# B-Cell Receptor Signalling in Chronic Lymphocytic Leukaemia

**Tom Butler**

A thesis submitted for the degree of Doctor of Philosophy at the University of London

Centre for Haemato-Oncology
Barts Cancer Institute

# 1.Abstract

Chronic Lymphocytic Leukaemia (CLL) cells may depend on B-cell receptor (BCR) signalling as well as other microenvironmental survival signals. A peculiarity of CLL is that cells preserve IgD signalling in the presence of reduced IgM signalling, a pattern mimicking anergic B-cells, and consistent with autoantigen exposure. This aim of this thesis was to examine the differing roles of IgM and IgD in CLL.

IgM and IgD expression was examined in CLL cells from peripheral blood (PB) and lymph node (LN). Co-expression of IgM and IgD was common, but levels of expression of IgD and IgM vary independently within these compartments, implying they may have differing roles. Most PB CLL samples underwent calcium (Ca) flux after IgD ligation, with evidence of relative IgM anergy. Incubation of cells for 24h *in vitro* partially restored IgM Ca flux, further supporting an anergic model. A pro-survival role of the BCR was suggested by the finding that BCR ligation was associated with reduced apoptosis *in vitro*.

Mechanistic differences of IgM and IgD signalling were examined using mass-spectrometry based phosphoproteomics. Six CLL samples were compared to five tonsil controls, and >5000 unique phosphopeptides identified. CLL and tonsil phosphoproteomes were compared. BCR induced changes in phosphoproteins and kinases differed between IgM and IgD ligation, and between CLL and tonsillar B-cells. Recognised and novel pathways included BCR, spliceosome and Notch signalling. Evidence in support of anergic signalling in CLL was observed.

Anergic IgM signalling is contrasted with IgD as a dynamic process, different in the tissue compartments of CLL. Phosphoproteomics offers a powerful tool for interrogating intracellular signalling, with phosphorylation networks characterizing pathway topology. BCR signalling in healthy B-cells has not previously been studied using this approach and comparisons with CLL highlight known as well as novel pathways that may well represent novel treatment targets.

**Table of Contents**

**List of Tables**

**List of Figures**

**List of Abbreviations**

| | |
|---|---|
| **Ag** | antigen |
| **AID** | activation-induced cytidine deaminase |
| **ANA** | antinuclear antigen |
| **AnV** | Annexin V |
| **APRIL** | a proliferation-inducing ligand |
| **BAFF** | B-cell-activating factor of the tumour necrosis factor family |
| **BCR** | B-cell receptor |
| **BH** | Benjamini-Hochberg |
| **BM** | bone marrow |
| **BTK** | Bruton's tyrosine kinase |
| **Ca** | Calcium |
| **CDK** | cyclin dependent kinase |
| **CDR** | complementary determining region |
| **Chk2** | checkpoint kinase 2 |
| **CK** | casein kinase |
| **CLL** | Chronic Lymphocytic Leukemia |
| **CML** | chronic myeloid leukaemia |
| **CSF1R** | Colony-Stimulating Factor Receptor Kinase |
| **CSR** | class-switch recombination |
| **CV** | coefficient of variation |
| **del** | deletion |
| **DLBCL** | diffuse large B-cell lymphoma |
| **DMSO** | dimethyl sulphoxide |
| **DNA** | Deoxyribonucleic acid |
| **dsDNA** | double stranded DNA |
| **EBV** | Eptsrein-Barr Virus |
| **EDAT** | ethylenediamine tetraacetic acid |
| **EGFR** | Epidermal Growth Factor Receptor |
| **EGTA** | ethylene glycol tetraacetic acid |
| **ERK** | extracellular signal-regulated kinase |
| **FDR** | False Discovery Rate |
| **FITC** | Fluorescein isothiocyanate |
| **FLMN** | Filamin A |
| **FR** | framework region |
| **GAPDH** | Glyceraldehyde 3-phosphate dehydrogenase |
| **GC** | germinal centre |
| **GSK3** | Glycogen synthase kinase-3 |
| **HCV** | hepatitis C virus |
| **HDGF** | heterogeneous nuclear ribonucleoprotein C |
| **HEL** | hen egg lysozyme |
| **HNRPC** | heterogeneous nuclear ribonucleoprotein C |
| **HPF** | high power fields |

| | |
|---|---|
| **Ig** | immunoglobulin |
| **IgD** | Immunoglobulin D |
| **IgG** | Immunoglobulin G |
| **IGH** | immunoglobulin heavy chain |
| ***IGHV*** | immunoglobulin heavy chain V gene |
| **IGK** | immunoglobulin light chain kappa |
| **IGL** | immunoglobulin light chain lambda |
| **IgM** | Immunoglobulin M |
| **IMAC** | immobilized metal affinity chromatography |
| **IMS** | industrial methylated Spirits |
| **InsR** | insulin receptor |
| **IPA** | Ingenuity Pathway Analysis |
| **ITAM** | immunoreceptor tyrosine-based activation motif |
| **ITIM** | Immunoreceptor tyrosine-based inhibitory motif |
| **KEGG** | Kyoto Encyclopaedia of Genes and Genomes |
| **LC-MS** | Liquid Chromatography Mass Spectrometry |
| **LC-MS/MS** | liquid chromatography–tandem mass spectrometry |
| **LN** | lymph node |
| **M-CLL** | CLL with mutated IGHV genes |
| **mAb** | monoclonal antibody |
| **MALT** | mucosa associated lymphoid tissue |
| **MAPK** | Mitogen activated protein kinase |
| **MBL** | monoclonal B-cell lymphocytosis |
| **MFI** | Median Fluorescence Intensity |
| **mRNA** | messenger RNA |
| **MS** | Mass Spectrometry |
| **mTOR** | mammalian target of rapamycin |
| **MYHIIA** | Myosin heavy chain 2a |
| **NF-kB** | Nuclear factor kappa-light-chain-enhancer of activated B cells |
| **NK** | not known |
| **OS** | overall survival |
| **PAK** | p21 protein (Cdc42/Rac)-activated kinase |
| **PB** | peripheral blood |
| **PCA** | principal components analysis |
| **PFS** | progression free survival |
| **PI** | propidium iodide |
| **PI3K** | phosphoinositide-3-kinase |
| **PKACa** | protein kinase, cAMP-dependent, catalytic, alpha |
| **PKC** | Protein kinase C |
| **PLC** | Phospholipase C |
| **PMA** | phorbol 12-myristate 13-acetate |
| **PTEN** | phosphatase and tensin homologue deleted on chromosome 10 |
| **PTM** | post-translational modification |
| **PTPROt** | protein tyrosine phosphatase receptor type O |
| **RA** | Rheumatoid Arthritis |
| **RF** | rheumatoid factor |

| | |
|---|---|
| **RNA** | Ribonucleic acid |
| **RPMI** | Roswell Park Memorial Institute |
| **RS6** | 40S ribosomal protein S6 |
| **S/Ser** | Serine |
| **SF3B1** | splicing factor 3b, subunit 1 |
| **SHIP-1** | phosphatidylinositol 5-phosphatase |
| **SHM** | Somatic hypermutation |
| **SLE** | systemic lupus erythematosus |
| **SLL** | small lymphocytic lymphoma |
| **ssDNA** | single stranded DNA |
| **STRING** | *Search Tool for the Retrieval of Interacting Genes/Proteins* |
| **SYK** | spleen tyrosine kinase |
| **T/Thr** | Threonine |
| **TCR** | T Cell Receptor |
| **TF** | transcription factor |
| **TJ** | Tight Junction |
| **TLR** | Toll-like Receptor |
| **U/UM-CLL** | CLL with unmutated IGHV genes |
| **WHO** | World Health Organization |
| **XIC** | extracted ion chromatogram |
| **Y/Tyr** | Tyrosine |
| **ZAP70** | 70 kDa zeta-associated protein |

**Acknowledgements**

There are many individuals within the Centre for Haemato-Oncology and Barts Cancer Institute who have supported me; I mention a few in person.

Firstly, thanks to my primary supervisor John Gribben, who conceived this project, and provided the space and freedom to pursue it in my own idiosyncratic way. He has always supported and guided me with a big heart and intellect, and has been extremely patient. I also thank David Taussig for his supervision; he never let me avoid difficult questions.

Particular thanks to Pedro Cutillas and Alex Montoya for introducing me to the world of phosphoproteomics, they are carrying out pioneering work in this field. To Andrew Clear, Abi Lee, Maria Calaminici, Rita Coutinho and Andrew Owen for expertise in immunohistochemistry and image analysis. To Sameena Iqbal and her tissue bank team, without which translational research is impossible; Lynn Haddon, Carol Jennings and David Williamson, without which research is impossible; Rob Petty and Jacek Marzec, who are less confused by statistical analysis than the rest of us; Simone Juliger, Christiana Kitromilidou, Sunil Iyengar, Lenushka Maharaj, Aine McCarthy and Li Jia, for sharing the challenges of western blotting; Farideh Miraki-Moud, Guglielmo Rosignoli, John Riches, Fabienne McClanahan, Samir Agrawal and Tim Farren for illuminating the mysteries of flow cytometry; Alan Ramsay, Emanuela Carlotti and Andrew Lister for the opportunity to educate; Lauren Wallis, Essam Ghazaly and Andrew Brash for discussions; Simon Joel for guidance. Thanks to Simon Hallam, Bryan Young and Floriana Manodoro for sharing an office. Huge thanks to Paul Greaves for keeping me sane and showing me what a good clinician scientist can do. My biggest thanks go to Andrew McWilliams for everything else.

I would like to thank all the patients who have provided me with their samples, and letting me care for some of them in clinic. Finally, I would also like to thank Cancer Research UK and the Barts Cancer Institute for funding this work.

**Individual Contributions**

I confirm that the bulk of work in this thesis is my own. Alex Montoya performed the high-performance-liquid chromatography mass-spectrometry of samples I prepared. Alex Montoya and Pedro Cutillas performed the analysis of the raw data from this process. Andrew Clear performed much of the immunohistochemistry, Rita Coutinho some of the image analysis.

# 2.Introduction

## 2.1 Chronic Lymphocytic Leukaemia (CLL)

Chronic Lymphocytic Leukaemia (CLL) is the most common leukaemia in the western world. It affects predominantly elderly individuals, with fewer than a third of patients being under the age of 60 at presentation,[1] and has a median age of diagnosis of 72 years[2]. The incidence of CLL is 3.9 per 100,000 people per year, the incidence in men is nearly twice that of women and there is geographical variation in CLL incidence, with high incidences in North America and Europe[1,2]. The disease presents in the lymph nodes without leukaemic involvement in 5% of cases and the disease is then known as small lymphocytic lymphoma (SLL). In the current World Health Organization (WHO) classification CLL and SLL represent different aspects of the same disease (CLL/SLL).

The diagnosis is made by the detection of a clonal population of small B-lymphocytes/B-cells in peripheral blood (PB) or bone marrow (BM), or by lymph node (LN) biopsy showing cells expressing the characteristic morphology and immunophenotype. CLL cells express CD19, dim CD20, CD5, CD23, and CD79a and weakly express surface IgM and IgD. Expression of CD38 is variable and has prognostic significance in this disease[3,4]. The international workshop on Chronic Lymphocytic Leukemia (iwCLL) revised guidelines require a B-cell lymphocytosis of greater than 5000/µl maintained for more than 3 months with the cells expressing the characteristic immunophenotype to reach a diagnosis of CLL[5]. CLL is increasingly an incidental finding in asymptomatic individuals when a lymphocytosis is found at the time of a routine blood count. Intriguingly, the increasing use of immunophenotyping has led to the identification of apparently healthy individuals with circulating clonal B-cells, often with the characteristic phenotype of CLL, but below the 5000/µL threshold demanded by CLL guidelines, and without any symptoms attributable to CLL. These cases have been catered for by the creation of a new entity, monoclonal B-cell lymphocytosis (MBL). The prevalence of MBL in the population is higher than that of CLL, with estimates of 3-5%[6]. On follow up, it has been noted that most of these individuals

never achieve the B-cell count threshold required for a formal diagnosis of CLL, and the majority of those that do develop CLL do not require subsequent treatment. A model of a universal asymptomatic precursor state has been supported by the finding of preceding MBL in almost all stored samples of patients latterly diagnosed with CLL[6]. This pre-leukaemic state is analogous to the model of monoclonal gammopathy of undetermined significance (MGUS) as an asymptomatic precursor of multiple myeloma, and of the models suggesting adenomas and carcinomas-in-situ as the precursors to full-blown epithelial tumours[7].

It is suggested that MBL precedes all cases of CLL[6] and strikingly, data suggest that CLL-like cells are detectable in almost all adults older than 70 years, and can be identified in a high proportion of those in younger age groups, provided sensitive enough techniques are employed (immunophenotyping of at least 50mL of peripheral blood)[8]. The process of ageing results in a narrowing of the B-cell repertoire,[9] and a speculative model may suggest an evolutionary continuum from oligoclonal B-cell senescence to monoclonal B-cell lymphocytosis to full-blown CLL. Oligoclonality is evident amongst low-count MBL cases, illustrating the transitions from polyclonality to oligoclonality *en route* to monoclonal MBL and CLL[10]. Oligoclonality or clonality of the T-cell repertoire has also been demonstrated in older individuals as well as in MBL[11] and CLL[12]. More recently, it has been suggested that the apparently non-malignant haemopoietic stem cells of patients with CLL have an inherent tendency to progress to CLL[13]. Transplant of CLL patients' purified $CD34^{+}CD38^{-}$ haemopoietic stem cells into mice recapitulated polyclonal haemopoiesis, as one might expect with healthy stem cells, but a bias towards B-cell differentiation was observed, as well as an eventual development of B-cell clonality with ageing. These B-cell clones were characteristic of CLL in terms of immunophenotype and a biased repertoire of immunoglobulin gene usage, but the actual immunoglobulin genes used differed from the original CLL clone. These provocative data require replication but may suggest CLL patients exhibit defects at earlier stages of B-cell differentiation than have been hitherto described.

## 2.2 Treatment of CLL

CLL follows a highly variable clinical course. Approximately 25% of patients require therapy at diagnosis due to bone marrow failure or to symptoms such as bulky adenopathy, organomegaly, fatigue or B-symptoms such as fevers, night sweats, weight loss or extreme fatigue. Some individuals diagnosed at an early stage remain asymptomatic for the rest of their lives, with their lifespan unaffected by CLL, while others develop an aggressive form of the disease. Treatment guidelines state that therapy should be reserved for those with advanced, symptomatic or progressive leukaemia[5]. Treatment is considered palliative due to the incurable nature of the disease with conventional chemotherapeutic agents and the often-advanced age of the patient. Several trials involving more than 2,000 patients with early disease have shown no survival benefit to treating patients with alkylating chemotherapeutic agents with early versus deferred therapy [14].

More recently, there has been improvement in the overall survival of CLL, particularly since the 1980s and in older age groups. This is not explained by lead-time bias due to the greater use of automated blood counters picking up early-stage disease[1]. There has been significant improvement over the past decade in the results of treatment of CLL with the use of combination chemotherapy and chemo-immunotherapy. This has resulted in improvements in response rates, complete remission (CR) rates and in progression free survival (PFS), but only recently has there been a demonstration of improved overall survival (OS) with chemo-immunotherapy[15].

The German CLL Study Group (GCLLSG) – led CLL8 study comparing the use of Fludarabine and Cyclophosphamide (FC) with Rituximab, Fludarabine and Cyclophosphamide (FCR) has shown that patients treated with FCR chemoimmunotherapy achieved a significant improvement in PFS and OS compared to patients treated with chemotherapy using FC alone[15]. Whilst trials showing improvements in survival with combination chemo-immunotherapy are heartening, there is still an unmet need: the CR rates in the CLL8 study were only 44% with FCR, and only 5% in patients with the high risk cytogenetic abnormality of 17p deletion. Toxicities were not negligible, and the trial was conducted in generally fit patients with

a median age of 61.[15] Failure to respond and relapse are still major challenges, and there is no standard of care for patients in this situation. The use of allogeneic stem cell transplantation has led to a long-term cure in the minority of patients for whom this approach is indicated. The efficacy, toxicity and cost of the various treatment regimes are highly variable.

In parallel with new therapeutic manoeuvres, there has been dramatic progress in the understanding of the basic biology of CLL, and the development of a panoply of prognostic factors that led many to hope that tailored therapy would improve responses. The knowledge gained with use of these prognostic factors has rarely led to clarification of the wisest plan for treatment of individual patients, but has provided important insights into CLL biology[16].

Newer agents are in development, and are showing promise in the laboratory and in initial trials[17-20]. An increasing understanding of the biology of the disease is driving these new treatments, but their appropriate use will need to be defined in clinical trials, and it is envisaged that no single agent will prove a 'magic bullet'. Because of CLL's incurable nature, and the often-advanced age of the patient at presentation, there is a need for better therapy, in terms of improved response rates and lower toxicities.

**2.3 Pathophysiology of CLL**

Cancers develop and progress via a combination of intrinsic cell changes interacting with extrinsic microenvironmental alterations[21,22]. There has been an increasing recognition that whilst a focus on intrinsic tumour cell abnormalities has led to profound insights, the surrounding tissues that do not form part of the malignant clone do have an important effect on tumour biology. These microenvironmental alterations are often considered secondary to intrinsic cell changes such as mutations in genes, with the malignant cell inducing abnormalities in surrounding non-malignant tissue such as stromal and immune cells. By inducing such perturbations in their microenvironment the cancer creates for itself a protective niche conducive to cancer cell survival and resistance to damage[21]. These mechanisms may also permit the evasion of the anti-tumour effects of the healthy immune system[23]. Conversely,

microenvironmental abnormalities may create or promote cell-intrinsic changes as a secondary effect. The tendency for squamous cell carcinomas (Marjolin Ulcers[24]) to arise in ulcers and old wounds has been recognised since the first century AD[25]. Persistent inflammation may promote the development of tumours by the continuous stimulation of cells by inflammatory mediators. Alternatively, an abnormal microenvironment may arise and provide a niche that allows cells with oncogenic mutations which would otherwise perish to 'seed' themselves in a favourable 'soil'[21]. It is difficult to establish causal primacy of microenvironment or cell-intrinsic abnormalities in the ultimate origins of malignancy, but is now clear that established tumours exhibit changes in both the malignant clone and the associated non-clonal tissues. It follows that therapy may need to target both mechanisms of tumourigenesis.

Much progress has been made in elucidating the cell-intrinsic genetic changes underlying CLL pathogenesis. A number of recurrent cytogenetic abnormalities have been identified, with aberrations detected by fluorescent in situ hybridization (FISH) being detected in most cases of CLL, and these convey prognostic significance[26]. The common chromosomal abnormalities observed include deletion of the long arm of chromosome 13 (del 13q), del 11q, trisomy 12, del 17p and del 6q[26,27]. These recurrent abnormalities point to the loci of candidate genes involved in pathogenesis. The most common abnormality is del 13q14, which occurs in 55% of cases, and has been associated with a favourable prognosis[26]. The first report linking microRNAs to cancer was in CLL, with the demonstration that two microRNA clusters, *mir-15a* and *mir16-1*, were located within the minimally deleted region at 13q14[28]. These microRNA deletions appear to act partly by enhancing expression of the anti-apoptotic protein Bcl-2. Mouse models suggest that the sizes of deletion at 13q14 influences the resulting CLL phenotypes, suggesting that more than simply these microRNAs are important[29].

Although it generally occurs in fewer than 10% of patients at diagnosis, del 17p is associated with rapid progression of disease, poor response to therapy and short survival, conveying the greatest impact on prognosis and treatment[26]. The deletion

involves the *TP53* locus at 17p13, and it is clear that mutations in the *TP53* gene can also contribute to disease progression and alter the sensitivity of CLL cells to chemotherapy agents[30,31]. Emerging data are demonstrating that concurrent *P53* mutation and 17p deletions occur more frequently than initially realised[32-37]. As well as predicting resistance to most commonly used therapeutic drugs, *P53* abnormalities often arise during disease progression, possibly as expansion of a previously unappreciated subclone[31]. Whilst guidelines on *P53* mutation analysis have been recently published[38], there is a currently unmet need in providing effective and non-toxic therapy for this group of patients.

Most recently, several studies employing modern high-throughput techniques to sequence CLL genomes have identified recurrent somatic mutations that may be presumed to be involved in pathogenesis[39,40]. Activating mutations of *NOTCH1* have been found in 5-15% of cases of CLL, with higher incidences in cases with poorer prognosis and chemorefractory disease[41], and lower frequency in the pre-leukaemic MBL[42,43]. Mutations in the splicing factor 3b, subunit 1 (*SF3B1*) are seen at a similar frequency, and are also associated with more aggressive disease[44,45]. These high-throughput sequencing methods also identified the known mutations of *P53* and *ATM* genes (5-15%) as well as new recurrent mutations in the *MYD88, XPO1, MAPK* and other genes at frequencies of a few percent each in large patient cohorts[40,44,46,47]. Whilst these recently discovered recurrent mutations will no doubt lead to new avenues for research and therapy, their low frequency highlights the fact that current methodologies suggest that there is no obvious unifying somatic DNA mutation that causes all or most cases of CLL. The scenario whereby each patient may have an infrequent or unique combination of mutations is becoming increasingly familiar in cancer research in general[48]. This leads to challenges in elucidating the causes and drivers of these cancers. Furthermore, designing sufficiently powerful randomised trials to test new therapies poses enormous challenges in an era of personalised genomic medicine, where each patient may have a unique combination of targetable mutations[48].

Single mutations have pleiotropic effects, and most cancers have multiple mutations. In leukaemias, there may be generally fewer mutations than in solid tumours (1 mutation/Mb in some haematologic cancers compared to 10-100 mutations/Mb in some melanomas[48]), but this still implies that several mutations interact and must be targeted in even the 'simpler' cancers. Whilst the ultimate causes of cancer are mutations in DNA, many insights have been gained into their effect on the malignant cell, manifested by abnormalities in epigenetics, gene expression, protein expression, apoptosis and proliferation. Some unifying principles may be derived when it is seen that apparently disparate mutations may influence only a few common pathways. In CLL, this may be illustrated by the finding of mutations of *MYD88, DD3X* and *MAPK*, each at low frequency, but all components of the same inflammatory pathway[44].

Further complexity arises when cell-intrinsic changes subsequent to mutations interact with microenvironmental abnormalities to produce the full phenotype of cancer[21], and it has been suggested that CLL has a particularly unique dependence on its microenvironment[22,49]. CLL cells are most obvious in the peripheral blood, and the ease of access to this bodily compartment explains why many researchers have focussed on cells taken from patients' blood samples. However, it has always been clear that CLL proliferates in the bone marrow (BM) and secondary lymphoid tissues such as the lymph nodes (LN) and spleen[50-52]. Whilst many insights into CLL pathogenesis have been gained by the study of the peripheral blood leukaemic lymphocytes, it is recognised that the most important events that drive the disease occur in these other tissues, and there are differences between leukaemic cells from these varying compartments, even within the same patient [53,54].

The CLL microenvironment is composed of T-cells, monocyte-derived nurse-like cells mesenchymal stromal cells, extracellular matrix components and soluble factors[22,55]. CLL cells cultured *in vitro* generally undergo apoptosis without additional support in the form of cytokines or stromal cell support[56]. Interactions with these cellular and other components appear to be necessary for CLL cell survival and proliferation. In addition, resistance to chemotherapeutic agents is conferred by stimuli received by these microenvironmental elements[22,57-59]. Trafficking and tissue retention are

regulated by chemokines and their receptors[60]. Signals from T-cells such as CD40 ligand, B-cell-activating factor of the tumour necrosis factor family (BAFF) and a proliferation-inducing ligand (APRIL) are essential for normal B-cell development and the generation of immune responses, and CLL cells are hyperresponsive to these T-cell-derived survival factors[61-63]. Once stimulated by their cellular neighbours and chemokines, CLL cells also recruit further accessory cells by the secretion of cytokines such as CCL3, demonstrating that the malignant cells actively shape their microenvironment[22,64]. Numerous other factors felt to influence CLL cell survival via their role in its microenvironment include CD38, CD49d, TLRs, CXCR4, and VLA-4[22,65-67]. Many of these factors were identified as important to CLL pathophysiology via the influence of their expression level on prognosis.

The role of T-cells in B-cell development appears to be corrupted by CLL in two fashions. Firstly, they may promote survival of the leukaemic cell. The microenvironment of the lymph node (LN) or bone marrow (BM) seems to provide anti-apoptotic and proliferative stimuli that result in the formation of characteristic proliferation centres or pseudofollicles, an entity not seen in other B-cell lymphomas[51]. It has been speculated that CLL cells within these histological structures recruit accessory cells to create a microenvironment that mimics the germinal centre in which B-cells respond to antigenic stimulus[68,69]. The *in vitro* stimulation of CD40 on CLL cells via CD40L on T-cells recapitulates the physiological immune response whereby T-cells support B-cell growth in the germinal centre, and is synergistic with B-cell Receptor (BCR) signalling[52]. Several anti-apoptotic pathways are induced, including survivin[70] and NF-kB[71] .

Furthermore, T-cells and related Natural Killer cells have an important role to play in eliminating tumours. CLL and other cancers can create an immune defect that permits escape of the tumour cell from immunosurveillance[72-75]. Our group has shown how CLL induces changes in T-cells, with notable changes in gene expression[76] and defective immune synapse formation[23]. Targeting the immune defect in CLL by use of immunomodulatory drugs such as lenalidomide reverses some of these abnormalities and has shown clinical efficacy in CLL[23,77].

Disrupting the interaction of the CLL cells with their microenvironment is beginning to show promising therapeutic effects in CLL[22,78]. Interruption of the CXCR4-CXCL12 axis can mobilise CLL cells from their protective microenvironmental niches in the tissues, rendering them more susceptible to attack with conventional chemotherapeutic agents[66]. Some of the more promising novel agents are small molecule inhibitors of pathways downstream of the B-cell receptor (BCR) [22,79,80]. The rationale for their use is based on the premise that one of the most important microenvironmental influences on CLL survival is received via the BCR. There is evidence to suggest that antigen in the microenvironment may act as a survival factor via its role as the BCR ligand. It is suggested that the malignant B-cell recapitulates the interaction of a healthy B-cell with its specific antigen, a process fundamental to the normal development of B-cells and the generation of an immune response. The role of the BCR in CLL is discussed further, but first the function of the BCR in normal B-cell development is described.

### 2.4 Normal B-cell development

To understand the classification and pathophysiology of malignant B-cells it has proved helpful to study the developmental stages of their normal counterparts[81]. B-cell development can be considered as occurring in two general stages: i) Antigen-independent differentiation of B-cell precursors from an haematopoietic stem cell (HSC) to naïve B-lymphocytes in the bone marrow (BM) and ii) the partially antigen-dependent maturation to memory/effector cells in secondary lymphoid tissues: lymph node (LN), mucosa associated lymphoid tissue (MALT), BM and spleen. Trafficking between these compartments is reflected in B-cell subsets present in peripheral blood (PB)[82]. The ultimate drive in B-cell development is the production of a flexible immune cell repertoire that is capable of recognising and eliminating a huge range of microbes that would otherwise cause tissue damage and death of the organism.

The stages in B-cell maturation have been mainly classified by immunohistochemical examination of the BM and LN[82]. Correspondingly, leukaemias and lymphomas have been classified according to similarity to normal B-cell morphology, location and immunophenotype in BM, LN and spleen[81,83]. Murine models have also been used to aid classification of normal and malignant B-cells, however it is clear that mouse B-cell

subsets do not precisely reflect those of B-cells in humans[84]. The cell of origin hypothesis suggests that a B-cell malignancy arises from transformation of these normal counterparts, and therefore the classification of leukaemias and lymphomas recapitulates the classification of B-cells during their ontogeny[81].

Key to understanding the development of healthy and malignant B-cells is consideration of the generation of antigen receptor diversity. The immune system is able to recognise a vast array of foreign antigens, mediated in B-cells by the BCR. The BCR complex is composed of two components: the recognition unit constituted by surface membrane-bound immunoglobulin (sIg) and the transmission unit composed of co-receptors CD79a (Ig$\alpha$, mb-1) and CD79b (Ig$\beta$, B29)[85]. The sIg is a heterodimer composed of 2 heavy chains (IGH) and 2 light chains (IGK or IGL) bonded by disulfide bridges. Each chain has 2 distinct parts: the N-terminal variable (V) domain responsible for antigen binding, and the C-terminal constant (C) domain responsible for membrane insertion and various effector functions. The C region is the determinant of the immunoglobulin isotype of the BCR or antibody. There are 5 different isotypes: IgM, IgG, IgA, IgE and IgD[86]. When inserted in the B-cell membrane they may exhibit differences in the downstream pathways activated by ligation of antigen, but for immunologists their more manifest role lies in the fact that B-cells at different stages of development express different isotypes and produce different classes of secreted antibody. The expression of IgM and IgD is generally associated with early stages of B-cell development, whilst IgG, IgA and IgE expressing cells are described as class-switched and reflect later developmental stages. Immunoglobulin secreted by B-cells forms antibody, and each antibody isotype has a different function that the immune system deploys depending on the nature of the antigenic challenge[82,87,88].

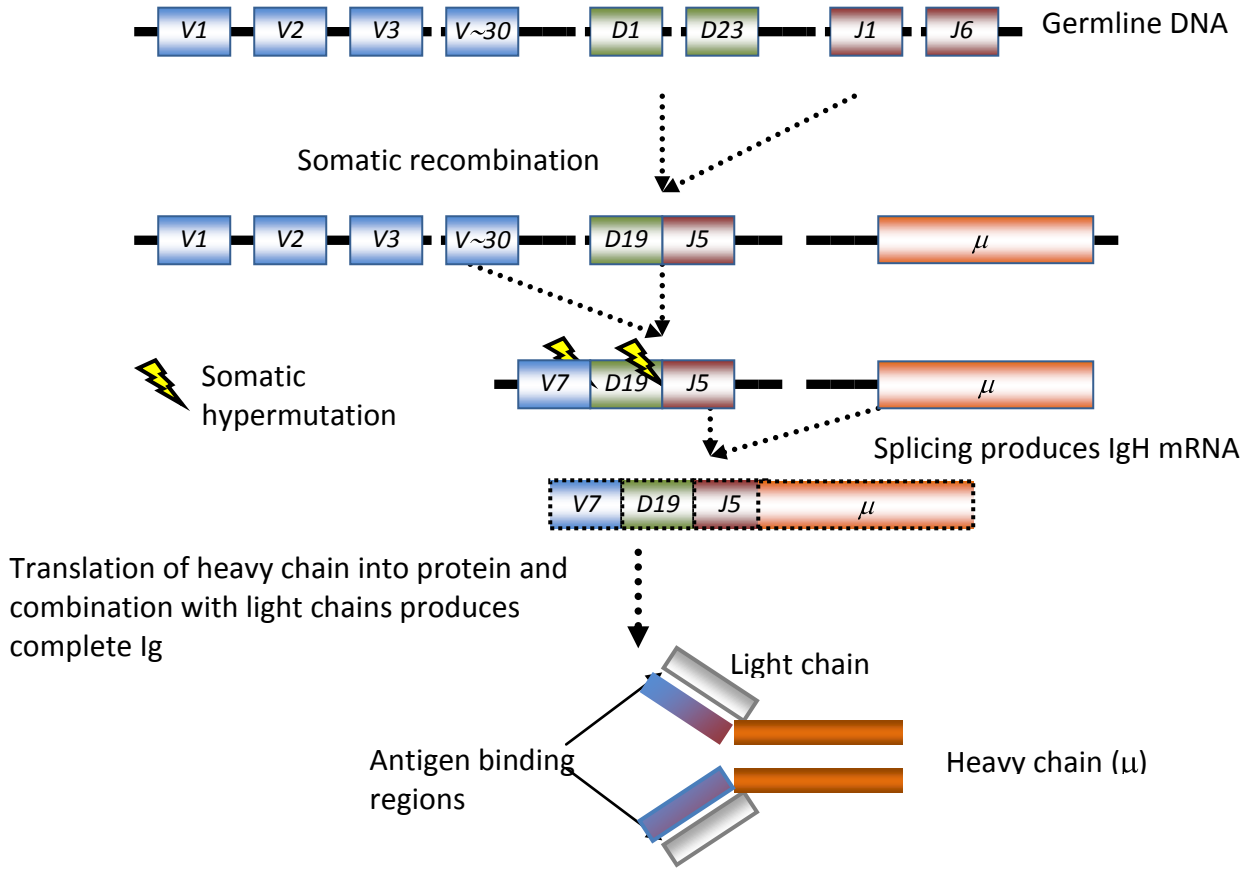
Fundamental to the generation of antigen receptor diversity in B-cells is the unique genomic organisation in the germline that permits rearrangement of the immunoglobulin gene loci during B-cell ontogeny[89]. The genes encoding the immunoglobulin heavy chain (IGH) are located on chromosome 14q32, span about 1.23Mb and are organised in 4 clusters *V*, *D*, *J* and *C* in a 5'-3' orientation. There are 123-129 *IGHV* genes on chromosome 14, depending on the haplotype, of which 38-46

are functional. Allele variants have been described for many *IGHV* genes, with some genes (e.g.*IGHV1-69* and *IGHV4-34*) having up to 13 alleles. There are 27 *IGHD* genes (23 functional) and 9 *IGHJ* genes (6 functional). Finally, there are 9 functional *IGHC* genes[90].

The light chain genes are located on chromosomes 2 and 22. The *IGK* locus is on 2p11, with 34-38 functional *IGKV* genes, 5 functional *IGHKJ* genes and a single *IGKC* gene. The *IGL* locus is at 22q11, with 29-33 functional *IGLV* genes and 4-5 functional *IGLC* genes[90].

The antigen-binding V domain of heavy chains (VH) is encoded by the variable (*V*), diversity (*D*) and joining (*J*) genes, while the V domain of light chains (VL) is encoded by *V* and *J* genes [90]. The combination of V,D,J and C genes during B-cell development results in the generation of messenger RNA (mRNA) encoding the heavy and light chains that compose the full immunoglobulin (Ig) molecule (see Fig.2.1)

**Figure 2.1: Generation of the primary B-cell repertoire**
**Variation in the antigen-binding region of the B-cell Receptor is generated by random DNA recombination of *V*, *D* and *J* gene segments. Further diversity is produced by introduction of mutations in the *V*, *D* and *J* genes by somatic hypermutation during the germinal centre reaction.**




The V domain contains 4 relatively conserved framework regions (FR) that maintain the overall structure of the domain. In between are 3 highly variable complementarity determining regions (CDRs) that form loops that directly interact with antigens. The CDR3 is at the junction of the V, D and J domains and has the highest variability and influence on antigen binding[90].

Assembly of the DNA encoding the Ig heavy and light chains occurs at particular stages of B-cell development. The stage of development is implied by the state of organisation of the *IG* loci, and the expression of various proteins involved in the Ig production process[89]. The first stage in the foetal BM and liver is antigen-independent.

It consists of the assembly, by DNA rearrangement, of the *V*, *D* and *J* genes to encode variable domain DNA for first heavy then light chains. These are then transcribed and translated into proteins that are assembled into Ig that is expressed on the cell surface as the pre-BCR and subsequent BCR. Only B-cell progenitors that successfully assemble and express surface BCR survive to subsequent stages, in a process of positive selection. A subsequent negative selection process then operates to eliminate autoreactive B-cells: central tolerance[91]. In order to eliminate B-cells with a BCR that is reactive against self-antigens, those B-cells whose BCR is strongly stimulated by antigen within the bone marrow are deleted from the B-cell repertoire[82]. Alternatively, immature B-cells that bind self-antigens in the BM also have an opportunity to rearrange their *IGL* loci in an attempt to eliminate BCR autoreactivity (receptor editing)[92]. If they fail to produce a non-autoreactive BCR after receptor editing, they will be deleted. Immature B-cells that successfully navigate this negative selection step become naïve mature B-cells, express the BCR as IgM and IgD, and leave the BM to circulate in the PB[82]. This process of central tolerance by negative selection cannot eliminate all autoreactive clones, and additional mechanisms such as anergy exist to control autoreactivity at later developmental stages. Indeed, it has been suggested that a degree of autoreactivity of the pre-BCR and BCR is necessary for positive selection[93]. Positive and negative selection may be perhaps considered within the context of a continuum of BCR autoreactivity[94].

The next developmental stage occurs in the periphery, within the secondary lymphoid organs. Naïve B-cells that have exited the BM circulate in the PB and enter the LN and spleen at the T-cell zone through high endothelial venules[95]. They recirculate between the PB and lymphoid tissues, and will die within several days if they fail to locate their cognate antigen (Ag). Binding of antigen to the BCR in the presence of appropriate accessory signals results in activation of the B-cell, resulting in proliferation and antibody production and the elimination of the offending antigen. Depending on the type and strength of BCR signals from their microenvironment they will become either plasma cells, marginal zone or follicular B-cells[82]. The immune system further refines the affinity and function of B-cells during the germinal centre (GC) reaction[96].

In the classical scenario of a B-cell binding a T-dependent antigen, follicular B-cells and helper T-cells migrate into the primary follicles of the lymphoid organs. These then mature into secondary follicles with a germinal centre[97]. The B-cells undergo intense proliferation as centroblasts in the dark zone of the germinal centre. Somatic hypermutation (SHM) introduces random point mutations, deletions and insertions in the variable regions of the immunoglobulin genes at rates of $10^{-3}$/base pair/division, about $10^{6}$ higher than the background spontaneous mutation rate[98]. Mutations extend 1.5-2kb from the promoter of the *V* gene, with a maximal frequency on the rearranged *V(D)J* genes, sparing the *C* region. Non-Ig genes such as *BCL-6* can also be targeted at a lower frequency, occasionally resulting in oncogenic mutations and translocations[99,100]. The initial step is mediated by activation-induced cytidine deaminase (AID), which deaminates C to U, followed by error-prone repair by generic repair mechanisms to produce sequence changes[97]. SHM preferentially targets particular sequence motifs, and can be silent (S) or result in replacement (R) of an amino acid at a given position. Mutations may occur in the framework region (FR) or the antigen-binding complementarity-determining region (CDR). Because higher affinity BCRs are selected for in the GC reaction, R mutations will tend to accumulate in CDRs, whilst S mutations in FRs are selected for in order to preserve the overall structure of the domain. Thus, raised R/S ratios in CDR compared to FR can imply evidence of antigen selection[90].

These mutations produce a wide diversity of BCRs, related to the original clone, but with varying binding affinities for antigen. B-cells then migrate to the light zone of germinal centres, where they compete with one another to bind antigen displayed on the surface of follicular dendritic cells[97]. High affinity for antigen results in survival signals and further proliferation, whilst relatively low affinity BCRs will lead to apoptosis, which is the fate of the majority of B-cells in the proliferating and diversifying clone. T-cell help is necessary for the GC reaction[97]. Each B-cell undergoes Darwinian selection by competing for survival signals received from antigen and T-cells. This affinity maturation process refines the binding of BCRs and secreted immunoglobulins to antigen, thereby enhancing the immune response to microbial invaders.

In parallel with affinity maturation, centrocytes also undergo isotype class-switch recombination (CSR) at the genomic level[97]. Class Switch of Ig isotype allows the production of antibodies that perform distinct effector functions[101]. CSR generally requires T-cell help through the engagement of CD40 receptors and the secretion of various cytokines such as IL-4 and IL-10[101]. The nature of the isotype depends on the type of antigen, the site of stimulation and the type of T-cell help provided. CSR is mediated by the deletional recombination of specific DNA sequences (switch sites) located upstream of each *C* region. The *IGHM* switch region is recombined to that of another *C* region and the intervening DNA is deleted by loop excision, leading to replacement of the original Cμ region with either Cγ, Cα or Cε[101]. This results in the same *VDJ* rearrangement being apposed to a new *C* region, resulting in production of a new isotype with the same antigen specificity. Its molecular control has many similarities with SHM, including dependence on AID. Whilst SHM and CSR classically and most frequently occur in the GC, the process can occur occasionally in marginal zone B-cells via a GC-independent route[102]. The CSR process results in either the production of high-affinity class-switched memory B-cells or differentiated plasma cells, which then secrete Ig as antibody[82].

The stochastic combinatorial process of V domain formation produces approximately $2x10^6$ different possible combinations of heavy and light chains. Further diversity is also added via a number of mechanisms such as modification of the junctions between *V, D* and *J* genes, and SHM. These mechanisms for diversity generation result in the production of $10^{11}$-$10^{12}$ different BCRs within one individual, sufficient to bind almost all conceivable antigens that may be encountered during that individual's lifetime[90]. The actual expressed repertoire within an individual might be more limited, with studies showing that the normal blood repertoire is biased with certain *V* genes being over-represented (such as *IGHV4-34, IGHV3-23*) and others under-represented. This bias may arise due to rearrangement efficiency, predilection for particular light and heavy chain pairings, selection at various stages, type of B-cell subset, chronic viral infection, age and genetic factors[103,104]. This bias must be taken into consideration when considering the repertoire of B-cells, but extensive data on healthy B-cell *V* gene usage are unavailable.

The complex process of Ig gene rearrangements leaves a unique signature in the genome of each B-cell. This can provide investigators with a read-out of the life history of an individual B-cell, confirm the clonal nature of a B-cell neoplasm, and imply the putative cell of origin of B-cell tumours[81]. The entire process of diversity generation is dependent on DNA damage and chromosomal rearrangements that occasionally result in mutations and translocations that may cause lymphoma[99,100].

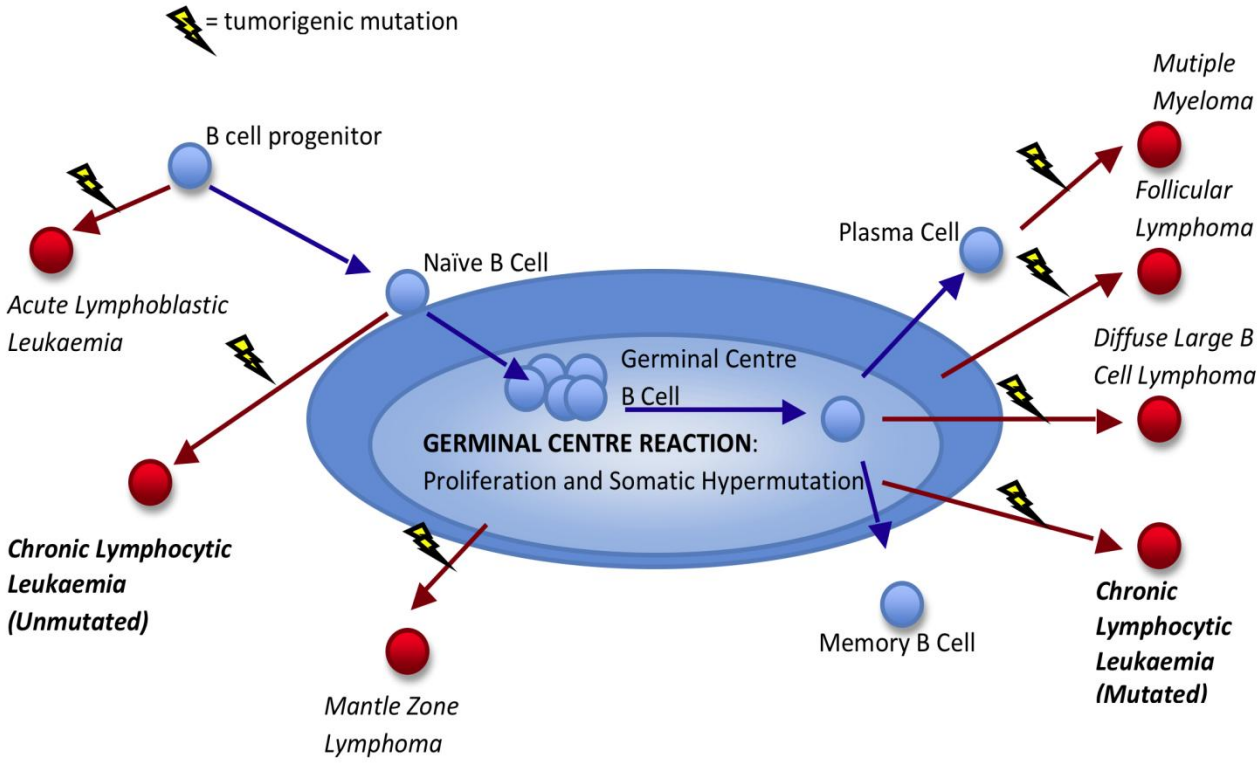
**2.5 Cellular origin of B-cell lymphomas and leukaemias**

Debates exist regarding the nature of a putative 'cell of origin' in cancer, the normal cell that acquires the first cancer-promoting mutation(s). It is important to distinguish this from the 'cancer stem cell', the putative cellular subset within a cancer that sustains malignant growth[105,106]. The classification of lymphomas and lymphoid leukaemias employs various techniques, including histology and immunophenotyping. In some lymphomas these approaches suggest an obvious healthy B-cell developmental stage as a counterpart, and by extension a presumed cellular origin, but in many cases the parallel between a lymphoma and an equivalent normal tissue cannot be drawn by tissue morphology or immunophenotyping alone. Analysis of the DNA encoding the *IGHV* genes has led to the ability to study the life history of an individual B-cell and an elucidation of the model of the equivalent normal 'cell of origin' of many lymphomas (see figure 2.2) **[81].** It is important to realise that whilst this model is highly successful in making sense of lymphoma ontogeny, it does not provide a complete picture of classification. Obviously, malignant cells will exhibit phenotypic and genomic differences from normal equivalents. Several entities known as 'gray-zone' lymphomas are not readily classified using this schema[107]. Furthermore, the similarity of phenotype and genotype of a lymphoma and a putative normal counterpart does not in itself constitute definitive proof of a tissue of origin[105]. What is generally assumed is that the B-cell will proceed along known developmental pathways at least until the initial transforming mutation, and possibly after this event. It might also be expected that if B-cell development does proceed after transformation, it will not necessarily proceed in a conventional fashion. To overlay further complexity, the process of SHM, classically thought to be a marker of the

germinal centre reaction, has been shown to also occur via other developmental pathways, albeit at low frequency[108]. Therefore, the cell of origin model does not provide a definitive cell of origin assignment in all lymphomas. This uncertainty is very apparent in the case of CLL[109-111].

**Figure 2.2: Cellular Origin of B-cell lymphomas and leukaemias (after Kuppers et al)[81]**

**The developmental stages of the normal B-cell are shown, with passage through the germinal centre and antigenic stimulation resulting in Somatic Hypermutation (SHM). The presence or absence of SHM and other markers can indicate the life history of an individual cell. The cell of origin hypothesis implies that B-cell lymphomas and leukaemias arise from a B-cell at a particular stage of development that has acquired mutations that convert a healthy B-cell into a neoplastic counterpart. The characteristics of the lymphoma cell are similar to the cell of origin, including the presence of SHM. The existence of CLL cells with and without SHM implies different cells of origin for these two subgroups, which also differ in terms of disease aggressiveness.**




When the *IGHV* genes of B-cell lymphomas were sequenced, most carried somatic mutations[81]. The types exhibiting SHM include follicular lymphoma, Burkitt lymphoma, diffuse large B-cell lymphoma (DLBCL), mucosa-associate lymphoid tissue (MALT) lymphomas, splenic lymphomas with villous lymphocytes, prolymphocytic leukaemias,

hairy cell leukaemias, lymphoplasmacytoid (LPC) lymphomas, multiple myeloma and some CLL[81]. The pattern of SHM is generally taken to indicate whether the lymphoma originates from a cell at a stage of development prior to the germinal centre (GC) reaction, or a post-GC cell. In follicular lymphoma, there is ongoing SHM within the tumour clone,[112] identifying it as a GC tumour. This also occurs in LPC lymphoma and occasionally in Burkitt, DLBCL and MALT lymphomas. Among mature B-cell lymphomas, unmutated *IGHV* genes are encountered in marginal zone lymphomas and some CLL, both $CD5^+$ B-cell tumours. This implies that they have arisen from B-cells that have not experienced the germinal centre.[81] The developmental stage of the B-cell that gives rise to CLL is unclear, as approximately half of CLL cases exhibit SHM (M-CLL) and half do not (U-CLL)[3,113]. The conventional threshold for assignment to the U-CLL group is that of >98% homology to germline *IGHV* gene sequences.

In summary, the study of rearranged variable-region genes in B-cell lymphomas/leukaemias has led to the model of lymphomas as B-cells transformed at different stages of B-cell development, with SHM implying experience of the germinal centre reaction[81]. The normal germinal centre processes of SHM and CSR may also produce translocations and mutations as an undesirable effect, leading to lymphoma pathogenesis. These findings have led to wide acceptance of the cell of origin model to explain some aspects of lymphomagenesis. However, because CLL appear to be comprised of two different subsets with and without SHM, the cell of origin is less obvious than with other lymphomas[109-111]. This may lead to difficulties in considering appropriate control healthy B-cells for comparison with CLL. This peculiarity of CLL also introduces the concept that the developmental routes whereby a B-cell produces a functional BCR are somewhat askew in CLL.

Further exploration of the configuration of the *IGHV* genes in these malignancies has led to the curious finding of a bias in the inherently stochastic process of *VDJ* gene rearrangements, and this is most apparent in CLL. This has led to the suggestion that antigenic selection may contribute to CLL pathogenesis[49,90], and many investigators therefore implicate the BCR as a critical molecule in CLL pathogenesis[49,114-117].

## 2.6 The role of Antigen in CLL pathogenesis

### *2.6.1 The BCR in lymphoid malignancies*

Inducible deletion of the BCR in mature cells leads to cell death[118]. The finding that BCR expression is essential for the survival of mature B-cells led to the concept of 'tonic' BCR signalling, whereby signals from an intact BCR are vital to resting B-cell existence, even in the absence of ligand binding[119-122]. Most neoplastic B-cells express surface immunoglobulin, and the pattern of SHM in their *IGHV* genes often indicates that they or their precursors were selected for antigen-receptor expression during the germinal centre reaction. Transforming chromosomal translocations to immunoglobulin loci are targeted to the non-rearranged allele[123], permitting expression of the BCR derived from the other allele. This has been taken as evidence that, like normal B-cells, most B-lymphoid malignancies require BCR expression[118]. The exception to this rule is Hodgkin lymphoma, where the malignant B-cells seem to arise from germinal centre B-cells that have lost the ability and requirement for expression of the BCR[81]. Even in this situation, the malignant cell may in some cases derive BCR-analogous survival signals from the Epstein-Barr Virus-encoded protein LMP2A that mimics an activated BCR[124].

Other features may also indicate a dependence of tumour growth on triggering by antigen. Indeed, a variety of infectious agents have been identified as associated with lymphoma, some in a directly causal fashion. It is estimated that worldwide, 16% of all cancers are caused by infectious agents[125]. The organisms include herpes viruses (Epstein Barr Virus is linked with Burkitt, central nervous system and Hodgkin lymphomas[126]; Human herpesvirus 8 is associated with Castleman's diseases and primary effusion lymphoma[127]), single-stranded RNA viruses (hepatitis C virus and splenic marginal zone lymphomas[128]), retroviruses (human T-cell leukaemia virus type 1 and adult T-cell lymphoma[129]), spirochaetal bacteria (*Borrelia burgdorferi* and cutaneous mucosa-associated tissue lymphoma[130]), and gastrointestinal bacteria (*Campylobacter jejuni* and immunoproliferative small intestinal disease[131]). There is a recognised association between the ocular anexae lymphoma subtype of marginal zone lymphoma and infection with *Chlamydia psitacci,* and recent data suggest that eradication of the infection is an effective treatment, in combination with

chemotherapy[132]. It is not suggested that all these microorganisms all exert their effects via BCR ligation, but that the principle of microbial lymphomagenesis is established.

A well-recognised association is between gastric marginal zone mucosa-associated (MALT) lymphoma and *Helicobacter pylori* infection[133], but again it is unclear as to the importance of BCR signalling as some effects are probably exerted via T-cells or more general inflammatory processes. Eradication of *H. pylori* at early stages of MALT lymphoma can result in disappearance of the tumour in 75% of cases[134], and the 25% of cases that do not respond to antibiotics are characterised by a clinically advanced stage and the t(11;18)(q21;q21) translocation. This translocation is felt to contribute to pathogenesis by activating the NF-kB pathway[135], known to be downstream of the BCR and other pathways. Therefore, it could be argued that whilst early stage gastric MALT lymphoma is dependent on microenvironmental signals from *H. pylori* infection, more advanced stages have acquired secondary lesions that result in constitutive activation of BCR signalling pathways, independent of antigenic stimulus. This concept is often invoked to suggest that more aggressive tumours in general have less dependence on microenvironmental signals than their more indolent counterparts, and that such independence may be gained by simulating microenvironmental support by constitutively activating pathways that are conventionally induced by external ligands[21,22]. This model is supported by the finding of multiple mutations in NF-kB signalling components in diffuse large B-cell lymphoma, which has been described as supporting lymphoma cell survival and proliferation by simulating 'chronic' BCR signalling[136-138].

It is therefore speculated that BCR signalling may contribute to lymphoma pathogenesis via a variety of mechanisms: tonic signalling in the absence of ligand is required for B-cell survival, ligand-dependent signalling provides a conduit for antigenic survival or proliferative signals from the microenvironment and chronic signalling may arise when pathways conventionally induced by ligand are constitutively activated. This highlights the centrality of the BCR to B-cells in both physiology and pathology.

### *2.6.2 Immunoglobulin genes in chronic lymphocytic leukaemia*

A variety of data have led many authors to posit the centrality of the BCR to CLL pathogenesis, mainly based on analyses of the genes encoding the antigen binding (V) regions of the BCR[49,90]. As well as suggesting divergent developmental pathways, the level of somatic hypermutation (SHM) in the *V* genes of individual CLL cases can also subdivide the disease into two patient groups with distinct clinical outcomes[3,113]. Some have therefore suggested that CLL, previously considered as a single entity, might best be considered as two distinct tumours, one arising from a cell at a stage prior to the germinal centre, and one arising after SHM has occurred[49]. The cases that have not undergone SHM (unmutated, U-CLL) have a more aggressive behaviour, as measured by shorter time to treatment and overall survival, than those that have undergone SHM (mutated, M-CLL). There are also differences in the pattern of *IGHV*, *IGKV* and *IGLV* gene utilization between the two groups, further implying separate developmental pathways without inter-conversion. The association of mutation status with prognosis is strong, and clinical trials are beginning to assess the use of this marker to guide therapy of CLL[16]. This association of SHM with prognosis is sometimes taken to be an indicator that normal immune mechanisms have an important influence over CLL biology, though the mechanistic details of how SHM affects disease aggressiveness in CLL are still obscure. More evidence for the role of the BCR in CLL has been provided by analysis of the *IGHV* gene repertoire used in CLL.

*IGHV* gene analysis in CLL reveals restrictions in which of the 38-46 functional *IGHV* genes a particular tumour expresses[90]. Similar analyses have found biases in gene use at the light chain loci, though most studies have focussed on the heavy chains. Usage of particular *IGHV* genes at frequencies higher than expected by chance is generally interpreted as an indicator of selective drive on a B-cell population via an antigen or superantigen[90]. Superantigens are mainly microbial products, often polymeric, which bind to framework regions of particular V domains[139,140]. For example, Staphylococcal protein A (SpA) binds to IGHV3 subgroup domains[140]. By binding outside the conventional antigen-binding site, superantigens can activate a large proportion of B-cells, with 30-40% activated by SpA. However, the restriction of the *IGHV* gene

repertoire in CLL is generally taken as evidence of antigenic rather than superantigenic drive as the antigen-binding regions exhibit marked peculiarities[90].

Restricted *IGHV* gene usage is a particular feature of CLL, the most striking being an increased usage of the *IGHV1-69* gene by U-CLL[113,141-143] and of *IGHV4-34* by M-CLL[113,142]. In contrast, this biased repertoire is not characteristic of follicular lymphoma, or of diffuse large B-cell lymphoma, though it does occur to a certain extent in some other mature B-cell neoplasms such as splenic marginal zone lymphoma[144] and mantle cell lymphoma[145].

The *IGHV1-69* gene bias is the most prominent in CLL. The expansion of B-cells using *IGHV1-69* has been noted in other conditions such as salivary gland extranodal marginal zone lymphoma of mucosa-associated tissue (salivary MALT)[146]. There is also a suggested link between *IGHV1-69* , hepatitis C virus (HCV) infection and splenic marginal zone lymphomas[147]. HCV can infect B-cells and cause chronic stimulation, with an increase in *IGHV1-69* expressing B-cells that respond to the HCV E2 antigen and expressed by the HCV-associated lymphomas[148]. HCV is also a major association of mixed cryoglobulinaemia (MC), which can proceed to lymphoma[149]. The IgM rheumatoid factors produced in this disease are commonly derived from *IGHV1-69* and tend to be combined with the *IGKV3-20* (A27) light chain, a combination also frequently found in CLL[150]. Suggestion of molecular mimicry linking autoimmunity and viral infection is indicated by the finding of a cross-reactivity between the IgG-Fc recognized by rheumatoid factors in MC and the NS3 protein of HCV[151]. It is not suggested that HCV plays a role in CLL, and the lack of homologous CDR3 sequences argues against specificity for HCV in CLL, but this is an example whereby stimulation of B-cells by antigen can lead to both *IGHV* gene bias and provide the circumstances for transforming events that may lead to lymphoma.

The second most utilised *IGHV* gene in CLL is *IGHV4-34*[90]. The *IGHV4-34* gene encodes intrinsically autoreactive antibodies because of universal recognition of the N-acetyllactosamine (NAL) antigenic determinant present in the I/i blood group antigen[152] and in the B-cell isoform of CD45[153]. A hydrophobic cluster of amino acids in

the FR1 of IGHV4-34 binds directly to the red cell I/i carbohydrate antigen[154], therefore this autoantigen acts as a superantigen. The I/i antigen may be expressed in oxidised apoptotic cells, and B-cells whose receptors bind to apoptotic cells may serve 'housekeeping' functions to remove cellular debris[155].

The *IGHV4-34* gene is over-expressed in cold agglutinin disease[156], and in 40% of primary central nervous system lymphomas[157-159]. All cases of cold agglutinin disease secrete IgM derived from this gene, and its induction by EBV accounts for the known association between EBV infection and transient cold agglutination[160]. A rise in IGHV4-34 antibodies is seen in acute infections with Epstein-Barr Virus (EBV)
and infection has major effects on the immune system of older people and may drive both CLL and the age-associated *IGHV4-34* expansion[126,161,162].

*IGHV4-34*-utilising B-cells are also increased in systemic lupus erythematosus (SLE), and encode anti-DNA antibodies[163,164]. *IGHV4-34* expression occurs at a high frequency in the B-cell progenitors of healthy individuals, however they are censored at multiple checkpoints during development[155,165]. To explain the low frequency of actual *IGHV4-34* usage in the mature healthy repertoire, it has been suggested that in normal B-cell development, *IGHV4-34* B-cells are excluded from the GC to prevent autoimmune disease[155]. It may be that this blockade is overcome in autoimmune disease and in particular infections.

The *IGHV4-34* gene is frequently used in CLL, particularly in M-CLL[166], perhaps reflecting the need for the inherently *IGHV4-34* sequences used by B-cell precursors of CLL to undergo SHM to lose autoreactivity and be permitted to enter the repertoire[142]. Support for this hypothesis was provided by Chiorazzi's group who cloned and expressed in vitro recombinant antibodies from M- and U-CLL B-cells and tested their reactivity to a panel of antigens[167]. U-CLL expressed highly polyreactive antibodies whereas most M-CLL B-cells had a more restricted reactivity profile. When M-CLL non-autoreactive CLL antibody sequences were reverted to their configuration that would have existed prior to SHM, they subsequently encoded polyreactive and autoreactive antibodies. They speculate that both U-CLLs and M-CLLs originate from self-reactive B-

cell precursors and that the influence of SHM on prognosis may be understood as actually decreasing BCR autoreactivity[167]. This may explain why M-CLL cases have a better prognosis than U-CLL: The process of SHM actually reduces polyreactivity and therefore potentially the range or binding affinity of available ligands for use as survival factors.

### *2.6.3 Stereotyped B-cell receptors in CLL*

Biases in *IGHV* gene use alone are evidence for antigenic selection, but further insight has been provided by a more detailed examination of the structures of the antigen-binding CDR3 domains that have the greatest influence on antigen binding. In the mid 1990s, two groups independently demonstrated that unrelated cases, especially those using the *IGHV1-69* gene, had distinctive VH CDR3 domains characterized by increased length and shared amino acid motifs[143,168]. There was also evidence of favoured combinations of genes, for example 90% of the *IGHV3-7* genes were associated with the *IGHJ4* gene, whereas 50% of the *IGHV1-69* and *IGHV4-34* genes were associated with the *IGHJ6* gene. Similar findings are apparent when other gene combinations are considered[104,169-173].

It seems that particular *V/D/J* and light chain gene combinations are favoured in different patients with CLL. These are termed 'stereotyped', and criteria for stereotypy have been described[174]. To fulfil these criteria, there must be usage of the same *IGHV/D/J* genes, the same *IGHD* reading frame, and VH CDR3 amino acid homology ≥60%.

Subsequent studies have extended these findings, with large international series suggesting that up to 30% of CLL cases (particularly U-CLL) exhibit restricted and stereotyped CDR3s and may be grouped into >100 subsets[111,142,175]. The largest series (2662 patients) found that the CDR3 sequences of the 30% of cases expressing stereotyped receptors formed 11 higher order clusters[111]. Some of these stereotyped sequences are homologous to those seen in mouse B-1 B-cells, leading the authors to make the suggestion that the 30% of CLL cases exhibiting stereotyped CDR3s may arise

from the human equivalent, via a T-independent route, which would explain the association of stereotypy with U-CLL.

There is some evidence for biased *IGHV* gene utilisation having influences on other areas of CLL biology. The risk of transformation of CLL to an aggressive lymphoma in Richter's Syndrome (RS) is 3-16%. Stereotyped CDR3s occur in 50% of RS cases, in particular the *IGHV4-39* subset[176](subset 8). *IGHV4-39* has no prognostic relevance for CLL progression *per se*, whereas the well-described poor prognostic factor *IGHV3-21* usage had no influence of risk for RS. There was also association with Trisomy 12, previously shown to associate with *IGHV4-39*[177]. Trisomy 12 is known to associate with RS in other studies, and with *NOTCH1* mutations[178]. An association of particular *IGHV* genes with microRNA abnormalities has also been suggested in CLL[179].

These examples of stereotypy in CLL are in contrast to the normal B-cell repertoire, where there is no reported preferential pairing between specific heavy and light chains. Comparison of CLL CDR3s with large public databases of non-CLL CDR3 sequences has shown only 0.7% sharing of homologous CDR3s between CLL and normal B-cells.[90] It has also been shown that VH CDR3 restriction was infrequent in other B-cell lymphomas, and resembled that of normal B-cells, suggesting that other lymphomas may arise stochastically from the normal B-cell repertoire[180]. It is not yet possible to predict specificity by sequence analysis alone, but comparisons with known antibody specificities has been performed and shown CLL cases with identical CDR3 structures to B-cells specific for autoantigens[181 182].

Since CLL occurs more commonly in older individuals, it has been suggested that there may be a shift in the normal B-cell repertoire with age, which would therefore be reflected in an equivalent shift in *V* gene usage in tumours that were stochastically selected from this ageing repertoire. There is no increase in *IGHV1-69* usage with age[183], however there is evidence of increased usage of *IGHV4-34* by the B-cells of healthy elderly indivduals[184]. This may be due to the reactivation of herpes viruses, and may offer a partial explanation of the high frequency of *IGHV4-34* usage in CLL[90]. Investigators have not yet conclusively answered these issues about pre-existing bias

in healthy B-cells. Could CLL arise stochastically from a B-cell subset that has an inherent restriction in its *IGHV* gene usage? There are early suggestions that some CLL stereotyped CDR3s have counterparts in the B-cells of healthy age-matched individuals, and it is suggested that these are the source population that may acquire transforming events that lead to some subgroups of leukaemia[103].

In summary, analysis of the gene sequences encoding the antigen-binding regions of the BCR has demonstrated marked biases in CLL, with some evidence of bias in similar lymphomas. The bias is even more apparent if the actual amino acid structure of these antigen-binding CDR3s is considered. These findings demonstrate that patients from opposite ends of the world may have B-cell receptors with identical antigen-binding regions, a finding that is highly unlikely given the stochastic nature of the generation of the antigen-binding capabilities of the BCR. The obvious implication is that these different cases of CLL are binding the same antigen. It must be noted that the same antigen is not suggested to be common to all CLL, as there are >100 subsets of stereotyped CDR3s, and it seems that no more than 30% of all CLL cases fall into one of these subsets[90]. Nonetheless, these findings have led to the suggestion that antigen is intimately involved in stimulating and/or selecting CLL cells, a corruption of the normal processes of B-cell ontogeny[49,114,117].

### *2.6.4 Antigens, Autoimmunity and CLL*

If an antigen is involved the selection of these B-cell clones, then the question arises as to the nature of the antigen[185]. Early findings suggested an association of IGHV stereotypy and autoantigen specificity[49,114,186]. The cells involved in autoimmunity and lymphoma are the same and the overlapping treatment options and similarities in biology have led many to try to marry these two fields[187]. Chronic antigen stimulation could plausibly favour malignant transformation of B-cells, akin to the relationship between *H.Pylori* and gastric MALT lymphoma. It has also been suggested that persistent autoantigen stimulation is involved in lymphomas that arise within salivary glands in Sjogren's syndrome[188], or thyroid tissue in Hashimoto's thyroiditis[189]. Pseudofollicles in CLL resemble similar structures observed in inflamed tissues in patients with the autoimmune diseases rheumatoid arthritis and multiple sclerosis[190].

However, there are difficulties, not least due to the heterogeneity of phenomena that are known as 'autoimmunity'.

Is autoimmunity a predisposing factor to the development of lymphoma? Retrospective epidemiological studies suggest that the development of lymphoid malignancies is more common in patients with autoimmune disease or with a family history of autoimmunity[191] [192-194]. There are a variety of associated autoimmune diseases, the most frequent being primary Sjogren's syndrome, rheumatoid arthritis and Systemic lupus erythematosus (SLE). There is also a higher risk associated with treatment with cytotoxic or biological agents, which may reflect disease activity, inflammation in general or the immunosuppressive effects of the agents themselves[191].

Are lymphomas associated with autoimmunity? Paraproteins in MM and the similar disorder Waldenstrom's Macroglobulinaemia may be autoreactive, particularly to nerve constituents, red cell antigens and clotting factors[190]. CLL and some other lymphoproliferative disorders are particularly associated with organ-specific autoimmune disease such as autoimmune haemolytic anaemia (AIHA), idiopathic thrombocytopenia (ITP) and autoimmune neutropenia[5]. However, in these cases, autoantibodies are produced by non-malignant 'bystander' B-cells, so we cannot draw conclusions about the autoreactivity of the BCR expressed by the malignant clone[195]. In the case of CLL, suppression and derangement of the normal B-cell populations leads to hypogammaglobulinaemia as well as autoantibody production. This may occur indirectly as an effect of CLL cells on T-cells[68]. It might be hypothesised that CLL produces a microenvironment conducive to the survival of autoreactive B-cells in general.

Cells from some CLL cases have the capacity to make autoantibodies *in vitro*, with the IgM on CLL cells often possessing rheumatoid factor (RF) activity, or binding activity for the Fc portion of human IgG[196]. Cultured CLL cells stimulated with phorbol 12-myristate 13-acetate (PMA) produce small amounts of monoclonal IgM that in one study reacted with a variety of self-antigens in 86% of cases[197]. The antigens include

IgG, ssDNA, dsDNA, histones, cardiolipin, and cytoskeletal determinants[198,199]. Collectively, early studies suggested that most CLL patients can produce polyreactive autoantibodies, albeit under somewhat artificial conditions, and most of these are IgM RFs. Similar IgM autoantibodies can also be detected in the sera of all individuals and so they have been termed 'natural autoantibodies'[200]. Natural antibodies have been implicated in the clearance of cellular debris, and there is some evidence for their role in clearing oxidised lipids and other harmful debris in protecting against atherosclerosis[201]. The obvious suggestion is that the IgM of CLL cells have natural antibody activity, however autoreactivity has been detected in 8 of 26 cases of follicular lymphoma[202], so this may not be a specific finding in CLL.

The early studies showing autoreactivity of CLL BCR were hampered by the small amounts of antibody secreted by CLL cells *in vitro*. By generating monoclonal antibodies (mAb) from the *IGHV* sequences of CLL cases using heterohybridoma and other techniques, several autoantigens have been better characterised. In one study recombinant antibodies from 28 U-CLL and 29 M-CLL cases were derived and compared to 31 antibodies derived from the $CD5^+$ cells of 3 healthy donors[167]. 20% of $CD5^+$ B-cells, 57% of M-CLL and 90% of U-CLL were reactive by indirect immunofluorescence assays (IFA) to panels of cell line antigens. There were common reactivity patterns within each CLL subset. Most antibodies recognised cytoplasmic structures, (often with very similar IFA patterns implying the same antigen) with a minority binding to nuclear components. In terms of polyreactivity against a panel of specific antigens, 3.2% of $CD5^+$ cells, 13.3% of M-CLL cells, and 79% of U-CLL cells expressed polyreactive mAbs. Polyreactivity was common in some U-CLL stereotyped subsets and not others. There was also high reactivity to oxidised Low Density Lipoprotein (ox-LDL) and oxidised Bovine Serum Albumin, neo-epitopes that may arise as a result of conjugation of metabolites of lipid peroxidation to membrane molecules. Oxidised LDL may have a role in the clearance of cell debris by natural IgM[201]. The mAbs with the broadest range of autoantigen binding were also those with the strongest binding to neoantigens. Apoptosis can make autoantigens accessible for recognition[203], and also create neo-epitopes by chemical modifications[203-205]. B-cells recognising these epitopes are often found in the B-1 cell compartment, the source of natural antibody.

Chiorazzi's group has performed further characterization of antigens, identifying non-muscle myosin heavy chain IIA (MYHIIA) as an autoantigen recognised by stereotyped subset 6 derived mAbs [206,207]. Subset 6 mAbs have a characteristic CDR3 sequence involving rearrangement of unmutated *IGHV1-69, IGHD3-16* and *IGHJ3* paired with a light chain with a characteristic sequence usually involving *IGKV3-20*. MYHIIA is a large cytoplasmic protein with functions in cell shape and movement. In order for the CLL BCR to bind MYHIIA, the authors suggested that it is exposed on the surface of apoptotic cells. Immunoglobulins from healthy volunteers also bound these apoptotic cells, and they suggest this reflects the presence of natural antibodies. U-CLL exhibited higher binding, as did patients with a poorer prognosis, with prognosis correlating more with binding than with mutation status. It is difficult to draw too many conclusions about prognostic impact from this small series of patients (n=24), but a different group has also correlated binding patterns to a panel of antigens and polyreactivity with prognosis in a group of 100 patients[208].

A Swedish group established 9 EBV-transformed CLL cell lines and harvested the mAbs secreted by them[209]. Autoantigen binding was exhibited for cytoskeletal and cytoplasmic antigens in various cell lines. The antibodies isolated bound several antigens, these were analysed by Matrix-assisted laser desorption/ionization (MALDI-TOF) mass spectrometric proteomics and protein tissue microarrays. Identified antigens were vimentin, filamin B, cofilin-1, PRAP-1, phosphorylcholine, cardiolipin, oxidised low-density lipoprotein, and *Streptococcus pneumoniae* polysaccharides. Vimentin is a cytoskeletal protein that is secreted as a stress response to bacterial insult and is exposed on apoptotic cells. Modified citrullinated vimentin (the Sa antigen) can generate autoantibodies in RA[210]. The survival-promoting effects of vimentin in an *in vitro* CLL culture system are reduced by the use of recombinant soluble anti-vimentin Ig to block antigen binding[114].

The finding of microbial antigens bound by CLL BCRs may also be relevant. BCRs from U-CLL utilising *IGHV1-69* seem to have a marked propensity to bind pUL32, a large Cytomegalovirus-encoded phosphoprotein[211]. There is also evidence of molecular

mimicry between *S.pneumoniae* polysaccharides and oxidised LDL[212] and between vimentin and phosphorylcholine (present in many bacterial cell membranes)[209]. These antigens are recognised by CLL mAb, and the demarcation between autoantigen and microbial antigen may not be strict. The risk of CLL is higher in individuals with a preceding infection (pneumonia or cellulitis), even when infections during the 5 year period prior to CLL diagnosis are excluded[213]. One interpretation is that infection may stimulate the growth of malignant B-cells, consistent with the antigenic stimulus hypothesis. An alternative explanation is that a degree of immunosuppression may well be apparent years before a diagnosis of CLL is made, increasing the risk of infection. Supporting this second hypothesis, recent data suggest that MBL itself is associated with a greater risk of hospitalization for infection[214]. A currently recruiting trial is assessing the efficacy of using antibiotics in the treatment of early stage CLL, based on the hypothesis of microbial stimulation of CLL progression[215].

Most recently, an intriguing study has raised a new possibility for the nature of the antigen for which the CLL BCR has specificity[216]. The authors of this study have previously characterised the binding specificities of recombinant CLL immunoglobulins[208]. They have also shown that the presence of constitutive signalling in developing B-cells is dependent on pre-BCR binding to invariant structural motifs within the pre-BCR itself[93]. As a synthesis of their previous work, this study suggests that in a limited number of characterised cases, the CLL BCR has specificity for the framework region of the BCR itself. Though they describe this as '…driven by antigen-independent cell-autonomous signalling', it might be better described as driven by antigen, but the antigen happens to be the BCR itself, on the same cell. This finding requires extension to a larger set of patients, but may provide much explanatory power. However, it makes certain other aspects of B-cell receptor signalling in CLL difficult to explain (discussed further in Chapter 8).

In summary, when the antigen-specificity of the BCRs of CLL patients are sought, a diverse range of putative antigens is identified, many of them autoantigens, These are often the products of cell apoptosis and oxidation or cytoskeletal components (such as vimentin, cofilin-1, MYHIIA and filamin B, often these are exposed on the surface of

apoptotic cells) [167,198,199,217]. Some individual cases of CLL (particularly U-CLL cases) appear to be polyreactive in that they bind multiple antigens, which may explain their more aggressive disease course[208] [167,198,199,217]. If the hypothesis that antigens stimulate the growth of CLL is correct, it is not inconceivable that each individual patient's leukaemia may be responsive to their own unique antigen(s). The existence of a limited number of subsets (albeit >100) implies that it is more likely that a narrow range of antigens are shared among patients[90,117]. Several questions remain unanswered. If antigenic selection is occurring, is this only at the early stages of MBL or CLL growth, or does it apply in more aggressive disease or even Richter's transformation? Generally, microenvironmental influences are felt to be less important for later stages of tumourigenesis, but there is no reason why they may not still have an effect. Does antigenic stimulation promote tumourigenesis by causing increased proliferation and therefore increasing the rate of mutagenesis? Do mutations have an effect by causing aberrantly enhanced responses to antigenic stimuli? It may be that MBL and CLL coincidentally arise from transformation of a B-cell subset that happens to exhibit a restricted repertoire due to the nature of antigens that the subset generally deals with, and that after transformation has occurred antigen ceases to be relevant to pathogenesis.

There is circumstantial evidence supporting ongoing antigen encounter in CLL. The phenotype [218]and gene expression profile[219] of CLL cells suggests similarity to memory B cells, as compared to naïve, tonsil GC and cord blood $CD5^+$ cells. Classical memory B cells, previously exposed to antigen, respond more rapidly to a second antigenic challenge, and have greater BCR avidity for antigens produced as part of the SHM process that occurs with T-cell help within the GC. Typically, there is also isotype switch to non-IgM/IgD sIg. This definition does not cover B cell responses to T-independent antigens, which can generate an enhanced number of cells recognizing a particular antigen without SHM or CSR. Immature B cells at earlier stages can also recognise antigen via their BCR. Therefore, 'antigen-experienced' does not equate to 'memory'. Whilst M-CLL could be analogous to classical memory B cells, U-CLL may well be derived from a B cell at an earlier stage, yet still have 'experienced' antigen binding without undergoing SHM.

There is some evidence of ongoing antigenic stimulation in CLL. One study analysed 71 CLL cases to search for intraclonal diversification (ID)[220]. Most cases did not display ID, but there was a subgroup of 13 IgG-switched cases in stereotyped subset 4 that displayed extensive ID. Others have found similar evidence of ID, and there are numerous reports of ongoing class-switch recombination (CSR) and AID expression in CLL subclones, indicating that the processes that operate in the germinal centre response to antigen are still extant in at least some cases of CLL[221,222]. AID and CSR generally seem to occur more often in U-CLL, and have been variably associated with prognosis. It may that ongoing BCR stimulation by antigen may promote mutagenesis by promoting proliferation, and Chiorazzi has suggested that AID may play an important role in genomic instability in CLL cases where it is highly expressed[223]. The role of AID in causing transforming mutations in other B-cell lymphomas is well recognised, the germinal centre being a dangerously mutagenic locale[100].

It might be expected that if CLL BCRs are autoreactive then autoimmune phenomena might be more apparent. Whilst there is a well-established association between autoimmune cytopenias and CLL, the autoreactive B-cells in these cases are not part of the CLL clone[224]. These are presumed to be dysregulated because of a general defect in eliminating autoreactive cells, perhaps because of T-cell defects[73,225], or an increase in B-cell survival factors such as BAFF, raised serum levels of which are known to associate with both CLL[226-228] and autoimmune disease[229,230]. Furthermore, it seems that 'healthy' B-cells also express autoreactive BCRs, albeit at a lower frequency than in CLL, and these may produce natural antibodies that are part of the body's mechanism for clearing cellular debris. There is an increase in autoantibody production with age, particularly for antibodies against cardiolipin, DNA, antinuclear antibodies (ANA) and rheumatoid factor (RF). However, this is not generally accompanied by a corresponding increase in autoimmune disease *per se*[231].

