Integrated Cellular and Plasma Proteomics of Contrasting B-cell Cancers Reveals Common and Systemic Tumour Signatures

Harvey E. Johnston[1,2], Matthew J. Carter[1], Kerry L. Cox[1], Melanie Dunscombe[1], Antigoni Manousopoulou[2,3], Paul A. Townsend[4], Spiro D. Garbis[2,3], Mark S. Cragg[1]*.

[1] Antibody and Vaccine Group, Cancer Sciences Unit, Faculty of Medicine, General Hospital, University of Southampton, Southampton SO16 6YD, UK

[2] Centre for Proteomic Research, Institute for Life Sciences, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK

[3] Clinical and Experimental Sciences Unit, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK

[4] Medical and Molecular Cancer Sciences, Paterson Building, Manchester Cancer Research Centre, Faculty of Biology, Medicine and Health, Manchester Academic Health Sciences Centre, University of Manchester, Manchester, M20 4BX

*Corresponding author: Mark S. Cragg. University of Southampton, Southampton General Hospital, Southampton, SO16 6YD, UK (Fax: +44 (0) 23 80704061; e-mail: msc@soton.ac.uk)

Running title: Cellular and plasma proteomics of B-cell tumours

Keywords: Eμ-*myc*, Eμ-*TCL1*, B-cell, oncogenesis, plasma, proteomics, biomarkers, lymphoma, leukaemia

Abbreviations

BL	Burkitt's lymphoma
CLL	Chronic lymphocytic leukaemia
DEP	Differentially expressed protein
Eμ	μ Ig heavy chain enhancer
Ig	Immunoglobulin
iTRAQ	Isobaric tags for relative and absolute quantitation
Rs	Regulation score (mean of $\log_2$(ratios)/(SD of $\log_2$(ratios) +1))
SEC	Size exclusion chromatography
SuPrE	Sub-proteome enrichment
TMT	Tandem mass tags
WT	Wildtype

Abstract

Approximately 800,000 leukaemia and lymphoma cases are diagnosed worldwide each year. Burkitt's lymphoma (BL) and chronic lymphocytic leukaemia (CLL), are examples of contrasting B-cell cancers; BL is a highly aggressive lymphoid tumour, frequently affecting children, whilst CLL typically presents as an indolent, slow-progressing leukaemia affecting the elderly. The B-cell-specific over-expression of oncogenes in Eμ-*myc* and Eμ-*TCL1* mice induce spontaneous malignancies modelling BL and CLL, respectively. Quantitative mass spectrometry proteomics and isobaric labelling were employed to examine the biology underpinning these contrasting B-cell tumours. Additionally, the plasma proteome was evaluated using sub-proteome enrichment to interrogate biomarker emergence and the systemic effects of tumour burden. Over 10,000 proteins were identified ($q<0.01$) of which 8270 cellular and 2095 plasma proteins were fully profiled with over 30% of each proteome demonstrating significant differential abundance for either tumour. Amongst the over-expressed B-cell tumour proteins were 126 potential therapeutic targets, 104 cell-surface proteins and >30 significantly enriched canonical pathways. Both tumour models highlighted over-expressed signatures of ribosome biogenesis, translation, cell-cycle promotion and chromosome segregation. Eμ-*myc* tumours specifically over-expressed several methylating enzymes and under-expressed many cytoskeletal components. Eμ-*TCL1* tumours specifically over-expressed ER stress response proteins and signalling components in addition to both subunits of the interleukin-5 (IL5) receptor. IL-5 treatment promoted Eμ-*TCL1* tumour proliferation, suggesting an amplification of IL5-induced AKT signalling by TCL1. Tumour plasma highlighted a striking tumour lysis signature, most prominent in Eμ-*myc* plasma, whilst Eμ-*TCL1* plasma contained dominant signatures of immune-response, inflammation, microenvironment interactions and displayed putative biomarkers of early-stage cancer. Integrated evaluation of tumour cell and plasma proteomics provided systemic insight not captured in isolation. This provided mechanistic insight into the biology behind the emergence of cancer biomarkers, suggesting that early- and late-stage cancers dominantly exhibit systemic and tumour lysis signatures, respectively. Overall, these findings provide a detailed characterisation of two contrasting B-cell tumour models, identifying common and specific profiles in both tumour cells and plasma.

Introduction

Burkitt's lymphoma (BL) and chronic lymphocytic leukaemia (CLL) represent the extremes of B cell cancers; BL is highly aggressive and frequently affects children, whilst CLL typically presents as an indolent, slow-progressing leukaemia in the elderly (1, 2).

BL is a hallmark myc-driven tumour; induced by chromosomal translocation of the transcription factor *myc* to immunoglobulin (Ig) enhancers (3). Myc influences approximately 15% of the genome, regulating processes such as cell proliferation, metabolism, adhesion, angiogenesis and de-differentiation (4-6). Such neoplastic-like traits make *myc* an aggressive oncogene, dysregulated or overexpressed in most cancers (7). Formal proof of its oncogenic properties were demonstrated when human myc was placed into the μ Ig heavy chain enhancer (Eμ) region of the mouse. These 'Eμ-*myc'* mice produced high-penetrance lymphomas within around 100 days, recapitulating several molecular and pathological aspects of BL (3, 8-10).

In contrast, CLL is a relatively indolent B-cell cancer, presenting with a $CD5^+CD19^+$ leukaemia (11, 12). Up to 90% of CLL cases express TCL1 – a protein involved in lymphocyte development (13-15). *TCL1* is suggested to promote cell survival and proliferation by amplification of AKT phosphorylation, induced by growth factor-, cytokine- and B-cell receptor-induced PI3K signalling (16, 17). The Eμ-*TCL1* mouse, overexpressing *TCL1* in B-cells again through the μ enhancer, was developed as a potential model of CLL (18). Eμ-*TCL1* mice recapitulate the expanded bone marrow, splenic and circulatory $CD5^+$ B-cell populations of CLL, detectable from around 3-5 months of age. The disease progresses to lethal splenomegaly and leukaemia at approximately 12 months (18, 19). Currently it is regarded as one of the most useful pre-clinical models of CLL (20, 21).

Together, these models provide the opportunity to study oncogenesis mediated by two contrasting oncogenes at opposing ends of the proliferative spectrum. Eμ-*myc* tumours, like BL, are highly aggressive, have rapid onset and form lymphoid tumours, while Eμ-*TCL1* mice, like CLL, present a relatively indolent and slow-developing leukaemia with secondary lymphoid organ involvement (3, 18).

In addition to cellular characterisations of tumours, plasma analysis holds the potential to identify additional biological signatures arising during oncogenesis. In particular, extracellular fluids can provide insight into the tumour-host dialogue between the immune system and micro-environment and report on the metabolic and homeostatic aberrations that tumours display. Further, combining cellular and plasma characterisation can provide greater insight into the mechanisms by which any biomarkers of disease may be entering the circulation.

Liquid chromatography coupled with mass spectrometry (LC-MS) proteomics currently provides the best means of establishing global differential protein expression profiles of cellular and plasma samples. These approaches are evolving rapidly and yielding ever-increasing proteome coverage (22, 23). Nonetheless plasma/serum proteomics still presents significant challenges due to the vast dynamic range of protein concentrations (24). While advances continue to be made with immunodepletion strategies (25-27), an alternative method termed, Sub-Proteome Enrichment by Size Exclusion Chromatography (SuPrE-SEC), offers an effective alternative. SuPrE-SEC uses size-dependent protein fractionation to deplete high-abundance proteins and enrich for low-abundance proteins. The resulting reduction of the protein concentration dynamic range facilitates a greater depth of LC-MS proteomics coverage (28, 29).

This investigation has applied multiplexed, LC-MS proteomics to the characterisation of plasma sub-proteomes and isolated B-cell material from Eμ-*myc*, Eμ-*TCL1* and WT mice. The resulting proteomes have been interrogated to identify common and tumour-specific signatures and to understand the combined cellular and extracellular characteristics of B-cell tumours.

Methods

Materials

Tris(hydroxymethyl)aminomethane (TRIS), SDS, Na_2EDTA, NH_4Cl, $NaHCO_3$, sodium deoxycholate (DOC), guanidine hydrochloride, glycine, HPLC and LC-MS grade ACN and formic acid (FA) and 100 μm cell sieves were purchased from Fisher Scientific. Tween20 (tween), sodium heparin, propidium iodide (PI), carboxyfluorescein succinimidyl ester (CFSE), asparagine, 2-

mercaptoethanol, octylphenoxypolyethoxyethanol (IgePalCA^{630}), triton x-100, protease inhibitors, ponceau S, acetic acid, methyl methanethiosulfonate (MMTS), tris(2-carboxyethyl)phosphine (TCEP), triethylammonium bicarbonate (TEAB), DMSO, hydroxylamine and ammonium hydroxide (NH_4OH) were purchased from Sigma.

Dulbecco's modified Eagle's medium (DMEM), glutamine, pyruvate, penicillin and streptomycin were purchased from Life Technologies. Mouse B-cell isolation kits and magnetic cell sorting columns were purchased from Miltenyi Biotech, Germany. 23 gauge needles and 1 ml syringes were purchased from BD Biosciences, Interleukin 5 (IL5) recombinant protein from Peprotech, Immobilon 0.45 µm pore PVDF membrane from Millipore and non-fat milk from Marvel. 30x DTT and 3x red loading buffer were purchased for Cell Signalling Technologies. Enhanced chemiluminescence reagents, tandem mass tags (TMT) 10-plex isobaric labelling reagents and 2 kDa M_w cut-off Slide-A-Lyzer dialysis cassettes and proteomics grade lys-c were purchased from Thermo Scientific. Proteomics grade trypsin was purchased from Roche. Isobaric tags for relative and absolute quantitation (iTRAQ) 8-plex reagents were purchased from ABSciex. Antibodies are detailed in **Supplementary methods**.

Animals

Mice were bred and maintained in-house with procedures carried out in accordance with home office licences PPL30/2450 and 30/2970 and PIL30/9925. Female Eµ-*myc* [C57BL/6J-TgN(Igh*Myc*)22Bri/J] hemizygous and Eµ-*TCL1* [C57BL/6J-TgN(Igh*TCL1*)22Bri/J] hemizygous mice were used. The Eµ-*myc* and Eµ-*TCL1* transgenes were detected with PCR. Eµ-*myc* primers (annealing temperature; 55°C): 5'- CAG CTG GCG TAA TAG CGA AGA G -3' and 5'- CTG TGA CTG GTG AGT ACT CAA CC -3' ~900 bp product, Eµ-*TCL1* primers (annealing temperature; 58°C): 5'- GCC GAG TGC CCG ACA CTC -3' and 5'- CAT CTG GCA GCA GCT CGA -3' ~250 bp product. Eµ-*TCL1* mice were screened monthly for count and percentage of $B220^+$ $CD5^+$ B-cells in the blood, being considered terminal either with >80% leukaemic cells or palpated splenomegaly >3 cm. Eµ-*myc* mice were checked daily and were considered terminal, typically upon visible lymph node tumour presentation.

Experimental Design and Statistical Rationale

Sample pooling, detailed in **Figures 1C** and **S1,** was used to accommodate comparative experiments in single isobaric label sets. Tumourous spleens were collected from 4 Eμ-*myc* and 4 Eμ-*TCL1* mice with terminal tumour presentation. For non-tumour controls, 6 spleens, handled as two pools of 3 spleens, were collected from 6 week old and 200 day wildtype (WT) littermates and 6 week-old Eμ-*myc* and Eμ-*TCL1* mice with no signs of tumour or splenomegaly (summarised in **Figure 1** and **S1**). Non-tumour pools were analysed with 6 samples per isobaric label, while tumour samples were analysed as biological replicate pools of two tumours.

