10/02/15

Genetics and Prognostication in Splenic Marginal Zone Lymphoma: Revelations from Deep Sequencing

Running Title: The clinical significance of gene mutations in SMZL

Marina Parry[1,*], Matthew JJ Rose-Zerilli[1,*], Viktor Ljungström[2,*], Jane Gibson[3], Jun Wang [4], Renata Walewska[5], Helen Parker[1], Anton Parker[5], Zadie Davis[5], Anne Gardiner[5], Neil McIver-Brown[5], Christina Kalpadakis[6], Aliki Xochelli[7], Achilles Anagnostopoulos[8], Claudia Fazi[9], David Gonzalez de Castro[10], Claire Dearden[10], Guy Pratt[11], Richard Rosenquist[2], Margaret Ashton-Key[1], Francesco Forconi[1], Andrew Collins[12], Paolo Ghia[9], Estella Matutes[13], Gerassimos Pangalis[14], Kostas Stamatopoulos[7,8], David Oscier[1,5], Jonathan C Strefford[1]

[1]Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK,[2]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Sweden, [3]Centre for Biological Sciences, University of Southampton, Southampton, UK, [4] Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, London, UK. [5]Department of Pathology, Royal Bournemouth Hospital, Bournemouth, UK, [6]Department of Hematology, School of Medicine, University of Crete, Heraklion, Greece, [7]Institute of Applied Biosciences, Center for Research and Technology, Thessaloniki, Greece, [8]Hematology Department and HCT Unit, G. Papanicolaou Hospital, Thessaloniki, Greece, [9]Division of Molecular Oncology, Department of Onco-Haematology, IRCCS Istituto Scientifico San Raffaele, Fondazione Centro San Raffaele, Università Vita-Salute San Raffaele, Milan, Italy, [10]Heamato-oncology Unit, Division of Molecular Pathology, Institute for Cancer Research, Sutton, UK, [11]School of Cancer Studies, University of Birmingham, Birmingham, UK; Department of Haematology, Heart of England NHS Foundation Trust, Birmingham, UK, [12]Genetic Epidemiology and Bioinformatics, Faculty of Medicine, University of Southampton, UK,

[13]Haematopathology Unit, Hospital Clinic, Barcelona University, Villarroel, Barcelona, Spain, [14]Department of Hematology, Athens Medical Center, Athens, Greece.

* These authors contributed equally to this manuscript

Correspondence to: Prof Jonathan C Strefford, Cancer Genomics, Cancer Sciences, Somers Cancer Research Building, Southampton General Hospital, Tremona Road, Southampton SO16 6YD. Tel: 44 23 8079 5246. E-mail: JCS@soton.ac.uk

Text word count: 3777

Abstract word count: 186

Number figures/tables: 3/3

Number references: 39

Statement of Translational Relevance

This multinational study identifies genomic and immunogenetic factors with prognostic significance in splenic marginal zone lymphoma (SMZL), a rare B cell non Hodgkins lymphoma. Genomic mutations in *TP53*, *KLF2*, *NOTCH2* and *TNFAIP3* were found collectively in over 40% of cases and 13% utilised IGHV genes with no somatic hypermutation (SHM). *TNFAIP3* mutations were associated with an increased risk of high grade transformation. IGHV genes lacking SHM, *KLF2* and *NOTCH2* mutations were associated with shorter time to first treatment, while *TP53* and *MYD88* mutations were predictors of short and long overall survival, respectively. In contrast to cytogenetic and FISH data, *NOTCH2* and *TP53* mutations remained independent factors of outcome in multivariate analyses which included the established prognostic markers: anaemia and thrombocytopenia. Genomic and immunogenetic data have the potential to aid diagnosis, and influence the timing and choice of treatment in SMZL.

Abstract

Purpose: Mounting evidence supports the clinical significance of gene mutations and immunogenetic features in common mature B-cell malignancies.

Experimental Design: We undertook a detailed characterization of the genetic background of splenic marginal zone lymphoma (SMZL), using targeted re-sequencing and explored potential clinical implications in a multinational cohort of 175 SMZL patients.

Results: We identified recurrent mutations in *TP53* (16%), *KLF2* (12%), *NOTCH2* (10%), *TNFAIP3* (7%), *MLL2* (11%), *MYD88* (7%) and *ARID1A* (6%), all genes known to be targeted by somatic mutation in SMZL. *KLF2* mutations were early, clonal events, enriched in patients with del(7q) and *IGHV1-2*04* B-cell receptor immunoglobulins, and were associated with a short median time-to-first-treatment (0.12 vs. 1.11 yrs; P=0.01). In multivariate analysis mutations in *NOTCH2* (HR 2.12, 95%CI 1.02-4.4, P=0.044) and 100% germline *IGHV* gene identity (HR 2.19, 95%CI 1.05-4.55, P=0.036) were independent markers of short time-to-first-treatment, while *TP53* mutations were an independent marker of short overall survival (HR 2.36, 95% CI 1.08-5.2, P=0.03).

Conclusion: We identify key associations between gene mutations and clinical outcome, demonstrating for the first time that *NOTCH2* and *TP53* gene mutations are independent markers of reduced treatment-free and overall survival, respectively.

Introduction

Splenic marginal zone lymphoma (SMZL) is a rare chronic B-cell lymphoproliferative disorder that predominantly affects elderly patients and involves the spleen, bone marrow, and usually the peripheral blood (1). The diagnosis is based on a combination of clinical, morphological, histopathological and immunophenotypic data, that serve to distinguish it from other splenic lymphomas (1). Additional distinctive biological features of SMZL include a remarkable bias to the expression of clonotypic B-cell receptors (BcR) utilizing the *IGHV1-2*04* gene and frequent deletions of chromosome 7q (2,3). The median survival of SMZL is around 10 years; 70% of patients require treatment for progressive, symptomatic disease and 5-10% undergo transformation into large B-cell lymphoma. Response rates to splenectomy, single agent Rituximab and Rituximab plus chemotherapy are high but approximately 40% of patients develop progressive disease within 5 years (4,5).

Easily measured, non disease-specific, parameters such as haemoglobin, platelet count, LDH, serum albumin and the presence of extrahilar lymphadenopathy have prognostic significance in multivariate analysis for overall survival and have led to the introduction of scoring systems and a prognostic index (6-8). In contrast, the value of biomarkers to predict outcome is much less clear. Unmutated immunoglobulin heavy variable (*IGHV*) genes, karyotypic complexity, *TP53* loss/mutation alone or in combination with del(8p), and del(14q) have all been suggested to have an adverse prognostic significance in univariate analyses but none have been confirmed in multivariate analyses (3,9-13). Candidate gene screening and more recently, whole genomic or whole exomic sequencing (WES) studies in small patient cohorts have identified recurrent mutations of genes involved in NOTCH, BcR, Toll-like receptor (TLR) and NF-κB signaling pathways, chromatin remodelling and the cytoskeleton (14-17). However, targeted resequencing of larger patient cohorts has resulted in conflicting data on the incidence and prognostic significance of *NOTCH2* mutations while little is known about the clinical significance of other gene mutations (14,15).

These observations highlight the need for larger studies to determine a more comprehensive picture of the clinical significance of gene mutations in SMZL. Accordingly, using a targeted re-sequencing approach, we screened for mutations in the largest cohort of well-characterised SMZL cases published to date [n=175] and identified a number of gene mutations that contribute to reduced outcome in SMZL. Most notably we demonstrate for the first time that previously known gene mutations (*NOTCH2* and *TP53*) are independent markers of poor survival.

Material and methods

Patients and samples

Table 1 describes our cohort of 175 SMZL patients from eight centres across Europe, all meeting established diagnostic criteria (18). The mean time from diagnosis to sampling was 3.2 years (0-24, SD = 4.7). Mantle cell lymphoma (MCL) was excluded in CD5+ve cases using FISH and conventional cytogenetics. Splenic lymphoma/leukemia unclassifiable (SLLU) was precluded either by splenic histopathology or by omission of SLLU-variant cases with distinctive cytology, such as those with splenic diffuse red pulp lymphoma (SDRL). Each transformation event was diagnosed histologically. Informed patient consent was obtained according to the declaration of Helsinki and the study was ethically approved by the local REC.

