X | DDR/CC |
| TP53 | 7157 | Southampton_Cancer_Genomics | hg19 | 17 | 7578202 | 7578202 | + | Missense_Mutation | SNP | A | C | novel | 4784 | AMPLICON | Illumina MiSeq | p.V216G | DDR/CC |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4071 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4072 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719187 | 61719187 | + | Missense_Mutation | SNP | C | A | novel | 4101 | AMPLICON | Illumina MiSeq | p.D624Y | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4178 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4310 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4419 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4631 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4008 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4014 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4062 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4101 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4103 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4104 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4121 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | G | novel | 4145 | AMPLICON | Illumina MiSeq | p.E571A | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4282 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4307 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | A | novel | 4310 | AMPLICON | Illumina MiSeq | p.E571V | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4351 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | A | novel | 4367 | AMPLICON | Illumina MiSeq | p.E571V | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4375 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719462 | 61719462 | + | Missense_Mutation | SNP | T | C | novel | 4454 | AMPLICON | Illumina MiSeq | p.H574R | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4510 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | G | novel | 4519 | AMPLICON | Illumina MiSeq | p.E571A | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | G | novel | 4528 | AMPLICON | Illumina MiSeq | p.E571A | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4546 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4566 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719186 | 61719186 | + | Missense_Mutation | SNP | T | C | novel | 4631 | AMPLICON | Illumina MiSeq | p.D624G | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4638 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719472 | 61719472 | + | Missense_Mutation | SNP | C | T | novel | 4760 | AMPLICON | Illumina MiSeq | p.E571K | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719471 | 61719471 | + | Missense_Mutation | SNP | T | A | novel | 4761 | AMPLICON | Illumina MiSeq | p.E571V | RNA processing |
| XPO1 | 7514 | Southampton_Cancer_Genomics | hg19 | 2 | 61719186 | 61719186 | + | Missense_Mutation | SNP | T | C | novel | 4762 | AMPLICON | Illumina MiSeq | p.D624G | RNA processing |
| ZFPM2 | 23414 | Southampton_Cancer_Genomics | hg19 | 8 | 106813404 | 106813404 | + | Missense_Mutation | SNP | G | A | novel | 4047 | AMPLICON | Illumina MiSeq | p.C365Y | Wnt |
| ZFPM2 | 23414 | Southampton_Cancer_Genomics | hg19 | 8 | 106815750 | 106815750 | + | Missense_Mutation | SNP | C | T | novel | 4047 | AMPLICON | Illumina MiSeq | p.A1147V | Wnt |
| ZFPM2 | 23414 | Southampton_Cancer_Genomics | hg19 | 8 | 106815207 | 106815207 | + | Missense_Mutation | SNP | G | A | novel | 4235 | AMPLICON | Illumina MiSeq | p.C966Y | Wnt |
| ZFPM2 | 23414 | Southampton_Cancer_Genomics | hg19 | 8 | 106815209 | 106815209 | + | Frame_Shift_Del | DEL | CT | C | novel | 4335 | AMPLICON | Illumina MiSeq | p.L967fs | Wnt |
| ZFPM2 | 23414 | Southampton_Cancer_Genomics | hg19 | 8 | 106815296 | 106815296 | + | Missense_Mutation | SNP | T | C | novel | 4335 | AMPLICON | Illumina MiSeq | p.S996P | Wnt |

## A.8 Table of all CLL, U-CLL and M-CLL power calculations

| Mutated gene recurrence (%) | All CLL (n=500) Power (%) | U-CLL (n=263) Power (%) | M-CLL (n=157) Power (%) |
|---|---|---|---|
| 20 | 100 | 100 | 100 |
| 10 | 100 | 100 | 100 |
| 5 | 100 | 97 | 76 |
| 3 | 96 | 65 | 32 |
| 2 | 69 | 29 | 11 |
| 1 | 12 | 3 | 1 |

## A.9 Waterfall plots of U-CLL and M-CLL




Waterfall plots showing the distribution of high confidence variants stratified by mutation type (key to the right) across the 25 genes for 263 U-CLL cases **(A)** or 157 M-CLL cases **(B)**, along with their mutation frequency (bar chart, left).

# A.10 Breakdown of 3 mutated intracellular signalling pathways

| Pathway 1 | Pathway 2 | Pathway 3 | Number of cases | Percentage |
|---|---|---|---|---|
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | RNA processing | 13 | 22.81 |
| Inflammatory pathway/B-Cell Receptor signalling | NOTCH signalling | RNA processing | 7 | 12.28 |
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | NOTCH signalling | 4 | 7.02 |
| DNA Damage Response/Cell Cycle Control | RNA processing | Wnt signalling | 4 | 7.02 |
| Chromatin Modification | DNA Damage Response/Cell Cycle Control | RNA processing | 3 | 5.26 |
| Autophagy | Inflammatory pathway/B-Cell Receptor signalling | Wnt signalling | 2 | 3.51 |
| Chromatin Modification | DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | 2 | 3.51 |
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | RNA processing | 2 | 3.51 |
| DNA Damage Response/Cell Cycle Control | NOTCH signalling | RNA processing | 2 | 3.51 |
| Inflammatory pathway/B-Cell Receptor signalling | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | RNA processing | 2 | 3.51 |
| Autophagy | DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | 1 | 1.75 |
| Autophagy | DNA Damage Response/Cell Cycle Control | RNA processing | 1 | 1.75 |
| Autophagy | Inflammatory pathway/B-Cell Receptor signalling | RNA processing | 1 | 1.75 |
| Chromatin Modification | DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | 1 | 1.75 |
| Chromatin Modification | DNA Damage Response/Cell Cycle Control | Wnt signalling | 1 | 1.75 |
| Chromatin Modification | Inflammatory pathway/B-Cell Receptor signalling | Wnt signalling | 1 | 1.75 |
| Chromatin Modification | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | RNA processing | 1 | 1.75 |
| Chromatin Modification | NOTCH signalling | Wnt signalling | 1 | 1.75 |
| Chromatin Modification | NOTCH signalling/MYC | RNA processing | 1 | 1.75 |
| DNA Damage Response/Cell Cycle Control | Inflammatory pathway/B-Cell Receptor signalling | NOTCH/MYC | 1 | 1.75 |
| DNA Damage Response/Cell Cycle Control | NOTCH signalling | NOTCH/MYC | 1 | 1.75 |
| MYC | Inflammatory pathway/B-Cell Receptor signalling | RNA processing | 1 | 1.75 |
| Inflammatory pathway/B-Cell Receptor signalling | Inflammatory pathway/B-Cell Receptor signalling/RNA processing | NOTCH signalling | 1 | 1.75 |
| Inflammatory pathway/B-Cell Receptor signalling | NOTCH signalling/MYC | Wnt signalling | 1 | 1.75 |
| Inflammatory pathway/B-Cell Receptor signalling | RNA processing | Wnt signalling | 1 | 1.75 |