For plasma, 6 samples were individually collected for each of the above tumour and non-tumour conditions, in parallel with B-cell isolation. In addition, 6 samples were isolated from Eμ-*TCL1* mice at a pre-terminal '30%' stage, when 30% $B220^+$ $CD5^+$ cells were present in the blood. To minimise bias these were selected from a cohort of 14 Eμ-*TCL1* mice, from every other mouse reaching the 30% threshold. Tumour and pre-terminal '30%' tumour plasma were also analysed as biological replicates, with 3 samples allocated to each label.

Each pair of tumour pools was compared with two WT sample pools, generating four ratios which could be evaluated for consistency and significance to ensure the reproducibility of quantitative results. Replicate ratios were analysed for consistency with an FDR-corrected t-test and considered significant with a p-value of <0.05. P-values were coupled with a measure of magnitude detailed in 'Quantitative and Statistical Analysis of MS Data', below.

B-cell Isolation

B-cells were isolated from single-cell suspensions using the mouse negative selection B-cell isolation kits and MACS columns according to the manufacturer's instructions, washed once at room temperature (RT) in red blood cell lysis buffer (155 mM NH_4CL, 10 mM $NaHCO_3$, 0.1 mM Na_2EDTA (pH 7.4) before 3 washes in PBS. B-cell isolation efficiency was assessed by flow cytometry immunostaining with CD19 and CD3, typically yielding >90% purity with less than 2% T-

cell ($CD3^+$) contamination (**Figure S1**). B-cell pellets were snap frozen and stored in liquid nitrogen prior to lysate preparation.

Plasma Isolation

To isolate plasma, maximise sample purity, minimise red blood cell lysis and accommodate for organ displacement from tumours, an adaption of bleeding from the inferior vena cava under terminal anaesthesia (30) was utilised. Using a 23 gauge needle, 700-1000 µl of blood was collected for each animal into 50 µg/ml sodium heparin in PBS. Blood was immediately placed on ice and centrifuged at 2000 g for 15 minutes at 4°C with plasma stored in liquid nitrogen. Samples were rejected if red blood cell lysis was visible by eye.

Plasma Sub-proteome Enrichment

20 µl was taken from each of the 6 replicate plasma samples described above to give pools of 120 µl. For tumour samples, two 120 µl pools of 3 x 40 µl were formed to provide biological replicates (summarised in **Figures 1** and **S1**).

SuPrE-SEC was adapted from that described previously (28). Each 120 µl plasma pool was diluted with 380 µl of 6 M guanidine hydrochloride in 10% methanol and separated by 3 KW804 SEC columns in series at 1.2 ml/min and 30°C. The low molecular weight sub-proteome was isolated during elution from 42-55 minutes (**Figure 1D**) and dialysed (2 kDa M_w cut-off) into ultrapure water (18.2MΩ cm^{-1}) with 5 exchanges into 5 litres. Protein was lyophilised and re-solubilised in 0.5 M TEAB with 0.05% SDS with 30 µg digested and labelled for MS analysis as described below.

Sample Preparation for MS

Snap frozen cell pellets were lysed on ice by trituration with a 23 gauge needle in 0.5 M TEAB with 0.05% SDS. Disrupted cells were further sonicated and lysates cleared at 16,000 g for 10 minutes at 4°C. 100 µg of cell lysate or 30 µg of plasma sub-proteome was reduced with 50 mM TCEP and alkylated with 200 mM MMTS, before digestion overnight at RT with a 30:1 ratio of proteomics grade trypsin. Plasma proteins were additionally digested for a further 2 hours at 37°C

with 100:1 proteomics grade Lys-c. Peptides were incubated with either iTRAQ 8-plex (B-cell peptides) or TMT 10-plex isobaric tags (plasma peptides) according to the manufacturer's instructions. Samples were then lyophilised and labelled peptides serially reconstituted, for each proteome, in 100 µl of 2% v/v ACN, 0.1% v/v NH_4OH.

Peptide Pre-fractionation

Peptides were resolved using high-pH (0.1% v/v NH_4OH) RP C8 chromatography (150 mm x 3 mm ID x 3.5 µm particle, XBridge, Waters) at 300 µl/min with a LC-20AD HPLC system (Shimadzu) maintained at 30°C, using the mobile phases (MP); A – 99.9% H_2O, 0.1% NH_4OH, B – 99.9% ACN, 0.1% NH_4OH. The two hour gradient was as follows; 0 minutes; 2% B, 10 minutes; 2% B, 75 minutes; 30% B, 105 minutes; 85% B, 120 minutes; 2% B. Fractions were collected in a peak-dependent manner and lyophilised.

Peptide Fraction Resolution and Characterisation by LC-MS/MS

Lyophilized peptide fractions were individually reconstituted in 2% ACN, 0.1% FA and ~500 ng of peptides loaded by a Dionex Ultimate 3000 (Thermo Scientific). Conditions used varied between the plasma and B-cell proteomes, detailed in supplementary methods. In summary, peptides were trapped by C18 and eluted over a reverse phase gradient, of which several lengths were used depending on the peptide fraction abundance and proteome. The total MS time for the iTRAQ-labelled B-cell proteome was ~200 hours, with the TMT-labelled plasma proteome analysed over ~250 hours.

Peptide elution was directly coupled to electrospray ionisation at 2.4 kV using a PicoTip nESI emitter (New Objective), and were characterised with an Orbitrap Elite Velos Pro mass spectrometer (Thermo Scientific). MS characterisation of eluting peptides was conducted between 350 and 1900 m/z at 120,000 mass resolution. The top 12 +2 and +3 precursor ions per MS scan (minimum intensity 1000) were characterised by tandem MS with high-energy collisional dissociation (HCD) (30,000 mass resolution (15,000 for the B-cell proteome), 1.2 Da isolation window, 40 keV normalised

collision energy) and CID (ion trap MS, 2 Da isolation window, 35 keV). Additionally the DMSO ion at 401.922718 was used as a MS lockmass (31).

MS Data Processing

Target-decoy searching of raw spectra data was performed with Proteome Discoverer software version 1.4.1.14 (Thermo Scientific). Spectra were subject to a two stage search, both using SequestHT (version 1.1.1.11), with Percolator used to estimate FDR with a threshold of $q \leq 0.01$. The first allowed only a single missed cleavage, minimum peptide length of 7, precursor mass tolerance of 5 ppm, no variable modifications and searched against the mouse UniProt Swissprot database (downloaded 01/15), supplemented with the human sequences for myc and TCL1. The second search used only spectra with $q > 0.01$ from the first search, allowed 2 missed cleavages, minimum peptide length of 6, searched against the human/mouse UniProt trembl database (downloaded 01/15), precursor mass tolerance of 10 ppm and a maximum of 2 variable (1 equal) modifications of; TMT or iTRAQ (Y), oxidation (M), deamidation (N,Q) or phospho (S,T,Y). In both searches, fragment ion mass tolerances of 0.02 Da and 0.5 Da were used for HCD and CID spectra, respectively. Fixed modifications of Methythio (C), TMT or iTRAQ (K and N-terminus) were used. PhosphoRS was used to predict the probability of specific phosphorylated residues. Reporter ion intensities were extracted from non-redundant PSMs with a tolerance of 20 ppm. To reduce ratio compression, peptide spectrum match data for proteins ($q<0.01$) were exported from Proteome Discoverer and submitted to Statistical Processing for Isobaric Quantitation Evaluation (SPIQuE) at spiquetool.com. This method weighted the contributions of each PSM quantitation to a protein's quantitation on the basis of PSM features (manuscript in preparation). For example, high-intensity peptides with low isolation interference were given a greater weighting factor. While the effects upon the ratios were minimal, the overall trend demonstrated efficient ratio decompression (**Supplementary methods**). An example of the effect on the ratio decompression to the human TCL1 protein, discretely expressed in Eμ-*TCL1* B-cells, is outlined in **Supplementary methods**.

The raw data and processed outputs have been deposited to the ProteomeXchange Consortium (32) via the PRIDE partner repository with the dataset identifier PXD004608.

Quantitative and Statistical Analysis of MS Data

Log_2 (ratios) were generated describing each sample pool relative to the two WT controls. To define those proteins with the greatest fold change for each biological state, an FDR-corrected one sample t-test was performed and average values determined. To derive a single, robust measure, representative of both the magnitude and consistency of differential expression, a ratio was defined between that of the mean and the standard deviation of the 4 log_2 ratios, termed the regulation score (Rs = mean/(standard deviation +1)). For proteins with no, or inconsistent differential expression, the Rs tends towards 0. The Rs also correlated more strongly with the p-value for each set of ratios **Supplementary methods**. In all instances, differentially expressed proteins (DEPs) were defined as those with an FDR-corrected p-value of <0.05 and an Rs threshold of >0.5 or <-0.5. While arbitrary, evaluation of this threshold demonstrated a minimum average fold change of 1.4, for which a fold change variation of just 6% was observed. This therefore fulfilled the power analysis principles described by Levin, 2011 (33), but on a case-by-case basis, by incorporating variation into each Rs via the SD as a denominator. The outliers for this threshold were additionally highlighted to demonstrate that consistent differential expression was still observed **Supplementary methods**. Other p-values were determined by a false discovery rate-corrected 1- or 2-sample, two tailed, t-test with no assumption of equal variance.

Commented [H1]: Adjusted for reviewer 2, point 3

Commented [CM2]: Should these not be in the sup methods files?? Or do you intend to include here also? If so remove sup methods details

Bioinformatics Analyses

Proteins reaching the thresholds outlined above were submitted to either Ingenuity Pathway Analysis (IPA) or Database for Annotation, Visualization and Integrated Discovery (DAVID). For DAVID analyses, the default settings were used for pathway and gene ontology (GO) term enrichment, with Benjamini-corrected p-values of <0.05 considered significant. For B-cell/tumour proteome enrichments, all quantified proteins were used as a background. For IPA analyses, default settings were used. Upstream regulator analysis was determined using all DEPs (Rs>0.5/<-0.5, p<0.05). Annotations of biomarkers and drug targets were conducted by IPA. Plasma protein deconvolution was conducted by removing any proteins annotated by IPA as predominantly cellular in addition to any proteins with more than 30 PSMs from the B-cell proteome.

SDS-PAGE and Western Blotting

Cells were lysed in radioimmunoprecipitation assay (RIPA) cell lysis buffer (0.15 M NaCl, 1% v/v octylphenoxypolyethoxyethanol (IgePalCA630), 0.5% w/v sodium deoxycholate (DOC), 0.1% w/v sodium dodecyl sulfate (SDS), 0.05 M TRIS (pH 8) and 1% v/v protease inhibitor) and as described previously (34). Lysate supernatant were isolated by centrifugation (16,000 g for 10 minutes at 4°C). Protein concentrations were determined by Bradford assay. Lysates were reduced in loading buffer with 2-mercaptoethanol at 95°C for 5 minutes. Lysates were resolved by SDS-PAGE and transferred to PVDF. Membranes were blocked in 5% (w/v) non-fat milk for 1 hour before probing with primary antibodies, as detailed in **Supplementary methods**. Expression was detected by incubation with a horseradish peroxidase-conjugated secondary antibody using enhanced chemiluminescence reagents and a ChemiDoc-It imaging system (UVP). GAPDH or tubulin were used as loading controls. Relative quantification of protein band intensity was determined using Image J, normalised against loading controls and ratios to WT B-cell lysates were derived. For Eμ-*myc* tumour validations, lysate protein concentrations were determined by bicinchoninic acid assay (BCA), reduced with DDT and diluted in 3x loading buffer. Western blots were visualised with fluorescent antibodies and an Odyssey Imaging System (Li-cor) and quantified using Image Studio 2.0 (Li-cor).