DNA was extracted from either peripheral blood [n=135], bone marrow [n=22], spleen [n=17] or lymph nodes [n=1]. Germline DNA was obtained from buccal cells or sorted T cells [n=25]. The Sequential DNA samples from 9 cases either diagnosed as clonal lymphocytosis of marginal-zone origin (19) [CBL-MZ, n=1] or SMZL [n=8] (mean of 4.3 yrs between samples, **Supp Table 1**) were evaluated to investigate the clonal evolution of key gene mutations.

Haloplex Re-sequencing and Sanger validation

189 DNA samples from 175 SMZL cases were analysed with Haloplex Target Enrichment system (Agilent Technologies) that enriched 2.39 Mb of genomic DNA for the coding regions of 49 genes known to be targeted by somatic mutations in SMZL, and an additional 719 genes with a postulated role in the pathophysiology of SMZL or other chronic B-cell lymphoproliferative disorders (**Supp Methods and Supp Table 2**). Independent analysis was performed by the University of Southampton and Uppsala University to allow the identification of high-confidence variants in our cohort (**Supp Methods**).

Using conventional Sanger sequencing, we validated 86 variants identified in a number of genes using the experimental conditions and primers described in **Supp Table 3**. Furthermore, we independently screened *NOTCH2* exon 34 in 145/175 SMZL, using primers from Rossi et al (14) .

Statistical analysis

Statistical analysis was performed using SPSS (v20). Time-to-first-treatment (TTFT), Event-free (EFS) and Overall survival (OS) as defined in the **Supp Methods**. Our cohort has 81% power to detect an Overall Survival 0.5 Hazard Ratio associated with *NOTCH2* mutations present in 26% of patients (as observed (14)). Results were determined to be statistically significant at the 5% level.

Results

Overview of re-sequencing data

The mean re-sequencing depth across our gene panel was 297-fold (range 129-702). More than 85% of all bases were covered at >50-fold. The analysis described herein, focuses on the biological and clinical importance of key recurrently mutated genes (**Figure 1A**) known to be somatically acquired in SMZL based on previously published data (14-17,20-22). For our data on other gene mutations, whilst many are annotated in the COSMIC database they could not be confirmed as somatically acquired due to the lack of patient germ-line material and were not taken forward for analysis (Detailed in **Supp Table 4**). This lack of germ-line material is not unexpected in an international retrospective cohort of rare tumours such as SMZL.

In our cohort, we identified recurrent mutations, at suitable frequency for accurate clinical correlations, in *TP53* [n=26 cases], *KLF2* [n=21], *MLL2* [n=20], *NOTCH2* [n=17], *TNFAIP3* [n=13], *MYD88* [n=12], *ARID1A* [n=10], *NOTCH1* [n=10] and *CREBBP* [n=9] (**Figure 1A and B, Supp Figure 1**). For validation, we employed Sanger sequencing and confirmed the presence of 86/86 selected variants in these genes and we showed 99% concordance between our Haloplex and Sanger sequencing of *NOTCH2* exon 34 (**Supp Table 5**). Furthermore, the Haloplex analysis of paired tumour/normal DNA samples [n=14], showed the presence of somatically acquired mutations in these genes [n=77] and, critically, no germ-line variants were identified in the genes we focus on herein (**Supp Table 6**). Therefore, our analytical data supports the somatic origin of mutations within these recurrent genes.

Mutation patterns and evolution

To obtain insight into the genomic context of these gene mutations, we submitted our data to two analytical packages. Firstly, we searched for pairwise gene correlations and mutually exclusive relationships between our 'known somatic' mutated genes using the mutation relation test (MRT, Genome Music) (23). There are 11 and 13 significant MRT co-occurring and mutually exclusive relationships between mutated genes, respectively (considering only relationships with P<0.001, **Figure 1C**), that demonstrated the following classes of gene mutation relationships: (1) a single and distinct independent gene mutation event, such as *MYD88*, where a mutation is invariably observed as an isolated event; (2) the presence of cancer drivers that have many mutually exclusive relationships, such as *NOTCH2, TP53* and *TNFAIP3*, and (3) a group of genes such as *KLF2* and *ARID1A* that have more co-occurring relationships, thus suggesting a synergistic function to promote tumorigenesis.

Secondly, we studied clonal evolution in SMZL, by differentiating between early, clonal events, and later, subclonal mutations. In order to do this, we initially performed integrative analysis of our Haloplex re-sequencing and SNP6 copy number data from our seven published WES SMZL cases (17), employing the ABSOLUTE algorithm (24) (**Supp Figure 2**). Using this approach, all our initial cases harboured a diploid genome, so we extended this analysis to include an additional 38 samples without copy number data but with purity information available from FACS analysis. Using this approach we were able to classify clonal or subclonal mutation in *KLF2*, *NOTCH2*, *TP53, CREBBP* and *TNFAIP3* (**Figure 2A and 2B**).

To extend this single time-point bioinformatics analysis, we also analysed a second DNA sample preceding or subsequent to SMZL diagnosis in 9 patients. For this analysis, we again only focused on variation in genes known to be targeted by somatic mutations in SMZL. We identified 12 variants, and after accounting for tumour purity, these data are outlined in **Supp Table 1 and Figure 2C**. Whilst the number of mutations per case was insufficient for comprehensive analysis, the following observations could be made: 1) 3 patients harboured mutations that remained fully clonal over the two timepoints (*ARID1A* [n=1] and *CREBBP* [n=2]) which supports the ABSOLUTE data for these genes, 2) 6 cases contained seven mutations where the normalized VAF increased over time, supporting the hypothesis that these genes are important in driving the disease, including four patients with *TP53* mutations that acquire a deletion of 17p (isochromosomes 17q) at the second timepoint (**Figure 2C**), and 3) three patients displayed either a mutation that became undetectable at timepoint 2 (*TNFAIP3* and *NOTCH2*), or one that remained at a low VAF (*NOTCH2*, 0.04% allele frequency), even with a concomitantly emerging *TP53* mutation (**Supp Table 1**).

Biologically significant mutations in SMZL

KLF2, or Krüppel-like factor 2, mutations were detected in 21/175 cases (12%, **Figure 1A**) and were distributed across the entire protein, with a cluster in the C2H2 domain (C terminus). A Q24X variant was identified in three patients, suggesting the presence of mutation hotspots. Mutations were often (43%) stop-gains or frame-shift variants (**Figure 1B**), suggesting an impact on protein function. All mutations tested were somatically acquired (n=9). From our ABSOLUTE analysis, all 11 mutations were defined as clonal (**Figure 2B**), and other recurrently mutated genes present in these cases were estimated to have lower CCFs than *KLF2* (e.g. *NOTCH2* and *TNFAIP3*, **Figure 2B**). *KLF2* mutations were significantly associated with del(7q) (53% vs. 11%; P=0.001), *IGHV1-2*04* gene usage (50% vs. 7%; P<0.001), and gene mutations including *NOTCH2*, *TNFAIP3* and *ARID1A* (all P<0.001). Together, these observations suggest that the potential cell survival advantage provided by an early *KLF2* mutation allows the acquisition of additional functionally synergistic gene mutations to promote tumourigenesis (**Figure 1C**).

We independently screened *NOTCH2* by Haloplex (mean gene coverage of 572-fold) and direct Sanger sequencing of exon 34, and identified 18 mutations in 17 patients, a frequency of 10% in our cohort (**Figure 1A**). We manually examined the *NOTCH2* sequence reads and found no evidence of any additional mutations below the resolution of our variant calling algorithm. As expected, the mutations were nonsense [n=9], frameshift [n=7] and missense [n=2] principally targeting the TAD and PEST domain encoded by exon 34 (**Figure 1B**). Several of our mutations (R2360*, R2400*) have been previously reported to result in over-expression of the Notch2 protein and active signalling (14). *NOTCH2* mutations were classified as sub-clonal or clonal (**Figure 2A**). We also identified *NOTCH1* mutations [n=10], several of which were truncating frameshift indels [n=2, P2514fs*4] or stopgain mutations [n=2] in exon 34.