## A.11 Significant associations from CLL4 Fisher's Exact testing

| P Value Rank | Tested mut driver | Clonality of tested mut driver | Association partner | Clonality of association partner | P Value | Q Value | Q>P | P<0.001 | P<0.01 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | del(17p) | All | TP53 | All | 3.57E-13 | 3.57E-06 | TRUE | TRUE | TRUE |
| 2 | TP53 | All | del17p | All | 3.57E-13 | 7.14E-06 | TRUE | TRUE | TRUE |
| 3 | del(17p) | Clonal | TP53 | All | 3.59E-13 | 1.07E-05 | TRUE | TRUE | TRUE |
| 4 | TP53 | All | del17p | Clonal | 3.59E-13 | 1.43E-05 | TRUE | TRUE | TRUE |
| 5 | del(17p) | Clonal | TP53 | Clonal | 2.23E-12 | 1.79E-05 | TRUE | TRUE | TRUE |
| 6 | TP53 | Clonal | del17p | Clonal | 2.23E-12 | 2.14E-05 | TRUE | TRUE | TRUE |
| 7 | del(17p) | All | TP53 | Clonal | 6.38E-09 | 2.50E-05 | TRUE | TRUE | TRUE |
| 8 | TP53 | Clonal | del17p | All | 6.38E-09 | 2.86E-05 | TRUE | TRUE | TRUE |
| 9 | IGHV | All | del13q | All | 8.99E-09 | 3.21E-05 | TRUE | TRUE | TRUE |
| 10 | del(13q) | All | IGHV | All | 8.99E-09 | 3.57E-05 | TRUE | TRUE | TRUE |
| 11 | +(12) | Clonal | del13q | All | 3.57E-08 | 3.93E-05 | TRUE | TRUE | TRUE |
| 12 | del(13q) | All | +(12) | Clonal | 3.57E-08 | 4.29E-05 | TRUE | TRUE | TRUE |
| 13 | +(12) | Clonal | del13q | Clonal | 1.17E-07 | 4.64E-05 | TRUE | TRUE | TRUE |
| 14 | del(13q) | Clonal | +(12) | Clonal | 1.17E-07 | 5.00E-05 | TRUE | TRUE | TRUE |
| 15 | IGHV | All | del11q | Clonal | 2.51E-07 | 5.36E-05 | TRUE | TRUE | TRUE |
| 16 | del(11q) | Clonal | IGHV | All | 2.51E-07 | 5.71E-05 | TRUE | TRUE | TRUE |
| 17 | +(12) | All | del13q | All | 1.77E-06 | 6.07E-05 | TRUE | TRUE | TRUE |
| 18 | del(13q) | All | +(12) | All | 1.77E-06 | 6.43E-05 | TRUE | TRUE | TRUE |
| 19 | +(12) | Clonal | BIRC3 | All | 2.08E-06 | 6.79E-05 | TRUE | TRUE | TRUE |
| 20 | BIRC3 | All | +(12) | Clonal | 2.08E-06 | 7.14E-05 | TRUE | TRUE | TRUE |
| 21 | +(12) | All | del13q | Clonal | 2.26E-06 | 7.50E-05 | TRUE | TRUE | TRUE |
| 22 | del(13q) | Clonal | +(12) | All | 2.26E-06 | 7.86E-05 | TRUE | TRUE | TRUE |
| 23 | IGHV | All | del13q | Clonal | 7.84E-06 | 8.21E-05 | TRUE | TRUE | TRUE |
| 24 | del(13q) | All | IGHV | All | 7.84E-06 | 8.57E-05 | TRUE | TRUE | TRUE |
| 25 | +(12) | All | BIRC3 | All | 1.70E-05 | 8.93E-05 | TRUE | TRUE | TRUE |
| 26 | BIRC3 | All | +(12) | All | 1.70E-05 | 9.29E-05 | TRUE | TRUE | TRUE |
| 27 | IGHV | All | del11q | All | 4.87E-05 | 9.64E-05 | TRUE | TRUE | TRUE |
| 28 | del(11q) | All | IGHV | All | 4.87E-05 | 1.00E-04 | TRUE | TRUE | TRUE |
| 29 | del(13q) | All | NOTCH1 | All | 8.42E-05 | 1.04E-04 | TRUE | TRUE | TRUE |
| 30 | NOTCH1 | All | del13q | All | 8.42E-05 | 1.07E-04 | TRUE | TRUE | TRUE |
| 31 | KRAS | Subclonal | RPS15 | All | 9.03E-05 | 1.11E-04 | TRUE | TRUE | TRUE |
| 32 | RPS15 | All | KRAS | Subclonal | 9.03E-05 | 1.14E-04 | TRUE | TRUE | TRUE |
| 33 | del(17p) | All | TP53 | Subclonal | 1.12E-04 | 1.18E-04 | TRUE | TRUE | TRUE |
| 34 | TP53 | Subclonal | del17p | All | 1.12E-04 | 1.21E-04 | TRUE | TRUE | TRUE |
| 35 | IGHV | All | XPO1 | All | 1.52E-04 | 1.25E-04 | FALSE | TRUE | TRUE |
| 36 | XPO1 | All | IGHV | All | 1.52E-04 | 1.29E-04 | FALSE | TRUE | TRUE |
| 37 | BIRC3 | Clonal | NOTCH1 | All | 2.47E-04 | 1.32E-04 | FALSE | TRUE | TRUE |
| 38 | NOTCH1 | All | BIRC3 | Clonal | 2.47E-04 | 1.36E-04 | FALSE | TRUE | TRUE |
| 39 | IGHV | All | XPO1 | Clonal | 2.78E-04 | 1.39E-04 | FALSE | TRUE | TRUE |
| 40 | XPO1 | Clonal | IGHV | All | 2.78E-04 | 1.43E-04 | FALSE | TRUE | TRUE |
| 41 | TP53 | Subclonal | FBXW7 | All | 3.26E-04 | 1.46E-04 | FALSE | TRUE | TRUE |
| 42 | FBXW7 | All | TP53 | Subclonal | 3.26E-04 | 1.50E-04 | FALSE | TRUE | TRUE |
| 43 | +(12) | All | SF3B1 | Clonal | 3.51E-04 | 1.54E-04 | FALSE | TRUE | TRUE |
| 44 | SF3B1 | Clonal | +(12) | All | 3.51E-04 | 1.57E-04 | FALSE | TRUE | TRUE |
| 45 | +(12) | Clonal | BIRC3 | Subclonal | 4.80E-04 | 1.61E-04 | FALSE | TRUE | TRUE |
| 46 | BIRC3 | Subclonal | +(12) | Clonal | 4.80E-04 | 1.64E-04 | FALSE | TRUE | TRUE |
| 47 | del(13q) | All | NOTCH1 | Clonal | 6.35E-04 | 1.68E-04 | FALSE | TRUE | TRUE |
| 48 | NOTCH1 | Clonal | del13q | All | 6.35E-04 | 1.71E-04 | FALSE | TRUE | TRUE |
| 49 | del(17p) | Subclonal | MED12 | All | 7.22E-04 | 1.75E-04 | FALSE | TRUE | TRUE |
| 50 | MED12 | All | del17p | Subclonal | 7.22E-04 | 1.79E-04 | FALSE | TRUE | TRUE |
| 51 | +(12) | Clonal | BIRC3 | Clonal | 7.96E-04 | 1.82E-04 | FALSE | TRUE | TRUE |
| 52 | BIRC3 | Clonal | +(12) | Clonal | 7.96E-04 | 1.86E-04 | FALSE | TRUE | TRUE |
| 53 | del(13q) | Clonal | NOTCH1 | All | 7.99E-04 | 1.89E-04 | FALSE | TRUE | TRUE |
| 54 | NOTCH1 | All | del13q | Clonal | 7.99E-04 | 1.93E-04 | FALSE | TRUE | TRUE |
| 55 | del(13q) | All | FBXW7 | All | 8.14E-04 | 1.96E-04 | FALSE | TRUE | TRUE |
| 56 | FBXW7 | All | del13q | All | 8.14E-04 | 2.00E-04 | FALSE | TRUE | TRUE |
| 57 | TP53 | Subclonal | BRAF | All | 8.74E-04 | 2.04E-04 | FALSE | TRUE | TRUE |
| 58 | BRAF | All | TP53 | Subclonal | 8.74E-04 | 2.07E-04 | FALSE | TRUE | TRUE |
| 59 | KRAS | All | RPS15 | All | 1.04E-03 | 2.11E-04 | FALSE | FALSE | TRUE |
| 60 | RPS15 | All | KRAS | All | 1.04E-03 | 2.14E-04 | FALSE | FALSE | TRUE |
| 61 | SF3B1 | Clonal | POT1 | Clonal | 1.19E-03 | 2.18E-04 | FALSE | FALSE | TRUE |
| 62 | POT1 | Clonal | SF3B1 | Clonal | 1.19E-03 | 2.21E-04 | FALSE | FALSE | TRUE |
| 63 | NOTCH1 | Subclonal | FBXW7 | All | 1.43E-03 | 2.25E-04 | FALSE | FALSE | TRUE |
| 64 | FBXW7 | All | NOTCH1 | Subclonal | 1.43E-03 | 2.29E-04 | FALSE | FALSE | TRUE |
| 65 | +(12) | All | BIRC3 | Subclonal | 1.48E-03 | 2.32E-04 | FALSE | FALSE | TRUE |
| 66 | BIRC3 | Subclonal | +(12) | All | 1.48E-03 | 2.36E-04 | FALSE | FALSE | TRUE |
| 67 | SF3B1 | Clonal | POT1 | All | 1.53E-03 | 2.39E-04 | FALSE | FALSE | TRUE |
| 68 | POT1 | All | SF3B1 | Clonal | 1.53E-03 | 2.43E-04 | FALSE | FALSE | TRUE |
| 69 | del(13q) | Clonal | FBXW7 | All | 1.58E-03 | 2.46E-04 | FALSE | FALSE | TRUE |
| 70 | FBXW7 | All | del13q | Clonal | 1.58E-03 | 2.50E-04 | FALSE | FALSE | TRUE |
| 71 | del(17p) | Subclonal | TP53 | Subclonal | 1.64E-03 | 2.54E-04 | FALSE | FALSE | TRUE |
| 72 | TP53 | Subclonal | del17p | Subclonal | 1.64E-03 | 2.57E-04 | FALSE | FALSE | TRUE |
| 73 | +(12) | Clonal | SF3B1 | Clonal | 2.11E-03 | 2.61E-04 | FALSE | FALSE | TRUE |
| 74 | SF3B1 | Clonal | +(12) | Clonal | 2.11E-03 | 2.64E-04 | FALSE | FALSE | TRUE |
| 75 | +(12) | All | SF3B1 | All | 2.21E-03 | 2.68E-04 | FALSE | FALSE | TRUE |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 76 | SF3B1 | All | +(12) | All | 2.21E-03 | 2.71E-04 | FALSE | FALSE | TRUE |
| 77 | +(12) | All | BIRC3 | Clonal | 2.21E-03 | 2.75E-04 | FALSE | FALSE | TRUE |
| 78 | BIRC3 | Clonal | +(12) | All | 2.21E-03 | 2.79E-04 | FALSE | FALSE | TRUE |
| 79 | del(13q) | Clonal | NOTCH1 | Clonal | 2.25E-03 | 2.82E-04 | FALSE | FALSE | TRUE |
| 80 | NOTCH1 | Clonal | del13q | Clonal | 2.25E-03 | 2.86E-04 | FALSE | FALSE | TRUE |
| 81 | del(13q) | All | KRAS | All | 2.30E-03 | 2.89E-04 | FALSE | FALSE | TRUE |
| 82 | KRAS | All | del13q | All | 2.30E-03 | 2.93E-04 | FALSE | FALSE | TRUE |
| 83 | SF3B1 | All | POT1 | Clonal | 2.60E-03 | 2.96E-04 | FALSE | FALSE | TRUE |
| 84 | POT1 | Clonal | SF3B1 | All | 2.60E-03 | 3.00E-04 | FALSE | FALSE | TRUE |
| 85 | RPS15 | All | POT1 | Subclonal | 2.76E-03 | 3.04E-04 | FALSE | FALSE | TRUE |
| 86 | POT1 | Subclonal | RPS15 | All | 2.76E-03 | 3.07E-04 | FALSE | FALSE | TRUE |
| 87 | BIRC3 | Clonal | NOTCH1 | Clonal | 2.81E-03 | 3.11E-04 | FALSE | FALSE | TRUE |
| 88 | NOTCH1 | Clonal | BIRC3 | Clonal | 2.81E-03 | 3.14E-04 | FALSE | FALSE | TRUE |
| 89 | +(12) | All | IGHV | All | 2.81E-03 | 3.18E-04 | FALSE | FALSE | TRUE |
| 90 | IGHV | All | +(12) | All | 2.81E-03 | 3.21E-04 | FALSE | FALSE | TRUE |
| 91 | BIRC3 | Subclonal | XPO1 | Clonal | 3.09E-03 | 3.25E-04 | FALSE | FALSE | TRUE |
| 92 | XPO1 | Clonal | BIRC3 | Subclonal | 3.09E-03 | 3.29E-04 | FALSE | FALSE | TRUE |
| 93 | del(13q) | All | KRAS | Clonal | 3.26E-03 | 3.32E-04 | FALSE | FALSE | TRUE |
| 94 | KRAS | Clonal | del13q | All | 3.26E-03 | 3.36E-04 | FALSE | FALSE | TRUE |
| 95 | IGHV | All | KRAS | All | 3.29E-03 | 3.39E-04 | FALSE | FALSE | TRUE |
| 96 | KRAS | All | IGHV | All | 3.29E-03 | 3.43E-04 | FALSE | FALSE | TRUE |
| 97 | KRAS | All | BRAF | Clonal | 3.49E-03 | 3.46E-04 | FALSE | FALSE | TRUE |
| 98 | BRAF | Clonal | KRAS | All | 3.49E-03 | 3.50E-04 | FALSE | FALSE | TRUE |
| 99 | +(12) | Clonal | SF3B1 | All | 3.73E-03 | 3.54E-04 | FALSE | FALSE | TRUE |
| 100 | SF3B1 | All | +(12) | Clonal | 3.73E-03 | 3.57E-04 | FALSE | FALSE | TRUE |
| 101 | BRAF | All | CHD2 | All | 4.06E-03 | 3.61E-04 | FALSE | FALSE | TRUE |
| 102 | CHD2 | All | BRAF | All | 4.06E-03 | 3.64E-04 | FALSE | FALSE | TRUE |
| 103 | IGHV | All | RPS15 | All | 4.33E-03 | 3.68E-04 | FALSE | FALSE | TRUE |
| 104 | RPS15 | All | IGHV | All | 4.33E-03 | 3.71E-04 | FALSE | FALSE | TRUE |
| 105 | +(12) | All | BRAF | Clonal | 4.74E-03 | 3.75E-04 | FALSE | FALSE | TRUE |
| 106 | BRAF | Clonal | +(12) | All | 4.74E-03 | 3.79E-04 | FALSE | FALSE | TRUE |
| 107 | +(12) | Clonal | KRAS | All | 4.82E-03 | 3.82E-04 | FALSE | FALSE | TRUE |
| 108 | KRAS | All | +(12) | Clonal | 4.82E-03 | 3.86E-04 | FALSE | FALSE | TRUE |
| 109 | KRAS | Subclonal | BRAF | Clonal | 4.85E-03 | 3.89E-04 | FALSE | FALSE | TRUE |
| 110 | BRAF | Clonal | KRAS | Subclonal | 4.85E-03 | 3.93E-04 | FALSE | FALSE | TRUE |
| 111 | +(12) | Clonal | KRAS | Subclonal | 5.06E-03 | 3.96E-04 | FALSE | FALSE | TRUE |
| 112 | KRAS | Subclonal | +(12) | Clonal | 5.06E-03 | 4.00E-04 | FALSE | FALSE | TRUE |
| 113 | SF3B1 | All | POT1 | All | 5.29E-03 | 4.04E-04 | FALSE | FALSE | TRUE |
| 114 | POT1 | All | SF3B1 | All | 5.29E-03 | 4.07E-04 | FALSE | FALSE | TRUE |
| 115 | IGHV | All | NFKBIE | Subclonal | 5.30E-03 | 4.11E-04 | FALSE | FALSE | TRUE |
| 116 | NFKBIE | Subclonal | IGHV | All | 5.30E-03 | 4.14E-04 | FALSE | FALSE | TRUE |
| 117 | NOTCH1 | All | FBXW7 | All | 5.72E-03 | 4.18E-04 | FALSE | FALSE | TRUE |
| 118 | FBXW7 | All | NOTCH1 | All | 5.72E-03 | 4.21E-04 | FALSE | FALSE | TRUE |
| 119 | del(17p) | Subclonal | SETD2 | All | 5.72E-03 | 4.25E-04 | FALSE | FALSE | TRUE |
| 120 | SETD2 | All | del17p | Subclonal | 5.72E-03 | 4.29E-04 | FALSE | FALSE | TRUE |
| 121 | del(13q) | Clonal | KRAS | Clonal | 5.74E-03 | 4.32E-04 | FALSE | FALSE | TRUE |
| 122 | KRAS | Clonal | del13q | Clonal | 5.74E-03 | 4.36E-04 | FALSE | FALSE | TRUE |
| 123 | IGHV | All | del13q | Subclonal | 5.95E-03 | 4.39E-04 | FALSE | FALSE | TRUE |
| 124 | del(13q) | Subclonal | IGHV | All | 5.95E-03 | 4.43E-04 | FALSE | FALSE | TRUE |
| 125 | RPS15 | All | NFKBIE | All | 6.48E-03 | 4.46E-04 | FALSE | FALSE | TRUE |
| 126 | NFKBIE | All | RPS15 | All | 6.48E-03 | 4.50E-04 | FALSE | FALSE | TRUE |
| 127 | BIRC3 | Subclonal | XPO1 | All | 6.88E-03 | 4.54E-04 | FALSE | FALSE | TRUE |
| 128 | XPO1 | All | BIRC3 | Subclonal | 6.88E-03 | 4.57E-04 | FALSE | FALSE | TRUE |
| 129 | RPS15 | Clonal | NFKBIE | All | 7.06E-03 | 4.61E-04 | FALSE | FALSE | TRUE |
| 130 | NFKBIE | All | RPS15 | Clonal | 7.06E-03 | 4.64E-04 | FALSE | FALSE | TRUE |
| 131 | BIRC3 | All | XPO1 | All | 7.62E-03 | 4.68E-04 | FALSE | FALSE | TRUE |
| 132 | XPO1 | All | BIRC3 | All | 7.62E-03 | 4.71E-04 | FALSE | FALSE | TRUE |
| 133 | BIRC3 | All | SF3B1 | Clonal | 7.68E-03 | 4.75E-04 | FALSE | FALSE | TRUE |
| 134 | SF3B1 | Clonal | BIRC3 | All | 7.68E-03 | 4.79E-04 | FALSE | FALSE | TRUE |
| 135 | del(11q) | Clonal | DDX3X | Clonal | 7.70E-03 | 4.82E-04 | FALSE | FALSE | TRUE |
| 136 | DDX3X | Clonal | del11q | Clonal | 7.70E-03 | 4.86E-04 | FALSE | FALSE | TRUE |
| 137 | +(12) | All | NOTCH1 | Clonal | 8.66E-03 | 4.89E-04 | FALSE | FALSE | TRUE |
| 138 | NOTCH1 | Clonal | +(12) | All | 8.66E-03 | 4.93E-04 | FALSE | FALSE | TRUE |
| 139 | BIRC3 | All | NOTCH1 | All | 9.23E-03 | 4.96E-04 | FALSE | FALSE | TRUE |
| 140 | NOTCH1 | All | BIRC3 | All | 9.23E-03 | 5.00E-04 | FALSE | FALSE | TRUE |
| 141 | RPS15 | Clonal | POT1 | Subclonal | 9.59E-03 | 5.04E-04 | FALSE | FALSE | TRUE |
| 142 | POT1 | Subclonal | RPS15 | Clonal | 9.59E-03 | 5.07E-04 | FALSE | FALSE | TRUE |
| 143 | del(11q) | All | DDX3X | All | 9.65E-03 | 5.11E-04 | FALSE | FALSE | TRUE |
| 144 | DDX3X | All | del11q | All | 9.65E-03 | 5.14E-04 | FALSE | FALSE | TRUE |
| 145 | IGHV | All | SF3B1 | Clonal | 9.75E-03 | 5.18E-04 | FALSE | FALSE | TRUE |
| 146 | SF3B1 | Clonal | IGHV | All | 9.75E-03 | 5.21E-04 | FALSE | FALSE | TRUE |