Cell Culture

Cells derived from Eμ-*myc* or Eμ-*TCL1* tumours were cultured in DMEM supplemented with 2 mM glutamine, 1 mM pyruvate, 45 units/ml penicillin, 45 μl/ml streptomycin, 200 μM asparagine, 50 μM 2-mercaptoethanol and 10% FCS. Eμ-*myc* tumours were cultured a density of $5x10^7$ cells/ml at 37°C in 10% CO_2. Eμ-*TCL1* cells were cultured at a density of $5x10^6$ cells/ml at 37°C in 5% CO_2. Cells were cultured for 24 hours prior to treatment. Unless otherwise specified, IL5 was used at a concentration of 100 ng/ml. Serum starvation was conducted for 4 hours in the above media with 0.5% BSA replacing FCS.

Flow Cytometry

Cells were stained with either the manufacturer's recommended concentration, or 10 µg/ml, of antibody (**Supplementary methods)** for 30 minutes in the dark, washed and analysed by flow cytometry with a FACScan or FACScalibur (BD) (35, 36). Relative expression was determined using the geometric means. Cell cycle status was assessed by hypotonic PI (50 µg/ml PI, 0.1% w/v sodium citrate, 0.1% w/v triton x-100) incubated at 4°C for 15 minutes. PI fluorescence was measured by FL2 on linear scale at the lowest flow rate. Cell division was tracked by CFSE dilution of cells stained with a concentration of 5 µM for 15 minutes at room temperature.

Results

Quantitative proteomics of Eµ-*myc* and Eµ-*TCL1* tumours and plasma

To characterise the global proteome expression of Eµ-*myc* and Eµ-*TCL1* tumours, an 8-plex iTRAQ experiment was designed incorporating terminal tumour, pre-malignant and age-matched WT B-cells (**Figure 1**). Sample pooling in combination with biological replicates enabled the averaging of biological variability of several samples within a single isobaric tag experiment. Splenic tumours samples, due to a greater anticipated variability, were pooled in two pairs of two tumours from each model; derived from a total of 8 terminal mice (**Figure S1**). Non-tumour controls consisted of samples pooled from 6 animals. For plasma proteomics, controls were again characterised as pools of 6 samples, with 2 pools of 3 plasma samples for each tumour condition. Additionally, plasma from pre-terminal Eµ-*TCL1* tumours was analysed, derived from mice reaching a $CD5^+$ $B220^+$ leukaemia threshold of 30%, termed '30%' samples. To accommodate these additional samples, TMT 10-plex isobaric labels were used.

The tumours selected for proteomics analysis, detailed in **Figure S1A-F**, were representative of the range of characteristics observed for the model. The median terminal presentations, for example (**Figure S1A**), of 92 and 318 days for Eµ-*myc* and Eµ-*TCL1* tumours, respectively, agreed with earlier reports (3, 8, 18, 19).

Splenic B-cells from tumours and controls were lysed, trypsin digested and assigned to the 8 isobaric labels of iTRAQ 8-plex (**Figure 1**) to provide relative protein expression quantitation. The labelled B-cell peptides were pooled and characterised by two-dimensional (2D) LC-MS identifying 9260 proteins ($q<0.01$). 8270 proteins were relatively quantitated across all 8 sample pools. The depth of proteome coverage and quantitation is summarised in **Figure S1H**.

Plasma samples were subjected to an adapted form of SuPrE-SEC (28), reducing the dynamic range and facilitating deeper proteome coverage by the exclusion of the majority of high-abundant proteins (**Table S1, Figure S1G**). 2D LC-MS characterisation of the plasma sub-proteome provided relative quantitation for 2095 of the 2688 identified proteins ($q<0.01$). Peptide spectrum matches (PSMs) to the higher-M_w protein albumin (n=3835, M_w=68.7kDa), for example, were of far lower abundance than transthyretin (n=7769, M_w=15.8kDa) or apolipoprotein A-II (n=7129, M_w=11.3kDa). Given approximate plasma concentrations of 40mg/ml, 260µg/ml and 300µg/ml for these proteins (37), respectively, this represented around a 275-fold enrichment of low M_w proteins.

Reproducibility and validation of quantitative proteomics results

The relative quantitative proteomes demonstrated tumour-dependent hierarchical clustering identifying common trends of differential tumour expression compared to both WT controls (**Figure 2A**). Evaluation of the reproducibility between tumour sample pools by linear regression highlighted the strong correlation between the 5 pairs of independent biological replicates (**Figure 2B**). Additionally, an overlap of 1288 proteins were quantitated in both the B-cell and plasma proteomes.

Individual protein expression was considered for a number of candidates to confirm the successful characterisation and analysis of the quantitative proteome. The results highlighted the anticipated model-specific over-expression of the *myc* and *TCL1* human transgenes (**Figure 2C**, **S2A**,**B**), in addition to characteristic CD5 over-expression and B220 under-expression for Eµ-*TCL1* tumours. Several DEPs were validated by antibody-based methods to further verify the relative quantitative accuracy of the proteomics; 10 for Eµ-*myc* tumours and 7 for Eµ-*TCL1* tumours (**Figure S2G,H**). The $\log_2$ ratios to WT of these validations were plotted alongside the quantitative proteomics

results (**Figure 2D**). Additionally, previously reported DEPs in Eμ-*myc* or Eμ-*TCL1* tumours were compared to the proteomics results (**Figure S2D**,**F**). A correlation was also observed with a study describing relative mRNA expression in Eμ-*myc* tumours (38) **(Figure S2J).** Overall, the quantitative proteomics correctly identified over- or under-expression for the vast majority of the 38 DEPs observed by antibody-based quantitations. This comparison also highlighted reporter ion ratio compression, a commonly reported feature of isobaric tag quantitation (39, 40). Anticipated plasma proteins, such as IgM over-abundance in Eμ-*TCL1* tumours, was also observed (**Figure S2I**). These results therefore supported the overall reliability of the DEP observations and the characterisation and analysis by which they were generated.

Proteomic characterisation of B-cell tumour phenotypes

Proteins were considered DEPs when both tumour pools demonstrated a clear, consistent over- or under-expression to both 6-week and 200-day WT samples. Rather than an average of the four derived ratios for each tumour, potentially misrepresentative of variable findings, a ratio of average to standard deviation (mean/(SD+1)) was calculated, termed the regulation score (Rs). This provided a single value expressive of both magnitude and consistency for either, or both, tumours relative to WT B-cells. In combination with FDR-corrected p-values, the Rs allowed careful selection of the most confidently and consistently DEPs (detailed in **Supplementary methods**).

The range of significantly DEPs for each tumour were represented by volcano plot, comparing Rs to $-\log_{10}$ (p-values) (**Figure 3A, S3A**). This analysis highlighted the broad extent of DEP signatures observed in each tumour, with approximately 3000 and 1500 DEPs in Eμ-*myc* and Eμ-*TCL1* tumours, respectively. The 6 week Eμ-*myc* and Eμ-*TCL1* B-cells exhibited signatures broadly indistinguishable from terminal Eμ-*myc* tumours and WT B-cells, respectively (**Figure S3B**). When compared to one another, a common signature became apparent between Eμ-*myc* and Eμ-*TCL1* tumours (**Figure 3B**) demonstrating that many of the Eμ-*myc* tumour-over-expressed proteins were also over-expressed in Eμ-*TCL1* tumours, but to a lesser magnitude. Approximately 700 and 200 proteins were considered over- and under-expressed in the tumours of both models, respectively (**Figure 3C**). Examples of the 10 most consistently DEPs from this common signature were

Commented [H3]: Included to address reviewer 2, point 1

Commented [CM4R3]: But weren't they weaker?? If so adda sentence to say that

highlighted (**Figure 3D**); and annotated as predicted cell surface markers with potential as immunotherapy targets; drug targets, annotated by IPA; and novel DEP with no previous published links with cancer (detailed further in **Figure S3C-G**). Amongst the DEP common to both tumours were, for instance signatures of cell cycle upregulation, such as the three kinesin proteins, KIF11, KIF20A and KIF23 and an overall trend of cell surface protein under-expression such as CD23, CR2, CD200, CD40, IgG receptor FCGRT, CD38, CD22 and IL21R. The most consistently over-expressed surface proteins, HMMR/CD168, has previously been associated with B-cell cancers (41, 42).

Bioinformatics reveals signatures common to contrasting B-cell tumours

While individual DEPs can suggest functional insight, approaches simultaneously considering all DEPs, potentially offer a broader understanding of biological mechanisms. Accordingly, topological, proteome-wide expression patterns were investigated by bioinformatics.

Gene ontology (GO) term enrichment was used to identify processes overrepresented by the DEPs observed in both tumours (**Figure 4A, Table S3**). Amongst the over-expressed proteins GO term enrichment identified strong signatures of mechanisms relating to cell growth and proliferation, including; 'ribosome biogenesis' (n=106, $p=1.8x10^{-27}$), 'translation' (n=135, $p=1.2x10^{-16}$), 'chromosome segregation' (n=81, $p=6.0x10^{-14}$) and 'cell cycle' (n=229, $p=2.5x10^{-11}$). Amongst the under-expressed proteins were trends of several immune function-related terms, such as lymphocyte activation (n=41, $p=8.8x10^{-7}$), immunoglobulin mediated immune response (n=15, $p=2.1x10^{-5}$) and BCR signalling (n=11, $p=2.9x10^{-4}$). Additionally, several processes related to immune evasion, differentiation and growth inhibition were enriched.

To visualise these processes and highlight the proteins contributing to the enriched GO terms, networks were displayed descriptive of individual proteins and their inter-relationships, defined by STRING (43). **Figure 4B** demonstrates a network generated from proteins annotated with 'chromosome segregation' highlighting the scale and complexity of the dysregulation of this process seen in in both tumours. 'Ribosome biogenesis' was also illustrated in this way (**Figure 4C**), highlighting those over-expressed proteins responsible for the term's enrichment. A cluster of

interacting ribosome proteins was identified, alongside several assembly regulators. Several canonical pathways, including ribosome-, translation-, and cell cycle-related pathways demonstrated significant enrichment in both tumours (**Figure S4A**). A network detailing all under-expressed proteins revealed a highly inter-related series of interactions highlighting major histocompatibility complex proteins, interferon response proteins and B-cell-related signalling molecules (**Figure 4D**).

To simultaneously analyse both proteome-wide over- and under-expression, upstream regulators were evaluated using IPA. Upstream regulator activity was inferred by the comparison of anticipated downstream expression profiles of proteins with those expression profiles derived by quantitative proteomics for both tumours (**Table S3, Figure S4B-E**). Protein functionality was then inferred from a combination of the resulting upstream regulator activation z-scores – a value proportional to overall predicted activation (44) – and the proteomics-determined differential abundance (**Figure 4E**). Several regulators not quantitated by proteomics, including miRNA were also inferred to be activated or inactivated (**Figure S4C**,**E**).