We identified recurrent mutations in *MYD88* [12/175 cases, 7%] and *TNFAIP3* [13/175 cases, 7%], genes involved in Toll-like receptor and NF-κB signalling. Of the 12 MYD88 mutations, 7 and 2 were the gain-of-function L265P or S219C variants, respectively (25). Mutations in *MYD88* were single and distinct events, mutually exclusive from mutations in *TP53* and *NOTCH2*. Twenty-one *TNFAIP3* mutations were identified in 13 patients (**Figure 1B**), 15 of which would result in truncation of the A20 protein. One of these mutations (E361X) has been shown to abrogate the ability of A20 to negatively regulate NFKB-signalling (26). Mutations co-existed with *KLF2* ($P<0.001$) mutations but showed a reverse association with *NOTCH2* ($P<0.001$) and *TP53* ($P<0.001$).

Mutations of *TP53* and *ARID1A*, both involved in cell cycle control and DNA damage response, were identified in 26/175 (16%) and 10/175 (6%) patients, respectively. We defined 28 missense [n=18], nonsense [n=5], frameshift [n=2] and splicing [n=3] *TP53* mutations, largely annotated in COSMIC (27/28, **Figure 1B**), in 26 patients who tended to have deletions of 17p ($P=0.003$) and a complex karyotype ($P<0.001$, **Figure 1C**). Finally, we confirm our previous study by demonstrating recurrent mutations in *CREBBP* [n=9] (17). All our *CREBBP* mutations appear to be early genetic events as they were classified as fully clonal (**Figure 2A**) akin to the situation in follicular lymphoma (27); two of our mutations were the Y1450C variant previously identified in DLBCL, which has been shown to compromise the protein's ability to acetylate BCL6 and p53 (28).

Clinical significance of mutations in SMZL

Initially we looked for associations between gene mutations and clinical and laboratory features measured routinely in clinical practice (**Figure 3A**). Patients with *KLF2* and *NOTCH2* mutations were at higher risk of receiving treatment including splenectomy (OR=4.51, 95%CI 1.68-12.10; $P=0.002$ & OR=1.16, 95%CI 1.08-1.25; $P=0.007$). Histological evidence of transformation to large B-cell lymphoma was reported in 19/175 (11%) patients; these patients were more likely to have 100% germline IGHV gene identity (40% vs. 10%, $P=0.04$) and exhibited a significantly shorter overall

survival (9.0 vs 16.5 yrs; P=0.04) in comparison to non-transformed cases. The only mutated gene associated with transformation was *TNFAIP3* (32% vs. 4%, P=0.002).

Follow-up outcome data were available for 164, 117 and 169 patients for TTFT, EFS and OS, respectively. First, we demonstrated the clinical relevance of our cohort by testing for the prognostic significance of previously documented clinical and laboratory features (**Table 2**). We then performed univariate analysis of the gene mutations against TTFT, EFS and OS (**Table 2**). Genes associated with reduced TTFT were; 1) *KLF2* (HR 1.93, 95%CI 1.16-3.32, P=0.01) where wild-type and mutant patients exhibited median TTFT of 1.11 and 0.12 years, respectively, and 2) *NOTCH2* (HR 2.13, 95%CI 1.26-3.58, P=0.003) where wild-type and mutant patients exhibited median TTFT of 0.94 and 0.09 years, respectively (**Figure 3B and C**). Gene mutations associated with shorter EFS included *TP53* (HR 2.17, 95%CI 1-4.74, P=0.05) with median EFS of 3.11 and 0.98 years for wild-type and mutated patients, respectively. Finally, we tested the impact of gene mutations on OS and showed reduced survival for *TP53* (HR 2.16, 95%CI 1.05-4.42, P=0.032) mutations with a median OS of 12.21 and 16.03 years for mutant and wild-type cases, respectively, and the reverse for *MYD88* mutated individuals (HR 0.04, 95%CI 0.01-2.48, P=0.02) (**Figure 3D and E**).

***NOTCH2* and *TP53* mutations were independent risk factors for reduced TTFT and OS**

Those gene mutations shown to be associated with reduced outcome in univariate analysis were tested using multivariate Cox proportional hazard analysis. Along with the presence of gene mutations, other variables included in the analysis were age at diagnosis, haemoglobin levels, platelets and lymphocyte counts. We developed these models for TTFT, EFS and OS as they permitted the relative prognostic value of gene mutations to be assessed in a large, informative group of patients in the context of the most available clinical data (**Table 3**). Our multivariate EFS model identified age at diagnosis, lymphocyte count and low platelet count as independent risk factors, however *TP53* became non-significant in this analysis. We show that in addition to haemoglobin levels, both *NOTCH2* (HR 2.12, 95%CI 1.02-4.4, P=0.044) and 100% germline *IGHV* gene identify (HR 2.19, 95%CI 1.05-4.55, P=0.036) are independent risk factors for TTFT. Furthermore, we show that the presence of *TP53* mutation is an independent risk factor for OS (HR 2.36, 95%CI 1.08-5.20, P=0.03).

Discussion

The primary aim of this study was to determine the clinical significance of somatically acquired gene mutations in SMZL, identified in the current and previously reported studies (14-17,22). Notably, we were able to identify key associations between gene mutations and clinical outcome, demonstrating for the first time that *NOTCH2* and *TP53* gene mutations are independent markers of poor outcome.

The main strengths of the present study were the cost-effective resequencing approach which enabled screening of a large number of candidate genes at high sequencing depth and, most importantly, the size of the cohort in a rare lymphoma, enabling us to overcome limitations befalling previous studies evaluating the clinical significance of clinical and genetic biomarkers in SMZL. Indeed, the lack of a treatment naïve clinical trial cohort, historical use of splenectomy for diagnosis, inclusion of non-splenectomised cases who might have SLLU and the indolent nature of the disease, where in an elderly population many patients die from unrelated causes, all underline the need for caution in interpreting outcome data in SMZL. We sought to minimize the effect of these factors in a number of ways: (1) by confining the study to centres with expertise in SMZL, we could ensure expert diagnostic review especially for cases diagnosed prior to the currently-accepted diagnostic criteria, (2) treatments included a limited range of modalities, predominantly splenectomy, alkylating agents and rituximab, and splenectomy was considered to be a therapy regardless of the indication (**Supp Table 7**), and (3) the use of multiple survival endpoints enabled the impact of prognostic markers on disease biology as well as the overall survival of an elderly patient cohort to be assessed.

In additional to confirming the presence of mutations in *TP53*, and in genes involved in NOTCH, BcR, TLR, NF-κB signaling and in chromatin modifiers (14-17,22), we identified recurrent heterozygous inactivating mutations in *KLF2*, a member of the Krüppel-like family of transcription factors with roles in cell differentiation, proliferation, activation and trafficking (29), in 12% of analyzed cases. *KLF2* was included in our re-sequencing experiments due to reanalysis of our published WES data (17), that showed evidence of mutations in 4/7 cases in spite of the low sequence read-depth present at this locus. During the preparation and submission of this manuscript, two studies independently identified recurrent *KLF2* mutations in SMZL, at a frequency higher than in our study (22,30), which is likely to be a reflection of the patient cohort analyzed in our current study, as we identified a lower frequency of del(7q) and *IGHV1-2*04* in our cohort, compared to other large studies (3,22). Interestingly, in mice, KLF2 deficiency is associated with a failure to maintain B-1 B cells, expansion of the marginal zone B-cell pool and expression of marginal zone characteristics by follicular B cells (31-33). These observations may reasonably be considered as indicating a role of *KLF2* mutations in the natural history of SMZL, an argument also supported by their significant enrichment among SMZL cases with clonotypic BcR utilizing the *IGHV1-2*04* gene. This alludes to acquisition and/or selection

of *KLF2* mutations in a context of particular signaling via specific BCRs with distinctive immunogenetic features, similar to what has been observed in other B-cell malignancies, most notably in stereotyped subsets of CLL (34). In 11 cases we were able to study the clonal architecture of *KLF2* mutations: in each case the *KLF2* mutations were clonal and were associated with other subclonal mutations, often involving other clinically significant genes such as *NOTCH2* and *TNFAIP3.* The invariable association of clonal *KLF2* mutations with other mutations involving different pathways, and deletions of 7q suggests that the former may have a pro-survival function and additional mutations may be necessary for disease progression. It will be of interest to determine the incidence of *KLF2* mutations and genomic complexity in cases of clonal B-lymphocytosis of marginal-zone origin (CBL-MZ) (19), especially in those cases progressing to SMZL.