## A.12 Significant variables – Schoenfeld residuals

**A**

**OS – Non proportional hazards**

| Variable | *P* Value |
|---|---|
| Clonal *ATM* | 0.0357 |
| Clonal *SF3B1* | 0.0226 |
| Clonal *TP53* | 0.000771 |
| Subclonal *RPS15* | 0.0402 |
| *ATM* | 0.0181 |
| Pathogenic *ATM* | 0.0407 |
| *TP53* | 0.000535 |
| FDR Treatment | 0.0214 |
| Age (<60yrs) | 0.0494 |
| Binet Stage B | 0.0161 |
| Binet Stage C | 0.0109 |

**B**

**PFS – Non proportional hazards**

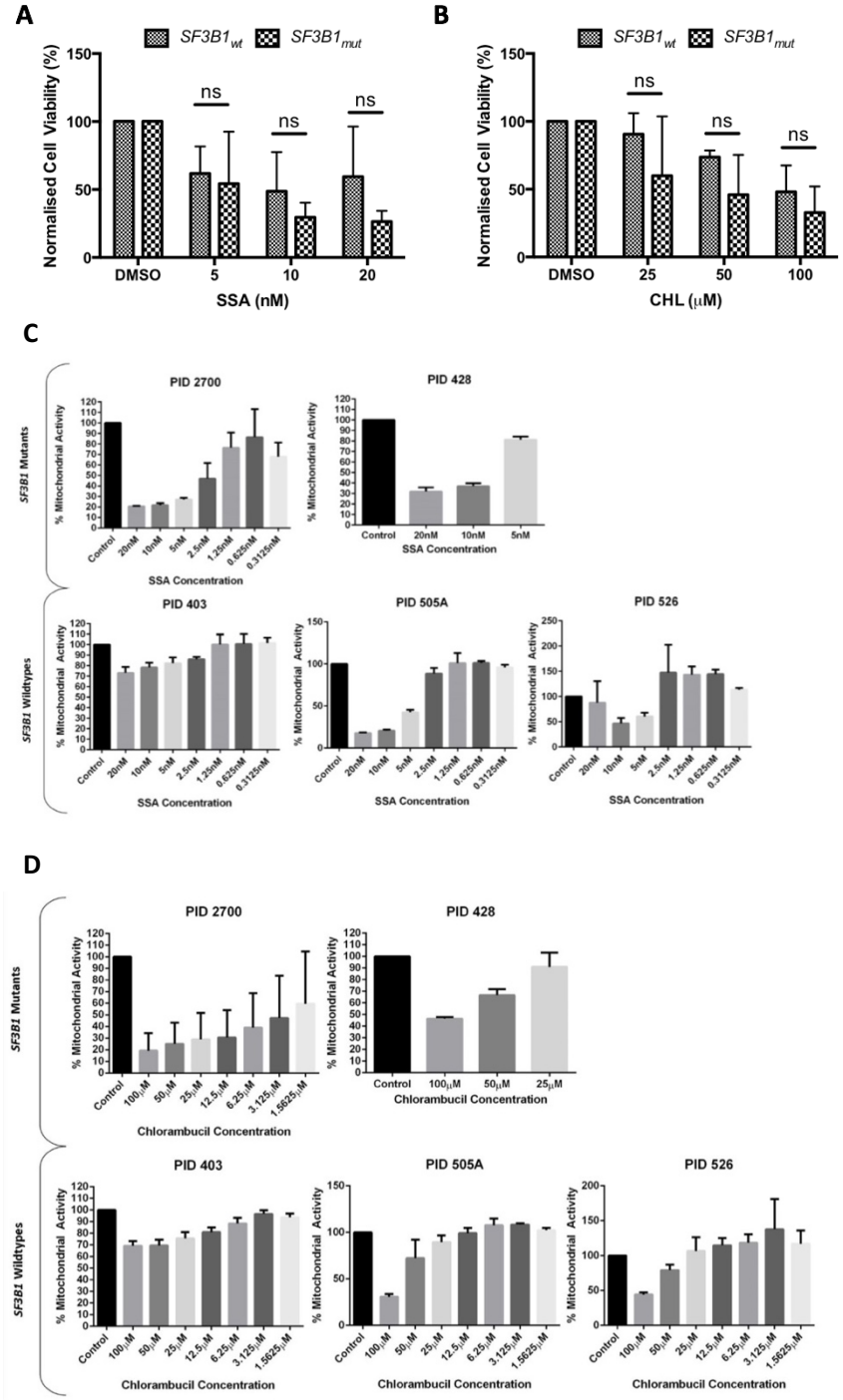
| Variable | *P* Value |
|---|---|
| Clonal *HIST1H1E* | 0.0486 |
| Clonal *MYD88* | 0.0429 |
| Clonal *NBEAL2* | 0.0417 |
| Clonal *SETD2* | 0.0464 |
| Clonal *TP53* | 1.33E-05 |
| Subclonal *BIRC3* | 0.0025 |
| Subclonal *NFKBIE* | 0.0345 |
| *BIRC3* | 0.0351 |
| *DDX3X* | 0.0157 |
| *FBXW7* | 0.0411 |
| *NFKBIE* | 0.00682 |
| *SAMHD1* | 0.0252 |
| *TP53* | 0.000318 |
| FC Treatment | 0.00171 |
| Gender (Male) | 0.013 |

Tables of variables with significant departures from 0 in Schoenfeld residual analysis for OS **(A)** and PFS **(B)**.

## A.13 Cohort characteristics of SSA/MMB & UMI-77 Southampton CLL cases

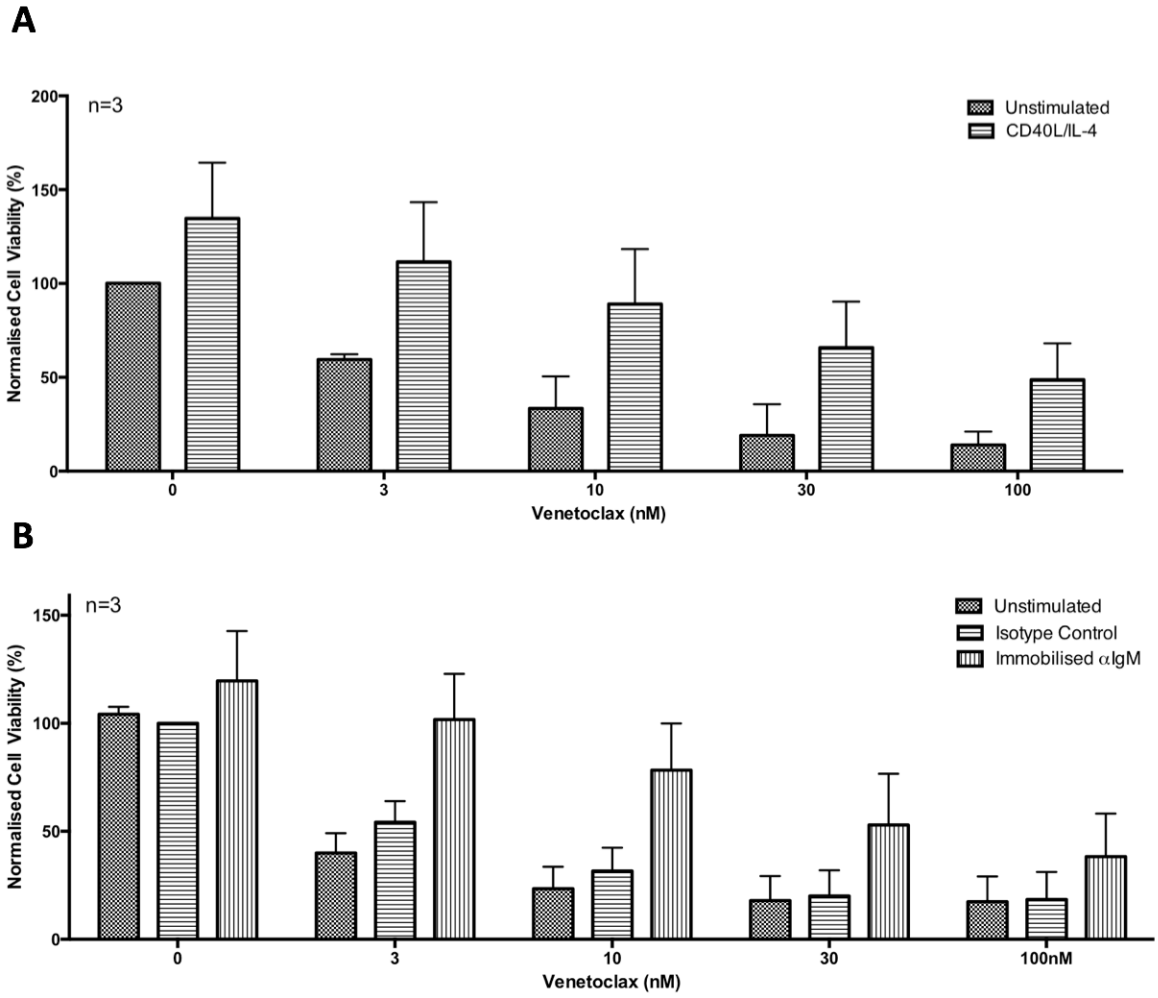
| Cohort | Southampton CLL ID | *IGHV* Mutational Status | ZAP70+ | CD38+ | CD49d | IgM iCa% | sIgM | *SF3B1* mutated |
|---|---|---|---|---|---|---|---|---|
| SSA/MMB | 322 | ND | - | - | - | - | - | 1 |
| | 403 | M (IgG) | - | - | - | - | - | 0 |
| | 428 | U | - | - | - | - | - | 1 |
| | 494 | M | - | - | - | - | - | 0 |
| | 505A | U | 14 | 16 | - | - | - | 0 |
| | 513 | U | - | 1 | - | 38 | - | 0 |
| | 526 | U | 40 | - | - | 6 | - | 0 |
| | 2700RBH | U | 2 | 81 | - | - | - | 1 |
| | 4189RBH | - | 6 | 50 | - | - | - | 1 |
| | 684 | M | 3 | 7 | - | 7 | 18 | - |
| UMI-77 | 163B | U | 64 | 32 | - | 8 | - | - |
| | 409D | U | 23 | 24 | 75 | 64 | 132 | - |
| | 456A | M | | 4 | - | 29 | - | - |
| | 465A | M | 30 | 12 | - | 48 | - | - |
| | 466 | U | 4 | 3 | 11 | 12 | 45 | - |
| | 483A | M | 1 | 1 | 18 | 2 | 108 | - |
| | 496A | M | 0.4 | 0 | 12 | 5 | 67 | - |
| | 563C | M | 0 | 0 | 13 | 81 | 62 | - |
| | 674C | U | 35 | 35 | 51 | 88 | 562 | - |
| | 680 | M | 1 | 83 | 40 | - | 7 | - |
| | 684 | M | 3 | 7 | 20 | 7 | 18 | - |
| | 690 | M | 10 | 0 | 18 | 1 | 18 | - |
| | 706 | M | - | 5 | 60 | 86 | 827 | - |
| | 716 | M | 1 | 0 | 21 | 85 | 123 | - |
| | 729 | M | 0 | 1 | 22 | 44 | 58 | - |
| | 733 | M | 0 | 10 | 12 | 5 | 33 | - |
| | 734 | M | 0 | 1 | 12 | 44 | 35 | - |
| | 735 | M | 0 | 0 | 17 | 27 | 48 | - |
| | 739 | U | 1 | 81 | 10 | 76 | 131 | - |
| | 742 | M | 0 | 1 | 7 | 45 | 49 | - |
| | 751 | M | 10 | 64 | 49 | 87 | 2235 | - |
| | 765 | M | 2 | 1 | 10 | 32 | 28 | - |
| | 766 | U | 11 | 0 | 7 | 78 | 92 | - |
| | 771 | U | 46 | 38 | 39 | 73 | 160 | - |
| | 774 | U | 23 | 0 | 24 | 27 | 179 | - |
| | 775 | U | 29 | 7 | 9 | 57 | 169 | - |
| | 776 | U | 21 | 15 | 8 | 39 | 113 | - |
| | 777 | U | 6 | 97 | 85 | 51 | 1529 | - |