Myc over-expression and inferred activation (Rs=1.11, z=8.09, $p=6.6x10^{-26}$), alongside other proliferative drivers and oncogenes (eg. E2F3, MYCN, RABL6 (Rs>0.5, z>4)) described several key, functional regulators influencing the neoplastic B-cell phenotype (**Figure 5I**, **Table S3**). The combined over-expression and inferred inactivation of TP53 (Rs=0.97, z=-3.76, $p=1.6x10^{-34}$) and retinoblastoma 1 (RB1) (Rs=0.59, z=-2.86, $p=2.3x10^{-21}$), represented an anticipated evasion of tumour suppressor pathways. Interestingly, RB1-like 1 (RBL1) (Rs=0.92, z=-4.47, $p=2.2x10^{-16}$), a lesser-established tumour suppressor, had greater over-expression and inferred inhibition than its better-studied family member.

It was additionally possible to compare differential protein phosphorylation to infer activation or inhibition (**Figure S4F-H**). HDAC2 over-expression in Eμ-*myc* tumours, for instance, was accompanied by a signature of downstream inactivation (**Figure S4B**) and a decreased phosphorylation of S394 (**Figure S4H**), a modification previously associated with the activation of HDAC2 **(45)**. Similarly, in Eμ-*TCL1* tumours, RBL1 over-expression was accompanied by apparent inactivation and decreased T385 phosphorylation. Differential phosphorylation of BCR signalling

pathway components demonstrated an increased phosphorylation of downstream proteins such as NF-κB, OCT2 and ETS1 (**Figure S4J**).

Characteristics specific to Eμ-*myc* B-cell tumours

While the common tumour signature in both models was strong, several tumour-specific differences were also apparent, particularly in the more aggressive Eμ-*myc* tumours. To evaluate tumour-specific expression, proteins were filtered to include those which exhibited significant over- or under-expression in one tumour model ($Rs>0.5/<-0.5$, $p<0.05$) but not the other ($Rs>0.25/>-0.25$). This identified 572 and 537 proteins with specific over- or under-expression in Eμ-*myc* tumours, respectively. The 10 most over- and under-expressed proteins in addition to cell surface proteins were illustrated to highlight examples of this specific differential expression pattern in Eμ-*myc* tumours (**Figure 5A**).

The over- and under-expressed proteins specific to Eμ-*myc* tumours were evaluated, as for Figure 4A, for GO term enrichment (**Figure 5B**, **S3**). The majority of GO terms observed for both tumours were also observed for the Eμ-*myc*-specific proteins, further highlighting trends of increased proliferation and growth of cells. Exceptions included trends of epigenetic-related processes and under-expressed cytoskeletal processes.

Most notably specific to over-expressed Eμ-*myc* tumour proteins was the enrichment of the term 'methylation' ($n=50$, $p=0.03$). A network was formed from these 50 proteins (**Figure 5C**), highlighting over-expressed methylation enzymes relating to gene expression; both epigenetically (histone methylation, $n=18$) and post-transcriptionally (RNA methylation, $n=10$). CpG methylation appeared widely dysregulated with 2/3 methylation (DNMT1, DNMT3B) and 2/3 demethylation (TET2, TET3) enzymes over-expressed.

Eμ-*myc* tumours demonstrated a broad downregulation of processes relating to the cytoskeleton, for example, 'actin cytoskeleton organisation' ($n=49$, $p=1.2\text{x}10^{-6}$), as illustrated in a network (**Figure 5D**). This highlighted the under-expression of 6 actin-related protein 2/3 complex subunits (ARPC) and 3 coronin proteins. Three inter-related proteins from this network, coronin 1a

(COR1A), lymphocyte cytosolic protein 1 (LCP1) and actin regulatory protein CAPG, were evaluated for relative expression in Eμ-myc tumours, relative to non-tumour B-cells by Western blot (**Figure 5E**). In each instance, under-expression was observed in line with that of the quantitative proteomics findings (compared in Figure 2D). Additionally, myc over-expression was evaluated alongside two further cytoskeleton-related proteins; a broadly functioning regulator of cytoskeletal activity, myosin-9 (MYH9) and a cytoskeletal-membrane junction protein, moesin (MOE) again demonstrating anticipated differential expression.

Characteristics specific to Eμ-*TCL1* B-cell tumours

Protein expression specific to Eμ-*TCL1* tumours was also evaluated, with the top 10 DEPs and membrane proteins presented (**Figure 6A**). As validation this highlighted the discrete expression of human TCL1 (**Figure S2A**) in Eμ-*TCL1*-derived B-cell samples. GO term enrichment identified upregulated processes relating to intracellular compartments, such as ER stress, vesicular transport and glycosylation. Extracellular interactions, including signalling, locomotion and adhesion appeared as the strongest under-expressed process in Eμ-*TCL1* tumours. A network formed from Eμ-*TCL1*-specific over-expressed proteins (**Figure 6C**) highlighted clusters of associated proteins contributing to enriched GO terms, such as ER and golgi proteins. Several signalling proteins were also specifically over-expressed, such as phosphatidylinositol kinases and phosphatases. Associated with these signalling proteins were the alpha and beta subunits of the interleukin 5 receptor (IL5Rα and IL5Rβ), also apparent amongst the top 10 membrane proteins highlighted in **Figure 6A**.

Interleukin 5 receptor is over-expressed by Eμ-*TCL1* B-cell tumours

The over-expression of the IL5R was first evaluated by plotting the individual iTRAQ ratios for each unique peptide matching IL5Rα and IL5Rβ (**Figure 6D**, **S5A-E**). For both IL5R subunits a clear, specific over-expression was apparent. Furthermore, the signature for IL5 activity was apparent through various bioinformatics analyses, including IPA regulator analyses (**Figure S4C**, **S5F**). A signature additionally became apparent in Eμ-*TCL1* terminal and pre-terminal plasma indicative of an

IL5-induced proliferation signature, as well as plasma signatures of quantitated proteins and inferred regulators potentially upstream of IL5 induction (**Figure S5G-I**).

Given the strength of this evidence, the surface expression of IL5Rα, the subunit specific to IL5 recognition, was evaluated by flow cytometry for terminal and pre-tumour Eμ-*TCL1* B-cells, relative to WT B-cells (**Figure 6E, S5J-K**). While IL5Rα expression was observed across a wide distribution for WT B-cells, for pre-tumour Eμ-*TCL1* B-cells, the emerging $CD5^+$ population was observed with high IL5Rα expression. The Eμ-*TCL1* tumour cells exhibited the same $CD5^+IL5R\alpha^+$ expression as the emerging pre-tumour Eμ-*TCL1* B-cell population.

Interleukin 5 drives proliferation in Eμ-*TCL1* B-cell tumours via AKT

The observation of substantial over-expression of both IL5Rα and IL5Rβ, strongly suggested a functional role for the IL5R and IL5 in driving Eμ-*TCL1* tumours. To investigate and functionally validate the IL5R over-expression, terminal tumours from Eμ-*TCL1* mice were treated with 0, 10 and 100 ng/ml of IL5 for 48 hours *in vitro*. Firstly, cell density indicated a dose-dependent expansion induced by IL5 (**Figure 7A**). CFSE labelling confirmed that cells were proliferating in a dose-dependent manner with as many as 70% of cells being post-mitotic after 48h of 100 ng/ml IL5 treatment (**Figure 7B**). Furthermore, evaluation of cell cycle phases by hypotonic PI staining demonstrated that almost three times as many cells (14.5%) were in $S/G_2/M$ phase after 100 ng/ml IL5 treatment, compared with no treatment (4.9%) (**Figure 7C**).

Given the proposed role of TCL1 in amplifying AKT signalling (16, 17), AKT activation by IL5 was investigated (**Figure 7D**). Serum starvation and the subsequent addition of IL5 induced AKT phosphorylation in a dose-dependent manner at doses of greater than 5 ng/ml. Downstream signalling was also apparent, with dose-dependent phosphorylation of the S6 ribosomal protein and S6 kinase, indicative of mTOR activation.

Plasma proteomics reveals signatures of tumour lysis and immune response

To investigate the overall plasma proteome signature, those proteins quantified commonly or discretely within the B-cell and plasma proteomes (**Table S2**) were dissected based on canonical

protein localisations (**Figure 8A**). The vast majority (82%) of the 1288 proteins observed as common to both proteomes, were annotated with a canonical cellular localisation, compared to just 40% of proteins identified discretely in the plasma proteome. Furthermore, evaluation of the relative abundances in terminal tumour plasma revealed a striking signature of over-abundant cell-derived proteins (**Figure 8B**). The tumour origin of this plasma signature was additionally indicated by GO term enrichment (**Figure 8C**), with highly similar trends to that of the B-cell tumours (**Figure 4A**), and an inter-proteome correlation between the approximate protein abundances (**Figure S6A**). Additionally proteins were highlighted with biomarker applications (**Figure S6H**), specific tumour plasma signatures (**Figure S6J-P**), differentially abundance relative to approximate relative plasma concentrations (**Figure S6Q**), over- and under-abundance in both the tumour and plasma proteomes (**Figure S6R**) and with potential as biomarkers derived from functional tumour regulators (**Figure S6S-T**).

To explore differential protein abundance resulting from the extrinsic response to the tumour, the plasma proteome was subject to deconvolution, removing proteins with a clear cellular origin. The remaining proteins, termed the 'lysis-free' plasma proteome (**Table S2**, **Figure S6F**), provided a signature descriptive of an immune response. Terms including 'defence response' ($n=16$, $p=4.0\times10^{-5}$) and 'extrinsic apoptotic signalling pathway via death domain receptors' ($n=6$, $p=1.4\times10^{-4}$) were identified as enriched amongst these over-abundant 'lysis-free' plasma proteins (**Figure 8C**). To detail this lysis-free signature further, a network was generated from all over-abundant lysis-free plasma proteins (**Figure 8D**). Protein clusters included chemotaxis regulators Ccl2, Ccl9 and Ccl21a, wound and inflammation response proteins, fibrinogens α, β and γ and the hyaluronan-binding, inter-alpha-trypsin inhibitor proteins, ITIH1-ITIH4.

Comparison between the terminal plasma signatures for each tumour model highlighted the dominance of this tumour lysis signature in Eμ-*myc* tumours, while Eμ-*TCL1* terminal plasma contained a dominant signature of extracellular proteins (**Figure 8B**). In both cases these signatures were observed for the opposing tumour, but to a lesser extent. GO term enrichment for over-abundant plasma proteins specific to Eμ-*myc* and Eμ-*TCL1* terminal tumours illustrated model-specific

signatures of tumour lysis and extracellular proteins, respectively. Eμ-*myc* tumour plasma specific GO terms had a strong resemblance to that of the B-cell tumours, highlighting cellular processes similar to both the tumours and the common plasma signature (**Figure 4A and 8C**). Over-abundant plasma proteins specific to terminal Eμ-*TCL1* tumours enriched for several GO terms related to immune processes, such as 'immune response' (n=17, $p=1.5x10^{-4}$) and 'lymphocyte activation' (n=12, $p=4.8x10^{-4}$).

Finally, consideration was given to proteins emerging in the pre-terminal '30%' Eμ-*TCL1* plasma (**Figure 8F**). Plasma proteins demonstrating a correlation between over-abundance and tumourigenesis were plotted, highlighting the extracellular immune-response proteins haptoglobin (Hp) and ITIH1, both also observed in the lysis free signature. Several additional proteins demonstrating over-abundance in pre-terminal '30%' Eμ-*TCL1* plasma are detailed in **Figure S6I**.

Discussion

B-cell tumours have been intensively investigated by genomics and transcriptomics in recent years, advancing clinical and biological understanding (46-49). Such information, is however sometimes limited, especially in a functional context, due to the weak correlations observed between mRNA and protein expression (50, 51). A comparison between our proteomics data and a former Eμ-*myc* tumour mRNA expression dataset **(Figure S2J**) highlighted an example of this, demonstrating an anticipated but limited correlation (52). Furthermore, genomics and transcriptomics are limited in capturing organism-wide tumour biology from acellular samples such as plasma.