Consistent with the role of the NOTCH2-Delta-like 1 ligand pathway in normal marginal zone development (35), and the previously reported finding of *NOTCH2* mutations in SMZL (14,15), we found recurrent mutations in exon 34 of *NOTCH2* in 10% of cases. This compares to an incidence of 21% (14), 25% (15) and 7% (16) in previously reported series, probably reflecting differences in sample size and cohort composition. We also detected recurrent mutations in other NOTCH pathway genes including *NOTCH 3* and *4* and *SPEN*. However, their role in the pathogenesis of SMZL is unclear so we have not focused on them specifically.

We evaluated the prognostic significance of somatically acquired mutations on TTFT, EFS and OS. Short TTFT was associated with mutations in *KLF2, NOTCH2* and *ARID1A*; short EFS with mutations in *TP53;* and, short OS with mutations in *TP53*, whereas *MYD88* mutations were associated with a longer OS. Although the study by Rossi *et al* indicated that *NOTCH2* mutations were associated with a prolonged 5-year OS and progression-free survival following first line treatment (14), our data based on a substantially larger cohort shows that *NOTCH2* mutation are linked to reduced outcome, an observation corroborated by Kiel *et al* who also noted an association with shorter time from diagnosis to either relapse, transformation or death, albeit in a much smaller cohort of SMZL cases [n=46] (15). Additional studies of larger patient cohorts will be required to validate the clinical importance of *NOTCH2* mutations. Cases with a *MYD88* mutation exhibited longer OS and comparable clinical and laboratory features to other cases with SMZL apart from a higher incidence of low level IgM paraproteins, detected in 8/9 cases with available data. The poor prognostic significance of *TP53* mutations is consistent with previously reported data on *TP53* abnormalities.

Our multivariate analysis demonstrated for the first time in SMZL that both genetic and immunogenetic parameters retained prognostic significance in a model that included age, haemoglobin, platelet count and lymphocyte count. We chose to base our multivariate analysis on the study of Salido *et al*, as this is the only large study to include base-line clinical variables with chromosomal features (3). While the independent prognostic significance of age, anaemia and

thrombocytopenia were expected and consistent with many previous studies, the biological basis for the impact of a lymphocyte count of < $4x10^9$/l, noted in an early (36) but not in more recent studies (3,37), requires further investigation. Interestingly, our study suggests that gene mutations, such as those targeting *NOTCH2* and *TP53* have more clinical utility than cytogenetic features, such as karyotypic complexity, 14q aberrations and *TP53* deletions that did not retain prognostic significance in previous reports (3). Specifically, *NOTCH2* and truly unmutated *IGHV* genes (but not unmutated *IGHV* genes using a 98% cut-off) were independent markers of TTFT and *TP53* of OS. Given the historical use of splenectomy to both diagnose and treat SMZL patients, our associations with TTFT should be considered with a note of caution.

Since transformation to a high grade lymphoma is usually associated with resistance to treatment and very poor survival, we were also interested to see if any genomic abnormalities were associated with an increased risk of transformation, as noted for *NOTCH1* mutations in CLL (38). In our study, *TNFAIP3* mutations together with truly unmutated clonotypic *IGHV* genes were all found at a higher frequency in cases that subsequently transformed. Further studies comparing the genomic landscape of paired chronic phase and clonally-related transformed samples, as performed in CLL and FL (27,39), will be required to determine the drivers of transformation.

In summary, we show that gene mutations and immunogenetic features have prognostic significance in a large and well-characterized cohort of patients with SMZL. Additional studies will be required to confirm our findings and to determine the functional consequences of these mutations, the incidence and importance of copy number and epigenetic abnormalities in gene silencing and the clinical value of mutation screening in the differential diagnosis and management of SMZL.

Acknowledgements

The authors gratefully acknowledge all patients who contributed to this study. The authors would also like to thank Professor Nick Cross for help during the preparation of this manuscript.

Grant Support

This work was funded by Leukaemia and Lymphoma Research, the Kay Kendall Leukaemia Fund, Cancer Research UK, Wessex Medical Research, the Bournemouth Leukaemia Fund, the Swedish Cancer Society and the Swedish Research Council.

Authorship contributions

MP wrote the paper, performed research and analysed the data, MRZ performed research, analysed the data, performed statistical analysis and wrote the paper, VL analysed and interpreted the data, JG analysed and interpreted data, JW analysed and interpreted the data, RW collected, analysed and interpreted data, HP performed research and analysed the data, AP performed research, ZD performed research, AG performed research, CK collected, analysed and interpreted data, ES collected, analysed and interpreted data, AX collected, analysed and interpreted data, CF collected, analysed and interpreted data, DGdC collected, analysed and interpreted data, CD collected, analysed and interpreted data, GP collected, analysed and interpreted data, RR analysed the data, MAK collected, analysed and interpreted data, FF collected, analysed and interpreted data, AC performed statistical analysis, PG collected, analysed and interpreted data, EM collected, analysed and interpreted data, GP collected, analysed and interpreted data, KS collected, analysed and interpreted data, DO designed the research, collected, analysed and interpreted data, and wrote the paper, JCS designed the research, analysed the data and wrote the paper.

The authors declare no conflict of interests.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition. Lyon, France: World Health Organization; 2008.

2. Bikos V, Darzentas N, Hadzidimitriou A, Davis Z, Hockley S, Traverse-Glehen A, et al. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. Leukemia 2012;26(7):1638-46.

3. Salido M, Baró C, Oscier D, Stamatopoulos K, Dierlamm J, Matutes E, et al. Cytogenetic aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. Blood 2010;116(9):1479-88.

4. Kalpadakis C, Pangalis GA, Angelopoulou MK, Sachanas S, Kontopidou FN, Yiakoumis X, et al. Treatment of splenic marginal zone lymphoma with rituximab monotherapy: progress report and comparison with splenectomy. Oncologist 2013;18(2):190-7.

5. Lenglet J, Traullé C, Mounier N, Benet C, Munoz-Bongrand N, Amorin S, et al. Long-term follow-up analysis of 100 patients with splenic marginal zone lymphoma treated with splenectomy as first-line treatment. Leuk Lymphoma 2014;55(8):1854-60.

6. Montalbán C, Abraira V, Arcaini L, Domingo-Domenech E, Guisado-Vasco P, Iannitto E, et al. Risk stratification for Splenic Marginal Zone Lymphoma based on haemoglobin concentration, platelet count, high lactate dehydrogenase level and extrahilar lymphadenopathy: development and validation on 593 cases. Br J Haematol 2013;159(2):164-71.

7. Montalban C, Abraira V, Arcaini L, Domingo-Domenech E, Guisado-Vasco P, Iannitto E, et al. Simplification of risk stratification for splenic marginal zone lymphoma: a point-based score for practical use. Leuk Lymphoma 2013:[Epub ahead of print].

8. Arcaini L, Lazzarino M, Colombo N, Burcheri S, Boveri E, Paulli M, et al. Splenic marginal zone lymphoma: a prognostic model for clinical use. Blood 2006;107(12):4643-9.

9. Algara P, Mateo MS, Sanchez-Beato M, Mollejo M, Navas IC, Romero L, et al. Analysis of the IgV(H) somatic mutations in splenic marginal zone lymphoma defines a group of unmutated cases with frequent 7q deletion and adverse clinical course. Blood 2002;99(4):1299-304.