Autoreactive BCRs are therefore surprisingly common in health and disease, and autoreactivity alone is insufficient as a simple explanation for BCR abnormalities in CLL. To understand this complexity further, it may be helpful to consider how autoreactive B-cells are dealt with by the immune system.

## 2.7 B-cell Anergy

Given the mechanism of generation of antigen receptor diversity, a means for inducing tolerance for self-antigens at the same time as recognising a huge range of foreign antigens is necessary. At an early stage in their development, T and B-cells are subject to either clonal deletion or receptor editing in the bone marrow: 'central tolerance'[232]. As many as 75% of newly generated B-cells are autoreactive, and whilst many of these are deleted in the process of central tolerance, an estimated 40% of recent emigrants from the bone marrow are autoreactive and must be silenced to prevent development of autoimmune disease[233]. Additional autoreactive cell silencing mechanisms operate in the periphery, and include regulatory T-cells, idiotypic regulatory networks and anergy[234].

Anergic cells have a phenotype suggesting developmental arrest at the transitional stage, however this does not automatically imply that developmental arrest has occurred as mature cells also acquire this phenotype and become anergic upon encounter with antigen in the absence of second co-stimulatory signal[232]. Anergic B-cells persist in the periphery for 5 days and continue to express the BCR, albeit at reduced levels. However, BCR has lost its ability to transduce activating signals upon antigen binding. Anergy requires more than simply the absence of co-stimulatory signals, with maintenance of the unresponsive state requiring chronic occupancy of a significant proportion of receptors by antigen[232]. Chronic BCR binding would be characteristic of a ubiquitously expressed autoantigen, and in the absence of co-stimulatory signals that indicate the presence of a foreign invader, leads to attenuation of further signalling to produce a desirable state of anergy to the autoantigen[234].

The mechanisms of anergy have been elucidated in animal models. Goodnow et al expressed hen egg lysozyme (HEL) as a neo-autoantigen in mice[235,236]. Utilising mice in which transgenic HEL was coupled to a heavy metal-responsive promoter could induce the expression of HEL induced at various developmental stages. These ML5 mice were then crossed with the MD3 mouse line that are transgenic for expression of a BCR with known high affinity for HEL. The presence of HEL did not lead to deletion in these mice; in fact HEL-specific B-cells were present in the periphery in almost normal numbers at

immature stages, but these were unable to mount effective responses to the HEL antigen. B-cells that became anergic in the presence of soluble HEL were deleted when the MD3 mouse was crossed to a mouse expressing HEL in a membrane-bound form[237], highlighting the importance of the context in which antigen is encountered. Occupancy of <5% of receptors led to development of B-cells that were 'ignorant' of antigen, 5-45% occupancy to anergy, whereas higher occupancy led to deletion[236].

Constant BCR occupancy is required to maintain anergy, and anergy is reversible. The transfer of anergic B-cells into an environment lacking HEL results in loss of the anergic phenotype and the reacquisition of the ability to produce antibody responses to HEL[238]. The anergic state is associated with reduced expression of IgM, as might be expected if the cell is seeking to reduce responses to antigen via reduction in BCR expression. Unexpectedly, IgD is maintained at levels similar to naïve cells, a finding that has been widely replicated yet not explained. These transgenic models have been criticised for their unphysiologic nature[239], and so evidence of anergic cells have been sought in other models. A population of $IgM^{low}IgD^{high}$ (An1) B-cells with other markers of anergy exists in the normal B-cell repertoire of wild-type mice that have been suggested to be naturally occurring anergic cells[236,240].

Given that so many newly formed B-cells are autoreactive[233,239], anergy is probably a major mechanism of tolerance induction. This may explain the presence of a surprisingly high frequency of autoreactive BCRs in the repertoire in the absence of overt autoimmune disease. A logical question is whether these anergic cells form a source of autoreactive B-cells that may subsequently cause autoimmune disease. There is some evidence that this may be the case as the An1 population is reduced in autoimmune disease mouse models[234]. Whilst murine models have been increasingly well characterised, much less is known about the human counterparts. The human equivalent of these anergic cells appears to be $IgD^{+}IgM^{lo/-}CD27^{-}$, have unmutated *IGHV* genes and constitute 2.5% of peripheral blood B-cells[241]. They are unresponsive to BCR stimulation as measured by calcium flux and protein tyrosine phosphorylation. This population is also enriched for B-cells that recognize ssDNA and Hep-2 cell antigens, well-known autoantigenic culprits in autoimmune disease.

Various features of An1 cells reduce their participation in immune responses. One is reduced lifespan. The half-life of follicular B-cells is 40 days, as opposed to 5 days for anergic B-cells[242,243]. The reduced lifespan is only apparent when non-anergic B-cells are also present to compete for B-cell activating factor (BAFF), an important survival factor for peripheral B-cells[244]. Anergic cells appear to have reduced BAFF receptor expression or reduced BAFF receptor signalling competence[230]. As might be predicted, over-expression of BAFF leads to autoimmunity[229]. BAFF and its related molecule APRIL may play an important role in CLL cell survival[227,245], yet the picture is confusing as BAFF and APRIL levels may be raised in CLL, yet the BAFF-Receptor levels are reduced and reports are conflicting. Other mechanisms of decreased survival involves high expression of the proapoptotic BCL family member BIM in anergic B-cells, and anergy is lost in anti-HEL mouse B-cells lacking BIM[246]. Anergic B-cells may also undergo apoptosis initiated by FAS signalling by encounter with $FASL^{+}CD4^{+}$ T-cells[247]. In contrast, it is well known that CLL cells exhibit a defect in apoptosis, with higher expression of anti-apoptotic BCL family members[248]. A model of CLL as an equivalent of anergic B-cells has been suggested[115,116,249]. It would be consistent with the concept of CLL cells as chronically binding autoantigens *in vivo*. One curious feature of both anergic and CLL cells is that they generally exhibit low expression of IgM whilst preserving the expression of IgD. Why this occurs is unknown, but if the BCR in CLL is to be studied, then it must be considered in both its expressed isotypes: IgM and IgD.

## 2.8 IgD

### *2.8.1 Introduction*

When antigen binds the V region of the BCR, the signal is propagated to the interior of the cell via the constant (C) region of the immunoglobulin molecule[122,250]. The structure of the C region determines the isotype of the BCR. The properties of this signal (and the effector properties of the antibody synthesized by the same B-cell) are dependent on the isotype. Each B-cell clone encodes one V region with specificity for a single antigenic epitope (or several similar epitopes in polyreactivity). One of the means by which a B-cell can modify its response to the same antigen is to switch isotype, with this switch resulting in altered signalling or altered effector function of

secreted antibody[101]. A curious feature of B-cells in the early stage of maturation is that they have dual expression of both IgM and IgD, and this is true of most cases of CLL[251-253]. Membrane IgD shares many properties with its more famous cousin IgM. Why both normal and malignant B-cells express two different isotypes with identical antigen-binding regions is still a mystery, with phraseology of review articles on IgD function often employing the words 'enigma' and 'riddle'[254,255].

Evolutionarily, the immunoglobulins first appeared in jawed vertebrates around 500 million years ago[256]. IgD is only expressed in teleosts (bony fish) and mammals, though there is a homologue (IgW) in cartilaginous fish and more homologs are being discovered as genomes of various species are sequenced. Its widespread and long-standing usage suggests some evolutionary advantage. Whilst IgM remains stable over evolutionary time, IgD has shown greater structural plasticity and may be expressed predominantly as a transmembrane or secreted molecule depending on the species studied. This has led to suggestions that IgD has been preserved as a structurally flexible locus to complement the functions of IgM[255]. The $C_H$ gene encoding the human $\delta$ chain is located in the IgH cluster at 14q32 on chromosome 14. It may be a result of duplication of the c$\mu$ gene during evolution[257]. It spans 10kb and has eight exons.

IgD was initially discovered in 1964 after investigation of a novel paraprotein in multiple myeloma sera[258]. The reasons for naming the new immunoglobulin 'D' included its '*d*ifference' from the other immunoglobulins, and the fact that IgA, IgB (a name for murine immunoglobulins, now fallen out of use) IgE and IgG were already used. The letter C has no Greek equivalent, and so IgC was never employed[86]. Subsequently, a large number of studies identified it as a major immunoglobulin that constitutes the BCR in both mice and humans[259].

IgD represents 0.25% of the total serum immunoglobulin and has a wide range of concentrations between individuals irrespective of age, but dependent on infection history[255]. Antigenic challenge can produce specific IgD, and IgD with autoantigen specificity has been identified[260]. In addition to blood, IgD is also present in various secretions. IgD does not bind to neutrophils or monocytes, does not cross the placenta, and does not activate complement strongly. Both $CD4^+$ and $CD8^+$ T-cells have

putative IgD receptors that can ligate both aggregated serum secreted IgD (secIgD) and surface IgD during antigen presentation by B-cells[259]. Stimulation of basophils and other cells with secIgD leads to the release of many cytokines that may favour B-cell survival and activation, including BAFF[261]. It therefore seems to have a role in promoting inflammation and links the adaptive and innate immune systems.

### *2.8.2 Structure of IgD molecule*

The heavy chains of IgD on the surface of B-cells are covalently linked by only one disulphide bridge, which allows a higher degree of mobility for the antigen-binding sites compared to other immunoglobulin isotypes[262]. The high carbohydrate content (9-14%) of $\delta$ chains explains why its molecular mass (175kDa) is higher than that of $\gamma$ chains. The structure of C$\delta$1 and C$\delta$2 domains is similar to other isotypes, whilst the C$\delta$3 domain is different with the lack of several proline residues and by the presence of two N-linked carbohydrates. Evidence suggests that these structural features explain many of the differing biological effects of IgD compared to other isotypes. Despite having the same V gene regions, it has been shown that the binding properties of IgM and IgD may be different because of the structural features of the receptor, with IgM binding having a higher functional affinity at higher antigenic concentrations[263]. The co-expression of IgM and IgD may therefore facilitate the recognition of the same antigen at different concentrations, or in different structural arrangements.

IgD appears to preferentially associate with $\kappa$ light chains, while the secreted form associates with $\lambda$ chains [262]. It is not obvious why this is, though various hypotheses have been suggested[250,251]. The intracytoplasmic tail of IgD is particularly short, is identical to that of IgM in humans and contains only three residues (Lys-Val-Lys).

Surface IgD is anchored in the membrane via its transmembrane hydrophobic domain, and is non-covalently linked to heterodimers of Ig$\alpha$ (CD79a) and Ig$\beta$ (CD79b). Due to different carbohydrate contents, IgM- and IgD-associated Ig$\alpha$ and Ig$\beta$ chains have different molecular masses[264]. As with the other sIg, Ig$\alpha$ and Ig$\beta$ play an essential role in IgD BCR signal transduction. Whilst IgM produced in the absence of Ig$\alpha$-Ig$\beta$ is retained in the endoplasmic reticulum through its interaction with chaperone proteins

such as calnexin, IgD may be transported to the membrane in the absence of Ig$\alpha$-Ig$\beta$ and calnexin is not required in this process[265].

IgM and IgD may be associated with different molecules[264-266]. Two accessory proteins, the BCR-associated proteins (BAP) 29 and BAP 31 are also part of the IgD BCR complex, whereas BAP 32, 37 and 41 are associated with IgM[267]. These may be involved in functional differences, however the nature of these molecules has not been fully investigated since these early reports.

In some contexts, an alternative and minor form of IgD is linked to the plasma membrane by a glycosyl-phosphatidylinositol (GPI) anchor[268]. It is sensitive to phospholipase C and hence may be released in soluble form. The GPI-linked isoform of IgD normally forms a minority of IgD, but selectively activates cAMP-dependent pathways, which may support calcium flux via canonical BCR signaling[269].

### *2.8.3 Regulation of IgD synthesis*

The methods of regulation of IgD synthesis vary according to developmental stage of the B-cell. Two mechanisms result in IgD expression: Alternative splicing of a common IgM and IgD mRNA transcript and Class Switch Recombination (CSR).

Differential splicing of long primary mRNA transcripts is the main mechanism of simultaneous expression of IgM and IgD. It is not certain why immature B-cells only express IgM. In the mouse, central immature B-cells contain no $\delta$ mRNA, whereas peripheral immature B-cells express amounts of $\delta$ mRNA similar to that found in mature B-cells. How this differential splicing occurs is uncertain. There is a specific region between the $\mu$ and $\delta$ genes, which attenuates the transcription complex and forces IgM mRNA production by default. Binding of lineage specific regulatory proteins may alter the ratio of processing of the primary mRNA transcripts[270,271]. The 200 bp att attenuator site is located 1650 bp downstream to the $\mu$ polyadenylation site in the $\mu$-$\delta$ intron and contains cis-acting repressors of $\delta$ gene transcription. Deletion of att leads to long $\mu$-$\delta$ transcripts, but these are processed into mature $\delta$ transcripts only in

mature B-lymphocytes[270]. The μ-δ intron also contains a binding site for the myc-associated zinc finger protein MAZ that terminates δ transcription[272].

Post-transcriptional controls involve translational and post-translational modifications. The density of IgD on mature B-cells is often much higher than that of IgM in spite of lower levels of the δ mRNA compared to μ mRNA[273]. This is partly explained by a more rapid turnover of IgM than IgD and lower stability of μ RNA compared to that of δ mRNA. The loss of IgD expression after mature B-cell activation is accompanied by a sharp drop of δ mRNA content and an increase in μ chain half-life[274]. There is usually no detectable IgD expression by the memory B-cell stage[275].

IgD expression can also occur via Class Switch Recombination (CSR). The absence of a canonical switch region 5' to the δ gene would seem to rule out the classical class switch from μ to δ analogous to the switch from μ to other isotypes. However, deletional recombination of cμ has been demonstrated in murine myeloma, human Hairy Cell Leukaemia B-cells and normal tonsillar B-cells with a region containing G-rich pentameric repeats located within the μ-δ intron able to recombine with the Sμ region and act as a switch region[276].[276] B-cells from tonsillar germinal centres have the highest frequency of IgD expression in health and are enriched in cells that exhibit IgD class switching and a have a high level of V gene somatic mutations[277]. These IgD-switched B-cells predominantly use λ light chains and can differentiate into plasma cells. Since most secreted IgD is λ light chain restricted, it is likely that these switched cells are the source of most secreted IgD. Also, these switched cells use a subset of V genes associated with autoreactivity and have evidence of receptor editing[278].

One study found that >50% of human IgD switched lymphocytes were autoreactive, as they encoded antibodies that bound antigens such as antinuclear antigens (ANAs), single- and double-stranded DNA and were often polyreactive[279]. It may be that the cryptic Cδ switch region is targeted in autoreactive B-cells, or that these cells are inhibited from switching to other isotypes. Alternatively, selection of IgD switched cells may occur after class switch, with survival of autoreactive clones, suggesting a role for survival signals via IgD. High levels of SHM are also exhibited by IgD switched B-cells

compared to $IgG^+$ and $IgM^+$ cells. This SHM may indicate either antigenic selection or an attempt to reduce the polyreactivity of autoreactive BCRs[259].

The superantigenic binding of certain bacterial proteins to the carbohydrate moieties of constant domains is peculiar to IgD. The Fc fragment of IgD binds nonspecifically to bacteria, including *Moraxella catarrhalis* and *Haemophilus influenzae*, and streptococcal groups A, C and G[280,281]. *Moraxella catarrhalis* binds to IgD and has a strong mitogenic effect. The specific *Moraxella* IgD-binding protein (MID) has been isolated, and it is stably expressed by all strains of *M. catarrhalis* tested[282]. It interacts with C$\delta$1 region of the heavy chain and acts as a superantigen. *M Catarrhalis* can be found in the mantle zone of human tonsils, and it is suggested that this accounts for the high prevalence of $IgD^+$ cells in the nasopharynx.

Serum IgD is often increased in patients with autoimmune diseases, such as rheumatoid arthritis and systemic lupus erythematosus (SLE)[255], consistent with the finding that many IgD class-switched cells are autoreactive in healthy individuals[279]. As further implication of their role in autoimmunity, selective depletion of $IgD^+$ B-cells in a mouse model of autoimmune arthritis leads to amelioration of the condition[283].

### *2.8.4 IgD in B-cell development*

IgD, coexpressed with IgM, is the major antigen receptor isotype on naive peripheral B-cells[259]. In general, IgD expression is confined to this stage of differentiation, whilst IgM has a somewhat greater range. During development, IgM is the initial BCR isotype synthesized, and is later expressed during the GC reaction, whilst IgD is downregulated. IgD is expressed once a B-cell has emigrated from the marrow to populate spleen, lymph nodes and intestinal mucosa. When B-cells leave the bone marrow they are termed transitional type 1 (T1) cells, and are particularly susceptible to negative selection (deletion on receptor ligation). It has been suggested that IgD rescues cells from IgM-induced apoptosis during tolerance induction and possibly the negative selection of immature B-cells results from low IgD expression[284]. At the end of the T1 stage they start to express IgD as well as IgM and become more resistant to negative selection. Though it is sometimes assumed that most autoreactive clones are

eliminated by this stage, low-affinity self-reactive clones can still be readily detected and IgD expression may promote their survival.

The T2 stage is less susceptible to negative selection and is associated with high levels of IgM and IgD expression and protection by anti-apoptotic factors, as well as different coupling to signalling pathway components (such as the involvement of PKCβ and Btk)[285]. The levels of BCR expression can guide B-cell development into different subsets[286]. Other signalling pathways such as TLR4, TLR9, CD40, BAFFR, BCMA and TACI also have an effect on the survival of T2 cells.

Geisberger et al suggest that IgD, by its expression at the T2 stage of development, supports the survival of B-cells at this stage, and that there must be a difference in signalling between IgM and IgD because the combined signalling at this stage would lead to a very large signal and elimination of the B-cell clone[254]. Perhaps by modulating the ratio of IgD and IgM the B-cell can regulate its own survival by modulating the total signal received by both isotypes and therefore altering the thresholds of negative and positive selection. Upon antigen stimulation during the germinal centre reaction, mature B-cells rapidly lose IgD before undergoing terminal maturation toward IgM plasma cell or class switching[287].

### *2.8.5 Signalling differences between IgD and IgM*

Evolutionary conservation, differential expression at certain stages of B-cell development and co-expression with IgM at others implies that the IgD and IgM isotypes of the BCR differ functionally. In general the similarities are pronounced, with the δ heavy chain compensating for the μ heavy chain in early B-cell development, and vice versa[288-290]. Dual knockout of the μ and δ genes leads to arrest at the pre-B cell stage, as no functional BCR can be produced[291]. δ knockout mice exhibit a slight reduction in peripheral mature B-cells[289], and delayed affinity maturation during T-dependent antigen responses[288]. Mice with μ gene deletion exhibit some defects in CSR, with slightly reduced B-cell numbers[291,295]. This suggests that IgM and IgD can largely substitute for each other during development, but that optimal B-cell maturation requires both IgM and IgD.

As discussed, downregulation of IgM compared to IgD is a feature of autoimmune mice models and anergic B-cells[236]. In immunoglobulin-transgenic mice co-expressing hen egg lysozyme (HEL)-specific IgM and IgD, the response to constitutive exposure to HEL (simulating autoantigen) results in downregulation of IgM with relative preservation of IgD[235,292]. The decrease of IgM is explained by persistent engagement of BCR by antigen resulting in downregulation of the BCR to prevent overstimulation of the B-cell and subsequent possible autoreactivity, but this does not explain why IgD is not decreased.

IgD and IgM selective knockouts were explored further by crossing with mice that expressed B-cells specific for hen-egg lysozyme (HEL)[292]. Both IgM and IgD-only cells could mount comparable T-independent and T-dependent responses to acute exposure to HEL antigen, with IgD only mice actually generating anti-HEL antibody with greater efficiency. Constitutive expression of HEL in a membrane-bound form led to deletion of HEL-specific cells from the repertoire, as might be expected via central tolerance. Similar to previous models, constitutive soluble HEL allowed maturation up to the transitional stage, but both IgD- and IgM-only B-cells were rendered anergic. As with other anergic models, IgM was downregulated 10-100-fold, whilst IgD was only downregulated by 5-fold in IgD-only cells. Despite both isotypes able to mediate the functional process of anergy, there is still a perplexing difference in the level of receptor downregulation during this process, even when the other isotype is absent. IgM appears to have a greater plasticity of expression level as compared to IgD.

Early explorations of signal transduction found no differences between IgM and IgD signalling[293], yet subsequent studies have found subtle differences. Binding of antigen by IgD can induce phosphorylation of multiple substrates that is earlier, stronger and more prolonged than after IgM binding in a study on a myeloma cell line[294]. In both a lymphoma-derived cell line and healthy human peripheral blood B cells, only anti-IgM induced growth arrest and cell death, in spite of similar tyrosine phosphorylation, calcium flux, activation of PKC and expression of transcription factors such as c-fos and Egr-1[295].

There is increasing evidence that mobility in membranes and interactions with the cytoskeleton have a critical role in regulating BCR signaling[296]. Single molecule imaging techniques suggest that IgD diffuses more slowly than IgM[297], and that the movement and clustering of BCR complexes is dependent on ezrin and the actin cytoskeleton. Similar techniques also show that IgD exists in pre-clustered complexes compared to the more diffuse, unclustered IgM[298].

Interactions between IgM and IgD signalling may occur. It can be shown that engagement of one receptor isotype results in unresponsiveness of the reciprocal isotype, probably by uncoupling of Ig from Ig$\alpha$-$\beta$ which lasts for more than 24 hours[299]. Other early results suggested that ligation of IgM but not IgD can induce negative responses[300-303] (growth arrest, anergy, apoptosis), which is consistent with a model whereby antigen exposure at an early (IgM only) stage of development leads to deletion or apoptosis, whereas mature cells (expressing IgM and IgD) generally undergo proliferation after encountering antigen. The teleological rationale for this would be that antigen encountered at an early stage is likely to be autoantigen, whereas mature B-cells would be expected to more often encounter foreign pathogens that require a robust response. IgD signalling in this model counteracts 'negative' signals delivered via IgM. In other models both IgM and IgD can transmit negative (apoptosis-inducing) signals in immature cells in irradiated mice, whilst IgM and IgD engagement in mature cells generally results in proliferation[293,294].

The differing role of signalling via IgD or IgM is not clear. Discrepancies may be explained by the differing developmental stages of the B-cells used as models, as well as differing receptor density. Similarities between IgM and IgD signaling are often more pronounced than their differences. The reason why B-cells have dual expression of IgM and IgD at various stages and how BCR signalling via these two isotypes differs still remains somewhat of a mystery. Nonetheless, the conservation of these isotypes throughout evolution and the sometimes-divergent responses to stimulation do suggest that they have differing roles in B-cell development and physiology.

### *2.8.6 IgD in CLL*

It is an established finding that CLL expresses surface immunoglobulin at reduced density as compared with healthy B-cells and other B-cell neoplasms[304]. This 'weak' or 'dim' expression as detected by flow cytometry is a peculiarity of CLL that has been incorporated into diagnostic criteria, but never satisfactorily explained[5,49]. The other established finding is that CLL generally co-expresses IgD and IgM. The exact proportions and combinations of IgD and IgM vary by method of detection and publication, and are discussed further in Chapter 5.

Few researchers have examined the effect of IgD signalling in CLL. Lanham et al's main aim was to measure the outcome of IgM signalling, as judged by global protein phosphorylation and Syk phosphorylation[305]. Of 40 CLL samples, 15 (38%) did not respond to IgM cross-linking. All samples expressed IgM, and IgD. Of the 15 samples tested that did not respond to anti-IgM, 10 (66%) were responsive to anti-IgD and 12 (80%) were responsive to anti-CD79a. The 3 samples that did not respond to anti-CD79 were nonresponsive to either anti-IgM or anti-IgD. The study was used as an illustration of dysfunctional IgM signalling, with preservation of IgD signalling in CLL[305]. The same group have published similar data demonstrating preserved IgD signalling in the presence of decreased IgM signalling in another study[116]. Intriguingly, this second study suggested that IgM expression and signalling capacity could be restored by incubation of cells *in vitro*, analogous to the scenario with anergic B-cells. This is discussed further in Chapter 6.

Zupo et al[306] examined 10 CD38 positive CLL cases, as it had been suggested that these cases have a better response to BCR ligation[307]. They suggest that IgM crosslinking results in CLL apoptosis, as have many other researchers, though published data are often conflicting on this outcome. They also suggest that IgD ligation results in less apoptosis as compared to IgM ligation, and in some conditions where cocktails of cytokines are employed, IgD ligation can result in differentiation to plasma cells. These studies do suggest that IgD and IgM have differing roles in CLL, and are discussed further in Chapter 6.

### *2.8.7 IgD Summary*

Membrane IgD is a somewhat mysterious molecule, sharing many properties with its more famous cousin IgM. It has not been established why CLL and naïve B-cells express IgD, often in combination with IgM. It is known that BCR stimulation via IgD has generally more effect on measures of BCR signalling in CLL. The outcomes of IgM stimulation in CLL are notoriously heterogeneous, as measured by the effect on cell survival (discussed further in Chapter 6). What is known is that IgD is more reliably expressed in CLL, despite the fact that most investigators have studied the effects of BCR stimulation via IgM, probably due to an assumption that IgM is the canonical surface immunoglobulin isotype in normal B-cells.

An analogy between CLL cells and anergic autoreactive B-cells has been suggested (discussed further in Chapter 6), as reflected by a downregulation of IgM and preservation of IgD expression, a pattern which has yet to be explained in either healthy or malignant B-cells. Whatever the underlying causes of the differences between IgD and IgM expression and signalling in healthy and malignant B-cells, any consideration of the BCR in CLL is obliged to consider both IgM and IgD. If we accept the hypothesis that antigen binding to the BCR of CLL cells may exert a selective or growth-promoting effect on CLL then we must expect that this antigen binds both to IgD and IgM since both isotypes will have identical V-regions with the same epitope specificity. In fact, given the generally greater expression of IgD we might expect that this interaction was more important than that of antigen-induced BCR signalling via IgM. Therefore, to study BCR signalling in CLL, we cannot avoid consideration of the signals produced by both IgD and IgM ligation. It is possible that such studies will also illuminate some areas of IgD function in normal physiology.

## 2.9 B-Cell Receptor Signalling

Hitherto, we have discussed the role of the BCR in CLL without considering the events occurring after the B-cell contacts its cognate antigen. The mechanisms of BCR signalling have been extensively studied in a variety of models, and different publications emphasise different aspects, depending on the scientific paradigms of the authors. Below is an attempt to summarise some of what is known about general aspects of BCR signalling, as well as an attempt to discuss BCR signalling in CLL, and in anergy.

B-cells can respond to multivalent antigens in solution by BCR clustering and capping, with potential endpoints of proliferation and differentiation if appropriate support is received. Whilst much information about intracellular cascades has been elucidated, it is only recently that techniques in live cell imaging have enabled investigators to begin to establish how antigen binding to BCR ectodomains is translated across the membrane to trigger signalling cascades, and the importance of B-cell recognition of antigen on the surface of antigen-presenting cells[246,317]. Even so the underlying mechanisms of signal transmission by the BCR across the plasma membrane are still somewhat of a mystery

The BCR comprises membrane-bound immunoglobulin (either IgM, IgD, sIgG, sIgA or sIgE) and a disulphide-linked heterodimer composed of Ig$\alpha$ (CD79a) and Ig$\beta$ (CD79a). These are transmembrane proteins with intracellular domains that each contain an immunoreceptor tyrosine-based activation motif (ITAM). On antigen binding, the BCR is phosphorylated on its ITAM tyrosines initially by LYN kinase, then by SYK kinase recruited to phosphorylated ITAMs via its SH2 domain. These events result in the initiation of several downstream signalling cascades.[250]

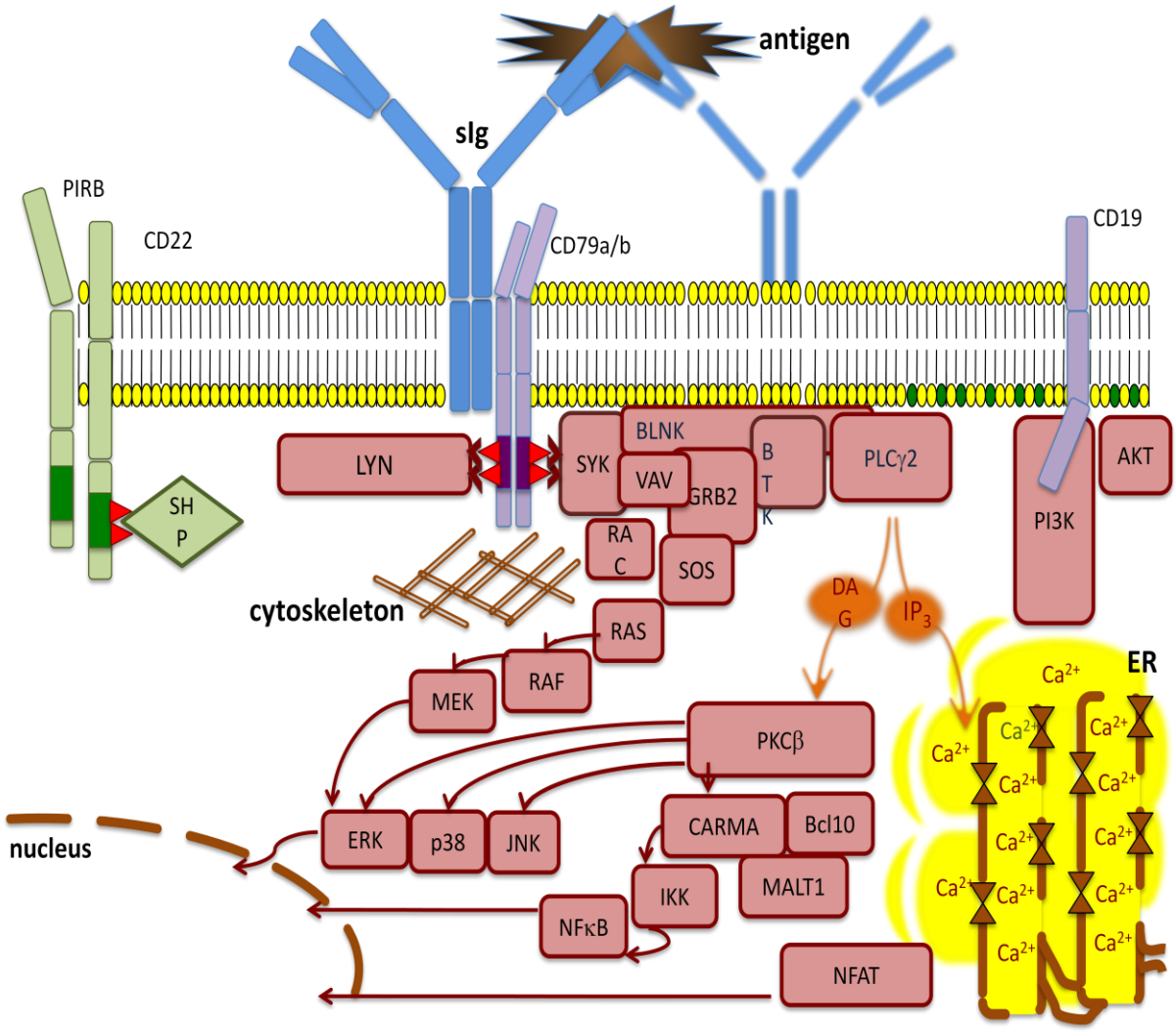
In the absence of antigen, B-cells are often described as ‘resting’. This may be a misnomer, as there is evidence that the cells are hovering near a tipping point of activation[308]. The expression of the BCR is essential for all stages of B-cell development, leading to the hypothesis that the BCR provides low-level ‘tonic’ signals

that promote survival, either spontaneously or on interaction with ubiquitous microenvironmental ligands[120]. The phosphoinositide-3-kinase (PI3K) signalling pathway is crucial for the survival of mature B-cells and this pathway may be connected to the BCR in resting cells by the GTPase TC21 (also known as RRAS2), which connects the catalytic subunit of PI3K to the unphosphorylated ITAM of the BCR[121,309]. Generally, however, the nature of the pathways underlying tonic BCR signalling is poorly understood as yet.

Phosphorylation of ITAMs by LYN initiates the formation of the signalosome, an assembly of intracellular signalling molecules such as the kinase SYK, phospholipase C$\gamma$2, PI3K, Bruton's tyrosine kinase (BTK), B-cell linker (BLNK) and VAV[310]. This signalosome allows the coordinated regulation of downstream cellular events such as calcium flux, antigen internalization and changes in gene expression (See Figure 2.3). Because the recognition of physically constrained antigen involves alteration of B-cell morphology, the cytoskeleton is seen to play a more important role than previously appreciated[311,312].

Cambier suggests that the BCR could be considered as a switch in three states[313]. Unoccupied receptors would have unphosphorylated ITAMs, acute occupancy of the BCR by antigen leads to biphosphorylation of ITAMs, activation of SYK and downstream pathways that lead to immune activation of the B-cell. Only once the two SH2 domains of SYK are engaged by dual ITAM phosphorylation is SYK activated by tyrosine phosphorylation[314]. Chronic receptor occupancy, as occurs in anergy, leads to predominantly ITAM monophosphorylation and therefore predominantly engagement of the LYN tyrosine kinase. As well as having a role in positive regulation of BCR signalling via activation of SYK, anergic B-cells display increased Ig$\alpha/\beta$ ITAM phosphorylation, but only monophosphorylation[234,313]. This may allow the inhibitory functions of LYN signalling to dominate[315]. Consistent with this, the LYN substrates SHIP-1 and DOK are constitutively phosphorylated in anergic B-cells[234], and these are involved in the negative regulation of downstream pathways. It is likely that other LYN targets such as CD22 and the phosphatases SHP-1 and SHP-2 also participate in the maintenance of anergy[234].

**Figure 2.3: B-cell Receptor Signalling pathways**
**Binding of antigen to cognate B-cell receptor leads to conformational changes that cause phosphorylation of cytoplasmic ITAM residues of CD79a/b by LYN kinase. Subsequent activation of SYK, other downstream kinases and linker proteins amplify these initial events, leading to the formation of a multimolecular signalosome complex, in association with clustering of multiple BCRs. This leads to events such as release of $Ca^{2+}$ from the endoplasmic reticulum (ER), cytoskeletal changes, and gene expression changes. Negative regulation of signalling is provided by phosphatases such as SHP that associate with co-receptors such as CD22 and PIRB. Amplification of signalling can be provided by co-receptors such as CD19. Pathways derived from multiple papers.[116,117,119,246,316,317,320,325-328] Details on individual components are in the text.**




## 2.10 Abnormalities of B-cell Receptor Signalling in CLL

The downstream components of BCR signalling have been studied to an extent in CLL, and we briefly consider the components of normal BCR signalling and abnormalities

seen in CLL and other lymphoid malignancies. In general, crosslinking of the BCR by monoclonal antibodies has been used to simulate antigenic encounter, a technique that has also been used by studies of normal BCR signalling, though in some other contexts it has been possible to use specific antigen in studies of non-malignant B-cells, classically in the HEL mouse models[316]. Partly inspired by the model of abnormal BCR signalling in CLL, several novel small molecule inhibitors of some of the BCR pathway components have been developed, in particular targeting BTK, PI3K, SYK and LYN. These are discussed further below, and the remarkable clinical efficacy displayed by some of these compounds is perhaps one of the strongest arguments for the role of abnormal BCR signalling in CLL.

### 2.10.1 Mutation status

In terms of CLL mutation status, most studies have suggested that BCR signalling is more potent in U-CLL, whether measured by calcium flux, apoptosis, proliferation or the phosphorylation of various downstream molecules[305,317-322]. However, because of the association between mutation status and ZAP70, CD38 or sIg expression, the increase in signalling capacity may be more associated with these other confounding factors[116,318].

### 2.10.2 CD79a/CD79b

One reason for poor BCR signalling in CLL may be low levels of CD79a/b, with either resultant low sIg expression, or poorer signalling[323]. Dissociation of the CD79a/b subunits from the sIg may be a mechanism whereby BCR signalling is decreased in anergic B cells[299]. Splice variants of CD79b have been noted to be used preferentially in CLL, and some have noted constitutive CD79a ITAM phosphorylation[317,324]. Recently, CD79a/b has been shown to play a critical role in diffuse large B cell lymphoma cell survival, via chronic activation of BCR downstream pathways such as NF-kB, with activating mutations in CD79b ITAMs detected in a high proportion of primary lymphoma cells of this subtype[325].

### 2.10.3 Calcium flux

Initial BCR signalling events typically activate phospholipase C (PLC), which results in abrupt rises in cytoplasmic calcium concentration ($[Ca^{2+}]i$ flux). Calcium flux is often

taken to be a surrogate marker or *sine qua non* of BCR signalling. The abrupt rise in intracellular $Ca^{2+}$ concentration ($[Ca^{2+}]i$) modifies a diverse range of intracellular processes[326]. Early reports suggested that 50-75% of CLL cells are able to display $[Ca^{2+}]i$ flux in response to crosslinking IgM[327,328]. The patterns of calcium flux response to IgM ligation have been likened to those seen in anergic B-cells[329].

Anergic B-cells exhibit elevated baseline intracellular $[Ca^{2}]i$ and BCR ligation fails to evoke large transient increase in $[Ca^{2+}]i$ [234]. In CLL, PLC phosphorylation after BCR ligation is heterogeneous, with greater phosphorylation in U-CLL.[312,332,342]

Calcium flux in CLL and anergic B-cells is discussed further in Chapter 6.

### 2.10.4 LYN tyrosine kinase

Several studies have consistently shown elevated LYN expression and constitutive activation in CLL cells[305,317,330]. Dasatinib is an src and abl kinase inhibitor that inhibits LYN and induces CLL cell apoptosis. It has clinically significant effects in CLL[331,332].

### 2.10.5 Spleen tyrosine kinase (SYK)

SYK mutations have been found in diverse cancer cells[333]. Sequencing of 40 CLL patients did not reveal any mutations or unusual frequencies of SNPs in SYK **[334].** CLL cells exhibit constitutive Tyr525/526 autophosphorylation and an increase on BCR ligation, which is reduced by a SYK inhibitor[335,336]. CLL cases with proliferative responses to BCR ligation have higher levels of SYK than anergic cells[329], whilst others found that SYK expression was homogeneous in CLL**[337]**. BCR and SYK activation also modulate cell adhesion and chemotaxis of B cells[338-340], and SYK inhibition can reduce the protective effect of the microenvironment in CLL, suggesting that SYK may play a role other than downstream of the BCR[341]. Fostamatinib is a prodrug of R406 under clinical development for treatment of rheumatoid arthritis and immune thrombocytopenic purpura. A phase I/II trial showed reasonable safety and response rates, with a combined PR and CR rate of 55% in pre-treated CLL/SLL patients [342], however further trials in lymphoma are not currently planned. Patients treated with Fostamatinib exhibited an initial lymphocytosis that has been since recognised as a BCR antagonist class effect. It is likely that it results from disruption of the CXCR4-

CXCL12 axis and other adhesion factors (perhaps including BCR) in the bone marrow and lymph node, leading to mobilization of B-cells into the peripheral blood.

### 2.10.6 ZAP-70

In a study comparing the gene expression of U-CLL with M-CLL, it was found that only a small number of differentially expressed genes separate the two groups,[343] the most specific being a gene encoding a 70 kDa zeta-associated protein (ZAP-70)[344]. Most mutated cases are ZAP-70 negative and unmutated cases ZAP70 positive. B cells generally lack ZAP-70, using the homologous kinase SYK for BCR signal transduction. SYK and ZAP70 have partially overlapping roles in early lymphocyte development[345]. ZAP-70 expression appears to be higher in sites of tissue proliferation of CLL[54]. ZAP-70 expression is associated with increased capacity for BCR signalling in CLL, as measured by global and SYK phosphotyrosine levels after IgM crosslinking [346]. Given that CLL expresses SYK as well as ZAP-70, and that SYK has an approximately 100-fold greater intrinsic PTK activity than ZAP-70[347], it is unclear what role ZAP-70 plays in BCR stimulation. It may be that it simply associates with other factors such as mutation status that enhance BCR signalling, though some have suggested that BCR signalling capacity is more associated with ZAP-70 than mutation status[334]. ZAP-70 negative cells transduced with an adenoviral vector encoding ZAP-70 increased BCR signalling capacity[334]. However, it has been subsequently shown that the enhancement of BCR signalling was independent of ZAP-70's kinase activity[348], suggesting that ZAP-70 may act as a scaffolding protein that perhaps enhances BCR translocation into lipid rafts, or reduces BCR internalization.

### 2.10.7 Bruton's tyrosine kinase (BTK)

The initial step in activation of BTK is recruitment to the plasma membrane through interaction between its pleckstrin homology (PH) domain and PIP3, a product of PI3K activity[349]. BTK, together with SYK, phosphorylates and activates PLC$\gamma$2 after stimulation. BTK is generally overexpressed in CLL, though not always constitutively active, and novel BTK inhibitors are amongst the most promising new therapeutic agents in CLL[350,351]. Ibrutinib (formerly PCI-32765) has shown remarkable efficacy in relapsed/refractory patients with CLL, with few adverse effects, and trials combining it

with other agents are ongoing. Again, the class effect of lymphocytosis was noted, coupled with shrinkage of lymph nodes.

### 2.10.8 Phosphoinositide 3-kinase (PI3K)

Phosphatidylinositol-3-kinase (PI3K) stimulation and resultant production of phosphatidylinositol-3,4,5-triphosphate (PIP3) occurs in parallel to SYK activation. This occur by multiple mechanisms, including direct interaction with LYN[352], or the adaptors BCAP and CD19[353]. CD19 also mediates further activation of LYN, and recruits other positive signal mediators such as BTK and VAV[354]. Various effectors, including VAV, BTK, PLC$\gamma$, AKT and SOS, are recruited by their pleckstrin homology (PH) domains to the membrane by binding PIP3. These effectors are therefore brought into close proximity to the BCR, upstream activators such as SYK and downstream targets such as PIP2. Generally, all the principal signalling pathways downstream from the BCR are dependent on PI3K and the accumulation of PIP3.

CLL cells receive survival signals via CD40 and BCR ligation, which are abolished by use of a PI3K inhibitor[62]. Constitutive PI3K activation has been noted in CLL [355,356] .

The Inositol phosphatase PTEN negatively regulates PI3K activity by degrading PIP3 and terminating signalling. PTEN is reduced in approximately half of CLL cases[357]. Activation of the PI3K pathways in CLL ultimately leads to upregulation of Mcl-1, whilst inhibition of PI3K leads to cell death via loss of Mcl-1[62,320,356]. The PI3K$\delta$ isoform-selective inhibitor GS1101 (formerly CAL101) has showing promising in vitro and clinical effects in CLL[358,359]. Clinical experience with GS1101 is perhaps the most mature in terms of trials, and results in combination with other agents are beginning to be reported[360].

AKT is one of the most important downstream effectors of PI3K, and mice with constitutively active AKT develop lymphomas[361]. Some studies in CLL have detected the presence of phosphorylated AKT in unstimulated CLL cells[320,362,363], while others have not, despite PI-3K activity[62,115,355,356]. One study used sonication coupled with lysis to reveal constitutive Akt activity in all CLL cases studied, coupled with sensitivity to Akt inhibitors, indicating that the method used for Akt activity estimation is crucial[362].

### 2.10.9 Protein kinase C (PKC)

The PKC family of enzymes are activated by the presence of $Ca^{2+}$, diacylglycerol (DAG) or other factors and they function in an array of cellular processes that can be specific

for a particular cell type. In B cells, PKCζ, PKCβ, PKCδ and PKCε play important roles in regulating signals generated by the BCR[364]. PKCβ is a classical PKC that can be activated by $Ca^{2+}$, DAG and phospholipids and is important for NF-κB activation upon BCR engagement[365]. PKCβII comprises 0.53% of total cellular protein in CLL, far in excess of the other PKC isoforms[364]. PKC-β inhibition with enzastaurin results in CLL cell apoptosis[366]. In the Tcl-1 transgenic mouse (which develops a CLL-like leukaemia), knockout of PKC-β abrogates the leukaemia phenotype[367]. There is pronounced variability in expression between CLL cases, with no correlation with *IGHV* gene mutation status or CD38 expression, but higher levels in patients with 11q and 17p deletion. There is also a correlation with stage, with higher levels in advanced Binet stage. CLL cells with higher PKCβII activity have lower levels of BCR-induced $Ca^{2+}$ flux[364]. Pre-treatment with the PKC inhibitor LY379196 restores $Ca^{2+}$ flux, and bryostatin (activator of PKCβII) leads to falls in levels of $Ca^{2+}$ flux in cells with low PKCβII activity. It has been suggested that this variation in enzyme activity may explain the heterogeneity of BCR signalling (as measured by $Ca^{2+}$ flux) in CLL. Abrams et al suggest that PKCβII may act as a feedback loop that keeps cell activation below a proapoptotic level[364].

### 2.10.10 Mitogen activated protein kinase pathways (MAPK): RAS/RAF/ERK

The MAPK cascades constitute a group of signal transduction pathways involving successive phosphorylation of coupled serine/threonine or dual specificity kinases.

There are four main groups of MAPK pathways: The p38 MAPK pathway, the extracellular signal-regulated kinase (ERK) family, the c-Jun NH2-terminal kinase (JNK) family and the ERK5 kinase family. The ERK and JNK pathways are relevant to BCR signalling[368]. Suboptimal antigen stimulation, as occurs with soluble autoantigens, can induce tolerogenic signalling that is characterized by the sole activation of ERK and NFAT[369]. In CLL, there appears to be variable ERK phosphorylation, with some authors reporting constitutive phosphorylation[317,319,370]. Farnesyltransferase inhibitors block RAS activity and ERK phosphorylation, resulting in CLL cell apoptosis[371]. The effects are mostly seen in combination with standard chemotherapeutic agents and are thought

to partly result from downregulation of downstream ERK targets such as Bcl-2, Bcl-xL and inhibitor of apoptosis proteins (IAPs).

Simultaneous activation of AKT, ERK and JNK is required for CLL cell proliferation induced by CpG ODN[372]. In both anergic B cells and CLL, antigen stimulation induces low calcium oscillations and activates NFAT and ERK, but does not activate NF-κB or JNK[369]. Whilst BCR ligation induces AKT and ERK activation in most CLL cases, JNK is not activated in the majority, and incomplete degradation of the NF-kB inhibitor IkB is observed [320]. Thus CLL signalling exhibits poor activation of the immunogenic JNK and NF-kB pathways, and some authors suggest that that constitutive activation of MAP kinase signalling pathway in the absence of AKT activation along with NFAT activation is similar to that seen in anergic B cells[115].

### 2.10.11 Nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB)

The signalling pathway controlling BCR-induced regulation of the NF-κB transcription factor plays a critical role in B cell activation and in lymphomas. Genetic defects in the NF-κB pathway lead to various immune deficiencies, whereas aberrant activation leads to autoimmunity and cancer[373]. CLL cells have high basal levels of nuclear NF-κB (p65, cRel and p50 subunits) compared with normal B cells[374]. This is variable between different patients, and correlates with *in vitro* cell survival and resistance to Fludarabine[71]. CLL cells with higher NF-κB levels are more sensitive to NF-κB inhibition[375]. The PI3K/AKT pathway is important in CLL cell survival, and is an upstream activator of NF-κB[363]. Glycogen synthase kinase-3β (GSK3β) regulates NF-κB-mediated BCL2 and XIAP expression in CLL, and inhibitors of this kinase prevent NF-κB binding to target gene promoters[376]. Other pathways in CLL that stimulate NF-κB signalling include TNF-related apoptosis-inducing ligand (TRAIL)[377] and BAFF[63].

Contact with stromal cells and soluble factors in the bone marrow appear to be important in CLL cell survival, and microenvironmental stimuli are one way in which NF-κB pathways may become activated[56]. Contact with BM stromal cells has a greater effect on CLL NF-κB activation than stromal cell-conditioned medium alone[378], though

soluble factors produced by stromal cells such as CXCL12 and BAFF do have an impact on this pathway[379,380].

Activating mutations of the NF-κB pathway are known in other leukaemias, myeloma and lymphoma, though not in CLL. Although proteasome inhibitors were initially thought to selectively target the NF-κB pathway, their effects are multiple, which may explain their disappointing results in treating patients with CLL. The cyclin dependent kinase inhibitor flavopiridol has some efficacy in CLL, and there are suggestions that it has an effect via NF-κB pathway inhibition[381].

**2.10.12 Phosphatases: SHIP, SHP-1, CD45**

Multiple feedback mechanisms modulate BCR signalling. Phosphatases are intimately involved in regulating BCR signalling, yet the picture is complex as each phosphatase has multiple targets, and individual phosphorylation sites may be dephosphorylated by multiple phosphatases[382]. PIP3 accumulation is inhibited by the SH2-domain containing phosphatidylinositol 5-phosphatase (SHIP-1) and by phosphatase and tensin homologue deleted on chromosome 10 (PTEN). The SHP-1 phosphatase also plays a role in feedback inhibition by dephosphorylating Igα, Igβ, SYK, VAV, and BLNK via recruitment to the Immunoreceptor tyrosine-based inhibitory motifs (ITIMs) on CD22 and CD72[383]. CLL is characterized by its low expression of CD22 compared to normal B cells[384]. Autoantibody production is greater in SHIP-1 and SHP-1 knockout mice compared to PTEN knockouts, therefore it is suggested that the former play a more important role in maintaining anergy[234]. Anergic cells contain reduced levels of PIP3, which may be a result of high SHIP-1 and PTEN activity. SHIP-1 is phosphorylated upon BCR stimulation and recruited to LYN via the DOK adaptor[234].

One study showed expression of SHP-1 and SHIP-1 in all 49 CLL cases studied by western blotting. SHIP-1 expression and constitutive phosphorylation were generally higher in ZAP70$^-$ cases, and BCR crosslinking caused increased SHIP-1 phosphorylation in ZAP70$^-$ cases only[385]. SHP1 hypermethylation has been detected in CLL. The protein tyrosine phosphatase receptor type O (PTPROt) is suppressed by methylation in various solid cancers, and a truncated version of PTPROt may be involved in

lymphomagenesis. PTPROt is methylated and suppressed in a majority of CLL, and re-expression induces apoptosis. The repression of this phosphatase in CLL may suggest enhancement of BCR signalling, as SYK is one of its substrates.[386] LYN and ZAP-70 are also substrates of PTPROt, and it may also play a role in BCR signalling and SYK activity in DLBCL[387]. Recent work has suggested that the PTPN22 phosphatase is highly expressed in CLL, and that pro-survival pathways may be preferentially activated after BCR ligation is CLL as a result of PTPN22 activity linking the BCR to Akt[388]. Interestingly, genetic variants of PTPN22 have been associated with increased risks of autoimmunity[389].

### 2.10.13 CD38 and other co-receptors

CD38 is both a cell surface enzyme (ectoenzyme) and a surface receptor that has been much studied in CLL[390-393]**.** CD38 expression may influence BCR signalling in CLL. In one study of 20 patients with CLL[307], two groups were distinguished by CD38 expression. IgM levels were the same between the two groups. Low CD38 expressors did not exhibit calcium flux. Crosslinking CD38 did not produce a Ca flux, nor alter BCR-induced Ca flux. Apoptosis occurred in high CD38 expressors stimulated by BCR, but not the low CD38 group. Others have found similar relationships between CD38 positivity and BCR signalling capacity, though it is difficult to tease out whether enhanced BCR signalling is due to the expression of Ig, CD38, ZAP70 or mutation status[305,319].

Another co-receptor that may be relevant in CLL is CD40. In the initiation of an adaptive immune response, multiple signals are necessary. The primary signal is engagement of the antigen-specific receptor (BCR in B cells), but generally speaking secondary stimulatory signals are required in order for the B cells to become activated in response to antigen. The exemplar costimulatory pathway is the interaction of CD40 on B cells with CD40 ligand (CD40L, CD154) on T-cells[394]. CD40L may be expressed by the T-cells in the microenvironment or by the CLL cells themselves[395,396]. Increased expression of CD40L may also be a consequence of enhanced NF-κB activity, as CD40L is a target gene of NF-κB[397]. The NF-κB activity resultant on CD40L stimulation is able to block the pro-apoptotic effect of fludarabine[398]. Both intracellular and surface expression of CD79a, CD79b and IgM is increased by CD40 stimulation in CLL[399]**.** BCR

ligation leads to pERK and pAKT in half of patients, whereas CD40L/IL-4 treatment produces phosphorylation in all cases. [400]. Apoptosis inhibition by CD40 ligation has been seen[395,401-403], but the effects of CD40L are not simple, with some studies finding that CD40L stimulation sensitizes cells to rituximab[404] or SYK inhibition-induced apoptosis [336].

Other microenvironmental influences have effects on BCR signalling, and are in turn affected by BCR ligation. BCR engagement modulates B cell responsiveness to several chemokines, including CXCL12[338,405], suggesting that BCR signalling has an effect on B cell migration, as one might expect for a cell that seeks antigen. SYK activation is important for cell adhesion and chemotaxis of healthy and CLL B cells[338-340,406], as well as in BCR signalling. Toll-like Receptors (TLRs) are important for detection of inflammation and link the adaptive and innate immune systems. BCR stimulation can upregulate TLR expression, and signals from the BCR and TLRs generally synergise via common NF-kB pathways to promote B cell survival[407-409].

### 2.10.13 Summary: BCR signalling

Normal BCR signalling is complex, and not all pathways have been fully elucidated. Many of the most important pathways exhibit abnormalities in CLL, and yet there is marked heterogeneity of findings within and between studies of BCR signalling in CLL, suggesting that factors as yet unconsidered are crucial. Perhaps some of the more persuasive arguments for the centrality of the BCR to CLL pathogenesis derive from the data on therapeutic targeting of these pathways. Small molecule inhibitors of pathway components such as SYK, PI3K and BTK are showing promising activity in patients with CLL and other lymphomas. An understanding of normal BCR signalling pathway mechanisms has enabled researchers to study corresponding abnormalities that influence CLL and other diseases of B-cells such as autoimmunity. A crucial ‘measure’ of BCR signalling is the output of such signalling: the effect on the B-cell itself after BCR ligation. The teleological ‘drive’ for BCR signalling is that of elimination of pathogenic microbes via the humoral arm of the adaptive immune system. Recognition of microbial antigens via the BCR ideally leads to activation and proliferation of specific B-cells that subsequently act as antigen presenting cells or produce specific antibody.

Conversely, in order to avoid the *'horror autotoxicus'*[410] of autoimmunity, B-cells encountering autoantigen via their BCR should undergo anergy or clonal deletion via apoptosis. Therefore, apoptosis or proliferation may be seen as polar alternative outcomes of BCR signalling. These differing cellular fates consequent on BCR ligation have been studied in CLL, and are discussed in detail in Chapter 6. However, the main determinant of such outcomes remains the pattern of activated intracellular pathways after BCR ligation. Whilst exploring these pathways has led to insights into CLL pathogenesis and is leading to new treatments, understanding is sill fragmentary. In particular, the role of IgD as compared to IgM is ill defined, and the mechanistic differences (if any) between IgD and IgM signalling are as yet obscure. In order to gain a more complete picture of the mechanisms of BCR signalling, it may be helpful to consider how new models of intracellular signalling in general are proving useful.

## 2.11 Principles of analysing signalling pathways

### *2.11.1 Challenges to existing models of signalling pathways*

New technologies such as high-throughput genomic sequencing are revealing hitherto unexpected genomic complexity in cancers. Despite their seeming phenotypic homogeneity, tumours of one tissue type have highly heterogeneous sets of defects in dozens of different genes in different patients[411], and intra-tumour heterogeneity within each patient is rapidly becoming the rule rather than the exception[412]. Leukaemias are often cited as examples of relatively 'simple' cancers that exhibit a limited repertoire of driver mutations compared to their solid counterparts, but the number of oncogenic mutations is still estimated to be at least ten in even the simpler leukaemias[413], though chronic myeloid leukaemia (CML) and its dependence on BCR-ABL may be an informative exception. Similarly, RNA interference studies show that a large number and wide spectrum of gene products contribute to tumour phenotype[411]. Thus, it seems that a simple relationship between genomic aberrations and tumour phenotype remains elusive[413].

In line with this, studies of CLL have not uncovered a single genetic defect present in more than 15% of unselected cases, suggesting that this leukaemia also has a complex genetic landscape[40,44,47]. One group has specifically targeted the BCR signalling

pathway in a search for novel mutations. They sequenced 301 known BCR-signalling and NF-kB related genes in 10 patients using modern high-throughput techniques, and found only 4 non-synonymous mutations in 3 patients (in *KRAS, SMARCA2, NFKBIE and PRKD3*)[414]. One study sequenced 70 of the known 95 tyrosine kinase genes in 95 CLL patients, no somatic mutations were observed[415]. Similarly, another group sequenced 515 kinases from 23 CLL patients and found only 6 non-synonymous mutations (*WEE1*, *NEK1*, *BRAF*, *KDR*, *MAP4K3* and *TRPM6)*[416]. Because of its potential for therapeutic targeting, the BRAF gene was sequenced in an additional 120 cases, with only 4 *BRAF* mutations in this validation set.

Moreover, even when a specific gene is identified as important to tumour pathology, manipulation of that gene product does not often lead to an avenue for effective treatment because of the complex consequences propagated through downstream transcriptional, translational and posttranslational circuits[411]. Despite the complexity of genomic aberrations in cancer, some general organizing principles have been suggested. One emerging principle is that the same limited cohort of key pathways governing cell phenotypes may be altered by disparate genetic defects[417]. In CLL, this may be illustrated by the finding of mutations of *MYD88, DD3X* and *MAPK*, each at low frequency, but all in the same inflammatory pathway[44]. A pathway-centric approach is a promising therapeutic strategy since a pathway (or multiple pathways) can be targeted via any of its components rather than targeting each mutated oncogene product which might require a unique approach tailored to each patients' repertoire of mutations[48]. An example of this is the use of the mammalian target of rapamycin (mTOR) inhibitors such as temsirolimus that are effective in patients with renal cell carcinomas. The predominant genetic defect in these tumours is loss of the Von-Hippel Lindau tumour suppressor, rather than any mutations related to the mTOR gene itself [418].

However, even this pathway-centric approach is incomplete, because of the increasing realisation of the importance of cross-talk among regulatory pathways[419]. Any given molecular component of a pathway can be identified to be associated with or interact with multiple signalling, transcriptional regulation, metabolic or cytoskeletal pathway

components[411]. Thus, pathways can never be considered in isolation of one another, as alteration of one leads directly or indirectly to changes in others. It follows that cancer and other diseases are most productively conceived of as a dysregulation of a multipathway intracellular network[411,420]. Moreover, these cell-intrinsic networks connect to components beyond the malignant cells themselves, to networks that represent the tumour microenvironment.

It has been claimed that cells are best conceived of as complex systems rather than as a collection of linear pathways[421]. The study of systems consisting of tens of thousands of interacting components is very complicated, and simplifying abstractions are necessary. Networks are one way in which disciplines as diverse as sociology, epidemiology, molecular biology and physics are attempting to model complex systems[421]. The main goal is to build networks from large datasets and link biological function to the topology of these networks. Many authors have suggested that a corresponding shift in methodology is required: from understanding dysregulated linear pathways based on reductionist techniques to considering these interlinked networks using data derived from high-throughput methods. Paradoxically, it is hoped that by taking a systems biology approach to analysing large datasets, truly descriptive models of tumour signalling pathways will help unify the often disparate genomic lesions underlying cancer. The true power of these new models will be tested by their ability to predict therapeutic targets and model drug resistance.

### *2.11.2 High-throughput methodologies for modelling signalling networks*

An obvious technique that generates large amounts of data that might benefit from consideration as part of a network model is gene expression microarray analysis. One group combined >200 transcriptional profiles from follicular lymphoma, Burkitt lymphoma, mantle cell lymphoma and normal B cells[422]. Interactions were defined by co-expression variation correlates. An interesting finding was that 80% of the 65,000 network interactions identified appeared to be common across tumours and normal B cells, implying a network 'backbone' that operates consistently across various cellular backgrounds. Nonetheless, hundreds of interactions appeared to be tumour-specific, cutting across diverse pathways, and dysregulated pathways often occurred

independent of mutations in components of those pathways. The basal gene expression profile of CLL cells is similar to antigen-stimulated B-cells.[219,343] By studying gene expression at time points after BCR ligation, our group identified a gene expression pattern over time common to CLL and healthy B-cells[423] and this result was subsequently confirmed by others[336]. Gene expression analysis has driven many of the analytical and computational tools in systems biology[424].

Although the causes of cancer lie ultimately in genetic and epigenetic changes, the manifestations are at the protein level, causing the cancer phenotype[425]. Therefore, to understand cell function, we need to understand cellular proteomes: the set of expressed proteins in a given cell or tissue at any one time. Moreover, for functional understanding, we need to consider more than just the expression level of proteins, but also post-translational modifications (PTMs) that regulate protein function. The growing interest in functional proteomics is not only influenced by new developments in complex network analysis, but also by advances in high-throughput proteomic techniques[425-427]. Mass spectrometry techniques now enable the routine analysis of several thousand proteins and PTMs per experiment. PTMs are extensively studied dynamic processes in cell signalling, playing fundamental roles in regulating signalling pathways and encoding information. The most important PTM in terms of protein function and signalling is reversible phosphorylation of tyrosine, serine and threonine residues. About 30% of human proteins contain covalently bound phosphate at any one time [428], and it could be suggested that phosphorylation is the most common 'word' in the intracellular signalling 'language'.

Post-translational modifications modify proteins and alter enzymatic activity, binding affinity and conformation[428]. Protein phosphorylation on serine (85% of protein phosphorylation), threonine (10-15%) and tyrosine (1-2%) residues is reversible and can produce rapid changes in protein tertiary structure that in turn affect protein function and cellular signalling pathways[429]. Despite its low relative abundance, regulated tyrosine phosphorylation is critically involved in most membrane-proximal signal transduction processes, and so is best characterised. This is also reflected by the number of diseases that are associated with mutations in tyrosine kinases.

Phosphorylation can regulate a variety of important functions, including subcellular localization, degradation and stabilization. PTMs control the binding of transcriptional regulatory elements to their regulatory sequence elements and alter the actions of RNA processing enzymes as well as the classically conceived regulation of biochemical activities of enzymes such as kinases[430].

Phosphorylation is a reversible post-translational modification mediated by protein kinases and reversed by phosphatases. The kinases are highly conserved in eukaryotes, and are sub classified into three major subfamilies, based on the residue that they modify (Serine/threonine, Tyrosine and dual specificity kinases). Tyrosine kinases appear to have evolved later than the Ser/Thr kinases[430]. In humans, the genome contains 518 protein kinases and ~140 phosphatases[431,432]. There are a number of diseases that result from mutations in particular protein kinases and phosphatases[430]. With most proteins having multiple phosphosites and most kinases and phosphatases numerous substrates, phosphorylation cannot be understood as pair-wise interactions in linear pathways. Phosphorylation networks are defined by their kinases, phosphatases and substrates. Furthermore, analysis of isolated kinase-substrate relationships in a simplified *in vitro* assay fails to capture the regulatory molecules, subcellular localisation and other influences that operate *in vivo*[433].

Drugs that inhibit protein kinases or phosphatases are being developed to target cancer and other diseases. The overexpression of growth factor receptor tyrosine kinases or their mutation to constitutively active forms is important in the pathogenesis of several cancers. An example of a particularly successful targeting of a constitutively active kinase is imatinib, which is a potent inhibitor of the Abelson tyrosine kinase (abl) that is abnormal in CML. The use of imatinib has revolutionised the treatment of this disease[434].

One of the best-characterized phosphoproteomes is that downstream of the Epidermal Growth Factor Receptor (EGFR)[435,436]. Growth factors such as epidermal growth factor (EGF) activate the classical mitogen-activated protein kinase (MAPK) cascade to cause cell proliferation. The uncontrolled activation of this pathway can

cause cancer[437]. Time-course studies have detailed the multiple phosphorylations that occur after EGF ligation. Increasing levels of particular EGFR mutants shifts signalling away from ERK and STAT3 towards the PI3K pathway, and induces phosphorylation and transactivation of the hepatocyte growth factor MET. Based on this finding, combination of EGFR and MET inhibitors synergised to kill tumour cells in vitro, a rational combination based on analyses of phosphoproteomic signalling networks[435].

The particular combination of kinase and phosphatase activation determines the topology of intracellular signalling networks. The understanding of these networks in cancer cells will be necessary to understand how therapies that target tumour cells affect signalling, and how resistance to therapy arises by bypassing particular network elements. Inhibitors of kinases are promising in their efficacy and lack of toxicity, often due to their selective targeting of kinases. Phosphoproteomics will be instrumental in determining the modes of action of these kinases, as well as the off-target effects and mechanisms of resistance of these kinases. This understanding should lead to better treatments for patients.

Phosphoproteomic studies of CLL and BCR signalling in cell lines have already identified a number of signalling pathways, but techniques have evolved whereby each experiment can generate greater number of phosphopeptides for consideration than early studies. The newer techniques have not been applied to the study of BCR signalling in any primary B-cells, malignant or otherwise.

**2.12 Summary**

CLL is currently incurable using conventional therapy, and there is a need for more effective and less toxic treatments. Small molecule inhibitors of BCR signalling pathways are beginning to show remarkable efficacy. These new therapies are based on the premise that the BCR is central to CLL pathophysiology. Much evidence points towards antigen, in particular autoantigen, as a pro-survival ligand in the CLL microenvironment. Various studies have highlighted abnormalities of BCR signalling pathways in CLL, with some likening CLL cells to anergic B-cells. Conventional conceptions of signalling as linear molecular pathways may be inadequate to

encompass all these abnormalities. The BCR in CLL is manifest as two isotypes, IgM and IgD, and IgD signalling is relatively unexplored in this field.

# 3. Aims and Objectives

Using primary CLL/SLL patient samples from the Barts Cancer Institute Tissue Bank, I aimed to explore the role of IgM and IgD in this disease, and hopefully establish new avenues for targeted therapeutic manoeuvres. The approach taken was as follows:

1. Establish and confirm the pattern of expression of the IgM and IgD BCR isotypes in peripheral blood, bone marrow and lymph node CLL cells.
This is outlined in Chapter 5.

2. Establish and clarify the outcome of BCR signalling via IgD and IgM in CLL, as determined by immediate events such as calcium flux and later outcomes such as cell proliferation and apoptosis.
This is outlined in Chapter 6.

3. Establish the mechanisms of BCR signalling via IgD and IgM in CLL and healthy B-cells by using the novel techniques of phosphoproteomics.
This is outlined in Chapter 7.

# 4 Materials and Methods

## 4.1 Patient Material

The patients for whom tissue and clinical data were available for immunohistochemistry and peripheral blood samples were identified from the St Bartholomew's Hospital (Barts Health NHS Trust) database. Established in 1967, it holds clinical details of all patients who have received treatment at Barts. This database is linked to a second database of stored tissue samples of patients who have consented for storage of material intended for research. Details of data stored, and the methods used by clinical laboratories are in Section 5.2. Tissue samples were stored in accordance with guidelines from the Human Tissue Authority. Ethical approval for the use of anonymised patient samples and data to perform this research was provided by the East London and City Health authority (Ethics approval number 05/Q0605/140)

## 4.2 Immunohistochemistry of Small Lymphocytic Lymphoma (SLL) lymph node sections

### *4.2.1 Introduction*

Paraffin embedded blocks of lymph node and spleen fragments were cut into sections 4μM thick. Antigen specific staining was carried out using commercial antibodies and optimised for the tonsil and SLL lymph node sections. Immunohistochemistry relies on the specific binding of antibody to the target antigen for detection. The signal is amplified by subsequent binding of a multimeric compound capable of binding horseradish peroxidase. The presence of the antibody-enzyme complex is determined by addition of a substance (Diamino Benzoic Acid, DAB) that is chemically altered by the enzyme, producing a dark brown/black colour when acted upon by peroxidase enzyme.

Staining was optimised for temperature, duration of reagent exposure and antibody concentration on tonsil sections, and subsequently larger numbers of slides were stained using an automatic staining system consisting of a slide rack, robot arm with

nozzles to dispense reagents, pump system and reservoirs of bulk reagents and a dispensing rack for the various primary antibodies and secondary amplification/colourimetric reagents, controlled by an incorporated software package (Dako Autostainer Plus). As a counterstain, nucleus-specific blue/purple dye haematoxylin was used.

### *4.2.2 De-waxing & dehydration*

1. Place slides in oven overnight at 60$^{o}$C in plastic racks
2. Remove paraffin using xylene: 5 minutes suspension, shake off excess, 5 minutes in second xylene pot
3. Transfer to industrial methylated Spirits (IMS) for 2 minutes
4. Block endogenous peroxidase activity by placing slides into 2% hydrogen peroxide/IMS for 2 minutes x2
5. Transfer to IMS for 2 minutes
6. Rinse in running tap water

### *4.2.3 Antigen Retrieval*

7. Heat 3L of antigen unmasking solution in a pressure cooker to 100$^{o}$C. Once boiling, immerse dewaxed, dehydrated slides in the boiling solution
8. Increase heat to 120-130 $^{o}$C and leave for 10 minutes
9. Remove from heat and cool slides with running tap water for 5 minutes
10. Place slides in trough of wash buffer. Slides are now ready for staining

### *4.2.4 Immunohistochemistry staining using Dako Autostainer Plus*

11. Mark slides using hydrophobic marker pen around edge of section & cover in wash buffer
12. Program Autostainer software specifying number of slides, reagents and incubation times. Top up Autostainer reagents
13. Start Autostainer. Process takes 2-3 hours
14. Rinse in tap water for 5 minutes
15. Counter stain in haematoxylin for 5 minutes, rinse in tap water for 5 minutes
16. Plunge into acid alcoholx7 and then into tap water for 30 seconds
17. Dip in Scott's solution x30, then into tap water for 5 minutes

18. Dehydrate slides through 3xIMS
19. Clarify in xylene
20. Apply DPX mountant and cover slip. Air Dry and label

An additional haematoxylin-eosin slide was prepared for morphology.

### *4.2.5 Reagents*

**Table 4.1: List of reagents and antibodies used in immunohistochemistry**

| Reagent | Supplier |
|---|---|
| Antigen Unmasking Solution | Vector Laboratories |
| Hydrogen Peroxide | VWR |
| Wash Buffer | Dako |
| Xylene | Sigma |
| Industrial Methylated Spirits | VWR |
| Haematoxylin Gill 2 | Merck |
| Scotts Solution | VWR |
| DAB buffer | Biogenix |
| Liquid DAB | Biogenix |
| SS Label Polymer HRP | Biogenix |
| DPX Mountant | Sigma |

| Antibody | Dilution | Supplier |
|---|---|---|
| Ki-67 (MIB-1) | 1:2000 | Dako |
| IgD monoclonal (clone DRN1C) | 1:1000 | Novocastra |
| IgM monoclonal (clone 8H6) | 1:1000 | Novocastra |

### *4.2.6 Image Analysis*

Slide Images were captured using the 3D HISTECH scanner microscope and Panoramic Viewer software (3D HISTECH Ltd). Cells were counted manually using the Panoramic Viewer Marker Counter function. Further Details are available in Chapter 5.

## 4.3 Peripheral Blood samples

Patient peripheral blood (PB) samples were collected after informed consent was obtained. Samples were collected into sterile clinical grade vacutainers containing ethylenediamine tetraacetic acid (EDTA). All samples were processed in laminar flow hoods designed for cell culture. Enrichment of mononuclear cells was performed by density centrifugation:

Whole blood samples in EDTA were layered onto Ficoll-Paque (Miltenyi) in sterile 15mL falcon tubes and centrifuged at 1500rpm for 25 minutes. The peripheral blood mononuclear cell layer aspirated and was then washed in RPMI-1640 medium (PAA) and resuspended in 90% foetal bovine serum (PAA)/10% dimethyl sulphoxide (Fisher)

and stored in liquid nitrogen. Storage was in the St Bartholomew's Hospital Tissue Bank, Centre for Haemato-Oncology, John Vane Science Centre, Charterhouse Square, London.

### 4.4 Tonsil Samples

Tonsils from patients requiring tonsillectomy were used after obtaining informed consent from patients or parents. Fresh tonsils were disrupted using forceps and scalpel in sterile conditions and suspended in RPMI. Suspensions were filtered through a 70μM mesh, washed twice in RPMI and enriched for CD19 positive B-cells using negative selection by magnetic beads. The AutoMACS pro automated selection machine (Miltenyi biotec) was used in conjunction with B-cell isolation kit II (Miltenyi biotec). Non-B cells are labelled with CD2, CD14, CD16, CD36, CD43, and CD235a were depleted from cell suspension by binding to beads. Following this, cells were washed in RPMI, resuspended in 90% foetal bovine serum (PAA)/10% dimethyl sulphoxide (Fisher) and stored in liquid nitrogen. Purity of B-cells was confirmed after thawing, with CD19 positivity of >90% in all cases.

### 4.5 Cell Lines

The RL lymphoma cell lines, readily available in our laboratory, was used for optimisation of techniques, as it produced robust calcium flux on IgD or IgM crosslinking. This is derived from an undifferentiated B-Non Hodgkin Lymphoma (www.dsmz.de). Culture and stimulation of cells were performed on cell suspensions at 0.5-1x$10^6$/mL.

### 4.6 Cell Culture

#### *4.6.1 Culture Medium and Conditions*

All cells were cultured in Roswell Park Memorial Institute (RPMI)-1640 medium (PAA) supplemented with 10% heat-inactivated Fetal Bovine Serum (PAA), 2 mM L-glutamine, and 1 mM sodium pyruvate (Invitrogen, Carlsbad, CA) supplemented with gentamicin (Gibco) at a final concentration of 50μg/mL. Cell culture was performed at 37°C in a humidified atmosphere containing 5% $CO_2$.

Frozen samples were thawed, washed twice in RPMI-1640 and resuspended in RPMI-1640 medium (PAA) supplemented with 10% heat-inactivated Fetal Bovine Serum (PAA), followed by determination of cell numbers and viability. Cells were cultured at 37°C in a humidified atmosphere containing 5% $CO_2$for 2-3 hours prior to experiments. Generally, experiments were performed at 0.5-4 x$10^6$/mL, further details in relevant sections.

### *4.6.2 Determination of Cell Viability*

Cell viability was assessed by the trypan blue dye exclusion method using the automated V-Cell XR (Beckman Coulter). Only cells with viability >90% were used in experiments.

### *4.6.3 Stimuli*

BCR crosslinking with soluble F(ab')$_2$ specific to surface IgD or IgM was used (Cambridge Biosciences) at a concentration of F(ab')$_2$ of 1-20μg/mL for initial optimisation. Subsequently, a concentration of 10μg/mL was utilised as stimulus, as this typically produced robust calcium flux. Control comprised the pH 8.2 100mM borate buffered saline (Sigma). Buffer was reconstituted in sterile distilled water, filtered (0.2μM sterile mesh) and heated at 90$^o$C for one hour. 0.5mg/mL Fetal Bovine Serum (PAA) was added and the buffer stored at 4$^o$C. In some cases, soluble F(ab')$_2$ specific to surface IgG (Cambridge Biosciences) was also employed as a control.

### *4.6.4 PI3K inhibitor treatment of cells*

CAL101 (GS1101) was obtained from Active Biochem and resuspended in dimethyl sulphoxide (DMSO). For cell treatments, CAL101 or equal amounts of DMSO vehicle were added to a final concentration of 10μM.

## 4.7 Flow Cytometry

### *4.7.1 Introduction*

Flow cytometry permits the simultaneous detection of different molecules present on cells in liquid suspension flowing in a stream within a flow cytometer. Lasers of different frequencies illuminate cells; filters and detectors can distinguish different

wavelengths of fluorescence. Monoclonal antibodies coupled to fluorochromes can determine the presence or absence of specific antigens on the cell surface or within cells after fixation and permeabilisation. The relative fluorescence intensity detected on single cells provides an estimation of the number of molecules on the cell. The use of DNA binding dyes such as propidium iodide (PI) can distinguish live from dead cells (as dead cells expose DNA once plasma membrane integrity is lost), and also determine the amount of DNA in single cells.

Flow Cytometry was performed on the FACScalibur flow cytometer equipped with a 15mW argon ion laser emitting at 488nm and a 635 nm. Isotype controls consisted of fluorochromes conjugated to non-specific antibodies to provide an estimate of background staining. These were used to define cells as negative for a particular marker, when suitable internal control negative populations were unavailable. Flow cytometric data were analysed using CellProquest and FloJo software.

### *4.7.2 Antibodies*

**Table 4.2 List of antibodies (all from BD Biosciences)**

| Antigen | Isotype | Fluorochrome |
|---|---|---|
| CD19 | IgG1k | FITC |
| CD5 | IgG2ak | APC |
| IgD (clone IA6-2) | IgG2ak | FITC |
| IgM (clone G20-127) | IgG1k | APC |
| IgM (clone G20-127) | IgG1k | FITC |
| CD27 | IgG1k | APC |

### *4.7.3 Determination of CD19, CD5, CD27, Immunoglobulin M (IgM) and D (IgD) expression on primary CLL cells*

Wash cells and resuspend in 50μL 2% PBS/Fetal Bovine Serum (FBS) (PAA) + 0.1% sodium azide (Sigma). Add Fc blocker (human gamma globulins, HAG, Sigma) (50μL per million cells). Incubate for 20 min at $4^0$C. Add 5μL Ab, incubate for 30 min at $4^0$C.

1. *IgD-FITC+ IgM-APC*
1. *CD19-FITC + CD5-APC*
2. *IgD-FITC + CD27-APC*
3. *CD19-FITC alone*
4. *Relevant isotype controls*

Wash with PBS/FCS/azide. Resuspend in 400μL PBS/FCS/azide. Maintain at $4^0$C. For selected CD19 staining, add 1μL of stock Propidium Iodide (10mg/mL, Sigma) solution. Fluorescence at wavelengths of 530 and 630nm was detected after excitation at 488nm on a BD FACScalibur flow cytometer.