## A.14 SSA reduces cell viability in CLL cells at low nanomolar concentrations in comparison to chlorambucil




**A** CLL samples (n=5) were treated 5-20nM SSA for 24 hours (*SF3B1* mutant n=2, *SF3B1* wild type n=3). Cell viability was measured by MTT mitochondrial activity and the percentage of viable cells normalised to the vehicle control. **B** CLL samples (n=5) were treated with 25-100μM CHL for 24 hours. Full range of concentrations per sample for SSA (**C**) and CHL (**D**). Cell viability was measured by MTT mitochondrial activity and the percentage of viable cells normalised to the vehicle control. Bars represent mean ± s.d.

## A.15 The effect of microenvironmental signals on venetoclax induced apoptosis




CLL samples (n=3) were pre-treated with either CD40L/IL-4 **(A)** or immobilised αIgM **(B)** for 6 hours, followed by 3-100nM venetoclax treatment for 24 hours. Cell viability was measured by Annexin V/PI flow cytometry for CD40L/IL-4 stimulated cells, and AV APC for immobilised αIgM stimulated cells. Cell viability was normalised to the percentage of viable cells of the vehicle control of unstimulated cells (CD40L/IL-4) or isotype control (immobilised αIgM).

# List of References


1. WHO. All Cancers (excluding non-melanoma skin cancer) Estimated Incidence, Mortality and Prevalence Worldwide in 2012 [Internet]. 2012 [cited 2014 Dec 5]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

2. Cancer Research UK. Cancer Statistics for the UK [Internet]. Cancer Research UK Website. 2012. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics#heading-Three

3. UK CR. Chronic lymphocytic leukaemia (CLL) incidence statistics. Cancer Research UK Website. 2014.

4. Butler T, Gribben J. Biologic and Clinical Significance of Molecular Profiling in Chronic Lymphocytic Leukaemia. Blood Rev. 2010;(24):135–41.

5. Strefford JC. The genomic landscape of chronic lymphocytic leukaemia: Biological and clinical implications. Br J Haematol. 2015;169(1):14–31.

6. Sutton LA, Rosenquist R. The complex interplay between cell-intrinsic and cell-extrinsic factors driving the evolution of chronic lymphocytic leukemia. Semin Cancer Biol [Internet]. Elsevier Ltd; 2015;34:22–35. Available from: http://dx.doi.org/10.1016/j.semcancer.2015.04.009

7. Oscier DG, Rose-Zerilli MJJ, Winkelmann N, Gonzalez de Castro D, Gomez B, Forster J, et al. The clinical significance of NOTCH1 and SF3B1 mutations in the UK LRF CLL4 trial. Blood [Internet]. 2012 Oct 18 [cited 2013 Jan 16]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/23086750

8. Rossi D, Khiabanian H, Spina V, Ciardullo C, Bruscaggin A, Famà R, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. Blood [Internet]. 2014;123(14):2139–48. Available from: http://www.haematologica.org/cgi/doi/10.3324/haematol.2015.136051

9. Skowronska A, Parker A, Ahmed G, Oldreive C, Davis Z, Richards S, et al. Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom leukemia research fund chronic lymphocytic leukemia 4 trial. J Clin Oncol. 2012;30(36):4524–32.

10. Jeromin S, Weissmann S, Haferlach C, Dicker F, Bayer K, Grossmann V, et al. SF3B1

mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. Leukemia [Internet]. Nature Publishing Group; 2014;28(1):108–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24113472

11. Young E, Noerenberg D, Mansouri L, Ljungström V, Frick M, Sutton L-A, et al. EGR2 mutations define a new clinically aggressive subgroup of chronic lymphocytic leukemia. Leukemia [Internet]. 2016;(August):1–8. Available from: http://www.nature.com/doifinder/10.1038/leu.2016.359

12. Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. Cancer Discov [Internet]. 2014; Available from: http://www.ncbi.nlm.nih.gov/pubmed/24920063

13. Parker H, Rose-Zerilli M, Larrayoz M, Clifford R, Edelmann J, Blakemore S, et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. Leukemia [Internet]. 2016;(April):1–8. Available from: http://www.nature.com/doifinder/10.1038/leu.2016.134

14. Ljungstrom V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, et al. Whole expme sequencing in relapsig chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. Blood. 2016;127(8):1007–16.

15. Rodríguez D, Bretones G, Quesada V, Villamor N, Arango JR, López-Guillermo A, et al. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. Blood [Internet]. 2015;126(2):195–202. Available from: http://www.bloodjournal.org/content/126/2/195.abstract

16. Schultz M. Rudolf Virchow. Emerg Infect Dis [Internet]. 2008 Sep [cited 2014 Dec 5];14(9):1480–1. Available from: http://wwwnc.cdc.gov/eid/article/14/9/08-0667_article.htm

17. Hanahan D, Weinberg R. The hallmarks of cancer. Cell [Internet]. 2000 [cited 2014 Nov 11];100:57–70. Available from: http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.elsevier-e82bf1bf-57fd-3911-add9-8e1592a777a9/c/00816839.pdf

18. Hanahan D, Weinberg R a. Hallmarks of cancer: the next generation. Cell [Internet]. Elsevier Inc.; 2011 Mar 4 [cited 2012 Nov 1];144(5):646–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21376230

19. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. Science (80- ). 2015;349(6255):1483–9.

20. World Health Organisation. Prevention. (Cancer control : knowledge into action : WHO guide for effective programmes ; module 2.) 1.Neoplasms – prevention and control. 2.Health planning. 3.National health programs – organization and administration. 4.Health policy. 5.Guidelines. I.W. Who [Internet]. 2007;1–56. Available from: http://apps.who.int/iris/bitstream/10665/43575/1/9241547111_eng.pdf

21. De Bont R, van Larebeke N. Endogenous DNA damage in humans: A review of quantitative data. Mutagenesis. 2004;19(3):169–85.

22. Ogino S, Gulley ML, den Dunnen JT, Wilson RB. Standard Mutation Nomenclature in Molecular Diagnostics. J Mol Diagnostics [Internet]. 2007;9(1):1–6. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1525157810603546

23. Thiagalingam S, Laken S, Willson JK V, Markowitz SD, Kinzler KW, Vogelstein B, et al. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci U S A [Internet]. 2001;98(5):2698–702. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=30201&tool=pmcentrez&rendertype=abstract

24. Weinberg R a. Oncogenes, antioncogenes, and the molecular bases of multistage carcinogenesis. Cancer Res. 1989;49:3713–21.

25. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio S a JR, Behjati S, Biankin A V, et al. Signatures of mutational processes in human cancer. Nature [Internet]. 2013;500:415–21. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3776390&tool=pmcentrez&rendertype=abstract

26. Smith AG, Howell D, Patmore R, Jack a, Roman E. Incidence of haematological malignancy by sub-type : a report from the Haematological Malignancy Research Network. Br J Cancer [Internet]. Nature Publishing Group; 2011;105(11):1684–92. Available from: http://dx.doi.org/10.1038/bjc.2011.450

27. Grigoropoulos N, Petter R, Van't Veer M, Scott M, Follows G. Leukaemia update. Part 1; diagnosis and maagement. Br med J. 2013;1660(March):366.

28. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, et al.

Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute–Working Group 1996 guidelines. Blood [Internet]. American Society of Hematology; 2008;111(12):5446–56. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2972576&tool=pmcentrez&rendertype=abstract

29. Rozman C, Montserrat E. Chronic Lymphocytic Leukaemia. N Engl J Med. 1995;(333):1052–7.

30. Döhner H, Stilgenbauer S, Benner a, Leupolt E, Kröber a, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med [Internet]. 2000 Dec 28;343(26):1910–6. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3175782&tool=pmcentrez&rendertype=abstract

31. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med [Internet]. 2011 Dec 29;365(26):2497–506. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22150006

32. Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature [Internet]. 2011 Jul 7 [cited 2012 Oct 4];475(7354):101–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3322590&tool=pmcentrez&rendertype=abstract

33. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabanian H, Ma J, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med [Internet]. 2011 Jul 4 [cited 2012 Oct 21];208(7):1389–401. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3135373&tool=pmcentrez&rendertype=abstract

34. Orchard JA, Ibbotson RE, Davis Z, Wiestner A, Rosenwald A, Thomas PW, et al. ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. 2004;363:105–11.

35. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK, Hamblin BTJ. Unmutated Ig V H Genes Are Associated With a More Aggressive Form of Chronic Lymphocytic Leukemia. 1999;1848–54.

36. Hamblin TJ, Orchard JA, Gardiner A, Oscier DG, Davis Z, Stevenson FK. Immunoglobulin V

genes and CD38 expression in CLL To the editor : 2000;2455–7.

37. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan X-J, et al. Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. Blood [Internet]. 2012 May 10 [cited 2014 Oct 16];119(19):4467–75. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3392073&tool=pmcentrez&rendertype=abstract

38. Chiorazzi N, Rai K, Ferrarini M. Chronic lymphocytic leukemia. New Engl J … [Internet]. 2005 [cited 2014 Oct 23];352:804–15. Available from: http://www.nejm.org/doi/full/10.1056/NEJMra041720

39. Morton LM, Wang SS, Cozen W, Linet MS, Chatterjee N, Davis S, et al. Etiologic heterogeneity among non-Hodgkin lymphoma subtypes. Blood [Internet]. 2008 Dec 15 [cited 2014 Oct 23];112(13):5150–60. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2597610&tool=pmcentrez&rendertype=abstract

40. Arena GD, Musto P. Monoclonal B-cell lymphocytosis. 2014;8(9):75–9.

41. LORTHOLARY P, BOIRON M, RIPAULT P, LEVY JP, MANUS A BJ. CHRONIC LYMPHOID LEUKEMIA SECONDARILY ASSOCIATED WITH A MALIGNANT RETICULOPATHY: RICHTER'S SYNDROME. Nouv Rev Fr Hematol. 1964;Sep-Oct(4):621–44.

42. Hurley JN, Fu SM, Kunkel HG. Chromosome abnormalities of leukaemic B cells in chronic lymphocytic leukaemia. Nature. 1980;283(3):76–8.

43. Langer-Safer PR, Levine M, Ward DC. Immunological method for mapping genes on Drosophila polytene chromosomes. Proc Natl Acad Sci U S A [Internet]. 1982;79(14):4381–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6812046%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC346675

44. Seiffert M, Dietrich S, Jethwa A, Glimm H, Lichter P, Zenz T. Exploiting biological diversity and genomic aberrations in chronic lymphocytic leukemia. Leuk Lymphoma [Internet]. 2012;53(6):1023–31. Available from: http://www.tandfonline.com/doi/full/10.3109/10428194.2011.631638

45. Hallek M, Fischer K, Fingerle-Rowson G, Fink AM, Busch R, Mayer J, et al. Addition of

rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: A randomised, open-label, phase 3 trial. Lancet. 2010;376(9747):1164–74.