10. Gruszka-Westwood AM, Hamoudi RA, Matutes E, Tuset E, Catovsky D. p53 abnormalities in splenic lymphoma with villous lymphocytes. Blood 2001;97(11):3552-8.

11. Fresquet V, Robles EF, Parker A, Martinez-Useros J, Mena M, Malumbres R, et al. High-throughput sequencing analysis of the chromosome 7q32 deletion reveals IRF5 as a potential tumour suppressor in splenic marginal-zone lymphoma. Br J Haematol 2012;158(6):712-26.

12. Hockley SL, Else M, Morilla A, Wotherspoon A, Dearden C, Catovsky D, et al. The prognostic impact of clinical and molecular features in hairy cell leukaemia variant and splenic marginal zone lymphoma. Br J Haematol 2012;158(3):347-54.

13. Rinaldi A, Mian M, Chigrinova E, Arcaini L, Bhagat G, Novak U, et al. Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact on the clinical outcome. Blood 2011;117(5):1595-604.

14. Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. J Exp Med 2012;209(9):1537-51.

15. Kiel MJ, Velusamy T, Betz BL, Zhao L, Weigelin HG, Chiang MY, et al. Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma. J Exp Med 2012;209(9):1553-65.

16. Martínez N, Almaraz C, Vaqué JP, Varela I, Derdak S, Beltran S, et al. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. Leukemia 2014:[Epub ahead of print].

17. Parry M, Rose-Zerilli MJ, Gibson J, Ennis S, Walewska R, Forster J, et al. Whole exome sequencing identifies novel recurrently mutated genes in patients with splenic marginal zone lymphoma. PLoS One 2013;8(12):e83244.

18. Matutes E, Oscier D, Montalban C, Berger F, Callet-Bauchu E, Dogan A, et al. Splenic marginal zone lymphoma proposals for a revision of diagnostic, staging and therapeutic criteria. Leukemia 2008;22(3):487-95.

19. Xochelli A, Kalpadakis C, Gardiner A, Baliakas P, Vassilakopoulos TP, Mould S, et al. Clonal B-cell lymphocytosis exhibiting immunophenotypic features consistent with a marginal zone origin: is this a distinct entity? Blood 2013:[Epub ahead of print].

20. Rossi D, Deaglio S, Dominguez-Sola D, Rasi S, Vaisitti T, Agostinelli C, et al. Alteration of BIRC3 and multiple other NF-κB pathway genes in splenic marginal zone lymphoma. Blood 2011;118(18):4930-4.

21. Novak U, Rinaldi A, Kwee I, Nandula SV, Rancoita PM, Compagno M, et al. The NF-{kappa}B negative regulator TNFAIP3 (A20) is inactivated by somatic mutations and genomic deletions in marginal zone lymphomas. Blood 2009;113(20):4918-21.

22. Piva R, Deaglio S, Famà R, Buonincontri R, Scarfò I, Bruscaggin A, et al. The krüppel-like factor 2 transcription factor gene is recurrently mutated in splenic marginal zone lymphoma. Leukemia 2014:[Epub ahead of print].

23. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, et al. MuSiC: identifying mutational significance in cancer genomes. Genome Res 2012;22(8):1589-98.

24. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotech 2012;30(5):413-21.

25. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. Nature 2011;470(7332):115-9.

26. Honma K, Tsuzuki S, Nakagawa M, Tagawa H, Nakamura S, Morishima Y, et al. TNFAIP3/A20 functions as a novel tumor suppressor gene in several subtypes of non-Hodgkin lymphomas. Blood 2009;114(12):2467-75.

27. Okosun J, Bödör C, Wang J, Araf S, Yang CY, Pan C, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat Genet 2014;46(2):176-81.

28. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 2011;471(7337):189-95.

29. Pearson R, Fleetwood J, Eaton S, Crossley M, Bao S. Kruppel-like transcription factors: A functional family. Int J Biochem Cell Biol 2008;40:1996-2001.

30. Clipson A, Wang M, de Leval L, Ashton-Key M, Wotherspoon A, Vassiliou G, et al. KLF2 mutation is the most frequent somatic change in splenic marginal zone lymphoma and identifies a subset with distinct genotype. Leukemia 2015:[Epub ahead of print].

31. Hart GT, Wang X, Hogquist KA, Jameson SC. Krüppel-like factor 2 (KLF2) regulates B-cell reactivity, subset differentiation, and trafficking molecule expression. Proc Natl Acad Sci U S A 2011;108(2):716-21.

32. Hoek KL, Gordy LE, Collins PL, Parekh VV, Aune TM, Joyce S, et al. Follicular B cell trafficking within the spleen actively restricts humoral immune responses. Immunity 2010;33(2):254-65.

33. Hart GT, Peery SL, Hamilton SE, Jameson SC. Cutting edge: Krűppel-like factor 2 is required for phenotypic maintenance but not development of B1 B cells. J Immunol 2012;189(7):3293-7.

34. Strefford JC, Sutton LA, 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 2013;27(11):2196-9.

35. Saito T, Chiba S, Ichikawa M, Kunisato A, Asai T, Shimizu K, et al. Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. Immunity 2003;18(5):675-85.

36. Troussard X, Valensi F, Duchayne E, Garand R, Felman P, Tulliez M, et al. Splenic lymphoma with villous lymphocytes: clinical presentation, biology and prognostic factors in a series of 100 patients. Groupe Francais d'Hématologie Cellulaire (GFHC). Br J Haematol 1996;93(3):731-6.

37. Parry-Jones N, Matutes E, Gruszka-Westwood AM, Swansbury GJ, Wotherspoon AC, Catovsky D. Prognostic features of splenic lymphoma with villous lymphocytes: a report on 129 patients. Br J Haematol 2003;120(5):759-64.

38. Algrin C, Brisou G, Beldjord K, Mounier N, Berger F, Baseggio L, et al. Prognostic Impact Of Somatic NOTCH2 Mutation In Splenic Marginal Zone Lymphoma Blood - ASH Abstracts 2013;122(21):4247.

39. 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 2013;210(11):2273-88.

40. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell 2013;152(4):714-26.

Figure 1. Distribution of recurrent gene mutations across patients, gene mutation maps and gene by gene associations for 175 SMZL cases.

(A) Heatmap of the distribution of gene mutation in our cohort

(B) Schematic diagram of the protein targeted by key mutations in SMZL, with their key functional domains. The symbols and colour denote the type of mutation. Mutations annotated in COSMIC v68 database are in bold text.

(C) Associations between genetic and immunogenetic features of our SMZL cohort. Shows the pairwise associations amongst significantly mutated genes, genetic and immunogenetic features (labelled as 'Genomic Feature') across 175 SMZL cases. Genes are annotated within key pathways known to be important in the pathogenesis of mature B-cell malignancies. The number of mutations (n) for each gene mutation in the analysis is shown. An association is shaded based on the significance and only gene by genomic feature (top matrix) and gene by gene (bottom matrix) associations with a p-value of <0.01 (Chi-squared/Fisher's exact test) or <0.001 (Mutation Relation Test, Genome MuSiC analysis) are included, respectively.