### *4.7.4 Calcium Flux*

Cells at $1x10^6$/mL were incubated with 1 mM Fluo-3 AM and 0.02%(v/v) Pluoronic F-127 (Molecular Probes) for 30 minutes at 37°C with 5% $CO_2$.[438] Cells were washed in RPMI and resuspended at $0.5x10^6$/mL in RPMI and kept on ice. For early experiments, the medium in which the cells were suspended for analysis was varied. Either RPMI or phosphate-buffered saline (PBS) was used. Ethylene glycol tetraacetic acid (EGTA, similar to EDTA but with greater selectivity for chelating $Ca^{2+}$ over $Mg^{2+}$) added to the medium was used to chelate extracellular $Ca^{2+}$.

Suspended cells were warmed to 37°C for at least 15 minutes prior to stimulation. Fluorescence at a wavelength of 530nm was detected after excitation at 488nm on a BD FACScalibur flow cytometer. Ionomycin (Sigma-Aldrich) at a concentration of 1μM was used to produce calcium flux as a positive control. Borate buffer was used as negative control. Stimuli were Goat $F(ab')_2$ Anti-Human IgD, IgM or IgG (Cambridge Biosciences) warmed to 37°C. The tube containing 1mL of suspended cells was removed, 100 μL stimulus added with 2x pipetting to mix and replaced on the platform, all within 5 seconds. The tube was also surrounded by a water bath (beaker) at 37°C during acquisition.

Flow cytometric data were analysed using CellProquest and FloJo software. A bivariate plot of fluorescence intensity at the desired wavelength (in our case utilising the FL1 sensor) against time was produced. In optimisation experiments, crosslinking of CD10 and CD20 was performed with unconjugated rabbit IgG (Dako). In order to provide a quantitative measure to enable comparison between experiments, threshold intensity was chosen based on unstimulated cells' fluorescence at the $85^{th}$ centile, and percentage of cells rising above the threshold were calculated (see results for details). The percentage of cells undergoing calcium flux 1 minute after stimulation was

measured, 15% was subtracted and was divided by the same measurement derived from ionomycin to yield a calcium flux as a proportion of maximum.

### *4.7.5 Apoptosis assays using Annexin V and Propidium Iodide*

Cells were stimulated as indicated in the relevant sections. Positive controls were produced by exposure to UV light for 5 minutes 24h prior to assay. Apoptosis was determined at baseline, 24 and 48 hours. The Apotarget Annexin V-FITC/Propidium Iodide apoptosis assay kit (Invitrogen) was used to stain cells and Fluorescence at wavelengths of 530 and 630nm was detected after excitation at 488nm on a BD FACScalibur flow cytometer. Flow cytometric data were analysed using CellProquest and FloJo software. Plots of Annexin V *vs* PI staining were divided into quadrants based on populations seen. These were used to characterise cells into viable, early apoptotic and late apoptotic/necrotic populations.

### *4.7.6 Cell cycle analysis using Ki67 and Propidium Iodide*

Cells were stimulated as indicated. Cell cycle stage was determined at baseline, 24 and 48 hours. Cells were fixed and permeabilised in paraformaldehyde and Triton (BD biosciences) and stained with Ki67-FITC monoclonal antibody (Invitrogen) and Propidium Iodide (PI, Sigma). Fluorescence at wavelength of 530 and 630nm was detected after excitation at 488nm on a BD FACScalibur flow cytometer. Flow cytometric data were analysed using CellProquest and FloJo software. Doublet discrimination was performed based on FL3-W. Plots of Ki67-FITC *vs* DNA content/PI were divided into quadrants based on isotype controls. These were used to divide cells into $G_0/G_1/G_2/S$ phases of the cell cycle.

## 4.8 Phosphoproteomics by Mass Spectrometry

### *4.8.1 Introduction*

In order to generate a global view of serine, threonine and tyrosine phosphorylation within a sample, enrichment techniques are used. The label free approach for enrichment and quantification of phosphopeptides used in this study has been previously published[439,440]. CLL cells were stimulated as indicated, and rapidly chilled

and washed before freezing as cell pellets. A parallel sample was taken for western blotting. Calcium flux was performed an hour before proteomic samples were stimulated to ensure viability and capacity for signalling. The pellet was then lysed in Urea containing phosphatase inhibitors. Lysate containing 0.5mg was then denatured, cysteines alkylated, and the sample digested to fragment peptides using trypsin. Samples were desalted and then enriched using Titanium dioxide ($TiO_2$) beads. The phosphorylated peptides were then dried before analysis by liquid chromatography-mass spectrometry. Data analysis was carried out using MASCOT (phosphopeptide identification) and PESCAL (quantification based on normalised peak height) software. Triplicates for each experimental condition were compared. Further analysis details are available in Chapter 7.

### *4.8.2 Bradford Assay for protein concentration*

Bradford (Biorad) working solution was prepared at 1:5 dilution, and assay performed using flat-bottomed 96-well plates (Fisher) in a PolarStar plate reader (BMG Labtech). Absorbance at 595nm was measured. Comparison with a standard curve derived from solutions of Bovine serum albumin was used as comparison, with Pearson correlation coefficient >0.99.

### *4.8.3 Stimulation*

Cells at $4x10^6$/mL were suspended in medium and stimulated with control borate vehicle, anti-IgM and anti-IgD $F(ab')_2$ at 10μg/mL. Triplicates of $15\text{-}20x10^6$ cells were harvested after 5 minutes and rapidly chilled in Phosphate Buffered Saline with phosphatase inhibitors (1 mM $Na_3VO_4$, 1 mM NaF). An additional aliquot of $5\text{-}10x10^6$ cells for western blotting validation were taken off into a separate pellet.

### *4.8.4 Cell lysis*

- Base Lysis Buffer composition is as follows:

8M Urea in 20mM HEPES (ph8) with 1 mM $Na_3VO_4$, 1 mM NaF, 2.5 mM $Na_4P_2O_7$, 1 mM β-glycerol-phosphate

- The volume of buffer required is as follows: 1ml of Lysis Buffer: 10 x $10^6$ cells. Chill protein lo-bind eppendorfs

a. Add 1 ml of Lysis Buffer preparation to each cell pellet and vortex the sample to thoroughly resuspend and homogenize the sample. Quick spin.
b. Cell suspension can then be sonicated (Soniprep 150 MSE) at 50% intensity, 3 times for 15 seconds.

   Ensure that the tube containing the cell suspension is immersed in ice. This will ensure that heat build up from sonication does not adversely affect phospho-protein content.
c. Centrifuge cell suspension at 20,000g for 10 minutes at 5$^{o}$C, and recover supernatant to a protein lo-bind tube.
d. Bradford assay for total protein count can then be performed and sample(s) can be progressed to the next stage after normalisation to 500μg of protein. Make up to 1mL total volume with 20mM HEPES.

### *4.8.5 Denaturation*

- *Reduction of Disulphide bridges:*

a. Add 10μl of 1M Dithiothreitol (DTT) to sample(s).
b. Vortex sample well and incubate at room temperature for 15 minutes in the dark.

- *Alkylation of cysteines:*

a. Add 40μl of 415mM Iodoacetamide (IAM) to sample(s).
b. Vortex sample well and incubate at room temperature for 15 minutes in the dark.

### *4.8.6 Digestion*

- Protein digestion is achieved with the use of immobilized Trypsin beads (TLCK-trypsin 20 TAME units/mg, Sigma). The following indicates the number of beads required: 80μl of Trypsin beads for 500μg of protein. *Approx 160μL of 50% slurry. Spin down and measure volume to make 80μL per sample*

a. *Trypsin bead conditioning*: Transfer beads to a fresh tube and centrifuge at 2000g for 5 minutes at 5$^{o}$C.
b. Discard supernatant and re-suspend beads in 20mM HEPES (double the volume i.e. if sampled 80μl of beads add 160μl of HEPES buffer). Vortex suspension thoroughly.
c. Centrifuge beads at 2000g for 5 minutes at 5$^{o}$C. Discard supernatant.
d. Repeat this process 2 more times.

e. Resuspend beads , add the same volume of 20mM HEPES as originally sampled i.e. 80µl.
f. Dilute the sample 1 in 4 with 20mM HEPES to a final volume of 4ml in 50mL falcon
g. Add 80µl of conditioned beads to each sample and incubate at 37$^{O}$C for 16 hours with shaking.

shake tube to suspend beads prior to overnight incubation. Beads can have a tendency to settle if allowed to solution allowed to rest.

***4.8.7 De-salting***

a. Chill centrifuge. Acidify digest to 1% of the final volume with Trifluoroacetic Acid (40µl for a 4ml digest).
b. Centrifuge samples at 2000g for 5 minutes at 5$^{O}$C and transfer supernatant to a fresh tube. Sample can now be processed further.

The resultant peptide solutions were desalted by solid phase extraction (SPE) using Oasis HLB extraction cartridges (Waters UK Ltd., Manchester, UK) according to manufacturer instructions with some modifications

*Assemble vacuum manifold as follows:*

- Insert waste container into glass chamber and then seal with the vacuum/taps lid.
- Connect the assembled manifold to a vacuum tap or pump. Ensure that the connection is tight.
- Insert OASIS cartridge (one cartridge/sample) into the tap (s) on the lid and press firmly to ensure a tight fit.
- Close the valve next to the dial and apply the vacuum. Adjust vacuum to approximately 5 mmHg.

*De-salt sample (s) as follows:*

1. Wash cartridge with 1ml of 100% ACN (LC-MS grade). Add the ACN and purge leaving a little solution behind.
2. **NB**: it is very important to maintain the cartridge fully solvated. When purging, don't allow the solution to pass through completely, as this may introduce air into the cartridge. This may impact on peptide binding and elution at a later stage.

3. Wash cartridge with 1ml of 98% @$H_2O$ (+ 2% ACN, 0.1% TFA). Purge and then top up with a further 0.5 ml and purge. This is to ensure complete removal of ACN prior to addition of protein digest sample.
4. Add 1ml of protein digest sample to the cartridge and purge the sample slowly (this is to enhance the binding of peptides to the reversed-phase material in the cartridge). Continue to top up the sample as it passes through the cartridge until all the sample has passed through. Do this for all samples and the go to step 5.
5. Wash cartridge with 1ml of 98% @$H_2O$ (+ 2% ACN, 0.1% TFA). Purge solution.
6. Elute peptides with 0.5ml of 1M Glycolic Acid (+ 80% ACN, 5% TFA). Purge slowly so that as much peptide is recovered as possible. Eluent can now be used for further work i.e. $TiO_2$ phosphopeptide enrichment.

### *4.8.8 Phospho-peptide enrichment with $TiO_2$ beads*

Phosphopeptide enrichment was performed using a $TiO_2$ protocol adapted for label free quantitative proteomics.

**TiO2 bead resuspension**

- Re-suspend $TiO_2$ beads (50% slurry, GL Sciences Inc., Japan) by adding 1ml of 1% TFA to glass vial containing 500µg of beads. Seal the vial and homogenize thoroughly before use.

**Solutions**

Prepare the following solvents before starting the enrichment. Rinse out glass beakers. Use glass pipette for elution solution.

**Equilibration Solution:** 1M Glycolic Acid in 80% ACN/5% TFA

**Wash solution 1:** 50% ACN.

**Wash solution 2:** 20mM Ammonium Acetate in 50% ACN (pH 6.8). Prepare fresh just before the enrichment.

**Elution solution:** (5% $NH_4OH$) 250µl of $NH_3.H_20$ stock solution + 750µl of 50% ACN. Prepare fresh just before the enrichment. Check pH is 10-11. Use glass vial.

**Enrichment**

1. Adjust all sample volumes to 1000µl with **Equilibration Solution**. Ensure enough beads.
2. Add 25µl of re-suspended $TiO_2$ beads to the OASIS eluted fraction(s).
3. Incubate sample(s) at room temperature for 5 minutes with rotation.
4. ***Spin column equilibration***

   (PepClean C-18 Spin Columns, Thermo Scientific, Rockford, IL)

- Remove column(s) caps and place column(s) in ***standard 2ml*** eppendorfs.
- Apply 300µl of **Equilibration Solution** to the very top of the walls of the column(s). This is to ensure that the reverse phase RP-C18 material (powdery material in the column) is fully exposed to the solution and that it all gathers above the filter at the bottom of the column.
- Centrifuge column(s) at room temperature for 30 seconds at 2000rpm.
- Discard flow through.
- Repeat step 4.

5. After 5 minutes of incubation, apply 500µl of $TiO_2$ beads /sample solution (ensure sample(s) is/are well homogenized) to an equilibrated spin column.
6. Centrifuge column(s) at room temperature for 30 seconds at 2000rpm. Discard flow through. Repeat steps 5 & 6 with the remaining 500µl of sample.
7. ***Wash 1: Removal of non-phosphorylated peptides***. Apply 300µl of **Equilibration Solution** to spin column(s). Centrifuge column(s) at room temperature for 30 seconds at 2000rpm. Discard flow through.
8. ***Wash 2:*** Apply 300µl of **Wash Solution 1** to spin column(s). Centrifuge column(s) at room temperature for 30 seconds at 2000rpm. Discard flow through.
9. ***Wash 3: Removal of strongly acidic non-phosphorylated peptides***. Apply 300µl of **Wash solution 2** to the spin column(s). Centrifuge column(s) at room temperature for 30 seconds at 2000rpm. Discard flow through.
10. **IMPORTANT:** transfer spin column(s) to fresh **2ml protein lo-bind** eppendorfs.
11. ***Phospho-peptide elution:*** Apply 50µl of **Elution Solution** directly to the RP-C18/$TiO_2$ layer in the column(s). Allow to absorb for 1 minute. Centrifuge column(s) at room temperature for 30 seconds at 2000rpm. Keep flow through.
12. Repeat step 11 two more times and combining all 3 elutions in the same tube.

13. **IMPORTANT:** Centrifuge sample at 10000rpm for 1 minute and check that there is no deposition present at the bottom. If there is, remove supernatant to fresh protein lo-bind tube and proceed to step 13. Deposition may indicate that $TiO_2$ beads have managed to get through the RP-C18 and filter layers. The presence of beads during acidification and drying may radically impact on phospho-peptide recovery!
14. ***Acidification:*** Acidify sample to 10% **Formic Acid** final concentration (15µl of neat FA to 150 µl eluted peptides).
15. Place samples in SpeedVac dryer until fully dry. (36-48h)
16. Add 13µL of reconstitution buffer/enolase. Leave 1h at 4$^{o}$C.
17. Spin 5min 13,0000 at 4$^{o}$C. Take 12µL clear supernatant and put into labelled eppendorfs in MS inserts. Put cap on. Dispose/snap freeze remainder.
18.

### *4.8.9 Nanoflow-Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) (performed with A Montoya and P Cutillas)*

For LC-MS/MS analysis, dried phosphopeptides were dissolved in 12 µl of 0.1% TFA containing a known concentration of enolase standard and analyzed in a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK) connected online to a nanoflow ultra-high pressure liquid chromatography (nanoAcquity, Waters). This ultra-high pressure liquid chromatography delivered a flow rate of 5 µl/min (loading) and 400 nL/min (gradient elution) with an operating back pressure of about 3000 psi. Separations were performed in a BEH 100 µm × 100 mm column (Waters). The mobile phases were solution A: 0.1% formic acid in LC-MS grade water; and solution B: 0.1% formic acid in LC-MS grade acetonitrile. Gradient runs were from 1% B to 35% B in 100 min followed by a 5 min wash at 85% B and a 7 min equilibration step at 1% B. Full scan survey spectra (m/z 350–1600) were acquired in the Orbitrap with a resolution of 60,000 at m/z 400. A data dependent analysis was employed in which the five most abundant multiply charged ions present in the survey spectrum were automatically mass-selected, fragmented by collision-induced dissociation (normalized collision energy 35%), and analyzed in the LTQ. Thus, Full-MS scans were followed by a maximum of five MS/MS scans (m/z 50–2000) resulting in a maximum

duty cycle of 2.5 s. Because chromatographic peaks were about 30 s at the base, these settings ensured that there were at least 10 data points per extracted ion chromatogram (XIC). Dynamic exclusion was enabled with the exclusion list restricted to 500 entries, exclusion duration of 40 s and mass window of 10 ppm.

### *4.8.10 MS Data Analysis*

Searches were automated with Mascot Daemon (v2.2.2; Matrix Science, London, UK). The parameters included, choosing trypsin as digestion enzyme with two missed cleavage allowed, carbamidomethyl (C) was set as fixed modification, and Pyro-glu (N-term), Oxidation (M) and Phospho (STY) were variable modifications. Data sets were searched with a mass tolerance of ±7 ppm and a fragment mass tolerance of ±800 mmu. Hits were considered significant when they had an Expectation value <0.05 (as returned by Mascot). False discovery rates were ~2% as determined by decoy database searches.

An in-house script was used to extract Mascot results, which were then placed in Excel files for further analysis. For peptides with multiple potential phosphorylation sites, the delta score between the first and second hits reported by Mascot was used to identify the correct position[441]

Pescal[442] was used to automate the generation of extracted ion chromatograms (XIC) and to calculate the peak heights and areas. Because of undersampling and the stochastic nature of peak selection for fragmentation in data dependent acquisition experiments, MS/MS data was not obtained from all the phosphopeptides in all the runs. To overcome this issue, phosphopeptides identified by Mascot with a statistical significant threshold were placed in a database of peptides quantifiable by LC-MS. PESCAL was used to quantify the intensities of the peptides present in the database across all the samples. PESCAL uses the m/z and retention time of the selected peptides to construct extracted ion chromatograms (XICs) for the first three isotopes of each ion. This applies restrictions on the molecular mass, retention time, charge, and isotope distribution, which permits the identification of the LC-MS elution profiles corresponding to the studied phosphopeptides with high confidence (false positive

discovery rate < 5%). Windows for XIC construction were 3 ppm and 5 min for m/z and retention time, respectively. The intensity values could then be calculated by determining the peak height and areas of each individual XIC. The resulting quantitative data were parsed into Excel files for further normalization and statistical analysis. Peptide intensities were normalized to the total chromatogram intensity and expressed as a percentage. Phosphopeptide enrichment was >85% for all samples.

**4.9 Western Blotting**

5-10 x$10^6$ cells were stimulated with vehicle, anti-IgM and anti-IgD F(ab')$_2$ at 10μg/mL in parallel to samples used for proteomics. These were then rapidly chilled in PBS with 1 mM $Na_3VO_4$, 1 mM NaF and washed twice in chilled PBS with 1 mM $Na_3VO_4$, 1 mM NaF. Cell pellets were snap frozen on dry ice and maintained at -80$^oC$ until cell lysis.

The RL cell line was used to provide control protein. Cells at a concentration of 1x$10^6$/mL were stimulated with 1mM pervanadate to produce non-specific phosphorylation. Unstimulated RL cell samples were also used.

*For 100mM pervanadate:*

*800μL* $Na_3VO_4$ *(100mM)*

*8.8μL 30%* $H_2O_2$

*15 min standing, protect from light. NB:turns yellow*

*dilute to final conc of 1mM.*

For lysis, 200 μL of Cell- Lytic cell lysis solution (Invitrogen) plus 1 mM $Na_3VO_4$, 1 mM NaF, 2.5 mM $Na_4P_2O_7$, 1 mM β-glycerol-phosphate was mixed with pellets for 15 min on ice, then spun at 14,000 rpm at 4$^0C$ for 10 minutes. Clear supernatant was taken off and protein concentration determined using the Bradford Assay (Biorad).

20-30μg of protein was used for per well. Proteins were separated using sodium dodecyl sulphate -containing polyacrylamide gel electrophoresis (SDS-PAGE), 4-12% Bis-Tris Nupage gels (Invitrogen) at 140V. To visualize the molecular weight of the proteins a standard containing 9 pre-stained protein bands in the range of 20 to 220kDa was placed in the 1st & last well of each gel (Novex Magic Mark XP Protein

Standard, Invitrogen). After electrophoresis, the separated proteins were transferred onto a Polyvinylidene fluoride (PVDF) membrane by Western Blotting,[443] using the iBlot dry method (Invitrogen). Transfer was at 20V, one for 7 minutes (standard transfer), one for 9 minutes (optimised for transfer of high molecular weight proteins).

Blocking with 5% BSA in Tris-buffered saline (TBS) pH 7.6 plus 0.1% Tween-20 (TBST) was used to reduce non-specific binding of the primary antibody. Probed proteins and phosphoproteins were chosen based on known signalling pathways and previous publications as described (See Table 4.3)

### *4.9.1 Primary Antibodies*

**Table 4.3: Primary antibodies used in western blotting.**

| (Phospho)protein | Species | Modification | Supplier | Dilution |
|---|---|---|---|---|
| Filamin A | rabbit | S2152 | Abcam | 1:1000 |
| 40S ribosomal protein S6 | rabbit | S235/S236 | NE Biolabs | 1:2000 |
| Lymphocyte-specific protein 1 | rabbit | S204 | Anaspec | 1:500-1:1000 |
| Myosin regulatory light polypeptide 9 | mouse | S20 | NE Biolabs | 1:500-1:1000 |
| Myosin-9 | rabbit | S1943 | ECM biosciences | 1:500-1:1000 |
| Wiskott-Adrich syndrome protein | rabbit | Y291 | Abcam | 1:500-1:2000 |
| Syk | rabbit | Y352 | NE Biolabs | 1:500-1:1000 |
| Akt | rabbit | S473 | NE Biolabs | 1:500-1:1000 |
| Stathmin | rabbit | S16 | Santa Cruz | 1:500-1:2000 |
| Btk | rabbit | T223 | NE Biolabs | 1:1000 |
| **Total Protein** | **Species** | | **Supplier** | **Dilution** |
| GAPDH | rabbit | | NE Biolabs | 1:2000 |
| Filamin A | rabbit | | Abcam | 1:1000 |
| 40S ribosomal protein S6 | rabbit | | NE Biolabs | 1:2000 |
| Akt | mouse | | NE Biolabs | 1:1000 |
| Btk | rabbit | | NE Biolabs | 1:1000 |
| Syk | rabbit | | NE Biolabs | 1:1000 |

A secondary antibody conjugated to Horseradish peroxidase (Dako) against the relevant species of primary antibody was added at. 1:2,000 dilution. This produces a colorimetric marker by enzymatically degrading the ECL plus developing agent (GE healthcare), the product of which was photographed using the Fujifilm Luminescent Image Analyzer.

Images were stored as TIFF files. Densitometry was performed using the ImageJ software (rsb.info.nih.gov). Images were analysed by optical density, rectangles highlighted the central one third of the band, and area of the peak measured[444]. Further details are in the relevant section.

### 4.10 Statistics

The GraphPad Prism 5 (GraphPad software) and Excel 2008 (Microsoft Corporation) were used to analyse data. Details of analyses are indicated in relevant results sections.

# 5. Results-Immunoglobulin Expression in CLL

## 5.1 General Introduction and Objectives

In order to determine the role of the B-cell receptor and its IgM and IgD isotypes in CLL, we explored expression of the receptor isotypes in different contexts. We sought to examine the expression of IgM and IgD in peripheral blood, bone marrow and lymph node CLL/SLL cells from patients seen at Barts Cancer Centre. Association with prognosis or other biological characteristics was determined in order to ascertain the effect, if any, on biology of the disease. Expression was determined in three contexts:

1. Peripheral Blood IgD and IgM expression in CLL clinical database (section 5.2)
2. Peripheral Blood IgD and IgM expression in stored tissue samples (section 5.3)
3. Bone marrow and lymph node IgD and IgM expression (section 5.4)

## 5.2 Peripheral Blood IgD and IgM expression: CLL clinical database

### *5.2.1 Introduction*

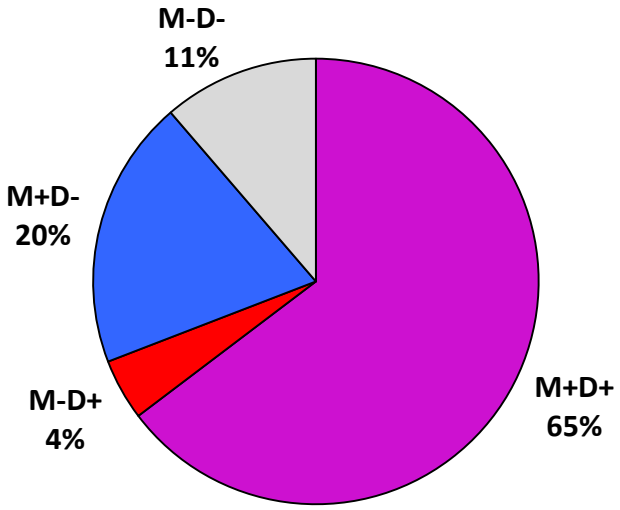
Various studies have examined the expression of IgM and IgD in CLL by flow cytometric immunophenotyping, and attempted to associate them with other factors, including prognosis. The published studies span a number of decades and a variety of techniques, making direct comparison of studies challenging. In addition, the reporting of methods and results in each paper varies such that details of techniques for immunophenotyping, criteria for positivity and other methodological aspects are often not explicit or unavailable from the published papers. Nonetheless, some commonalities can be observed.

In general, IgD and IgM co-expression is found to be a common pattern ($IgD^+M^+$), with $IgD^+M^-$ or $IgD^-M^+$ expression noted at lower frequencies by some authors, though it is often unclear as to how these criteria for positivity were established. Table 5.1 below is an attempt to summarise published results from all papers examining IgD and IgM expression in CLL, where the data regarding the relative proportions in the IgD and IgM characterising subsets are explicit. Figure 5.1 graphically represents the mean values derived from all series.

**Table 5.1: Published series examining IgD and IgM expression in CLL.**
**The lowest row 'Total' takes the number of cases in all published series to derive a mean proportion in each group. Note that some papers do not separate $IgD^+IgM^-$ cases, and in these cases a value of zero is shown.**
**n=number of patients in study**

| Paper | *n* | $IgD^+M^+$ % | $IgD^+$ $M^-$ % | $IgD^-M^+$ % | $IgG^+$ % | Other | notes |
|---|---|---|---|---|---|---|---|
| **Rudders et al (1976)[253]** | *26* | 15 | 0 | 85 | 19 | 1/26 IgA & IgG | All $IgD^+M^+$, cases associated with IgG or IgA |
| **Hamblin & Hough (1977)[445]** | *60* | 25 | 0 | 47 | 0 | 17/60 light chain only | No changes over time (up to 12y) |
| **Baldini et al (1985)[446]** | *76* | 54 | 0 | 38 | 0 | 1 $IgA^+$ | $IgD^-M^+$cases worse prognosis |
| **Kimby et al (1985)[447]** | *66* | 38 | 0 | 33 | 8 | 10 light chain only | $IgD^-M^+$cases in advanced disease |
| **Hamblin et al (1987)[448]** | *200* | 43 | 0 | 34 | 3 | 1 IgA | $IgD^-M^+$ cases longer survival |
| **Geisler et al (1991)[449]** | *540* | 73 | 7 | 12 | 6 IgG/A | 7% no Ig | IgD-no prognostic impact. $IgM^+$ lower overall survival. 30% cut-off for positives |
| **Shen et al (2001)[450]** | *76* | 72 | 21 | 4 | 1 IgG/A | | $IgD^+M^-$ only associated with favourable prognostic factors |
| **Marantidou et al (2010)[451]** | *184* | 88.6 | 0 | 0 | 11.4 | | IgG cases mostly IGHV mutated |
| **Total** | ***1228*** | **63.8** | **4.4** | **19.3** | **5.7** | **5.5** | |

**Figure 5.1: IgD and IgM expression in published studies**
**Summary means of all 1228 cases in 8 published series in each subgroup. Approximately half the $IgD^-M^-$ cases express IgG or IgA, but in many cases light chain only is detected (though in some series IgG/IgA was not sought)**




The 540 patients in the Geisler et al paper dominate these data. In this series, it is sometimes difficult to ascertain the precise methods used. Mean time between diagnosis and immunophenotyping was 10 months (7-17 months). IgM and IgD staining (proportion of cells positive) followed a continuous distribution, and they select an arbitrary cut-off of >30% as a threshold for positivity. The effects of expression on prognosis as assessed by overall survival was not significant for IgD or IgM positivity using the 30% cut-off, but only for IgM mean fluorescent intensity as a continuous variable, but details on how this conclusion was derived are unavailable. There was no attempt to correlate with other biological factors.

Only two publications have examined the pattern of immunoglobulin expression in the recent era where CD38, ZAP70 and *IGHV* mutation status were known to be important prognostic factors in CLL. Shen et al divided cases (n=76) into $IgD^+IgM^+$, $IgD^+IgM^-$ and $IgD^-$ groups, and suggest an association between $IgD^+IgM^-$ cases and low white blood count and CD38 negativity, and by implication good prognosis[450]. The primary aim of Marantidou et al's study was to explore the use of AID splice variants in CLL (N=195), and there is a brief mention of 184 cases that had immunoglobulin isotype data available. Of these, there is no differentiation between IgD and IgM expression (all cases are 'IgMD'), and the IgG positive cases (11%) mostly had mutated *IGHV* genes.

What might be concluded from these various studies is that IgD expression (~70%) is common, as is IgM expression (~85%), with dual expression forming the largest single subgroup. However, methods varied between series, in particular many papers did not attempt to distinguish $IgD^+IgM^+$ cases from $IgD^+IgM^-$ or $IgD^-IgM^+$ cases, cases often designated 'IgMD'. Critically, the levels of expression for a definition of positivity are generally not mentioned, and where they are, it seems that an arbitrary threshold for positivity is chosen, as the level of expression is a continuous variable[449]. There may or may not be an association with prognosis, more often publications have associated the pattern of expression with other known prognostic markers, rather than to associate with prognosis *per se*.

A few more recent studies have considered the expression of IgD and IgM in conjunction with IgG or other class-switched isotypes[452,453]. These studies do not compare IgD and IgM expression as the methods involved generally examine mRNA expression to detect $\mu$–$\delta$ transcripts, and it is known that the regulation of IgD and IgM expression is post-transcriptional[255]. Conventionally, class switching will result in the recombinational deletion of DNA encoding the $\mu$ and $\delta$ chains that are required for IgM and IgD surface expression. Therefore, IgG or IgA expression is in theory incompatible with IgD/IgM expression. Many of the studies mentioned above (Table 5.1) did suggest co-expression of IgD/IgM and IgG, though the expression was assessed for each isotype staining the bulk population, rather than triple positivity on single cells. More often authors suggest that the IgG-positive cells reflect a minority sub-clone, with the majority being IgM/IgD positive[454,455].

One explanation for the expression of IgG in this situation is a technical one. B-cells express Fc receptors whose function is to bind to the constant regions of immunoglobulin, typically IgG. Therefore, CLL cells are often coated with a layer of adsorbed IgG or IgA that mimics expression of those isotypes, whereas it actually reflects polyclonal IgG or IgA present in the patient's plasma. The authors of the first study reporting this in 1975 observed that incubating at pH 5 or in standard culture media for 20 minutes at 37 $^oC$ removed this adsorbed IgG without changing expression

of IgM[456]. This may well explain much of the IgG positivity of CLL in early and subsequent studies.

However, others claim that multiple isotype expression is genuinely possible, and has been observed in cell lines, malignant B-cells and during the primary immune response[457]. Many studies have reported the expression of mRNA for $\mu$ and $\delta$ mRNA as well as $\gamma$ mRNA in CLL, consistent with IgG/M/D co-expression[453]. An explanation for this is intraclonal diversity, and one group suggest that different isotype transcripts utilise the same VDJ segment mRNA, but that mutations had occurred, highlighting the occasional process of ongoing SHM/CSR within CLL, particularly within CLL that expresses IgG[458]. This is supported by several other studies[459,460], though some authors have claimed that true dual IgM and IgG expression in a single cell can still occur[455,461]. It is difficult to definitively establish whether CLL cells that express IgM/IgD genuinely co-expresses IgG or other isotypes, or whether such findings are due to technical issues or intraclonal diversity. Clearly, caution is required when interpreting surface immunoglobulin isotype data.

### *5.2.2 Methods*

Patients seen at the Barts Cancer Centre are invited to consent for storage of their clinical details and tissue storage for the purposes of research. Data gathered included dates of Diagnosis, follow-up, treatment and death. Clinical characteristics such as cell count, lymphadenopathy, immunophenotyping, cytogenetics and results of specialised tests are also included. A variety of assays are performed for diagnostic and treatment purposes, the details of which follow.

### *5.2.3 Immunophenotype by flow cytometry*

*The clinical immunophenotyping laboratory of Barts Health NHS trust performed assays.* All patients with suspected CLL have immunophenotyping performed by flow cytometry. Samples with suspected CLL are gated on $CD19^{+}CD5^{+}$ cells. B-cell clonality is determined by light chain restriction (>10:1 bias $\kappa/\lambda$ or $\lambda/\kappa$). An immunophenotypic 'CLL score'[304] is allocated to confirm diagnosis. The percentage of $CD19^{+}CD5^{+}$ cells expressing surface IgM, IgD and IgG is determined, and positivity defined as >30% of cells expressing the immunoglobulin isotype. Internal negative control lymphocytes

are used as a threshold for positivity. The following is from the immunophenotyping standard operating procedure for immunoglobulin analysis on suspected B-cell lymphoproliferative disorders:

*'This tube is used to identify the IgH chain expression. The gate is usually set according to internal negative control lymphocytes. Frequently however, some cells in the internal negative gate will stain positive and overall, there is more variation between specimens than is usually seen for most other antigens investigated. The CD19+ population may also appear a little higher in background fluorescence compared to other (negative) lymphocytes. True IgH expression is usually seen as a clear shift above background negative fluorescence. As a general rule, a cut-off of 25-30% should be used for reporting IgH positivity.'*

Typically, immunoglobulin heavy chain expression is 'weak' or 'dim' in CLL, in fact this is one of the diagnostic criteria[5]. Weak Ig expression is generally poorly defined, but is locally stated to be present if expression is within one logarithm of non-expressing lymphocytes. Results where >5% CLL cells are ZAP-70 positive are recorded as positive cases. CD38 positivity is defined at a 30% cut-off.

### *5.3.2 Cytogenetics*

*The clinical cytogenetics laboratory of Barts Health NHS trust performed assays.*

Cytogenetic data were included, with the results of G-banding in addition to Fluorescent in situ hybridisation (FISH) for the commonly found abnormalities seen in CLL[26]: deletions of 13q, 11q, 17p, and 6q; trisomy 12 cyclin D1, (as mantle cell lymphoma is considered as a possible diagnosis with atypical morphology or CLL score $\leq$3/5).

### *5.2.4 Other data*

Other markers such as immunoglobulin *IGHV* gene mutation status (*assay performed by Royal Free NHS trust laboratories*), tissue biopsy and bone marrow biopsy were recorded where available. The dates and types of treatment, and response to treatment were recorded. Treatment Free survival is defined as the time between diagnosis and first treatment.

### *5.2.5 Statistical analysis*

Data were tabulated in Microsoft Excel and statistical comparisons were made using GraphPad Prism. Unless otherwise indicated, all statistical comparisons were made using tests designed for use with non-parametric data, as in most cases data was not normally distributed, or could not be assumed to be so. Probability of differences in means being due to chance was calculated using Mann-Whitney test[462] for unpaired data and Wilcoxon test[463] for paired data. Correlations were calculated using Spearman's test. Where patient groups were segregated by IgD/IgM positivity or negativity, the differences in overall survival and treatment free survival were assessed using the Kaplan Meier survival method[464]. Censorship occurred at date of last follow up or date at which event (treatment or death) occurred. Time to that point was measured from date of diagnosis. The log rank test was used to compare survival distributions and generate a probability for the differences in survival being due to chance[465].

### *5.2.6 Characteristics of patients*

We attempted to estimate how representative the patients within the Barts database were of the general population. Almost all requests for immunophenotyping from local networks covering approximately 4 million people in NE London and southern Essex are processed at the Royal London Hospital Flow Cytometry laboratory. The database held details on 1,182 with a diagnosis of CLL/SLL.

Median age at CLL diagnosis was 67.6 years (range 25-99). Published data give median ages at diagnosis of 72 years (US)[2] or 71 years (H*aematological Malignancy Research Network website, www.hmrn.org:* UK data*)*. The slightly younger median age at diagnosis may reflect the tertiary nature of our local referral centre, or the local population, or other factors. Strikingly, the minimum age at diagnosis was 25, with 16/1182 (1.4%) patients having an age at diagnosis less than 35 years. Published data from the comprehensive US SEER database looking at CLL/SLL diagnoses 1987-2004 identified 0.095% of patients <25 years with CLL/SLL and 0.68% <35 years[2].

### *5.2.7 Crude Incidence of CLL*

The local population served by the immunophenotyping flow cytometry laboratory comprises approximately 1.5 million people (NE London Cancer Network), but since the flow cytometry laboratory receives samples from South Essex and other locations outside the cancer network, a wider population of 3.7 million people (based on numbers in the region) is partially covered. See Table 5.2.

**Table 5.2: Crude Incidence of CLL**
**The local population served by the network divides the number of new cases detected by the local flow cytometry laboratory. It would be expected that all diagnoses within the Cancer network population are made in the tertiary centre, but since referrals from outside the network are also made, the denominator local population is larger. Both incidence rates are reported.**

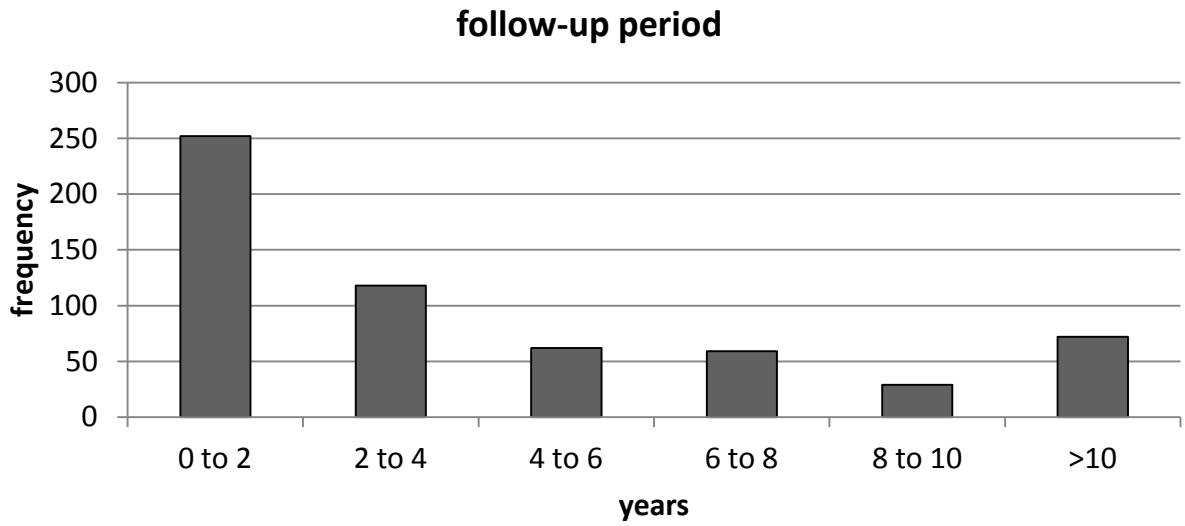
| Year | New diagnoses | Incidence per 100,000 per year (Cancer network) | Incidence per 100,000 per year (wider population) |
|---|---|---|---|
| 2009 | 130 | 8.67 | 3.51 |
| 2008 | 138 | 9.20 | 3.73 |
| 2007 | 121 | 8.07 | 3.27 |
| 2006 | 116 | 7.73 | 3.14 |
| 2005 | 105 | 7.00 | 2.84 |
| 2004 | 76 | 5.07 | 2.05 |
| 2003 | 80 | 5.33 | 2.16 |
| 2002 | 59 | 3.93 | 1.59 |

The crude incidence of CLL/SLL for the last three years is between 3.27-9.20 per 100,000 per year, compared to published crude incidences of 4.2 per 100,000 patient years[466]. The slightly lower incidence estimates may be explained by a younger median age of the population in our area, but may also be explained by incomplete data gathering. It is difficult to determine true incidence rates as the total population determining the denominator is unknown. The lower incidence prior to 2005 probably reflects changes in vigilance, setting up of cancer networks and the increased use of flow cytometry in recent years. Given that the estimate of incidence of new cases of CLL is comparable to national figures, it might be reasonably claimed that coverage of the local population is reasonably comprehensive and the patients in our clinical network under consideration are representative of the general population.

### *5.2.8 Median Follow up*

The median follow-up for patients was only 2.86 years (range 0-35 years), as data gathering processes were more comprehensive in the past 5 years (see figure 5.2). This limits the generalisability of conclusions derived from such a short follow up period.

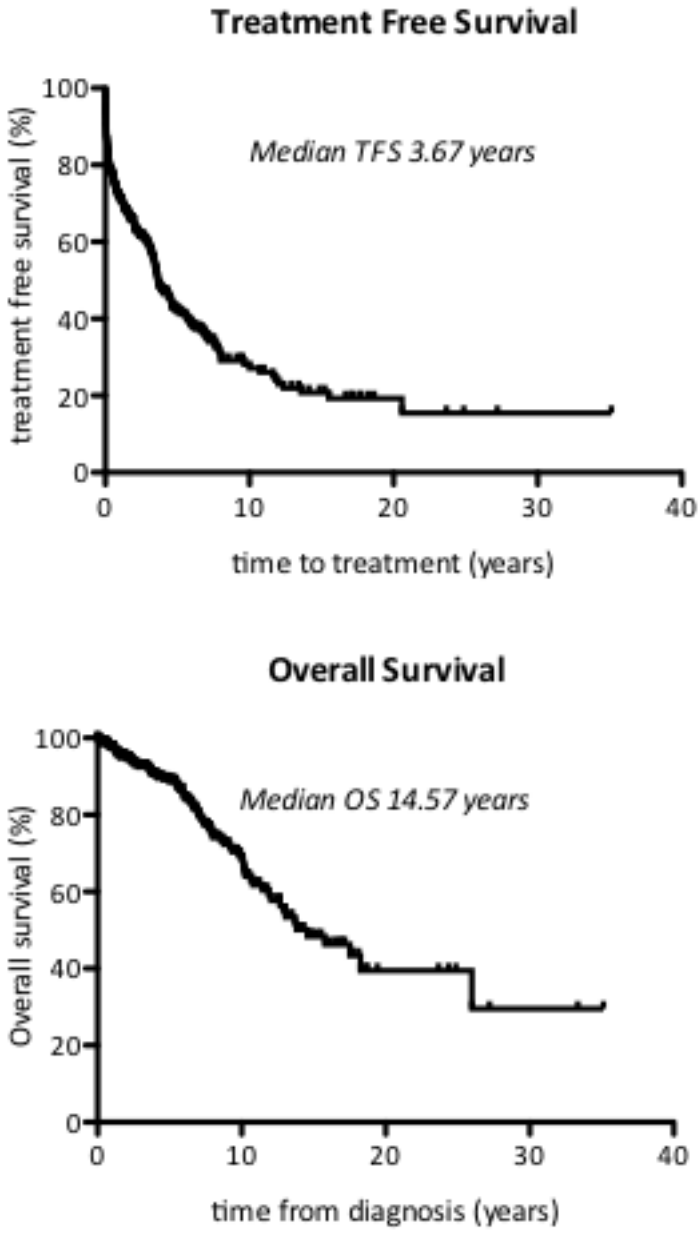
**Figure 5.2: Follow up period of patients**



***Treatment free and overall survival***

Data available on 863/1182 patients, 54% received treatment at some point during follow up (figure 5.3), with many patients not requiring treatment lost to follow up.

**Figure 5.3: TFS and OS of patients with available data**




***5.2.9 Cytogenetics***

Of 1,182 patients with CLL/SLL, only 377 (32%) had cytogenetic data using the standard fluorescent in situ hybridisation (FISH) panel used for lymphoproliferative disorders.

In the original Dohner et al[26] publication, cytogenetic subgroups were classified according to a hierarchy. The five major categories were defined as follows: patients with a 17p deletion; patients with an 11q deletion but not a 17p deletion; patients with 12q trisomy but not a 17p or 11q deletion; patients with a normal karyotype; and patients with a 13q deletion as the sole aberration. Table 5.3 compares the groupings of our patients with that published by Dohner et al:

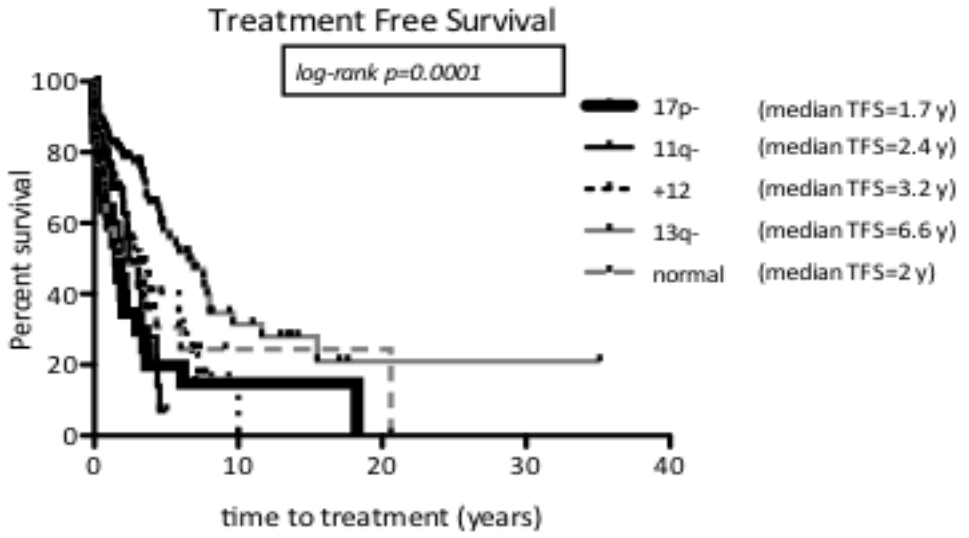
**Table 5.3: Patients in Barts database with cytogenetic data**
**Classified according to hieracrchical group and compared to Dohner et al paper.**

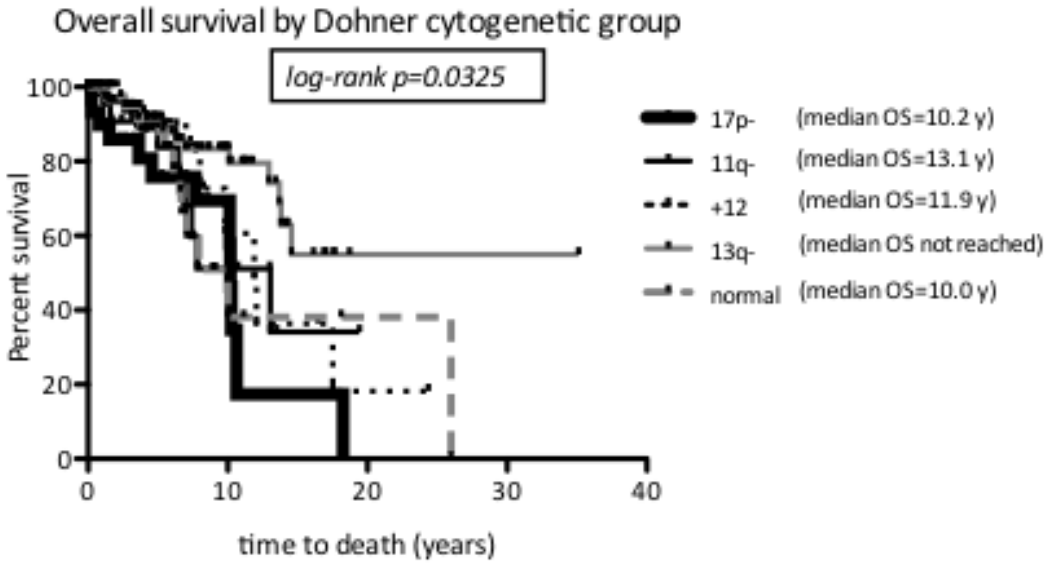
| Cytogenetic subgroup | Number | % in Barts database | Dohner et al % |
|---|---|---|---|
| normal | 85 | 23 | 18 |
| 13qdel alone (inc biallelic) | 164 | 44 | 36 |
| tri 12 (no 11q or 17p del) | 66 | 18 | 14 |
| 11q (no 17p del) | 27 | 7 | 17 |
| 17p | 35 | 9 | 7 |

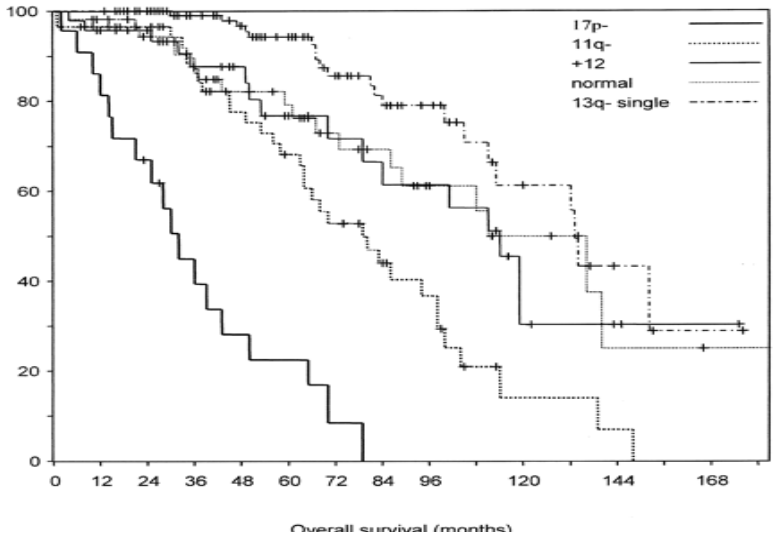
The proportions in each hierarchical cytogenetic subgroup in our group of patients appear similar to that expected. The relative prognostic impact of the different cytogenetic groups is also similar to that found by Dohner et al (Fig.5.4).

The poor prognostic effect of 17p deletion is seen, with a median survival of 123 months for 17p deletion *vs* not reached for 13q deletion. Median overall survival for the patients with 17p deletion in Dohner et al was only 32 months. Studies subsequent to Dohner et al have tended to find longer survival in 17p deleted subgroups[467], which may partly explain our findings.

**Figure 5.4: OS &TFS by cytogenetic subgroup**
**Below is the published OS by subgroup from the original Dohner et al paper**



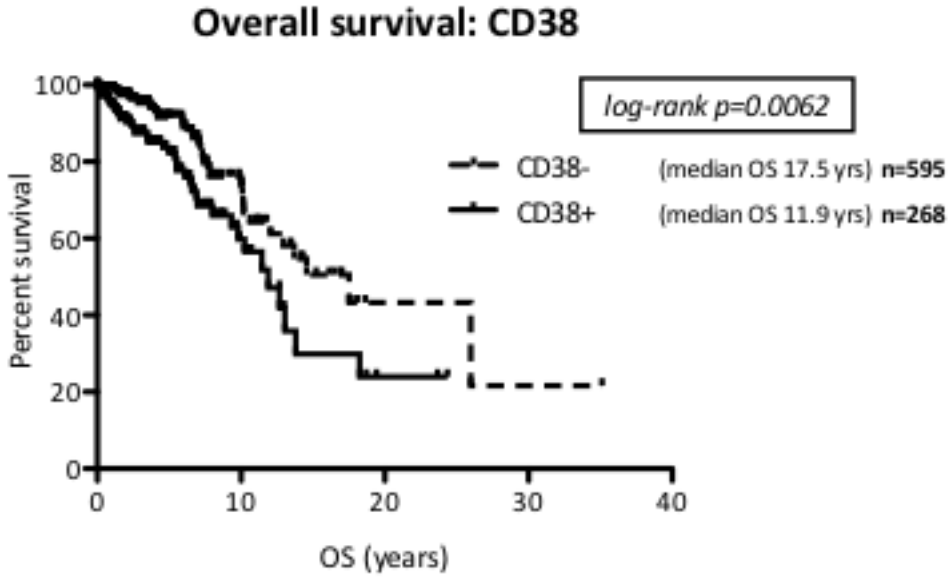



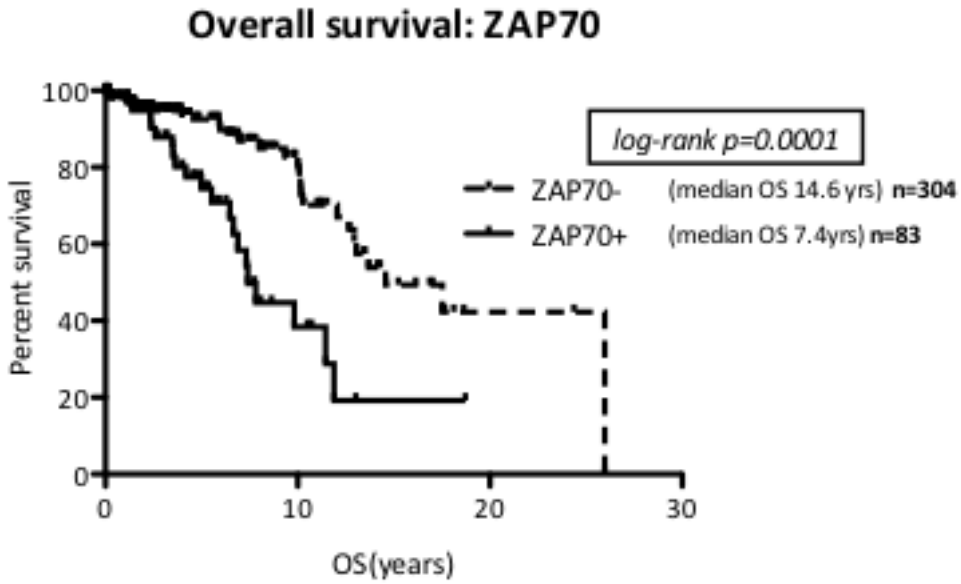
### *5.2.10 Prognostic markers*

The association of immunophenotypic markers with prognosis was derived. Both CD38 (at 30% expression) and ZAP70 positivity (at 5% expression) were associated with reduced overall survival (Figure 5.5):

**Fig 5.5: Overall survival by immunophenotypic prognostic marker group**
**CD38**



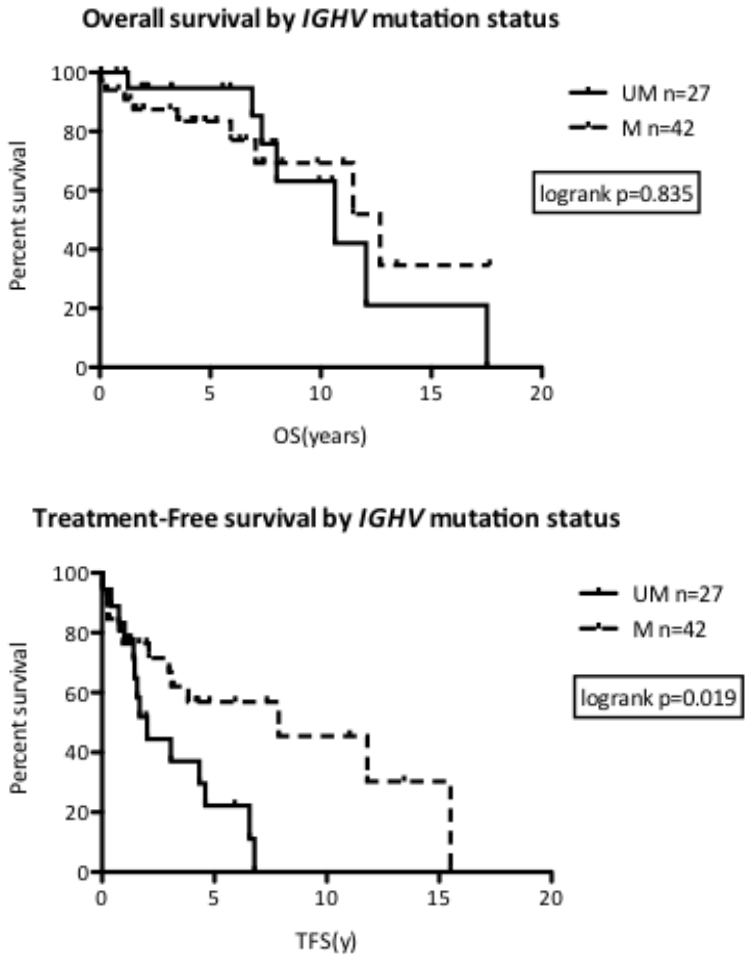

**ZAP70**




***IGHV* gene mutation status and Prognosis**

*IGHV* gene mutation status data were only available for 69/1182 patients. There was no effect on OS, but a significant effect on TFS (Fig 5.6):

**Fig 5.6:Survival by *IGHV* gene status**




In order to confirm known associations between prognostic markers, categorical data were compared using Fisher's exact test. Despite low numbers of patients with overlapping prognostic factor data available, the known associations of *IGHV* gene mutation and CD38 positivity (n=65, p=0.041) and ZAP70 positivity (n=56 p=0.0022) were confirmed in our series. In summary, published prognostic factors maintained their predictive value and known associations in our data set, suggesting that the patients in our database were comparable to published series.

### *5.2.11 Results: Peripheral Blood IgD and IgM expression in CLL clinical database*

Measurement of surface IgM has historically been part of the routine immunophenotyping panel when CLL or another B-lymphoproliferative disorder (BLPD) is suspected. Measurement of IgD was initiated in a few cases in 2003, and became routine in 2008. Repeated measurements of surface immunoglobulin isotypes (Igs) have been made in a limited number of cases. Most cases have had Ig determination on one occasion only, most consistently of IgM. There are many instances where IgD has not been determined. There are also cases where the assay failed. Repeated determination of Ig seems to give mainly consistent results, though there are a number of cases where isotype expression appears to have changed over time. Whether this is due to variability in the assay, or true variability in disease characteristics, is difficult to ascertain.

847 patients had IgM, and 204 patients had IgD determined on at least one occasion. Figures 5.7 and 5.8 demonstrate the range of expression:

**Figure 5.7: Surface expression of IgD as determined by clinical immunophenotyping service.**

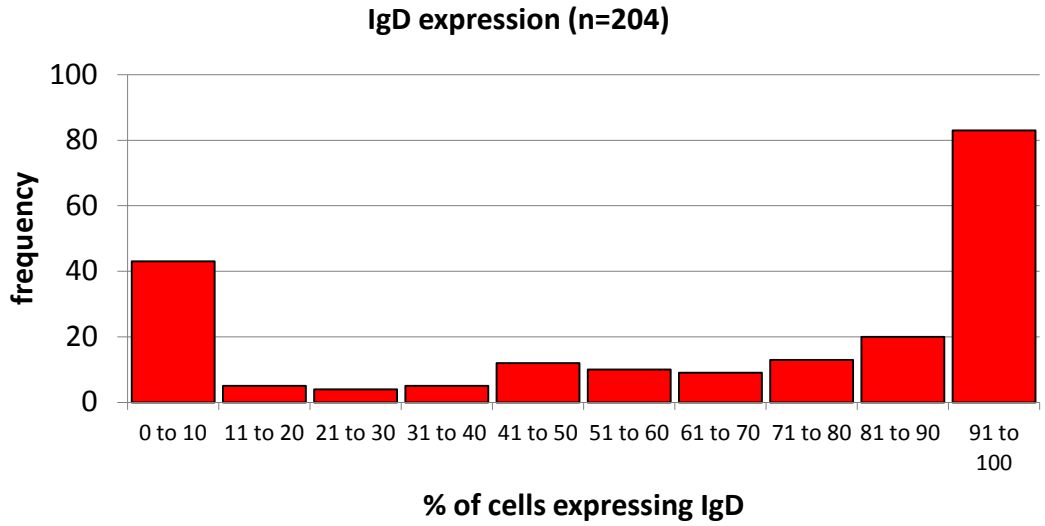
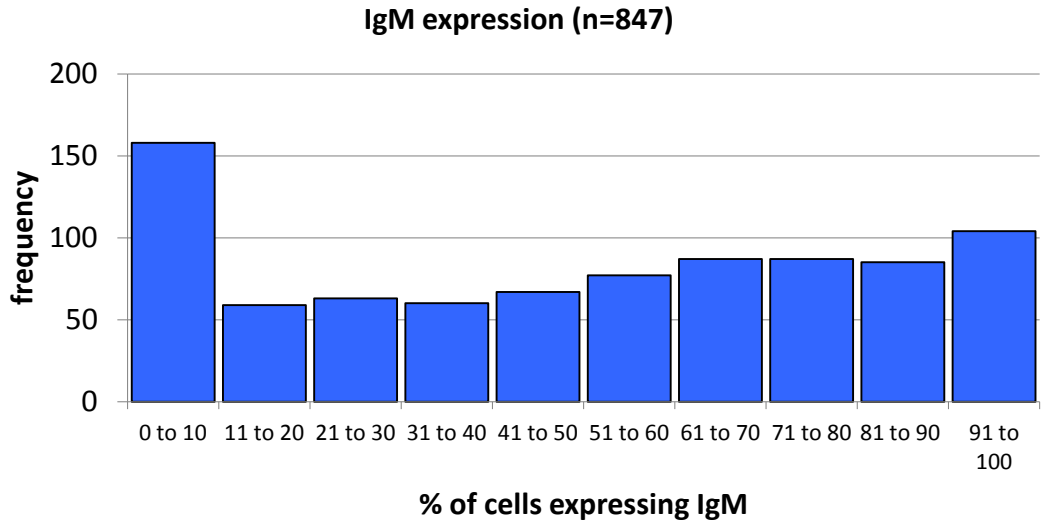



**Figure 5.8: Surface expression of IgM as determined by clinical immunophenotyping service.**



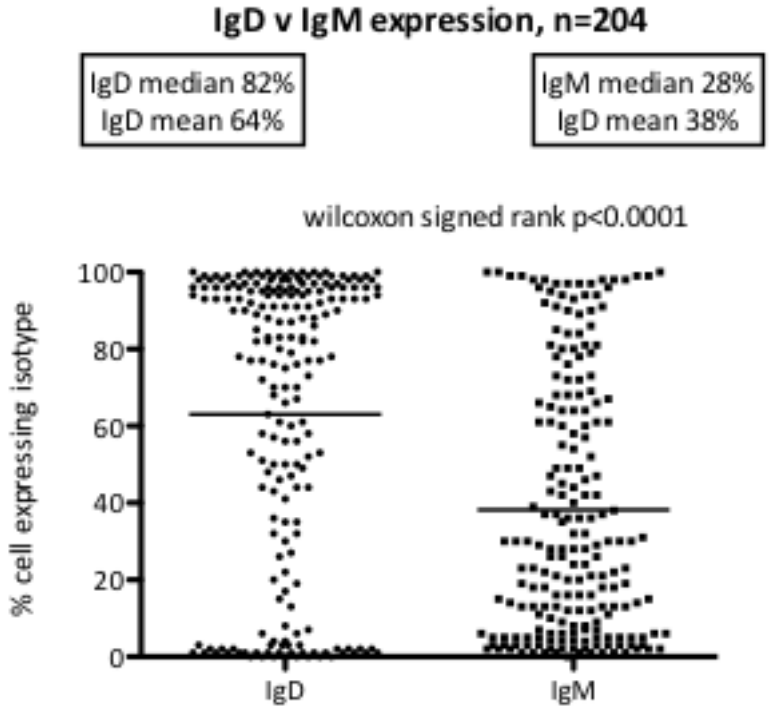

To assign a case of CLL to a categorical classification based on positivity of expression of an isotype, it is necessary to decide on a threshold value of expression. The local clinical immunophenotyping service sets a threshold of ≥30% of $CD5^{+}CD19^{+}$cells expressing a particular isotype. There is, however, a clear continuum of expression seen for IgM, and to an extent with IgD. It could also be claimed that a cut-off of 10% might be employed, based on a visual inspection of the data. The cut-off of 10% mainly alters the proportion of cases that are defined as IgM positive/negative, and the majority of patients would still be classed as either IgM or IgD positive (see table 5.4).

**Table 5.4: Proportion of patients classified by two different positivity thresholds**

| Threshold for positivity | 30% of cells expressing | 10% of cells expressing |
|---|---|---|
| Percentage of patients $IgD^+$ | 74.5 | 78.9 |
| Percentage of patients $IgM^+$ | 66.9 | 81.4 |

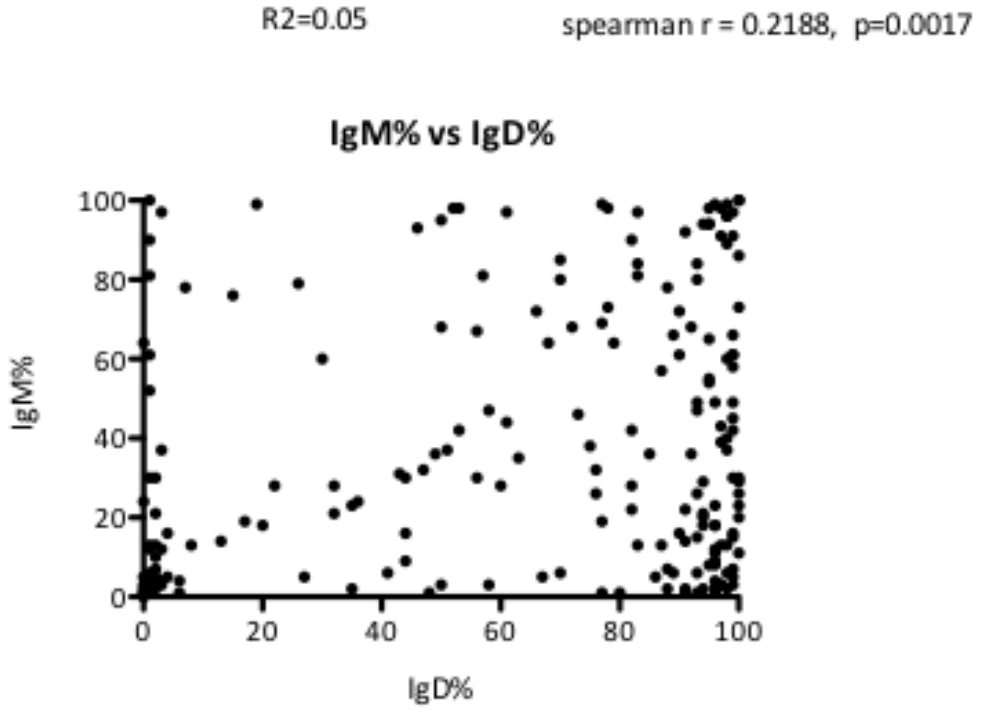
When the percentage of cells expressing IgD is compared to IgM in the same patient, it seems that IgD expression is significantly higher than that of IgM (Fig 5.9)

**Figure 5.9: expression of IgM vs expression of IgD. The horizontal bars represent the mean values**




By plotting the expression of IgM against the expression of IgD (see figure 5.10), it can be seen that there is a weak correlation between IgM and IgD, but with linear regression showing an $R^2$ of 0.05, the relationship between the two is extremely weak and unlikely to be of biological significance.

**Figure 5.10: Expression of IgM vs IgD for the 204 patients with both determined**

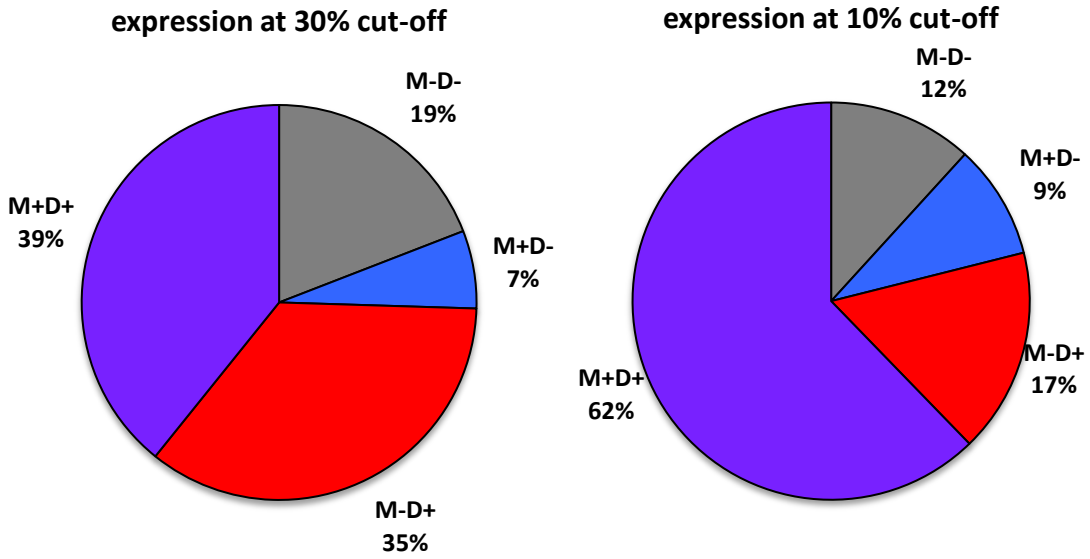



By using a threshold at ≥30% or ≥10% of $CD5^{+}CD19^{+}$cells expressing the relevant isotype, positivity and negativity for the various isotypes is defined (see Table 5.5 and figure 5.11). This classifies patients into subgroups of IgM/IgD positivity, considering simultaneous expression of IgM and IgD in the 204 cases where this has occurred:

**Table 5.5: Frequencies by subgroup classification**

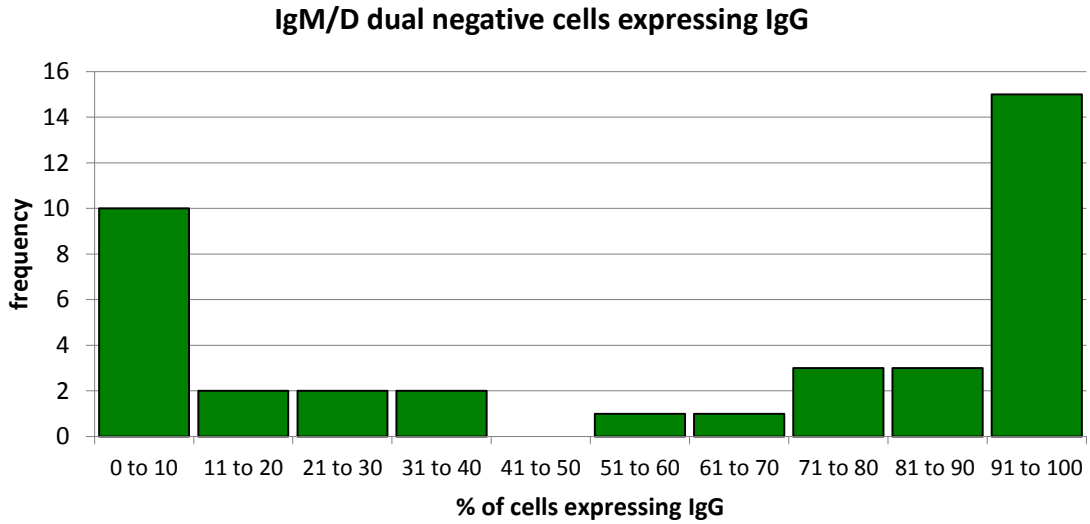
| | **Frequency at 30% cut-off** | **%** | **Frequency at 10% cut-off** | **%** |
|---|---|---|---|---|
| **M-D-** | 39 | **19.1** | 24 | **11.8** |
| **M+D-** | 13 | **6.4** | 19 | **9.3** |
| **M-D+** | 72 | **35.3** | 34 | **16.7** |
| **M+D+** | 80 | **39.2** | 127 | **62.3** |

**Figure 5.11: Proportion by subgroup classification**




Depending on the cut-off employed, there were 39 (30% cut-off) or 24 (10% cut-off) defined as $IgM^{-}IgD^{-}$. Of these cases defined as negative for both IgM and IgD, 79% (at 10% cut-off of IgG expression) and 67% (at 30% cut-off) were positive for IgG (see Fig. 5.12). Presumably, those that did not express IgG, IgM or IgD would be positive for IgA or IgE, as all had clonal expression of surface light chains demonstrated. An alternative explanation is that the assay failed or Ig expression was too low for detection.

**Figure 5.12: Proportion of IgG expressing cells in $IgD^{-}IgM^{-}$ CLL cases**




### *5.2.12 Stability of isotype expression*

Because it is anecdotally observed that the immunoglobulin isotype appeared to change over time in a small number of cases, an attempt to estimate the stability of

this measurement was made. The number of times (in percentage of all cases) the isotype was determined is given below (Table 5.6):

**Table 5.6: Repeated measurement of immunoglobulin isotype**

| Number of measurements | IgM (total 847) % | IgD (total 204) % |
|---|---|---|
| **1 occasion** | 83.5 | 93.0 |
| **2 occasions** | 11.3 | 6.5 |
| **≥3 occasions** | 5.2 | 0.47 |

In total 140 (16.5%) of patients had IgM measured on more than one occasion, and 15 (7%) had IgD measured on more than one occasion. By assigning a threshold of 30% for positivity or negativity in determination of isotype, we can calculate the number of cases that had different classifications of isotype expression on different occasions, i.e. the proportion of all cases that appeared to change their surface immunoglobulin isotype over time (Table 5.7).

**Table 5.7: Changes in isotype expression on repeated measurement**

| Change in isotype? | IgM (total 140) % | IgD (total 15) % |
|---|---|---|
| **No change on repeat measurement** | 72.9 | 86.7 |
| **Change between 2 occasions** | 27.1 | 13.3 |

Thus we see that a sizeable minority appear to change their isotype expression from positive to negative or *vice versa* (at least for the 30% threshold). In general, this change between sample dates was small, for example the mean change in sIgM expression between samples was only 8%, and the majority (74%) of cases that changed their classification from positive/negative had a less than 30% change in expression levels. Perhaps this change is more marked with IgM.

### *5.2.13 Isotype expression in relation to diagnostic date*

The proximity of the Ig measurement to the diagnostic date was also determined (see Table 5.8).

**Table 5.8: Chronological separation between date of diagnosis of CLL and date of immunoglobulin isotype determination.**
**Shown are the numbers (n) and percentages (shaded) of patients in each group**

| | IgM n | **IgM %** | IgD n | **IgD %** |
|---|---|---|---|---|
| **Date of Ig=diagnosis date** | 677 | **79.8** | 136 | **63.6** |
| **Ig < 30 days from diagnosis date** | 9 | **1.1** | 2 | **0.9** |
| **Ig 30 to 360 days after diagnosis date** | 32 | **3.8** | 11 | **5.1** |
| **Ig 1-2 years after diagnosis date** | 19 | **2.2** | 10 | **4.7** |
| **Ig >2 years after diagnosis date** | 111 | **13.1** | 55 | **25.7** |

Therefore, approximately 80% of IgM measurements were performed at diagnosis. This is lower with IgD determination as the assay was introduced later.

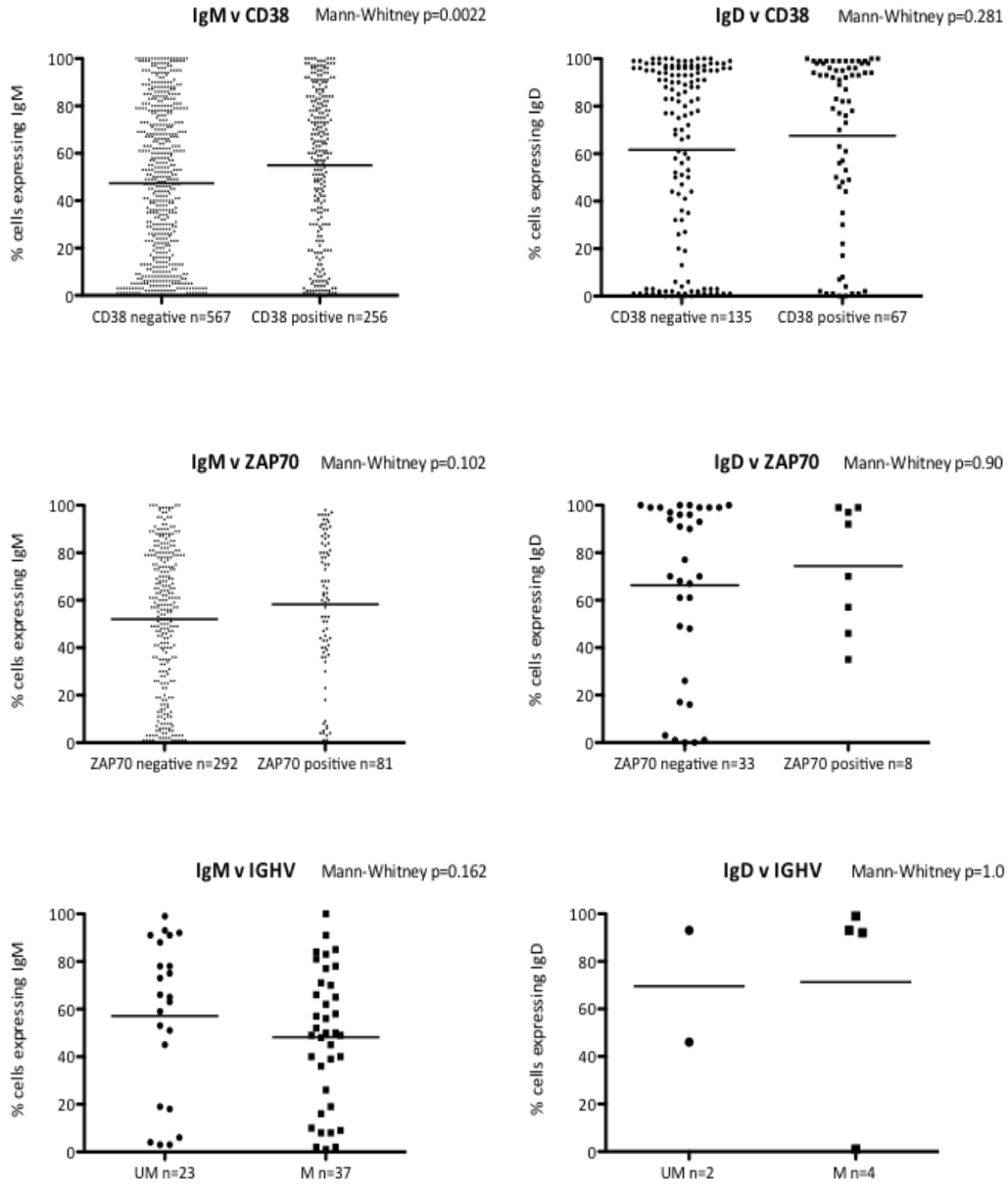
### *5.2.14 Isotype expression and prognosis*

The association of Ig expression as a continuous variable was compared with known prognostic factors (see Fig. 5.13). The mean level of IgM was slightly higher in CD38 positive cases, but no other associations were noted

When these data are considered as categorical variables, the association of CD38 positivity with IgM positivity is also significant when 30% threshold for positivity is used (Fisher's exact test, p=0.0057), but not at the 10% threshold. No other marker (CD38, ZAP70 or *IGHV* status) is associated with IgM positivity at 10% threshold, or with IgD positivity by either the 10% or 30% thresholds.