This study aimed to combine and implement recent advances in quantitative MS proteomics to comprehensively characterise the tumours and plasma of contrasting mouse B-cell cancer models; Eμ-*myc* and Eμ-*TCL1*. By comparing tumours driven by contrasting oncogenes, with differing phenotypes and rates of progression, any apparent common signatures could suggest the most conserved and essential B-cell cancer mechanisms. Model-specific signatures could provide insight into the molecular basis of the discrete cancer phenotypes and oncogenes, potentially of relevance to BL and CLL. Simultaneous plasma characterisation offered potential insight into tumour-host dialogue, systemic cancer impacts and biology of biomarker emergence. Furthermore, greater

understanding of these models offers an opportunity to appreciate their strengths and weaknesses in pre-clinical applications.

Overall, isobaric-labelled LC-MS proteomics of B-cell and plasma samples (**Figure 1**) provided biologically reproducible (**Figure 2A**, **B**), high-depth (**Figure S1H**) and representative (**Figure S1**) characterisations of the Eμ-*myc* and Eμ-*TCL1* tumour models. The accuracy of the quantitative results was validated by anticipated transgene expression (**Figure 2C**), Western blotting and flow cytometry (**Figures 2D, S2G,H**) alongside comparisons to previously published protein expressions (**Figure S2D,F**) and mRNA expression data (**Figure S2J**). Interestingly, a brief analysis revealed that those proteins over-expressed without corresponding mRNA over-expression were enriched with proteins annotated with the term 'translation'.

A common B-cell tumour signature in divergent models

The expression correlation between the contrasting Eμ-*myc* and Eμ-*TCL1* tumours (**Figure 3**) was indicative of a common B-cell tumour signature which was proportional to aggression. Indeed, this signature strongly highlighted canonical tumour characteristics such as up-regulated cell proliferation and growth (**Figure 4**). Furthermore, frequent observations of these over-expressed proteins in several human cancers (**Figure S3C**) highlighted potential relevance to non-B-cell cancers. This was particularly noteworthy, given the many clinically relevant cell surface proteins and drug targets characterised (**Figure 3E**, **S3F,G**). HMMR/CD168, for instance, has previously been considered as an immunotherapy target in B-cell cancers (41, 53). Similarly, amino acid and zinc transporter over-expression could offer targets of metabolic inhibition or immunotherapy. The observation of proteins, and even putative proteins, in this common B-cell tumour signature with no prior links to cancer offers several new hypotheses (**Figure S3E**). DDX49 and MAK16, for instance, are poorly characterised proteins over-expressed in all 4 tumour pools, and could be inferred to have a role in ribosome biogenesis, given their homology, orthologous functions and the prevalence of ribosome biogenesis upregulation in both tumours.

Down-regulation of immune functions were indicative of the loss of B-cell characteristics, not essential and potentially even inhibitory towards tumour development. Constitutive Ig synthesis, for instance, would provide a negative selective pressure on resource-limited tumour cells. MHC component under-expression suggested a clear mechanism of immune evasion, such as that described previously (54, 55).

Model specific signatures of Eμ-*myc* and Eμ-*TCL1* tumours

Eμ-*myc*- and Eμ-*TCL1*-specific DEPs (**Figures 5-6**) highlighted distinct expression patterns responsible for the contrast in tumour phenotypes; such as the vast protein dysregulation and proliferation produced by the aggressive and pleiotropic nature of myc. The contrast between these models was also apparent prior to tumourigenesis (**Figure S3B**), with little to no tumour signature in 6-week Eμ-*TCL1* B-cells, whilst Eμ-myc 6-week old B-cells were almost indistinguishable from terminal tumours. Although none of these 6-week Eμ-myc mice presented with splenomegaly, metastases or ill health, it remains possible that tumour development was present at an early stage in one or more of the mice. Eμ-*myc*-specific DNA and histone methylation dysregulation (**Figures 5B,C**) suggests dynamic epigenomic instability as a basis of myc-induced tumour characteristics, previously observed in B-cell lymphomas (56, 57). This potentially drives accelerated trait acquisition via non-mutational evolution (58).

Broad under-expression of actin cytoskeletal organising components was potentially indicative of defects to normal B-cell migration and adhesion (**Figure 5D**), identifying a possible mechanism promoting lymph node metastases; frequently observed for Eμ-*myc* tumours. These proteins may also have been under-expressed as a result of redundant immune-related functions such as those described in **Figure 4D**. Of the three inter-connected actin cytoskeletal organising proteins validated with under-expression (**Figure 5E**), all three demonstrated both immune and cytoskeletal function. Cor1A promotes F-actin disassembly and cell motility (59) and mutations have been shown to impair lymphocyte function (60, 61). LCP1 regulates actin bundling and B-cell development (62, 63), while LCP1 knock out impaired B-cell migration (64). CAPG caps actin filaments with $CAPG^{-/-}$ mice exhibiting immune defects (65). Given observations of the pro-metastatic nature of these

proteins in non-lymphocyte tumours (66, 67), this indicated metastasis may be a more passive event for Eμ-*myc* tumours. Over-expression of adhesion components such as integrin beta 1 may have facilitated this. Non-canonical cytoskeletal component AHNAK, a titin-related and cancer-associated cytoskeletal protein, shown to be essential for migration and invasion (68); was over-expressed in Eμ-*myc* tumours, potentially highlighting an alternative metastatic pathway. Additionally other cytoskeletal components such as those directing chromosome segregation were over-expressed. Validation of two further under-expressed proteins extended the observation of cytoskeletal dysregulation, with moesin and myosin-9 annotated with roles in leukocyte migration. Together these observations suggest a trend for further investigation.

The Eμ-*TCL1*-specific features, such as ER stress response and aberrant glycosylation upregulation (**Figure 6B-C**) have previously been observed (69). The proteomics may therefore offer some specific novel examples of contributing proteins. Furthermore, integrin-α2 under-expression, transmembrane protein processing, post-translational modification, transport and membrane lipid composition offer several potential mechanisms by which circulatory, leukaemic cells formation may be promoted in Eμ-*TCL1* tumours.

Minimal BCR pathway under-expression, relative to Eμ-*myc* tumours, indicates a role for antigen-receptor signalling, observed previously (70), in Eμ-*TCL1* tumours (**Figure S5J**). This highlights potential examples of negative regulation or redundant functions for tumour growth and survival eg. BCL6, FOXO1 and CD45. Upregulated phosphorylations, such as PI3Ks and FOXO1, specifically in Eμ-*TCL1* tumours may also illustrate tumour mechanisms (**Figure 5F-J**, **Table S4**).

Interleukin 5 receptor over-expression and signalling in Eμ-*TCL1* tumours

Signalling dysregulation (**Figure 6C**) and increased expression and phosphorylation of PI3Ks (**Figure S4J**) suggested the presence of strong receptor signalling underlying Eμ-*TCL1* tumourigenesis. IL5R over-expression (**Figure 6A,C-E, S5**) strongly implied a role for the IL5:IL5R signalling pathway in Eμ-*TCL1* tumours. While not previously described in Eμ-*TCL1* mice, other

mouse CLL-like B-cell tumours have demonstrated IL5 sensitivity (71-74), in addition to a CLL-like leukaemia arising in mice with constitutive over-expression of IL5 (75).

The dose-dependent proliferation of Eμ-*TCL1* tumour cells upon IL5 treatment (**Figure 7A-C**) provided additional, functional validation of IL5R over-expression. IL5R signalling has been suggested to induce activation of Lyn, JAK2, Syk, BTK, NF-κB and PI3K (76-78). Given the suggested role of AKT amplification by TCL1, PI3K signalling was a likely candidate for the effectuation of the IL5 response. Indeed, the PI3K catalytic subunit type 2 beta (PIK3C2B), alongside other PI3K isoforms, appeared as the most over-expressed of the IL5R downstream signalling molecules. Dose-dependent AKT phosphorylation upon IL5 treatment confirmed a role for the PI3K pathway, highlighting TCL1 as a potential factor in the amplification of IL5R signalling (**Figure 7D-E**).

The observation of IL5RA expression by the expanded $CD5^+$ peritoneal and splenic B-cell populations in young Eμ-*TCL1* mice (**Figure S5J**) strongly suggests a precursor cell population with characteristics of B-1 B-cells. Given the induction of B-cell proliferation observed with IL5 (79, 80) and that $IL5^{-/-}$ mice are deficient for $CD5^+$ B-1 B-cells (81), IL5R signalling emerges as a probable component driving Eμ-*TCL1* tumourigenesis from $CD5^+$ $IL5RA^+$ B-1 B-cells.

The role for IL5 in B-cells, and B-cell tumours, appears to not be conserved between species (80, 82) highlighting a limitation in the recapitulation of human CLL with previous reports suggesting IL5 induces spontaneous apoptosis in these cells (83). However, the similarities of IL5R signalling in Eμ-*TCL1* tumours with IL4 signalling in CLL, suggest continued potential for Eμ-*TCL1* mice in modelling many aspects of the human disease.

Plasma proteomics and systemic tumour signatures

The >250-fold, depletion of high-abundant proteins by SuPrE-SEC and the identification of a wide range of tumour and systemic proteins (**Table S2**, **Figure 8**, **S6**) suggested the low M_w sub-proteome provided an effective approach to plasma analysis. The capture of the degradome and

peptidome, portions of plasma containing fragmented proteins and potential biomarkers, contributed to the success of this method (84-86).

The over-abundance of tumour-derived proteins in terminal plasma strongly suggested tumour lysis as a dominant mechanism (**Figures 8A-D**, **S6A,D-E**), a process described in aggressive therapies and lymphoma (87-89). Lysis product over-abundance correlated with tumour aggression and higher rates of cell death, most clearly in Eμ-*myc* tumours. A correlation between total protein abundance in the B-cell proteome and over-abundance in the plasma was observed, regardless of tumour over-expression (**Figure S6G-I**). Several of these proteins were annotated biomarkers (**Figure S6H**) highlighting tumour-lysis products in late-stage, aggressive tumours, as the dominant biomarker signature. It is likely that exosomes and apoptotic blebs also contribute to this signature.

The systemic immune response signature observed for terminal Eμ-*TCL1* plasma suggested that slower tumour development elicited a greater inflammatory response, perhaps due to gradual accumulation, chronic inflammation and a loss of immune regulation (**Figures 8E**). An anti-tumour immune signature emerged for both models, when deconvoluted to give the 'tumour lysis-free' plasma proteome, which suggested several potential systemic biomarkers (**Figure 8C,D**).

The contrasting signatures of tumour lysis and immune response in Eμ-*myc* and Eμ-*TCL1* terminal plasma respectively, was reflective of the vastly differing rates of tumour development. However, in both models the signatures were observed, to a lesser extent, in the opposing tumour.

Several systemic signatures were also observed upon integration of the B-cell tumour and plasma proteomes. Over-abundance of the carriers of the extracellular matrix component hyaluronan ITIH1-4 (**Figure 8E**), substantial tumour over-expression of CD168 (HMMR, hyaluronan-mediated motility receptor) (**Figure 3D**) and links between hyaluronan, inflammation and tumour growth (90) suggest hyaluronan as a component of the B-cell tumour microenvironment. Given their high concentrations in plasma (**Table S1**, **Figure S6Q**) (37), ITIH proteins hold potential as biomarkers of B-cell tumours and aberrant hyaluronan metabolism and transport. Fibronectin (FN1), recognised in the promotion of cancers (91), was observed as marginally over-abundant and inferred to be

functional (**Figure S6S**). Given the over-expression of FN1-binding integrin β1 on tumours, this may represent another microenvironment signature of pro-oncogenic interactions with B-cell tumours.