46. Zenz T, Eichhorst B, Busch R, Denzel T, Habe S, Winkler D, et al. TP53 Mutation and Survival in Chronic Lymphocytic Leukemia. J Clin Oncol [Internet]. 2010;28(29):4473–9. Available from: http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2009.27.8762

47. Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet [Internet]. 2012 Jan [cited 2012 Oct 6];44(1):47–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22158541

48. Austen B, Powell JE, Alvi A, Edwards I, Hooper L, Starczynski J, et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. Blood. 2005;106(9):3175–82.

49. Rossi D, Fangazio M, Rasi S, Vaisitti T, Monti S, Cresta S, et al. Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. Blood. 2012;119(12):2854–62.

50. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Nonlinear partial differential equations and applications: Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci [Internet]. 2002;99(24):15524–9. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.242606799

51. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, et al. The DLEU2/miR-15a/16-1 Cluster Controls B Cell Proliferation and Its Deletion Leads to Chronic Lymphocytic Leukemia. Cancer Cell [Internet]. Elsevier Ltd; 2010;17(1):28–40. Available from: http://dx.doi.org/10.1016/j.ccr.2009.11.019

52. Ten Hacken E, Burger J a. Microenvironment interactions and B-cell receptor signaling in Chronic Lymphocytic Leukemia: Implications for disease pathogenesis and treatment. Biochim Biophys Acta [Internet]. Elsevier B.V.; 2015;1863(3):401–13. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0167488915002451%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/26193078

53. Ramsay AD, Rodriguez-Justo M. Chronic lymphocytic leukaemia – the role of themicroenvironment pathogenesis and therapy. Br J Haematol. 2013;162:15–24.

54. Nowakowski GS, Hoyer JD, Shanafelt TD, Zent CS, Call TG, Bone ND, et al. Percentage of smudge cells on routine blood smear predicts survival in chronic lymphocytic leukemia. J Clin Oncol [Internet]. 2009 Apr 10 [cited 2014 Mar 15];27(11):1844–9. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2668708&tool=pmcentrez&rendertype=abstract

55. Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, Jaffe ES. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. Blood. 2011 May;117(19):5019–32.

56. Kalpadakis C, Pangalis G a, Sachanas S, Vassilakopoulos TP, Kyriakaki S, Korkolopoulou P, et al. New insights into monoclonal B-cell lymphocytosis. Biomed Res Int. Hindawi Publishing Corporation; 2014 Jan;2014:258917.

57. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature [Internet]. 2015; Available from: http://www.ncbi.nlm.nih.gov/pubmed/26200345

58. Rawstron AC, Shanafelt T, Lanasa MC, Landgren O, Hanson C, Orfao A, et al. Different biology and clinical outcome according to the absolute numbers of clonal B-cells in monoclonal B-cell lymphocytosis (MBL). Cytometry B Clin Cytom. 2010 Jan;78 Suppl 1(May):S19-23.

59. KR R. A critical analysis of staging in CLL. In: UCLA Symposium and Cellular Biology New Series. 1987. p. 253–64.

60. Muller P a J, Vousden KH. P53 Mutations in Cancer. Nat Cell Biol [Internet]. Nature Publishing Group; 2013 Jan [cited 2014 Mar 19];15(1):2–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23263379

61. Rose-Zerilli MJJ, Forster J, Parker H, Parker A, Rodri AÉ, Chaplin T, et al. ATM mutation rather than BIRC3 deletion and/or mutation predicts reduced survival in 11q-deleted chronic lymphocytic leukemia: Data from the UK LRF CLL4 trial. Haematologica. 2014;99:736–42.

62. Riches JC, O'Donovan CJ, Kingdon SJ, McClanahan F, Clear AJ, Neuberg DS, et al. Trisomy 12 chronic lymphocytic leukemia cells exhibit upregulation of integrin signaling that is modulated by NOTCH1 mutations. Blood. 2014;123(26):4101–10.

63. Molica S. A systematic review on Richter syndrome: what is the published evidence? Leuk

Lymphoma. 2010;51(March):415–21.

64. Mccann JJ, Lo TM, Webster D a, Lo M. Cross-linking of DNA by Alkylating Agents and Effects on DNA Function in the Chick Embryo. 1971;(NOVEMBER):1573–9.

65. Pettitt AR. Mechanism of action of purine analogues in chronic lymphocytic leukaemia. Br J Haematol. 2003;121:692–702.

66. Hallek M, Dc W. Signaling the end of chronic lymphocytic leukemia : new frontline treatment strategies. 2013;3723–34.

67. Catovsky D, Richards S, Matutes E, Oscier D, Dyer M, Bezares R, et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. Lancet. 2007;370:230–9.

68. Bosanquet A, Johnson S, Richards S. Prognosis for Fludarabine therapy of chronic lymphocytic leukaemia based on ex vivo drug response by DiSC assay. Br J Haematol. 1999;106:71–7.

69. Stilgenbauer S, Schnaiter A, Paschka P, Zenz T, Rossi M, Döhner K, et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: Results from the CLL8 trial. Blood. 2014;123:3247–54.

70. Edelmann J, Holzmann K, Miller F, Winkler D, Bühler A, Zenz T, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood [Internet]. 2012 Oct 9 [cited 2012 Oct 23];120(24):4783–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23047824

71. Warner JL, Arnason JE. Alemtuzumab use in relapsed and refractory chronic lymphocytic leukemia: a history and discussion of future rational use. Ther Adv Hematol [Internet]. 2012;3(6):375–89. Available from: http://journals.sagepub.com/doi/10.1177/2040620712458949

72. Burger JA, Tedeschi A, Barr PM, Robak T, Owen C, Ghia P, et al. Ibrutinib as Initial Therapy for Patients with Chronic Lymphocytic Leukemia. N Engl J Med [Internet]. 2015;373(25):2425–37. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1509388

73. Furman RR, Sharman JP, Coutre SE, Cheson BD, Pagel JM, Hillmen P, et al. Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia. N Engl J Med [Internet]. 2014;370(11):997–1007. Available from:

http://www.nejm.org/doi/10.1056/NEJMoa1315226

74. Stilgenbauer S, Eichhorst B, Schetelig J, Coutre S, Seymour JF, Munir T, et al. Venetoclax in relapsed or refractory chronic lymphocytic leukaemia with 17p deletion: a multicentre, open-label, phase 2 study. Lancet Oncol [Internet]. 2016;17(June):768–78. Available from: http://dx.doi.org/10.1016/ S1470-2045(16)30019-5

75. Landau D a, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell [Internet]. Elsevier Inc.; 2013 Feb 14 [cited 2014 Feb 20];152(4):714–26. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3575604&tool=pmcentrez&rendertype=abstract

76. Caldecott KW. DNA single-strand break repair and spinocerebellar ataxia. Cell. 2003;112:7–10.

77. Stracker TH, Petrini JHJ. The MRE11 complex: starting from the ends. Nat Rev Mol Cell Biol [Internet]. Nature Publishing Group; 2011;12(2):90–103. Available from: http://dx.doi.org/10.1038/nrm3047

78. Lavin MF, Kozlov S. ATM activation and DNA damage response. Cell Cycle. 2007;6(8):931–42.

79. Johnson RD, Jasin M. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. EMBO J. 2000;19(13):3398–407.

80. Jackson SP. Sensing and repairing DNA double-strand breaks. Carcinogenesis. 2002;23(5):687–96.

81. San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. Annu Rev Biochem. 2008;77:229–57.

82. Raa GD, Derks IAM, Navrkalova V, Skowronska A, Moerland PD, Laar J Van, et al. The impact of SF3B1 mutations in CLL on the DNA-damage response. 2015;(August 2014):1133–42.

83. Galluzzi L, Vitale I, Abrams JM, Alnemri ES, Baehrecke EH, Blagosklonny M V, et al. Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. Cell Death Differ [Internet]. 2012 Jan [cited 2013 Feb 27];19(1):107–20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3252826&tool=pmcentrez&r

endertype=abstract

84. Youle RJ, Strasser A. The BCL-2 protein family: opposing activities that mediate cell death. Nat Rev Mol Cell Biol. 2008;9(january):47–59.

85. Fulda S, Debatin K-M. Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy. Oncogene. 2006;25:4798–811.

86. Schimmer AD, Munk-pedersen I, Minden MD, Reed JC. Bcl-2 and Apoptosis in Chronic Lymphocytic Leukemia. Curr Treat Options Oncol. 2003;(4):211–8.

87. Faderl S, Keating MJ, Do K a, Liang SY, Kantarjian HM, O'Brien S, et al. Expression profile of 11 proteins and their prognostic significance in patients with chronic lymphocytic leukemia (CLL). Leukemia [Internet]. 2002;16(6):1045–52. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12040436

88. Smith a. J, Dai H, Correia C, Takahashi R, Lee S-H, Schmitz I, et al. Noxa/Bcl-2 Protein Interactions Contribute to Bortezomib Resistance in Human Lymphoid Cells. J Biol Chem [Internet]. 2011 Mar 22 [cited 2013 Mar 18];286(20):17682–92. Available from: http://www.jbc.org/cgi/doi/10.1074/jbc.M110.189092

89. Kipps TJ, Eradat H, Grosicki S, Catalano J, Cosolo W, Dyagil I, et al. A phase 2 study of the BH3 mimetic BCL2 inhibitor navitoclax (ABT-263) with or without rituximab, in previously untreated B-cell chronic lymphocytic leukemia. Leuk Lymphoma [Internet]. 2015;8194(November):1–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25797560

90. Billard C. BH3 Mimetics: Status of the Field and New Developments. Mol Cancer Ther [Internet]. 2013;12(9):1691–700. Available from: http://mct.aacrjournals.org/content/12/9/1691%5Cnhttp://mct.aacrjournals.org.gate1.inist.fr/content/12/9/1691.full.pdf%5Cnhttp://mct.aacrjournals.org.gate1.inist.fr/content/12/9/1691.full.pdf+html?frame=header%5Cnhttp://mct.aacrjournals.org.gate1.inist.fr/conte

91. Silva MCC, Bodor DL, Stellfox ME, Martins NMC, Hochegger H, Foltz DR, et al. Cdk Activity Couples Epigenetic Centromere Inheritance to Cell Cycle Progression. Dev Cell [Internet]. Elsevier Inc.; 2012;22(1):52–63. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1534580711004667

92. Tomlinson R, Ziegler T, Teerawit S, Terns R, Terns M. Cell Cycle-regulated Trafficking of

Human Telomerase to Telomeres. Mol Biol Cell. 2006;17(2):955–65.

93. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer. 2009;9(3):153–66.

94. Tsukada N, Burger J a., Zvaifler NJ, Kipps TJ. Distinctive features of “nurselike” cells that differentiate in the context of chronic lymphocytic leukemia. Blood. 2002;99(3):1030–7.

95. Bruno S, Ledda B, Tenca C, Ravera S, Maria A, Mazzarello AN, et al. Metformin inhibits cell cycle progression of B-cell chronic lymphocytic leukemia cells. 2015;6(26).

96. Landau DA, Tausch E, Taylor-weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations driving CLL and their evolution in progression and relapse. 2015;

97. Christian BA, Grever MR, Byrd JC, Lin TS. Flavopiridol in chronic lymphocytic leukemia: a concise review. Clin Lymphoma Myeloma [Internet]. Elsevier Inc.; 2009;9 Suppl 3(September):S179–85. Available from: http://dx.doi.org/10.3816/CLM.2009.s.009%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/19778838

98. Walsby E, Pratt G, Shao H, Abbas AY, Fischer PM, Bradshaw TD, et al. A novel Cdk9 inhibitor preferentially targets tumor cells and synergizes with fludarabine. Oncotarget [Internet]. 2014;5(2):375–85. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3964214&tool=pmcentrez&rendertype=abstract

99. Woyach JA, Johnson AJ, Byrd JC. The B-cell receptor signaling pathway as a therapeutic target in CLL. 2012;120(6):1175–84.

100. Fearon DT, Carroll MC. REGULATION OF B LYMPHOCYTE RESPONSES TO FOREIGN AND SELF-ANTIGENS BY THE CD19 / CD21 COMPLEX. Annu Rev Immunol. 2000;18:393–422.

101. Cheng PC, Dykstra ML, Mitchell RN, Pierce SK. A role for lipid rafts in B cell antigen receptor signaling and antigen targeting. J Exp Med [Internet]. 1999;190(11):1549–60. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2195743&tool=pmcentrez&rendertype=abstract

102. Buchner M, Fuchs S, Prinz G, Pfeifer D, Bartholomé K, Burger M, et al. Spleen tyrosine kinase is overexpressed and represents a potential therapeutic target in chronic lymphocytic leukemia. Cancer Res. 2009;69(13):5424–32.