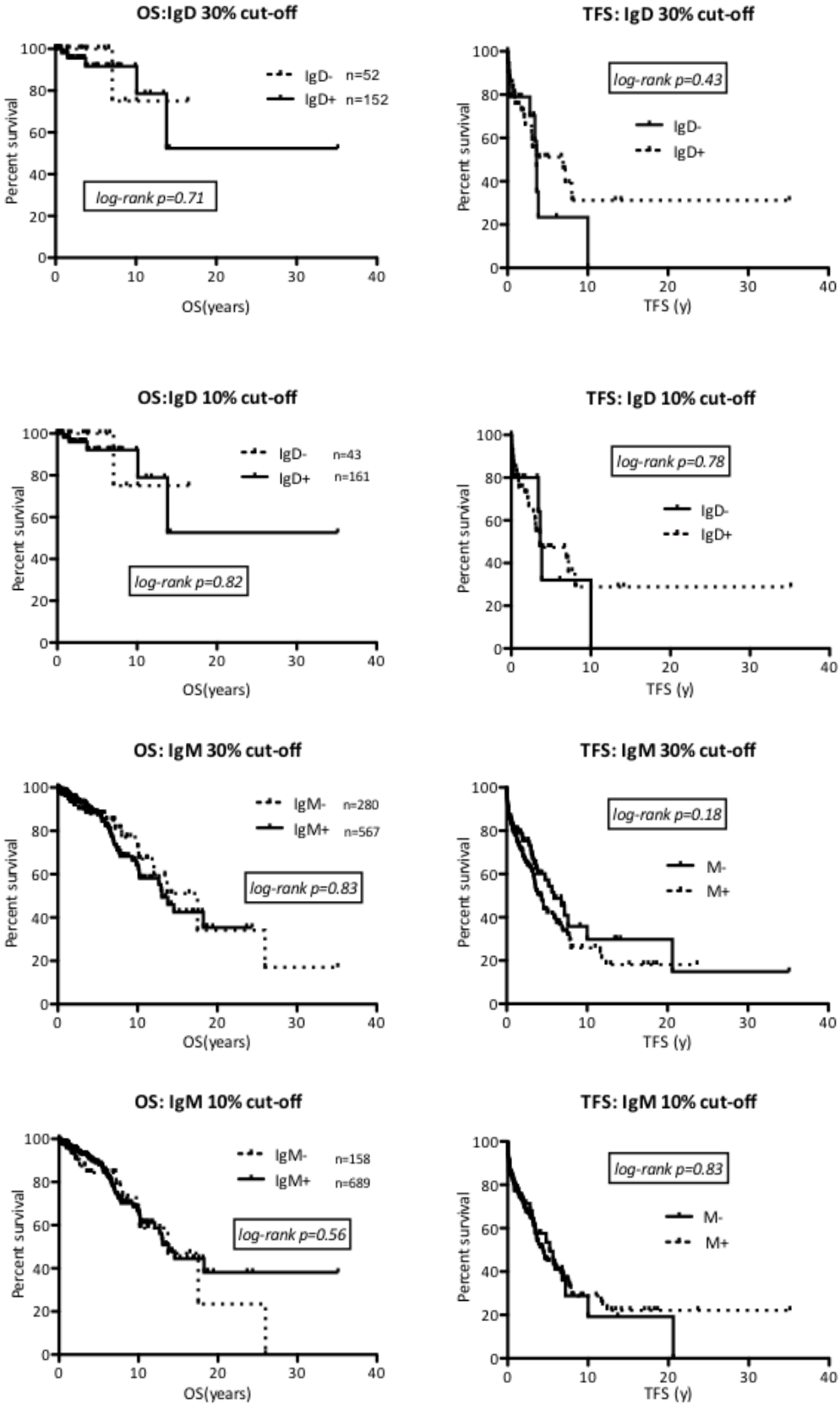
Simultaneous cytogenetic and Ig data were available for 238 IgM and 52 IgD patients. There is a significant association between deletion of 11q and IgM positivity at 10% threshold (p=0.022), but not at 30% threshold (p=0.19). Similarly, trisomy 12 is associated with IgM negativity at 10% threshold (p=0.005) but not at 30% threshold (p=0.06). There is no significant association between del 17p or del13q and IgM positivity, and no associations between IgD positivity and any cytogenetic group.

**Figure 5.13: Expression levels of IgM and IgD in various groups defined by published prognostic markers.**
**Horizontal bars represent the mean.**




The effect of Ig isotype on prognosis was determined. All 204 patients with IgD data had survival data available. There appears to be no effect of Ig isotype on overall survival (see below) or treatment-free survival (Fig. 5.14)

**Figure 5.14: Influence of IgD and IgM expression as a categorical variable on prognosis as judged by Overall Survival (OS) and Treatment Free survival (TFS)**



When the IgD and IgM levels are used to divide patients into $IgM^-IgD^-$, $IgM^+IgD^-$, $IgM^-IgD^+$ and $IgM^+IgD^+$ subgroups, there is no significant influence on survival at either 10% or 30% thresholds. When $IgM^-IgD^-$ patients (12-19% of total) are compared to patients with at least one isotype expressed, there is no difference in OS or TFS at either threshold for positivity.

### *5.2.15 Summary and Discussion: Peripheral Blood IgD and IgM expression in CLL clinical database*

Detection of monoclonal surface Ig (sIg) is a prerequisite for CLL diagnosis. Clonality is ascertained by light chain restriction, implying that sIg is present in all cases. Classically, IgM and IgD are expressed on CLL cells, often weakly/dimly. We have examined a large series (n=1182) of patients seen at Barts Cancer Centre. Data on IgM, IgD, other biological characteristics and prognosis is not available for all patients, but some conclusions may be drawn. The database appears reasonably comprehensive in terms of coverage of all new patients within the local network, and clinical characteristics such as prognostic markers appear to be consistent with published series. It is derived from the patients seen in a large tertiary centre (as are many published series), so may not be representative of the general population.

204/1182 patients had IgD and IgM expression data available. The majority had expression levels ascertained close to the diagnostic date. A sizeable minority seemed to exhibit variation of expression over time as judged by a change in positive/negative status. Follow-up is also short, and altogether these factors weaken any conclusions that might be drawn about the use of IgD and IgM as a prognostic factor. Survival analysis suggests that the level of expression of IgD and IgM had no effect on overall survival or treatment free survival. This data on the lack of prognostic impact of IgD and IgM expression on PB CLL cells is in general agreement with several other studies (see 5.2.1). This has two implications.

Firstly, PB IgM and IgD levels cannot be used as a biomarker to predict prognosis. Over the past two decades, there have been a large number of publications providing

evidence for one factor or another having some prognostic impact in CLL[16]. Of the hundreds of prognostic markers, only disease stage and *P53* gene abnormalities have shown any real use as a biomarker/prognostic factor to guide therapy in a clinical context, though prospective trials are attempting to incorporate factors such as *IGHV* gene mutation status. The aim of this analysis is not to add to the list of potential biomarkers but to gain insight into CLL biology.

This leads to the second implication: Expression level of the BCR in peripheral blood CLL cells has little influence on disease biology. Many of the prognostic factors in CLL have been investigated in depth in an attempt to gain greater insights into tumourigenesis, with varying success. The finding that expression level of a molecule conveys prognostic impact does not automatically imply its importance in disease biology, though it is generally taken as supportive evidence for such a role. Also, the lack of prognostic impact of expression level does not automatically imply lack of effect of that molecule on disease biology. Many authors have provided evidence that suggests the centrality of the BCR to CLL pathogenesis. Since it is felt that proliferation of the CLL clone occurs in the microenvironment of the lymph node and bone marrow, it may be that peripheral blood CLL cells do not reflect events in the microenvironment, which may also account for the lack of prognostic impact of PB BCR expression levels.

Despite this, several conclusions can be drawn from our data. The majority of patients express a BCR isotype of IgD/IgM specificity on CLL cells at some level. It may be that a more sensitive assay would detect lower expression levels. We have some data that IgM (and to an extent IgD) expression changes over time within a patient. This may be due to variation of assays, or may be an interesting effect of biology. More relevantly, we suggest IgD expression is generally more common than IgM expression, with dual $IgM^{+}IgD^{+}$ cells forming the biggest group. This implies that any studies that seek to examine the role of the BCR in CLL cannot ignore IgD. Most studies have looked at the effect of signalling via IgM. Accompanying a reduced IgM expression, there is a lack of correlation between IgM and IgD expression within patients, suggesting that IgM and IgD expression levels are regulated independently, which is puzzling if their function is

identical, as many authors imply. Furthermore, the clonal nature of the BCR requires that both IgD and IgM bind the same antigen, so independent regulation may imply differing responses to antigen. Conversely, this differential regulation may be taken as evidence against ongoing antigen encounter, at least in PB cells.

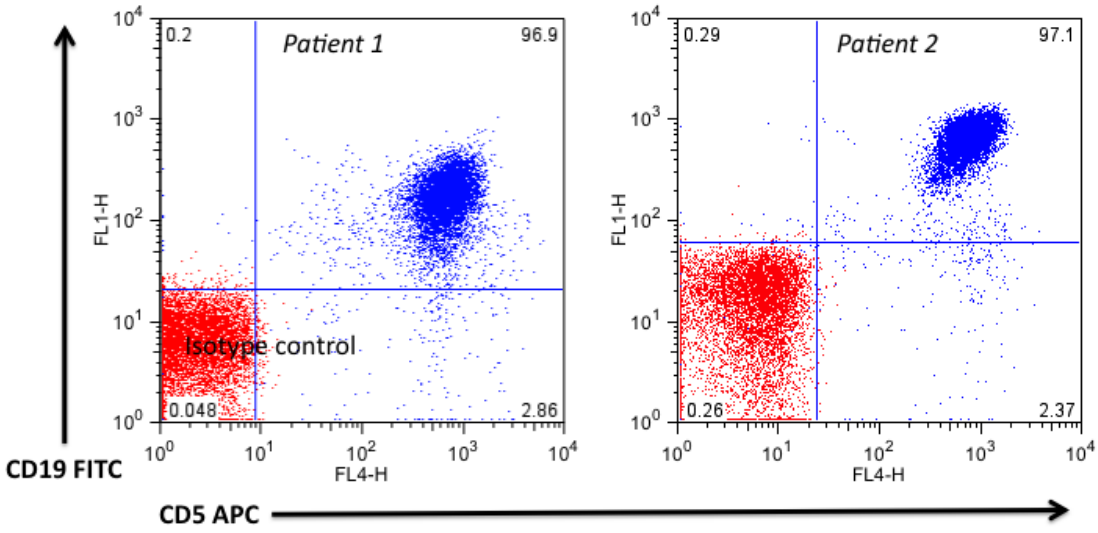
We use these and published data to guide the further examination of IgD and IgM expression and signalling *in vitro.*

## 5.3 IgM and IgD expression in peripheral blood: mononuclear fractions from tissue bank

### *5.3.1: Methods*

IgD and IgM expression was assessed on thawed stored CLL samples that had been separated by density-gradient centrifugation and frozen. Patient samples were selected based on high WBC, untreated and to reflect a range of recognised prognostic factors. Each sample was assessed by CD5/CD19 dual staining to assess purity. Generally, studies were only performed on samples from untreated patients with >90% CLL B-cells as assessed by CD5 and CD19 dual staining. Example flow cytometric plots are shown below (Fig 5.15).

**Figure 5.15: Example cytometric plots from assays to assess $CD19^{+}CD5^{+}$ dual positive cells as a marker of CLL B-cell purity.**
**In red are plots from isotype controls, with quadrants drawn at 99$^{th}$ centile for these. In blue are the dual stained samples. Each quadrant contains percentages of dual positive cells present in each quadrant.**



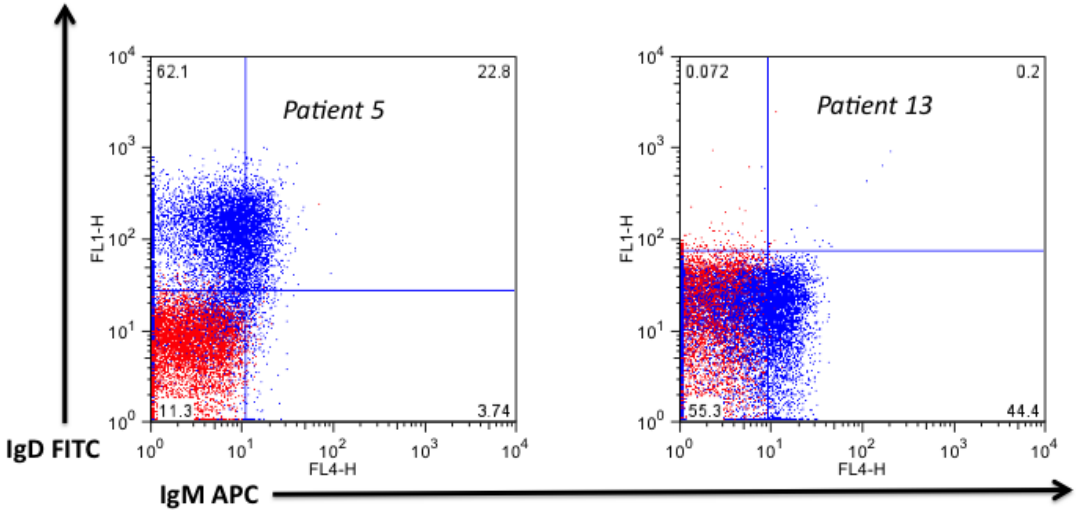

20 samples were studied extensively *in vitro*, and their characteristics are below. (Table 5.9)

**Table 5.9: Characteristics of patients.**
**UPN: Unique patient number assigned for the purpose of experiments. P signifies peripheral blood source.**
***IGHV*: Immunoglobulin *IGHV* gene mutation status (M=mutated, <98% homology to baseline, U=unmutated, ≥98% homology to baseline)**
**Cytogenetics: FISH panel for CLL performed**
**CD38: peripheral blood flow cytometric expression (1=>30%, 0=<30%)**
**ZAP70: peripheral blood flow cytometric expression (1=>5%,0=<5%)**
**$CD19^+CD5^+$: percentage of mononuclear cells dual positive for CD5/CD19 by flow cytometry**
**X: assay not performed or failed.**

| UPN | *IGHV* | cytogenetics | CD38 | ZAP70 | $CD19^+CD5^+$ |
|---|---|---|---|---|---|
| P1 | M | del13q | 0 | 0 | 96.9 |
| P2 | X | del13q | 0 | 0 | 97.1 |
| P3 | U | del13q | 1 | X | 87.9 |
| P4 | X | del13q | 0 | 0 | 96.5 |
| P5 | X | X | 0 | 0 | 94.2 |
| P6 | M | del13q | 0 | 0 | 91.8 |
| P7 | U | del13q del11q | 0 | 0 | 95.9 |
| P8 | M | del13q | 0 | 0 | 98.2 |
| P9 | X | del13q | 1 | 1 | 95.6 |
| P10 | X | +12 del17p | 1 | 1 | 90.6 |
| P11 | X | normal | 1 | 1 | 96.3 |
| P12 | X | X | X | X | 96.7 |
| P13 | X | del13q | 0 | 0 | 98.4 |
| P14 | X | del13q | 1 | 1 | 96.5 |
| P15 | M | del13q | 0 | 0 | 96.8 |
| P16 | U | del17p | 1 | X | 93.2 |
| P17 | M | del13q | 1 | 0 | 91.7 |
| P18 | M | del13q | 0 | X | 92.9 |
| P19 | M | del13q | 0 | 0 | 93.5 |
| P20 | U | del13q | 1 | X | 91.4 |

IgM and IgD dual staining was performed. The expression levels of IgD and IgM did not separate clear populations from isotype controls, and semi-quantification was performed based on quadrants drawn for isotype controls at the $99^{th}$ centile of fluorescence intensity (see Fig. 5.16). A percentage of cells positive for IgM and IgD is derived using this method, but as the fluorescence intensity is a product of the proportion of cells expressing the immunoglobulin and the expression by the individual cells, this percentage cannot necessarily be assumed to be the percentage of cells that are genuinely 'positive'.

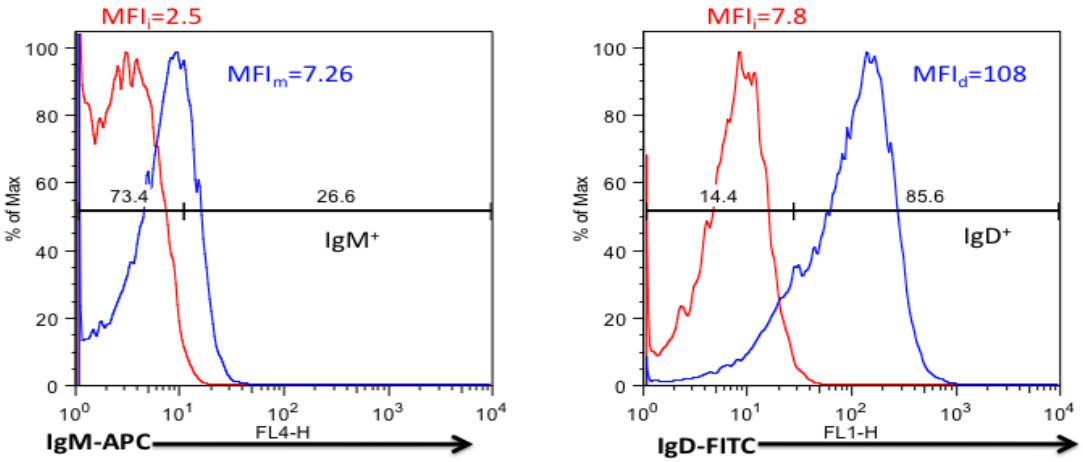
**Figure 5.16: Example flow cytometric plots of IgD positive and IgM positive cases.**




An alternative analysis method would be report the Median Fluorescence Intensity (MFI) difference between isotype control and IgD/IgM staining (sometimes referred to as Specific MFI). Shown below (Fig 5.17) are the frequency histograms for patients 5 and 13.

To provide a specific MFI, the $MFI_{d/m}$ is divided by the $MFI_i$. The results are shown below as compared with the percentage above isotype threshold method.
(Fig. 5.18).

**Figure 5.17: Histograms showing staining for IgM/IgD as compared to isotype control at 99$^{th}$ centile.**
**Shown are the percentages of cells positive above the isotype control, and the Median Fluorescence Intensity (MFI) for isotype ($MFI_i$) and stained ($MFI_{m/d}$)**
**Patient 5**




**Patient 13**

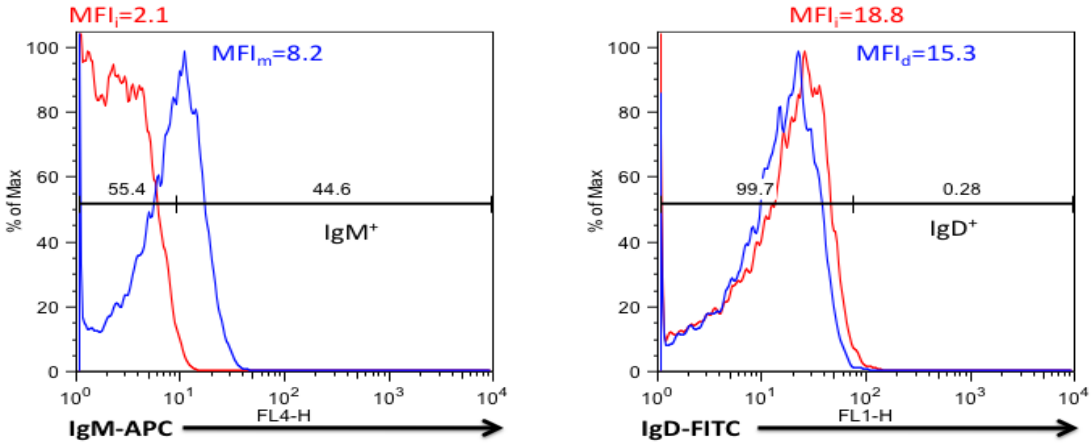
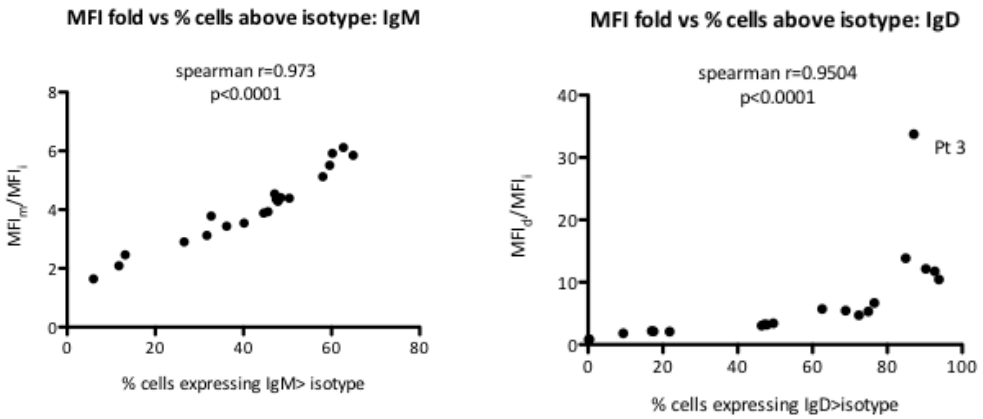


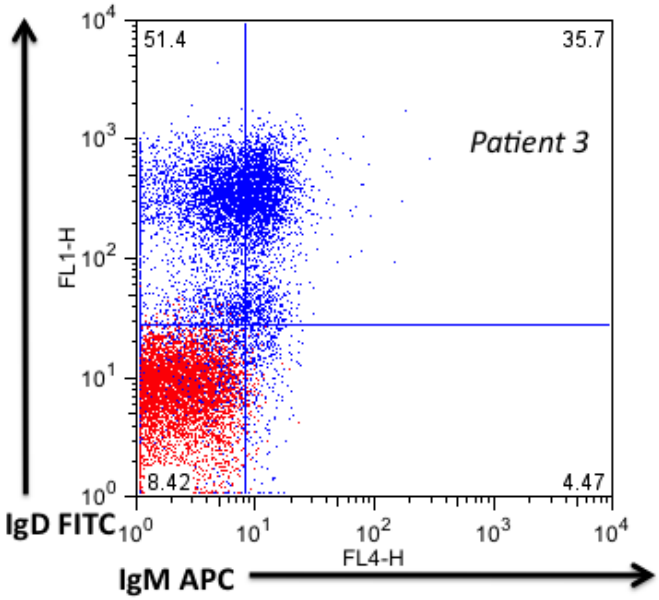

**Figure 5.18: Comparison of specific Median Fluorescence Intensity method to percentage of cells above 99$^{th}$ centile of isotype control.**



It is seen that there is a clear correlation between these two measures. In general, the *in vitro* quantifications by the specific MFI fold change and percentage above isotype are highly correlated with each other, and there seems little basis to suggest choosing one measure rather than another, so we have henceforth used percentage cells above isotype control. Because most comparisons using these data will rely on non-parametric statistics that consider rank comparisons rather than the quantifications, an absolute quantification of Ig level may not be important for many subsequent *in vitro* assays.

There is also a clear outlier in patient 3, with a much higher IgD MFI as compared to the percentage of cells above the isotype threshold. The reason for this is shown below (Fig. 5.19). There are two populations of $IgM^{+}IgD^{+}$ cells, and the larger population has a clearly increased IgD-related MFI, with some with MFI $>10^{3}$, which is unusual when compared to the other patients. It may be that this CLL aberrantly expresses high levels of IgD (which would be described as 'bright' in some publications), and this will need to be taken into consideration.

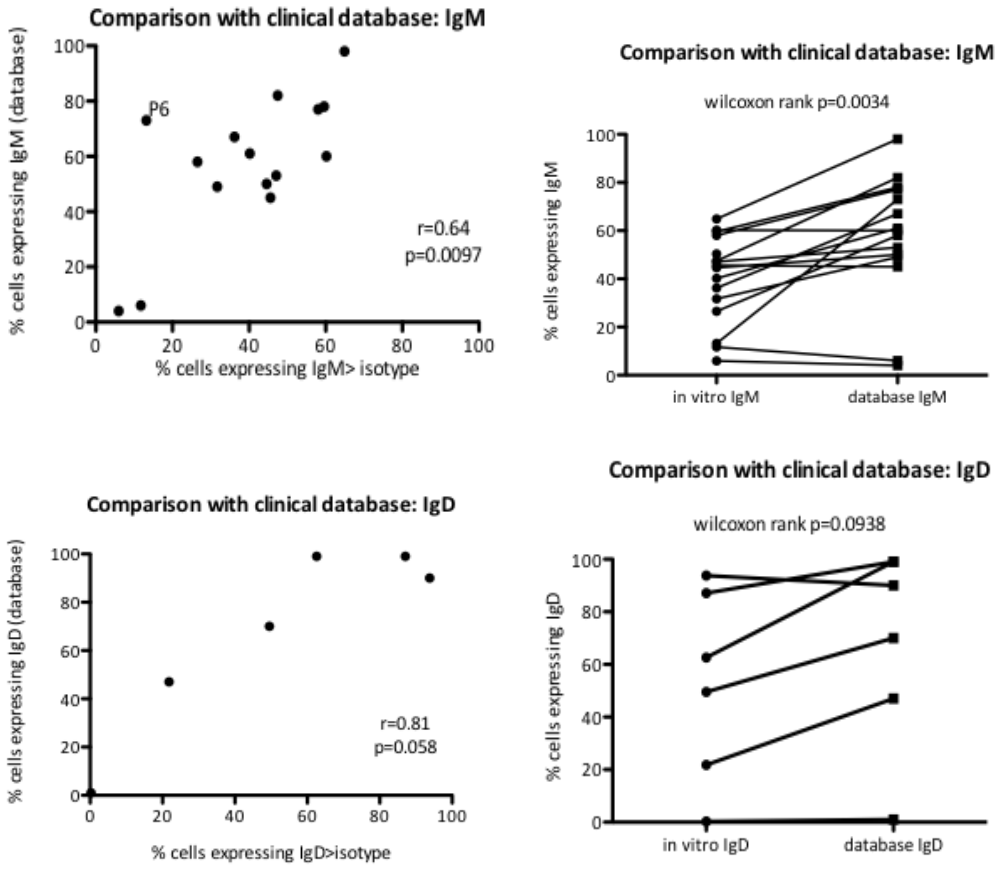
**Figure 5.19: Flow cytometric plots of IgD and IgM expression for patient 3**




In an attempt to validate the method further, the percentage> isotype method was compared with the expression levels from the clinical database. There was limited overlap, with 15/20 having IgM expression levels measured by both techniques, but only 6/20 for IgD. The time between samples used for clinical immunophenotyping and

samples taken for freezing in the tissue bank ranged from 0.5-8.7 years (mean 3.6 years). There was generally an agreement between the percentage positive cells measured via the *in vitro* method and the results from the database, as assessed by Spearman correlation, however the *in vitro* method appears to provide lower values than that in the database (particularly for IgM). The comparisons are shown below (Fig 5.20):

**Figure 5.20: Comparison of IgM and IgD expression levels by flow cytometry by two methods.**
**Expression in database (percentage positive cells as compared to negative internal control lymphocytes) *vs* expression as compared to isotype control *in vitro*. The *in vitro* method of determining IgM levels produced lower values when compared to database values, and this significant difference in means is maintained when outlier P6 is removed from the analysis.**



### *5.3.2: Results: IgM and IgD expression in vitro*

Within this group of 20 patients for which frozen samples were used, there was no correlation between levels of expression of IgM and IgD, and no significant difference in expression (figs 5.21 & 5,22):

**Figure 5.21: Comparison of IgD and IgM expression by flow cytometry for thawed samples.**

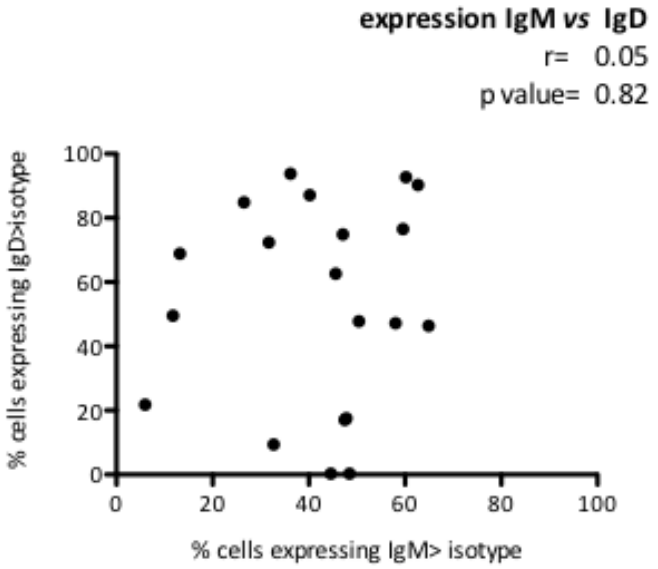
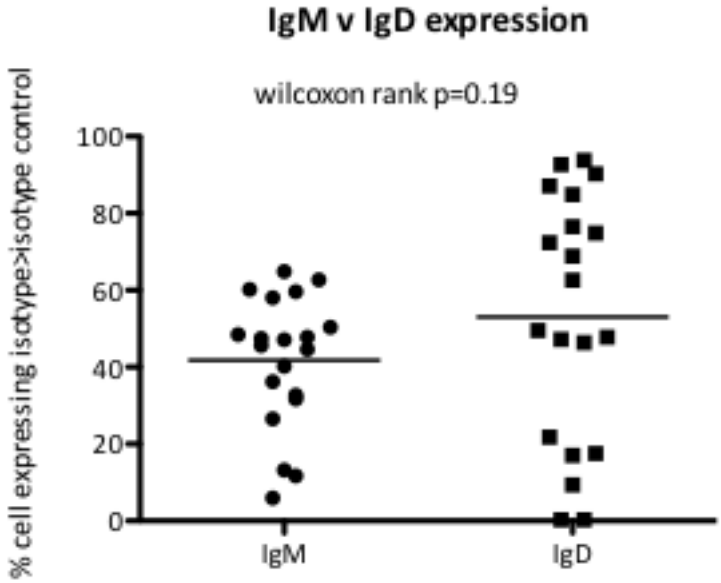



**Figure 5.22: Comparison of IgM and IgD expression in thawed samples. Horizontal bars represent means.**




There was no pattern of association with *IGHV* mutation status, CD38 expression or ZAP70 expression (data not shown).

### *5.3.3 Summary and Discussion: IgM and IgD expression in peripheral blood: mononuclear fractions from tissue bank*

To summarise, 20 samples from untreated patients were selected based on high tumour proportion as judged by $CD5^{+}CD19^{+}$ fraction. Levels of IgD and IgM expression were determined by flow cytometry and quantified based on comparison to isotype controls. This method had reasonable correlation with flow cytometric determination of IgD/IgM staining as assessed by the clinical immunophenotyping laboratory using a somewhat different method. However, there seemed to be a consistently lower estimation of the level of IgM positive cells with the *in vitro* method as compared with levels on the database. Various explanations may be offered for this inconsistency. There is a large time difference between the dates of samples taken from patients (up to 8.7 years), and so if IgM expression changed over time a difference may be detected. However, there is still a correlation of expression levels using the two techniques even with this time difference, and thus a change in IgM level over time would require a consistent fall in IgM expression over time if this explanation were correct.

A more likely explanation is that the *in vitro* method (comparison to isotype control) is not directly comparable to the methods used by the clinical laboratory (comparison to internal negative lymphocytes), and consistently underestimates IgM expression. The freezing process may also have an influence. Whilst percentage positive greater than isotype control may be presumed to correlate with the number of CLL cells that are genuinely positive for that immunoglobulin, it does not necessarily follow that this percentage measure is equivalent. Furthermore, the number of molecules per cell may be higher or lower (reflected in 'bright' and 'weak' expression on individual cells), which will be reflected by our method that does not separate a discrete population of cells from isotype control. What might reasonably be assumed is that the *in vitro* method used provides an estimate that reflects both the number of cells positive for an immunoglobulin isotype and the expression level of that isotype on individual cells. It could therefore be argued that one patient has 'more' IgD or IgM than another, when the same method is used.

These 20 characterised samples were used for further *in vitro* studies of the effect of ligating IgD or IgM (see Chapter 6).

## 5.4 Bone marrow and lymph node IgD and IgM expression

### *5.4.1 Introduction*

As discussed, the microenvironment of many tumours has an influence on malignancy, and there is evidence that this is particularly important in CLL. Whilst the peripheral blood compartment exhibits the most obvious (and conveniently accessible to researchers) involvement with leukaemia, it is generally thought that the proliferation of CLL cells occurs within the secondary lymphoid tissues: The lymph nodes, spleen and bone marrow. Collectively, these form the CLL microenvironment[468,469]. The historical view of CLL was as an accumulative disease whereby long-lived malignant B-cells proliferate and die slowly, however seminal studies using heavy water labelling have shown that the proliferative rates of CLL cells are higher than was initially appreciated and that a substantial portion of the malignant clone (0.1-1%) divides daily[470]. This proliferation is higher in the lymph nodes than the bone marrow.

One view of CLL cells in the peripheral blood may simply is as a 'spill over' from the more important lymph node/bone marrow compartments where proliferation and microenvironmental interactions occur[469]. This does not necessarily imply that peripheral blood cells do not re-enter these compartments and proliferate, as happens with healthy B-cells. Many of the more recent publications involving *in vitro* studies of CLL cells have attempted to provide various ligands that simulate microenvironmental stimuli, including chemokines such as CXCl12, cell adhesion molecules such as VLA-4, T-cell signals such as CD40, other factors such as BAFF, as well as cellular elements such as nurse-like cells and other stromal cells[22]. Generally, these studies have managed to prolong CLL cell survival in culture, and mediate resistance to chemotherapeutic agents, but inducing proliferation of primary CLL cells *in vitro* has proved more challenging. Various publications have compared CLL cells found in the bone marrow and lymph node microenvironment to paired peripheral blood samples from the same patient. CD38[471] and Zap70 expression[54] and an anti-apoptotic balance

of BCL2-family member expression[472] is known to be higher in the lymph nodes of patients as compared to paired PB samples. Evidence of NF-kB signalling and of BCR signalling in the microenvironment is also becoming apparent[53].

Early studies highlighted the existence of pseudofollicles/proliferation centres in SLL, with perhaps 40% of SLL lymph nodes having this architectural pattern, the remaining cases displaying a diffuse infiltrate[473]. Pseudofollicles are generally defined as aggregates of lymphoid cells, which are larger and paler staining than surrounding lymphocytes. They often have more abundant cytoplasm, with most having nuclei morphologically similar to surrounding lymphocytes. However, the pseudofollicles are also enriched for prolymphocytes with larger nuclei, often Ki67 positive, providing the pseudofollicle with its alternative name of proliferation centre[51]. The size and abundance of these pseudofollicles has been associated with prognosis by some authors, as has Ki67 positivity, though publications are inconsistent about the relationship between Ki67 expression and prognosis[51,474-476]. One group have data suggesting that the cells within pseudofollicles are enriched for cytogenetic abnormalities as compared to surrounding non-pseudofollicle cells[477].

Peripheral blood CLL cells generally express Ki67 at a low level (generally <2%)[468], and proliferation is presumed to occur within the LN and BM compartments (as with normal lymphocytes) where Ki67 levels are of the order of 2-30%[53,474]. The bone marrow in CLL is known to demonstrate a pseudofollicular pattern in some instances, but generally, this is only in later stage disease[478]. Studies using heavy water to examine CLL cell proliferation suggest that bone marrow is perhaps not a major proliferative compartment in CLL patients with stable disease, as compared to lymph nodes [479].

Pseudofollicles are also associated with markers of activation of the NF-kB pathway[469,480]. A recent paper examined the gene expression profiles and selected protein phosphorylation (by western blot and phospho-flow cytometry) of paired peripheral blood (PB), bone marrow (BM) and lymph node (LN) CLL cells[53]. Comparing gene expression of BM, LN and PB, they suggest that greater differences are found

between lymph nodes and the other two compartments. Gene Set Enrichment analysis of the lymph node upregulated genes suggested higher expression of genes involved in BCR signalling and NF-kB signalling pathways. Stimulation of PB CLL cells with anti-IgM *in vitro* upregulated similar BCR pathway genes, and this was more apparent in U-CLL. Anti-IgM stimulation also caused phosphorylation of the proximal BCR signalling pathway kinase SYK, and increased SYK phosphorylation was noted in LN cells as compared to PB. They suggest that the differences observed were supportive evidence of BCR signalling within the lymph node (LN) microenvironment. Whilst bone marrow (BM) did exhibit similar changes, the differences between PB and BM were less than that between PB and LN, leading them to suggest that the LN is a more critical compartment in terms of CLL proliferation[53]. If antigen is stimulating CLL survival and/or proliferation, this is likely to occur within the lymph node microenvironment rather than the peripheral blood, recapitulating the normal immune response of B-cells to antigen.

There are few data examining sIg expression in LN compartments. One recent study of 72 cases identified many markers that distinguished pseudofollicles from interfollicular areas, but state that IgM and IgD do not differentiate between these areas[481]. Anecdotally, our histopathologists report occasional differences in IgM and IgD expression between pseudofollicle and interfollicular areas. No studies have compared peripheral blood to lymph node or bone marrow compartments in terms of Ig expression.

### *5.4.2 Methods: Bone marrow CLL immunophenotyping*

In a limited number of cases (n=6), paired bone marrow and peripheral blood samples had flow cytometry performed whereby IgD and IgM expression could be compared. Flow cytometric methods were the same as for peripheral blood studies, and were performed by the clinical flow cytometry laboratory. Only 1 of 6 patients had the BM and PB samples taken on the same day. The period between the BM and PB analysis ranged up to 20 months. In general, PB immunophenotyping was performed prior to BM, which was generally performed when therapy was planned.

PB and BM IgD was available for 4 of 6 patients. Table 5.10 summarises their characteristics.

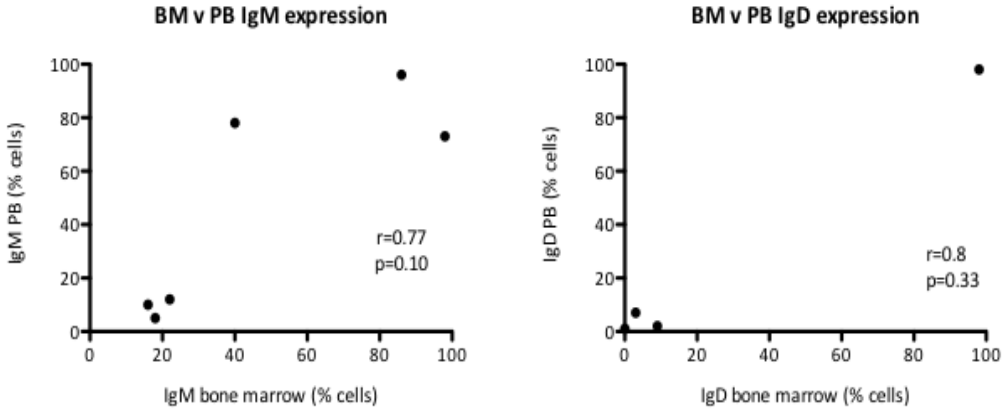
**Table 5.10: Characteristics of paired bone marrow v peripheral blood samples**

| Patient | Date of PB | Date of BM | Period elapsed between BM & PB sample (days) | IgD% BM | IgD% PB | IgM% BM | IgM% PB |
|---|---|---|---|---|---|---|---|
| **BM1** | 29/3/2010 | 29/3/2010 | 0 | 0 | 1 | 22 | 12 |
| **BM2** | 29/1/2010 | 6/4/2010 | 67 | 3 | 7 | 40 | 78 |
| **BM3** | 24/5/2009 | 1/7/2009 | 38 | 9 | 2 | 16 | 10 |
| **BM4** | 30/7/2008 | 31/3/2010 | 609 | 98 | 98 | 86 | 96 |
| **BM5** | 12/9/2007 | 30/8/2007 | -13 | ? | ? | 18 | 5 |
| **BM6** | 10/8/2009 | 22/12/2008 | -231 | ? | 78 | 98 | 73 |

### *5.4.3 Results: Bone marrow CLL immunophenotyping*

Fig. 5.23 compares expression levels in paired samples.

**Figure 5.23: Comparison of IgD and IgM expression in paired bone marrow and peripheral blood (PB) samples.**
**Wilcoxon rank tests suggest no difference of means between PB and BM for both IgD and IgM.**




As is seen, where PB IgM/IgD expression is low/high, BM expression is also low/high. Sample size is small which presumably results in lack of significant p-values for this correlation, and limited generalisability. However, it could be claimed that the correlation of the limited number is consistent with BM and PB IgM/IgD expression being similar.

### *5.4.4 Methods: SLL lymph node immunohistochemistry*

As part of the tissue bank of Barts Cancer Institute, patients are invited to consent to storage and use of tissue taken for diagnostic or therapeutic purposes. 240 SLL/CLL patients had available paraffin-embedded tissue blocks from lymph node biopsies/splenectomy biopsies. There were 204 CLL patients with PB IgD and IgM data. There were only 19 with both, and 10 lymph node/spleen tissue blocks were readily identified. All sections were stained with Haematoxylin-Eosin (H&E), Ki67, IgM and IgD. Sections of reactive tonsils removed from children/young adults were used as positive controls for IgD and IgM staining. Further details are available in Chapter 3. Images were captured using the 3D HISTECH scanner microscope and Panoramic Viewer software. Cells were counted manually using the Panoramic Viewer Marker Counter function.

The characteristics of the 10 patients with lymph node biopsies/sections of spleen and peripheral blood immunophenotyping for both IgD and IgM are shown in Table 5.11.

**Table 5.11: Characteristics of paired peripheral blood (PB) and lymph node (LN) samples from patients with SLL/CLL.**
**The date of PB is the date at which IgD and IgM analysis was available. Shaded = treated prior to biopsy**

| Patient | Date of PB | Date of LN/ splenectomy | Period elapsed between LN & PB sample (days) | Treated? | *IGHV* | Cytogenetics | CD38 |
|---|---|---|---|---|---|---|---|
| L1 | 24/11/2008 | 31/10/2008 | -24 | 01/08/2008 | X | normal | 0 |
| L2 | 21/01/2009 | 03/06/2009 | 133 | 01/02/2007 | M | del13q del17p | 1 |
| L3 | 12/11/2008 | 12/03/2008 | -245 | 30/01/2006 | X | del13q | 0 |
| L4 | 21/01/2009 | 06/10/2010 | 623 | 01/11/2008 | X | del17p +12 | 0 |
| L5 | 22/01/2009 | 22/09/2010 | 608 | 01/05/2010 | X | X | 0 |
| L6 | 18/03/2009 | 25/11/2009 | 252 | 01/01/2005 | X | del11q del17p | 1 |
| L7 | 10/11/2008 | 19/01/2009 | 70 | 05/09/2003 | X | del11q | 1 |
| L8 | 28/01/2009 | 26/01/2009 | -2 | 04/12/2010 | X | X | 0 |
| L9 | 17/03/2010 | 20/08/2008 | -574 | 01/10/2007 | X | del13q del17p | 0 |
| L10 | 10/03/2010 | 12/05/2010 | 63 | 01/05/2007 | X | del13q | 0 |

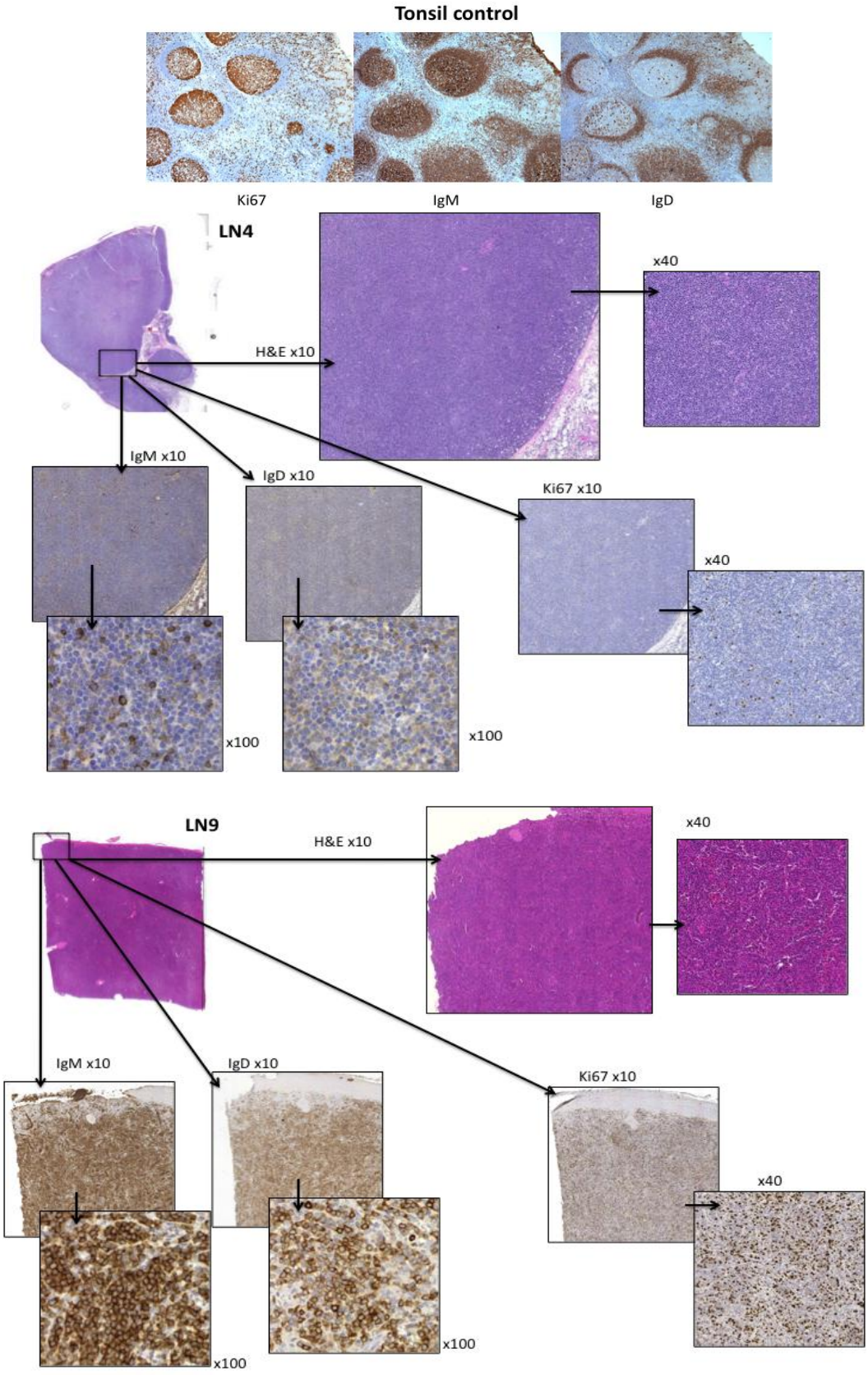
Paraffin-embedded sections were taken and stained with H&E (for morphology), CD20, IgD, IgM and Ki67 (see Fig. 5.24). Pseudofollicles could be clearly identified by H&E staining in 3 of 10 sections (L2, L3, L10). In the other sections, the pattern of infiltration was either diffuse or suggestive of coalescent pseudofollicles/somewhat heterogeneous. CD20 staining was almost universally diffuse and uniform, but LN2 and

LN10 both had the suggestion of greater CD20 staining within pseudofollicles as compared with interfollicular areas, though the majority of cells within interfollicular areas were also CD20 positive.

Ki67 was strongly positive within the pseudofollicles of LN10, but other samples were diffusely positive, albeit with some heterogeneity in pattern. Whilst IgM and IgD staining was more apparent within the pseudofollicles of LN10, this was not true of the other samples with prominent pseudofollicles (LN2 and LN3).

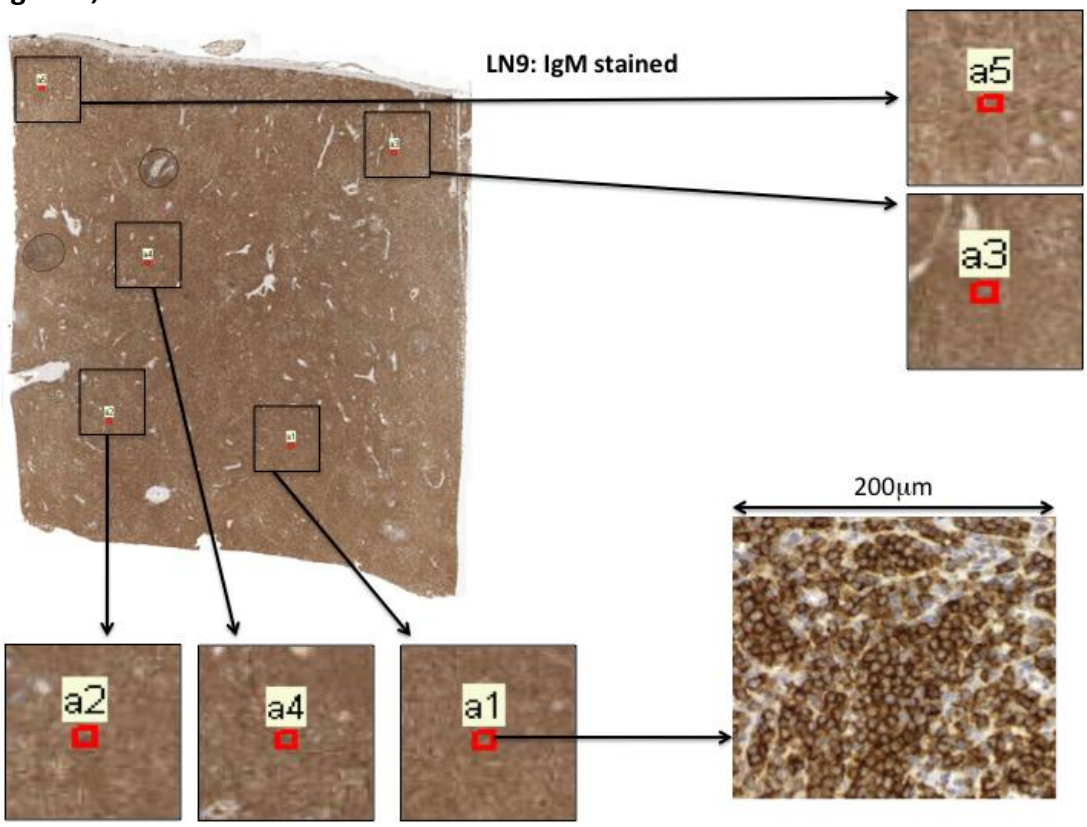
It is not possible to definitively answer the question as to differential IgD/IgM expression between pseudofollicle and interfollicular areas given these samples. Pseudofollicles were discernible in only 3 samples, the CD20 staining within these samples was somewhat heterogeneous, and IgD and IgM differential pseudofollicular staining was only present in one sample. Therefore, it is not possible to state that IgD and IgM staining differed between pseudofollicles, in fact the data suggest otherwise, with 2 of 3 samples exhibiting pseudofollicular pattern having no difference between the two compartments.

**Figure 5.24: Representative examples of H&E, Ki67, IgD, IgM staining.**
**Tonsil staining patterns for Ki67, IgM and IgD are shown for comparison. LN4 had a diffuse infiltrate with low levels of Ki67, IgD and IgM. LN9 had a widespread yet somewhat heterogeneous infiltrate that was strongly positive for all 3 markers.**



An attempt was made to quantify the levels of expression for IgM, IgD and Ki67. For each section, five high power fields (HPF at x100 magnification, 200μmx200μm square) were selected from each section. These were evenly spaced and contained 400-800 cells per HPF. Illustration of the method is shown below in Fig. 5.25

**Figure 5.25: Selection of 200μmx200μm for cell counting**
**Example is of IgM stained LN9. All cells within each area are counted as positive or negative, and the scores collated.**

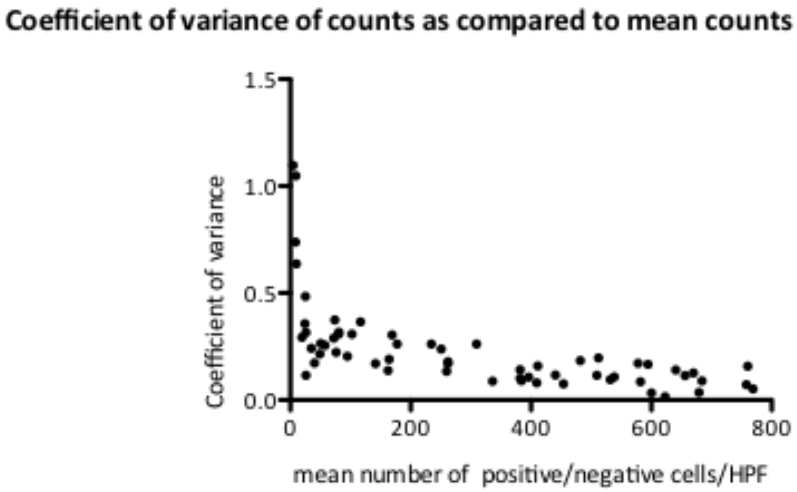



Where pseudofollicles were prominent, we included pseudofollicular and interfollicular areas in the same HPF. All positive and negative stained cells were counted within each field.

It is presumed that the more cells are counted, the greater the accuracy in estimation of the true proportion of cells positive or negative for a particular marker. Five HPFs were chosen as a pragmatic compromise between accuracy and time taken to manually count cells. In an attempt to estimate the variability, the coefficient of variation (CV, standard deviation divided by mean) was calculated for each of the 5 HPF counts. The mean CV was 0.23 (23%). When CV is plotted against mean count per

sample, it can be seen that the high CVs (and variability between counts in each HPF) are all in samples with low proportion of positive/ negative cells, as one might expect. The majority of CVs are <50% (Figure 5.26).

**Figure 5.26: Coefficient of variation of counts from 5 high power fields (HPFs) as compared to mean counts of positive/negative cells.**
**Values for IgM/IgD/Ki67 positive/negative cell counts combined. The high CVs at low mean cells/HPF is due to variability being enhanced by low numbers of positive/negative cells.**

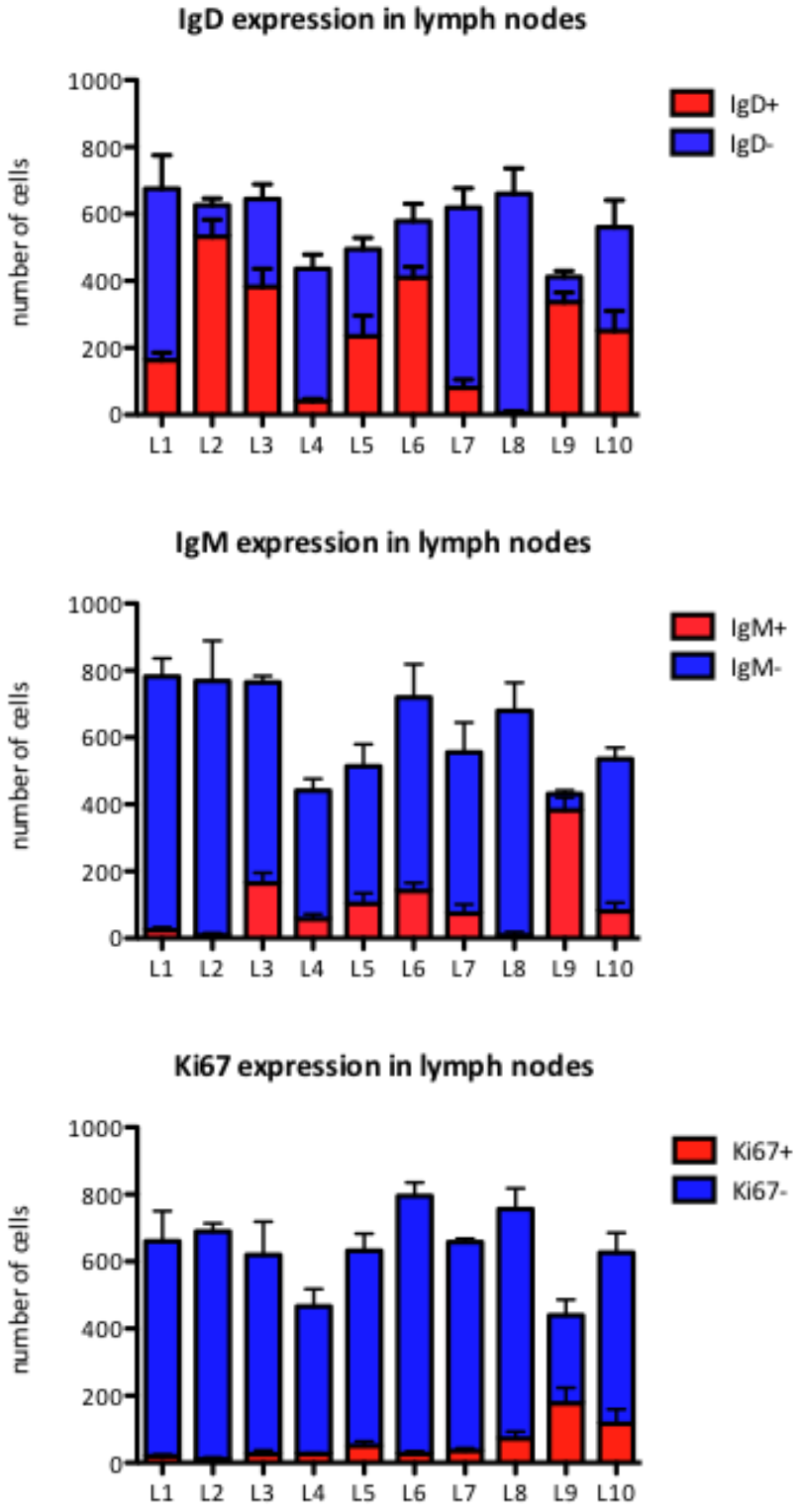



Further validation of the method was provided by a second person with experience in such methods (Dr Rita Coutinho) performing a manual count of the same areas for a limited number of HPFs (LN4, LN, LN10). Comparison of the percentages positive for each of these 15 validation areas was made. Wilcoxon signed rank p-values were all >0.05, indicating that there was no significant difference in median proportion of Ki67, IgD or IgM cells between the two manual counts, and Spearman correlation coefficients were 0.86-0.99 ($p<0.001$ for all correlations) for each comparison.

### *5.4.5 Results: SLL lymph node immunohistochemistry*

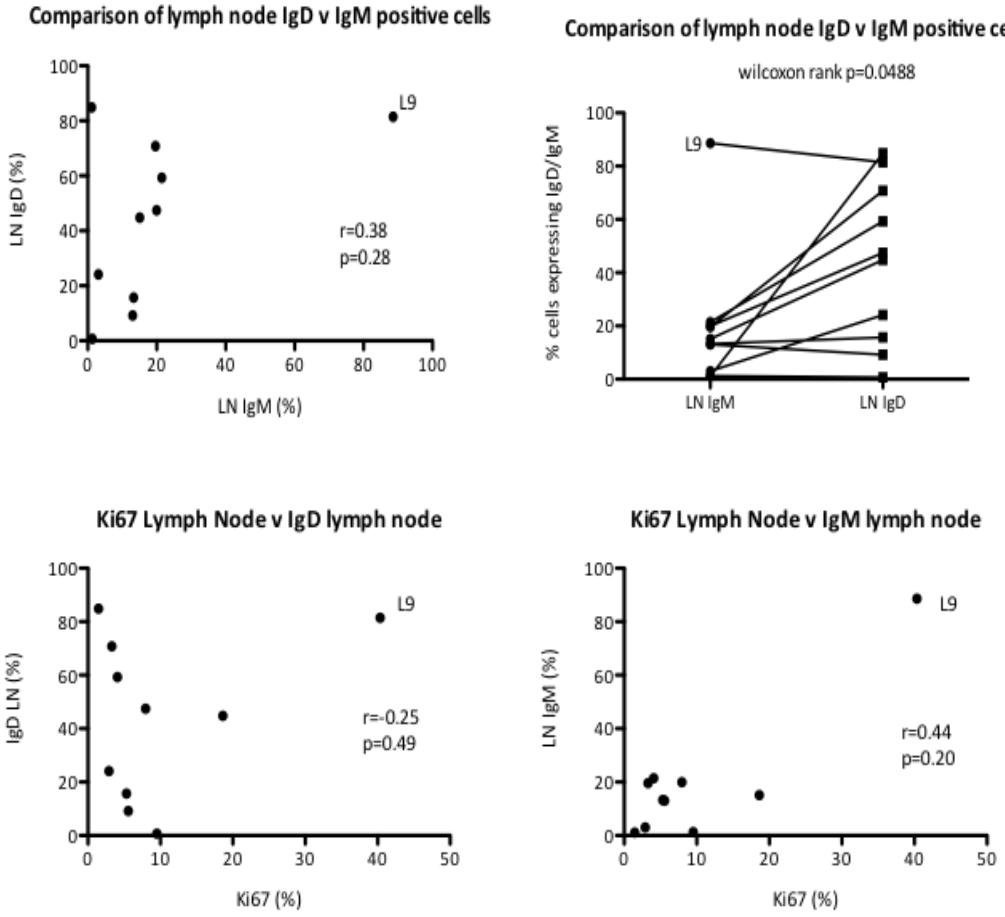
Results are summarised in Fig. 5.27:

**Figure 5.27: Frequencies of IgD, IgM and ki67 positive cells within each sample. The bars show the mean of 5 high power fields (HPFs) and the error bars show the standard deviation.**



There was no correlation between lymph node IgM and IgD percentage, or with Ki67 and IgD/IgM. There is a suggestion that IgM expression (mean 30%) within lymph nodes was lower than IgD expression (mean 44%) (Fig. 5.28):

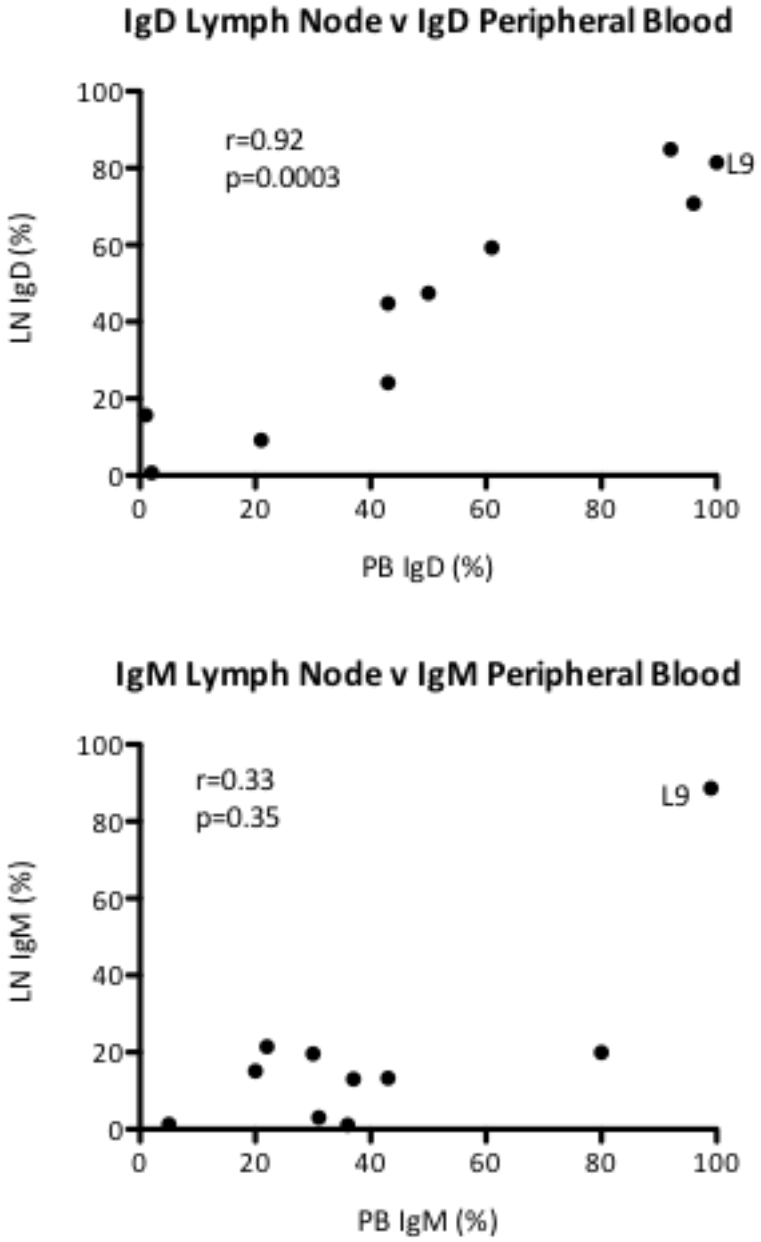
**Figure 5.28: Comparison of IgD, IgM and Ki67 positive cells within each patient. Horizontal bars represent means. The outlier L9 is discussed in the text**




A notable outlier is sample L9, with very high levels of IgM (88%) IgD (81%) and Ki67 (40%). This sample was a section of spleen from a patient who had received multiple lines of chemotherapy prior to biopsy, but had no evidence of Richter's transformation.

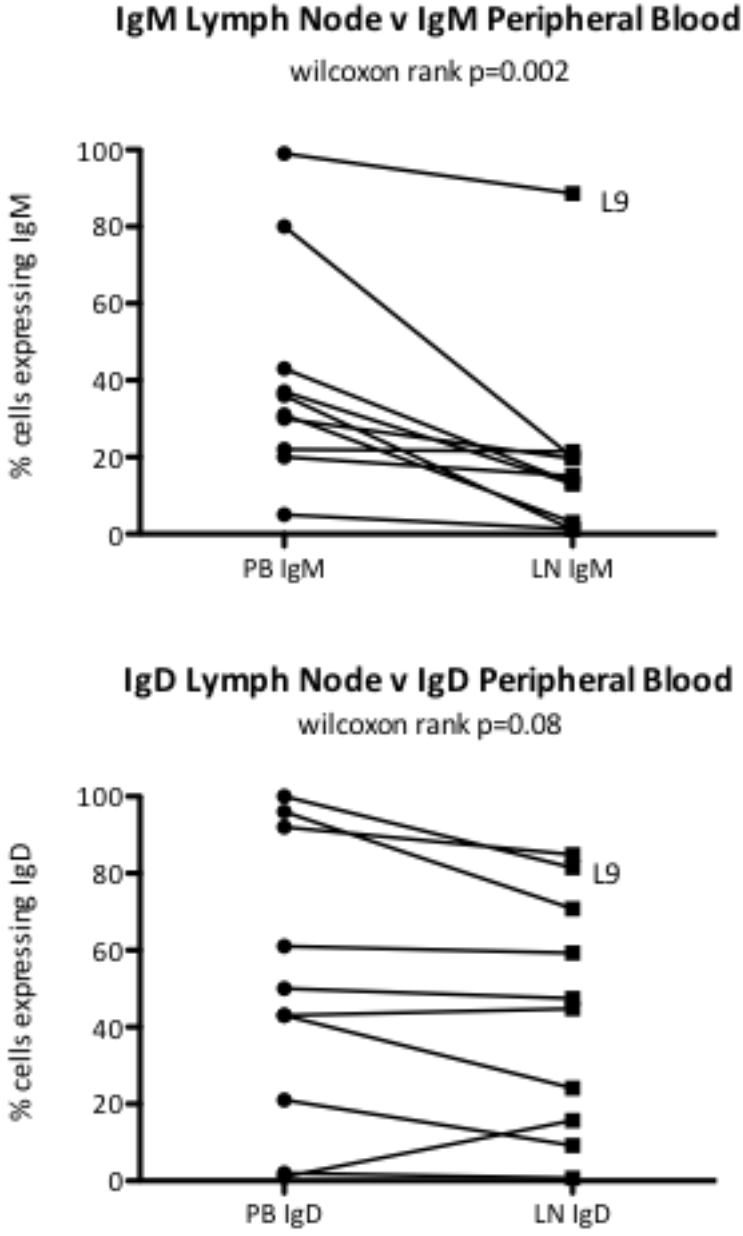
Comparison with PB IgD and IgM expression from the clinical database was made. There was significant correlation between peripheral blood IgD and lymph node IgD (as a percentage of all cells), but not between lymph node and peripheral blood IgM (Fig 5.29):

**Figure 5.29: Comparison of Lymph nodes (LN) and peripheral blood (PB) expression**



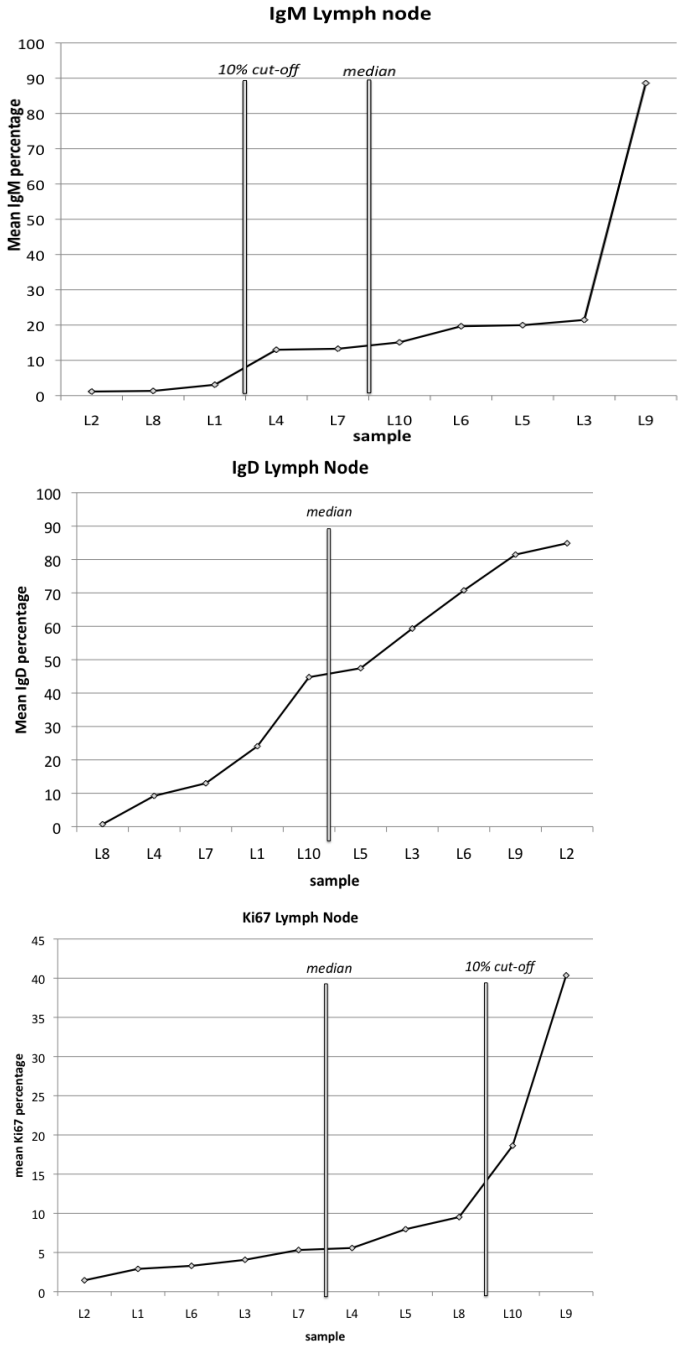

Comparison of expression between PB and LN suggests that IgM expression in the lymph node is slightly lower than that of the peripheral blood (mean 14% difference), whereas this is not the case with IgD (Fig. 5.30):

**Figure 5.30: Comparison of expression levels between PB and LN**



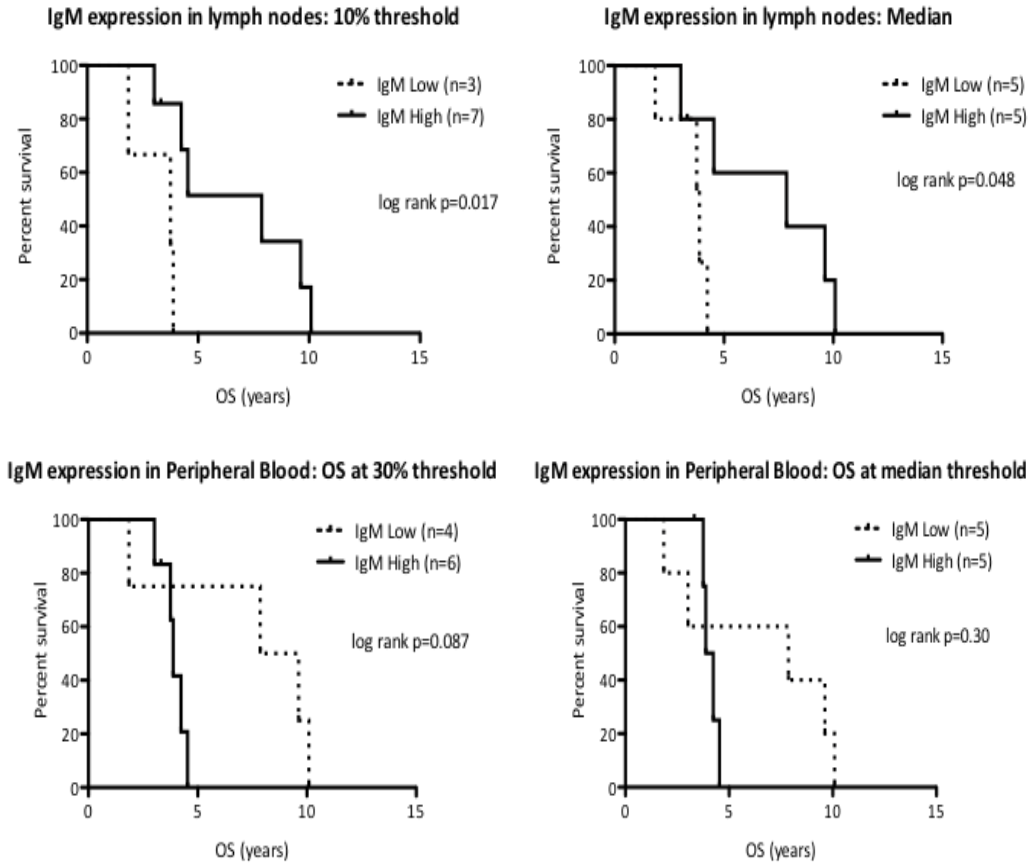

Finally we attempted to correlate disease biology with the expression of IgM, IgD and Ki67 in the lymph node. To do this we classified samples into 'High' and 'Low' expression groups. Survival curves were then plotted for each group. There is no obvious threshold for classifying 'High' and 'Low' and so the data were inspected visually (see Fig. 5.31)

**Figure 5.31: Determination of expression threshold for survival analysis**
**Samples were ordered according to the percentage positivity of each marker in order to determine an obvious threshold for labelling a sample as High or Low. For IgM & Ki67, there seemed to be a 'step' at approximately 10%, and this was selected as a threshold. For IgD, there was no obvious transition, and so the median was used as a threshold. The median was also used as an alternative threshold for IgM & Ki67 High/Low.**



These thresholds (10% or median) were used to classify the samples into High and Low expression for the purposes of survival analysis (Fig. 5.32). There appeared to be a significant effect of IgM expression on overall survival from diagnosis, both at median threshold and 10% threshold, with low IgM levels predicting a decreased overall survival. This is in contrast to the survival of the same patients segregated into different groups by peripheral blood IgM expression.

**Figure 5.32:Survival by IgM expression in lymph node and peripheral blood**




Treatment Free Survival for IgM was not different for either threshold, nor was TFS or OS for median or 10% thresholds when groups were segregated by IgD or Ki67 levels.

### *5.4.6 Summary and Discussion: IgD and IgM expression in the Bone marrow and lymph node microenvironment*

The expression of IgD and IgM in 10 patient samples was evaluated and compared to paired PB samples. Immunophenotyping by immunohistochemistry and flow cytometry used different antibody clones, yet are essentially modifications of the same technique, so it could be claimed that the percentage cells derived from one method might be comparable to that derived by a different method. There are other inherent weaknesses – there were often long periods of time (up to 1.7 years) between PB and LN samples, with some patients receiving treatment between samples, which may be expected to limit interpretation. Lymph node biopsy or splenectomy in patients with CLL is generally performed in advanced stages, or when patients are refractory to treatment, though some biopsies are performed as part of an operation or to investigate lymphadenopathy, with an unexpected finding of SLL. Many of them had received treatment prior to LN biopsy. Also, there is a high rate of poor prognostic markers such as 17p deletion in this group. Therefore, this group of patients is enriched for late stage, treated and poor prognosis cases, and cannot be expected to be typical of CLL as a whole. Biopsies are performed for a reason in CLL, and it is unethical to perform biopsies in early stage patients who do not require treatment. Nonetheless, few studies have examined this question and the studies do provide new information.