Functional biomarker analysis for over-abundant plasma proteins with inferred functionality in tumours highlighted S100A6, a previously proposed biomarker (92, 93), and EIF4G1 (**Figure S6T**) suggesting candidates for non-invasive testing directly related to tumour function.

Pre-terminal Eμ-*TCL1* plasma additionally offered several interesting findings, describing a more clinically-relevant time point in tumourigenesis when a leukaemia is present, but is otherwise asymptomatic. The observation of overabundant pre-terminal '30%' Eμ-*TCL1* plasma proteins common to the terminal signature (**Figure 8E, S6I,O,P,R**) suggested the characterisation of progressive, early biomarkers of B-cell tumours. Components of the tumour lysis-free signature describing a systemic immune response were more prevalent, such as Hp and ITIH1. Proteins correlating with tumour progression, traceable to tumour lysis were, however, also present to some extent, suggesting that both lysis and immune signatures offer a source of biomarkers at earlier stages of tumour development. The more detailed evaluation of the multiple stages of tumour progression, of these and other tumour models, could therefore offer considerable insight into the dynamics of biomarker emergence, better informing biomarker discovery and application.

Concluding remarks

In conclusion, this study has detailed the protein expression changes present across highly contrasting B-cell tumours, highlighting dominant features of proliferation and growth. Eμ-*myc* tumours demonstrated a specific trend of epigenomic instability and cytoskeletal component under-expression, while Eμ-*TCL1* tumours specifically exhibited ER stress, dysregulated signalling, and IL5-driven proliferation. Many specific targets of kinase inhibitors, nucleoside analogues and immunotherapy emerged for each tumour type. Plasma proteomics indicated lysis products as a major late-stage tumour biomarker signature, dominantly in the more aggressive Eμ-*myc* tumours, while Eμ-*TCL1* tumour plasma had a dominant signature of an anti-tumour immune response. Integration of plasma and B-cell proteomics data, alongside tumour lysis, highlighted IL5R, HMMR/CD168, ITIH

proteins, FN1, S100A6 and EIF4G1 as candidates for further investigation. These findings reinforce that aggressive, targeted, chemotherapy and immunostimulatory antibodies promise effective means of treating Eμ-*myc* and Eμ-*TCL1* tumours, respectively with the results offer exciting possibilities when combined and integrated with additional 'omics and functional data. Finally, these findings suggest that global, integrative evaluation of tumours can provide systemic insight into tumour biology not captured by the evaluation of cells or plasma in isolation.

Acknowledgements

We would like to thank Roger Allsopp and Derek Coates for raising funds for PAT that allowed the purchasing of the FT-MS proteomics platform at the University of Southampton. We would like to acknowledge the Biomedical Research Facility and their staff, the use of the IRIDIS High Performance Computing Facility and associated support services at the University of Southampton, and the PRIDE team in the submission of our data to the PRIDE database. The authors would also like to thank Cory White for SPIQuE development, Dr. Xunli Zhang for HPLC access and Theo Roumeliotis for proteomics training. Also thanks to Graham Packham for critical feedback on this work.

Author contributions

Conceptualization: HEJ, PAT, SDG, MSC. Methodology: HEJ. Validation: HEJ, MJC. Formal analysis: HEJ. Investigation: HEJ. Resources: HEJ, MD, KLC, AM, SDG. Writing – Original draft: HEJ, MSC. Writing – Review & Editing: HEJ, MSC, MJC, SDG, PAT. Visualisation: HEJ. Supervision: MSC, PAT, SDG. Funding Acquisition: MSC, PAT, HEJ.

Conflict of Interest Disclosure

MSC is a retained consultant for Bioinvent and has performed educational and advisory roles for Baxalta. He has received research funding from Roche, Gilead and GSK. PAT is cofounder of Karus Therapeutics, a NED for the Aptamer Group and has performed educational and advisory roles for Abcam, BioRelate, ReactaBiotech and Waters amongst others. He is also Adjunct Professor,

Singapore Institute of Immunology, A*Star, Singapore; and Honorary Professor, Department of Histology and Embryology, University of Athens, Greece.

References

1. Iversen U, Iversen OH, Bluming AZ, Ziegler JL, Kyalwasi S. Cell kinetics of African cases of Burkitt lymphoma. A preliminary report. Eur J Cancer. 1972;8(3):305-8.
2. Molyneux EM, Rochford R, Griffin B, Newton R, Jackson G, Menon G, et al. Burkitt's lymphoma. Lancet. 2012;379(9822):1234-44.
3. Adams JM, Harris AW, Pinkert CA, Corcoran LM, Alexander WS, Cory S, et al. The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature. 1985;318(6046):533-8.
4. Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F. The c-Myc target gene network. Seminars in cancer biology. 2006;16(4):253-64.
5. Zeller KI, Jegga AG, Aronow BJ, O'Donnell KA, Dang CV. An integrated database of genes responsive to the Myc oncogenic transcription factor: identification of direct genomic targets. Genome biology. 2003;4(10):R69.
6. Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. Science. 1990;250(4984):1149-51.
7. Nesbit CE, Tersak JM, Prochownik EV. MYC oncogenes and human neoplastic disease. Oncogene. 1999;18(19):3004-16.
8. Harris AW, Pinkert CA, Crawford M, Langdon WY, Brinster RL, Adams JM. The E mu-myc transgenic mouse. A model for high-incidence spontaneous lymphoma and leukemia of early B cells. The Journal of experimental medicine. 1988;167(2):353-71.
9. Mori S, Rempel RE, Chang JT, Yao G, Lagoo AS, Potti A, et al. Utilization of pathway signatures to reveal distinct types of B lymphoma in the Emicro-myc model and human diffuse large B-cell lymphoma. Cancer Res. 2008;68(20):8525-34.
10. Zhu D, Qi CF, Morse HC, 3rd, Janz S, Stevenson FK. Deregulated expression of the Myc cellular oncogene drives development of mouse "Burkitt-like" lymphomas from naive B cells. Blood. 2005;105(5):2135-7.
11. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner 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. 2008;111(12):5446-56.
12. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner 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. 2008;111(12):5446-56.
13. Herling M, Patel KA, Khalili J, Schlette E, Kobayashi R, Medeiros LJ, et al. TCL1 shows a regulated expression pattern in chronic lymphocytic leukemia that correlates with molecular subtypes and proliferative state. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK. 2006;20(2):280-5.
14. Kang SM, Narducci MG, Lazzeri C, Mongiovi AM, Caprini E, Bresin A, et al. Impaired T- and B-cell development in Tcl1-deficient mice. Blood. 2005;105(3):1288-94.
15. Said JW, Hoyer KK, French SW, Rosenfelt L, Garcia-Lloret M, Koh PJ, et al. TCL1 oncogene expression in B cell subsets from lymphoid hyperplasia and distinct classes of B

cell lymphoma. Laboratory investigation; a journal of technical methods and pathology. 2001;81(4):555-64.
16. Pekarsky Y, Koval A, Hallas C, Bichi R, Tresini M, Malstrom S, et al. Tcl1 enhances Akt kinase activity and mediates its nuclear translocation. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(7):3028-33.
17. Laine J, Kunstle G, Obata T, Sha M, Noguchi M. The protooncogene TCL1 is an Akt kinase coactivator. Molecular cell. 2000;6(2):395-407.
18. Bichi R, Shinton SA, Martin ES, Koval A, Calin GA, Cesari R, et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(10):6955-60.
19. Johnson AJ, Lucas DM, Muthusamy N, Smith LL, Edwards RB, De Lay MD, et al. Characterization of the TCL-1 transgenic mouse as a preclinical drug development tool for human chronic lymphocytic leukemia. Blood. 2006;108(4):1334-8.
20. Pekarsky Y, Drusco A, Kumchala P, Croce CM, Zanesi N. The long journey of TCL1 transgenic mice: lessons learned in the last 15 years. Gene Expr. 2015;16(3):129-35.
21. Simonetti G, Bertilaccio MT, Ghia P, Klein U. Mouse models in the study of chronic lymphocytic leukemia pathogenesis and therapy. Blood. 2014;124(7):1010-9.
22. Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, et al. Mass-spectrometry-based draft of the human proteome. Nature. 2014;509(7502):582-7.
23. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014;509(7502):575-81.
24. Nanjappa V, Thomas JK, Marimuthu A, Muthusamy B, Radhakrishnan A, Sharma R, et al. Plasma Proteome Database as a resource for proteomics research: 2014 update. Nucleic Acids Res. 2014;42(Database issue):D959-65.
25. Keshishian H, Burgess MW, Gillette MA, Mertins P, Clauser KR, Mani DR, et al. Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. Molecular & cellular proteomics : MCP. 2015;14(9):2375-93.
26. Jones KA, Kim PD, Patel BB, Kelsen SG, Braverman A, Swinton DJ, et al. Immunodepletion plasma proteomics by tripleTOF 5600 and Orbitrap elite/LTQ-Orbitrap Velos/Q exactive mass spectrometers. Journal of proteome research. 2013;12(10):4351-65.
27. Patel BB, Barrero CA, Braverman A, Kim PD, Jones KA, Chen DE, et al. Assessment of two immunodepletion methods: off-target effects and variations in immunodepletion efficiency may confound plasma proteomics. Journal of proteome research. 2012;11(12):5947-58.
28. Al-Daghri NM, Al-Attas OS, Johnston HE, Singhania A, Alokail MS, Alkharfy KM, et al. Whole serum 3D LC-nESI-FTMS quantitative proteomics reveals sexual dimorphism in the milieu interieur of overweight and obese adults. J Proteome Res. 2014;13(11):5094-105.
29. Garbis SD, Roumeliotis TI, Tyritzis SI, Zorpas KM, Pavlakis K, Constantinides CA. A novel multidimensional protein identification technology approach combining protein size exclusion prefractionation, peptide zwitterion-ion hydrophilic interaction chromatography, and nano-ultraperformance RP chromatography/nESI-MS2 for the in-depth analysis of the serum proteome and phosphoproteome: application to clinical sera derived from humans with benign prostate hyperplasia. Anal Chem. 2011;83(3):708-18.
30. Adeghe AJ, Cohen J. A better method for terminal bleeding of mice. Laboratory animals. 1986;20(1):70-2.
31. Hahne H, Pachl F, Ruprecht B, Maier SK, Klaeger S, Helm D, et al. DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. Nat Methods. 2013;10(10):989-91.

32. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol. 2014;32(3):223-6.
33. Levin Y. The role of statistical power analysis in quantitative proteomics. Proteomics. 2011;11(12):2565-7.
34. Blunt MD, Carter MJ, Larrayoz M, Smith LD, Aguilar-Hernandez M, Cox KL, et al. The PI3K/mTOR inhibitor PF-04691502 induces apoptosis and inhibits microenvironmental signaling in CLL and the Emicro-TCL1 mouse model. Blood. 2015;125(26):4032-41.
35. Hussain K, Hargreaves CE, Roghanian A, Oldham RJ, Chan HT, Mockridge CI, et al. Upregulation of FcgammaRIIb on monocytes is necessary to promote the superagonist activity of TGN1412. Blood. 2015;125(1):102-10.
36. Ivanov A, Cragg MS, Erenpreisa J, Emzinsh D, Lukman H, Illidge TM. Endopolyploid cells produced after severe genotoxic damage have the potential to repair DNA double strand breaks. J Cell Sci. 2003;116(Pt 20):4095-106.
37. Farrah T, Deutsch EW, Omenn GS, Campbell DS, Sun Z, Bletz JA, et al. A high-confidence human plasma proteome reference set with estimated concentrations in PeptideAtlas. Mol Cell Proteomics. 2011;10(9):M110 006353.
38. Rounbehler RJ, Fallahi M, Yang C, Steeves MA, Li W, Doherty JR, et al. Tristetraprolin impairs myc-induced lymphoma and abolishes the malignant state. Cell. 2012;150(3):563-74.
39. Ow SY, Salim M, Noirel J, Evans C, Wright PC. Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. Proteomics. 2011;11(11):2341-6.
40. Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS. Addressing accuracy and precision issues in iTRAQ quantitation. Mol Cell Proteomics. 2010;9(9):1885-97.
41. Snauwaert S, Vanhee S, Goetgeluk G, Verstichel G, Van Caeneghem Y, Velghe I, et al. RHAMM/HMMR (CD168) is not an ideal target antigen for immunotherapy of acute myeloid leukemia. Haematologica. 2012;97(10):1539-47.
42. Giannopoulos K, Li L, Bojarska-Junak A, Rolinski J, Dmoszynska A, Hus I, et al. Expression of RHAMM/CD168 and other tumor-associated antigens in patients with B-cell chronic lymphocytic leukemia. Int J Oncol. 2006;29(1):95-103.
43. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43(Database issue):D447-52.
44. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30(4):523-30.
45. Eom GH, Nam YS, Oh JG, Choe N, Min HK, Yoo EK, et al. Regulation of acetylation of histone deacetylase 2 by p300/CBP-associated factor/histone deacetylase 5 in the development of cardiac hypertrophy. Circ Res. 2014;114(7):1133-43.
46. Ochs RC, Bagg A. Molecular genetic characterization of lymphoma: application to cytology diagnosis. Diagnostic cytopathology. 2012;40(6):542-55.
47. Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, et al. The genetic landscape of mutations in Burkitt lymphoma. Nature genetics. 2012;44(12):1321-5.
48. Yustein JT, Liu YC, Gao P, Jie C, Le A, Vuica-Ross M, et al. Induction of ectopic Myc target gene JAG2 augments hypoxic growth and tumorigenesis in a human B-cell model. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(8):3534-9.
49. Xiao W, Tran B, Staudt LM, Schmitz R. High-throughput RNA sequencing in B-cell lymphomas. Methods Mol Biol. 2013;971:295-312.

50. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Molecular and cellular biology. 1999;19(3):1720-30.
51. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet. 2012;13(4):227-32.
52. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. Nature. 2011;473(7347):337-42.
53. Tabarkiewicz J, Giannopoulos K. Definition of a target for immunotherapy and results of the first Peptide vaccination study in chronic lymphocytic leukemia. Transplant Proc. 2010;42(8):3293-6.
54. Solana R, Romero J, Alonso C, Pena J. MHC class I antigen expression is inversely related with tumor malignancy and ras oncogene product (p21ras) levels in human breast tumors. Invasion Metastasis. 1992;12(3-4):210-7.
55. Paulson KG, Tegeder A, Willmes C, Iyer JG, Afanasiev OK, Schrama D, et al. Downregulation of MHC-I expression is prevalent but reversible in Merkel cell carcinoma. Cancer Immunol Res. 2014;2(11):1071-9.
56. Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, et al. Induction of tumors in mice by genomic hypomethylation. Science. 2003;300(5618):489-92.
57. Hopp L, Nersisyan L, Loffler-Wirth H, Arakelyan A, Binder H. Epigenetic Heterogeneity of B-Cell Lymphoma: Chromatin Modifiers. Genes (Basel). 2015;6(4):1076-112.
58. Li S, Garrett-Bakelman FE, Chung SS, Sanders MA, Hricik T, Rapaport F, et al. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. Nat Med. 2016;22(7):792-9.
59. Pollard TD, Borisy GG. Cellular motility driven by assembly and disassembly of actin filaments. Cell. 2003;112(4):453-65.
60. Foger N, Rangell L, Danilenko DM, Chan AC. Requirement for coronin 1 in T lymphocyte trafficking and cellular homeostasis. Science. 2006;313(5788):839-42.
61. Haraldsson MK, Louis-Dit-Sully CA, Lawson BR, Sternik G, Santiago-Raber ML, Gascoigne NR, et al. The lupus-related Lmb3 locus contains a disease-suppressing Coronin-1A gene mutation. Immunity. 2008;28(1):40-51.
62. Al Tanoury Z, Schaffner-Reckinger E, Halavatyi A, Hoffmann C, Moes M, Hadzic E, et al. Quantitative kinetic study of the actin-bundling protein L-plastin and of its impact on actin turn-over. PLoS One. 2010;5(2):e9210.
63. Todd EM, Deady LE, Morley SC. The actin-bundling protein L-plastin is essential for marginal zone B cell development. J Immunol. 2011;187(6):3015-25.
64. Todd EM, Deady LE, Morley SC. Intrinsic T- and B-cell defects impair T-cell-dependent antibody responses in mice lacking the actin-bundling protein L-plastin. Eur J Immunol. 2013;43(7):1735-44.
65. Parikh SS, Litherland SA, Clare-Salzler MJ, Li W, Gulig PA, Southwick FS. CapG(-/-) mice have specific host defense defects that render them more susceptible than CapG(+/+) mice to Listeria monocytogenes infection but not to Salmonella enterica serovar Typhimurium infection. Infect Immun. 2003;71(11):6582-90.
66. Foran E, McWilliam P, Kelleher D, Croke DT, Long A. The leukocyte protein L-plastin induces proliferation, invasion and loss of E-cadherin expression in colon cancer cells. Int J Cancer. 2006;118(8):2098-104.
67. Glaser J, Neumann MH, Mei Q, Betz B, Seier N, Beyer I, et al. Macrophage capping protein CapG is a putative oncogene involved in migration and invasiveness in ovarian carcinoma. Biomed Res Int. 2014;2014:379847.

68. Shankar J, Messenberg A, Chan J, Underhill TM, Foster LJ, Nabi IR. Pseudopodial actin dynamics control epithelial-mesenchymal transition in metastatic cancer cells. Cancer Res. 2010;70(9):3780-90.
69. Kriss CL, Pinilla-Ibarz JA, Mailloux AW, Powers JJ, Tang CH, Kang CW, et al. Overexpression of TCL1 activates the endoplasmic reticulum stress response: a novel mechanism of leukemic progression in mice. Blood. 2012;120(5):1027-38.
70. Iacovelli S, Hug E, Bennardo S, Duehren-von Minden M, Gobessi S, Rinaldi A, et al. Two types of BCR interactions are positively selected during leukemia development in the Emu-TCL1 transgenic mouse model of CLL. Blood. 2015;125(10):1578-88.
71. Slavin S, Strober S. Spontaneous murine B-cell leukaemia. Nature. 1978;272(5654):624-6.
72. O'Garra A, Barbis D, Wu J, Hodgkin PD, Abrams J, Howard M. The BCL1 B lymphoma responds to IL-4, IL-5, and GM-CSF. Cell Immunol. 1989;123(1):189-200.
73. Lohoff M, Sommer F, Rollinghoff M. Suppressive effect of interferon-gamma on the BCL1 cell-dependent interleukin 5 bioassay. Eur J Immunol. 1989;19(7):1327-9.
74. Lasky JL, Thorbecke GJ. Characterization and growth factor requirements of SJL lymphomas. II. Interleukin 5 dependence of the in vitro cell line, cRCS-X, and influence of other cytokines. Eur J Immunol. 1989;19(2):365-71.
75. Wen X, Zhang D, Kikuchi Y, Jiang Y, Nakamura K, Xiu Y, et al. Transgene-mediated hyper-expression of IL-5 inhibits autoimmune disease but increases the risk of B cell chronic lymphocytic leukemia in a model of murine lupus. Eur J Immunol. 2004;34(10):2740-9.
76. Sato S, Katagiri T, Takaki S, Kikuchi Y, Hitoshi Y, Yonehara S, et al. IL-5 receptor-mediated tyrosine phosphorylation of SH2/SH3-containing proteins and activation of Bruton's tyrosine and Janus 2 kinases. The Journal of experimental medicine. 1994;180(6):2101-11.
77. Coffer PJ, Schweizer RC, Dubois GR, Maikoe T, Lammers JW, Koenderman L. Analysis of signal transduction pathways in human eosinophils activated by chemoattractants and the T-helper 2-derived cytokines interleukin-4 and interleukin-5. Blood. 1998;91(7):2547-57.
78. Adachi T, Alam R. The mechanism of IL-5 signal transduction. Am J Physiol. 1998;275(3 Pt 1):C623-33.
79. Wetzel GD. Interleukin 5 regulation of peritoneal Ly-1 B lymphocyte proliferation, differentiation and autoantibody secretion. Eur J Immunol. 1989;19(9):1701-7.
80. Sanderson CJ. Interleukin-5: an eosinophil growth and activation factor. Dev Biol Stand. 1988;69:23-9.
81. Kopf M, Brombacher F, Hodgkin PD, Ramsay AJ, Milbourne EA, Dai WJ, et al. IL-5-deficient mice have a developmental defect in CD5+ B-1 cells and lack eosinophilia but have normal antibody and cytotoxic T cell responses. Immunity. 1996;4(1):15-24.
82. Clutterbuck E, Shields JG, Gordon J, Smith SH, Boyd A, Callard RE, et al. Recombinant human interleukin 5 is an eosinophil differentiation factor but has no activity in standard human B cell growth factor assays. Eur J Immunol. 1987;17(12):1743-50.
83. Mainou-Fowler T, Prentice AG. Modulation of apoptosis with cytokines in B-cell chronic lymphocytic leukaemia. Leukemia & lymphoma. 1996;21(5-6):369-77.
84. van Winden AW, van den Broek I, Gast MC, Engwegen JY, Sparidans RW, van Dulken EJ, et al. Serum degradome markers for the detection of breast cancer. Journal of proteome research. 2010;9(8):3781-8.
85. Karczmarski J, Rubel T, Mikula M, Wolski J, Rutkowski A, Zagorowicz E, et al. Pre-analytical-related variability influencing serum peptide profiles demonstrated in a mass

spectrometry-based search for colorectal and prostate cancer biomarkers. Acta Biochim Pol. 2013;60(3):417-25.
86. Yang J, Yang J, Gao Y, Zhao L, Liu L, Qin Y, et al. Identification of potential serum proteomic biomarkers for clear cell renal cell carcinoma. PLoS One. 2014;9(11):e111364.
87. Davidson MB, Thakkar S, Hix JK, Bhandarkar ND, Wong A, Schreiber MJ. Pathophysiology, clinical consequences, and treatment of tumor lysis syndrome. Am J Med. 2004;116(8):546-54.
88. Wossmann W, Schrappe M, Meyer U, Zimmermann M, Reiter A. Incidence of tumor lysis syndrome in children with advanced stage Burkitt's lymphoma/leukemia before and after introduction of prophylactic use of urate oxidase. Ann Hematol. 2003;82(3):160-5.
89. Otrock ZK, Hatoum HA, Salem ZM. Acute tumor lysis syndrome after rituximab administration in Burkitt's lymphoma. Intern Emerg Med. 2008;3(2):161-3.
90. Misra S, Hascall VC, Markwald RR, Ghatak S. Interactions between Hyaluronan and Its Receptors (CD44, RHAMM) Regulate the Activities of Inflammation and Cancer. Front Immunol. 2015;6:201.
91. Sponziello M, Rosignolo F, Celano M, Maggisano V, Pecce V, De Rose RF, et al. Fibronectin-1 expression is increased in aggressive thyroid cancer and favors the migration and invasion of cancer cells. Mol Cell Endocrinol. 2016;431:123-32.
92. Zhang J, Zhang K, Jiang X, Zhang J. S100A6 as a potential serum prognostic biomarker and therapeutic target in gastric cancer. Dig Dis Sci. 2014;59(9):2136-44.
93. Nishi M, Matsumoto K, Kobayashi M, Yanagita K, Matsumoto T, Nagashio R, et al. Serum expression of S100A6 is a potential detection marker in patients with urothelial carcinoma in the urinary bladder. Biomed Res. 2014;35(6):351-6.
94. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.