103. Contri A, Brunati AM, Trentin L, Cabrelle A, Miorin M, Cesaro L, et al. Chronic lymphocytic leukemia B cells contain anomalous Lyn tyrosine kinase, a putative contribution to defective apoptosis. J Clin Invest. 2005;115(2):369–78.

104. Lannutti BJ, Meadows S a, Herman SEM, Kashishian A, Steiner B, Johnson AJ, et al. CAL-101, a p110delta selective phosphatidylinositol-3-kinase inhibitor for the treatment of B-cell malignancies, inhibits PI3K signaling and cellular viability. Blood [Internet]. 2011;117(2):591–4. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3694505&tool=pmcentrez&rendertype=abstract

105. Brown J. PCI-32765, the First BTK (Bruton's Tyrosine Kinase) Inhibitor in Clinical Trials. Curr Hematol Malig Rep [Internet]. 2013;8(1):1–6. Available from: http://link.springer.com/article/10.1007/s11899-012-0147-9

106. Friedberg JW, Sharman J, Sweetenham J, Johnston PB, Vose JM, LaCasce A, et al. Inhibition of Syk with fostamatinib disodium has significant clinical activity in non-Hodgkin lymphoma and chronic lymphocytic leukemia. Blood. 2010;115(13):2578–85.

107. Gilbert S. Developmental Biology 6th Edition. 2000. http://www.ncbi.nlm.nih.gov/books/NBK10072/.

108. Artavanis-Tsakonas S. Notch Signaling: Cell Fate Control and Signal Integration in Development. Science (80- ) [Internet]. 1999 Apr 30 [cited 2012 Nov 1];284(5415):770–6. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.284.5415.770

109. Duncan AW, Rattis FM, DiMascio LN, Congdon KL, Pazianos G, Zhao C, et al. Integration of Notch and Wnt signaling in hematopoietic stem cell maintenance. Nat Immunol [Internet]. 2005 Mar [cited 2012 Nov 27];6(3):314–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15665828

110. Calvi LM, Adams GB, Weibrecht KW, Weber JM, Olson DP, Knight MC, et al. Osteoblastic cells regulate the haematopoietic stem cell niche. 2003;841–6.

111. Maillard I, Koch U, Dumortier A, Shestova O, Xu L, Sai H, et al. Canonical notch signaling is dispensable for the maintenance of adult hematopoietic stem cells. Cell Stem Cell [Internet]. 2008 Apr 10 [cited 2012 Dec 7];2(4):356–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18397755

112. Bray SJ. Notch signalling : a simple pathway becomes complex. 2006;7(September):678–

89.

113. Martinez Arias A, Zecchini V, Brennan K. CSL-independent Notch signalling: a checkpoint in cell fate decisions during development? Curr Opin Genet Dev [Internet]. 2002 Oct;12(5):524–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12200157

114. Kopan R, Ilagan MXG. Review The Canonical Notch Signaling Pathway : Unfolding the Activation Mechanism. 2009;

115. Guruharsha KG, Kankel MW, Artavanis-tsakonas S. The Notch signalling system : recent insights into the complexity of a conserved pathway. Nat Publ Gr [Internet]. Nature Publishing Group; 2012;13(9):654–66. Available from: http://dx.doi.org/10.1038/nrg3272

116. Hubmann R, Schwarzmeier JD, Shehata M, Hilgarth M, Duechler M, Dettke M, et al. Notch2 is involved in the overexpression of CD23 in B-cell chronic lymphocytic leukemia. Blood [Internet]. 2002;99(10):3742–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11986231

117. O'Neil J, Grim J, Strack P, Rao S, Tibbitts D, Winter C, et al. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. J Exp Med. 2007;204(8):1813–24.

118. Balatti V, Bottoni A, Palamarchuk A, Alder H, Rassenti LZ, Kipps TJ, et al. NOTCH1 mutations in CLL associated with trisomy 12. Blood. 2012;119(2):329–31.

119. Rosati E, Sabatini R, Rampino G, Tabilio A, Di Ianni M, Fettucciari K, et al. Constitutively activated Notch signaling is involved in survival and apoptosis resistance of B-CLL cells. Blood. 2009;113(4):856–65.

120. Cadigan KM, Peifer M. Wnt signaling from development to disease: insights from model systems. Cold Spring Harb Perspect Biol [Internet]. 2009;1(2):a002881. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20066091%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC2742092/pdf/cshperspect-JCT-a002881.pdf

121. Yost C, Torres M, Miller JR, Huang E, Kimelman D, Moon RT. The axis-inducing activity, stability, and subcellular distribution of ??-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. Genes Dev. 1996;10(12):1443–54.

122. Behrens J, von Kries JP, Kühl M, Bruhn L, Wedlich D, Grosschedl R, et al. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature.

1996;382(6592):638–42.

123. Itoh K, Brott BK, Bae G-U, Ratcliffe MJ, Sokol SY. Nuclear localization is required for Dishevelled function in Wnt/beta-catenin signaling. J Biol [Internet]. 2005;4(1):3. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=551520&tool=pmcentrez&rendertype=abstract

124. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. Beta-Catenin Is a Target for the Ubiquitin Proteasome Pathway. {Embo} J. 1997;16(13):3797–3804.

125. Lu D, Zhao Y, Tawatao R, Cottam HB, Sen M, Leoni LM, et al. Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2004;101(9):3118–23.

126. Gandhirajan RK, Staib PA, Minke K, Gehrke I, Plickert G, Schlösser A, et al. Small molecule inhibitors of Wnt/beta-catenin/lef-1 signaling induces apoptosis in chronic lymphocytic leukemia cells in vitro and in vivo. Neoplasia. 2010;12(4):326–35.

127. Hayden MS, Ghosh S. NF-κB, the first quarter-century: remarkable progress and outstanding questions. Genes Dev. 2012 Feb 1;26(3):203–34.

128. Furman RR, Asgary Z, Mascarenhas JO, Liou H-C, Schattner EJ. Modulation of NF- B Activity and Apoptosis in Chronic Lymphocytic Leukemia B Cells. J Immunol [Internet]. 2000;164(4):2200–6. Available from: http://www.jimmunol.org/content/164/4/2200.full

129. Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. Vol. 136, Cell. 2009. p. 701–18.

130. Shcherbakova I, Hoskins A a, Friedman LJ, Serebrov V, Corrêa IR, Xu M-Q, et al. Alternative spliceosome assembly pathways revealed by single-molecule fluorescence microscopy. Cell Rep [Internet]. 2013 Oct 17 [cited 2014 Feb 21];5(1):151–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24075986

131. Will CL. Identification of Both Shared and Distinct Proteins in the Major and Minor Spliceosomes. Science (80- ) [Internet]. 1999 Jun 18 [cited 2013 Jan 28];284(5422):2003–5. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.284.5422.2003

132. Turunen JJ, Niemelä EH, Verma B, Frilander MJ. The significant other: splicing by the minor spliceosome. Wiley Interdiscip Rev RNA. 2013;4(1):61–76.

133. Cass DM, Berglund JA. The SF3b155 N-Terminal Domain Is a Scaffold Important for Splicing †. 2006;(24):10092–101.

134. Moldón A, Query C. Crossing the exon. Mol Cell [Internet]. 2010 Apr 23 [cited 2014 Mar 15];38(2):159–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20417595

135. Schneider M, Will CL, Anokhina M, Tazi J, Urlaub H, Lührmann R. Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes. Mol Cell [Internet]. 2010 May 23 [cited 2014 Jan 31];38(2):223–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20417601

136. Gozani O, Potashkin J, Reed R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol Cell Biol [Internet]. 1998 Aug;18(8):4752–60. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=109061&tool=pmcentrez&rendertype=abstract

137. Schellenberg MJ, Dul EL, Macmillan AM. Structural model of the p14/SF3b155 branch duplex complex. 2011;155–65.

138. Seghezzi W, Chua K, Shanahan F, Gozani O, Reed R, Lees E. Cyclin E associates with components of the pre-mRNA splicing machinery in mammalian cells. Mol Cell Biol [Internet]. 1998 Aug;18(8):4526–36. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=109038&tool=pmcentrez&rendertype=abstract

139. Boudrez A, Beullens M, Waelkens E, Stalmans W, Bollen M. Phosphorylation-dependent interaction between the splicing factors SAP155 and NIPP1. J Biol Chem [Internet]. 2002 Aug 30 [cited 2014 Mar 15];277(35):31834–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12105215

140. Golas MM, Sander B, Will CL. Molecular Architecture of the Multiprotein Splicing Factor SF3b. 2003;980(2003):10–5.

141. Wang C, Chua K, Seghezzi W, Lees E, Gozani O, Reed R. Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. Genes Dev [Internet]. 1998 May 15;12(10):1409–14. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=316838&tool=pmcentrez&rendertype=abstract

142. Strefford JC, Sutton L, Baliakas P, Agathangelidis a, Malčíková J, Plevova K, et al. Distinct patterns of novel gene mutations in poor-prognostic stereotyped subsets of chronic lymphocytic leukemia: the case of SF3B1 and subset #2. Leukemia [Internet]. 2013;27(April):2196–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23558524

143. Gentien D, Kosmider O, Nguyen-Khac F, Albaud B, Rapinat a, Dumont a G, et al. A common alternative splicing signature is associated with SF3B1 mutations in malignancies from different cell lineages. Leukemia [Internet]. Nature Publishing Group; 2014 Jan 17 [cited 2014 Feb 3];(January):1–2. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24434863

144. Furney SJ, Pedersen M, Gentien D, Dumont AG, Rapinat A, Desjardins L, et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. Cancer Discov [Internet]. 2013 Oct [cited 2014 Mar 25];3(10):1122–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23861464

145. Larrayoz M, Blakemore SJ, Dobson RC, Blunt MD, Rose-Zerilli MJJ, Walewska R, et al. The SF3B1 inhibitor spliceostatin A (SSA) elicits apoptosis in Chronic Lymphocytic Leukemia cells through downregulation of Mcl-1. Leukemia [Internet]. Nature Publishing Group; 2015;(April):1–28. Available from: http://www.nature.com/doifinder/10.1038/leu.2015.286

146. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet [Internet]. Nature Publishing Group; 2011;12(1):7–18. Available from: http://www.nature.com/doifinder/10.1038/nrg2905

147. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res [Internet]. Nature Publishing Group; 2011;21(3):381–95. Available from: http://www.nature.com/doifinder/10.1038/cr.2011.22

148. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature [Internet]. Nature Publishing Group; 2012;481(7380):157–63. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22237106&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1038/nature10725

149. Grada A, Weinbrecht K. Next-Generation Sequencing: Methodology and Application. J Invest Dermatol [Internet]. Nature Publishing Group; 2013;133(8):e11-4. Available from: http://dx.doi.org/10.1038/jid.2013.248%5Cnpapers3://publication/doi/10.1038/jid.2013.2

150. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol. 2008;26(10):1135–45.

151. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet. 2008;9:387–402.

152. Shapland EB, Holmes V, Reeves C, Sorokin E, Durot M, Platt D, et al. Low-cost, high-throughput sequencing of DNA assemblies using a highly multiplexed Nextera process. ACS Synth Biol [Internet]. 2015;150425205656002. Available from: http://pubs.acs.org/doi/abs/10.1021/sb500362n

153. Ronaghi M. Pyrosequencing Sheds Light on DNA Sequencing Pyrosequencing Sheds Light on DNA Sequencing. 2001;(650):3–11.

154. Berglund EC, Lindqvist CM, Hayat S, Övernäs E, Henriksson N, Nordlund J, et al. Accurate detection of subclonal single nucleotide variants in whole genome amplified and pooled cancer samples using HaloPlex target enrichment. BMC Genomics [Internet]. 2013;14:856. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4046713&tool=pmcentrez&rendertype=abstract

155. White TB, McCoy AM, Streva VA, Fenrich J, Deininger PL. A droplet digital PCR detection method for rare L1 insertions in tumors. Mob DNA. 2014;5(1):30.

156. Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2.

157. Kaspers GJL, Veerman AJP, Pieters R, Zantwijk I Van, Wering ER Van. DRUG COMBINATION TESTING IN ACUTE LYMPHOBLASTIC LEUKEMIA USING THE MTT ASSAY. Leuk Res. 1995;19(3):175–81.

158. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet [Internet]. Nature Publishing Group; 2016;17(6):333–51. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27184599

159. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1 and ATM mutations in chronic lymphocytic leukemia. Blood [Internet]. 2016; Available from: http://www.ncbi.nlm.nih.gov/pubmed/26837699

160. Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, Timbs AT, et al. SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. Blood. 2014;123(7):1021–31.

161. Jethwa A, Hüllein J, Stolz T, Blume C, Sellner L, Jauch A, et al. Targeted resequencing for analysis of clonal composition of recurrent gene mutations in chronic lymphocytic leukaemia. Br J Haematol. 2013;163(4):496–500.