The intensity of staining of CLL LN samples was variable, and intensity of IgM and IgD staining per cell was less than that of tonsils, consistent with the concept of 'dim' or 'weak' Ig expression in CLL. Despite positive and negative cells being readily distinguishable, the level of Ig expressed on a positive cell of one patient may be less that that compared to another, and within patients there were some cells that expressed Ig more strongly than others.

There was no correlation of IgM/IgD with Ki67, a finding that suggests that the expression of BCR does not associate with the proliferation of CLL cells in the LN. A caveat to this is that the Ki67-positive cells may themselves express IgD/IgM at a

higher level than their Ki67-negative counterparts. To determine this would require double staining for Ki67/Ig on individual cells. This highlights the fact that only a small proportion of the clone is proliferating at any one time and the methods for IgD and IgM determination examine the bulk population.

A tentative association between shorter overall survival and low expression of IgM in lymph node is suggested by the present findings, but this requires investigation in a larger patient set. IgM expression is lower than IgD expression in LN. IgD expression in lymph node/spleen appeared to correlate with PB Ig expression in the same patient, however, there appeared to be a consistent lower percentage of IgM expressing cells in the LN compared to PB.

If genuine, these data demand a putative mechanism. The critical issue is the relationship between the circulating PB CLL compartment and the presumed proliferative LN compartment. If they were biologically equivalent and in equilibrium, then similar levels of Ig expression would be expected. To an extent, this is seen in the case of IgD: PB levels correlate and are not different from LN levels. However, IgM LN expression is not correlated with PB expression, is lower in LN, and this effect is perhaps associated with more aggressive disease. This suggests that either the LN cells are composed of a population that expresses lower IgM, or that cells reduce their IgM levels on migration to the LN (or both). There are at least two plausible mechanisms for this.

Firstly, if BCR signalling in terms of binding antigen in LN is occurring, as we and others have suggested, then it might be expected that CLL cells would adopt an anergic phenotype in the LN. This involves reducing IgM expression whilst maintaining IgD. The teleological argument for this, as for PB anergic cells, is that the cell is attempting to limit strong antigenic stimulation that might either cause apoptosis or result in autoreactive phenomena.

The second and not mutually exclusive explanation suggests that CLL cells utilise IgM as an aid to migration into its LN microenvironment. Normal B-cells recirculate

multiple times between the PB and LN compartments, searching for antigen. Once antigen is encountered in the LN (recognised via BCR signalling), the cell may become activated and cease its wanderings. This process is associated with changes in chemokine receptors to enable location within the nascent germinal centres and to associate with helper T-cells. It can be postulated that once a CLL cell locates antigen, IgM is downregulated as it has served its role. Chiorazzi's group has offered a similar speculative model to explain the role of CXCR4 in CLL cell migration[482]. It is known that PB CLL has higher levels of the CXCR4 chemokine receptor than LN counterparts, and that stimulation (for example via the BCR) results in lower CXCR4 expression. They suggest that the $CXCR4^{dim}$ CLL cells appear to have recently divided and exited the LN (as evidenced by higher Ki67 and deuterium enrichment in heavy water studies). The life cycle of cells continues in the PB, whereby cells upregulate CXCR4 in an attempt to return to the LN compartment where they will again encounter antigen and other survival signals[482]. The process of changing chemokine expression to modulate B-cell migration and localisation is well recognised, and perhaps IgM serves a similar chemotactic role.

These speculations do not explain why IgD does not exhibit a similar pattern, but the IgM/IgD expression pattern seen in anergic cells is analogous. The implication is that IgM and IgD expression is modulated differently in the PB and LN compartments.

**5.5 General Discussion: IgD and IgM expression in CLL**

There are a number of issues with methodology that are difficult to address. We attempt to compare IgM and IgD expression using immunophenotyping by three different methods. Each method derives a 'percentage of cells positive' value. Whilst it may be justifiable to claim that this may reflect genuine expression of IgD or IgM on cells, it is difficult to state exactly how exactly this can be interpreted, and whether different techniques may be directly compared. Are the cells deemed negative for IgD or IgM truly negative, or is the assay simply insensitive for low expression levels? B-cell survival is incompatible with absence of BCR expression, and so the question of whether $IgM^{-}IgD^{-}$ cells are genuinely 'negative' must be asked[118]. The *in vitro* flow cytometry assays may claim to detect dual $IgM^{+}IgD^{+}$ cells, but the other techniques

used do not. For example, when a method suggests that 75% of cells are positive for IgD or IgM, it does not imply that 75% are $IgM^{+}IgD^{+}$, though presumably some must be. These arguments apply to any study using monoclonal antibodies to detect antigens and that compare expression levels and positive/negative staining. We believe we have offered reasonable arguments as to the veracity of using the methods to determine IgD and IgM expression in CLL.

Given these issues with methods, any conclusions drawn should not be over interpreted. With these caveats in mind, we suggest that IgD and IgM expression levels are different from one another in PB and LN, with IgD levels being higher. Furthermore, we suggest that there are differences between IgM expression in the PB and LN CLL compartments. IgD expression in the LN appears to reflect that in the PB. These data suggest that IgD and IgM expression levels are differentially regulated, implying that they have differing functional roles.

# 6. Results: Outcomes of BCR ligation *in vitro*: Calcium Flux, apoptosis and cell proliferation

## 6.1 Introduction

B-Cell Receptor signalling is often discussed as if it is a unitary concept, with the capacity, or lack therefore, of B-cells to undergo activation after antigen binding or BCR crosslinking presented as an all-or-none phenomenon. What is implied is the quantification of a measure of BCR signalling, and different authors have used different indicators[383]. Most commonly, two general measures of BCR signalling are employed:

1. Immediate events in the seconds, minutes and hours after BCR ligation. This can involve visualisation of the mobility of single BCR molecules, or of multimolecular complexes involved in the B-cell immune synapse[250]. Calcium flux after BCR ligation is a rapid event that occurs within seconds and is often taken as a surrogate for the capacity to signal[326]. Other early events involve the phosphorylation of various known proximal BCR signalling pathway components including LYN, SYK and BTK, or of molecules further downstream including AKT, NFKB and ERK[483]. Later events (hours to days) include changes in gene expression and cellular morphology.
2. Outcome events involving the fate of the individual B-cell. This may imply cell proliferation or death by apoptosis, migration from one compartment to another, or interaction with other cells[122]. A wider context may include phenomena such as antibody production, tumour growth or death of the organism.

All of the above measures may be taken as 'readouts' of BCR signalling, but most commonly, calcium flux, protein phosphorylation and apoptosis/proliferation are used as surrogates for BCR signalling[122,250,308], and have been used to study the BCR in CLL[114,115,186,249,360,484,485]. The aim here was to examine and compare the outcomes of IgM and IgD signalling in primary CLL samples in order to confirm and extend published findings.

## 6.2 Calcium flux after BCR ligation in CLL

### 6.2.1 Introduction

In normal B-cells, initial BCR signalling events include abrupt rises in cytoplasmic calcium concentration ($[Ca^{2+}]i$), also known as calcium flux[326]. Calcium flux is often taken as a surrogate for BCR signalling capacity. Duration and degree of these calcium bursts is modified by the nature of antigen, the developmental stage of the B-cell and the presence or absence of co-signalling[326]. Calcium is one of the most important small molecule second messengers employed by cells for signal transduction. Eukaryotic cells maintain calcium concentration in their internal milieu 2-4 orders of magnitude below that of extracellular fluid (the typical concentration of which is 2mM)[326]. This is maintained via the regulation of voltage or ligand-dependent calcium channels and pumps, which permit cytoplasmic influx of calcium from the extracellular medium or from intracellular endoplasmic reticulum (ER) stores. The classic mechanism of activation involves agonist binding to a membrane receptor, causing a downstream intracellular signalling cascade that activates phospholipase C (PLC). Activated PLC then hydrolyses membrane phosphatidyl 4,5-bisphosphate ($PIP_2$) to inositol 1,4,5-trisphosphate ($IP_3$) and 2-diacylglyceraol (DAG). Specific $IP_3$-gated channels are present in the ER membrane, and activation causes a rapid influx of $Ca^{2+}$ into the cytoplasm. Sustained calcium influx is mediated by subsequent opening of surface membrane $Ca^{2+}$ channels to permit entry of extracellular $Ca^{2+}$ when intracellular stores are depleted. The burst of intracellular $Ca^{2+}$ concentration ($[Ca^{2+}]i$) then modifies a diverse range of intracellular processes[326].

Calcium signalling has been explored in the Hen-Egg Lysozyme (HEL) mouse model of anergy. Naïve (nontolerant) cells whose BCRs are ligated with HEL or anti-IgD/M show a biphasic calcium response: a large, transient rise (>1mM) followed by a smaller plateau (>500nM)[326]. Anergic B-cells, resulting from constitutive HEL exposure, exhibit elevated baseline $[Ca^{2+}]_i$, and BCR ligation fails to evoke calcium flux[234,369]. Single cell analysis shows that the mean elevation in $[Ca^{2+}]i$ is produced by asynchronous calcium oscillations (200-300nM range) which are decreased when these cells are transferred to mice that do not express HEL, indicating that this response is due to continued

stimulation of the self-reactive BCR[369]. Disengagement of antigen from the anergic BCR results in restoration of low basal [$Ca^{2+}$}i, reduction of constitutive ERK phosphorylation, extension of B-cell lifespan and restoration of the ability to respond to BCR ligation with the burst of [$Ca^{2+}$}i that characterises Ca flux[486].

Approximately 50-75% of CLL cells are able to display Ca flux in response to crosslinking IgM[327,328]. The ability of CLL cells to undergo Ca flux after IgM crosslinking has been associated with widespread tyrosine phosphorylation. In particular, phosphorylation of SYK is correlated with the ability to undergo Ca flux[337], and defective PLC phosphorylation may be a putative explanation for the absence of Ca flux in many cases[327], with greater phosphorylation in U-CLL[320]. There are also suggestions of greater calcium flux in CD38 positive cases[307]. Attenuated Ca flux after IgM ligation has been noted recurrently in CLL, and some studies suggest that Ca flux is generally required for subsequent downstream events such as phosphorylation of BCR pathway components, and alterations in apoptosis/proliferation[49,305-307,317,319,320,329]. In contrast, IgD signalling as judged by Ca flux seems to be relatively preserved[116,305,306]. Few publications have compared IgD and IgM signalling.

IgM and IgD signalling have been studied in CD38-positive cases in order to enrich for cells that demonstrate robust BCR signalling[306]. The data on IgM/D-induced Ca flux are not explicit in this study, although they suggest similar amplitude of Ca flux after anti-IgM and IgD ligation, but shorter duration after IgD signalling. They also suggest similar temporal patterns when the profile of SYK phosphorylation is monitored.

Professor Stevenson's group has studied IgD signalling more extensively. In one study of 40 CLL samples, all expressed IgM and IgD, yet 15 (38%) did not respond to IgM cross-linking as measured by Ca flux and SYK phosphorylation. Of these 15 non-responding samples, 10 (66%) were responsive to anti-IgD and 12 (80%) were responsive to CD79a crosslinking. The 3 samples that did not respond to anti-CD79a were also nonresponsive to both anti-IgM and anti-IgD. The study was used as an illustration of dysfunctional IgM signalling, with relative preservation of IgD and general downstream signalling in CLL[305]. They subsequently comment on their own

unpublished data that suggests that whilst IgD responses (as measured by Ca flux and ERK phosphorylation) are more common, they are transient compared to IgM[249]. The same group have published similar data demonstrating preserved IgD signalling in the presence of decreased IgM signalling in another study of a larger cohort of 112 patients[116]. They suggest that 53% of cases could be designated as 'responders' to IgM ligation, as judged by the presence of >5% of cells exhibiting Ca flux. Responders tended to be U-CLL, have high IgM expression and be CD38 and ZAP70 positive. By considering cases with discordant *IGHV* gene mutation status, CD38 and ZAP70 status, they suggest that the *IGHV* gene mutation status is a greater determinant of signalling capacity than CD38/ZAP70 expression. Higher IgM expression was associated with Ca flux, particularly within the M-CLL group. However, no marker completely dichotomized samples into responding or non-responding groups. In particular, cases with equivalent IgM expression frequently had divergent Ca flux responses, indicating that IgM expression alone does not determine Ca flux capacity, and that the IgM had become 'disconnected' from downstream signalling pathways. In contrast to this defective IgM signalling, IgD crosslinking produced Ca flux in 95% of cases, and the majority of cases exhibited higher levels of Ca flux when compared to IgM-induced Ca fluxes. Although the data is not shown, the authors state that co-addition of anti-IgM with anti-IgD does not affect the IgD-induced Ca flux.

Intriguingly, the findings of this study also suggest that IgM expression and signalling capacity can be restored by incubation of cells *in vitro*, analogous to the scenario with anergic B-cells. They selected a group (n=14) of responders and non-responders to IgM as judged by Ca flux and incubated them in medium alone for 24-72 hours, assessing recovery as an increase in >10% cells responding with Ca flux. In 10 of 14 cases there was a clear recovery of IgM-induced Ca flux that was generally evident by 24 hours. This included 6 of 7 initial non-responders, and 4 of the 7 responders also showing an improvement in Ca flux.

In the initial responders group, the 4 that increased their responses also showed an accompanying increase in expression of IgM (as judged by Median Fluorescence Intensity of the $CD5^{+}CD19^{+}$ population), with 2 cases increasing IgM but not increasing

Ca flux. In the non-responder group, most cases increased IgM expression, including the case that did not recover Ca flux capacity. In general, improvement in Ca flux ability was associated with increases in IgM expression, but this IgM increase was not sufficient to permit increase in Ca flux. IgD-induced Ca flux increased in one case, but was otherwise stable over 24-72h, and IgD expression tended to decrease during incubation.

Furthermore, when selected cases were incubated with anti-light chain to crosslink the BCR and cause internalization of the surface Ig, this resulted in an immediate reduction in IgM and blockage of Ca flux after IgM ligation, with partial recovery after incubation in vitro for 24h. In the 6 cases tested, IgM and Ca flux after 72 hours incubation exceeded that present in the original cells. These data suggest that CLL cells can increase IgM signalling capacity by re-expression of Ig after endocytosis. IgD levels and calcium flux followed a similar pattern after Ig internalization, though the levels of Ca flux/IgD expression did not exceed those seen at baseline.

Because of the similarities to the HEL mouse model of anergy, Mockridge et al suggest that anergy to IgM ligation is reversed *in vitro* by the removal of antigen, permitting re-expression of IgM, and/or 're-connection' of the downstream signalling apparatus to the expressed IgM. Not all cases manifest this phenomenon, and the authors speculate that in this situation a putative autoantigen is expressed by the CLL cells themselves, leading to the anergic phenomenon being perpetuated *in vitro*. This reversible model of IgM anergy with preserved IgD is a plausible mimic of mouse models of anergy, whereby removal of autoantigen results in re-expression of IgM and regaining of its ability to signal[236,486]. IgD expression and signalling is less affected by these manoeuvres, though it is known to exhibit limited downregulation and defective signalling capacity after chronic BCR engagement in mouse models[292].

In summary, many authors use Ca flux as a surrogate for BCR signalling, with other proximal events such as phosphorylation of proteins correlating with the burst of cytoplasmic ($[Ca^{2+}]i$, dependent on SYK-induced phosphorylation of PLC. Both anergic and CLL cells seem to have defective IgM BCR signalling, as measured by Ca flux. This is

often associated with reduced IgM expression, though defective IgM signalling is also exhibited when IgM expression is high. In contrast, IgD signalling in anergic and CLL cells appears relatively preserved. This IgM-anergic phenotype can be reversed by incubation *in vitro*, with the suggestion that removal of the putative chronic autoantigen exposure permits increased expression and/or capacity for signalling via IgM.

Similarly, I use Ca flux as a measure of early events in BCR signalling and studied the effect of IgM and IgD ligation.

## 6.2.2 Methods

### *6.2.3 BCR ligation*

*In vitro* methods that attempt to recapitulate the physiological response of B-cells to antigen follow two approaches. Either the known cognate antigen for that B-cell is used as a stimulus (e.g. Hen Egg Lysozyme in mouse models with BCRs of known specificity), or the BCR is crosslinked using monoclonal antibody to Ig heavy or light chains. With the unavailability of known antigen, studies in CLL have taken the latter approach. Table 6.1 summarises relevant publications.

**Table 6.1. Characteristics of the stimuli used to crosslink the BCR in various studies of CLL.**

| Reagent | Concentrations | Supplier | References |
|---|---|---|---|
| **Anti-IgM** | | | |
| Rabbit polyclonal anti-μ $F(ab')_2$ | 15μg/mL[327]<br>10ng/mL-100mg/mL[307]<br>40μg/mL.[324] | Jackson Immunoresearch laboratories[327], Accu specs[307] | **Michel et al 1993[327]**<br>**Zupo et al 1996[307]**<br>**Gordon et al 2000[324]** |
| Goat polyclonal antihuman IgM/μ $F(ab')_2$ | 10ng/mL-100μg/mL[306,307] (mostly 10μg/mL)<br>5 μg/mL[348]<br>10μg/mL [62,115,317,320-322,330,341,364,487]<br>20 μg/mL[116,305,364] | Southern Biotech[116,305-307,320,321,348,487]<br>Jackson Immunoresearch Laboratories[317,364]<br>Caltag Laboratories[115,330]<br>Sigma[322]<br>MP Biomedicals[341]<br>American Qualex[62]. | **Zupo 1996[307]**<br>**Bernal 2001[62]**<br>**Lanham 2003[305]**<br>**Petlickovski 2005[320]**<br>**Zupo 2000[6]**<br>**Allsup 2005[317]**<br>**Contri 2005[330]**<br>**Deglesne 2006[321]**<br>**Abrams 2007[364]**<br>**Gobessi 2007[348]**<br>**Mockridge 2007[116]**<br>**Muzio 2008[115]**<br>**Guarini 2008[322]**<br>**Gobessi 2009 [487]**<br>**Quiroga 2009[341]** |
| Rabbit anti-human μ | 10 μg/mL | Sigma[306] | **Zupo 2000[306]** |
| Anti IgM IgG | 10μg/mL | Tenovus labs | **Cragg 2002[488]** |
| Goat antihuman IgM/μ IgM | 10 μg/mL. | Southern Biotech | **Petlickovski 2005[320]** |
| Goat anti-human IgM/μ IgG | 20 μg/mL. | Southern Biotech | **Lanham 2003[305]** |
| Anti-human IgM | 10 μg/mL. | Not stated | **Minuzzo 2005[489]** |
| Goat anti IgM/IgD | 10μg/mL | Not stated | **Baudot 2009[335]** |
| Biotinylated goat antihuman IgM F(ab')2 followed by avidin | 2-10[329], 10-20[318,346,484]-, 20[423]μg/mL | Southern Biotech | **Lankester 1995[329]**<br>**Chen 2002[346]**<br>**Chen 2005[318]**<br>**Vallat 2007[423]**<br>**Chen 2008[484]** |
| Mouse anti- μ followed by goat anti-mouse Ig | 5μg/mL | Beckton Dickinson[307] | **Zupo 1996[307]** |
| **Anti-IgD** | | | |
| Goat antihuman IgD/δ $F(ab')_2$ | 20 μg/mL | Southern Biotech | **Mockridge 2007[116]** |
| Goat polyclonal anti-δ | 10ng/mL-100μg/mL (mostly 10μg/mL)[306] | Sigma[306] | **Zupo 2000[306]** |
| Rabbit anti-human δ | 10 μg/mL | Caltag[306] | **Zupo 2000[306]** |
| Biotinylated anti–δ , subsequently cross-linked with streptavidin | 2 to 10 μg/mL | Southern Biotech | **Lankester 1995[329]** |
| Goat anti-human IgD/μ IgG | 20 μg/mL. | Southern Biotech | **Lanham et al 2003[305]** |

The use of $F(ab')_2$ rather than complete antibody has been discussed by various authors, with the concern that the Fc component of antibodies may stimulate cells via Fcγ receptors, and this is well known to have a general inhibitory effect on B-cell

activation[307,320]. Despite this, studies comparing complete antibody or F(ab')$_2$ fragments have found no obvious differences in outcome, and in a limited number of experiments, I found no difference in Ca flux or apoptosis. The nature of the BCR crosslinking agent is discussed further in 6.4.7. I selected a stimulus of 10μg/mL of anti-sIg F(ab')$_2$ obtained from Southern Biotech for most experiments, as this seemed to be the most common published reagent.

### *6.2.4 Calcium Flux in Cell lines*

The RL lymphoma cell line was used to optimise Ca flux methods as it displayed robust Ca flux with both IgD and IgM crosslinking. Live cells are selected on FSC/SSC characteristics of lymphoid cells (figure 6.1) and fluorescence at 525nm (FL1 channel) is plotted against time (figure 6.2):

**Figure 6.1: Gating population of interest, RL cell line.**

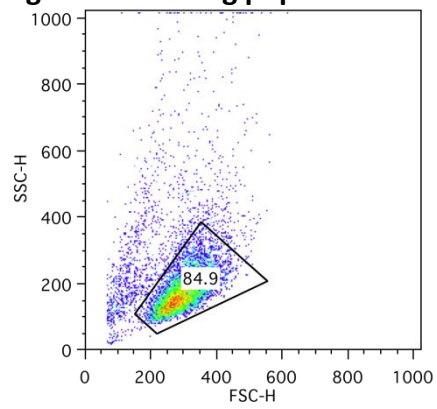
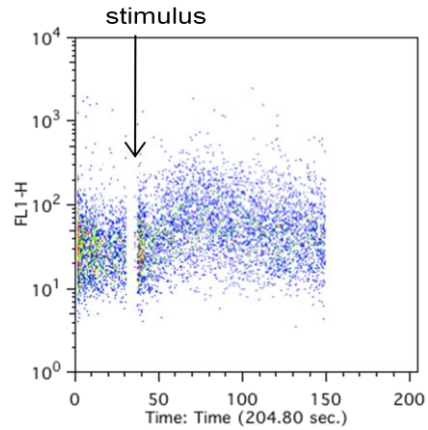



**Figure 6.2.: Fluorescence against time afters IgM stimulus.**
**A bivariate plot of median fluorescence intensity against time.**




This can be analysed further by presenting the data as a median intensity vs time using FloJo software. Figures 6.3-6.5 demonstrate the effect of the various media used for stimuli, as well as the effect of chelation of extracellular $Ca^{2+}$ using ethylene glycol tetraacetic acid (EGTA). The ionophore Ionomycin was used as a positive control in all

experiments as it produces a large and rapid rise in cytoplasmic $Ca^{2+}$ by allowing influx of calcium from the endoplasmic reticulum and the extracellular space. Roswell Park memorial Institute (RPMI) medium was used as a suspending medium for subsequent experiments as it was felt to be more 'physiological' than phosphate-buffered saline, and more robust Ca fluxes were exhibited.

**Figure 6.3: Use of F(ab')$_2$ fragments to crosslink IgM and induce calcium flux in RL cell line**
**In RPMI ($Ca^{2+}$ concentration 0.42 mM ) and PBS ($Ca^{2+}$ concentration 0 mM). Calcium flux is expressed as the median fluorescent intensity (MFI) in the FL1 channel of the flow cytometer**

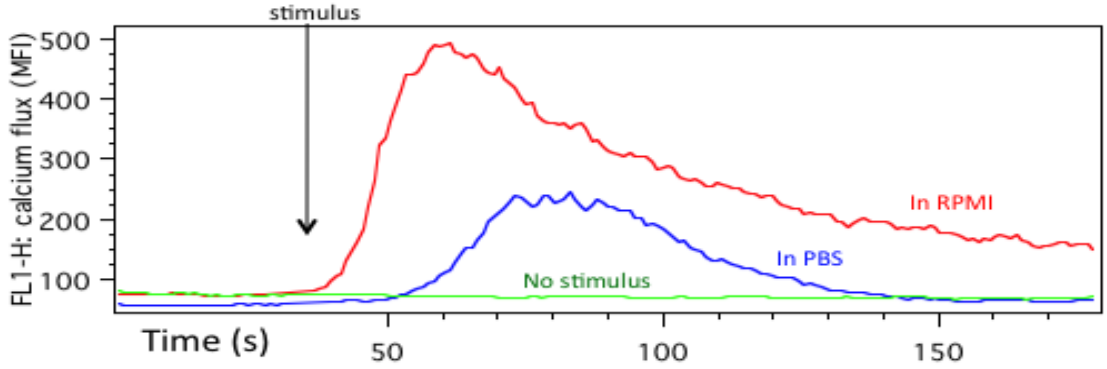



**Figure 6.4: Ionomycin induced calcium flux-median fluorescence intensity in RL cell line, in RPMI and PBS.**

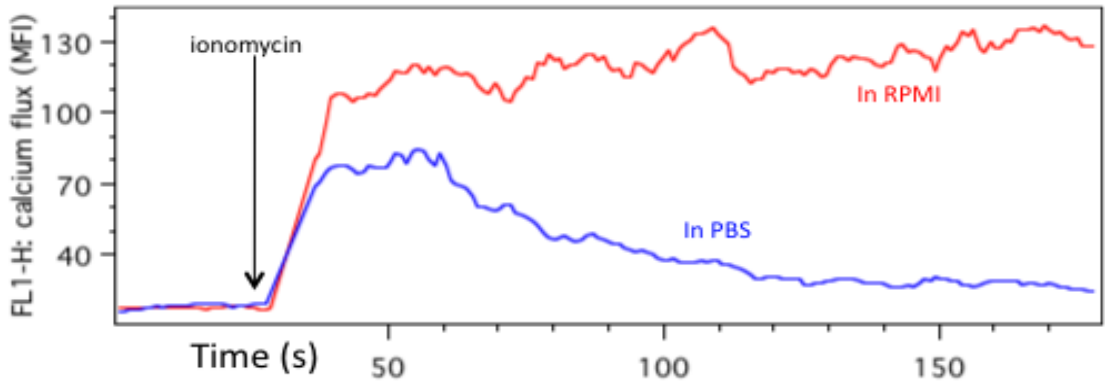
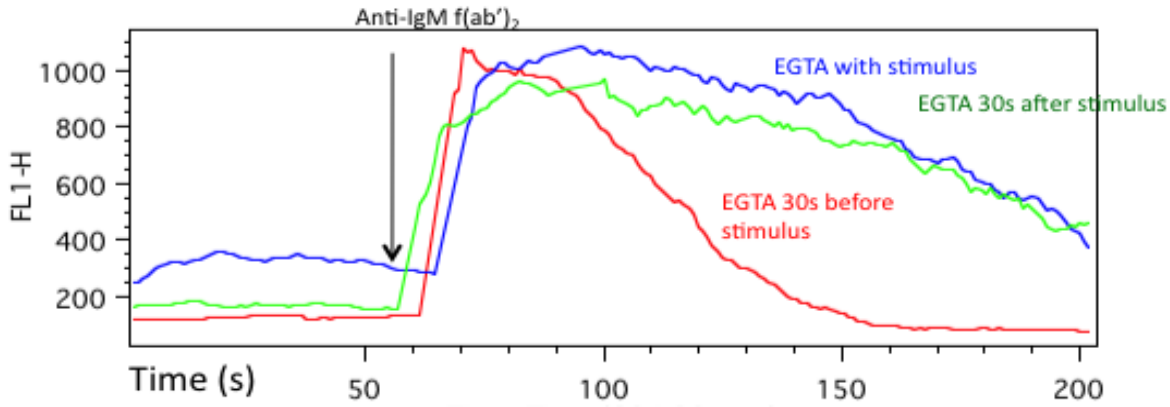
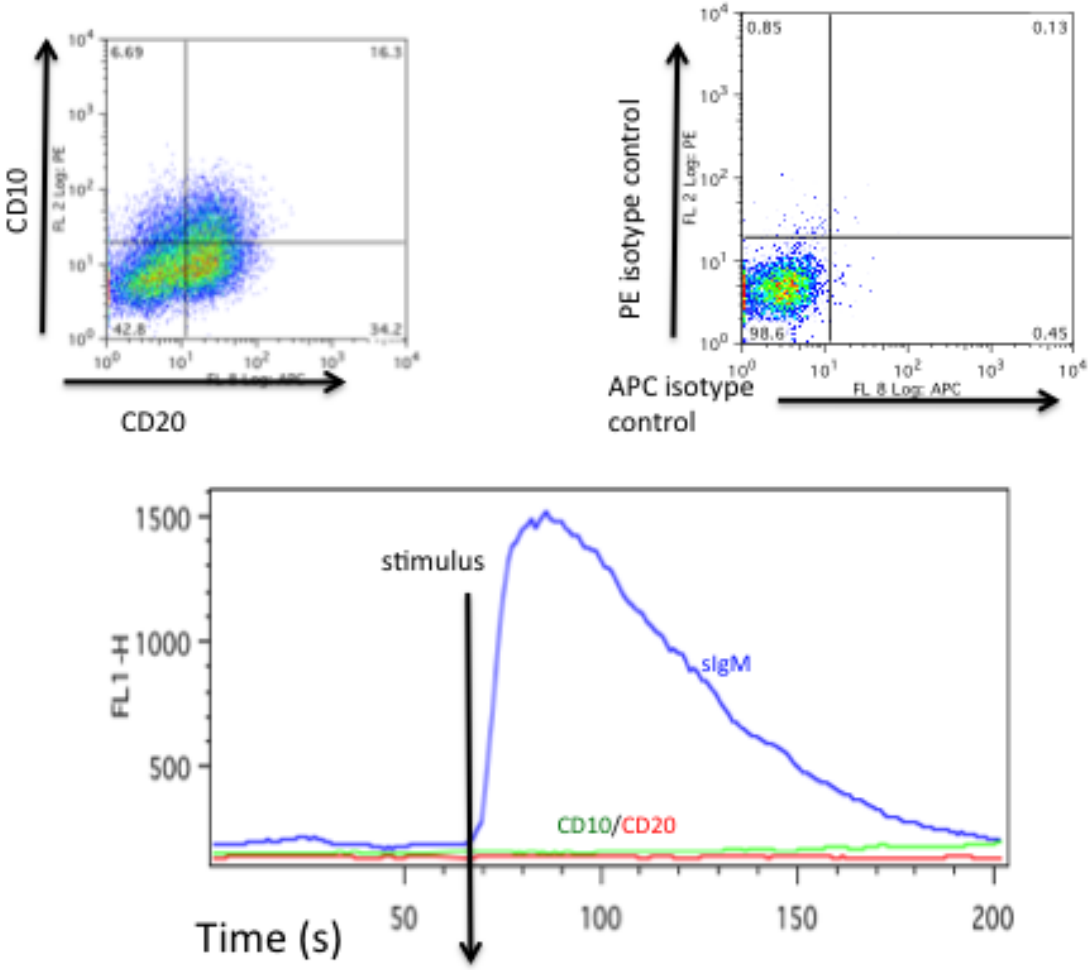



**Figure 6.5: Calcium flux in RL cells in the presence of EGTA (calcium chelator)**



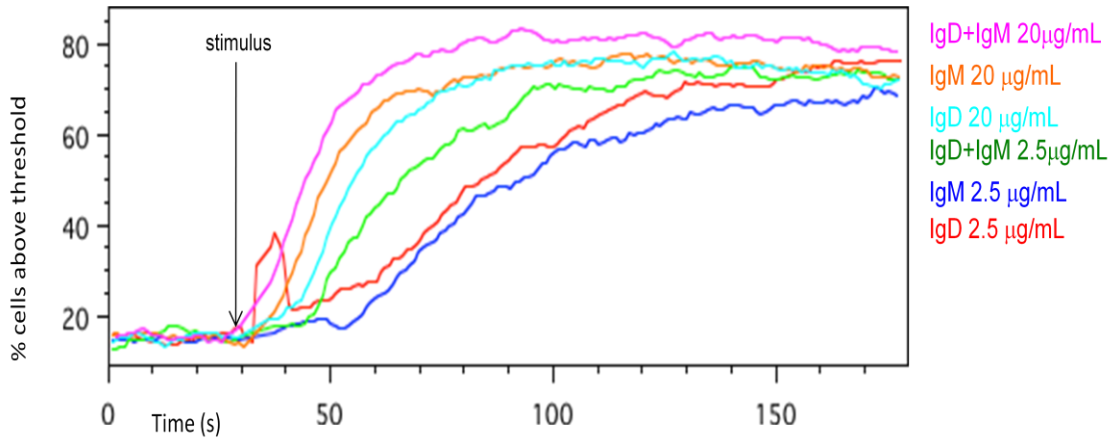
Despite their expression on the RL cell line, crosslinking with a monoclonal IgG specific to CD10 or CD20 did not cause $Ca^{2+}$ flux (Fig 6.6), suggesting that calcium flux was not due to non-specific crosslinking of surface receptors in general.

**Figure 6.6: Surface expression of CD10 and CD20 and crosslinking of sIgM, CD10 and CD20 and resultant calcium flux in RL cell line**




Most publications simply state the presence or absence of calcium flux, a qualitative judgement of whether the median fluorescence intensity (MFI) increases or not[306,326-328,364]. These qualitative assessments may be valid within an experimental setting but create difficulties in comparing results between experiments. In order to produce comparability of calcium flux between samples, Stevenson's group set a background threshold pre-stimulation at the 85$^{th}$ percentile, considering the percentage of cells that exhibit fluorescence above this baseline threshold after stimulus[116]. This then provides us with a measure of the proportion of cells responding (see figure 6.7).

**Figure 6.7: Calcium flux in RL after differing concentrations of anti-IgM and anti-IgD.**
**Shown in the accompanying table are the derived quantifications of calcium flux (time of peak fluorescence intensity and proportion of cells above threshold intensity)**




| Stimulus | Time of peak post stimulus (s) | Peak % of cells above threshold |
|---|---|---|
| Anti-IgD 2.5 μg/mL | 144 | 77.6 |
| Anti-IgM 2.5 μg/mL | 144 | 71.6 |
| Anti-IgD+IgM 2.5 μg/mL | 112 | 76.3 |
| Anti-IgD 20 μg/mL | 98 | 79.5 |
| Anti-IgM 20 μg/mL | 88 | 79.1 |
| Anti-IgD+IgM 20 μg/mL | 63 | 84.7 |

There appears to be a dose effect, with more rapid and maximal stimulation at a concentration of 20μg/mL F(ab')$_2$. Noticeable effects were seen at 1μg/mL (data not shown). Crosslinking IgD produced similar responses to crosslinking IgM.

After optimization in the RL cell line, the same methods were used to study Ca flux in characterized primary CLL cells. Cells were thawed, rested for 2-3 hours by incubation in medium alone at 4x10$^6$/mL before loading with Fluo3 dye and stimulation. The anti-IgM/D F(ab')$_2$ fragments were used at concentrations of 10μg/mL, as this had been

shown to produce robust and maximal responses, when they occurred, and is in keeping with most published methods. When IgD and IgM were added simultaneously, a mixture of stimuli at 5μg/mL was used, as this would equate to a total concentration of stimulus of 10μg/mL, and it is expected that IgM and IgD would compete for the same antigen at a fixed concentration *in vivo*.

The 20 patient samples characterised by IgM and IgD expression in Chapter 5 were studied. Ca flux was performed on samples from the same patient twice in 4/20 cases, and thrice in 3/20 samples. Three of these repeat samples were taken and frozen on different occasions. In all cases, the same pattern of Ca flux was exhibited (5 IgD>IgM, 2 IgM>IgD), suggesting that the patterns of Ca flux are a relatively stable phenomenon, at least in the time shortly after thawing. Where quantification for Ca flux is used (percentage cells Ca flux at 1 minute, see below), the mean of these repeated samples is taken.

## 6.2.5 Results

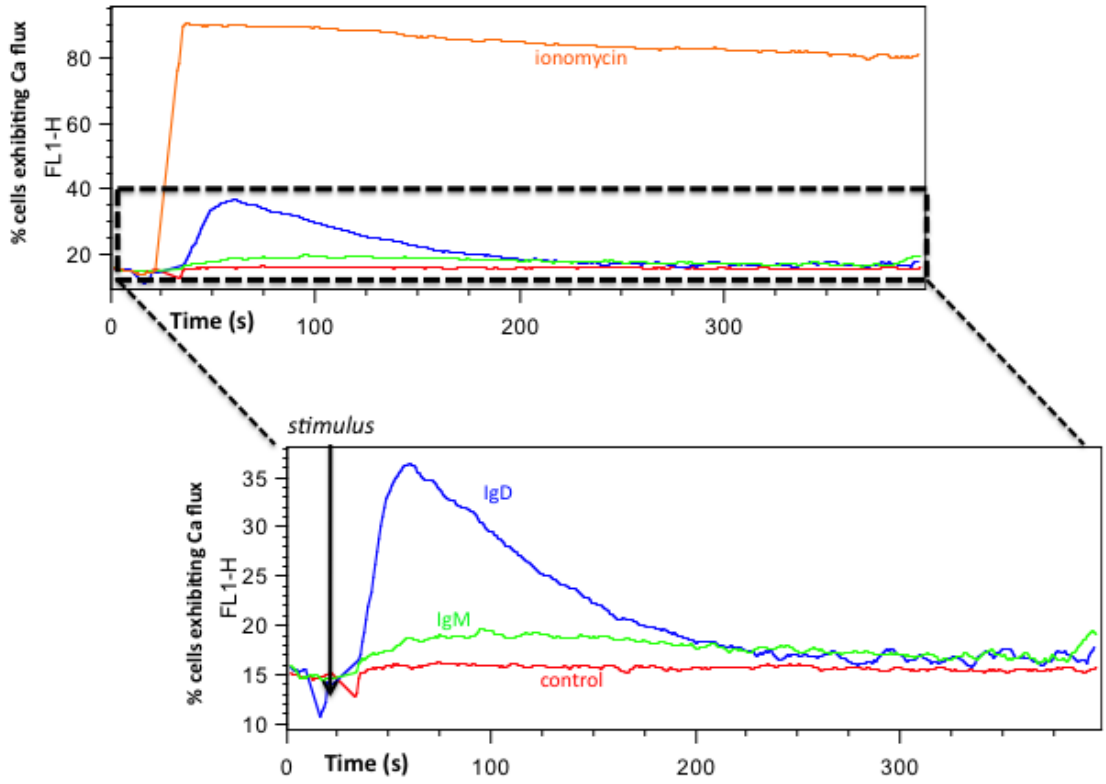
### *6.2.6 Calcium flux via IgD is greater than via IgM*

Representative calcium flux profiles are shown (Fig 6.8). In general, most (13/20) CLL followed a pattern of dual IgM- and IgD-induced Ca flux, with 11/20 having greater IgD-induced Ca flux, and 2 having IgM>IgD. 5/20 displayed no calcium flux to anti-IgM, despite IgD calcium flux. 2 samples (13 and 10) displayed no calcium flux to either anti-IgD or IgM. When combined IgD+IgM stimuli are used, Ca flux levels are intermediate between those exhibited after each stimulation alone.

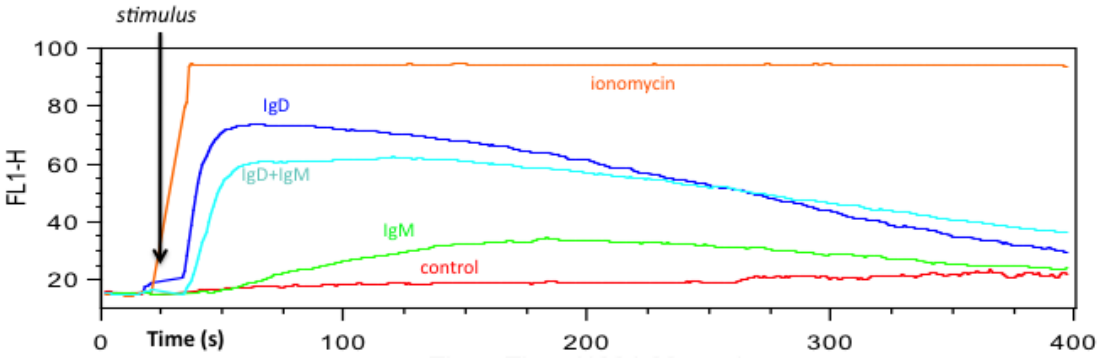
**Figure 6.8: Representative Ca flux profiles of primary CLL cells after IgD or IgM ligation**

**Compared with ionomycin or control (borate buffer vehicle). Note that the method uses a baseline at the 85$^{th}$ centile of unstimulated cells, therefore the baseline proportion of cells is 15%.**

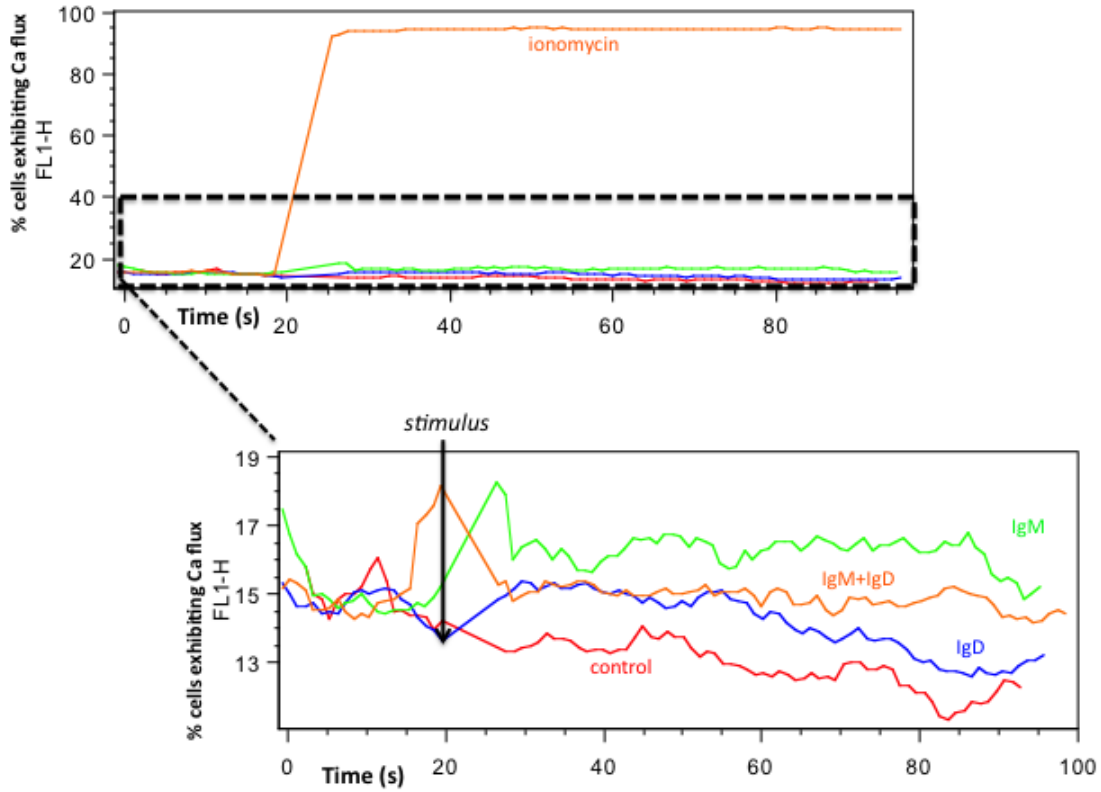
**P1-Exhibits calcium flux with IgD, low level with IgM**



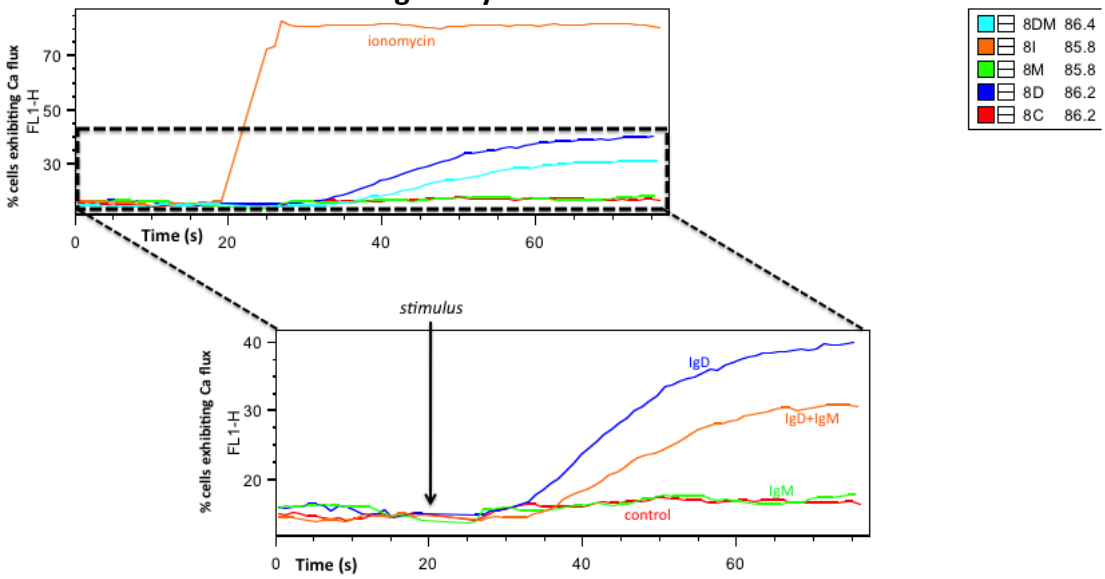

**P14-Exhibits calcium flux with IgD and IgM**



**P10-Exhibits calcium flux with ionomycin only**



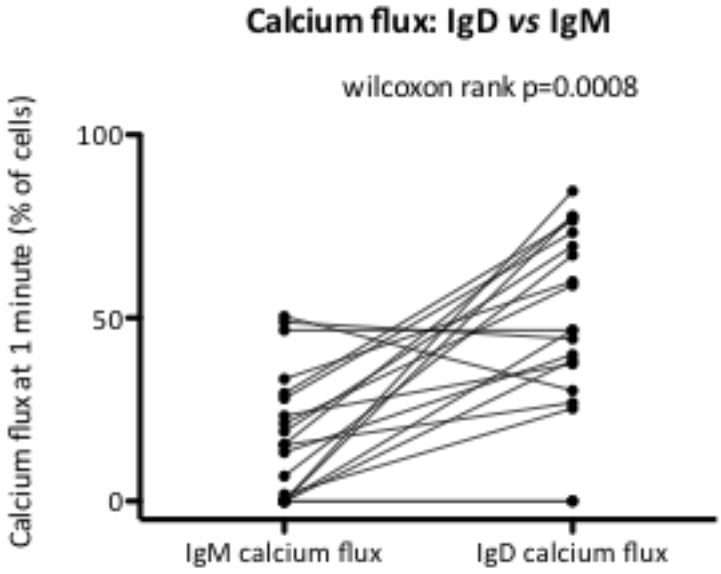

**P5- Exhibits calcium flux with IgD only**




In all samples, there was no obvious difference in the time course of IgD and IgM-induced Ca flux, though samples were not studied beyond 6 minutes, as the number of cells became limiting, and cooling of the medium was thought to have an influence at longer timepoints. In order to derive a quantification of Ca flux, the level (percentage of cells) of Ca flux at 1 minute after stimulus was taken. This was normalised into a

percentage of cells responding by subtracting the 15% baseline, and dividing by the percentage of cells undergoing Ca flux with ionomycin. The ionomycin value was generally 75-100%, and normalisation to this value is expected to compensate for differences in dye loading and viability. Figure 6.9 shows that the level of calcium flux for IgD was significantly higher for IgD stimulus as compared to IgM stimulus.

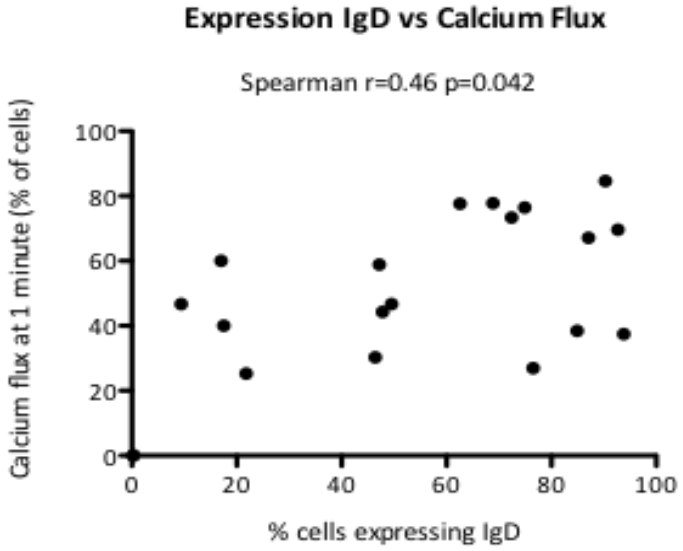
**Figure 6.9: Comparison of Ca flux after IgD and IgM ligation in 20 CLL samples. The proportion of cells exhibiting Ca flux 1 minute after stimulation was compared.**




### *6.2.7 Whilst IgD expression correlates with IgD-induced calcium flux, IgM expression and calcium flux do not*

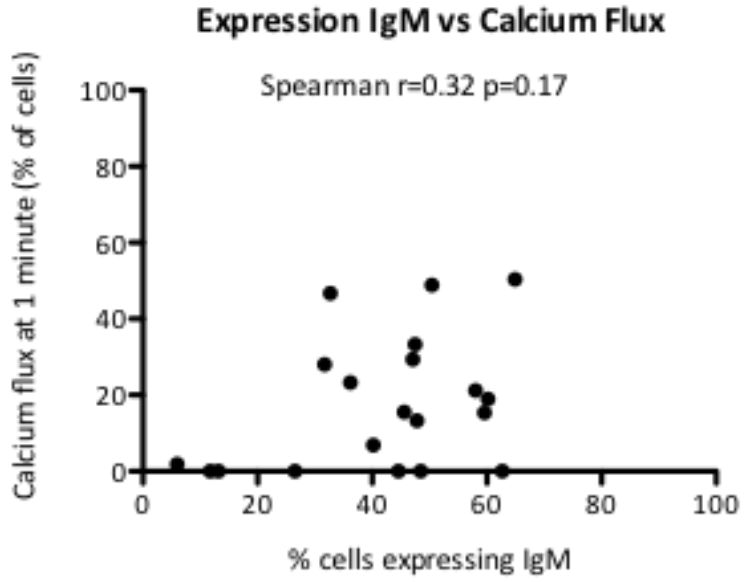
A prerequisite for BCR signalling is expression of the BCR. The levels of Ca flux were compared with the expression of the relevant BCR isotype for the 20 primary CLL samples studied. When the percentage cells exhibiting calcium flux at one minute after stimulation is plotted against the level of expression of the corresponding isotype (see Chapter 5.2), a significant correlation between expression of IgD and calcium flux after IgD ligation is apparent (fig 6.10).

**Figure 6.10: Calcium flux at one minute after stimulation with anti-IgD, compared to expression of IgD.**
**The two samples without IgD expression did not exhibit calcium flux. These same two samples expressed IgM but did not exhibit Ca flux after IgM crosslinking.**



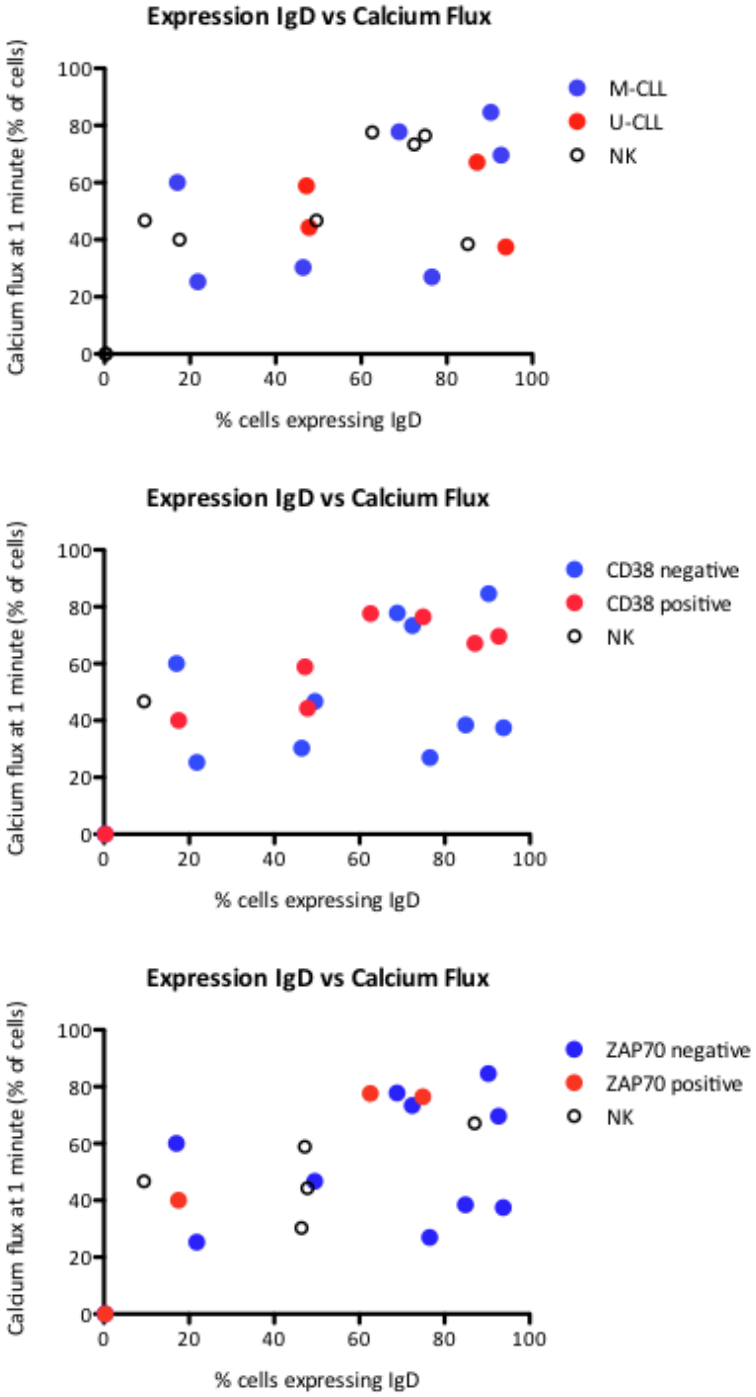

In contrast, 7/20 samples exhibited no IgM-induced calcium flux, despite some of these samples expressing relatively high IgM levels. Of the 13/20 samples exhibiting some level of Ca flux after anti-IgM stimulation, there was no clear relationship between IgM expression level and the level of calcium flux (see Fig 6.11)

**Figure 6.11: Calcium flux at one minute after stimulation with anti-IgM, compared to expression of IgM.**



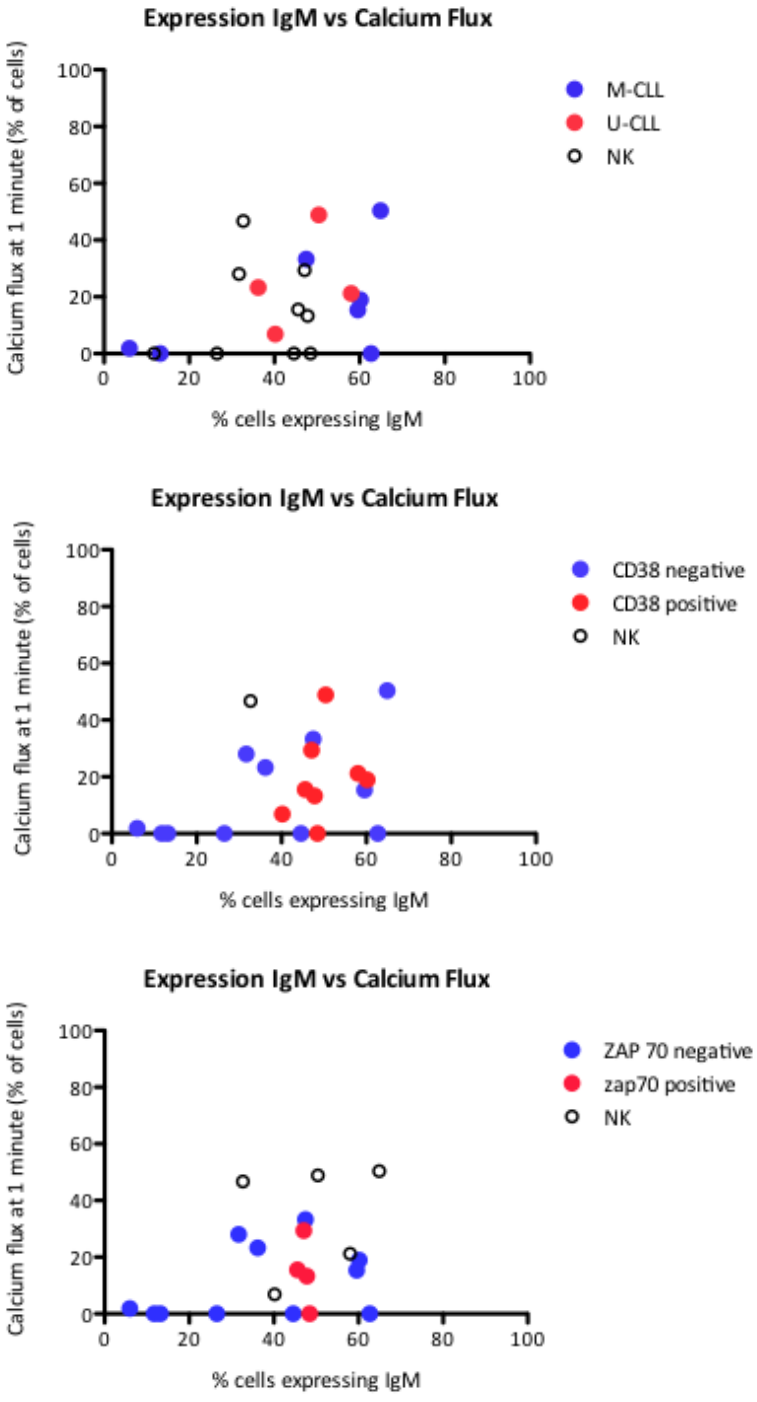
In order to associate the capacity to undergo calcium flux with known prognostic markers, samples were grouped by *IGHV*, CD38 and ZAP70 status. There was no obvious relationship between these markers and IgD expression/calcium flux (see fig 6.12), and the mean Ca flux was not different in these different prognostic groups.

**Figure 6.12: Calcium flux at one minute after stimulation with anti-IgD, compared to expression of IgD, by prognostic group.**
**Samples are coloured according to the indicated prognostic marker. NK= status not known**



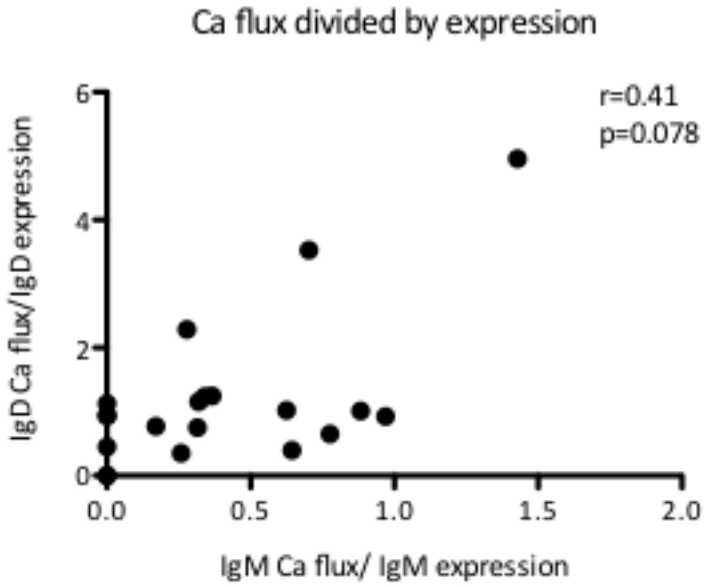
The same data are shown for IgM stimulation in Figure 6.13. Whilst the mean values of Ca flux after IgM ligation were higher in each of the poor prognostic groups (U-CLL, CD38 and ZAP70 positive), there is much overlap and none of these differences were significant at $p<0.05$ levels.

**Figure 6.13: Calcium flux at one minute after stimulation with anti-IgM, compared to expression of IgM, by prognostic group.**
**Samples are coloured according to the indicated prognostic marker.**



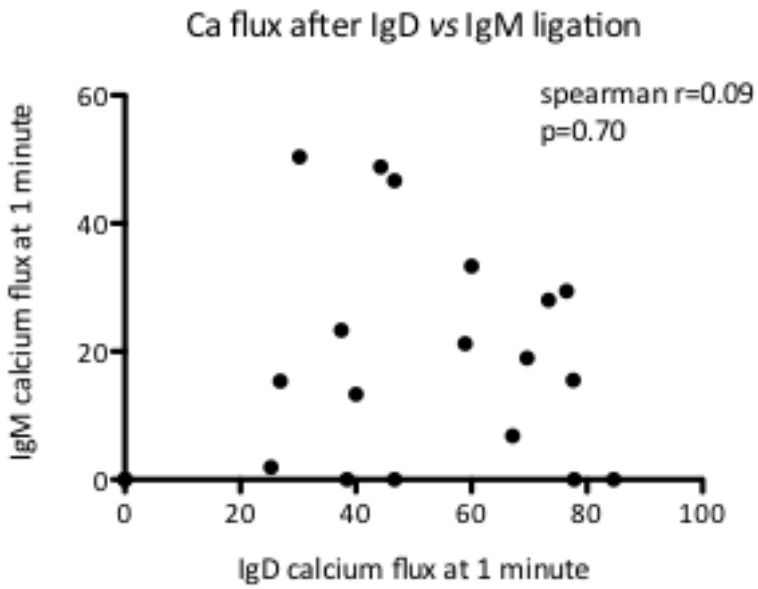
It may be that the greater capacity of cells to undergo Ca flux via IgD was simply related to the possibly higher expression of IgD (see Fig 5.22 in Chapter 5). However, there were several samples where clear expression of IgM was not associated with the capacity for IgM-induced Ca flux. Figure 6.14 calculates the amount of Ca flux divided by expression of the BCR isotype to demonstrate that Ca flux 'per molecule' was less for IgM.

**Figure 6.14: Ca flux divided by expression of IgM/IgM.**
**The percentage of cells undergoing Ca flux at 1 minute after BCR ligation is divided by the expression of the relevant BCR isotype. Note the axes indicate that IgD Ca flux is generally higher relative to the expression of IgD, and there are several cases with no IgM Ca flux. When the different prognostic subgroups are considered, there is no obvious relationship between the derived values of Ig Ca flux/ Ig expression (not shown).**




Also, the inherent capacity of cells to undergo Ca flux at all does not determine the variability of responses, as samples that undergo Ca flux after IgM are not those that have the greatest degree of Ca flux after IgD ligation, and vice versa (see Fig. 6.15).

**Figure 6.15: Ca flux at 1minute after IgD or IgM ligation.**
**There is also no relationship with prognostic markers (not shown).**




### *6.2.8 Restoration of IgM signalling on in vitro incubation*

4 samples were selected to see if *in vitro* incubation restored the capacity to undergo Ca flux after BCR ligation. They were selected based on their known Ca flux characteristics (see table 6.2):

1. **P14**: A sample exhibiting Ca flux after both IgD and IgM engagement.
2. **P1:** A sample exhibiting Ca flux with IgD and low (5.7% of cells) IgM Ca flux.
3. **P5**: A sample exhibiting IgD but no IgM Ca flux.
4. **P10:** A sample with absence of Ca flux with after IgD or IgM ligation.

**Table 6.2: Characteristics of 4 selected samples**

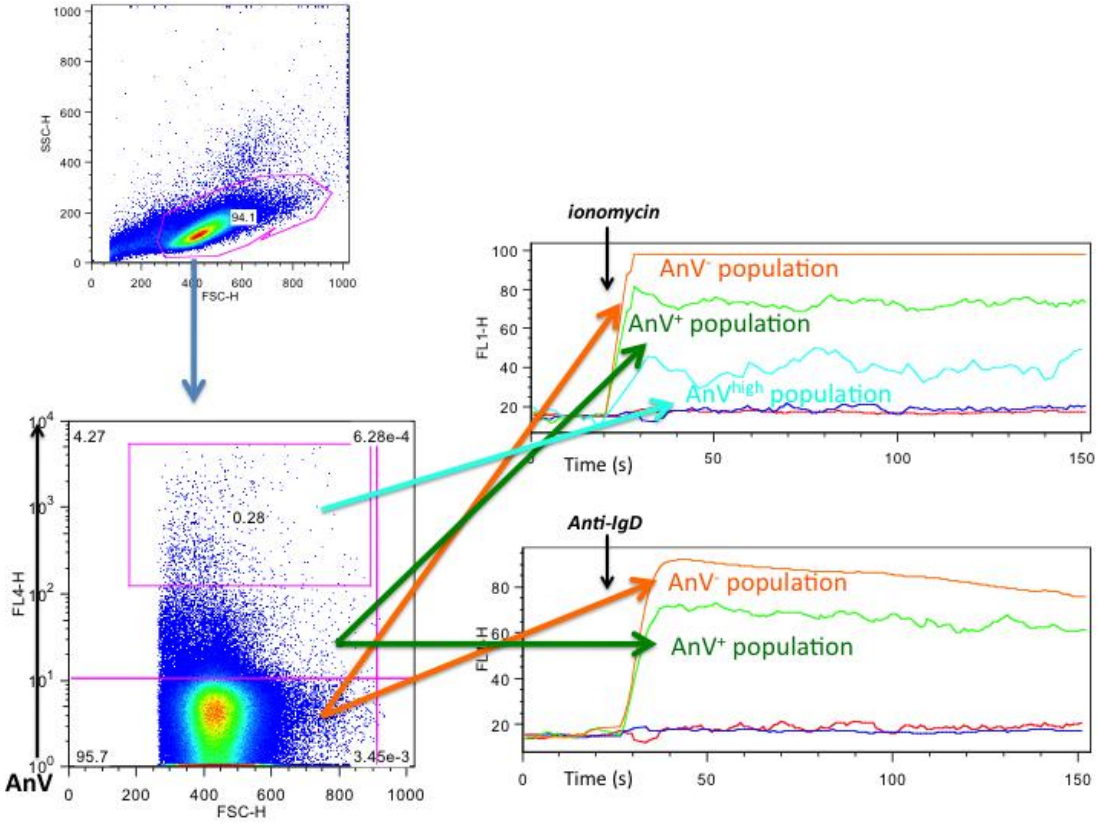
| Sample | Expression IgD (% cells) | Expression IgM (% cells) | Calcium flux IgD (% cells at 1 min. after stimulation, as % of ionomycin) | Calcium flux IgM (% cells at 1 min. after stimulation, as % of ionomycin) |
|---|---|---|---|---|
| P14 | 63 | 46 | 77 | 16 |
| P1 | 22 | 6 | 25 | 6 |
| P5 | 85 | 27 | 38 | 0 |
| P10 | 0.3 | 48 | 0 | 0 |

The samples were incubated for 24h in medium alone, or 'treatment' with vehicle borate buffer, IgM (10μg/mL), IgD(10μg/mL) or IgD+IgM (5μg/mL each). Expression of

IgD and IgM and apoptosis was assayed at t=0 and 24h. Ca flux was assayed at baseline and at 24h.

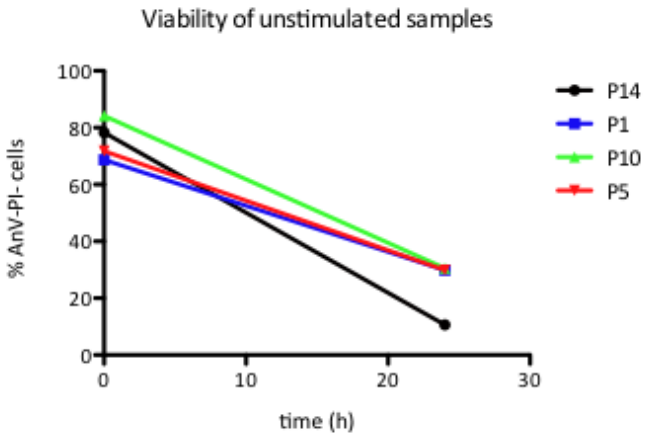
In order to evaluate Ca flux in viable cells only, cells were gated on a population negative for Annexin V staining, but many of the Annexin V positive ($AnV^+$) cells in early apoptosis also had the capacity for calcium flux after stimulation (Fig 6.16):

**Figure 6.16: Selection of viable population for calcium flux assays using AnV-Alexa Fluor 647.**
**An $AnV^-$ population was selected for calcium flux measurements, though it can be seen that many of the cells gated as apoptotic ($AnV^+$) could undergo calcium flux, albeit at reduced amounts. Even the $AnV^{high}$ cells could undergo a degree of calcium flux after ionomycin stimulus.**



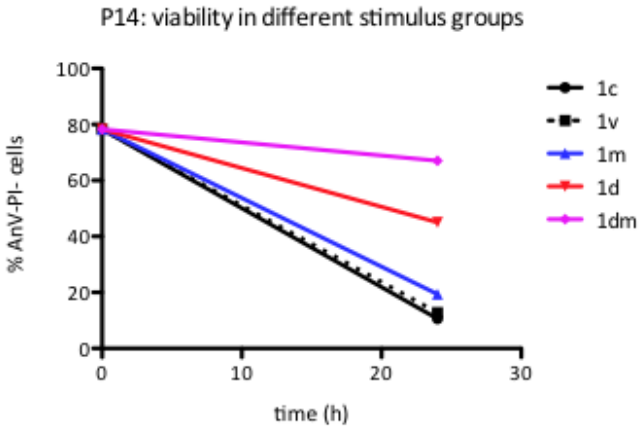

In general, apoptosis increased over time (see fig 6.17).

**Figure 6.17: Viability (annexinV Propidium Iodide dual negative AnV-PI-) cells in cells incubated in medium alone.**



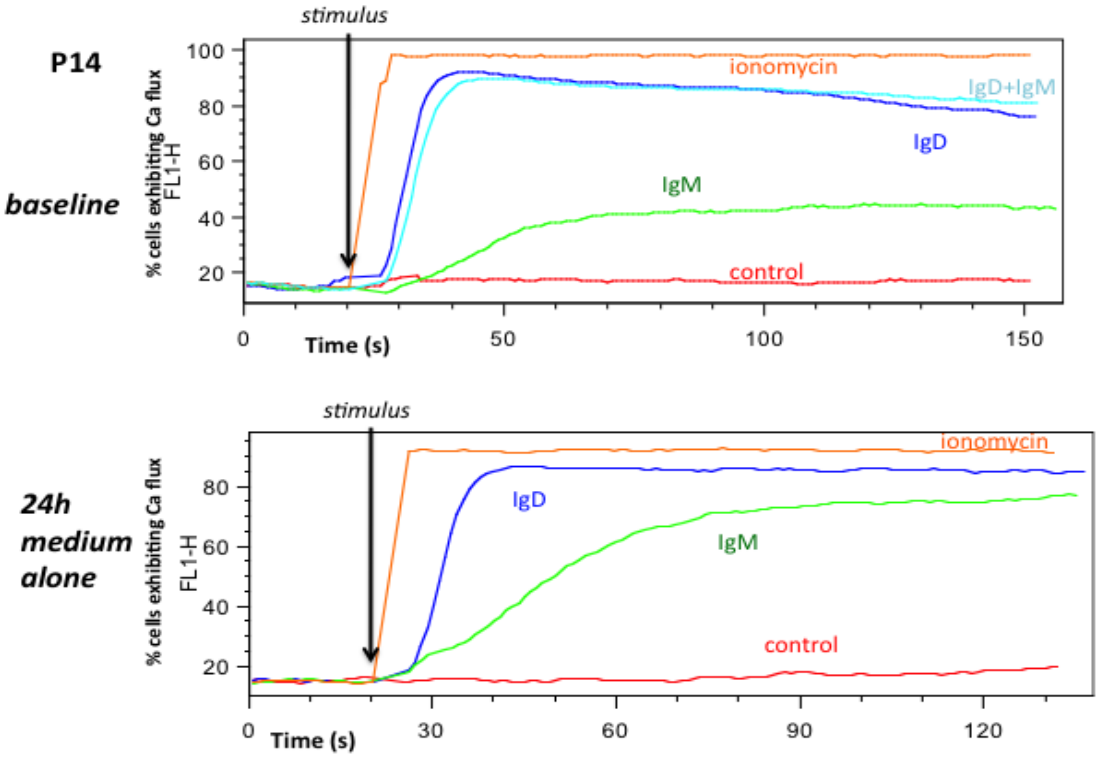

And the different treatments resulted in different levels of apoptosis at 24h:

**Figure 6.18 viability in P14 samples at baseline and 24hours, with different treatment stimuli:**
**1c=control medium alone**
**1v=borate buffer vehicle**
**1m=anti-IgM 10μg/mL**
**1d=anti-IgD 10μg/mL**
**1dm=anti-IgD 5μg/mL + IgD 5μg/mL**




The plots for each patient at baseline and after 24h are shown below. It can be seen that despite a fall in viability over 24h, the CLL cells exposed to medium alone still maintained a viable population that were able to undergo calcium flux in response to stimulus, and in fact, the response to IgM ligation appears to have increased (Fig 6.19):

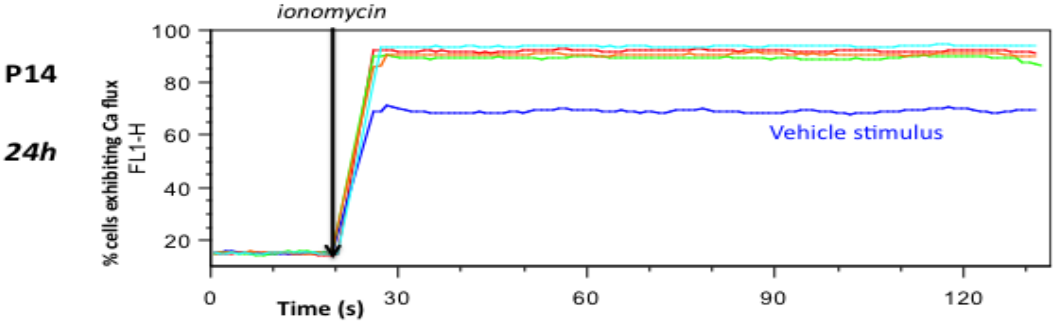
**Figure 6.19: Calcium flux at baseline and 24hours after culture in medium alone. Ca flux was measured after ionomycin, IgM, IgD, IgD+IgM stimulus or vehicle control. The same stimuli were used after 24h culture.**




The level of response to IgD stimulation appears to be similar (>80% of cells undergoing Ca flux) at 24h when compared to baseline.

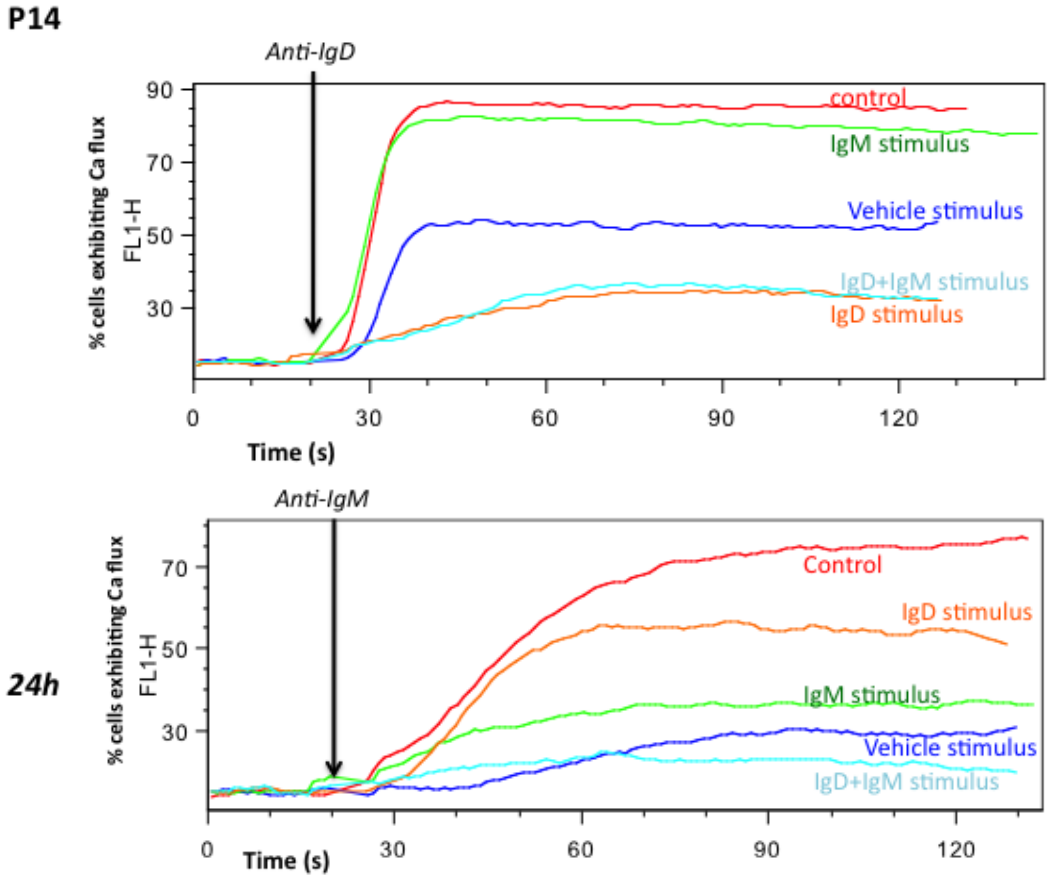
When the ionomycin responses for each treatment stimulus are compared, it seems that the vehicle borate buffer appears to impair calcium flux. This may be due to the low viability of these cells at 24h, but it might be expected that the similar viability of control (medium only) cells might display the same pattern (fig 6.20):

**Figure 6.20: Ionomycin responses after 24h exposure to different treatment stimuli. The ca flux response to ionomycin is shown for cells cultured in various conditions. The top lines represent culture with IgD, IgM, IgD+IgM or medium alone. The lower line ('Vehicle stimulus') represents culture in the presence of vehicle.**



Similar data can be plotted for the responses to IgM and IgD ligation after 24hours (figure 6.21):

**Figure 6.21: Calcium flux responses to anti-IgM and anti-IgD after differing treatment stimuli:**

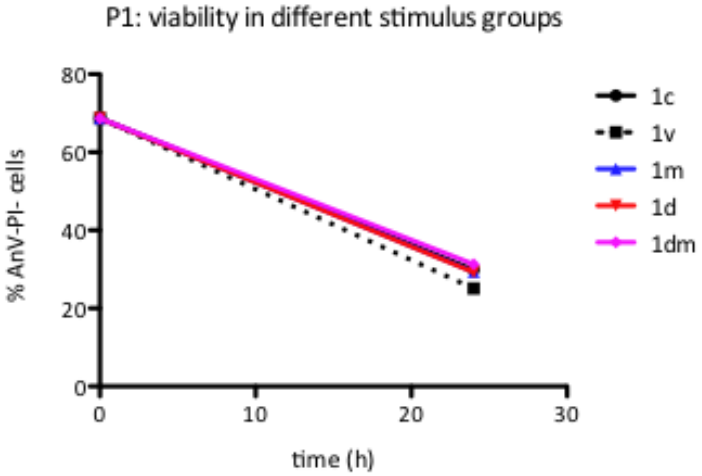



It appears that the response to IgD is impaired by 24h exposure to anti-IgD, and that IgM induced calcium flux is impaired by 24h of anti-IgM treatment. Furthermore, IgD treatment reduces calcium flux response to anti-IgM, whereas the converse is not true. These differences are not simply due to differences in viability, as IgD and IgD+IgM treatment resulted in greater viabilities, yet reduced Ca flux.