Figure legends

Figure 1. Eμ-*myc* and Eμ-*TCL1* model proteomic characterisation workflow. B-cells and plasma were isolated from splenocytes and blood, respectively, derived from terminal and pre-terminal Eμ-*myc* and Eμ-*TCL1* mice and WT controls (**Figure S1C**). Pooling was used to accommodate all samples within single isobaric-labelled experiments for B-cell (iTRAQ 8-plex) and plasma (TMT 10-plex) proteomics. Pools of 2, 3 or 6 samples were used for tumours, tumours plasma and pre-tumour controls, respectively. For all tumour samples, biological replicate pools were analysed. Plasma pools, totalling 120 µl per pool, were subjected to size exclusion chromatography (SEC) to isolate the low molecular weight sub-proteome (**Figure S1G**). B-cells were isolated by negative selection and were lysed, quantified and then 50 µg of protein pooled from replicates of each condition. Plasma and B-cell proteins were subjected to reducing conditions, cysteine alkylation, trypsin proteolysis and isobaric labelling and pooling of the generated peptides. The labelled peptides for each proteome were resolved by two-dimensional liquid chromatography and quantitatively characterised by mass spectrometry (MS). The MS data were then subject to target-decoy analysis and reporter ion quantitation to generate two quantitative proteomes.

Figure 2. Analysis and validation of the quality of quantitative proteomics data. A. Hierarchical clustering of all 8270 and 2095 fully profiled protein $\log_2$ (ratios) relative to the 6 week WT control (in addition to the 200 day WT control for tumour samples) using Cluster 3.0 and Euclidian distance to represent the topological similarities and differences for each sample. A Venn diagram highlights proteins identified in both proteomes. **B.** Linear regression highlighting the reproducibility of the $\log_2$ (ratios) relative to WT 6 week samples of the biological replicates analysed for each tumourous condition. **C.** Relative quantitative proteomics-derived fold changes of transgene-derived protein expression and characteristic Eμ-*TCL1* phenotypes, relative to WT B-cells. **D.** A summary of 17 Western blot and flow cytometry validations of Eμ-*myc* and Eμ-*TCL1* tumour protein expressions relative to WT B-cell controls, plotted alongside the proteomics-derived expression changes (as $\log_2$ (ratios)). These values relate to the validations detailed in **Figure S2G,H**.

Figure 3. Differential protein expression in Eμ-*myc* and Eμ-*TCL1* B-cell tumours. A. Volcano plots highlighting reproducible, significant differential protein expression in each B-cell tumour on the basis of the regulation score (Rs) and FDR-corrected p-values (one sample t-test). The Rs was calculated from the mean and standard deviation (SD) of the 4 $\log_2$ (ratios) of tumour protein expression relative to both WT controls (Rs= mean/(SD+1)). The number of proteins considered significantly ($p<0.05$) over-expressed ($Rs>0.5$) or under-expressed ($Rs<-0.5$) is detailed. **B.** A linear regression between the protein expression observed in Eμ-*myc* and Eμ-*TCL1* B-cell tumours, both relative to WT B-cells. **C.** Venn diagrams highlighting proteins considered over- and under-expressed

common to both Eμ-*myc* and Eμ-*TCL1* B-cell tumours. **D.** Heat maps of individual $\log_2$ (ratios to WT), for examples of proteins differentially expressed commonly in both tumours. The top 10 proteins falling into the following categories are presented; significantly over-expressed in both tumours with high confidence (>3 unique peptides and unique quantitations), cell surface expression, proteins annotated by ingenuity pathway analysis as drug targets, proteins for which no specific links to any type of cancer have previously been made, based on PubMed searching and proteins significantly under-expressed in both tumours.

Figure 4. Bioinformatic interrogation of commonly differentially expressed B-cell tumour proteins. A. Gene ontology (GO) term enrichment analysis for proteins over- and under-expressed in both tumours, visualised using Revigo (94) and summarised based on parent GO terms and Revigo-defined semantic space. **B-D.** StringDB networks highlighting interactions and relationships for proteins **B.** over-expressed in both tumours annotated with the term 'chromosome segregation' (GO:0007059) (n=81), **C.** over-expressed in both tumours annotated as 'ribosome biogenesis' (GO:0042254) (n=106), and **D.** those proteins under-expressed in both tumours relating to terms descriptive of immune regulation. **E.** Upstream regulator activation z-scores inferred by IPA for all consistently differentially expressed B-cell tumour proteins, plotted against tumour protein expression, relative to WT B-cells. A positive z-score indicates a signature of protein activation.

Figure 5. Eμ-*myc*-specific tumour characteristics. A. Heat maps of individual $\log_2$ (ratios to WT), for the top 10 proteins specifically over- and under-expressed in Eμ-*myc* tumours (Rs<0.25/>-0.25 for Eμ-*TCL1* tumours). Additionally the top 10 proteins with cell surface expression are shown. **B.** Upstream regulator activation z-scores (as Figure 4) for over- and under-expressed Eμ-*myc* tumour proteins. **C.** GO term enrichment analysis for differential expression specific to Eμ-*myc* tumours ($Rs_{Eμ\text{-}myc}$ >0.5/<-0.5, p<0.05, $Rs_{Eμ\text{-}TCL1}$ <0.25/>-0.25) as described in Figure 4. **D-E.** StringDB networks derived from Eμ-*myc*-specific differentially expressed proteins annotated with **D.** 'methylation' (GO:0032259) (n=50), and **E.** 'actin cytoskeleton organisation' (GO:0030036) (n=49). **F.** Western blot evaluation of the three 'actin cytoskeleton organisation'-annotated proteins, coronin 1a (COR1A), lymphocyte cytosolic protein 1 (LCP1) and actin regulatory protein CAPG, expression in Eμ-*myc* tumours, relative to non-tumour B-cells. Two addition cytoskeletal proteins, myosin-9 (MYH9) and moesin (MOE) are also evaluated alongside myc.

Figure 6. Eμ-*TCL1*-specific tumour characteristics. A. Heat maps for proteins and surface proteins specific to Eμ-*TCL1* tumours (Rs<0.25/>-0.25 for Eμ-*myc* tumours). **B.** Upstream regulator activation z-scores (as Figure 4) for over- and under-expressed Eμ-*TCL1* tumour proteins. **C.** GO term enrichment analysis for differential expression specific to Eμ-*TCL1* tumours ($Rs_{Eμ\text{-}TCL1}$ >0.5/<-0.5, p<0.05, $Rs_{Eμ\text{-}myc}$ <0.25/>-0.25). **D.** A StringDB network derived from Eμ-*TCL1*-specific differentially

expressed proteins demonstrating interaction scores of >0.7 highlighting processes in protein clusters. (ER; endoplasmic reticulum). **E.** Individual iTRAQ quantifications (as $\log_2$ (ratios to WT)) for the peptides uniquely matching to IL5RA and it receptor partner CSF2RB (IL5RB). **F.** Flow cytometry evaluation of IL5 receptor alpha subunit (IL5Rα) expression on two Eμ-*TCL1* tumours, compared to splenic B cells pooled and isolated from three 2 month old WT and Eμ-*TCL1* mice.

Figure 7. IL5 drives Eμ-*TCL1* tumour proliferation. A. Tumour cells from three spleens of terminal Eμ-*TCL1* mice were cultured at an initial density of $5x10^6$ cells/ml and treated in with IL5. Cell density was measured after 48h. **B.** Cell division analysis measured by CFSE staining of terminal Eμ-*TCL1* tumours treated *in vitro* with 0, 10 or 100 ng/ml IL5 for 48 hours **C.** IL5-treated Eμ-*TCL1* tumours (48h) were stained with hypotonic propidium iodide and DNA content measured with flow cytometry to determine the proportion of cells entering mitosis. **D.** Serum starved (4h) Eμ-*TCL1* tumour cells were evaluated for IL5 treatment dose response after 30 minutes by Western blotting for the phosphorylation of AKT, GSK3, ERK, SRC, S6 and S6K. **E.** Serum starved (4h) Eμ-*TCL1* tumour cells were treated with IL5 in the presence of either TGX-221 (a specific PI3Kβ inhibitor) or DMSO, and the phosphorylation of AKT measured by Western blotting. **F.** Eμ-*TCL1* tumour cells were incubated for 48h with or without IL5 in the presence of either TGX-221 (5μM) or DMSO and the percentage of non-viable cells measured by Annexin V/PI. **G.** Summary of potential signal transduction pathway components, overlaid with Eμ-*TCL1* tumour proteomics quantitations, highlighting PIK3C2B as a mechanism of IL5-induced AKT phosphorylation, amplified by TCL1 and inhibited by the PI3Kβ-specific inhibitor TGX-221.

Figure 8. Model-dominant plasma signatures of Eμ-*myc* and Eμ-*TCL1* tumours. A. Inter-proteome comparison detailing the canonical cellular localisations, annotated by Ingenuity Pathway Analysis (IPA), of proteins commonly or discretely quantified in plasma or B-cells. **B.** The relative plasma protein abundances to WT plasma for terminal Eμ-*myc* and Eμ-*TCL1* tumours, highlighting those proteins annotated as intracellular (cytosolic/nuclear proteins) and extracellular. **C.** GO term enrichment for plasma proteins over-abundant in both tumours, considering all plasma proteins and additional analysing only those extracellular proteins not traceable to the B-cell tumour proteome, termed the 'tumour lysis-free' plasma. **D.** StringDB networks highlighting interactions and relationships for proteins over-abundant in the 'tumour lysis-free' plasma for both tumours. **E.** GO term enrichment for plasma proteins over-abundant specific to each tumour ($Rs>0.5$ and $p<0.05$, $Rs_{other\ tumour}<0.5$). **F.** Proteins demonstrating progressive increases in abundance correlating with tumourigenesis in Eμ-*TCL1* mice.

Supplementary data

Supplementary methods

Figure S1. Proteomic sample characteristics. Details of the mice from which B-cell and plasma samples were derived. Plasma SuPrE-SEC traces. (Relating to **Figure 1**)

Figure S2. Western blot, flow cytometry and literature-derived validation and the characterisation of myc and TCL1 transgene products. (Relating to **Figure 2**)

Figure S3. Detailed results describing tumour protein differential expression (Relating to **Figure 3**)

Figure S4. Bioinformatics analyses detailing enriched pathways, upstream regulators, phosphopeptides and B-cell receptor signalling (Relating to **Figures 4-6**)

Figure S5. IL5Rα expression and IL5 signalling in Eμ-*TCL1* mice (Relating to **Figure 7**)

Figure S6. Plasma proteome analysis details and results (Relating to **Figure 8**)

Table S1. Full quantitative B-cell and plasma proteomes

Table S2. Comparative and 'lysis-free' proteomes

Table S3. Bioinformatics analyses; gene ontology enrichment analysis and upstream regulator analysis of both and individual tumours

Table S4. Phosphorylated peptides; demonstrating relative expression relative to WT phosphopeptide expression and parent protein expression in tumours