162. Herling CD, Klaum??nzer M, Rocha CK, Altm??ller J, Thiele H, Bahlo J, et al. Complex karyotypes and KRAS and POT1 mutations impact outcome in CLL after chlorambucil-based chemotherapy or chemoimmunotherapy. Blood. 2016;128(3):395–404.

163. Guièze R, Robbe P, Clifford R, De Guibert S, Pereira B, Timbs A, et al. Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL. Blood. 2015;126(18):2110–7.

164. Maura F, Mosca L, Fabris S, Cutrona G, Matis S, Lionetti M, et al. Insulin growth factor 1 receptor expression is associated with NOTCH1 mutation, trisomy 12 and aggressive clinical course in chronic lymphocytic leukaemia. PLoS One. 2015;10(3):1–12.

165. Lionetti M, Fabris S, Cutrona G, Agnelli L, Ciardullo C, Matis S, et al. High-throughput sequencing for the identification of NOTCH1 mutations in early stage chronic lymphocytic leukaemia: Biological and clinical implications. Br J Haematol. 2014;165(5):629–39.

166. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood [Internet]. 2008 Jun 15 [cited 2012 Jul 22];111(12):5446–56. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2972576&tool=pmcentrez&rendertype=abstract

167. Gonzalez D, Martinez P, Wade R, Hockley S, Oscier D, Matutes E, et al. Mutational Status of the TP53 Gene As a Predictor of Response and Survival in Patients With Chronic Lymphocytic Leukemia: Results From the LRF CLL4 Trial. J Clin Oncol [Internet]. 2011;29(16):2223–9. Available from: http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2010.32.0838

168. Chen YC, Liu T, Yu CH, Chiang TY, Hwang CC. Effects of GC Bias in Next-Generation-Sequencing Data on De Novo Genome Assembly. PLoS One. 2013;8(4).

169. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of Next-Generation Sequencing Systems. J Biomed Biotechnol [Internet]. 2012;2012:1–11. Available from: http://www.hindawi.com/journals/jbb/2012/251364/

170. Smith EN, Jepsen K, Khosroheidari M, Rassenti LZ, D'Antonio M, Ghia EM, et al. Biased estimates of clonal evolution and subclonal heterogeneity can arise from PCR duplicates in deep sequencing experiments. Genome Biol [Internet]. 2014;15(8):420. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25103687%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4165357

171. Lawrence MS, Stojanov P, Polak P, Kryukov G V, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature [Internet]. 2013;499(7457):214–8. Available from: http://dx.doi.org/10.1038/nature12213

172. Wu X, Tschumper RC, Jelinek DF. Genetic characterization of SF3B1 mutations in single chronic lymphocytic leukemia cells. Leukemia [Internet]. 2013;(May):1–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23685408%5Cnhttp://dx.doi.org/10.1038/leu.2013.155

173. Schwaederlé M, Ghia E, Rassenti LZ, Obara M, Dell' Aquila ML, Fecteau JF, et al. Subclonal evolution involving SF3B1 mutations in chronic lymphocytic leukemia. Leuk Off J Leuk Soc Am Leuk Res Fund, UK [Internet]. 2013 Jan 22 [cited 2013 Apr 18];(January):1–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23337928

174. Wang L, Brooks AN, Fan J, Wan Y, Gambe R, Li S, et al. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. Cancer Cell [Internet]. 2016;30:1–14. Available from: http://dx.doi.org/10.1016/j.ccell.2016.10.005

175. Ramsay AJ, Quesada V, Foronda M, Conde L, Martínez-Trillos A, Villamor N, et al. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. Nat Genet [Internet]. Nature Publishing Group; 2013;45(5):526–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23502782

176. Gunay-Aygun M, Falik-Zaccai TC, Vilboux T, Zivony-Elboum Y, Gumruk F, Cetin M, et al. NBEAL2 is mutated in gray platelet syndrome and is required for biogenesis of platelet alpha-granules. Nat Genet [Internet]. Nature Publishing Group; 2011;43(8):732–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21765412

177. Cullinane AR, Schäffer AA, Huizing M. The BEACH is hot: A LYST of emerging roles for

BEACH-domain containing proteins in human disease. Traffic. 2013;14(7).

178. Bergman LM, Morris L, Darley M, Mirnezami AH, Gunatilake SC, Blaydes JP. Role of the unique N-terminal domain of CtBP2 in determining the subcellular localisation of CtBP family proteins. BMC Cell Biol [Internet]. 2006;7:35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16999872

179. Piatigorsky J. Dual use of the transcriptional repressor (CtBP2)/ribbon synapse (RIBEYE) gene: How prevalent are multifunctional genes? Trends Neurosci. 2001;24(10):555–7.

180. Cremolini C, Di Maio M, Petrelli F, Berenato R, Loupakis F, Pietrantonio F. BRAF-mutated metastatic colorectal cancer between past and future. Br J Cancer [Internet]. Nature Publishing Group; 2015;33(11):1–2. Available from: http://www.nature.com/doifinder/10.1038/bjc.2015.316

181. Ascierto P a, Kirkwood JM, Grob J-J, Simeone E, Grimaldi AM, Maio M, et al. The role of BRAF V600 mutation in melanoma. J Transl Med. 2012;10(1):85.

182. Pekova S, Mazal O, Cmejla R, Hardekopf DW, Plachy R, Zejskova L, et al. A comprehensive study of TP53 mutations in chronic lymphocytic leukemia: Analysis of 1287 diagnostic and 1148 follow-up CLL samples. Leuk Res [Internet]. Elsevier Ltd; 2011;35(7):889–98. Available from: http://dx.doi.org/10.1016/j.leukres.2010.12.016

183. Simen BB, Yin L, Goswami CP, Davis KO, Bajaj R, Gong JZ, et al. Validation of a next-generation-sequencing cancer panel for use in the clinical laboratory. Arch Pathol Lab Med. 2015;139(4):508–17.

184. Smith T, Heger A, Sudbery I. UMI-Tools: Modelling sequencing errors in Unique Molecular Identifiers to improve quantitation accuracy. Genome Res [Internet]. 2017; Available from: http://genome.cshlp.org/content/early/2017/01/18/gr.209601.116.abstract

185. Collu GM, Hidalgo-Sastre A, Acar A, Bayston L, Gildea C, Leverentz MK, et al. Dishevelled limits Notch signalling through inhibition of CSL. Development [Internet]. 2012;139(23):4405–15. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3509734&tool=pmcentrez&rendertype=abstract

186. Parry M, Rose-zerilli MJJ, Gibson J, Ennis S, Walewska R, Parker H, et al. Whole Exome Sequencing Identifies Novel Recurrently Mutated Genes in Patients with Splenic Marginal Zone Lymphoma. 2013;8(12):1–5.

187. Bewick V, Cheek L, Ball J. Statistics review 12: survival analysis. Crit Care [Internet]. 2004;8(5):389–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15469602%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1065034

188. Dietrich S, Floegel A, Troll M, Kühn T, Rathmann W, Peters A, et al. Random Survival Forest in practice: a method for modelling complex metabolomics data in time to event analysis. Int J Epidemiol [Internet]. 2016;(September):dyw145. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27591264

189. Ishwaran H, Gerds TA, Kogalur UB, Moore RD, Gange SJ, Lau BM. Random survival forests for competing risks. Biostatistics. 2014;15(4):757–73.

190. Larrayoz M, Rose-Zerilli MJJ, Kadalayil L, Parker H, Blakemore S, Forster J, et al. Non-coding NOTCH1 mutations in chronic lymphocytic leukemia; their clinical impact in the UK CLL4 trial. Leukemia [Internet]. 2016;1(November 2016):1–4. Available from: http://www.nature.com/doifinder/10.1038/leu.2016.298

191. D. Schoenfeld. Partial Residuals for The Proportionnal Hazards Regression Model. Biometrika [Internet]. 1982;69(1):239–41. Available from: http://www.jstor.org/stable/2335876

192. Bourne HR, Sanders D a, McCormick F. The GTPase superfamily: a conserved switch for diverse cell functions. Nature. 1990;348:125–32.

193. Weber CK, Slupsky JR, Andreas Kalmes H, Rapp UR. Active ras induces heterodimerization of cRaf and BRaf. Cancer Res. 2001;61:3595–8.

194. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the BRAF gene in human cancer. Nature. 2002;417:949–54.

195. Pandzic T, Larsson J, He L, Kundu S, Ban K, Akhtar-Ali M, et al. Transposon Mutagenesis Reveals Fludarabine Resistance Mechanisms in Chronic Lymphocytic Leukemia. Clin Cancer Res [Internet]. 2016;22(24):6217–27. Available from: http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-15-2903

196. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: A retrospective consortium analysis. Lancet Oncol. 2010;11(8):753–62.

197. Dielschneider RF, Xiao W, Yoon J-Y, Noh E, Banerji V, Li H, et al. Gefitinib targets ZAP-70-expressing chronic lymphocytic leukemia cells and inhibits B-cell receptor signaling. Cell Death Dis [Internet]. 2014;5:e1439. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4649506&tool=pmcentrez&rendertype=abstract

198. Foà R, Del Giudice I, Cuneo A, Del Poeta G, Ciolli S, Di Raimondo F, et al. Chlorambucil plus rituximab with or without maintenance rituximab as first-line treatment for elderly chronic lymphocytic leukemia patients. Am J Hematol. 2014;89(5):480–6.

199. Li S, Symonds ALJ, Zhu B, Liu M, Raymond M V., Miao T, et al. Early growth response gene-2 (egr-2) regulates the development of b and t cells. PLoS One. 2011;6(4):1–9.

200. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee HJ, Sciammas R, et al. Multilineage Transcriptional Priming and Determination of Alternate Hematopoietic Cell Fates. Cell. 2006;126:755–66.

201. Unoki M, Nakamura Y. EGR2 induces apoptosis in various cancer cell lines by direct transactivation of BNIP3L and BAK. Oncogene. 2003;22:2172–85.

202. Matsushima-nishiu M, Unoki M, Ono K, Pten EE. Growth and Gene Expression Profile Analyses of Endometrial Cancer Cells Expressing Exogenous PTEN Growth and Gene Expression Profile Analyses of Endometrial Cancer Cells. 2001;3741–9.

203. Garbelli A, Beermann S, Di Cicco G, Dietrich U, Maga G. A motif unique to the human dead-box protein DDX3 is important for nucleic acid binding, ATP hydrolysis, RNA/DNA unwinding and HIV-1 replication. PLoS One. 2011;6(5).

204. Ojha J, Ayres J, Secreto C, Tschumper R, Rabe K, Dyke D Van, et al. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia Deep Sequencing Identifies Genetic Heterogeneity and Recurrent Convergent Evolution in Chronic Lymphocytic Leukemia. 2014;

205. Rossi D, Rasi S, Fabbri G, Spina V, Fangazio M, Forconi F, et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. 2016;119(2):521–30.

206. Mart A, Pinyol M, Navarro A, Aymerich M, Jares P, Juan M, et al. Mutations in TLR / MYD88 pathway identify a subset of young chronic lymphocytic leukemia patients with favorable outcome. Blood. 2014;123(June):3790–6.

207. Baliakas P, Hadzidimitriou A, Agathangelidis A, Rossi D, Sutton LA, Kminkova J, et al.

Prognostic relevance of MYD88 mutations in CLL: The jury is still out. Blood. 2015;126(8):1043–4.

208. Rossi D, Rasi S, Spina V, Bruscaggin A, Monti S, Ciardullo C, et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. 2013;

209. Dal Bo M, Bulian P, Bomben R, Zucchetto A, Rossi F, Pozzo F, et al. CD49d prevails over the novel recurrent mutations as independent prognosticator of overall survival in chronic lymphocytic leukemia. Leukemia. 2016;20th May [(November 2015):2011–8.

210. Ehrlinger J. ggRandomForests: Exploring Random Forest Survival. 2016; Available from: http://arxiv.org/abs/1612.08974

211. Wan Y, Wu CJ. SF3B1 mutations in chronic lymphocytic leukemia. Blood [Internet]. 2013 Apr 8 [cited 2013 Apr 10]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/23568491

212. Bonnal S, Vigevani L, Valcárcel J. The spliceosome as a target of novel antitumour drugs. Nat Rev Drug Discov [Internet]. 2012 Nov [cited 2013 Feb 27];11(11):847–59. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23123942

213. Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, et al. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. 2007;3(9):5–9.

214. Albert BJ, Mcpherson PA, Brien KO, Czaicki NL, Osman S, Li M, et al. Meayamycin inhibits pre-mRNA splicing and exhibits picomolar activity against multidrug resistant cells. Mol Cancer Ther. 2010;8(8):2308–18.