Using a combination of FITC-conjugated anti-IgM or IgD and propidium iodide, the expression of IgD and IgM in viable (PI negative) cells was monitored. The expression of IgD declined in all samples, but more so in anti-IgD or IgM+IgD treated samples. The expression of IgM seemed unaltered by treatments and incubation in medium alone *in vitro.*

Similar responses to *in vitro* culture were seen for P1 (Fig 6.22) and P5 (Fig 6.23), though with these samples, there was also an increase in response to IgD stimulation after incubation, perhaps because of the relatively low levels of IgD response at baseline. In P5, the response to anti-IgM was higher than to anti-IgD at 24h, which is remarkable considering the absence of Ca flux after IgM stimulation at baseline. IgD expression declined in all samples, whilst IgM expression was unchanged in P5 and increased slightly in P1 samples (from 6% to 10-15% in all treatments).

**Figure 6.22 P1 viability and Ca flux after incubation**
**The decrease in viability is similar for all treatments. There is a recovery in IgM and IgD calcium flux after 24h culture in medium alone. Vehicle treatment stimulus may be associated with decreased calcium flux to ionomycin, but this is not apparent with the other treatment stimuli. Ca flux response to anti-IgM at 24h is most impaired by 24h anti-IgM treatment, but also by treatment with anti-IgD, and a similar effect is seen with IgD ca flux responses.**



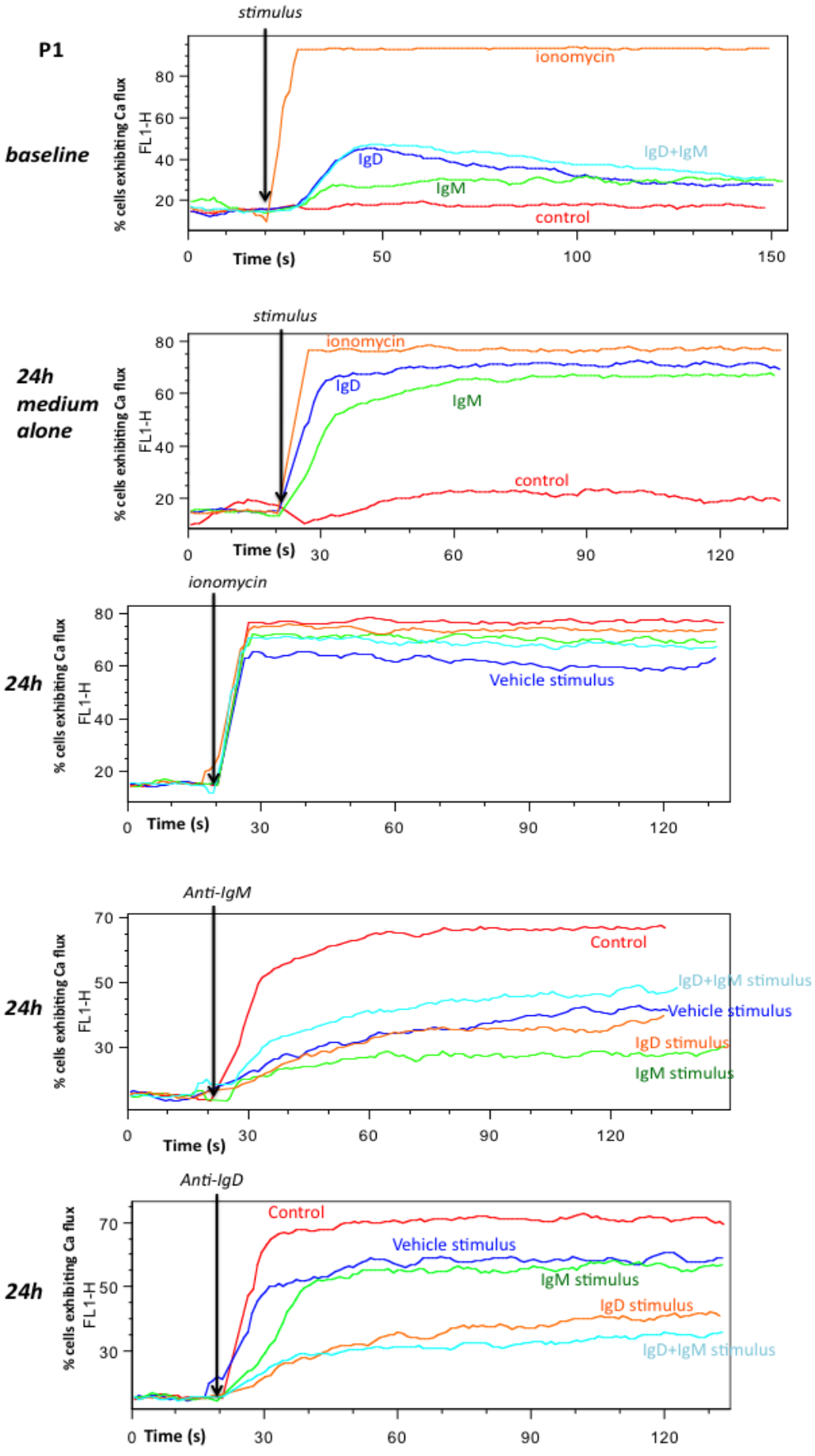
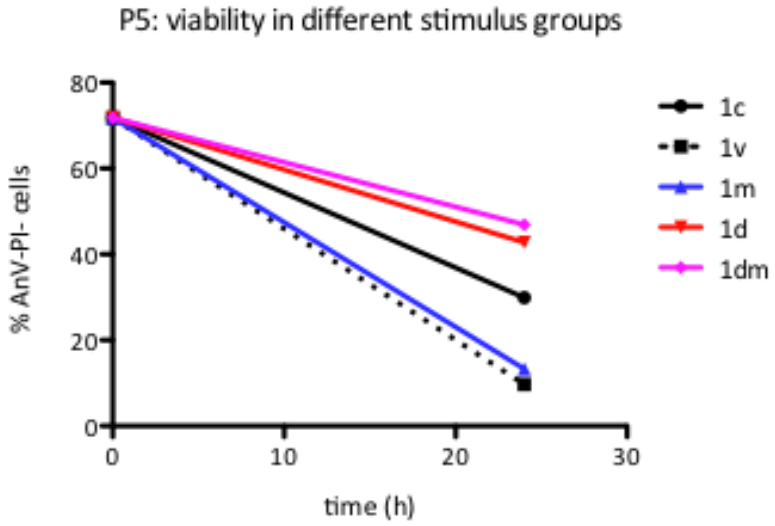
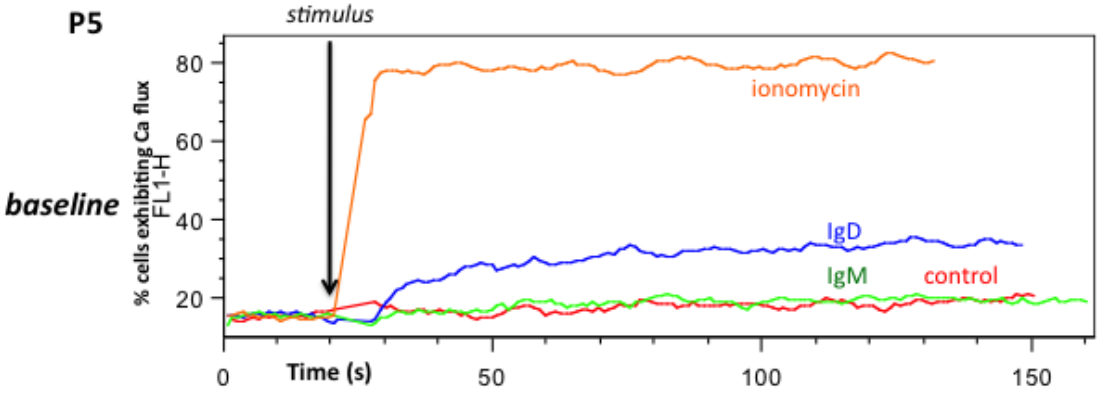
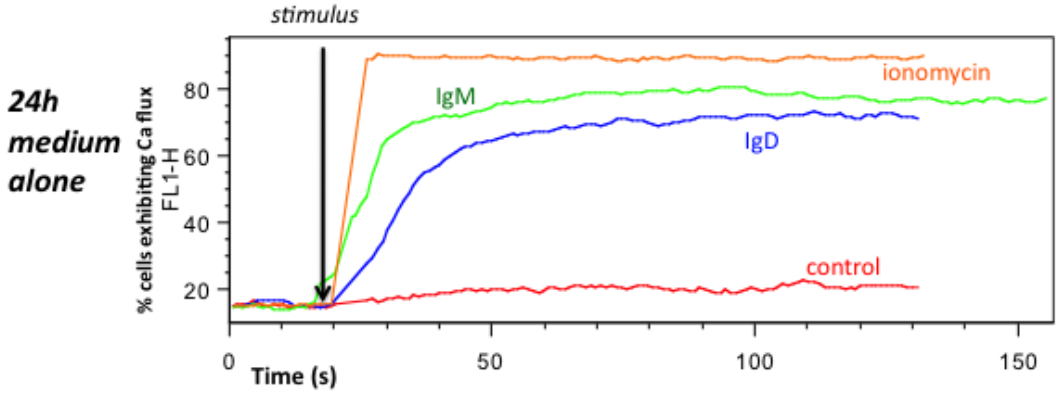
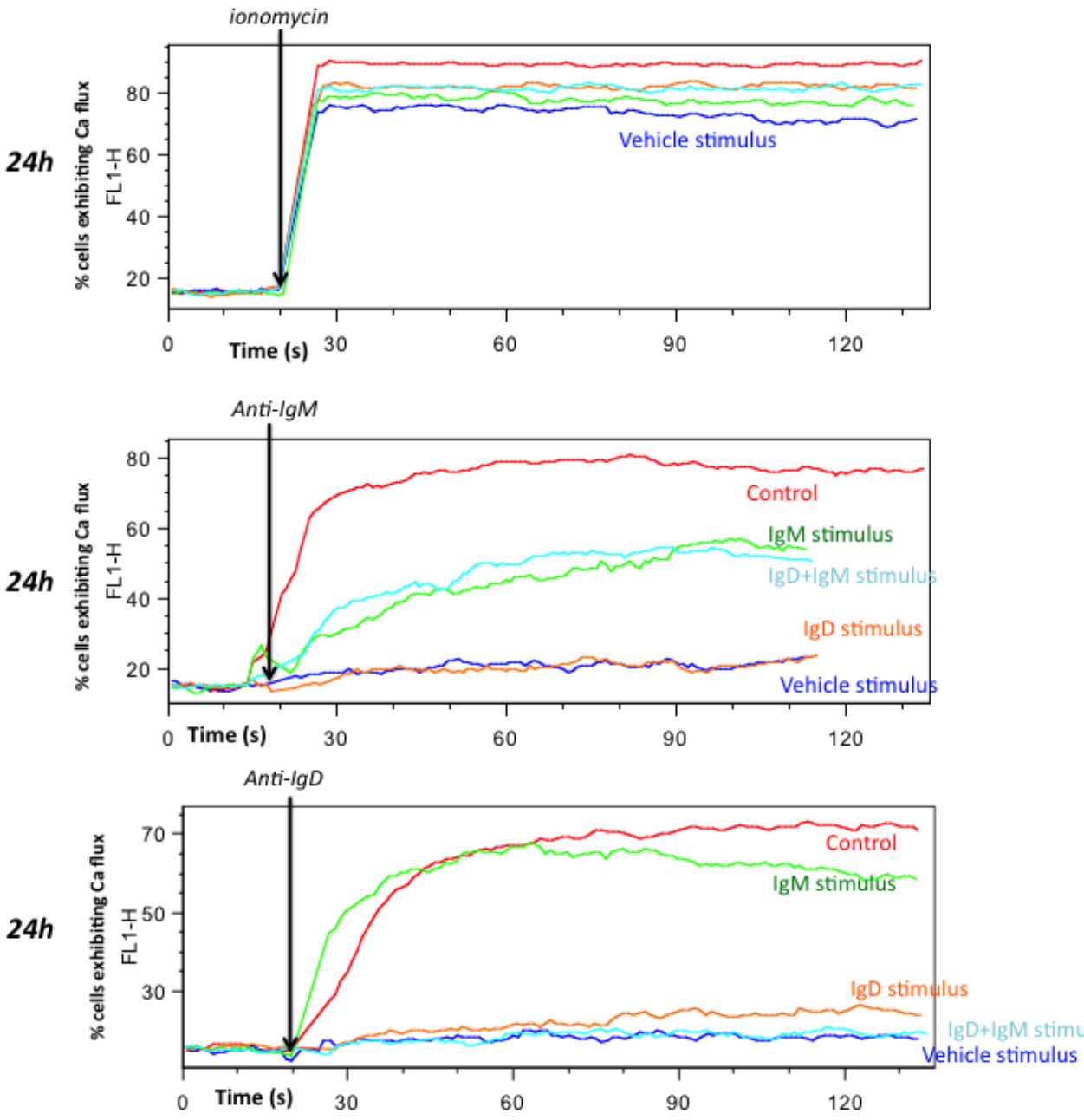

P1
baseline
stimulus
ionomycin
IgD+IgM
IgD
IgM
control
% cells exhibiting Ca flux
FL1-H
Time (s)
24h
medium
alone
stimulus
ionomycin
IgD
IgM
control
% cells exhibiting Ca flux
FL1-H
Time (s)
24h
ionomycin
Vehicle stimulus
% cells exhibiting Ca flux
FL1-H
Time (s)
24h
Anti-IgM
Control
IgD+IgM stimulus
Vehicle stimulus
IgD stimulus
IgM stimulus
% cells exhibiting Ca flux
FL1-H
Time (s)
24h
Anti-IgD
Control
Vehicle stimulus
IgM stimulus
IgD stimulus
IgD+IgM stimulus
% cells exhibiting Ca flux
FL1-H
Time (s)

**Figure 6.23: P5 viability and Ca flux after incubation**
**IgD+IgM or IgD treatment stimuli seems to result in a general increase in viability compared to the other treatments. The absence of response to anti-IgM at baseline is altered to high levels after 24h culture in medium alone. This recovery is impaired by vehicle or IgD treatment, though curiously not by IgD+IgM treatment. Also different is the recovery of ca flux response to anti-IgM after anti-IgM treatment, though the level is not as great as for medium alone.**



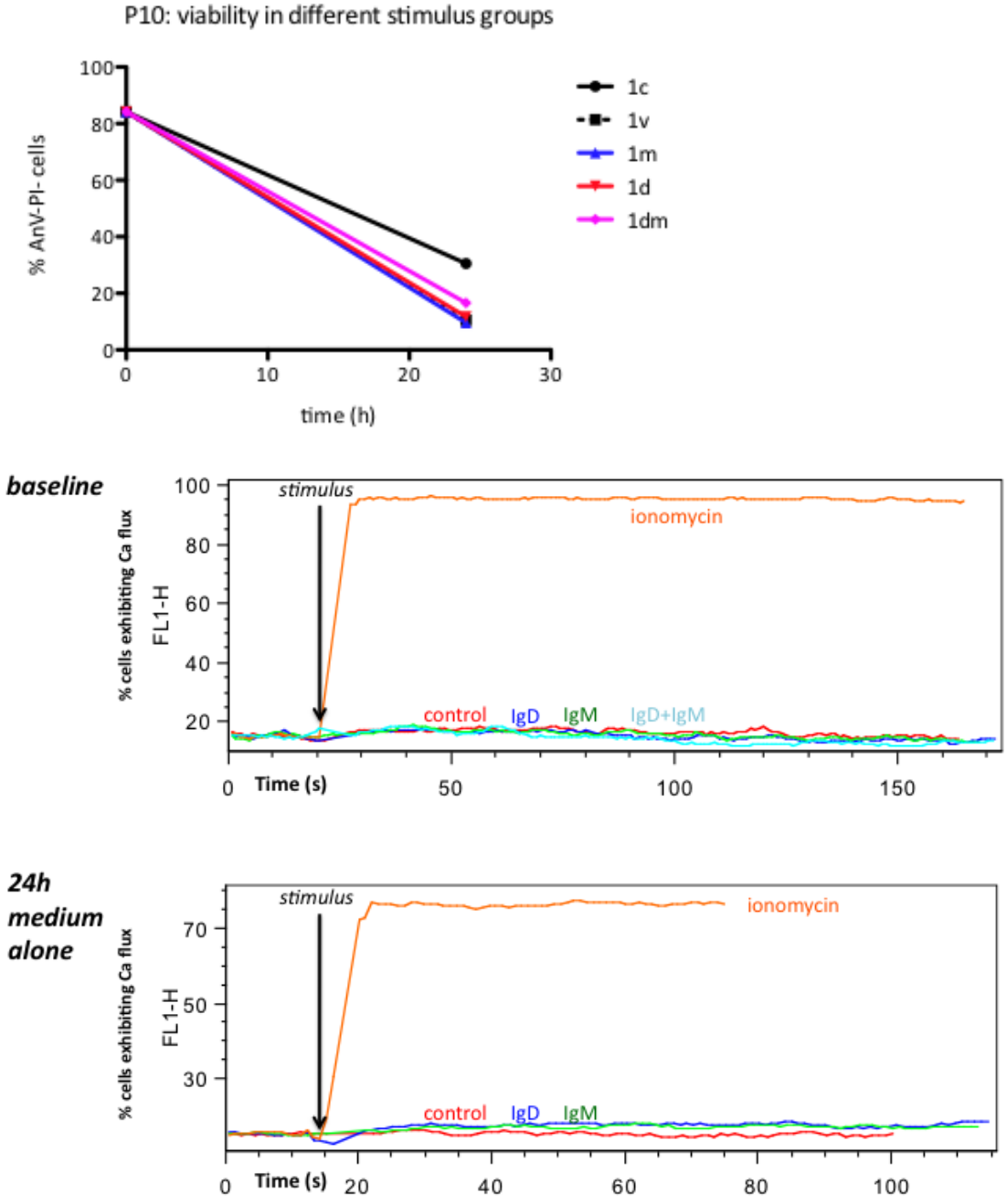






Finally, Sample P10 did not display any recovery in Ca flux after incubation *in vitro* (Fig 6.24). There was a slight increase in IgD expression (0% to 4%) over 24 hours, and a decrease in IgM expression.

**Figure 6.24: P10 viability and Ca flux after incubation**
**Treatment stimuli did not result in large differences in viability at 24h. The absence of response to IgM & IgD at baseline is unaltered after 24h culture in medium alone. Similarly, there is no recovery of ca flux with vehicle, IgD or IgM treatment.**



### 6.2.9 Summary and Discussion

Crosslinking the BCR with anti-IgM or IgD $F(ab')_2$ fragments produces robust Ca flux in the RL cell line and in the majority of primary CLL samples. The use of $F(ab')_2$ fragments is consistent with most of the published studies (see table 6.1), which used the same approach, and I used the same supplier that most frequently provided the reagents for these studies. However, the caveats and potentially less physiologic nature of soluble stimulus must be kept under consideration, particularly in the light of several studies that show increased survival and proliferation with anti-IgM antibodies that have been immobilised by conjugation to beads or culture plates (see below section 6.4.7)[320,321].

The calculation of $[Ca^{2+}]i$ in absolute terms (to provide a value in nM) is possible by a variety of methods (use of calcium buffers, or combination of dyes such as fluo-3 with Fura Red)[490]. I did not pursue an absolute quantification, as I sought to compare calcium flux within the context of studies that compared IgD and IgM flux, and amongst primary samples. Furthermore, the degree of fluorescence using the Fluo3 methods has a linear relationship with the cytosolic $[Ca^{2+}]i$ concentration when methods are compared[490]. The use of the $85^{th}$ centile is consistent with approaches used by others, and provides a quantitative measure of the proportion of responding cells within a sample. The relation of this to the Ca flux induced by ionomycin is felt to give an indication of the number of cells that have capacity to undergo Ca flux, taking into consideration variations in dye loading, temperature changes and cell viability. Whilst all samples exhibited viability greater than 90% using the trypan blue dye exclusion method (and most had levels >97%), the level of viability as estimated by dual AnV/PI negativity was less than this.

The capacity of most (18/20) CLL to exhibit Ca flux after IgD crosslinking was confirmed; and there was a correlation between the expression level of IgD and quantification of Ca flux. The two samples with zero IgD expression exhibited no Ca flux after crosslinking IgD, despite Ca flux with ionomycin. These samples also had absent Ca flux after IgM crosslinking (despite expression of IgM on >40% of cells). In contrast, reduced IgM-induced Ca flux was noted, both in terms of the number of samples showing any response (13/20) and the level of response of each sample, with

IgD Ca flux greater than IgM Ca flux in most samples studied. This was not simply an effect of the expression level of IgD being generally greater than IgM, as samples with high IgM expression did not necessarily exhibit Ca flux, and Ca flux after IgD ligation was often high in samples with relatively low IgD expression. Similarly, the absence of correlation between IgD and IgM Ca flux indicates that reduced IgM Ca flux is not simply due to defective Ca flux *per se*. This confirms the known phenomenon in CLL of capacity of the IgD BCR isotype to undergo the initial events in the BCR signalling cascade, coupled with an anergic pattern with IgM.

Furthermore, I have confirmed the findings of Mockridge et al in demonstrating the ability of CLL cells to regain capacity for signalling after *in vitro* culture in a limited number of samples[116]. This was partially suppressed by simulating antigen exposure. The reversal of ca flux signalling defects by *in vitro* incubation mimics anergic B-cell models. The addition of anti-IgD or -IgM $F(ab')_2$ fragments to the in vitro culture system, perhaps simulating the presence of antigen, partly inhibits this reversal of anergy, again similar to mouse models. These data confirm the findings of Mockridge et al, though I did not monitor signalling at further time points after 24h as viability in the studied samples was low. It would be desirable to extend these findings by evaluating recovery from anergy in more patient samples. For pragmatic reasons this was not pursued as large numbers of primary cells and reagents are required for each experiment, and Mockridge et al published larger numbers in any case. Together these data support the model of anergic BCR signalling in CLL: IgD expression and Ca flux is preserved, whilst IgM Ca flux is often absent despite high IgM expression.

In contrast to Mockridge et al, I did not find marked increases in BCR expression after *in vitro* incubation. My method differs from theirs in that I tried to gate expression on viable cells only, however one would expect any BCR expression increases to therefore be greater with my method. In their study, most increases in IgM were apparent at 48 hours rather than at the 24-hour timepoint I used, which may explain my findings. One sample (P10) that exhibited no expression of IgD seemed to slightly increase IgD expression to 3-4% after incubation; otherwise all samples exhibited a decrease in IgD expression after 24 hours. One sample (P1) exhibited a slight increase (6% initially, 10-

15% after 24h) in IgM expression, which accompanied an increase in Ca flux. IgM expression was similar in samples P5 and P10, and decreased slightly in P10. My data therefore suggest that changes in Ca flux at 24h are not accompanied by consistent changes in expression of IgM or IgD, implying that reversal of anergy is not due to increase of BCR expression. However, it could be claimed that comparing flow cytometric expression from experiments performed on cells with low viability on two different occasions may not be a valid quantification method, even if one considers the IgM/D expression method a valid one for relative quantification of BCR expression within a single experiment (see chapter 5).

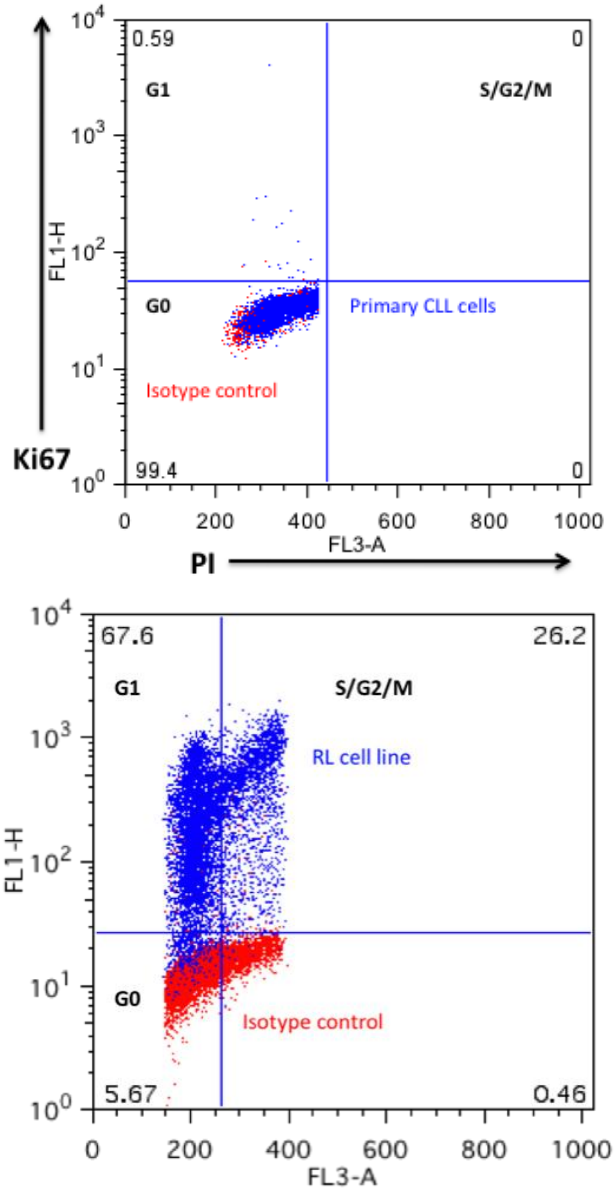
In summary, the early events in BCR signalling were evaluated by measuring Ca flux in 20 primary CLL samples. I confirmed the general capacity for IgD signalling in the presence of reduced IgM signalling, mimicking the anergic models of B-cell anergy. Incubation of cells *in vitro* for 24h can lead to recovery of IgM signalling capacity, and this phenomenon is seen to an extent with IgD. This is consistent with a model whereby the expression and signalling capacity of IgM (and IgD) is a dynamic phenomenon in CLL. This model suggests that antigen engagement *in vivo* may be occurring in the lymph node microenvironment, resulting in an anergic phenotype, and that removal of the putative antigen during *in vitro* incubation results in recovery of BCR signalling capacity.

However, I was unable to show consistent alterations in expression of IgM or IgD to accompany this reversal of anergy, which may suggest that the technique for measuring expression in the experiment is insufficiently robust, or that anergy is dependent on more than downregulated BCR expression. In support of the second interpretation, the absence of Ca flux after IgM ligation is common in CLL samples cases with apparent high IgM expression. I therefore suggest that anergy is partly a result of ‘disconnection’ of downstream signalling pathways from IgM.

## 6.3 Proliferation of CLL cells after BCR ligation

Cell cycle analysis of cell lines and primary leukaemia cells was performed (see figure 6.25). CLL samples were stimulated with anti-IgD or anti-IgM $F(ab')_2$ fragments and compared to vehicle control at 24 and 48 hours, in triplicates. CLL cells reliably remained in $G_0$ (89-99%) despite stimulus with IgM and IgD at 24 and 48 hours after stimulation. Only 5 patient samples were stimulated, as no effects were seen.

**Figure 6.25: Cell cycle analysis in RL cell line and primary CLL.**
**Representative examples. CLL cells were fixed, permeabilised and stained with anti-Ki67-FITC antibody and Propidium Iodide (PI). Cells positive for Ki67 alone were defined as in G1 stage of cell cycle, Ki67-PI dual positive as S/G2/M and dual negative in G0. CLL cells remain in Go, whereas the cell line is predominantly in G1 or S/G2/M stages. Quadrants based on isotype controls**



PB CLL cells are generally in a quiescent stage of the cell cycle with typically greater than 90% of cells in $G_0$[491]. Whilst BCR stimulus via IgM or IgD had differing effects on the RL cell line (IgM and IgD stimulation causing shift from $G_1$ and $G_2$/S into $G_0$, IgM>IgD), CLL cells reliably remained in $G_0$ despite different stimuli in 5/5 patient samples tested. This is generally in keeping with the published literature, whereby CLL cell entry into $G_1$ is only caused by BCR stimulation in conjunction with cytokine stimulation (such as IL-4), stromal cell support, or nonphysiological stimuli such as phorbol 12-myristate 13-acetate (PMA)[472,491-498]. In contrast to my own data and that of most of the published literature, there are a few publications where a significant fraction of CLL cells *in vitro* enter the cell cycle after BCR stimulus alone[319,475]. These discrepancies may be due to methodological differences, the characteristics of the patients, or other unknown factors. My data are in keeping with most of the literature. Further attempts to induce changes in CLL cell proliferation with BCR stimulus in other conditions were felt to be outside the scope of this study, and were not pursued.

## 6.4 Apoptosis after BCR ligation in CLL

### 6.4.1 Introduction

The term apoptosis was first used in 1972 to describe a morphologically distinct form of cell death[499]. Apoptosis has become acknowledged as a mode of 'programmed' cell death, occurring normally during development and ageing. It also acts as a homeostatic mechanism to maintain cell populations in tissues and as a defence mechanism when cells are damaged[500]. Several methods exist for the detection of apoptosis[500]. Annexin V assays have been most widely used in studies of CLL cell apoptosis[306,319,320,372,423,501-503].

Cells undergoing apoptosis translocate phosphatidylserine (PS) to the outer membrane under certain conditions by the action of scramblases[504]. Annexin V (AnV) is a 35-36kDa, calcium dependent phospholipid-binding protein that binds specifically to PS.[505] Conjugation of AnV to fluorophores such as Fluorescein Isothiocyanate (FITC) can allow

labelling of apoptotic cells using flow cytometry or microscopy. Addition of a non-permeable DNA binding dye such as Propidium Iodide (PI) can allow discrimination of apoptotic cells from necrotic/late apoptotic cells that have both exposed intracellular DNA and exposed PS. The technique has high sensitivity, being able to detect single apoptotic cells[500].

Multiple authors have examined apoptosis as an endpoint after BCR stimulation in CLL[62,307,317,319,320]. These have mostly addressed signalling through IgM, by crosslinking IgM with antibody or $F(ab')_2$ fragments. Given the data regarding the role of the BCR in CLL, the obvious hypothesis is that BCR stimulation will result in a decrease in apoptosis. However, this is a field that contains often-contradictory results. Some authors find that IgM stimulation results in decreased apoptosis when measured 24-48 hours after stimulation[62,320,341], whereas others find the opposite: a consistent induction of apoptosis with IgM stimulation[307,322,423]. To further complicate matters, some studies find reduced apoptosis with certain patient samples, no change with some and increased apoptosis with other patients[319]. These differing outcomes have been linked to prognostic factors such as *IGHV* gene mutation status, CD38 or ZAP70 positivity, and the apoptotic outcome after IgM stimulation may have prognostic impact in itself[319].

To an extent, these divergent responses to IgM crosslinking have been accounted for by the different protocols used, in particular the use of soluble *vs* immobilised anti-IgM antibodies, but these do not entirely explain the divergent responses. Many studies do not account for the presence or absence of proximal BCR signalling events such as Ca flux, and given the often anergic responses of CLL to IgM stimulation, these may well have an impact. These considerations are explored in greater depth in section 6.4.7.

Only one group has studied the effect of IgD stimulation on CLL apoptosis, but methodological details are occasionally obscure in the relevant paper[306]. Zupo et al studied 10 CD38 positive CLL patients, as they had previously shown that IgM-induced apoptosis and Ca flux was more prevalent in $CD38^+$ cells. They utilised both polyclonal goat/rabbit antibodies and $F(ab')_2$ fragments to IgM or IgD, but state that the use of antibody *vs* $F(ab')_2$ fragments made no difference. Their control was normal goat

immunoglobulin, but no details on the nature of this are provided. Anti-IgM consistently induced apoptosis at periods 24 hours to 5 days after stimulus, whereas anti-IgD resulted in no difference from control at early timepoints, and in some patients, IgD stimulation resulted in lower apoptosis after 3 days of culture. In three patients studied, there was no change in cell proliferation with either stimulus, including in combination with interleukins 2,4,6 or 10. However, the combination of interleukin 2 with IgD stimulation resulted in apparent plasma cell differentiation and production of secreted immunoglobulin after 5 days of culture, an effect not seen in control culture or IgM stimulation.

These studies suggest that the effects of IgD and IgM stimulation result from differing signalling pathways activated by the two isotypes, with IgD ligation activating pro-survival pathways and the authors attribute this effect to the neoplastic transformation of B-cells at a particular stage of differentiation where IgD produces a predominant anti-apoptotic effect as compared to IgM. Therefore, there are suggestions that IgD and IgM signalling play differing roles in CLL, at least in terms of cell survival in an *in vitro* system.

The same group have subsequently explored IgD signalling in a larger cohort of 106 patients[506], but methodological details are not stated in this study. Apoptosis was measured (by PI exclusion only) 48h after anti-IgD antibody. In this context, responses were divided into three groups, without a consistent pro-survival effect of IgD. Group I (n=33) displayed inhibition of spontaneous apoptosis, Group II (n=8) exhibited increased apoptosis, whilst group III (n=65) did not exhibit changes in apoptosis (greater than 20% change in viability). Groups I and II were combined and called 'responders', and this group had a significantly shorter treatment free survival compared to non-responders. This was also independent of ZAP70, CD38 or *IGHV* gene mutation status. Responders also had greater levels of tyrosine phosphorylation and surface IgD expression. The authors also state (but do not show data) that IgM ligation in the same cohort results in apoptotic responses that also correlate with prognosis, though this was not an independent prognostic factor on multivariate analysis. These findings are analogous to another study that linked the apoptotic response of IgM

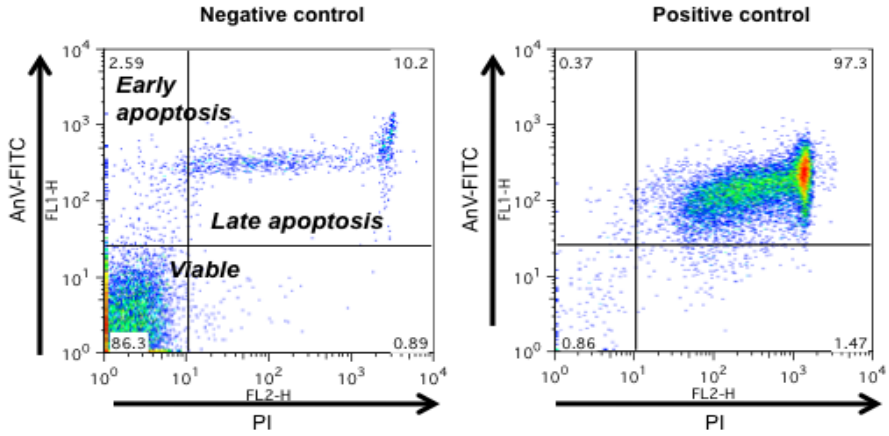
signalling to prognosis[319]. Nedellec et al studied 41 cases and showed that these could be divided into three groups depending on the outcome of IgM crosslinking: Group 1 (n=15) demonstrated no Ca flux and no changes in apoptosis. Group 2a (n=7) showed capacity for Ca flux but decreased apoptosis, whereas Group 2b (n=19) showed increased apoptosis in addition to Ca flux. Poor prognostic markers were associated with Group 2.

In summary, data on the effect of IgM or IgD stimulation in CLL are often contradictory. In a limited number of cases there may be a difference in terms of IgD *vs* IgM stimulation. The effects on cell survival have not been consistently linked to Ig expression or Ca flux responses. I aimed to clarify the role of IgD and IgM signalling in my characterised cohort of patients, and relate apoptotic responses to expression of these two isotypes and the Ca flux responses to ligation.

### 6.4.2 Methods

Primary CLL cells were thawed, rested for 2-3 hours and levels of apoptosis measured as a baseline. They were then stimulated with vehicle borate buffer, IgM (10μg/mL), IgD (10μg/mL) or IgD+IgM (5μg/mL each). Apoptosis was then measured at 24 and 48 hours after stimulation. Each sample is stimulated in triplicate. The AnV/PI method (see fig 6.26) classifies cells as either viable ($AnV^-PI^-$), early apoptosis ($AnV^+PI^-$) or late apoptosis ($AnV^+PI^+$). For statistical comparisons, the percentage of viable cells was compared.

**Figure 6.26: The Annexin V/Propidium Iodide (AnV/PI) method for evaluating apoptosis.**
**Shown is a representative plot.**




18 of the 20 characterised patients were studied using this method.

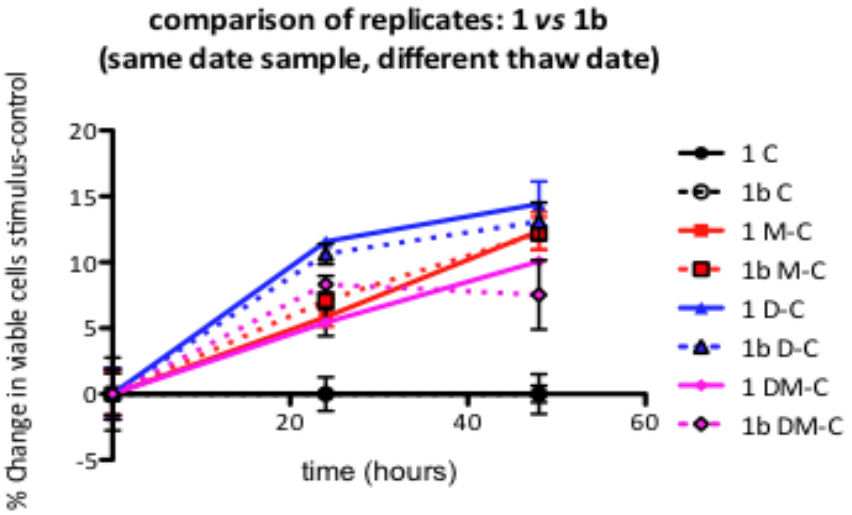
In order to estimate the inherent variability of this experimental setup, stimulations were repeated in two patient samples (based on sample availability). Two patients (P1 and P7) had repeated experiments, one repeat from the same sample thawed and stimulated on different occasions, the other from different samples taken from a patient 3 months apart (Fig 6.27). As shown, IgD and IgM stimulus generally resulted in significant improvements in viability, whilst there are no differences when the repeats are compared. The lack of differences was apparent when both more powerful yet less appropriate parametric (t-test) and non-parametric (Mann Whitney) comparisons were made.

**Figure 6.27: Comparison of apoptosis levels 24 and 48 hours after different stimuli in the same patient samples.**

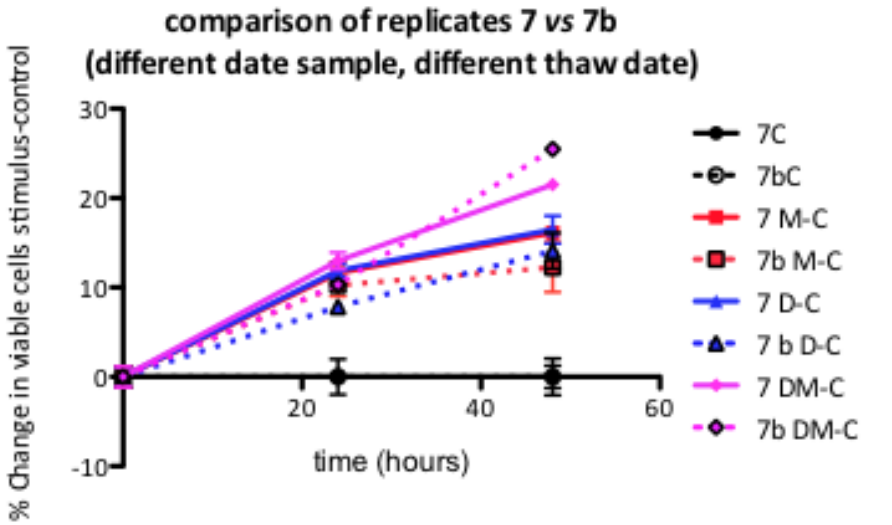
**Shown are the mean percentages of viable cells, subtracted from the values of vehicle control, with error bars representing standard deviation.**

**C= vehicle control, M-C= IgM stimulus *vs* control, D-C= IgD stimulus *vs* control, DM-C, IgM +IgD stimulus *vs* control**

**For each comparison (t-test) of the same stimulus in the repeats, there was no significant difference at p<0.05. Not shown are the p values comparing each stimulus to control, which are significant for all stimuli.**




*unpaired t-test: no significant differences between 7 and 7b at any time point*

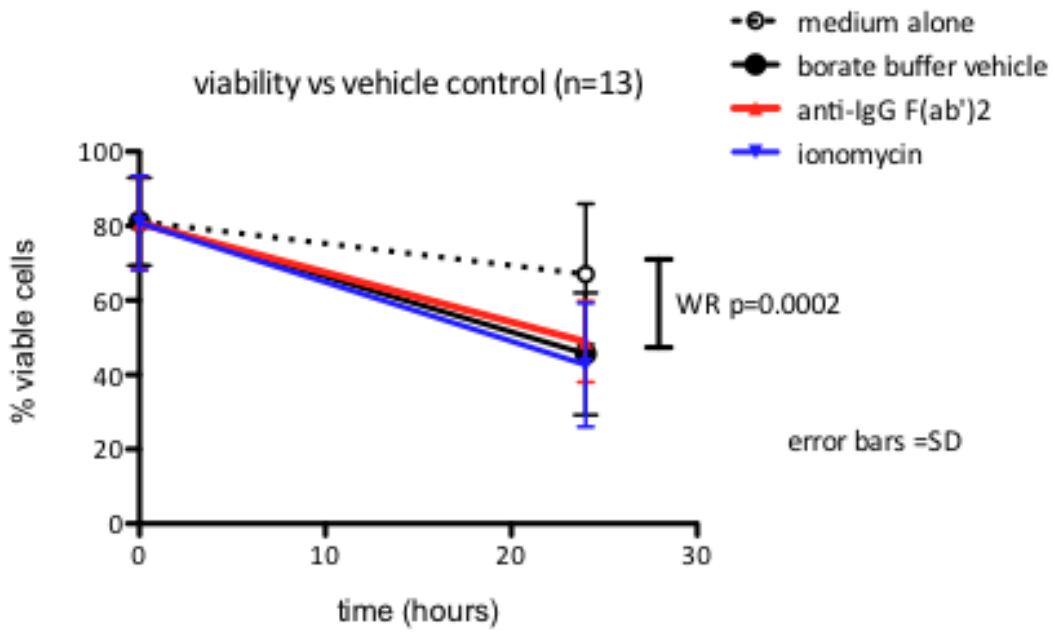



*unpaired t-test: no significant differences between 7 and 7b at any time point*

### 6.4.3 Results: Borate buffer vehicle control has an effect on CLL apoptosis

The anti-IgD and anti-IgM $F(ab')_2$ stimuli are suspended in sterile borate buffer composed of sodium tetraborate and boric acid at a pH of 8.2. The pH of RPMI is 8, which falls slightly when incubated in 5% $CO_2$. Upon addition of anti-IgM/D in borate buffer at 10μg/mL, there is a slight noticeable colour change in the medium, presumably as a result of pH change. There was therefore concern that the borate buffer may affect the physiology of the CLL cells. In order to verify this, the effect of stimulation with anti-IgD/M was compared with various control conditions. In many cases, comparison of anti-IgD/M stimulation with medium alone resulted in increased apoptosis. However, when sterile borate buffer was used as the control condition, this resulted in a general relative increase in viability after anti-IgD/M stimulus. There was therefore the suspicion that the borate vehicle has some toxicity. This was confirmed by experiments comparing different conditions. Control anti-human IgG goat $F(ab')_2$ in the same medium was compared to borate buffer, ionomycin in borate buffer, and medium alone. Each of the stimuli using borate buffer (buffer alone, ionomycin, anti-IgG) resulted in a significant decrease in viability (see figure 6.28). Since the effect of IgD/M stimulus would therefore be a combination of increased apoptosis as a result of toxicity of the vehicle in addition to any effect of IgD/M itself, all experiments subsequently used sterile borate buffer as a control. In early experiments, anti-IgG $F(ab')_2$ was also used, but one patient sample (not amongst the 20 characterised used) was noted to exhibit Ca flux after this stimulus (presumably as a result of IgG class switching or adsorbed IgG), and so this control condition was not used for apoptosis experiments. Note that there was no effect on Ca flux of the borate buffer.

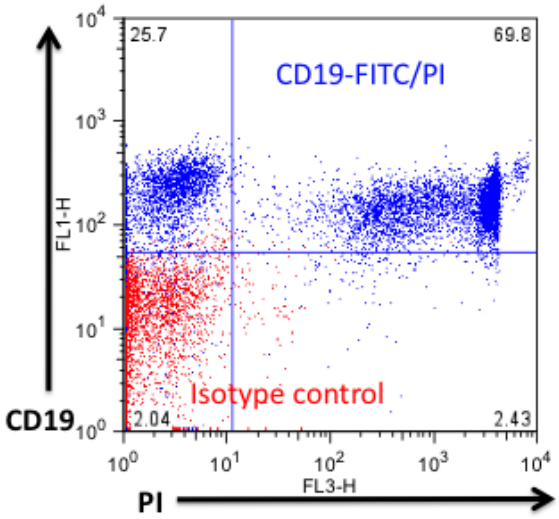
**Figure 6.28: Comparison of viability at 24h after stimulation with a variety of conditions.**
**Error bars show standard deviations.**

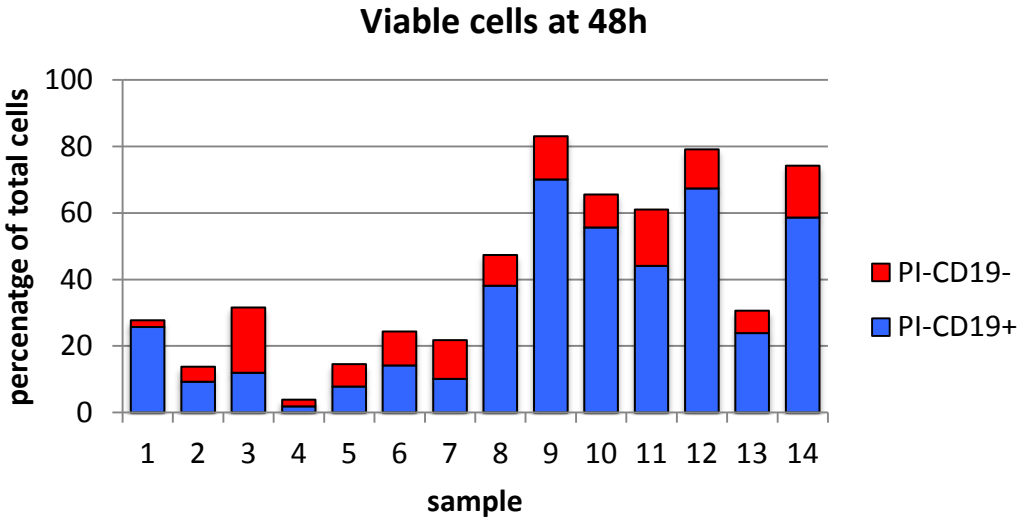



### 6.4.4 Results: Some residual viable cells after 48 hours of culture are non-B-cells

Whilst all samples were formed of a majority of $CD5^{+}CD19^{+}$ CLL B-cells at baseline (all >87%, the majority>95%), this may not be true after 24-48h in culture. To estimate the proportion of B-cells amongst viable cells remaining after 48h, staining with CD19-FITC and propidium iodide was also performed (in 14/18 samples), and the proportion of viable (PI negative) that were also CD19 positive was calculated (see figure 6.29). The mean proportion of $CD19^{+}$ viable cells at 48h was 69% (range 37-92%), indicating that whilst B-cells generally formed the majority of cells alive after incubation *in vitro*, in some cases the remaining viable population was often comprised of a large proportion of non-B-cells. This measure of viability will include AnV positive cells, and so is not the same measure of viability as used in most of my experiments examining apoptosis. Interestingly, PI positive cells are generally also CD19 positive.

**Figure 6.29: Estimation of proportion of viable cells that are also B-cells at 48h. Shown is a representative plot, with quadrants drawn at the 99$^{th}$ centile for isotype controls. Shown below are the proportions of viable (PI negative) cells at 48h, in blue are the viable B-cells also CD19$^{+}$, in red the viable CD19$^{-}$ cells.**



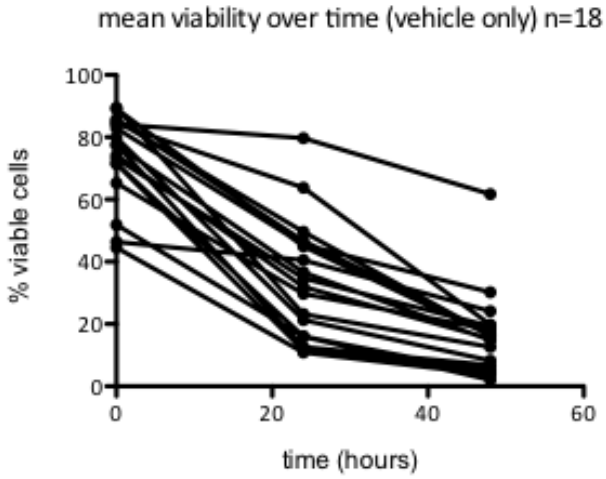




### 6.4.5 BCR stimulus via IgM and IgD results in decreased apoptosis

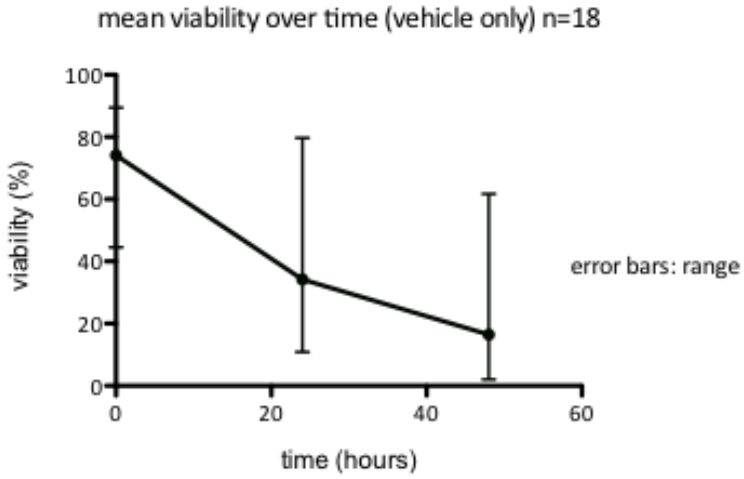
In general, stimulation with anti-BCR antibodies led to increase in viability.

Vehicle control stimulated cells had the typical decrease in viability over time (Fig 6.30):

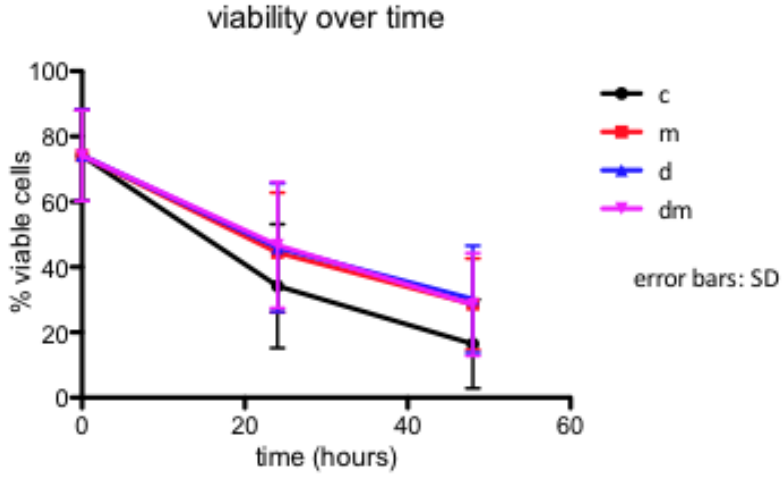
**Figure 6.30: Proportions of viable cells at different times after stimulation.**
**a) The mean values for all 18 samples in the control condition**




**b) the derived mean for these controls**

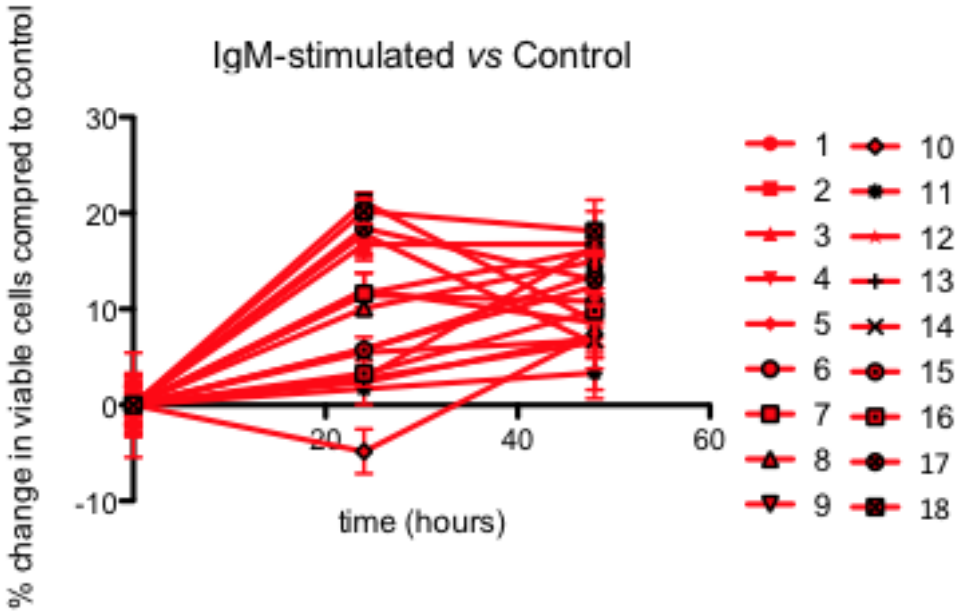



**c) the mean values of all 18 samples for each of the different stimuli (c=control vehicle, m=anti-IgM, d=anti-IgD, dm=anti-IgD+anti-IgM)**



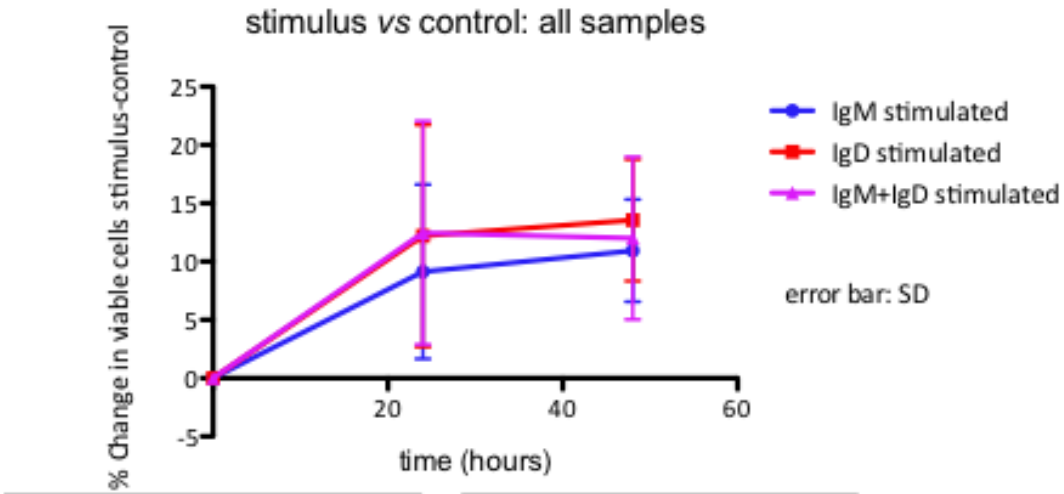
The differences between stimulated and control is highlighted when the percentage of cells is subtracted from the mean control value at that time point (see Fig 6.31). Heterogeneity of response is seen.

**Figure 6.31: The effect of IgM stimulation on viability.**
**Shown are the mean viabilities of each IgM-stimulated sample subtracted from the vehicle control means. Error bars represent standard deviations of each triplicate.**




The most apparent effect is that IgD or IgM stimulus results in a decrease in apoptosis as compared to vehicle control (fig 6.32). However, when IgD is compared to IgM stimulation, there is a small but significant difference, with IgD stimulation resulting in greater viability.

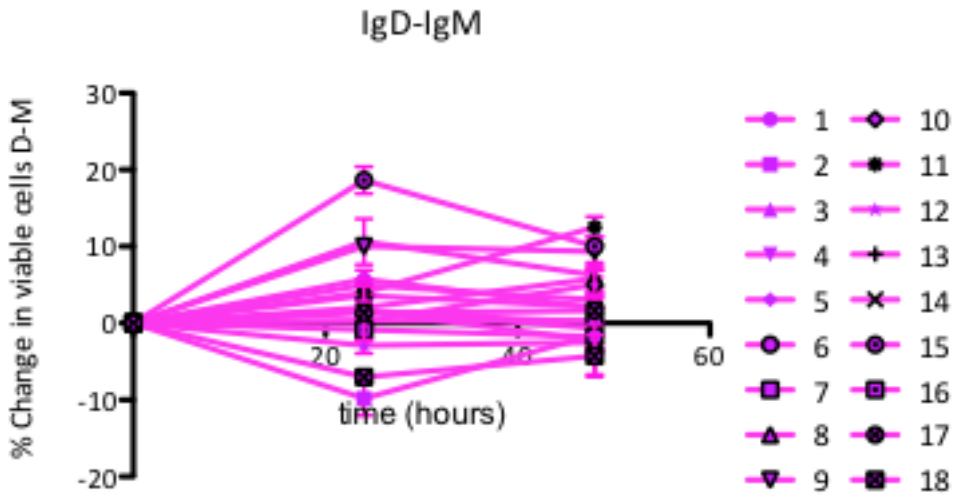
**Figure 6.32: Mean viabilities for all samples combined**
**Viability after stimulus subtracted by viability with vehicle control. P values are derived from comparisons using the wilcoxon signed rank test.**




| 24h | 48h |
|---|---|
| **IgM *vs* Control p=0.0006** | **IgM *vs* Control p=0.0002** |
| **IgD *vs* Control p=0.0006** | **IgD *vs* Control p=0.0002** |
| **IgD+M *vs* Control p=0.0004** | **IgD+M *vs* Control p=0.0003** |
| **IgM *vs* IgD p=0.018** | IgM *vs* IgD p=0.058 |
| **IgM *vs* IgD+M p=0.010** | IgM *vs* IgD+M p=0.28 |
| IgD *vs* IgD+M p=0.92 | IgD *vs* IgD+M p=0.37 |

This reflects the underlying heterogeneity of responses (fig6.33):

**Figure 6.33 Comparison of mean viabilities for each sample, IgD stimulated–IgM stimulated.**

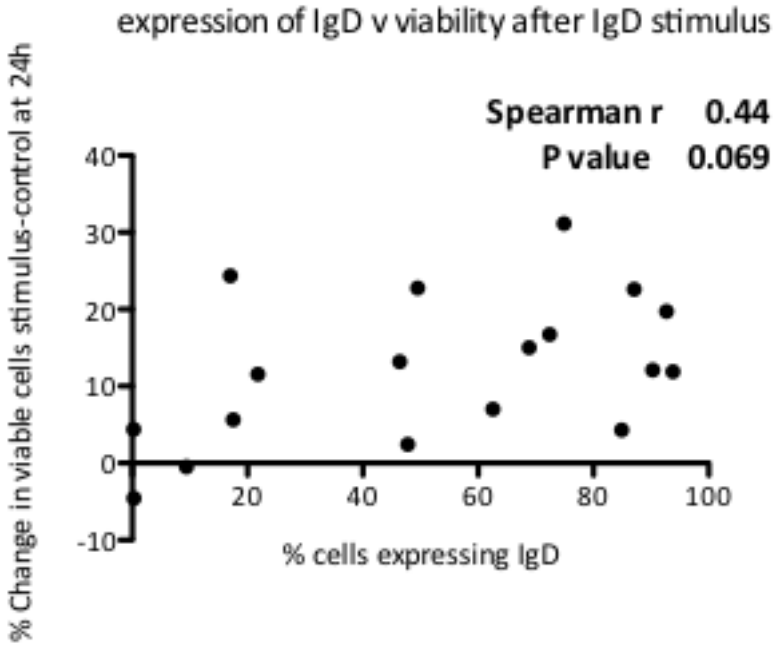



The reasons for this heterogeneity were explored further. Higher expression of IgD generally led to higher viability after IgD stimulus at 24h, with low/zero expression of IgD being associated with no change in viability compared to control (Fig 6.34).

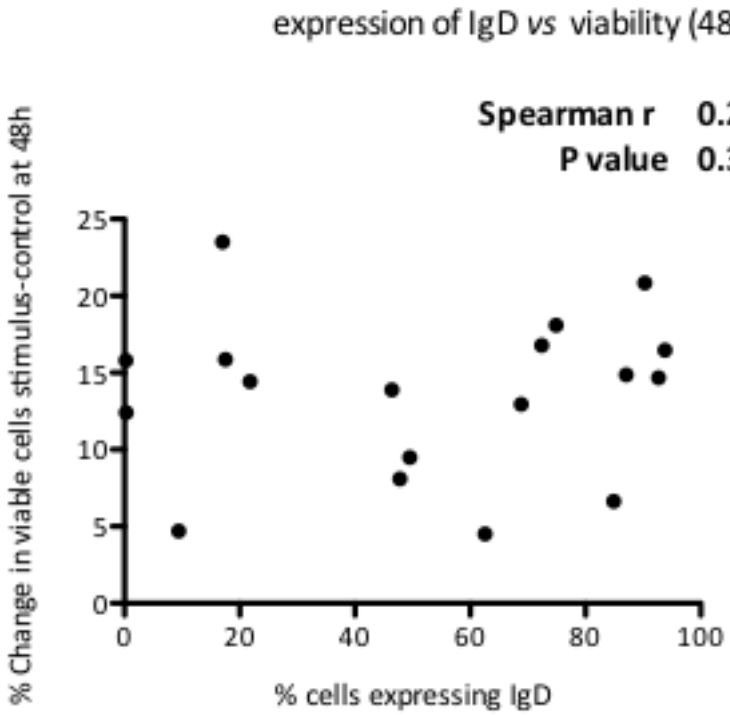
However, the correlation between IgD expression and viability after IgD stimulation was not significant at the $p<0.05$ level, and any relationship is weaker after 48h.

**Figure 6.34: Percentage expression IgD *vs* mean viability change after IgD stimulation compared to control**

**24h**

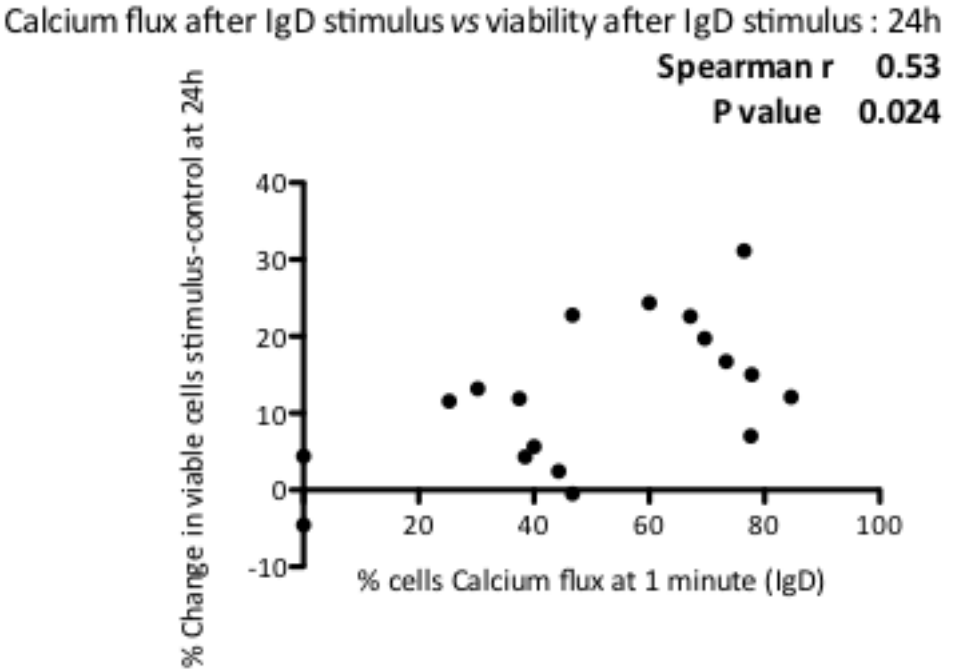



**48h**

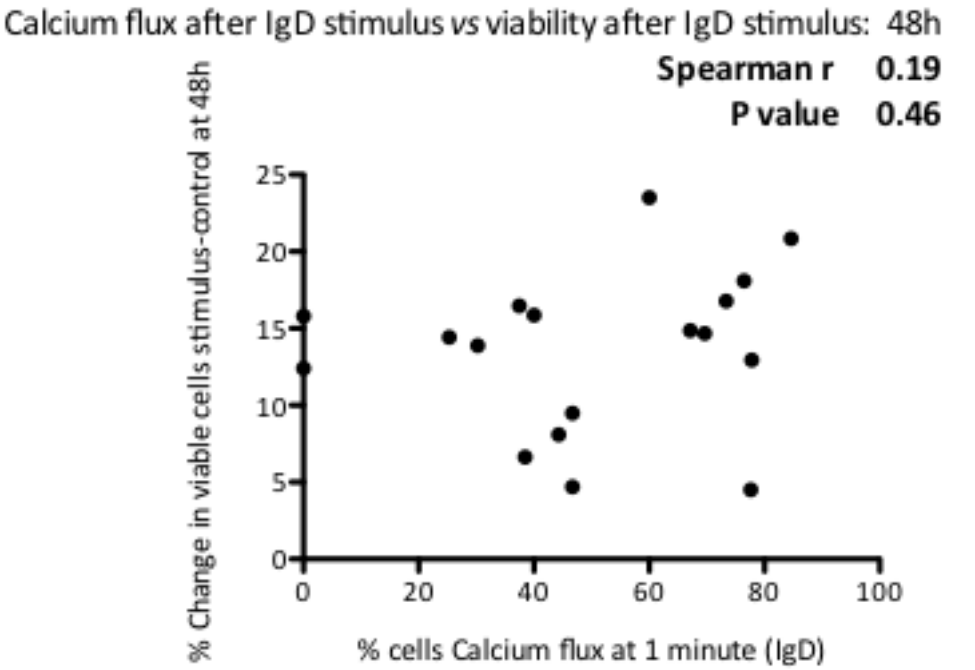


Despite this increased calcium flux after IgD stimulus is significantly correlated with viability at 24h (fig 6.35), though the relationship is lost after 48h:

**Figure 6.35: Mean change in viability after IgD stimulus compared to Ca flux after IgD stimulus**

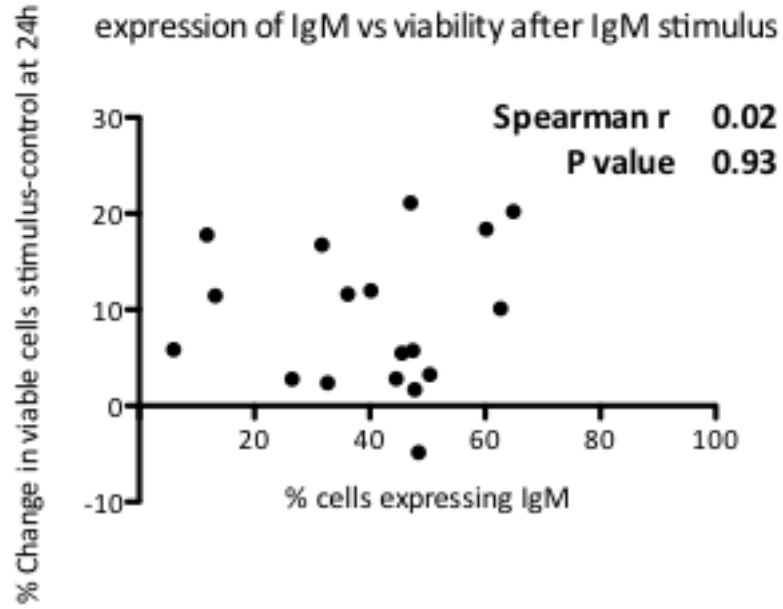


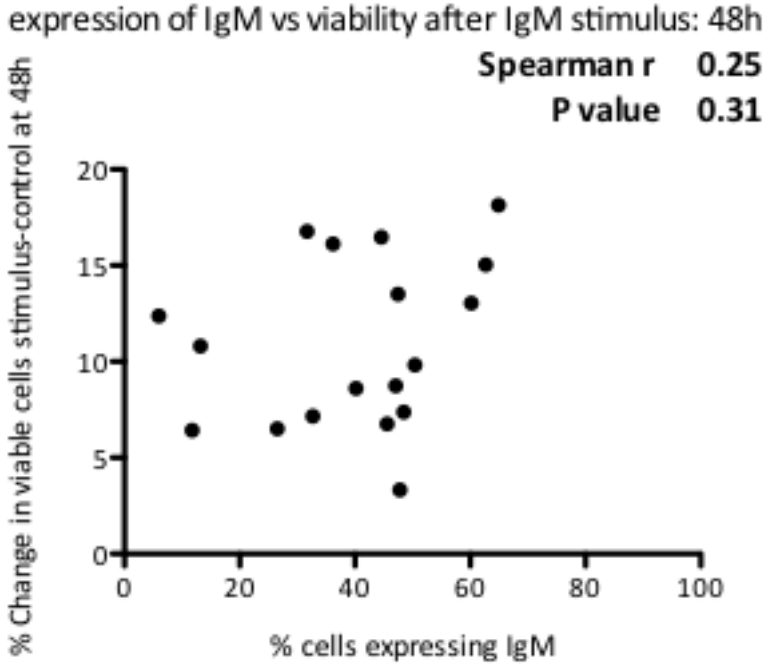

**48h**

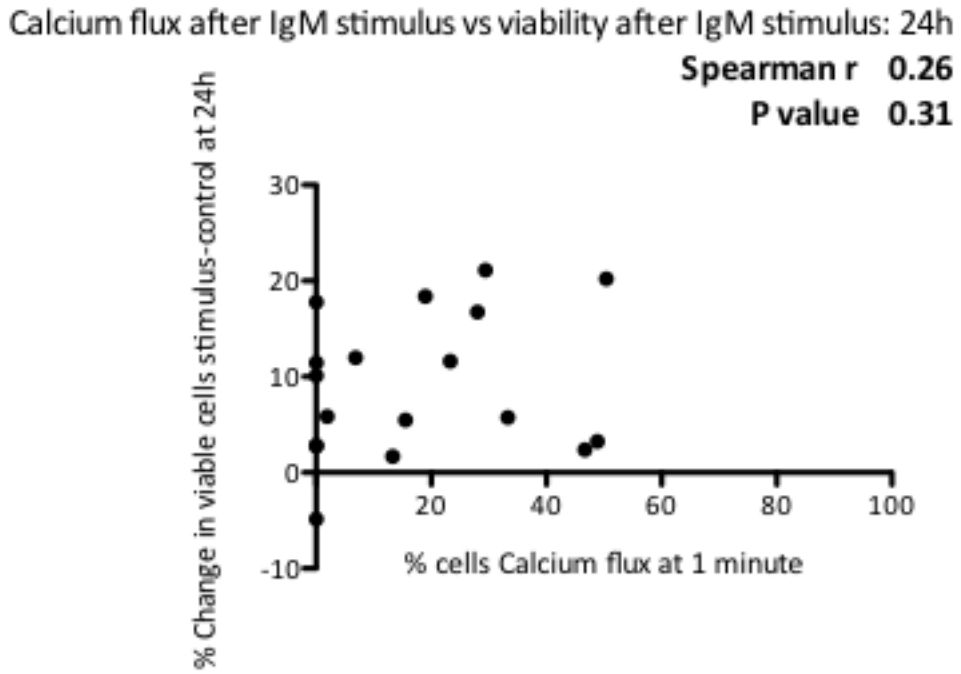


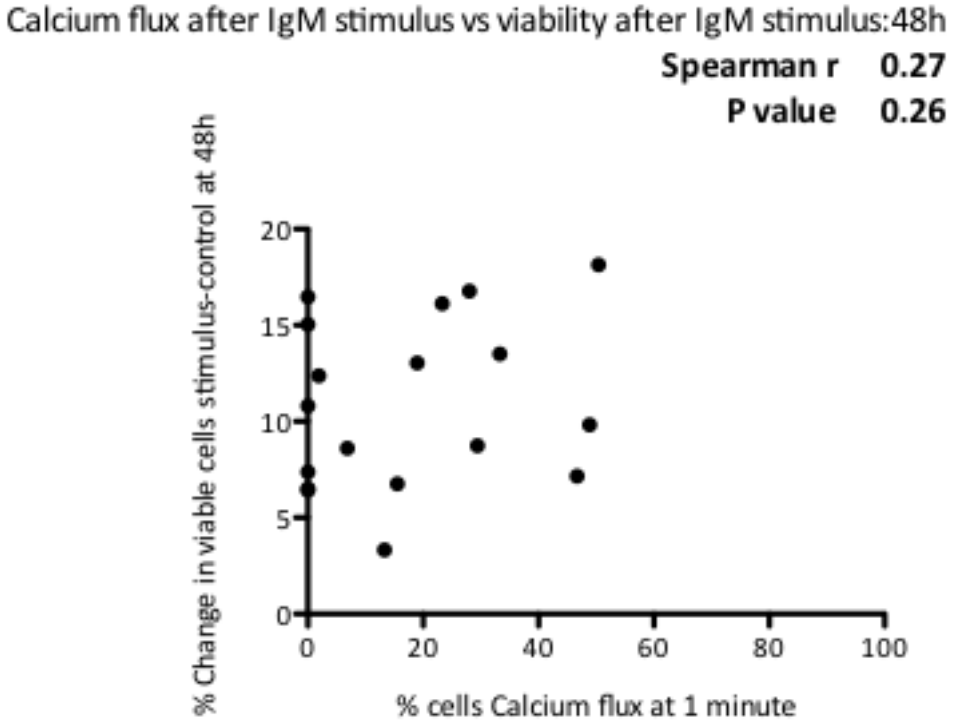
Conversely, IgM expression and calcium flux was not correlated with viability after stimulus and some cells increased viability in the absence of ca flux (Fig 6.36):

**Figure 6.36: Percentage expression IgM *vs* mean viability change after IgM stimulation compared to control.**
**Also shown is the mean change in viability after IgM stimulus compared to Ca flux after IgM stimulus**



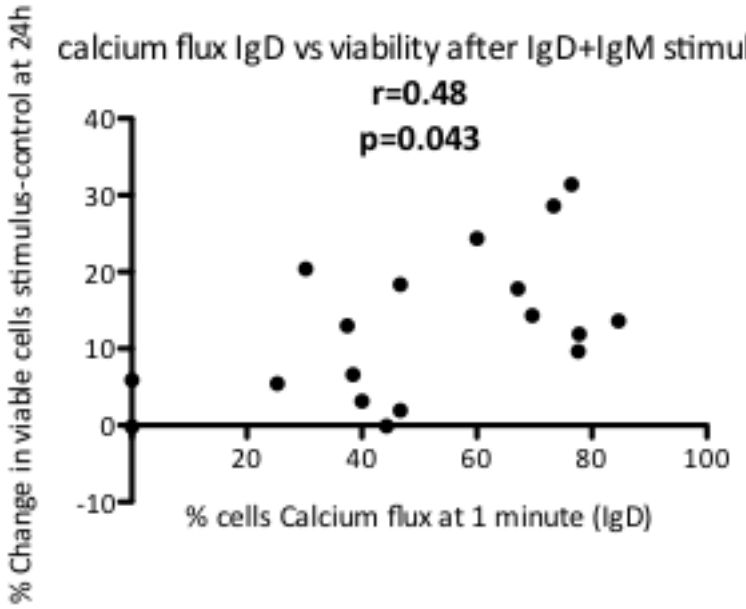



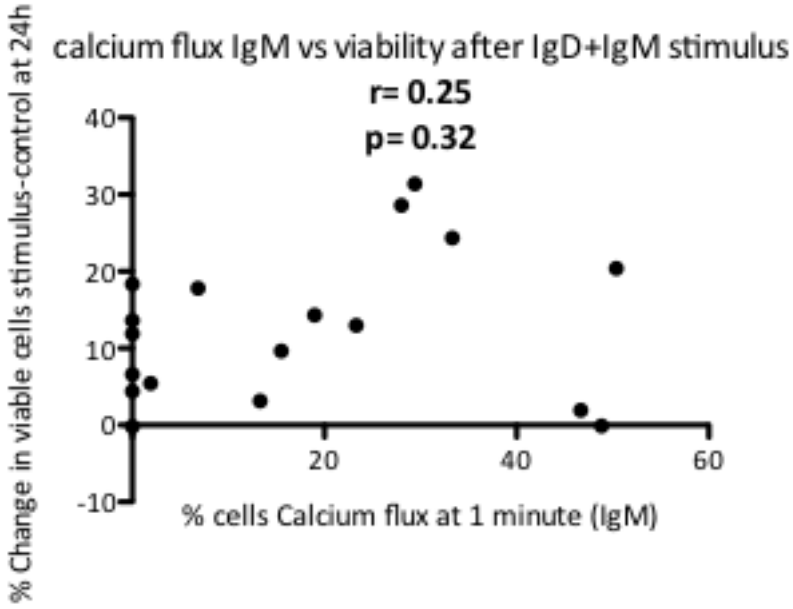



Finally, dual stimulation with anti-IgD+anti-IgM also led to increased survival, and this was more correlated with IgD signalling (Fig 6.37):

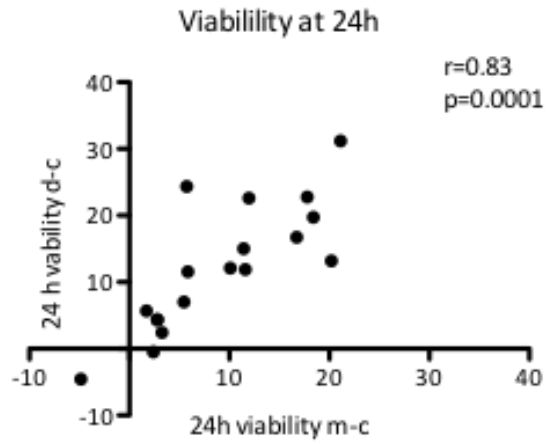
**Figure 6.37 Change in viability after IgD+IgM stimulus versus Ca flux after IgD or IgM stimulus**

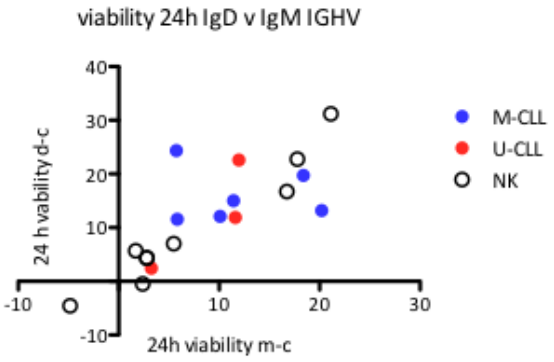


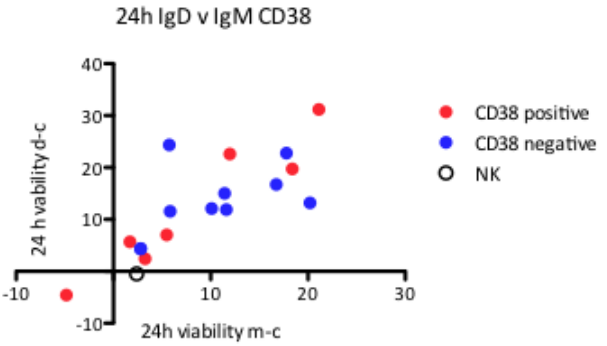



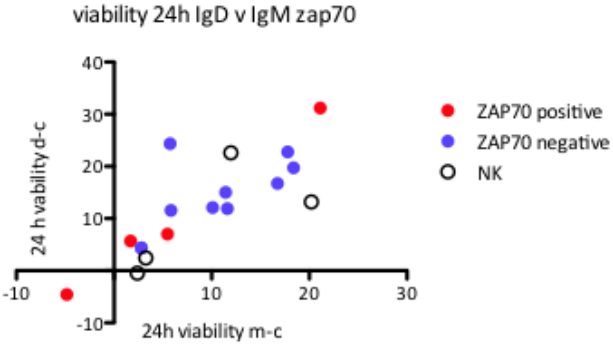

There was no obvious association between stimulus-induced viability and *IGHV* gene mutation, CD38 or ZAP70 status. Finally, there was a strong correlation (Spearman r=0.83, p = 0.0001) between the ability of a sample to respond to either IgM or IgD stimulus with an increase in viability (see Fig 6.38). This was not associated with *IGHV* mutation status, CD38 or ZAP70 positivity.