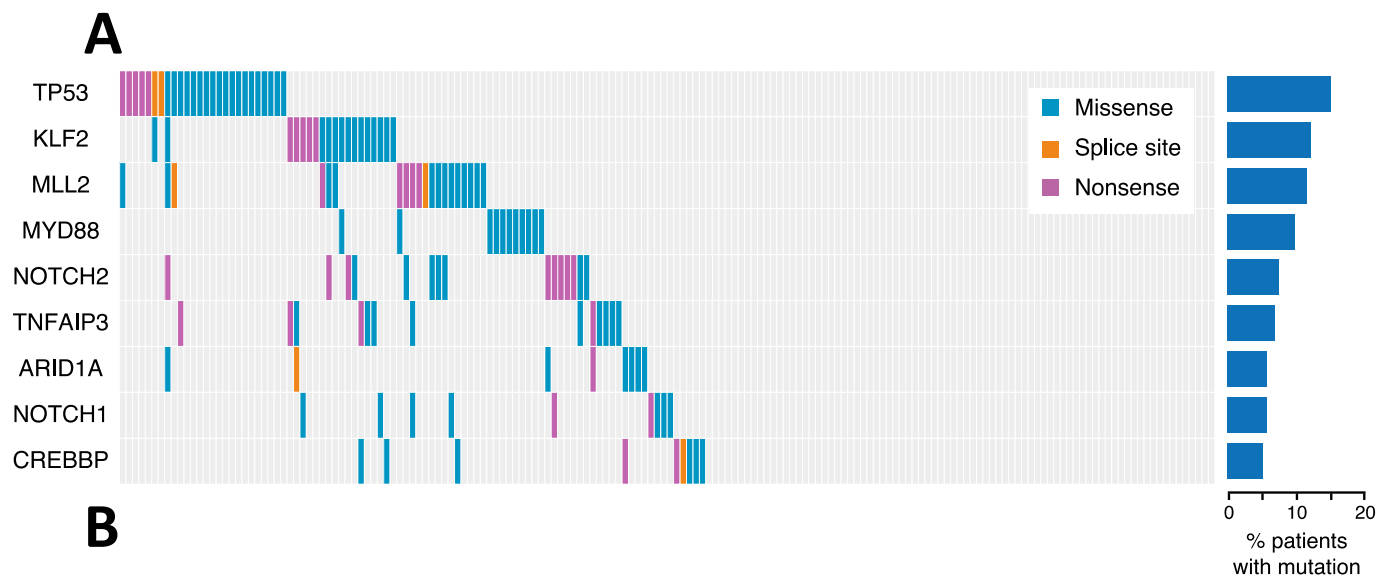
Figure 2. Clonal distribution and temporal analysis of gene mutations in SMZL.

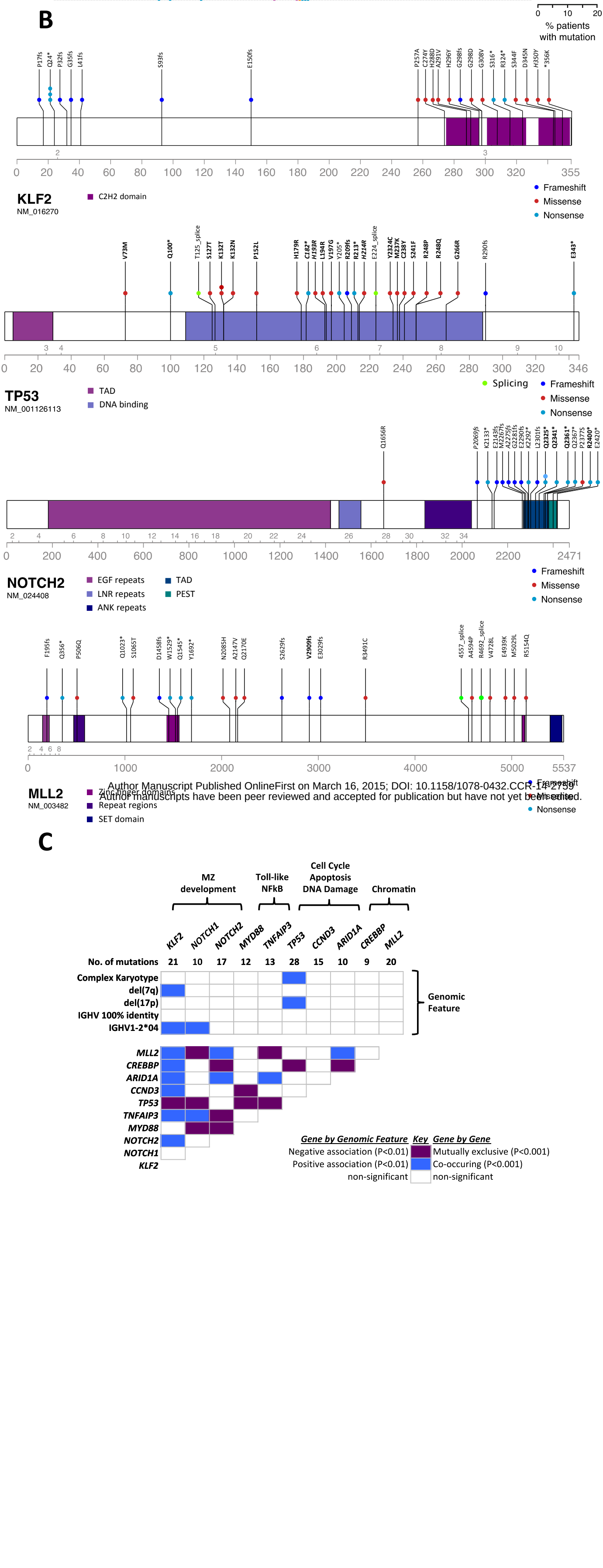
(A) Shows the distribution of the estimated proportion of tumour cells harbouring a mutation in 45 patients, based on the availability of purity information. Genes are displayed from left to right showing genes displaying more clonal and subclonal mutations, respectively, by using a binomial distribution based on the alternative allele read count, the total read count from cancer cells and an expected variant allelic fraction (VAF) of 0.45. For each gene, the bone-and-whisker plots show the adjusted ratio of observed VAF divided by the 50% of the purity estimate derived from CD19+ FACS data. The number of cases (n) for each gene mutation in the analysis is shown (bottom). (B) Shows the presence of clonal *KLF2* mutations in five SMZL patients with matched deep re-sequencing and SNP6 data available. No *KLF2* gene deletions were identified by SNP6 copy number data analysis. For each cases the cancer cell fraction (CCF) derived with the ABSOLUTE algorithm is shown for the *KLF2* variant and co-occurring mutations. This approach estimates the cancer cell fraction (CCF) harbouring a mutation by correcting for sample purity and local copy number changes, where mutations are classified as clonal if the CCF was >0.95 with a probability >0.5, and subclonal otherwise (40). (C) Temporal re-sequencing analysis of sequential time points in cases showing clonal expansion of gene mutations (7 of 12 mutations in 9 patients). The Y-axis shows the VAF for a given mutation after accounting for tumour purity.

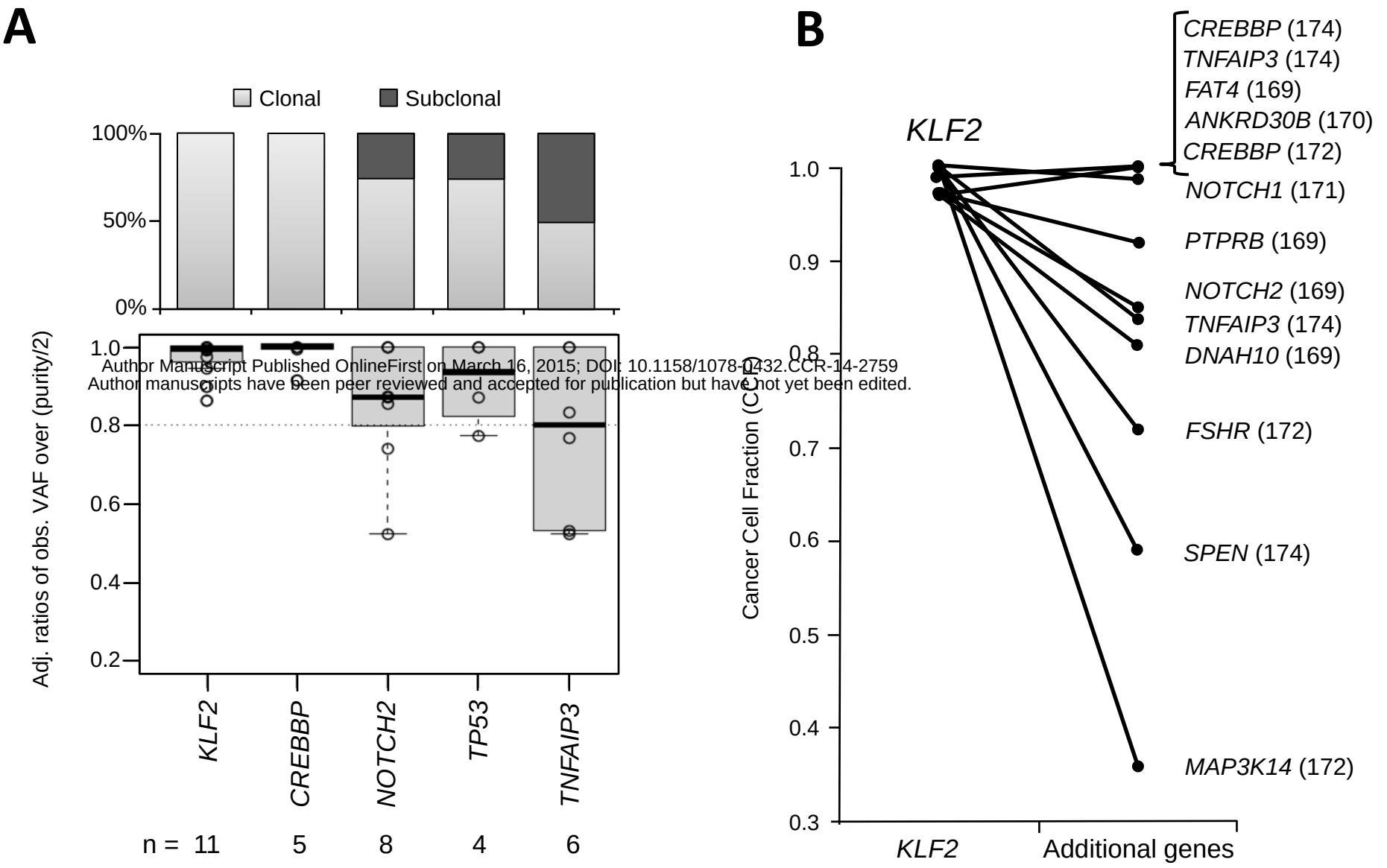
Figure 3. The clinical significance of gene mutations in SMZL.