215. Fan L, Lagisetti C, Edwards CC, Webb TR, Potter PM. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. ACS Chem Biol. 2011;6(6):582–9.

216. Folco EG, Coil KE, Reed R. The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. 2011;(617):440–4.

217. Muraki M, Ohkawara B, Hosoya T, Onogi H, Koizumi J, Koizumi T, et al. Manipulation of alternative splicing by a newly developed inhibitor of Clks. J Biol Chem. 2004;279(23):24246–54.

218. Hasegawa M, Miura T, Kuzuya K, Inoue A, Ki SW, Horinouchi S, et al. Identification of SAP155 as the target of GEX1A (Herboxidiene), an antitumor natural product. ACS Chem

Biol. 2011;6(3):229–33.

219. Sakai Y, Yoshida T, Ochiai K, Uosaki Y, Saitoh Y, Tanaka F, et al. GEX1 compounds, novel antitumor antibiotics related to herboxidiene, produced by Streptomyces sp. I. Taxonomy, production, isolation, physicochemical properties and biological activities. J Antibiot (Tokyo). 2002;55(10):855–62.

220. Xargay-torrent S, López-guerra M, Rosich L, Montraveta A, Roldán J, Rodríguez V, et al. The splicing modulator sudemycin induces a specific antitumor response and cooperates with ibrutinib in chronic lymphocytic leukemia. 2015;

221. Eskens F a LM, Ramos FJ, Burger H, O'Brien JP, Piera A, de Jonge MJ a, et al. Phase I pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107 in patients with advanced solid tumors. Clin Cancer Res [Internet]. 2013;19(22):6296–304. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23983259

222. Nakajima H, Sato B, Fujita T, Takase S, Terano H, Okuhara M. New antitumor substances, FR901463, FR901464 and FR901465. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities. J Antibiot (Tokyo). 1996;49(2663):1196–203.

223. Gao Y, Koide K. Chemical perturbation of Mcl-1 pre-mRNA splicing to induce apoptosis in cancer cells. ACS Chem Biol. 2013;8(5):895–900.

224. Laetsch TW, Liu X, Vu a, Sliozberg M, Vido M, Elci OU, et al. Multiple components of the spliceosome regulate Mcl1 activity in neuroblastoma. Cell Death Dis [Internet]. 2014;5(2014):e1072. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3944256&tool=pmcentrez&rendertype=abstract

225. Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M, et al. Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol [Internet]. 2007 Sep [cited 2013 Mar 8];3(9):570–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17643112

226. Moore MJ, Wang Q, Kennedy CJ, Silver P a. An alternative splicing network links cell-cycle control to apoptosis. Cell [Internet]. Elsevier Ltd; 2010 Aug 20 [cited 2014 Feb 28];142(4):625–36. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2924962&tool=pmcentrez&rendertype=abstract

227. Turner P V, Pekow C, Vasbinder MA, Brabb T. Administration of substances to laboratory animals: equipment considerations, vehicle selection, and solute preparation. J Am Assoc Lab Anim Sci [Internet]. 2011;50(5):614–27. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3189663&tool=pmcentrez&rendertype=abstract

228. Blunt MD, Koehrer S, Dobson R, Larrayoz M, Wilmore S, Hayman A, et al. The dual Syk/JAK inhibitor cerdulatinib antagonises B-cell receptor and microenvironmental signaling in chronic lymphocytic leukemia. 2016;(13):1–13. Available from: http://europepmc.org/abstract/med/27697994

229. Roberts AW, Davids MS, Pagel JM, Kahl BS, Puvvada SD, Gerecitano JF, et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. N Engl J Med [Internet]. 2016;374(4):311–22. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa1513257

230. Thijssen R, Slinger E, Weller K, Geest CR, Beaumont T, Van Oers MHJ, et al. Resistance to ABT-199 induced by microenvironmental signals in chronic lymphocytic leukemia can be counteracted by CD20 antibodies or kinase inhibitors. Haematologica. 2015;100(8).

231. Billard C. Design of novel BH3 mimetics for the treatment of chronic lymphocytic leukemia. 2012;872(February):2032–8.

232. Brumatti G, Ekert PG. Seeking a MCL-1 inhibitor. Cell Death Differ [Internet]. Nature Publishing Group; 2013;20(11):1440–1. Available from: http://dx.doi.org/10.1038/cdd.2013.114%5Cn10.1038/cdd.2013.114

233. Smit LA, Hallaert DYH, Goeij B De, Jaspers A, Kater AP. Differential Noxa / Mcl-1 balance in peripheral versus lymph node chronic lymphocytic leukemia cells correlates with survival capacity. 2016;109(4):1660–9.

234. Varadarajan S, Vogler M, Butterworth M, Dinsdale D, Walensky LD, Cohen GM. Evaluation and critical assessment of putative MCL-1 inhibitors. Cell Death Differ [Internet]. Nature Publishing Group; 2013;20(11):1475–84. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3792441&tool=pmcentrez&rendertype=abstract

235. Besbes S, Pocard M, Mirshahi M, Billard C. Critical Reviews in Oncology / Hematology The first MCL - 1-selective BH3 mimetics have therapeutic potential for chronic lymphocytic leukemia. 2016;100:32–6.

236. Kotschy A, Szlavik Z, Murray J, Davidson J, Maragno AL, Le Toumelin-Braizat G, et al. The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. Nature [Internet]. Nature Publishing Group; 2016;538(7626):477–82. Available from: http://www.nature.com/doifinder/10.1038/nature19830%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27760111

237. Abulwerdi F, Liao C, Liu M, Azmi AS, Aboukameel A, Sun D, et al. A Novel Small-Molecule Inhibitor of Mcl-1 Blocks Pancreatic Cancer Growth In Vitro and In Vivo. 2014;1(13):565–76.

238. Gao Y, Trivedi S, Ferris RL, Koide K. Regulation of HPV16 E6 and MCL1 by SF3B1 inhibitor i1. Gao, Y., Trivedi, S., Ferris, R. L. & Koide, K. Regulation of HPV16 E6 and MCL1 by SF3B1 inhibitor in head and neck cancer cells. Sci. Rep. 4, 6098 (2014).n head and neck cancer cells. Sci Rep [Internet]. 2014;4:6098. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4138523&tool=pmcentrez&rendertype=abstract

239. Sharma a, Singh K, Mazumder S, Hill BT, Kalaycio M, Almasan a. BECN1 and BIM interactions with MCL-1 determine fludarabine resistance in leukemic B cells. Cell Death Dis [Internet]. Nature Publishing Group; 2013;4(5):e628. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3674362&tool=pmcentrez&rendertype=abstract

240. White MJ, Mcarthur K, Metcalf D, Lane RM, Cambier JC, Herold MJ, et al. Apoptotic Caspases Suppress Type I IFN Production. Cell [Internet]. Elsevier Inc.; 2014;159(7):1549–62. Available from: http://dx.doi.org/10.1016/j.cell.2014.11.036

241. Matsushita K, Tomonaga T, Shimada H, Shioya A, Higashi M, Matsubara H, et al. An essential role of alternative splicing of c-myc suppressor FUSE-binding protein-interacting repressor in carcinogenesis. Cancer Res. 2006;66(3):1409–17.

242. Matsushita K, Kajiwara T, Tamura M, Satoh M, Tanaka N, Tomonaga T, et al. SAP155-mediated splicing of FUSE-binding protein-interacting repressor serves as a molecular switch for c-myc gene expression. Mol Cancer Res [Internet]. 2012 Jun [cited 2014 Mar 15];10(6):787–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22496461

243. Matsushita K, Tamura M, Tanaka N, Tomonaga T, Matsubara H, Shimada H, et al. Interactions between SAP155 and FUSE-binding protein-interacting repressor bridges c-Myc and P27Kip1 expression. Mol Cancer Res [Internet]. 2013 Apr 17 [cited 2013 Apr 25];

Available from: http://www.ncbi.nlm.nih.gov/pubmed/23594796

244. Scandurra M, Rossi D, Deambrogi C, Rancoita PM V, Chigrinova E, Mian M, et al. Genomic Profiling of Richter's Syndrome: Recurrent Lesions and Differences with de novo Diffuse Large B-Cell Lymphomas. Hematol Oncol. 2009;28:62–7.

245. Chigrinova E, Rinaldi A, Kwee I, Rossi D, Rancoita PM V, Jonathan C, et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. 2014;122(15):2673–82.

246. Scarfò L, Ghia P. Reprogramming cell death: BCL2 family inhibition in hematological malignancies. Immunol Lett [Internet]. Elsevier B.V.; 2013;155(1–2):36–9. Available from: http://dx.doi.org/10.1016/j.imlet.2013.09.015

247. Liu J, Chen G, Feng L, Zhang W, Pelicano H, Wang F, et al. Loss of p53 and altered miR15-a/16-1 MCL-1 pathway in CLL: insights from TCL1-Tg:p53(-/-) mouse model and primary human leukemia cells. Leukemia [Internet]. 2014 Jan [cited 2014 Mar 1];28(1):118–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23608884

248. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger J a, Blum K a, et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. N Engl J Med [Internet]. 2013;369:32–42. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3772525&tool=pmcentrez&rendertype=abstract

249. Kashyap MK, Kumar D, Villa R, La Clair JJ, Benner C, Sasik R, et al. Targeting the spliceosome in chronic lymphocytic leukemia with the macrolides FD-895 and pladienolide B. Haematologica [Internet]. 2015; Available from: http://www.haematologica.org/cgi/doi/10.3324/haematol.2014.122069

250. Pei B, Sisu C, Frankish A, Howald C, Habegger L, Mu XJ, et al. The GENCODE pseudogene resource. Genome Biol. 2012;13(ePub ahead of print).

251. Oncology AS of C. Major Milestones Against Cancer [Internet]. Cancer Progress. 2015. Available from: http://www.cancerprogress.net/timeline/major-milestones-against-cancer

252. American Cancer Society. 100 Years of Saving Lives [Internet]. cancer.org. 2015. Available from: http://www.cancer.org/fight/timeline

253. Rickert RC. New insights into pre-BCR and BCR signalling with relevance to B cell

malignancies. Nat Rev Immunol [Internet]. 2013;13(8):578–91. Available from: http://www.nature.com/doifinder/10.1038/nri3487

254. Fabbri G, Khiabanian H, Holmes AB, Wang J, Messina M, Mullighan CG, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. J Exp Med [Internet]. 2013;210(11):2273–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24127483

255. Messina M, Giudice I Del, Khiabanian H, Rossi D, Chiaretti S, Rasi S, et al. Genetic lesions associated with chronic lymphocytic leukemia. Blood. 2014;123(15):2378–88.

256. Goldin LR, McMaster ML, Rotunno M, Herman SEM, Jones K, Zhu B, et al. Whole exome sequencing in families with CLL detects a variant in Integrin Beta 2 associated with disease susceptibility. Blood [Internet]. 2016; Available from: http://www.bloodjournal.org/cgi/doi/10.1182/blood-2016-02-697771

257. Speedy HE, Kinnersley B, Chubb D, Broderick P, Law PJ, Litchfield K, et al. Germline mutations in shelterin complex genes are associated with familial chronic lymphocytic leukemia. Blood [Internet]. 2016;128(19):blood-2016-01-695692. Available from: http://www.bloodjournal.org/content/early/2016/08/15/blood-2016-01-695692.abstract

258. Rose-Zerilli MJJ, Gibson J, Wang J, Tapper W, Davis Z, Parker H, et al. Longitudinal copy number, whole exome and targeted deep sequencing of “good risk” IGHV-mutated CLL patients with progressive disease. Leukemia [Internet]. Nature Publishing Group; 2016;30(February):1–10. Available from: http://www.nature.com/doifinder/10.1038/leu.2016.10

259. Rajasagi M, Shukla SA, Fritsch EF, Keskin DB, Deluca D, Carmona E, et al. Systematic identi fi cation of personal tumor-speci fi c neoantigens in chronic lymphocytic leukemia. 2016;124(3):453–63.

260. Woyach J a, Furman RR, Liu T-M, Ozer HG, Zapatka M, Ruppert AS, et al. Resistance mechanisms for the Bruton’s tyrosine kinase inhibitor ibrutinib. New Eng J Med [Internet]. 2014;370(24):2286–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24869598

261. Rodríguez D, Ramsay AJ, Quesada V, Garabaya C, Campo E, Freije JMP, et al. Functional analysis of sucrase-isomaltase mutations from chronic lymphocytic leukemia patients. Hum Mol Genet. 2013;22(11):2273–82.

262. Dehm SM. Test-firing ammunition for spliceosome inhibition in cancer. Clin Cancer Res.

2013;19(22).