**Figure 6.38: Change in viability after IgD or IgM stimulus. Below is shown the data according to prognostic factor.**



### 6.4.6 Summary and Discussion: Apoptosis after BCR ligation

Apoptosis was assayed using the Annexin V-Propidium Iodide method and comparisons made using the percentage of viable cells present at 24 and 48 hours after stimulation with anti-IgD or IgM $F(ab')_2$ fragments. When frozen samples from the same patient are thawed and stimulated on different occasions, the same pattern of responses is seen, and this holds for samples from the same patient taken and frozen on different dates, though this repeat experiment was only attempted with two patients. This suggests that a particular patient sample responds in the same way.

Whilst many studies examine CLL cell apoptosis after BCR ligation, it is not always possible to compare methods. The most common methods involve Propidium Iodide (PI) exclusion, often in combination with Annexin V (AnV) staining and examination by flow cytometry. Many reports showed data comparing PI positive (designated apoptotic) *vs* PI negative (designated viable)[62,307,317]. Others compared AnV positive (designated apoptotic) *vs* AnV negative (designated viable)[62,321]. Some appeared to use combinations of AnV and PI, but it is impossible to judge from the published data whether they designate AnV positivity alone or $AnV^+PI^+$ cells as apoptotic.[317] Many studies did not account for the presence or absence of proximal BCR signalling events such as Ca flux, and given the often anergic responses of CLL to IgM stimulation, these may well have an impact

Though there is much variation and it is sometimes difficult to directly compare results, my findings regarding the general decline of CLL cell viability in medium alone are in keeping with published results[306,319,320,372,423]. Whilst I have studied a limited number of samples, the size of the cohort is similar or greater than those of published series

In general, stimulation with IgD, IgM or a combination of both led to small but significant increases in cell survival at 24 and 48 hours. A small difference between IgD and IgM is also seen, with IgD appearing to promote survival at greater levels than IgM, though the average effect is relatively small.

When the amount of improvement in viability is related to expression of the relevant isotype, there was no statistically significant correlation with IgD or IgM expression. A significant correlation exists between IgD-induced Ca flux and viability at 24 hours after IgD or IgD+IgM stimulation, the only suggestion that relates the capacity of BCR signalling via IgD to a role for IgD on promoting CLL cell survival. There are puzzling findings that improvements in viability after IgM stimulation may occur in cells that have low expression of IgM or exhibit no Ca flux with IgM. One explanation for this may be that Ca flux is not required for the pro-survival effects of IgM. However, given the findings of recovery of IgM Ca flux after incubation *in vitro*, an alternative explanation may be offered. I speculate that a period of incubation permits IgM signalling to recover over 24-48h, at which time binding to the $F(ab')_2$ stimulus may occur, resulting in a survival signal. This may occur to a lesser extent with IgD

What is clear is that both IgD and IgM stimulation exert a pro-survival effect. However, the degree of viability change does not correlate in a straightforward fashion with the expression of the BCR isotype, or the Ca flux induced after stimulation. Limited data suggest a relationship between the degree of IgD-induced Ca flux and decreased apoptosis.

Figure 6.38 suggests a correlation between responses to IgM and IgD in terms of improved viability. This suggests that a particular patient sample may have an inherent ability to respond to BCR stimulation. This is not correlated with the usual prognostic markers that are often invoked to explain differing responses to BCR ligation. Whilst microenvironmental signals are important, the intrinsic characteristics of the leukaemic cell will have a major impact on disease biology. A general ability to respond to BCR stimulation by whichever isotype may vary between patients.

There are some differences between IgM and IgD, with IgD appearing to exert a greater anti-apoptotic effect. Possible explanations for this include greater IgD expression relative to IgM; greater IgD Ca flux; unconsidered effects of the $F(ab')_2$ stimulus, such as the affinity of the antibody; the 2-3 hour 'rest' period after thawing

may have different effects on IgD and IgM signalling capacity. The phenomena of reversal of IgM (and to an extent IgD) anergy by incubation *in vitro* adds another level of complexity to this attempt to correlate the amount of BCR signalling with the amount of pro-survival effect of BCR signalling. A dose response relationship is often used to infer causality. Where this dose-response exists, these limited data suggest IgD may have a greater effect on survival than IgM, but it is difficult to draw strong conclusions from these data.

I suggest that there may be genuine differences in IgD and IgM signalling that account for these data, particularly considered in the light of the Ca flux data.

### *6.4.7: Choice of Stimuli*

The choice of control as a comparison appears to be critical. When IgD or IgM stimulation is compared to culture in medium alone, IgM and IgD often appear to induce apoptosis. In contrast, when the same stimulations are compared to borate buffer vehicle controls, IgD and IgM are seen to promote cell survival. Therefore, the borate vehicle used for dissolving the stimulating $F(ab')_2$ appears to have a degree of toxicity. It is clear that if this effect of vehicle is not taken into account, and medium alone is used as a control, then stimulation with IgD/IgM will often lead to induction of apoptosis. This is discussed further in relation to the published literature.

It is often considered straightforward to stimulate B-cells via their BCR, one simply cross-links sIg and marked changes in biochemical cascades and cellular behaviour occur. This simplistic model has been enormously successful in elucidating crucial signalling pathways in B-cells. Most studies in CLL have crosslinked sIg on B-cells by monoclonal antibody to sIg components, an approach that is well tested in studies of healthy B-cells.

How does the surface immunoglobulin that constitutes the BCR transduce the signal caused by antigen encounter across the membrane to the interior of the cell? It has often been thought that the simplest model for this is by crosslinking of multiple BCRs by polyvalent antigen, and this explains the more robust responses seen in vitro by

using multivalent antigens, but monovalent antigens are also able to stimulate B-cells as well. Other perturbations, perhaps involving conformational changes in immunoglobulin, lipid rafts or the cytoskeleton have been postulated[308,311]. It is likely that many different factors are crucial in mediating this initial step of BCR signal transduction, but despite much work in this area, the exact process is not clear-cut. Whilst crosslinking the BCR using anti-sIg antibodies clearly produces a robust response in B-cells that mimics effects seen when B-cells bind their cognate antigen, it may not recapitulate the situation *in vivo* as closely as is commonly assumed.

Some studies suggest that soluble antigen produces different responses from membrane-bound antigen, with immature cells encountering membrane-bound self-antigen undergoing apoptosis, whilst soluble self-antigens induce anergy[235,507]. Some argue that this is due to the multivalent nature of membrane-bound antigens providing a stronger BCR signal[508]. Soluble antigen is taken up and internalised by B-cells, which can continue activating cells via BCR signalling and endosomal Toll-like Receptors, as well as undergo further processing for presentation to T cells[509,510]. Full B-cell activation is dependent on both BCR signalling at the plasma membrane and T-cell help received when B-cells act as antigen-presenting cells, so internalisation of antigen is an important mechanism. Soluble antigen is a well-recognised potent immunogen, as small molecules such as toxins can readily enter follicles within lymph nodes, bind follicular B-cells and cause activation[250]. However, increasing attention is focussed on the recognition of antigen on the surface of membranes, in particular on the surface of antigen-presenting cells[511].

Antigen within the lymph node is often expressed on the surface of other cells such as macrophages, dendritic cells and follicular dendritic cells[250]. This generally produces more robust signalling, though this may be attributed to co-stimulation or adhesion molecules expressed by these antigen-presenting cells, rather than the spatial context *per se*. The formation of BCR microclusters on encountering membrane-bound antigen is dependent on CD19 and rearrangements of the cortical cytoskeleton[298].

Some authors have increased signalling strength by using anti-sIg monoclonal antibodies linked to streptavidin and crosslinked with biotin (though presumably these can subsequently be internalised), [329]. Others have used anti-IgM conjugated to the polysaccharide Dextran, which has been shown to produce a multivalent BCR stimulus that may have greater potency in activating B-cells[508,512], and which can induce pro-survival responses in CLL, as compared to medium alone or soluble anti-IgM $F(ab')_2$[513]. Interestingly, one study did not observe Ca flux after use of dextran-conjugated anti-IgM, a finding which has been noted in healthy mouse B-cells, and which suggests that Ca flux may not be necessary for mitogenic responses[514].

Some investigators have attempted to make their *in vitro* models of BCR stimulation in CLL more physiologic by immobilizing the anti-IgM to simulate membrane-bound antigen encounters. One early study noted CLL cell proliferation after use of Sepharose bead-coupled anti-IgM.[329] More recently, Petlickovski et al have used anti-IgM coupled to magnetic beads.[320] They noted marked differences of behaviour of CLL cells simulated with either soluble or immobilized anti-IgM, with greater pAkt and pERK with immobilized anti-IgM, and induction of CLL cell survival rather than apoptosis. They have found similar effects in more than one study.[487 372]

An alternative approach to immobilizing anti-IgM antibodies involves coating culture plates with the antibody before adding cells. [321] This form of immobilized $F(ab')_2$ and IgM has been shown to exert an antiapoptotic effect on CLL cells, compared to no stimulus or soluble antibody[**515** 400]. In a different protocol, Nedellec et al and Guarini et al used plates coated with anti-goat antibody, added CLL cells, then added 10μg/mL goat anti-IgM $F(ab')_2$. They suggest that this approach was an immobilised stimulus, though presumably some stimulus would have also occurred in a 'soluble' form. This may explain both these groups' findings: some cases exhibited a decrease in apoptosis and some an increase. Guarini (n=9) found an increase in apoptosis in M-CLL, and an increase in proliferation in U-CLL (increase in G1 from 1% to 2.7% after stimulation).

Despite following the same published protocol for coupling beads to anti-sIg monoclonal antibodies, I was unable to produce beads that appeared to bind cells

known to be IgM positive (RL cell line). Correspondingly, there was no apparent effect of 'conjugated' beads on calcium flux, apoptosis, or cell proliferation, compared to unconjugated beads, and so this approach was abandoned. Due to lack of published methodology, the coating of culture plates was not pursued. It is uncertain as to how to judge the success of the process. This lack of use of immobilised stimuli is a weakness in my work that will make comparisons with other studies somewhat difficult.

There is therefore a consensus that immobilised vs soluble anti-IgM may well account for some of the differences between the studies of BCR signalling in CLL, a finding that makes sense considering similar findings in healthy B-cells and considers the physiological context in which most antigens are encountered. However, it does not explain why some groups using soluble anti-IgM find an increase in viability whilst others find an increase in apoptosis. It may be helpful to consider the exact methods of BCR stimulation in each of the published studies.

Burger's group report an increase in viability after IgM crosslinking, compared to medium alone, an effect blocked by SYK inhibition [69,341]. They crosslink with 10μg/mL polyclonal anti-IgM $F(ab')_2$ supplied by MP biomedicals. This appears to be a lyophilized reagent that is reconstituted with medium before stimulus, therefore there is no effect of vehicle in addition to $F(ab')_2$, which may explain their consistent findings of improved viability after IgM stimulation.

Efremov's group have published several papers finding improved survival after IgM ligation with anti-IgM coupled to beads, in contrast to increased apoptosis with 10μg/mL anti-IgM $F(ab')_2$ in borate buffer[320,372,388], and a different group comparing soluble $F(ab')_2$ (borate buffer) vs plate bound $F(ab')_2$ have shown the same effect[321]. Similarly, Vallat et al found increased apoptosis with the anti-IgM $F(ab)_2$ coupled to biotin, again in borate buffer, as compared to medium alone[423].

Two groups have compared soluble anti-IgM $F(ab')_2$ with some form of vehicle control. Allsup et al found increased survival with IgM ligation when $F(ab')_2$ was compared to

an isotype control, both in phosphate-buffered saline[317]. Similarly Bernal et al found increased survival when anti-IgM F(ab')$_2$ was compared to polyspecific goat F(ab')$_2$. The company website (American Qualex) does not provide datasheets that provide information on the vehicle used in this study, and did not respond to enquiries.

In summary, authors find that soluble anti-IgM F(ab')$_2$ increases apoptosis when compared to incubation in medium alone, with the exception of Burger's group, who reconstituted lyophilized F(ab')$_2$ with medium. Analogous to my studies, the groups that compared some form of vehicle control found improved survival after IgM stimulation. Immobilised anti-IgM antibodies, either plate- or bead-bound, resulted in increased survival relative to both medium and soluble F(ab')$_2$. Some groups used a method that probably combined soluble with immobilised stimulus, and interestingly they found a mixture of responses. It therefore seems that the exact method of crosslinking IgM has a critical effect. This is probably in addition to a greater pro-survival effect of immobilised stimulus (though I suggest that a valid comparison has not been made). Additional factors such as CD38 and *IGHV* gene mutation status probably also have an effect, as various studies have found differences in the degree of apoptosis in differing prognostic groups. If this argument is accepted as valid, then most of the published literature is probably compatible with soluble anti-IgM stimuli causing an improvement in CLL survival.

These arguments also apply to my studies, though my main interest was in comparing BCR signalling via IgM and IgD. It does suggest that further work clarifying the exact method of BCR stimulus would be prudent.

### 6.4.8: Summary and Discussion: Outcomes of BCR ligation in vitro: Calcium Flux, apoptosis and cell proliferation

Teleologically, B-cells respond to BCR signalling with two divergent outcomes. The encounter of microbial antigen should result in pro-survival and proliferative stimuli to enable antigen-specific B-cells to mount an effective immune response. The encounter

of autoantigen ideally leads to an opposite response: pro-apoptotic BCR signals lead to deletion from the repertoire, or anergy results in the inability of autoreactive B-cells to mount an inappropriate response. Which of these divergent pathways a B-cell takes is dependent on a variety of factors. The nature of the antigen, binding affinity, binding valency and spatial arrangement is important. One critical factor is the presence of a second signal to accompany the primary BCR signal. Classically, this second signal is delivered by antigen specific helper T-cells, but additional signals are derived via other receptors such as Toll-like receptors that indicate inflammatory processes that may be caused by invading microbes. I have not considered these factors in my studies, attempting to focus on BCR stimulation alone. To an extent, some investigators have simulated these factors in conjunction with BCR stimulus in CLL.

Another critical factor affecting the response to BCR signal is the developmental stage of the B-cell. At an early stage in differentiation, BCR signalling is more likely to reflect binding of autoantigen, as microbial exposure is unlikely in bone marrow *in utero*. Therefore, BCR signals generally lead to deletion or receptor editing in the process of central tolerance. Mature cells are more likely to be expected to deal with foreign invaders, and so tend to respond to BCR ligation with proliferation. These differing outcomes are thought to be partly due to the pattern of activation of pathways downstream of the BCR, with dependence on pro-survival pathways that involve BTK, AKT and NF-kB[285]. It seems that the outcome of BCR ligation is therefore dependent on the particular pattern of pathways activated, and B-cells can be primed to respond with an apoptotic or proliferative response by altering which signalling cascades are connected to the BCR. To an extent, proximal BCR signalling events such as Ca flux are common to both pro- and anti-apoptotic pathways, but these proximal events are also modulated to determine these outcomes. A classic example is the attenuated Ca flux of anergic B-cells. Some mechanism has recognised that chronic antigen stimulation is occurring, and down-regulated the expression of IgM, and the ability of IgM to trigger Ca flux. The anergic B-cell is therefore thought to be primed to permit tolerance to autoantigen. This priming is reversed by removal of antigen. Paradoxically, IgD does not exhibit this effect, though there is a lesser degree of reduced IgD expression in mouse models of anergy[292]. There are various explanations. Perhaps IgD is less critical

than IgM in determining B-cell responses. The ability of IgD to compensate for the absence of IgM suggests that it can fulfil most of its cousin's roles. There are also certain immune defects in IgD-deleted mice, though these are relatively subtle. The maintenance of IgD throughout evolution suggests it plays some important role.

An alternative explanation is that IgD signalling is maintained in order to permit autoreactive cells to survive. It could be argued that ideally, all autoreactive B-cells should be deleted in the process of central tolerance. Why is anergy such a major tolerance mechanism? It may be that it simply deals with those cells that have escaped central tolerance. Alternatively, there may be some benefit to maintaining this pool of autoreactive B-cells. To an extent, newly generated BCRs must provide some signal in order for B-cells to survive, and it is well known that a weak level of autoantigen binding can provide such survival signals[94]. Also, in order to provide a repertoire that recognises the almost infinite variety of potential microbial antigens, some autoreactive B-cells are maintained that may subsequently be required to fight off an invader. This process goes awry during molecular mimicry, where overt autoimmune disease is subsequent to an infection and autoreactive B-cells attack autoantigens with similar epitopes to those present on microbial antigens[212]. Finally, there is a population of B-cells that produces natural autoantibody that may play an important role in clearing cellular debris such as apoptotic cells[516]. These explanations may provide a rationale for maintaining a population of autoreactive B-cells, rather than deleting them, with IgD perhaps providing a tonic signal.

In truth, no one has offered a convincing reason why B-cells co-express IgM and IgD, nor accounted for signalling differences between these two isotypes. What is known is that IgD and IgM are co-expressed in CLL, and that CLL resembles anergic B-cells in terms of lower IgM expression and attenuated ability to undergo proximal BCR events such as Ca flux. The anergic model is reinforced by the dynamic nature of IgM signalling, with reversible anergy after *in vitro* incubation. This is indirect supportive evidence for a model whereby chronic antigen binding is occurring *in vivo.* It may be that a degree of reversal of anergy may occur when CLL cells are in the peripheral blood, with re-encounter with antigen and IgM downregulation recurring on entry to

the lymph node microenvironment. This might account for cases of CLL that are able to undergo Ca flux after IgM ligation, as these circulating cells have had some time away from chronic antigen exposure, enabling them to upregulate IgM.

Teleologically, if the antigenic selection hypothesis of CLL is correct, one would expect BCR ligation to provide pro-survival signals. Many studies do in fact find that BCR ligation *in vitro* provides such a signal, and my findings agree with these studies. Those studies that find apoptosis after IgM ligation may be due to the nature of the stimulus used, both in terms of the relevant control for comparison, or soluble *vs* immobilised stimuli. The reversible nature of anergic IgM signalling also complicates the picture. I suggest that IgD and IgM can both deliver pro-survival signals in CLL. This is teleologically plausible, as the IgD and IgM isotypes will have identical antigen-binding regions, and therefore both bind the putative antigen. There is also the suggestion that IgD signalling may have a greater pro-survival effect than IgM, though the effect is small, and it is difficult to tease out whether this is due to higher expression of IgD, greater capacity to signal, an effect of the crosslinking reagent, or other factors. Despite their many similarities, the IgM and IgD isotypes of the BCR appear to play differing roles in CLL. The next Chapter describes attempts to describe mechanistic differences between IgD and IgM signalling.

# 7. Results-Phosphoproteomic analysis of BCR signalling

## 7.1 Introduction

Mass-spectrometry based phosphoproteomics is a technique that has great potential for mapping intracellular signalling pathways. Post-translational modifications (PTMs) change the conformation of proteins and thereby alter enzymatic activity, binding affinity and other functions[428]. The most important PTM is conceived to be protein phosphorylation on serine (85% of protein phosphorylation), threonine (10-15%) and tyrosine (1-2%) residues. This PTM is reversible and can produce rapid changes in protein tertiary structure that in turn affect protein function and cellular signalling pathways [429]. Despite its low relative abundance, regulated tyrosine phosphorylation is critically involved in many membrane-proximal signal transduction processes and is best characterised, though this may be due to the relative ease with which phosphotyrosine can be identified compared to phosphoserine/threonine. The number of diseases associated with mutations in tyrosine kinases also reflects the importance of phosphotyrosine[428]. Phosphorylation can regulate a variety of important functions, including subcellular localization, degradation and stabilization. It can control the binding of transcriptional regulatory elements to their regulatory sequence elements and alter the actions of RNA processing enzymes as well as the classically conceived regulation of biochemical activities of enzymes such as kinases[430]. Other PTMs such as ubiquitination, acylation and glycosylation are being explored by newer techniques. A brief summary of the high-throughput proteomic methodologies follows before the insights gained from the phosphoproteomic techniques are discussed.

### *7.1.2 Phosphoproteomics*

All proteomic experiments are faced by the same challenges: complexity, dynamic range and temporal dynamics.[428] The PhosphoSitePlus database currently (October 2012) lists >170,000 phosphorylation sites in a variety of species[517], and the numbers continue to increase[428]. Protein phosphorylation can occur extremely rapidly, receptor tyrosine kinase (RTK) phosphorylation often starting within seconds after stimulus application.[518] Tyrosine phosphorylations generally occur faster than serine/threonine

phosphorylation. On addition of epidermal growth factor to HeLa cells, most EGFR tyrosines become phosphorylated within 1 minute, some within seconds, while serine and threonine phosphorylation may require up to 10 minutes[519]. Despite the low abundance of tyrosine phosphorylation, it is very dynamic and can build extensive networks with interaction partners through binding motifs, such as src homology 2 (SH2) domains.[520]

There are diverse methods for analysis and quantification.[521] Previously, the identification of protein phosphorylation sites relied on the classical methods of *in vitro* and *in vivo* labelling with radioactive $^{32}$P-phosphate, peptide sequencing by Edman degradation, often followed by mutational analysis to verify the modified residue. Recent developments in Mass Spectrometry (MS) have had a dramatic impact on protein phosphorylation research. Thousands of phosphorylation sites can now be analysed in a large-scale manner and new techniques are reported regularly[522,523].

Phosphorylated proteins may number hundreds of millions to a few copies per cell, and many of the phosphorylation events that are associated with known signalling pathways may occur on proteins expressed at a relatively low level. Large scale quantitative analysis has shown that up to 30% of the phosphorylation sites can change in abundance within 30 minutes after application of a specific stress.[520] Most changes are less than tenfold, suggesting that phosphoproteome dynamics is expressed in changes of threshold levels rather than on-off phosphorylation. Several authors imply that phosphorylation only occurs on signalling proteins at low abundance, but phosphoproteomic datasets are not overrepresented in terms of low abundance proteins, and the profile of abundance distribution is similar for the proteome and the phosphoproteome[523]. It is generally implied that phosphorylated versions of a protein comprise a minority of all the molecules representing that protein.

Transient and substoichiometric changes in phosphopeptide levels are difficult to detect in whole cell lysates with millions of peptides. In order to generate a global view of serine, threonine and tyrosine phosphorylation that occurs at low stoichiometry

within a sample, enrichment techniques are used. Immunoprecipitation (IP) of tyrosine-phosphorylated proteins with high affinity antiphosphotyrosine antibodies provides a good yield and specificity[524]. The large size of phosphotyrosine makes this modified amino acid more immunogenic than phospho-serine or –threonine and therefore it has been easier to make anti-phosphotyrosine antibodies. Pan-specific antibodies for phosphoserine and phosphothreonine tend to be of lower affinity and yield unsatisfactory enrichment.[428]

Another commonly used method involves the use of immobilized metal affinity chromatography (IMAC), based on the high affinity of phosphate groups for metal ions such as $Fe^{3+}$, $Zn^{2+}$ and $Ga^{3+}$ [525]. Detection of phosphopeptides in the femtomole range can be achieved, and stable isotope labelling may enhance quantification.[526] Titanium dioxide ($TiO_2$) has emerged as the most common of the metal oxide affinity chromatography (MOAC)-based phosphopeptides enrichment. It has a greater capacity and convenience relative to IMAC resins. Phosphopeptide-enriched mixtures can then be analysed further. They may be resolved into various molecular weight fractions by liquid chromatography–tandem mass spectrometry (LC-MS/MS) for protein identification. Much of the initial work on MS-based phosphoproteomics was concerned chiefly with identifying novel phosphorylation sites, but in order to identify important changes in cellular networks, quantification of changes in phosphorylation is necessary.[428] Several MS-based quantification methods are available, often using labelling of amino acids with stable isotopes, which are then incorporated into cultured cells. These techniques are not generally applicable to primary cells, but other techniques such as spiking samples with a known concentration of phosphopeptides or labelling after enrichment steps have been used successfully[527]. A label-free method that quantifies peptides based on their ion intensity in MS spectra is proving to be a reasonably reliable quantitative means of determining relative levels of phosphopeptides within a sample[439].

Evaluating changes in phosphorylation across a large number of sites in a network is required to adequately describe the mechanisms by which phosphorylation controls cell biology. Rapid advances in these high-throughput techniques have not been

paralleled by equally powerful methods of analysing and interpreting the data generated, but the large-scale technologies have driven new approaches to analysis and interpretation. Phosphoproteomics by MS may represent an unbiased approach capable of monitoring cellular phosphorylation events in the absence of *a priori* knowledge. This would lead to a greater confidence in the interaction and signalling networks produced by high-throughput techniques as compared to an approach based on a narrow consideration of the curated literature.

### *7.1.3 Phosphoproteomics in studying haematological malignancies and immunoreceptor signalling*

A limitation of the phosphoproteomics techniques is the large numbers of cells required per sample, this is reflected in the widespread use of cell lines to provide sufficient numbers of cells for experiments. Primary tumour cells are generally hard to obtain, but studies of leukaemia have the advantage that large number of malignant cells exist in relatively high purity in the nucleated fractions of blood. A small number of phosphoproteomics studies have examined myeloid malignancies and have identified a few hundred phosphoproteins in acute and chronic myeloid leukaemia cell lines[528,529]. A recent study used IMAC and LC-MS/MS to identify 76 unique phosphoproteins in mantle cell lymphoma cell lines.[530] These data were correlated with information known on copy number gains obtained by Single Nucleotide Polymorphism-chip analysis and proteins involved in signal transduction cascades. Pathways involving NF–kB and PI3K were affected, and potential novel pathways in mitochondrial signalling were implicated. Advances in techniques are permitting the identification of ten of thousands of phosphoproteins.

The only phosphoproteomic study of CLL examined the consequences of CXCR4 signalling[531]. Analysis of 5 CLL samples generated approximately 1200 unique phosphosites, but with very little overlap between patients. The ability of this study to quantify phosphoproteins was limited, but several phosphoproteins thought to be relevant to CLL biology were identified, including LYN, SHIP-1 and HS1. Semi-quantitative data suggested the involvement of PDCD4 and HSP27 in CXCR4 signalling.

Dynamic signalling pathways in immune cells have been examined in a limited fashion using MS-based proteomics platforms. Phosphoproteins involved in signalling have been examined at various time points after stimulation of the FCεRI receptor in mast-cell lines.[532] One recent study has characterised 450 of 2000 proteins in murine T cells that were differentially regulated by T Cell Receptor stimulation[533]. This study highlighted novel mechanisms for HDAC7 in T Cell receptor signalling, highlighting the power of the technique. One study has examined BCR signalling in B-cell lines by examining phosphotyrosine changes.[534] In this study, BCR or TCR crosslinking was compared to cells with non-specific tyrosine phosphorylation produced by pervanadate. The B-cell line Namalwa was incubated with anti-sIgM for 3 minutes then washed, lysed, and purified by anti-phosphotyrosine antibody-conjugated sepharose column to enrich for phosphoproteins. TCR signalling was also examined. There was extensive overlap in identified phosphoproteins involved in signalling via both the TCR and BCR, with 128 of 725 proteins identified as being common to BCR and TCR signalling. Common proteins included tubulin isoforms, ribosomal proteins, chaperones, Cbl, Rho GTPases etc. Proteins not obviously directly involved in signalling such as tight junction protein Z02 and translational regulator eIF3 were also identified. There was also a stimulation of phosphorylation of RNA processing and endocytosis-related proteins.

Techniques have evolved whereby each experiment can generate ever-greater number of phosphopeptides for consideration. The phosphoproteomic techniques have not to date been applied to the study of BCR signalling in primary cells, leukaemic or healthy. It is hoped that by using phosphoproteomic methods, the mechanisms of BCR signalling will be elucidated in the absence of *a priori* assumptions about the nature of the pathways involved. The use of this technique has not reported in healthy B-cells, and so novel aspects of physiological BCR signalling may also be highlighted and comparison of the phosphoproteomes of healthy B-cells compared that that observed in leukaemia may be informative. In this study, the phosphoproteome of CLL cells is considered and compared to that of tonsil B-cells. The effect of BCR stimulation on the

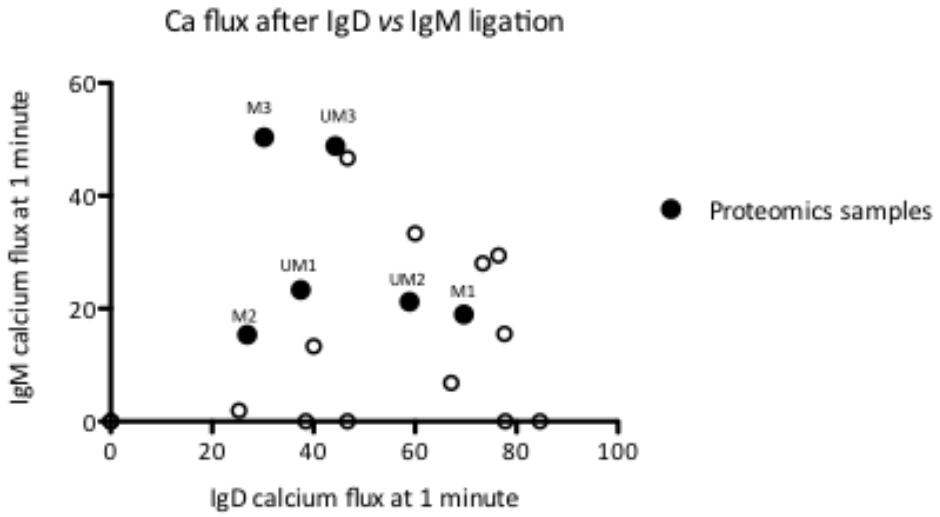
phosphoproteome is then considered, with a focus on the differences between IgD and IgM, and between CLL and Tonsil.

## 7.2 Methods

### *7.2.1 Samples*

To determine the phosphoproteome of CLL cells after BCR stimulus, samples were selected from untreated patients whose cells exhibited robust Ca flux with both IgM and IgD (Fig 7.1). This use of non-anergic samples enabled the study of IgD and IgM signalling without the complication of considering complete IgM anergy (absence of calcium flux) as a factor in this work.

**Figure 7.1 Calcium flux characteristics of samples used for proteomics.**
**Empty circles represent other samples tested (for further details see Chapter 6).**




To minimise handling and potential activation of cells by selection techniques that involve the use of monoclonal antibody binding, samples with known high proportions of malignant B-cells were used, as well as cases that had a range of prognostic markers in order to represent the known molecular heterogeneity of this disease. Patient characteristics of the samples used in this study are shown in table 7.1:

**Table 7.1 Characteristics of patients studied in phosphoproteomic experiments**
**UPN: Unique patient number (see chapter 5)**
***IGVH*: Immunoglobulin IGHV gene mutation status (M=mutated, <98% homology to baseline, U=unmutated, ≥98% homology to baseline)**
**Cytogenetics: FISH panel for CLL performed**
**CD38: peripheral blood flow cytometric expression (1=>30%, 0=<30%)**
**Ca flux: Percentage of cells undergoing Ca flux at 1 minute, as a proportion of calcium flux with ionomycin (see Chapter 6)**
**$CD19^+CD5^+$: percentage of cells dual positive for CD5/CD19 by flow cytometry**

| UPN | Proteomics ID | *IGHV* | Cytogenetics | CD38 | Ca flux IgM | Ca flux IgD | $CD19^+$ $CD5^+$ |
|---|---|---|---|---|---|---|---|
| **P7** | **UM1** | U | del13q del11q | 0 | 23 | 37 | 95.9 |
| **P20** | **UM2** | U | del13q | 1 | 21 | 59 | 98.3 |
| **P16** | **UM3** | U | del17p | 1 | 52 | 44 | 98.4 |
| **P17** | **M1** | M | del13q | 1 | 19 | 70 | 91.7 |
| **P19** | **M2** | M | del13q | 0 | 15 | 27 | 98.6 |
| **P18** | **M3** | M | del13q | 0 | 50 | 30 | 98 |

Tonsillar B-cells were used as a control, as they displayed robust Ca flux with IgD and IgM and provided sufficient numbers of cells for the analysis. Cell suspensions were derived from fresh tonsil, and negative selection for B-cells performed before freezing and storage. Apart from the negative selection process, cells were treated identically to the CLL samples. Purity was assessed by CD19 staining (Fig 7.2)

Table 7.2 demonstrates that the selection technique used to purify tonsillar B-cells achieves >90% purification, and figure 7.3 indicates that these are predominantly (94%) naïve B-cells (defined by CD27 negativity) that express IgD, or both IgM and IgD, in keeping with the results obtained with calcium flux after stimulation (Fig 7.4).

**Figure 7.2: CD19-determined purity of tonsillar B-cells**
**a) before and b) after negative selection for B-cells (representative plots). Isotype control (red) is compared to FITC-conjugated antibody specific for the B-cell marker CD19 (blue). Selection increases B-cell proportions to 95.3% of all cells**

**a** **b**

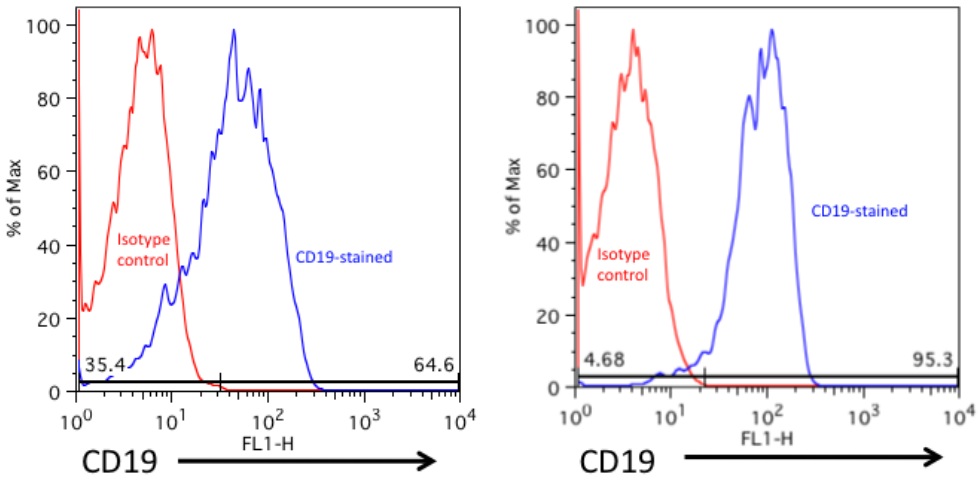



**Table 7.2: Characteristics of of negatively-selected tonsillar samples**
**Ca flux: Percentage of cells undergoing Ca flux at 1 minute, as a proportion of calcium flux with ionomycin (see Chapter 6)**
**$CD19^+$: percentage of cells positive for CD19 by flow cytometry**

| Tonsil number | $CD19^+$ | Ca flux IgM | Ca flux IgD |
|---|---|---|---|
| **1** | 95.3 | 57 | 71 |
| **2** | 97.9 | 50 | 64 |
| **3** | 96.9 | 83 | 75 |
| **4** | 94.3 | 56 | 67 |
| **5** | 92.1 | 74 | 84 |

**Figure 7.3: Surface isotype and CD27 expression of purified tonsillar B-cells (tonsil number 1).**
**Quadrants based on isotype controls. CD27 negativity is taken as a measure of naivety**

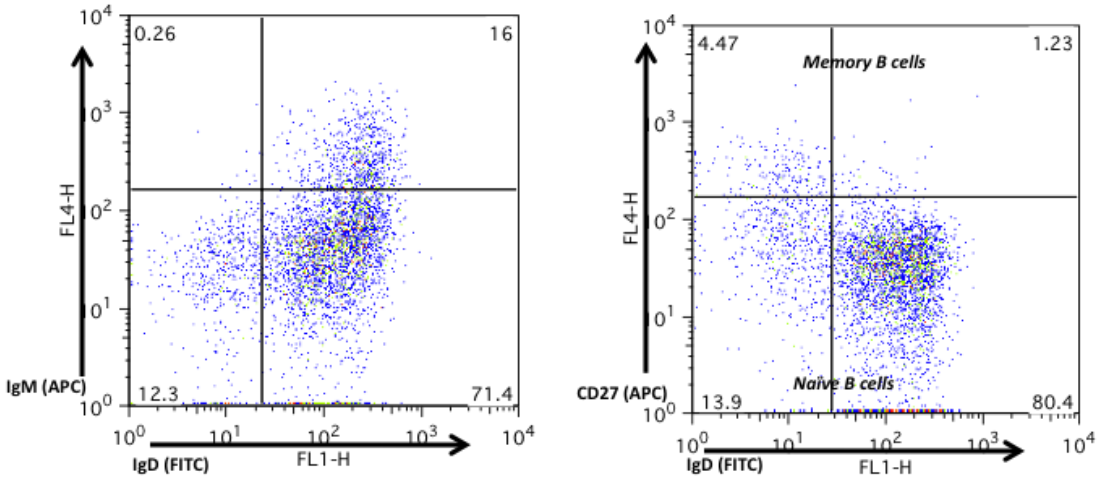
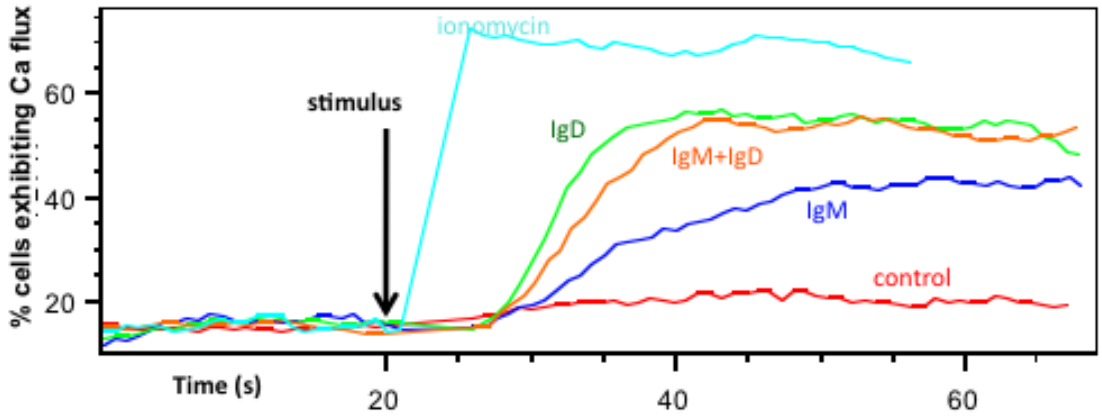
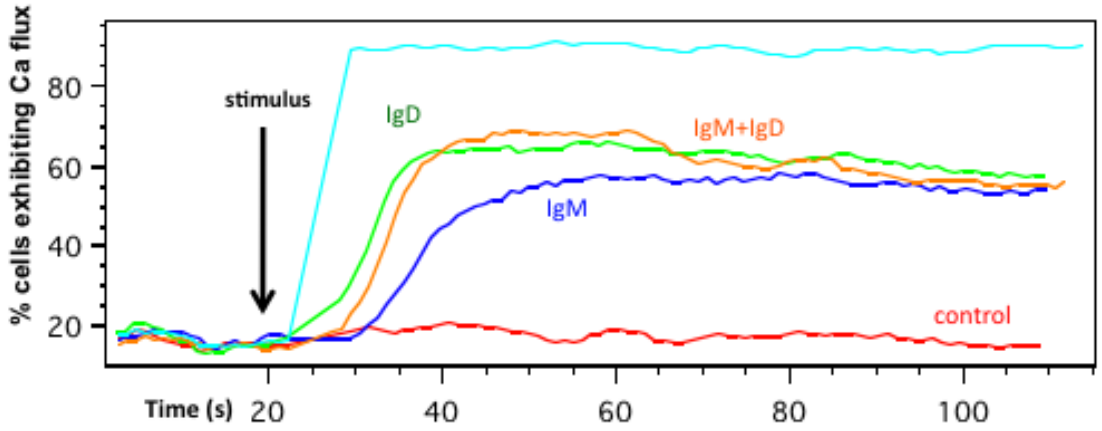


**Figure 7.4: Calcium flux responses in tonsillar B-cells (representative results).**

**T1**




**T4**




Upon thawing, all samples were divided into replicates and rested for 2-3 hours before experiments. All samples had viability >95% as assessed by trypan blue dye exclusion. An aliquot of each was assessed for their ability to undergo Ca flux on the same day as stimulation for the proteomics experiments to confirm viability and ensure the capacity for proximal BCR signalling events. Triplicates of 15-20x10$^6$ cells were used to provide sufficient protein for proteomics. An additional aliquot of 5-10x10$^6$ cells for western blotting validation were stimulated in parallel. Three replicates per condition were stimulated separately and simultaneously. This led to a total of 9 proteomics samples (triplicates of baseline, IgM and IgD-stimulated) per patient. These biological replicates (same source) are subsequently treated as technical replicates except where indicated. With 6 CLL and 5 Tonsil 'patient' samples, 99 samples for proteomics were processed.

Similar to previous methods (see Chapter 6), cells at $4x10^6$/mL were stimulated with 10μg/mL anti-IgM or IgD $F(ab')_2$ fragments, or control borate buffer. After 5 minutes at 37°C, samples were transferred to chilled falcon tubes, and washed twice in chilled phosphate buffered saline (PBS) containing phosphatase inhibitors. After washing, pellets were snap frozen in dry ice and kept at -80°C until lysis.

Cells were lysed in cold 8M Urea in 20mM HEPES (ph8) with additional phosphatase/kinase inhibitors ($Na_3VO_4$, NaF, β-Glycerol Phosphate, $Na_2H_2PO_4$). The Bradford assay was used to quantify protein concentrations. 500μg of protein was then denatured with dithiothreitol, cysteines alkylated with iodoacetamide and proteins digested over 16 hours with trypsin beads. The digest was then acidified with trifluoroacetic acid.

Samples were de-salted, followed by enrichment of phosphopeptides with Titanium Dioxide beads immobilised on reversed-phase C-18 resin spin columns. The phosphorylated peptides were dried and reconstituted in buffer for LC-MS analysis.
Liquid chromatography Mass spectrometry (LC-MS/MS) analysis on an orbitrap mass spectrometer was performed with our collaborator Alex Montoya. Data analysis was carried out using MASCOT (phosphopeptide identification) and PESCAL (quantification based on normalised peak height) software [440,535] by Alex Montoya and Dr Pedro Cutillas.

## 7.3 Results

### *7.3.1 Overall analysis strategy*

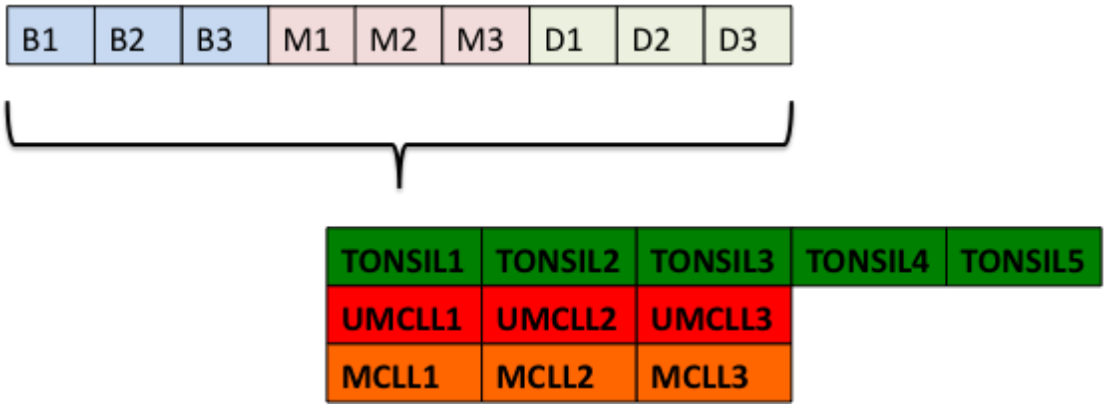
There are no established analysis methods for analysing high-throughput phosphoproteomic data. Where useful, methods from gene expression analysis have been employed[536-541]. The experiments provided quantitative data on large numbers of phosphorylation sites within several thousand proteins. Estimates of the degree of experimental 'noise' within the dataset were made, and appropriate analysis methods were specified. Following this, differences between leukaemic and healthy B-cells were considered, and the effects of B-cell Receptor stimulation. The phosphosites derived from these analyses were then characterised. One interpretation considers the nature and function of the phosphorylated proteins, based on the presumption that phosphorylation alters the function of a protein. Methods adopted from gene expression analysis consider the proteins as components within specified biological pathways, or their known alterations in disease states. This involves gene set enrichment analysis and comparison to curated databases designed for this function. Beyond this, the data can also provide an interpretation of the activities of kinases responsible for the phosphorylations themselves. A minority of phosphorylation sites have experimentally confirmed kinases, and where possible these were identified. Various algorithms for prediction of kinases exist, and these methods are pursued to an extent. An exhaustive account of all analyses performed is beyond the scope of this thesis, but I aim to provide example of analyses that represent both the potential and challenge of phosphoproteomic data analysis. The primary aims were to use the data to gain insights into CLL biology and BCR signalling in general.

### *7.3.2 Characteristics of data*

5088 unique phosphopeptide entities ('features') were identified. Each identified feature has various characterising data, as well as a derived quantification/expression, the sequence of amino acid residues, probability of misidentification, phosphorylation site (phosphosite) and other chemical modifications such as oxidation and formylation. Quantification/expression based on the peak height is used. Each sample is given a total phosphopeptide content of 1, which is the sum of all peak height quantifications

identified in that sample[439]. The expression values of each feature therefore reflect the fraction of the total that the feature contributes to the total signal of a sample. In discussion of the results, the Uniprot ID is referred to when describing a protein, with phosphorylations indicated by pY (phosphotyrosine), pS (phosphoserine) or pT (phosphothreonine) followed by a number indicating the phosphorylated residue. Triplicates, stimuli and patients are described as in Figure 7.5.

**Figure 7.5: Proteomics Replicates and Patients.**
**5 Tonsil and 6 CLL patient samples were used. Triplicates were stimulated with IgM (M), IgD (D) or vehicle control (Baseline, B).**




### *7.3.3 Proteomic quantifications: expression and 'zero' normalisation*

To better consider the expression levels, quantifications were converted to log base 2. As can be seen in fig 7.6, it is likely that values <-23 are likely to be equivalent to zero where, in essence, these values are likely to be 'noise' that reflects such low levels of peptides that they may be ignored in analyses. The analysis was applied separately to the CLL and tonsil data in Fig 7.7.

**Figure 7.6: $log_2$ expression frequency in all samples.**
**The second chart shows the frequencies around the minimum from the previous chart.**

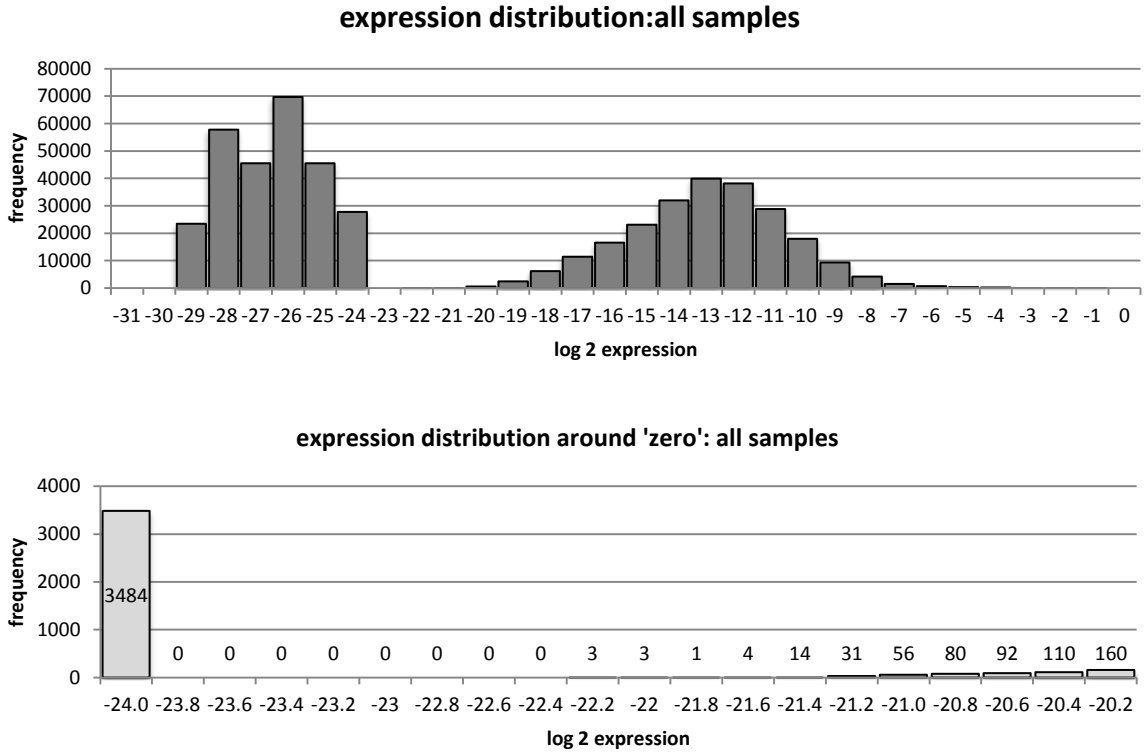
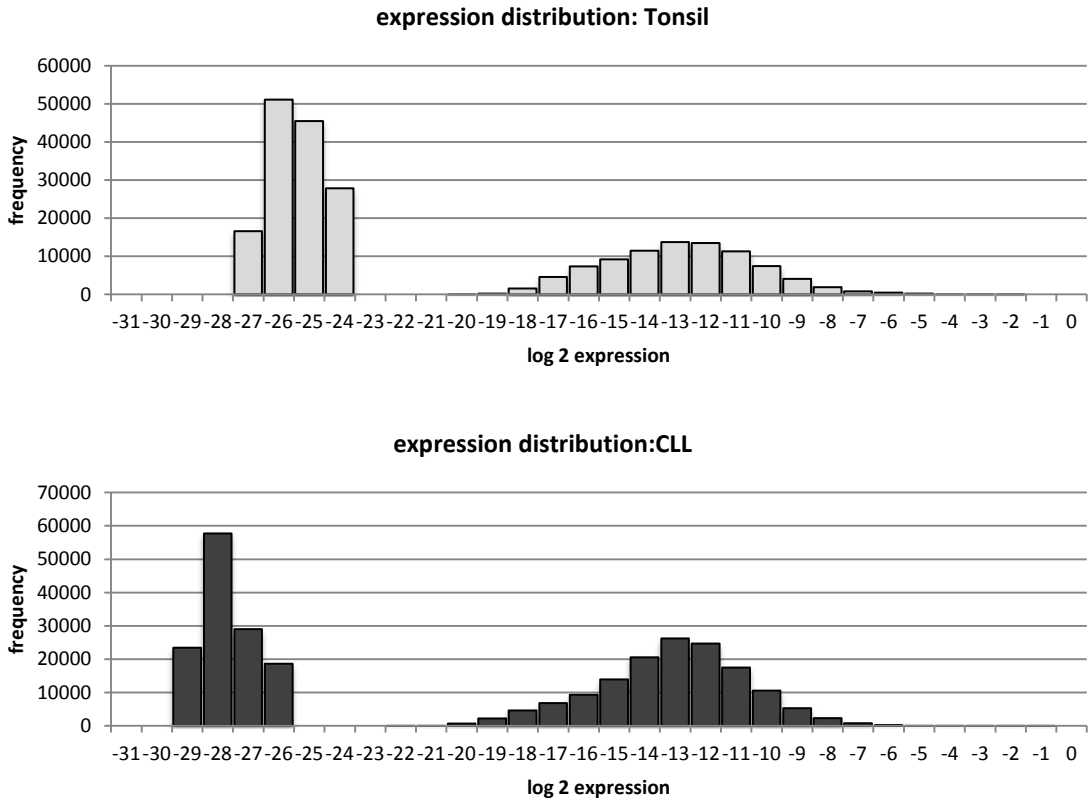



**Figure 7.7: Log expression values for tonsil and CLL samples separately**



Overall, 54% of all feature quantifications have expression values <-23, as not all samples contain every feature. For the CLL sample, this is 47%, for tonsil 62%.

Since it is likely that values below -23 are equivalent to zero for both CLL and tonsillar B cells, this cut-off was used to define a zero/minimum as $1.19209 \times 10^{-7}$ (-23 on $\log_2$ scale). Subsequent data analyses are performed using the data set with the data normalised ('zeroed') to this cut-off/zero threshold of -23. Apart from reducing 'noise', the rationale for this is that analysis of the expression differences (fold changes) between two samples will be strongly affected by lack of zero normalisation. The use of a cut-off/threshold for the minimum value allows proper comparisons in terms of fold changes between samples. The use of a minimum value instead of a zero avoids problems with dividing by zero. The technique is derived from gene expression analysis[541,542].

86/5088 (1.7%) features contain a tyrosine phosphorylation, 5034/5088 (98.9%) serine/threonine phosphorylation. This is similar to published data[428]. 1652 unique proteins were phosphorylated, with 3179 unique phosphosites producing the 5088 unique features. 1028/5088 had more than one feature per unique phosphosite, with one (HNRPC pS260) having 16 different features representing the same phosphosite. This is partly due to a phosphoprotein being trypsin digested into peptides of different fragment lengths, oxidised/formylated peptides derived during sample processing, and differing fragmentation within the mass spectrometer. This presents an analytical challenge when attempting to quantify a particular phosphosite as multiple features may represent it, each of which may provide an indicator of the abundance of the parent phosphoprotein.

It might be expected that the quantifications of different features derived from the same phosphosite would correlate with each other. Alternatively, it may be that trypsin digestion varies between different samples due to alterations in pH or other variables, such as chemical instability of particular peptides, ionization efficiencies of differing phosphopeptides or alterations in cleavage site produced by phosphorylation itself [543,544]. Therefore one might also expect some variation between samples as to

the relative levels of features derived from one phosphoprotein. Furthermore, some features may be misidentifications. Even when there is great confidence in identifications, it might be expected that quantifications based on low abundance features may be less accurate in terms of both identification and quantification.

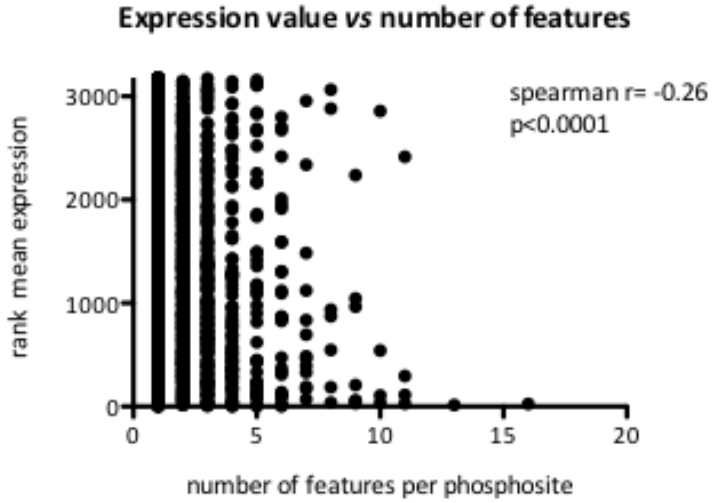
To estimate the similarities and differences between quantifications of different features derived from the same phosphoprotein, we examined correlations between multiple features derived from the same phosphoprotein. Of 3179 unique phosphosites, 1028 are represented by more than one feature. The distribution of these is below in table 7.3

**Table 7.3: Number of features representing each unique phosphosite.**
**Of 5088 features, 3179 unique phosphosites are present.**

| Number of features | Frequency | Notes |
|---|---|---|
| 16 | 1 | HNRPC pS260 |
| 15 | 0 | |
| 14 | 0 | |
| 13 | 1 | HDGF pS165 |
| 12 | 0 | |
| 11 | 4 | |
| 10 | 4 | |
| 9 | 7 | |
| 8 | 7 | |
| 7 | 13 | |
| 6 | 28 | |
| 5 | 43 | |
| 4 | 90 | |
| 3 | 211 | |
| 2 | 618 | |
| 1 | 2152 | |
| ***total*** | ***3179*** | |

Of note, the two phosphosites producing the most features were ranked 33$^{rd}$ (HNRPC) and 13$^{th}$ (HDGF) in terms of mean expression, which is consistent with a model whereby more abundant phosphoproteins produce greater numbers of features. This hypothesis is supported by a significant negative correlation between the number of features representing each phosphosite and the rank mean expression (se fig 7.8), though this is not invariably the case.

**Figure 7.8 Number of features per unique phosphosite *vs* the rank mean expression. The number of features representing a phosphosite is summed to yield a quantification of a phosphosite, and the mean value for all 9 samples taken and ranked to provide an estimate of the relative abundances of each phosphosite.**

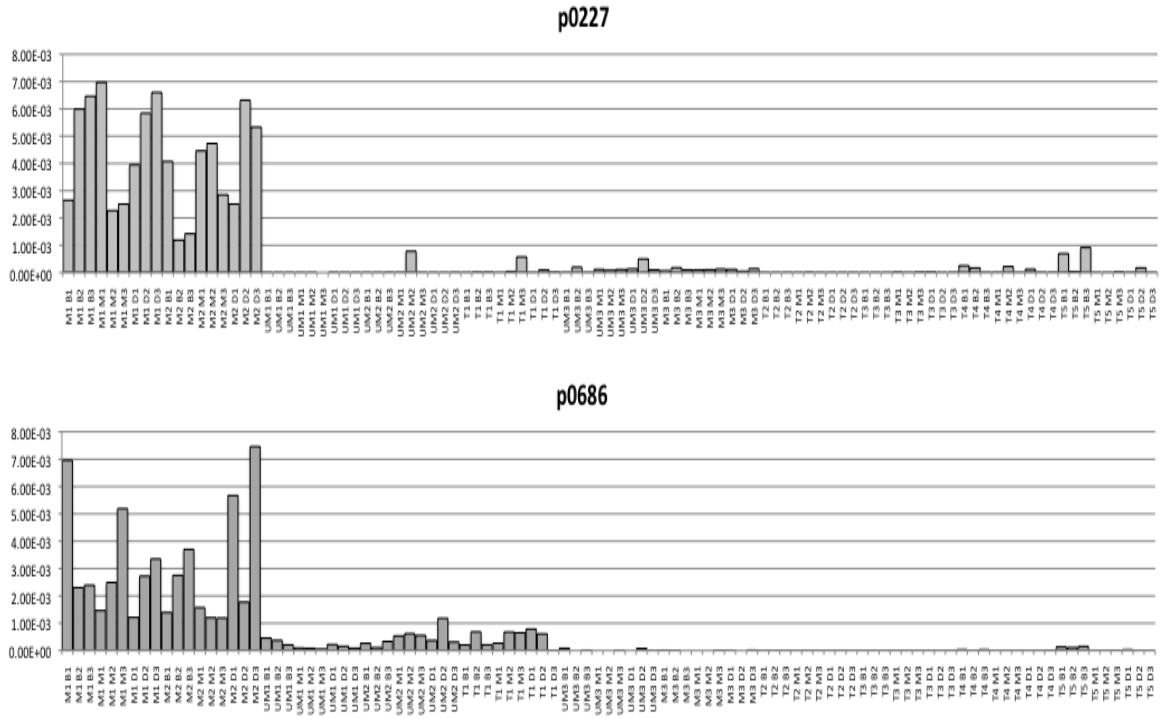



The two phosphosites represented by the highest number of features were considered in more detail. The phosphosite HNRPC pS260 was represented 16 times. The two most abundant phosphopeptides representing this are p0227 and p0686, quantifications of the same fragment:

MESEGGADD**S$^{p}$**AEEGDLLDDDDNEDRGDDQLELIK

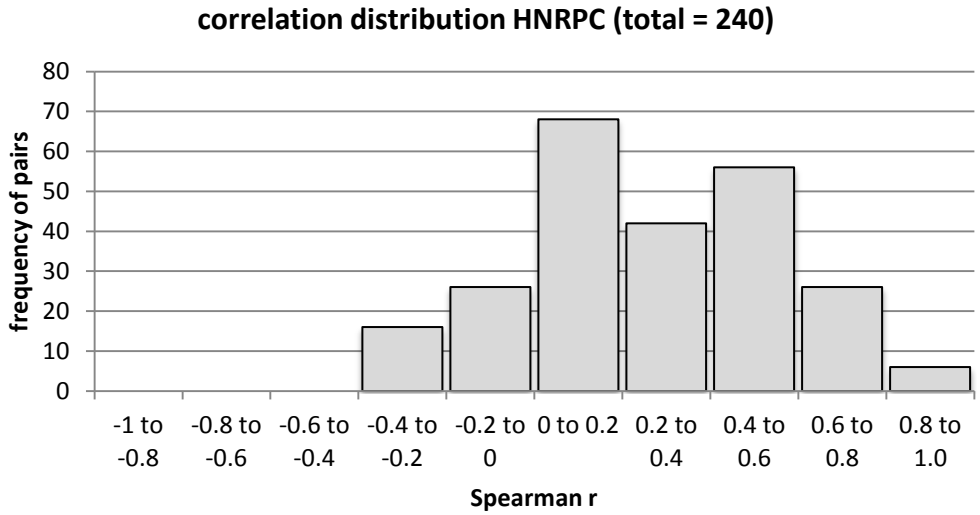
These differ only in that the initial methionine is oxidised. Their quantifications are similar but not identical across samples (see figure 7.9), with Pearson r 0.65 ($p<0.0001$), spearman r 0.34 ($p=0.0007$). The Spearman r is increased to 0.44 when zero normalised data are used.

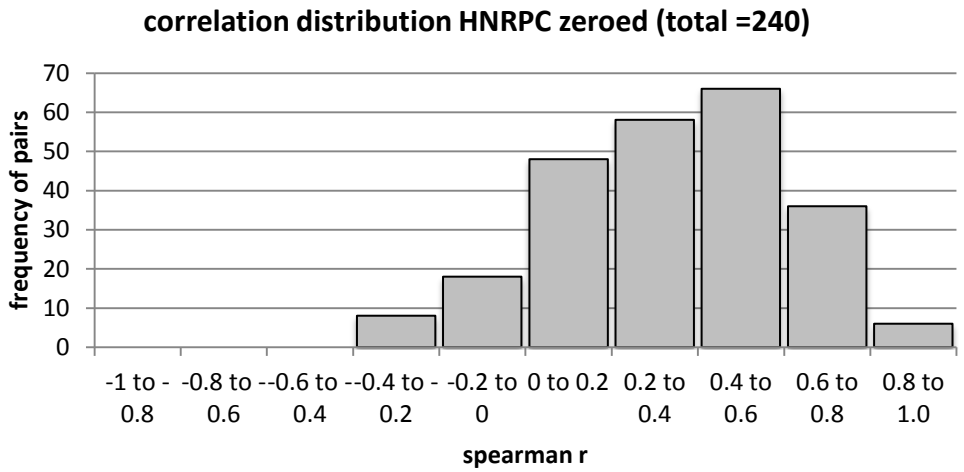
**Figure 7.9 Expression values of two most abundant features representing HNRNPC pS260**

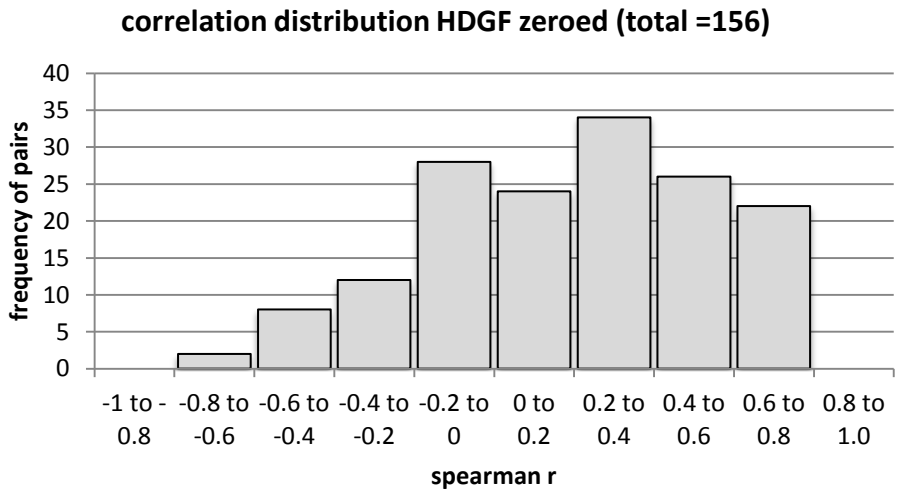



To formally quantify this similarity, the correlations between all 16 features were considered for all 99 samples. Of the 240 resulting correlations, most were positive, with 61% being significant at $p<0.05$ level (73% for zero normalised data). The distribution of correlation coefficients is represented in fig 7.10.

**Figure 7.10. Distribution of correlation coefficients of all features of HNRPC pS260. Each of the 16 feature abundances was correlated with the 15 other features to generate 240 spearman r correlation coefficients. Most were positive, and the distribution of values is shown for both zeroed and unzeroed data. The same analysis is shown for HDGF (zero-normalised )**

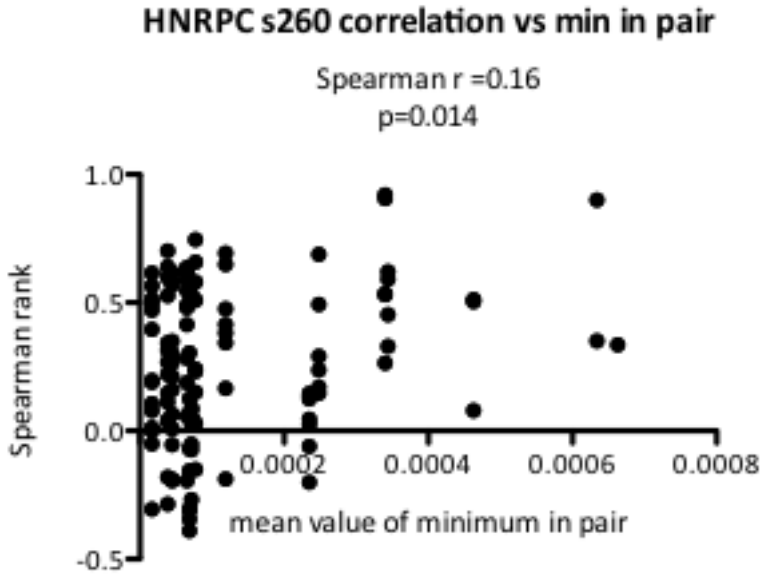


Therefore, whilst there is often a positive correlation between features within a sample indicating similar quantifications of features derived from the same phosphosite, the correlation is often zero or occasionally negative. To an extent, lower abundance features exhibit less correlation, as one might expect if noise/variability is more prominent with low abundance features (Fig 7.11), but this does not account for the majority of the poor correlations between features.

**Figure 7.11: Spearman r correlation coefficients for each feature representing HNRPC pS260.**
**For each of the 240 correlations calculated, the spearman r was compared to the expression value of the lower abundance feature of the correlation pair. The lower the expression value, the greater tendency to a negative correlation coefficient.**




Therefore, whilst the features representing a particular phosphoprotein are generally related to each other in terms of expression value, there is a large amount of variation, only partly due to low abundance features. One interpretation might be that the analysis inaccurately assigns the identification to each feature. Another is that these fragments are produced differently in different samples/experiments, for the reasons stated (Trypsin cleavage sites, differing pH). A combination of these effects may be possible. Regardless of the cause, it presents a challenge: it is not always clear which measure of abundance should be considered in analyses. The most abundant feature might most accurately represent the phosphoprotein abundance. However, the most

abundant feature in one sample/ experiment may not be the same as that in another. Furthermore, it can be seen that even highly abundant features do not always correlate with other high abundance features, indicating that at least one of them is non-proteotypic. A derived value that takes into account all feature abundances would be most appropriate. An obvious candidate would be the mean expression value of all features representing a particular phosphosite. However, a highly abundant phosphoprotein produces several fragments, so the mean value of each of multiple fragments would produce an inaccurate underestimation of the abundance of the parent phosphoprotein. We have chosen to take the sum of all features as a marker of phosphosite abundance, summing the expression value of each feature to produce an overall expression value for each phosphosite. The expression values of 5088 features were thus combined to 3179 individual unique phosphosites to give expression values for each phosphosite.

This method also partly compensates for the variability that is due to 'noise' at low abundance levels, with features of higher abundance contributing more to the derived value. Misidentifications will hopefully be counteracted by this combination of their data with true identifications that might be expected to occur at higher abundance.

We therefore have two datasets used in further analyses:
1. P for 'Primary', consisting of 5088 phosphosites, many of whom will be duplicates
2. J for 'Joint', consisting of 3179 unique phosphosites
It might be expected that the J datasets would have somewhat less variability/noise, but also that the results from the two datasets would largely overlap.

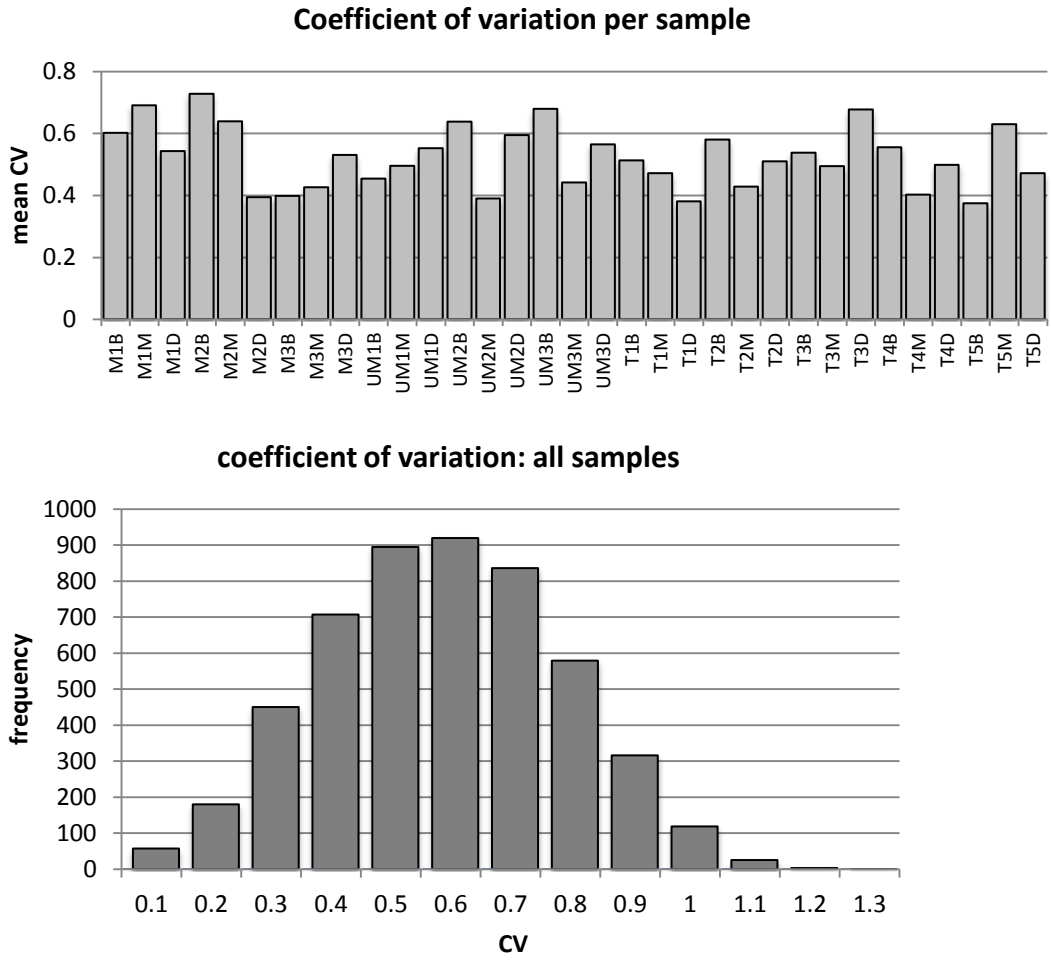
### *7.3.4 Variability and clustering*

Attempts were made to estimate the variability between replicates and samples using a variety of methods. The coefficient of variation was calculated, the correlation coefficients were calculated, and clustering was performed.

**1. Coefficient of variation.**

For each triplicate, the standard deviation of the expression values for a particular feature was divided by the mean value to yield the coefficient of variation (CV). The mean CV for each sample and all samples were calculated (Fig 7.12).