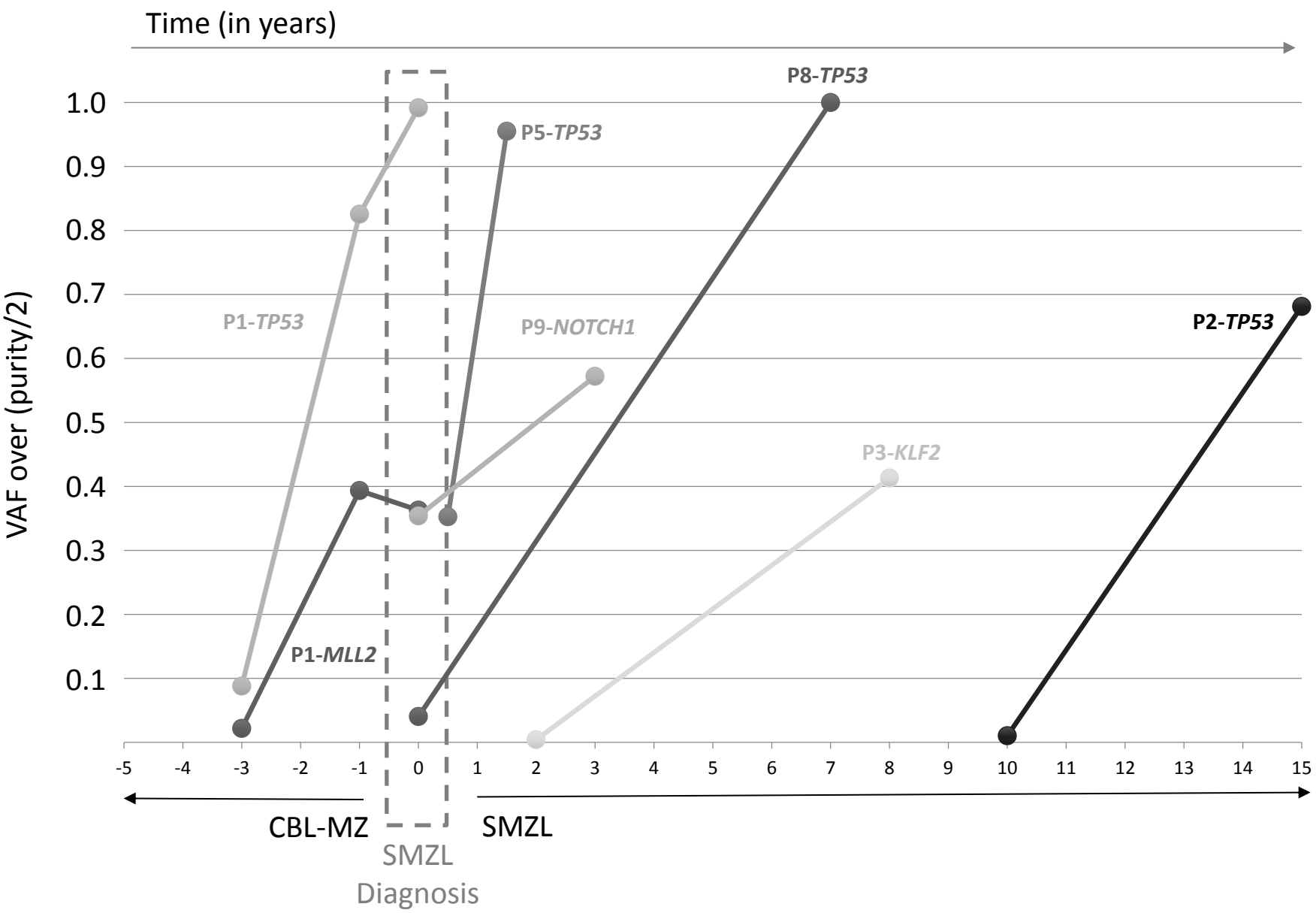
(A) Shows the associations between the presence of gene mutations and clinical features. Where possible genes are annotated within key pathways known to be important in the pathogenesis of

mature B-cell malignancies. The number of mutations (n) for each gene mutation in the analysis is shown. An association is shaded based on the significance and only associations with a p-value of <0.05 are included (Chi-squared/Fisher's exact test). (B) and (C) show KM plots for time to first treatment for patients with *KLF2* and *NOTCH2* mutations, respectively. (D) and (E) show overall survival KM plots for patients with *TP53* and *MYD88* mutations, respectively. For each KM Plot, the grey and black lines identify the wild-type and mutated patient groups, respectively. The P values are derived from Kaplan-Meier analysis with a log-rank test and median survival times with 95% confidence intervals.









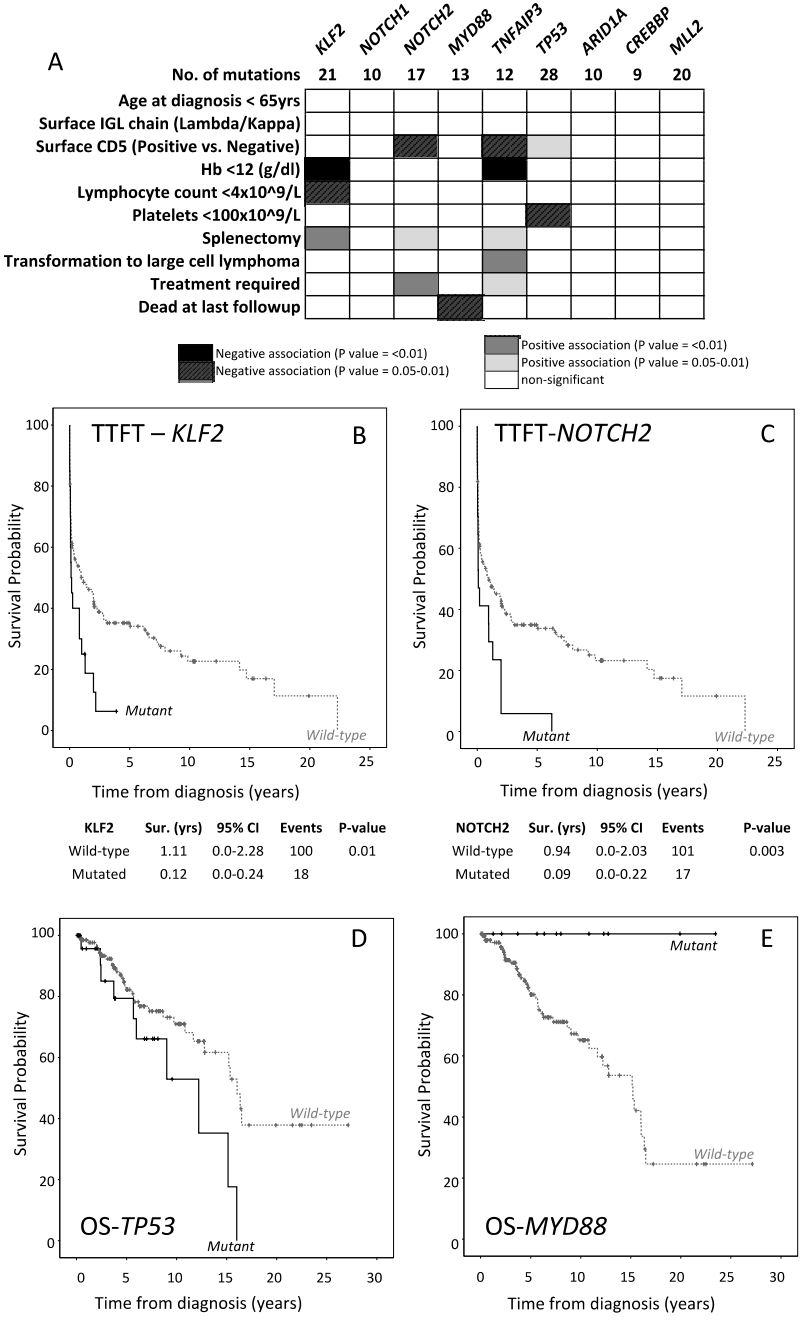


Table1: Patient Characteristics

Variable	Definition	N	%
Number of Patients	SMZL diagnosis	175	100%
Age at Diagnosis	mean (range)	68 (36-90)	
IGHV genes	*IGHV1-2*04*	16	13%
	not *IGHV1-2*04*	108	87%
del(7q) status from karyotype	del(7q)	17	19%
	normal	74	81%
Surface CD5 FACS result	CD5+	40	27%
	CD5-	108	73%
WBC	mean (range) ($x10^9$/L)	20 (0.5-158)	
Hb	<12 g/dl	73	44%
	≥12 g/dl	93	56%
Lymphocyte count	$<4x10^9$/L	41	27%
	$\geq 4x10^9$/L	110	73%
Platelets $<100x10^9$/L	$<100x10^9$/L	30	18%
	$\geq 100x10^9$/L	134	82%
high-risk FISH results	del(17p)	10	33%
	normal	20	67%
Splenectomy	YES	55	34%
	NO	109	67%
Transformation to large cell lymphoma	YES	19	17%
	NO	91	83%
TTFT status	treated (inc. Splenectomy)	122	74%
	Untreated	42	26%
Event-Free Survival status	Event (Death, transformation, 2nd Tx)	52	44%
	No Event	65	56%
Status (Alive or Dead at last followup)	Dead	43	25%
	Alive	126	75%

Footnote: Complex Karyotype was defined as ≥2 cytogenetically-visible clonal alterations; Tx: treatment

Table 2: Univariate survival analysis of recurrently mutated genes

	Variable	Description	Total	Events	Median (yrs)	95% CI	HR	95% CI	P-value
TTFT	*KLF2*	mutated	20	18	0.12	0.0-0.24	1.93	1.16-3.23	**0.01**
		unmutated	140	100	1.11	0.0-2.28			
	NOTCH2	mutated	17	17	0.09	0.0-0.22	2.13	1.26-3.58	**0.003**
		unmutated	143	101	0.94	0.0-2.03			
	Hb	<12 g/dl	70	63	0.1	0.04-0.17	2.75	1.87-4.02	**<0.001**
		>12 g/dl	84	51	2.73	0.0-7.14			
	Lymphocytes	<4 x 10^9/l	40	33	0.15	0.07-0.24	1.76	1.16-2.68	**0.007**
		>4 x 10^9/l	101	69	1.43	0.50-2.37			
	IGHV identity	100%	12	11	0.14	0.0-0.38	2.06	1.07-3.74	**0.027**
		< 100%	78	50	1.98	0.98-2.99			
EFS	*TP53*	mutated	15	8	0.98	0.04-12.22	2.17	1.00-4.74	**0.05**
		unmutated	84	32	3.11	2.35-6.20			
	Age	>65 yrs	53	26	6.82[a]	4.45-9.20	2.09	1.07-4.08	**0.028**
		<65 yrs	45	14	12.69[a]	9.19-16.18			
	Platelet count	< 100 x 10^9/l	19	11	2.92	2.03-3.80	1.99	0.98-4.02	**0.052**
		> 100 x 10^9/l	78	28	6.91	4.47-9.34			
OS	*TP53*	mutated	26	10	12.21	5.28-19.14	2.16	1.05-4.43	**0.032**
		unmutated	134	33	16.03	14.64-17.43			
	MYD88	mutated	12	0	-[b]	-[b]	-[c]	-[c]	**0.02**[d]
		unmutated	148	43	-				
	Age	>65 yrs	103	37	10.36[a]	9.0-11.76	6.37	2.55-15.87	**<0.001**
		<65 yrs	56	6	22.65[a]	19.38-25.91			
	Hb	<12 g/dl	68	24	9.01	2.90-15.12	2.69	1.45-4.99	**0.001**
		>12 g/dl	87	18	16.35	14.99-17.70			

Footnote: Log-Rank P-values. [a] Mean survival value as median not reached. [b] No events in *MYD88* mutated cases and median survival times not presented, follow-up time ranged from 1.25 to 19.9 years. [c] HR and 95% CI cannot be reliable calculated as there are no events in *MYD88* mutated group. [d] Log-Rank P-value for Chi-squared value reported for the *MYD88* OS Kaplan-Meier analysis (**See Figure 3E**).

Table 3: Multivariate survival analysis of recurrently mutated genes

Variable	TTFT		
	HR	95% CI	P-Value
Hb<12g/dl	2.28	1.32-3.96	0.003
IGHV 100% identity	2.19	1.05-4.55	0.036
NOTCH2	2.12	1.02-4.40	0.044
	EFS		
	HR	95% CI	P-Value
Plts<100x10^9L	3.75	1.68-8.41	0.001
Lymphocytes <4x10^9L	0.41	0.17-0.96	0.04
Age at diagnosis <65yrs	0.45	0.21-0.96	0.038
	OS		
	HR	95% CI	P-Value
Hb<12gdl	2.18	1.12-4.23	0.02
Lymphocytes <4x10^9L	2.35	1.11-4.97	0.03
Age at diagnosis <65yrs	0.09	0.03-0.27	<0.001
TP53	2.36	1.08-5.20	0.03

Footnote: TTFT Multivariate: 83 cases with 56 events; 92 cases with missing data. EFS Multivariate: 82 cases with 35 events; 93 cases with missing data. OS Multivariate: 134 cases with 38 events; 38 cases with missing data. Backwards-step regression was employed, including the following clinical variables (Hb<12gdl, Plts<100x10^9L, Lymphocytes <4x10^9L, Age at diagnosis <65yrs) and the representative gene status variables significantly associated with treatment, event and survival outcome in univariate analysis (Table 2). The TTFT model also included *IGHV* 100% identity, *KLF2* and *NOTCH2* mutation status. EFS and OS also included *TP53* mutation status. Variables removed from the backwards-step regression are not shown.

AACR American Association for Cancer Research

Clinical Cancer Research

Genetics and Prognostication in Splenic Marginal Zone Lymphoma: Revelations from Deep Sequencing

Marina Parry, Matthew J Rose-Zerilli, Viktor Ljungström, et al.

Clin Cancer Res Published OnlineFirst March 16, 2015.