**Figure 7.12: Coefficients of variation calculated for each triplicate of the 5088 features in P dataset.**
**The mean value for each sample is plotted, and the frequency distribution of all CVs is also shown**

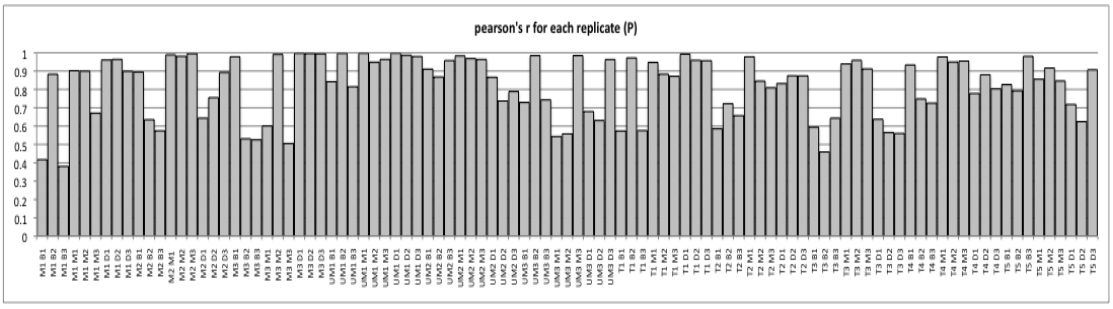


The mean CV for all samples was 0.52 (52%). The same analysis on the data before 'zeroing' yielded a mean CV of 0.64, lending support the use of zero normalisation as a means to reduce 'noise'. There was no reduction of mean CV by consideration of only the top 50% abundant features (implying that variability is not due to more 'noise' at low expression values). The J dataset had a mean CV of 0.53.

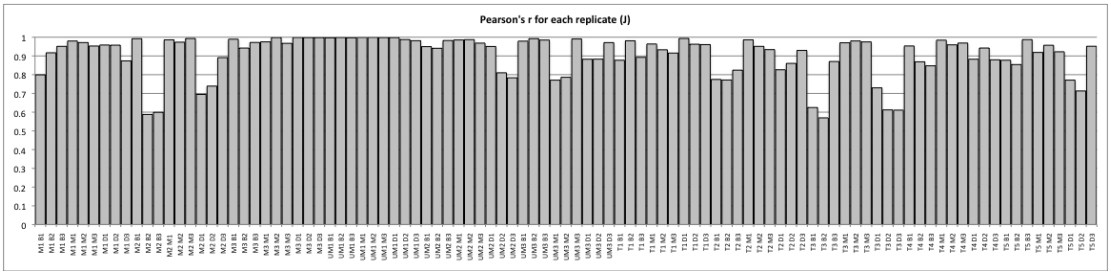
**2. Correlation coefficients**

Pearson's r was calculated for each replicate pair, with each sample yielding three values (Fig 7.13). For example, patient M1 had three baseline samples: M1B1, M1B2, M1B3. The correlation coefficents for M1B1 vs M1B2, M1B2 vs M1B3 and M1B3 vs M1B1 were calculated. Mean pearson's r for all replicate pairs was 0.82, which is unchanged by zero nomalisation. Correlation was not higher when the top 50% abundant features were considered. The mean correlation coefficient was 0.91 with the J dataset.

**Figure 7.13: Correlation coefficients for each sample.**
**Pearson r was calculated for each replicate pair within an experimental condition. Each coefficient is plotted for each sample pairing.**
**A Primary dataset**




**B Joint dataset**



Pearson's correlation is an analysis that should be applied to normally distributed data, and the low values may reflect this. We were unable to perform Spearman's r on the large dataset, and so clustering techniques were employed to assess similarities between samples.

**3. Clustering**

Two clustering techniques were used, hierarchical clustering using the Cluster software[537], and Principal Components analysis using Metaboanalyst software[545] with standard parameters derived from gene expression analysis[541].

**Unsupervised Hierarchical clustering**

The expression values were logged, centred by median values and distance calculated using the spearman rank correlation method. Average linkage was used for hierarchical clustering. The means of triplicates were calculated and the same clustering algorithm was applied to the means of each triplicate. Heatmaps and hierarchical trees were generated using TreeView software[546]. See Figs 7.14 and 7.15

**Figure 7.14 Heatmap and unsupervised hierarchical clustering of all 99 samples (P dataset). Note that the hierarchical tree is enlarged and rotated by 90$^{o}$ to facilitate viewing.**

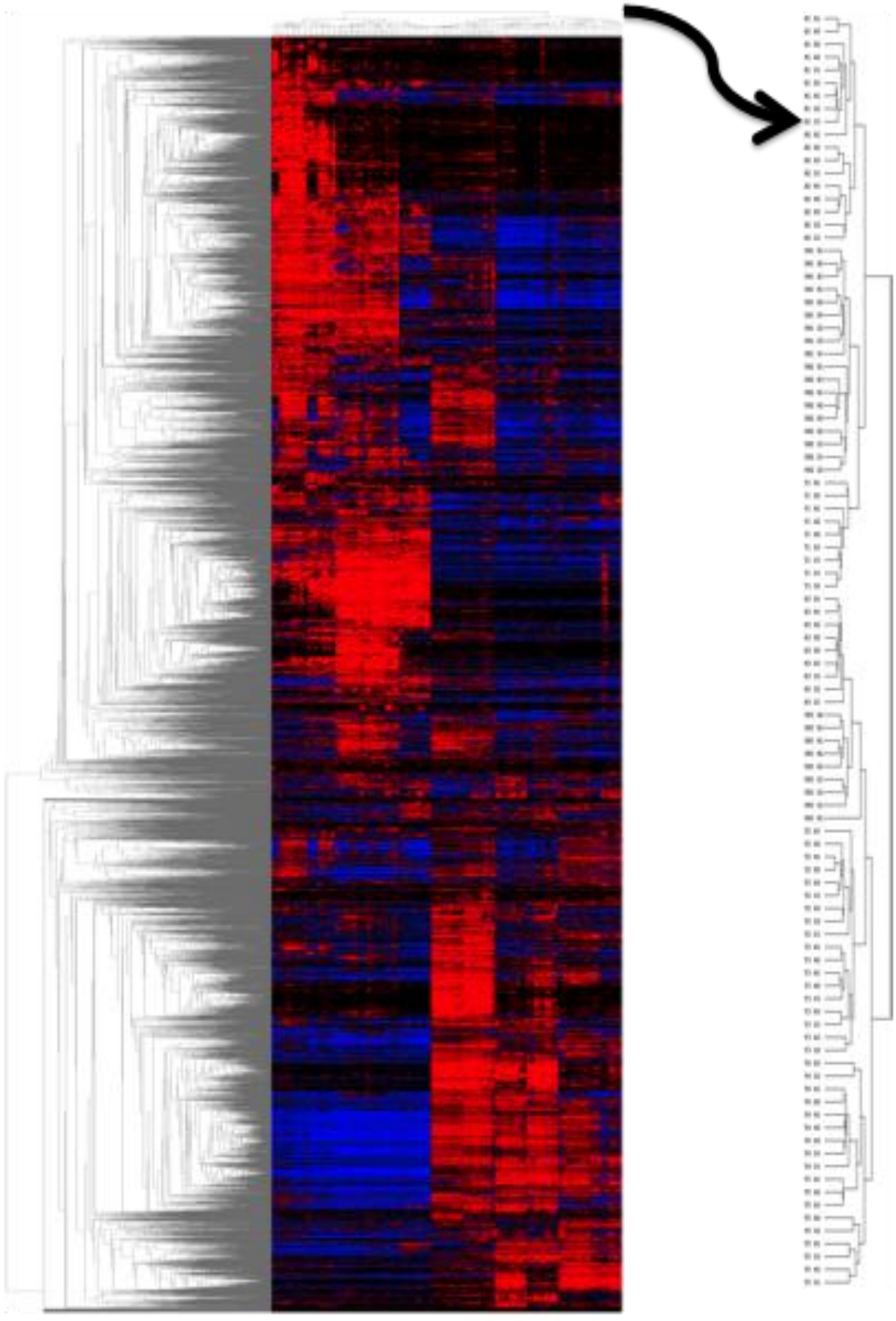
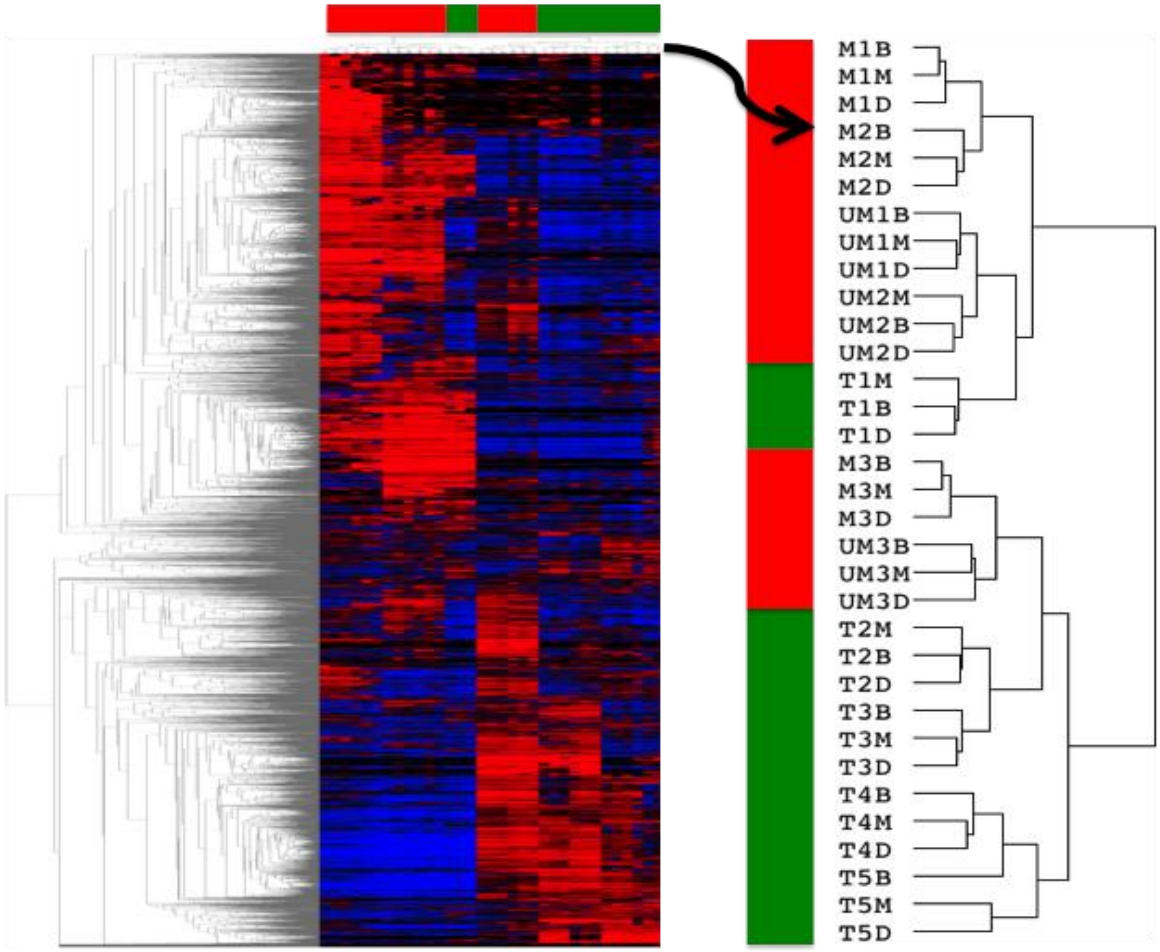


**Figure 7.15 Heatmap and hierarchical clustering of 33 means of triplicates (P dataset). Note that the hierarchical tree is enlarged and rotated by 90° to facilitate viewing. CLL samples are coloured red and tonsil green.**
**M=CLL *IGHV* mutated, UM= CLL *IGHV* unmutated T=Tonsil**
**B=baseline, M= IgM stimulated, D= IgD stimulated.**

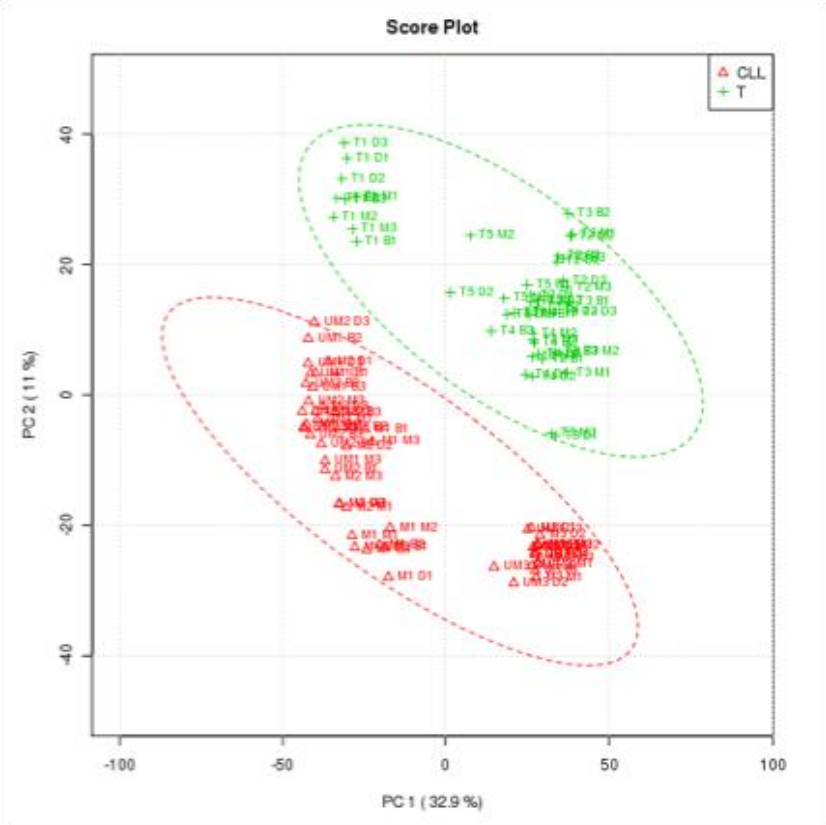



Whilst there is clear clustering by patient, replicates do not appear to cluster together, there is no clustering of baseline, IgM- and IgD-stimulated samples, and there is overlap between the CLL and tonsil samples using this method, though these are partially separated. Hierarchical clustering was identical between joint and primary datasets.

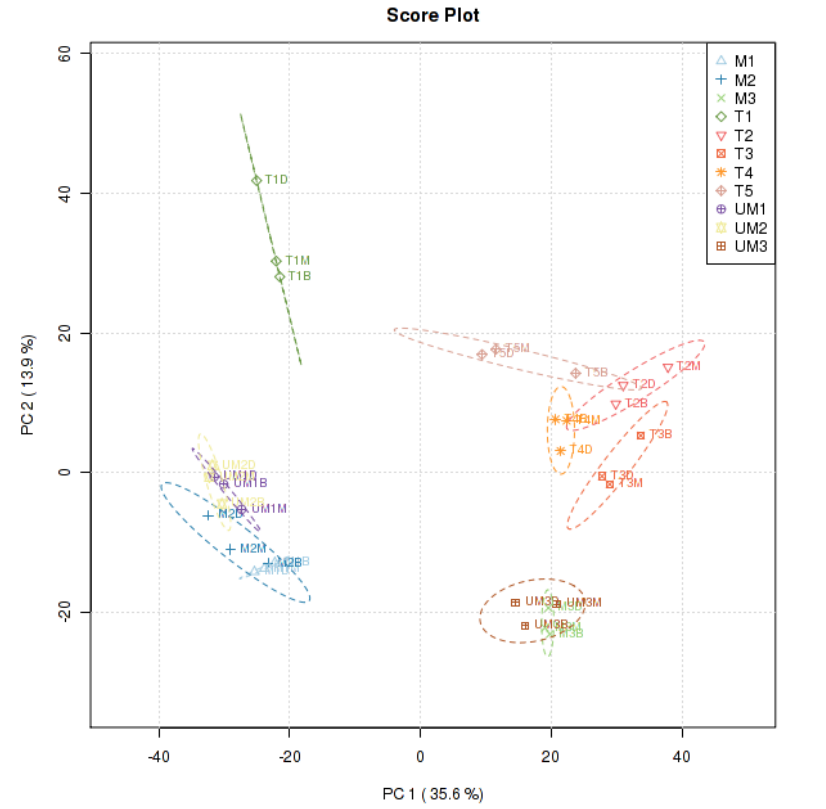
**Principal Components Analysis**

The MetaboAnalyst software provides a principal components analysis (PCA) algorithm often used for mass spectrometric data. Using standard techniques, there was good separation of Tonsil and CLL samples when considered together, though individual replicates within samples did not generally group together using PCA. Figure 7.16-7.17 shows representative examples.

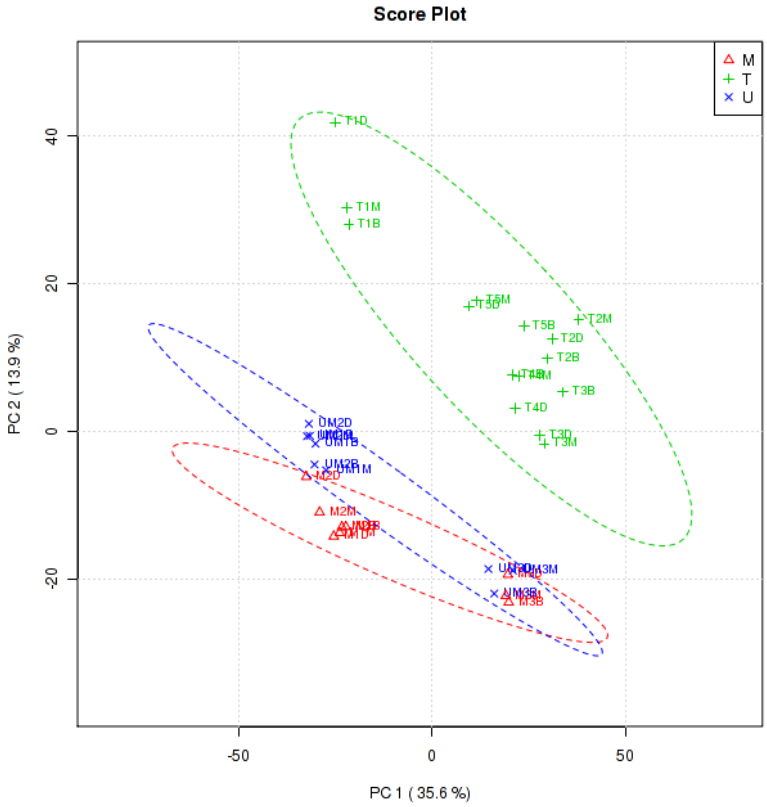
**Figure 7.16: PCA of P dataset.**
**Ellipses represent 95% confidence intervals.**
**A All samples grouped by CLL *vs* tonsil**




**B all means grouped by sample**



**C all means grouped by *IGHV* mutation status**



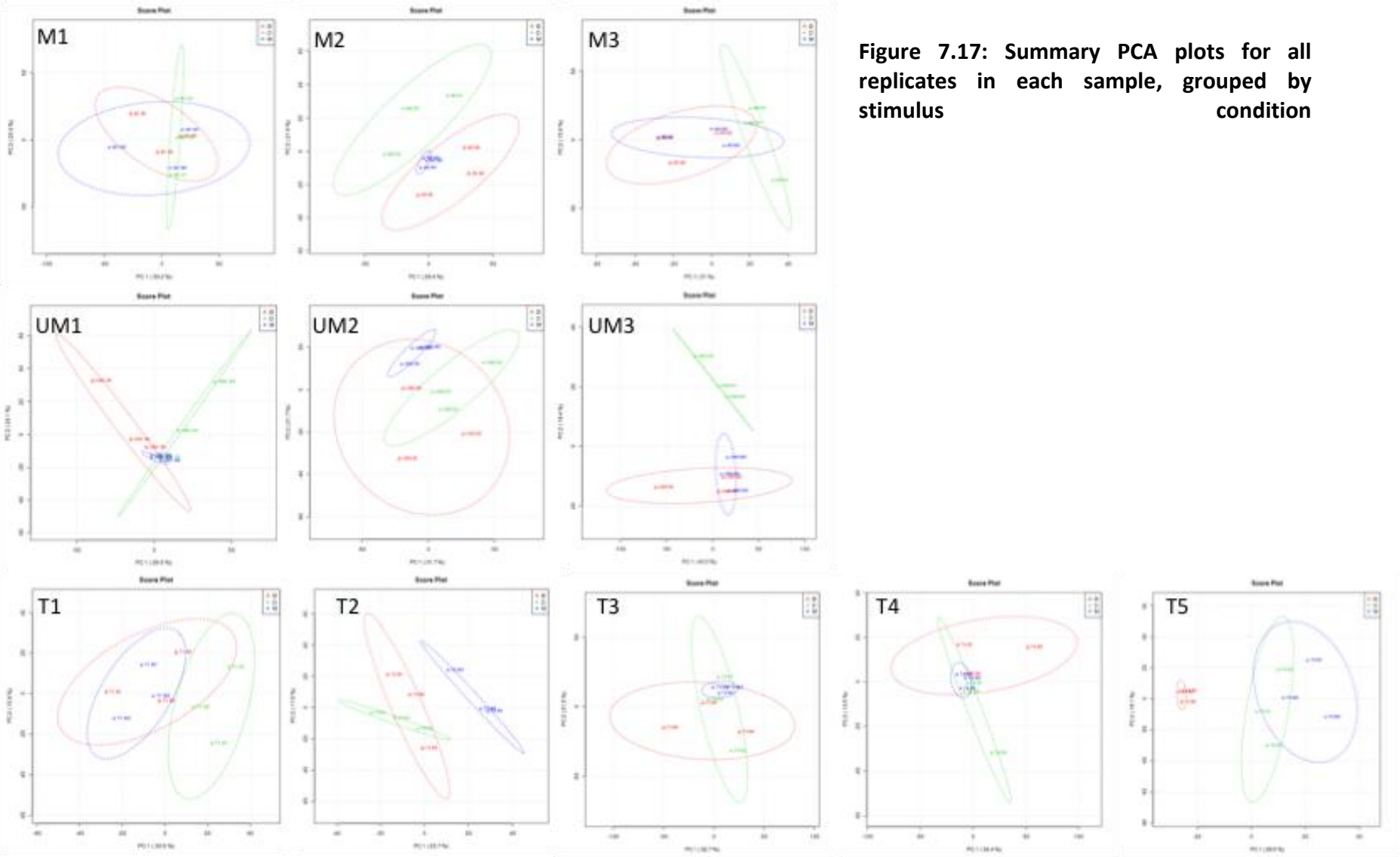
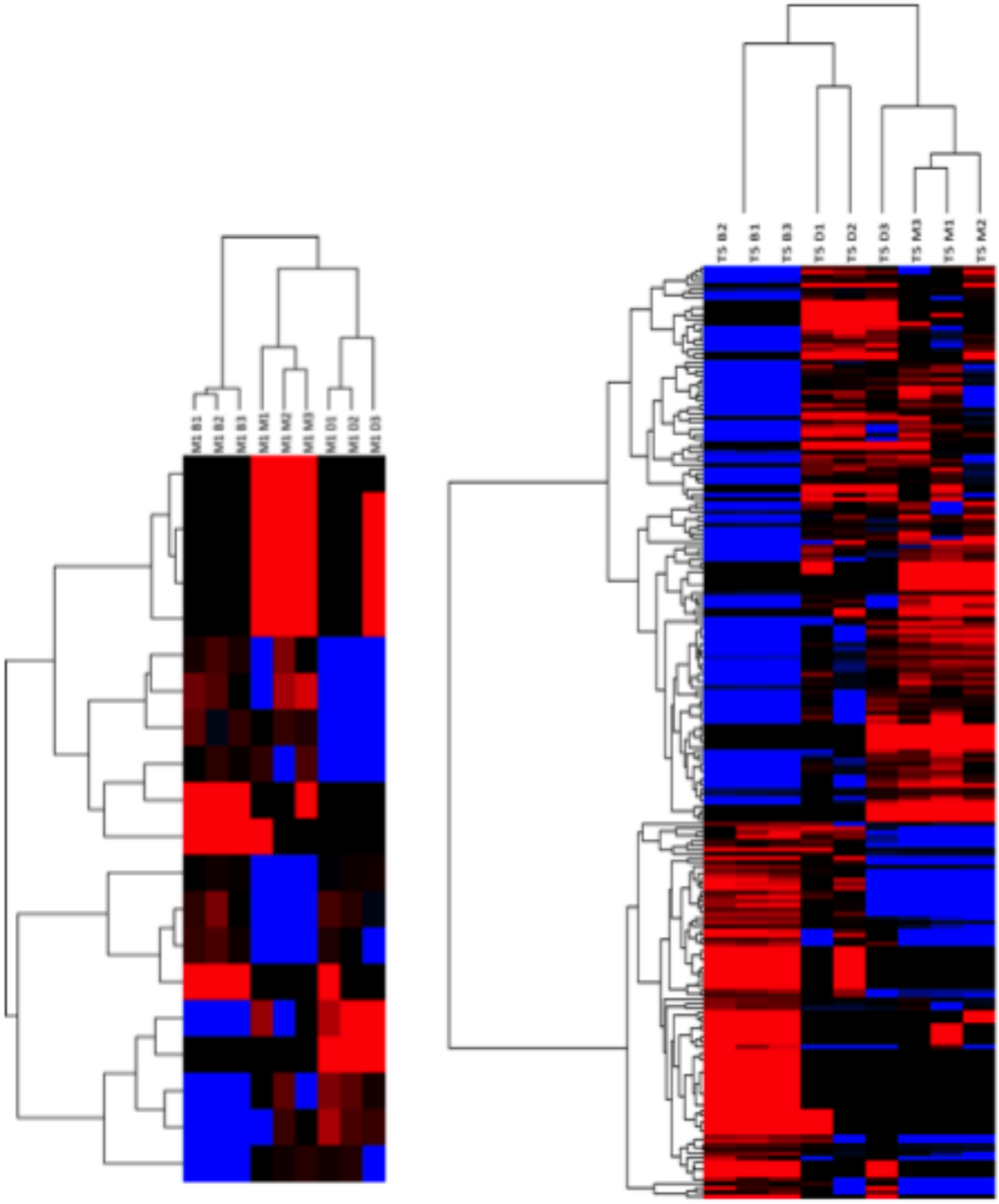



**Figure 7.17: Summary PCA plots for all replicates in each sample, grouped by stimulus condition**

It seems that although there is high overlap between replicates and stimuli, there is clear separation of 'patients' from each other. Joint and primary datasets produced the same patterns.

The lack of clustering of replicate conditions with each other may seem concerning, but this is often seen in gene expression data. Most of the clustering is dominated by which patient the sample comes from. There is also a large amount of 'noise', as discussed. We would expect only a small proportion of the total phosphosites to alter on stimulation of the BCR, and this is indeed the case (see section 7.4). Less than 2% of the total phosphosites were significantly altered by BCR stimulation, and the pattern was often unique to each patient. Therefore, it is not surprising that the replicate variability is dominated by the 98% of phosphoproteins not consistently altered by BCR stimulation. When supervised clustering is performed on each of the patients with only the phosphoproteins significantly altered by BCR stimulation, then clear clustering is seen. Figure 7.18 demonstrates this graphically. Full details are present in section 7.4, but these data are shown here to explain the lack of replicate clustering when all data are considered. The effect of BCR stimulus is small, but statistically significant.

**Figure 7.18: Supervised clustering of patients M1 and T5**
**The replicates of these samples do not cluster together when all phosphosites are considered by unsupervised clustering (see figure 7.17). When only phosphosites significantly altered by IgM or IgD stimulation are considered then clear separation of the replicates by stimulus is seen. Patient M1 had 10 phosphosites significantly altered by IgM stimulation, and 10 by IgD stimulation, with no overlap. Patient T5 had 145 phosphosites altered by IgM stimulation, 104 by IgD, with an overlap of 47 phosphosites. Shown are the supervised clusters for the IgM- and IgD-altered phosphosites combined. See section 7.4 for details.**



### *7.3.5 Characteristics of phosphoproteins*

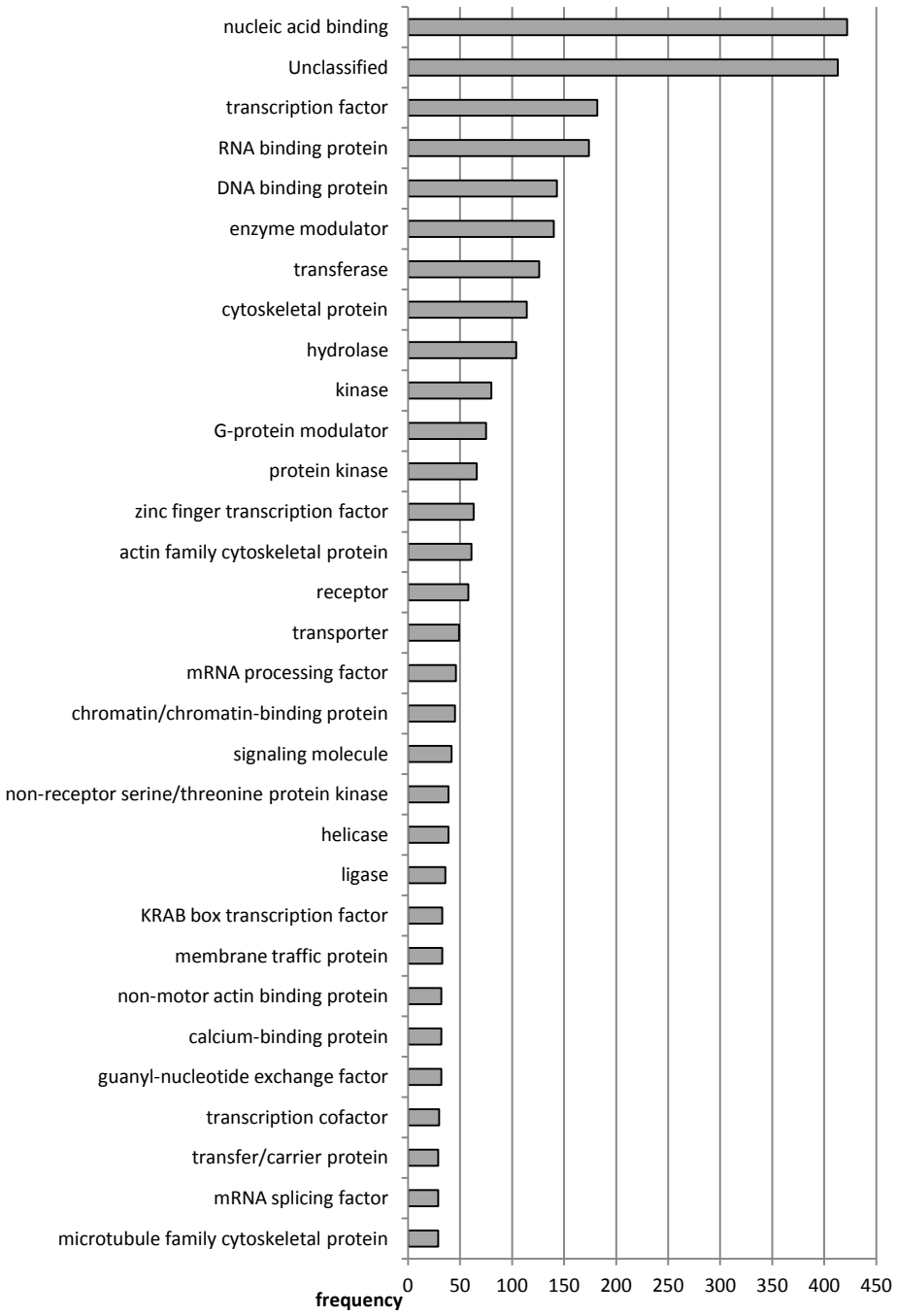
The nature of the detected phosphoproteins was considered. The abundance of a phosphoprotein was estimated by summing all of its component phosphosites. The twenty most abundant phosphoproteins were calculated by taking the mean for all samples, and are shown below in table 7.4.

**Table 7.4: nature of most abundant phosphoproteins in all samples.**
**The mean expression values of all samples for each phosphosite were calculated. Different phosphosites for the same protein were then summed. The expression values derived are as a fraction of total proteins.**

| Uniprot ID | Protein | Expression | Function |
|---|---|---|---|
| H12 | Histone H1.2 | 0.112 | Histone |
| LSP1 | Lymphocyte-specific protein 1 | 0.042 | intracellular F-actin binding protein |
| LIPS | Hormone-sensitive lipase | 0.037 | lipase |
| K1C13 | Keratin, type I cytoskeletal 13 | 0.031 | cytoskeletal protein |
| TR150 | Thyroid hormone receptor-associated protein 3 | 0.019 | Plays a role in transcriptional coactivation. |
| BCLF1 | Bcl-2-associated transcription factor 1 | 0.019 | Death-promoting transcriptional repressor |
| HMGA1 | High mobility group protein HMG-I/HMG-Y | 0.016 | transcription regulation/RNA processing |
| DDX51 | ATP-dependent RNA helicase DDX51 | 0.015 | biogenesis of 60S ribosomal subunits |
| ARF5 | ADP-ribosylation factor 5 | 0.012 | GTP-binding protein |
| ZN774 | Zinc finger protein 774 | 0.011 | May be involved in transcriptional regulation |
| FA40A | Protein FAM40A | 0.011 | regulation of cell morphology and cytoskeletal organization |
| CALX | Calnexin | 0.010 | quality control apparatus of the ER by the retention of incorrectly folded proteins |
| NMDZ1 | Glutamate [NMDA] receptor subunit zeta-1 | 0.008 | glutamate-gated ion channels with high calcium permeability |
| TEBP | Prostaglandin E synthase 3 | 0.007 | promoting disassembly of transcriptional regulatory complexes |
| PPE2 | Serine/threonine-protein phosphatase with EF-hands 2 | 0.007 | May function as a calcium sensing regulator of ionic currents |
| CN159 | UPF0317 protein C14orf159, mitochondrial | 0.006 | ? |
| SHOX | Short stature homeobox protein | 0.006 | Controls fundamental aspects of growth and development. |
| ARP3B | Actin-related protein 3B | 0.006 | organization of the actin cytoskeleton |
| H13 | Histone H1.3 | 0.004 | Histone |
| MCCA | Methylcrotonoyl-CoA carboxylase subunit alpha | 0.003 | ? |

When the nature of the phosphoproteins is considered in terms of their molecular function as judged by Protein ANalysis THrough Evolutionary Relationships (PANTHER) molecular function[547], the largest classes of phosphoproteins function were nucleic acid binding, unclassified and transcription factor (fig 7.19).

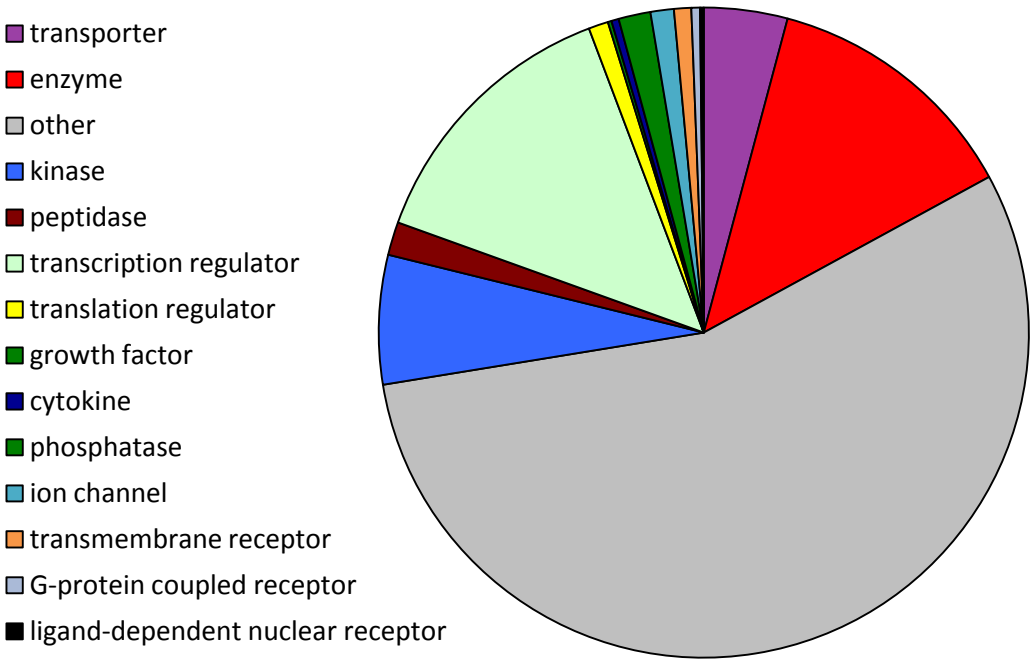
**Figure 7.19: PANTHER protein classifications of all phosphoproteins. Frequency of molecules in the 30 most frequent protein classes.**



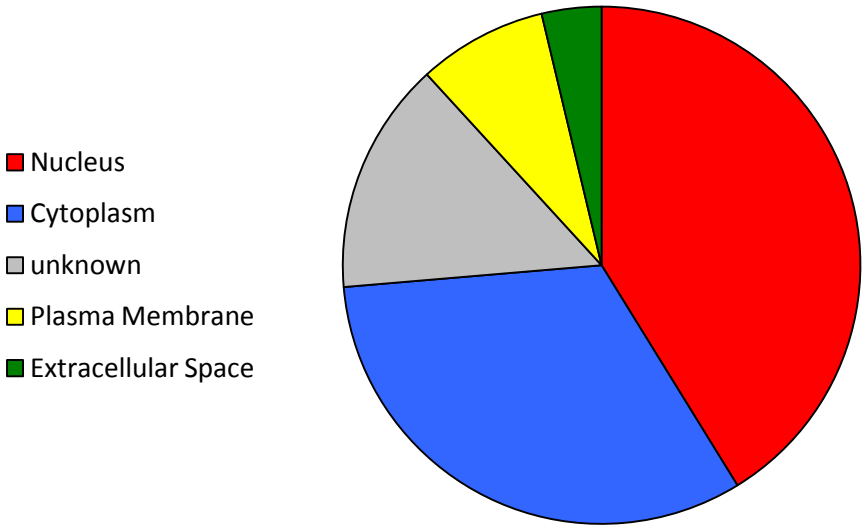
The use of Ingenuity Pathway Analysis (IPA) has a somewhat simpler classification system, and also permits assignment to particular cellular localisation (figure 7.20)

**Figure 7.20 IPA classification of molecular function and cellular localisation for all detected phosphoproteins**

**A**



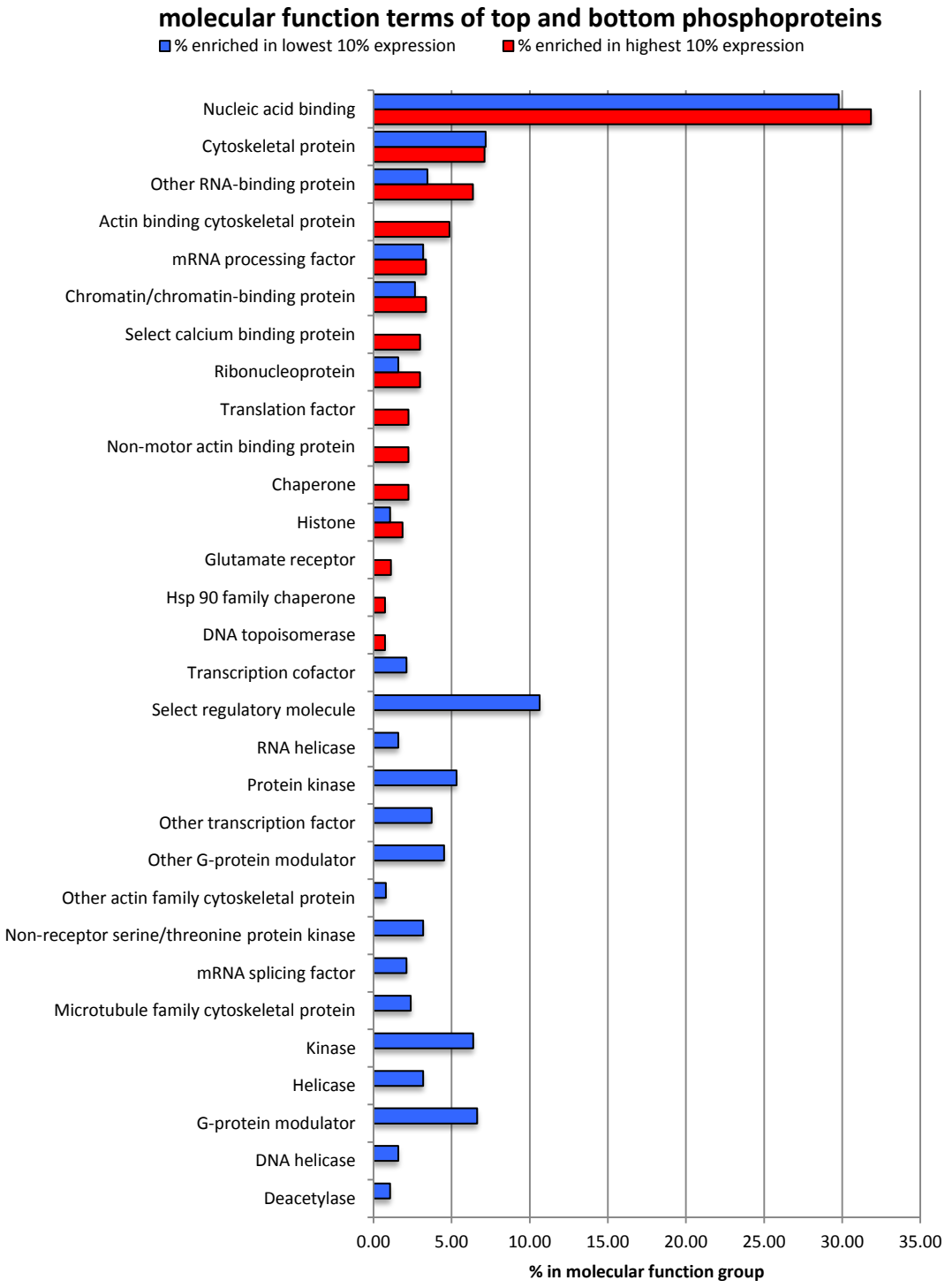

**B**




The top 10% most abundant proteins were compared to the 10% least abundant proteins in terms of molecular function (Fig 7.21). The more abundant proteins were

enriched for RNA binding proteins, cytoskeletal components, and histones, though these were also present at low abundance. The least abundant proteins tended to have enzymatic or regulatory functions.

**Figure 7.21: Most and least abundant 10% of phosphoproteins.**
**Taken from the mean expression values of baselines for CLL and tonsil combined.**
**The molecular function of those proteins was classified using PANTHER.**



### *7.3.6 Ingenuity Pathway Analysis of detected phosphoproteins*

To gain more insight into the functions of the phosphoproteins present in the baseline CLL and tonsil samples, IPA was used to provide an estimation of which known cellular pathways are active based on the detected phosphoproteins, analogous to gene expression analysis (Table 7.5). Interestingly, several pathways were significantly enriched, including the B-Cell receptor pathway (fig. 7.22). Because of the general similarities between CLL and Tonsil expression profiles (93% of phosphoproteins with non-zero mean expression in tonsil samples were also in CLL), the results are similar.

**Table 7.5 Top 10 most significantly enriched cellular pathways.**
**Ingenuity (IPA) software used Fisher's exact test to calculate p-value of enrichment.**
**The ratio represents:**
***number of phosphoproteins in our dataset/total number of proteins in pathway***
**A: The results from analysis of mean baseline CLL samples**
**B: The results from Tonsils**
**Fig 7.22: The highlighted proteins from the ERK/MAPK & BCR IPA pathways**

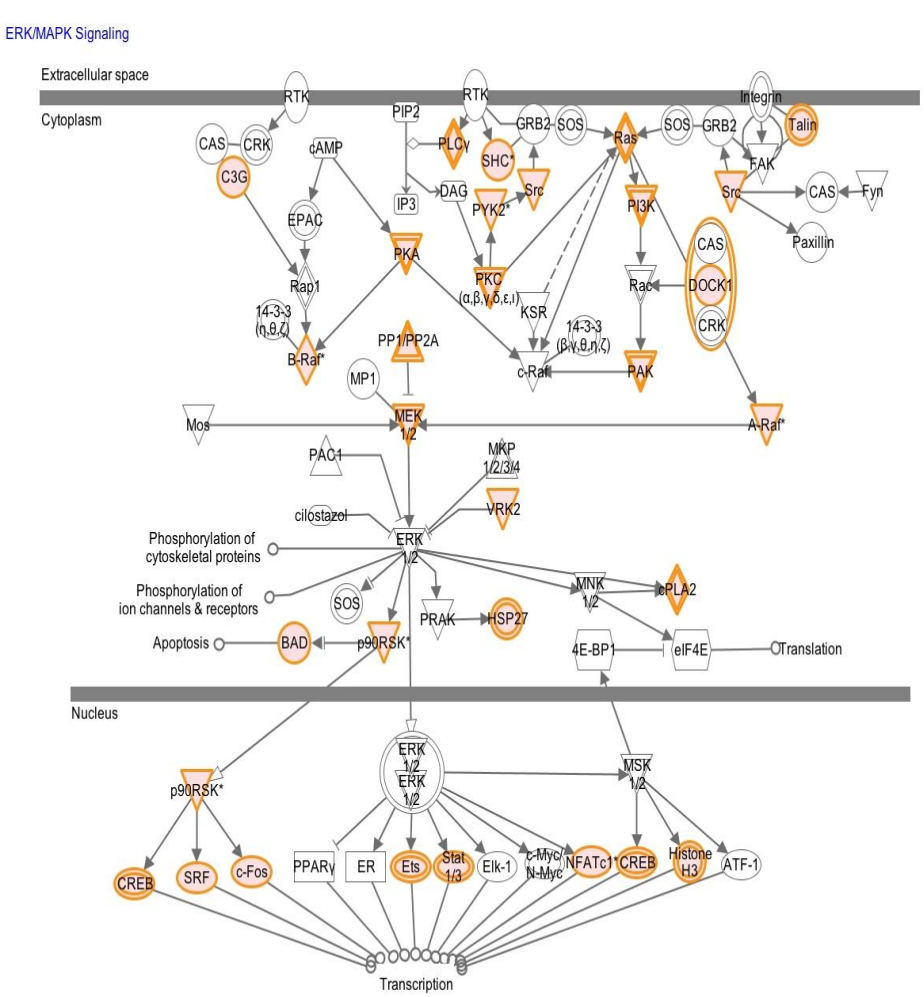
**A: CLL baseline pathways**

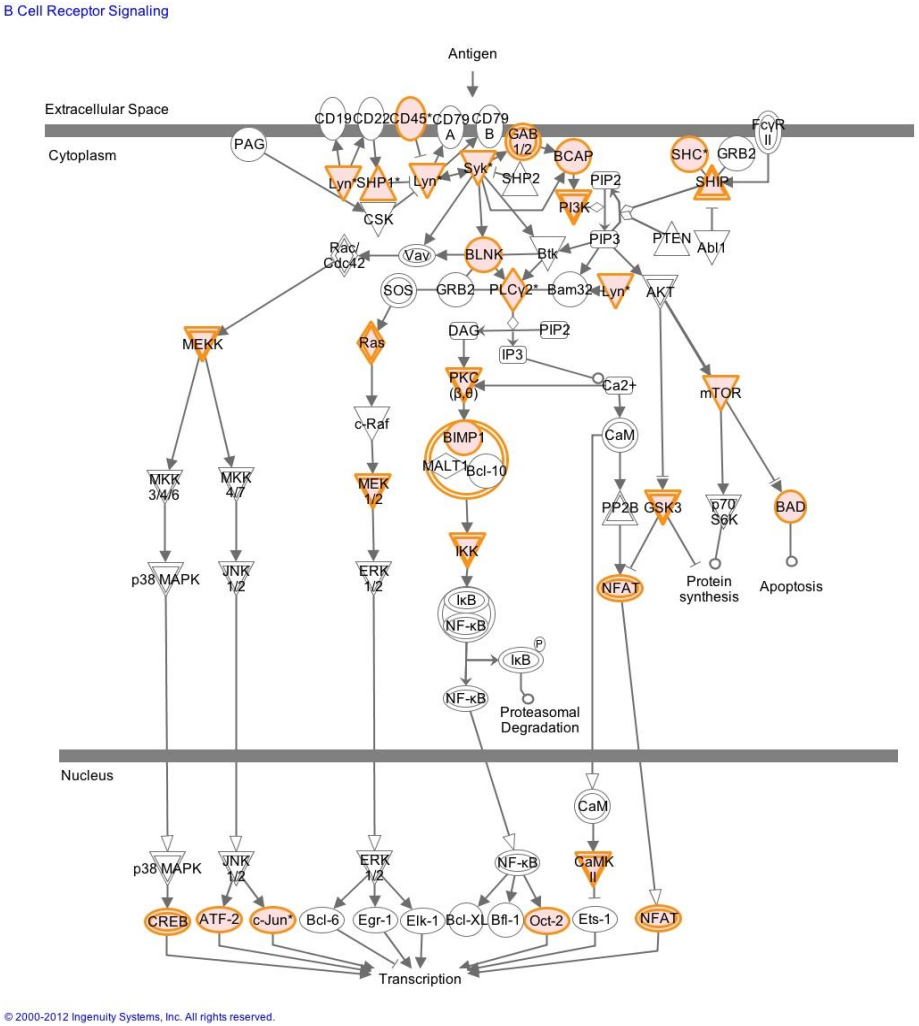
| **Ingenuity Canonical Pathways** | **p-value** | **Ratio** |
|---|---|---|
| ERK/MAPK Signalling | 1.62E-07 | 42/199 |
| DNA Methylation and Transcriptional Repression Signalling | 3.24E-07 | 11/23 |
| **B Cell Receptor Signalling** | **6.17E-07** | 34/157 |
| Protein Kinase A Signalling | 1.20E-05 | 61/384 |
| Synaptic Long Term Potentiation | 1.26E-05 | 26/115 |
| Telomerase Signalling | 1.70E-05 | 23/99 |
| Telomere Extension by Telomerase | 1.95E-05 | 8/16 |
| 14-3-3-mediated Signalling | 2.00E-05 | 25/114 |
| PI3K Signalling in B Lymphocytes | 3.89E-05 | 27/128 |
| EIF2 Signalling | 4.17E-05 | 32/182 |
| mTOR Signalling | 5.62E-05 | 33/190 |
| Phospholipase C Signalling | 6.03E-05 | 40/244 |

**B: Tonsil baseline pathways**

| Ingenuity Canonical Pathways | p-value | Ratio |
|---|---|---|
| ERK/MAPK Signalling | 8.13E-08 | 41/199 |
| DNA Methylation and Transcriptional Repression Signalling | 1.66E-07 | 11/23 |
| **B Cell Receptor Signalling** | **4.07E-07** | 33/157 |
| Telomere Extension by Telomerase | 1.17E-05 | 8/16 |
| Synaptic Long Term Potentiation | 1.20E-05 | 25/115 |
| Telomerase Signalling | 1.91E-05 | 22/99 |
| Protein Kinase A Signalling | 2.75E-05 | 57/384 |
| Estrogen Receptor Signalling | 3.55E-05 | 26/133 |
| Insulin Receptor Signalling | 7.24E-05 | 25/131 |
| DNA Double-Strand Break Repair by Non-Homologous End Joining | 7.24E-05 | 7/16 |

**Fig 7.22 ERK/MAPK and B-Cell Receptor pathways from IPA.**
**Highlighted in orange are the components of the pathway that are significantly enriched in the CLL dataset (all phosphoproteins detected).**



### *7.3.7 Kyoto Encyclopaedia of Genes and Genomes (KEGG) Pathway analysis*

The use of KEGG as an alternative method to indicate biological pathways may be complementary to Ingenuity. Table 7.6 shows the demonstrated KEGG pathways derived from the same datasets as Table 7.5

Below (Fig. 7.23) is the KEGG pathway with the BCR components from the CLL samples highlighted:

**Figure 7.23 KEGG BCR pathway.**
**Phosphoproteins enriched in CLL baseline samples are highlighted in red.**

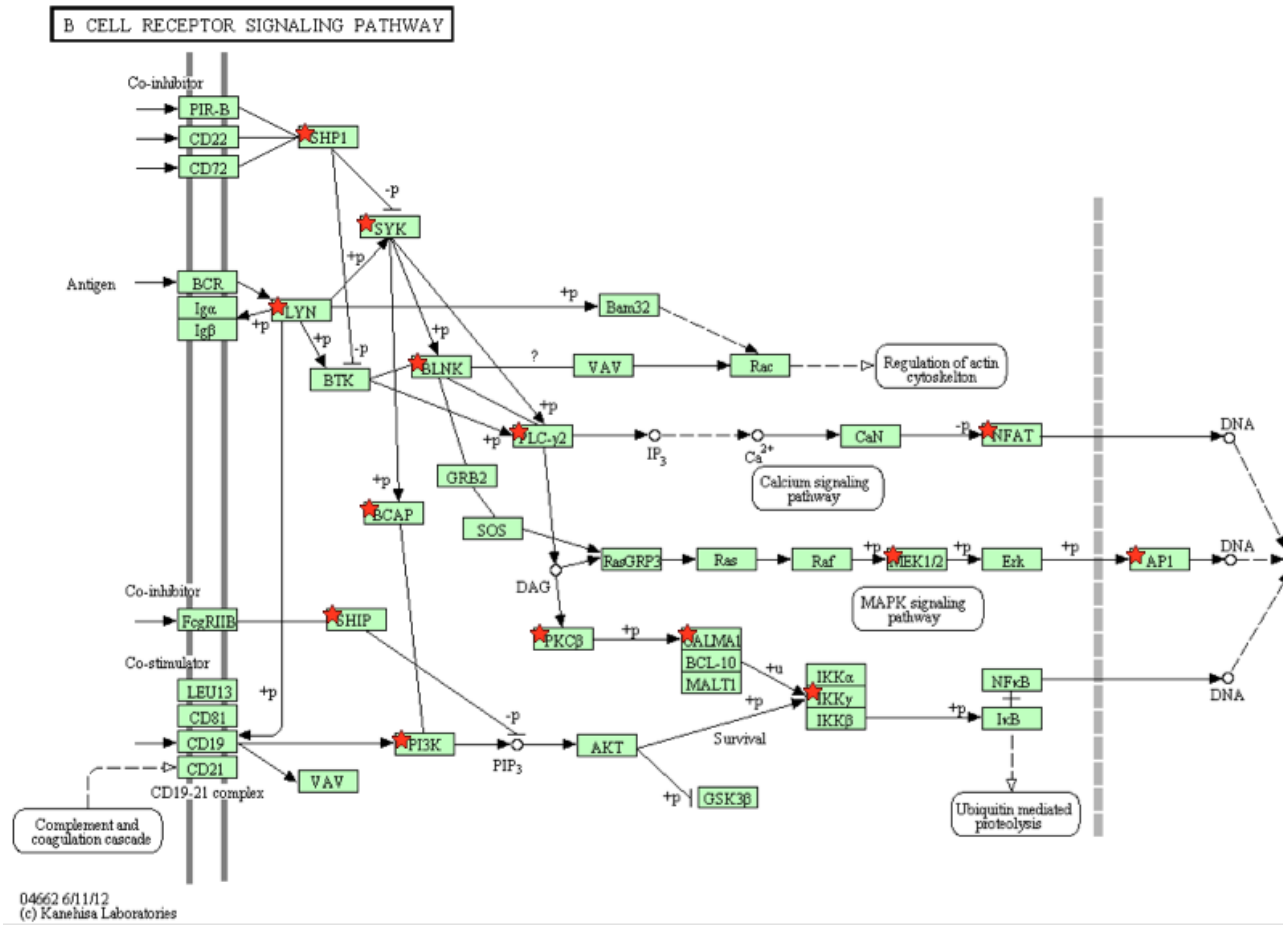


**Table 7.6 KEGG pathways significantly enriched in CLL and Tonsil baseline samples. P-values are from the modified Fisher's exact test. The ratio represents:** ***number of phosphoproteins in our dataset/total number of proteins in pathway***

**A CLL**

| KEGG pathway | p-value | ratio |
|---|---|---|
| Spliceosome | 0.0000 | 39/126 |
| **B cell receptor signalling pathway** | **0.0004** | **17/75** |
| Ribosome | 0.0007 | 18/87 |
| Fc gamma R-mediated phagocytosis | 0.0008 | 19/95 |
| Long-term potentiation | 0.0013 | 15/68 |
| Nucleotide excision repair | 0.0029 | 11/44 |
| Insulin signalling pathway | 0.0039 | 22/135 |
| Endocytosis | 0.0054 | 27/184 |
| Fructose and mannose metabolism | 0.0206 | 8/34 |
| Chronic myeloid leukemia | 0.0212 | 13/75 |
| MAPK signalling pathway | 0.0237 | 33/267 |
| Glycolysis / Gluconeogenesis | 0.0262 | 11/60 |
| ErbB signalling pathway | 0.0284 | 14/87 |
| GnRH signalling pathway | 0.0332 | 15/98 |
| Non-small cell lung cancer | 0.0340 | 10/54 |
| Glioma | 0.0355 | 11/63 |
| Phosphatidylinositol signalling system | 0.0432 | 12/74 |
| Pathogenic Escherichia coli infection | 0.0462 | 10/57 |

**B Tonsil**

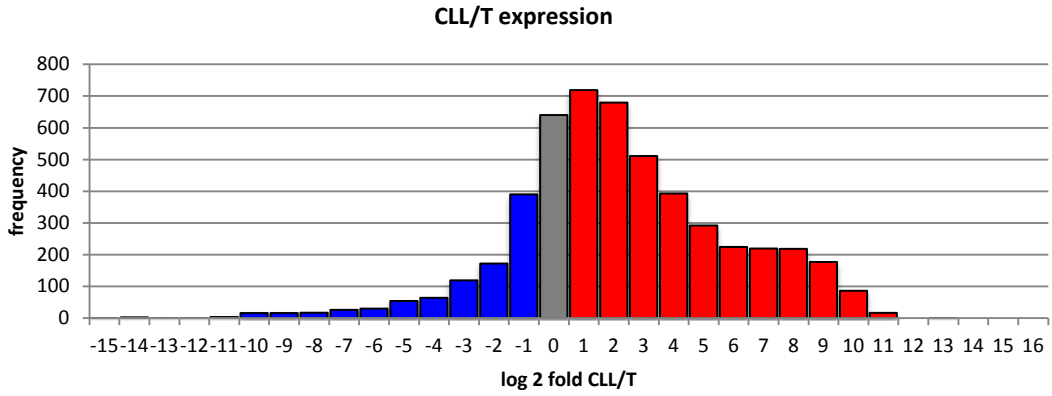
| KEGG pathway | p-value | ratio |
|---|---|---|
| Spliceosome | 0.0000 | 39/126 |
| Ribosome | 0.0003 | 18/87 |
| Fc gamma R-mediated phagocytosis | 0.0009 | 18/95 |
| **B cell receptor signalling pathway** | **0.0016** | **15/75** |
| Long-term potentiation | 0.0019 | 14/68 |
| Nucleotide excision repair | 0.0058 | 10/44 |
| Insulin signalling pathway | 0.0077 | 20/135 |
| Endocytosis | 0.0147 | 24/184 |
| ErbB signalling pathway | 0.0159 | 14/87 |
| Non-small cell lung cancer | 0.0218 | 10/54 |
| Glioma | 0.0221 | 11/63 |
| VEGF signalling pathway | 0.0288 | 12/75 |
| Pathogenic Escherichia coli infection | 0.0301 | 10/57 |
| Focal adhesion | 0.0373 | 24/201 |
| GnRH signalling pathway | 0.0386 | 14/98 |
| MAPK signalling pathway | 0.0393 | 30/267 |
| Fructose and mannose metabolism | 0.0443 | 7/34 |

The suggestion of enrichment of components of various pathways detected in baseline samples has several possible interpretations. These proteins may coincidentally be generally more abundant or more phosphorylated than others and are therefore more likely to be detected by the mass-spectrometry technique. It would be expected be that B-cells have high expression of proteins (and therefore phosphoproteins) known to be important in B-cell physiology, and are enriched for BCR pathway components. To support this, a comparison of the CLL and Tonsil proteins with the SAGE database of tissue expression of known proteins[548] suggested that both lists were enriched for proteins present in leucocytes, with the most significantly enriched tissue being 'lymph node Large B cell Lymphoma' ($p=9.4x10^{-29}$). An *in silico* comparison with similar data from non-B-cell tissues may confirm this. Constitutive phosphorylation of these pathway components may indicate either constitutive activation of the pathways, or indicate that there is a constitutive level of phosphorylation of unstimulated pathways. Alternatively a degree of baseline pathway activity may indicate the pathway is 'primed' for stimulation. This phosphorylation in the absence of overt stimulation could reflect simply the abundance of the pathway components. The high number of detected RNA processing proteins is reflected by the spliceosome KEGG signalling pathway enrichment.

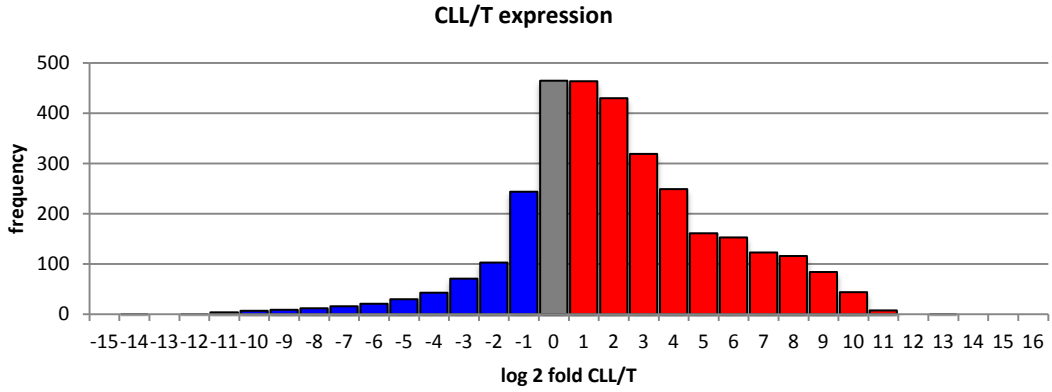
### *7.3.8 CLL vs Tonsil Comparisons*

In order to compare the phosphoproteome of leukaemia and healthy B-cells, the mean expression values within the baseline only conditions were compared**.** The mean expression levels of all CLL samples were divided by the mean values for all tonsils to provide a fold change CLL/T. These were then logged to base 2. Their distribution is a positively skewed normal distribution (see fig 7.24), indicating that perhaps more phosphosites are present in CLL cells compared to tonsils, interpreted as a greater range of phosphosites present in CLL, rather than overall greater abundance of phosphoproteins in CLL, though this is also possible.

**Figure 7.24: Distribution of CLL/T fold changes in baseline samples.**
**Red indicates higher expression in CLL compared to tonsil, blue lower.**
**A P dataset**




**B J dataset**




Because the data were approximately normal in distribution, unpaired t-tests were performed to highlight significant differences between CLL and tonsil, followed by False Discovery Rate (FDR) of 0.05 correction using the Benjamini-Hochberg (BH) method[538,539].

In order to reach significance, a particular phosphosite generally required a high fold change to be present in all CLL samples as compared to all Tonsil samples. A phosphosite highly upregulated in 5/6 CLL samples was therefore generally ignored using our approach. Significant mean fold changes ranged from 2.5-1400, with the median being 75-fold for the P dataset, 45-fold for the J dataset. There were numerous differentially expressed phosphosites, mainly upregulated (increase in CLL relative to tonsil). Shown below (Fig 7.25) are the volcano plots for all phosphosites, and the effect of Benjamini-Hochberg correction. Figure 7.26 gives an impression of the relative expression values in the most significantly altered phosphosite.

**Figure 7.25 Volcano plots of p-value *vs* fold change (P dataset).**
**A: Unpaired t-tests were performed to calculate a p-value for significant differences; B: p values after benjamini-hochberg correction. $Log_2$ fold change (mean expression in CLL samples divided by mean expression in tonsil samples) is also calculated**

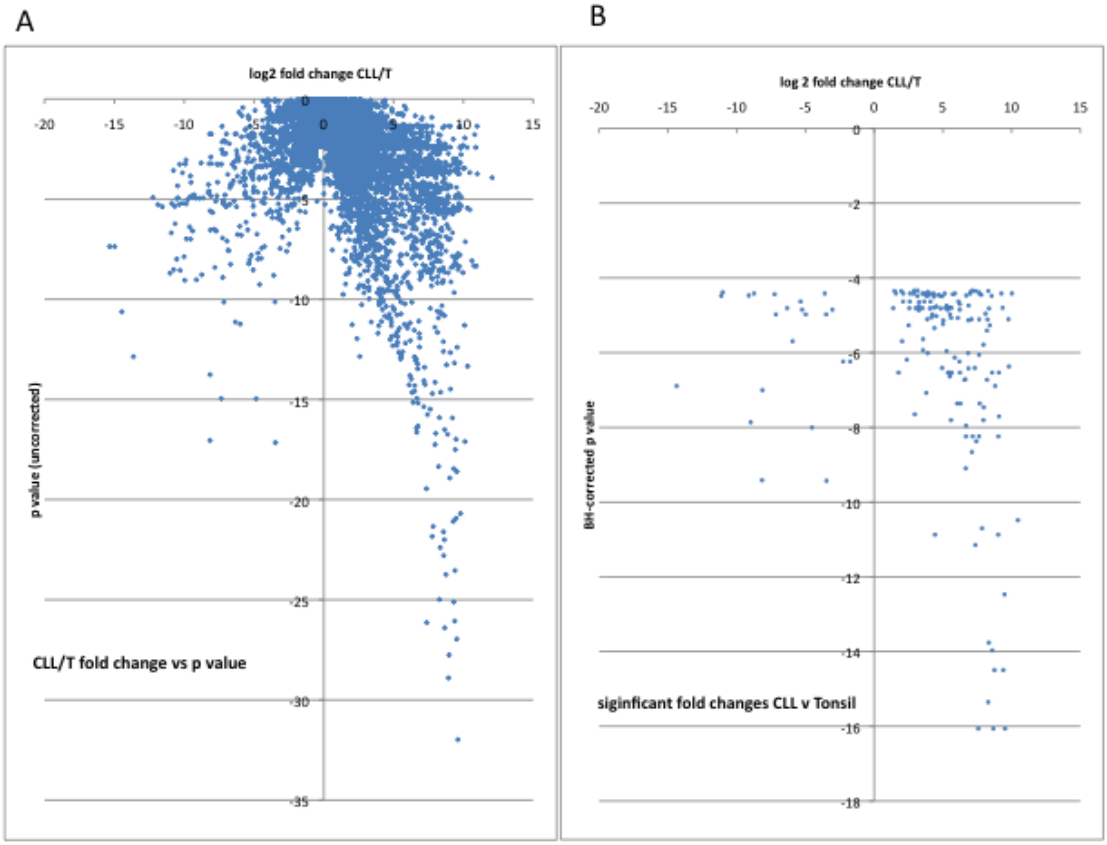
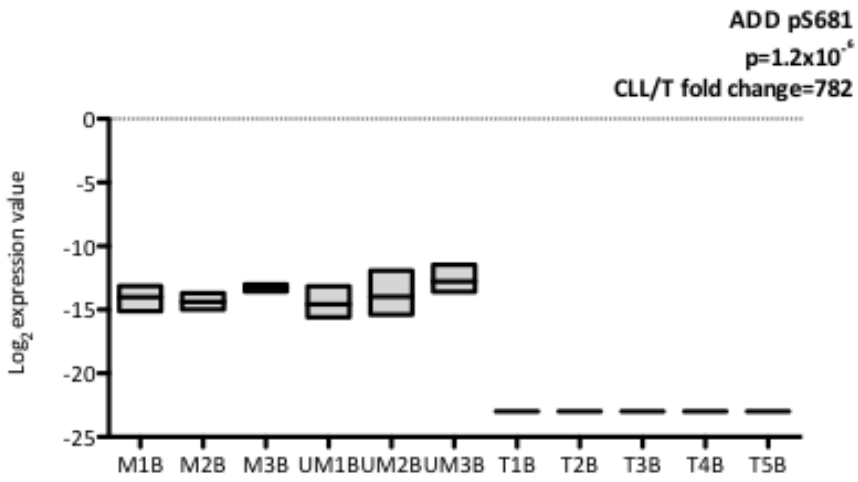



**Figure 7.26: Example mean expression values for the most significantly differentially regulated phosphosite**
**(see table 7.7). Boxes represent range of expression values of triplicates. Horizontal bar indicates mean.**



184/5088 (3.6%) phosphosites were differentially expressed between CLL and Tonsil in primary datasets, 169/3179 (5.3%) in joint datasets. These were not enriched for phosphotyrosine, as 1.1% were tyrosine and 99.5% ST (as compared to 1.7% and 98.9% in all baseline samples). There was an overlap of 113 phosphosites between the P and J datasets.

Table 7.7 shows the phosphosites exhibiting the greatest fold changes.

**Table 7.7: Top and bottom 10 most differentially phosphorylated phosphosites in P dataset.**

**Shown is the Uniprot ID, protein name, the phosphosite, log fold change, log corrected p-value of t-test, function, and upstream kinase. ?kinase indicates that the kinase-substrate relationship is not confirmed, but predicted using Scansite.**

**Top 10** **(upregulated in CLL v Tonsil)**

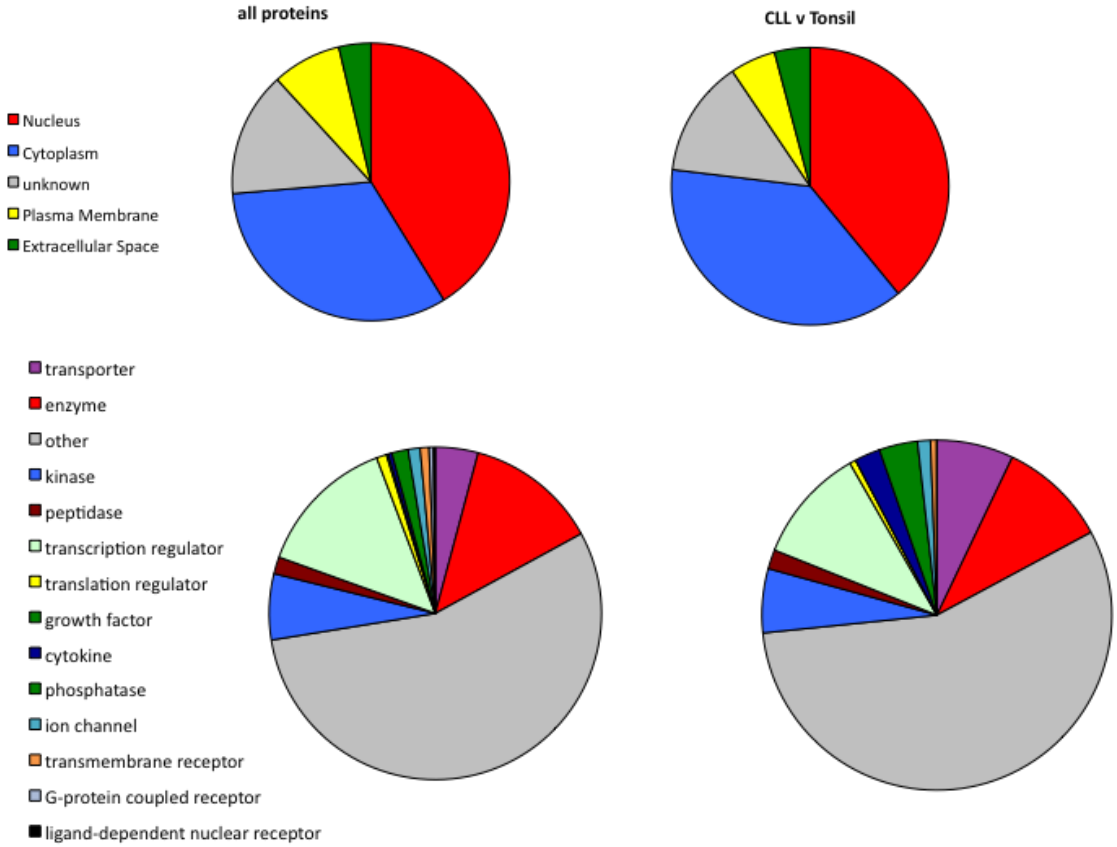
| Uniprot | protein | phosphosite | $log_2$ CLL/T | $log_2$ BH p value | function | kinase |
|---|---|---|---|---|---|---|
| **SAMN1** | SAM domain-containing protein SAMSN-1 | SAMSN1 p-S23 | 10.30 | -7.16 | Negative regulator of B-cell activation. Promotes RAC1-dependent membrane ruffle formation and reorganization of the actin cytoskeleton. Stimulates HDAC1 activity. Regulates LYN activity by modulating its tyrosine phosphorylation | ?AKT |
| **VIME** | Vimentin | VIM p-S26 | 10.12 | -9.80 | cytoskeletal component | PAK1 |
| **K1143** | Uncharacterized protein KIAA1143 | KIAA1143 p-S50 | 10.09 | -5.83 | ? | ?CK2 |
| **I2BP2** | Interferon regulatory factor 2-binding protein 2 | IRF2BP2 p-S175 | 9.79 | -12.86 | transcriptional repressor. | ?CDK1 |
| **ADDG** | Gamma-adducin | ADD3 p-S681 | 9.61 | -19.82 | Membrane-cytoskeleton-associated protein that promotes the assembly of the spectrin-actin network. Binds to calmodulin. | ?CDK1/CDC2/CDK5 |
| **SAMN1** | SAM domain-containing protein SAMSN-1 | SAMSN1 p-S90 | 9.55 | -6.54 | Negative regulator of B-cell activation. Promotes RAC1-dependent membrane ruffle formation and reorganization of the actin cytoskeleton. Stimulates HDAC1 activity. Regulates LYN activity by modulating its tyrosine phosphorylation | ?CAMK2/CK2/CLK2 |
| **AHNK** | Neuroblast differentiation-associated protein AHNAK | AHNAK p-S5400 + Formyl | 9.53 | -10.97 | May be required for neuronal cell differentiation | ?PKC |
| **CQ085** | Uncharacterized protein C17orf85 | C17orf85 p-S25 | 9.53 | -16.81 | ? | ?ERK1 |
| **DOC10** | Dedicator of cytokinesis protein 10 | DOCK10 p-T1440 | 9.48 | -13.03 | guanine nucleotide exchange factor (GEF) | ?AKT |
| **ACINU** | Apoptotic chromatin condensation inducer in the nucleus | ACIN1 p-S561 | 9.48 | -9.78 | Component of a splicing-dependent multiprotein exon junction complex | ?PKA |

**bottom 10** **(upregulated in Tonsil v CLL)**

| Uniprot | protein | phosphosite | $log_2$ CLL/T | $log_2$ BH p value | function | kinase |
|---|---|---|---|---|---|---|
| **SMC4** | Structural maintenance of chromosomes protein 4 | SMC4 p-S41 | -4.59 | -4.39 | SMC-4 is a core subunit of condensin I and II, large protein complexes involved in chromosome condensation | ?CDK1 |
| **FA40A** | Protein FAM40A | FAM40A p-S335 | -4.84 | -8.19 | Plays a role in the regulation of cell morphology and cytoskeletal organization. | ?GSK3B |
| **DDX21** | Nucleolar RNA helicase 2 | DDX21 p-S89 + Formyl | -5.98 | -5.66 | Can unwind double-stranded RNA (helicase) and can fold or introduce a secondary structure to a | ?CDK1 |
| **BCLF1** | Bcl-2-associated transcription factor 1 | BCLAF1 p-S531 + Formyl | -6.33 | -5.62 | single-stranded RNA (foldase). Functions as cofactor for JUN-activated transcription. Involved in rRNA processing | GSK3B |
| **KBTBA** | Kelch repeat and BTB domain-containing protein 10 | KBTBD10 p-S589 p-S592 | -7.15 | -4.95 | Required for pseudopod elongation in transformed cells. Substrate-specific adapter of an E3 ubiquitin-protein ligase complex which mediates the ubiquitination and subsequent proteasomal degradation of target proteins | ?CLK2 |
| **ROA2** | Heterogeneous nuclear ribonucleoproteins A2/B1 | HNRNPA2B1 p-S102 + Formyl | -7.35 | -8.19 | Involved with pre-mRNA processing | ?CLK2/AURB |
| **RL14** | 60S ribosomal protein L14 | RPL14 p-S139 | -8.15 | -7.36 | ribosome component | ?CDK1/CDK5 |
| **RAN** | GTP-binding nuclear protein Ran | RAN p-S135 + Formyl | -8.15 | -9.64 | GTP-binding protein involved in nucleocytoplasmic transport | ?PAK4/PLK1 |
| **K1C13** | Keratin, type I cytoskeletal 13 | KRT13 p-S427 | -13.64 | -6.73 | cytoskeletal component | ?CDK1/CDC2/CDK5 |
| **K1C13** | Keratin, type I cytoskeletal 13 | KRT13 p-S427 + Oxi + Formyl | -14.46 | -5.27 | cytoskeletal component | ?CDK1/CDC2/CDK5 |

The differentially phosphorylated protein classes were similar to those seen in all samples, with possibly a higher proportion of proteins with cytosolic localisation (Fig 7.27)

**Figure 7.27: Most abundant IPA molecular localisation function classes in differentially regulated phosphoproteins, CLL as compared to Tonsil. J dataset**




Kinases and phosphatases within each list of differentially regulated phosphosites were considered. 11/184 (P dataset) and 10/169 (J dataset) phosphosites were on kinases. 6/184 and 4/169 were phosphatases (Table 7.8). Of these CSF1R pY561 is an activating autophosphorylation, and PTPN6 (SHP-1) pY536 is an activating phosphorylation downstream of InsR, Lck & Src kinases. The other phosphorylations do not have published functions on the PhosphoSitePlus website[549].

**Table 7.8 phosphosites on kinases & phosphatases different between CLL & Tonsil.**
**J=mean fold change CLL/T phosphosites in J dataset.**
**P= mean fold change CLL/T phosphosites in J dataset**
**X=not significantly different in that dataset**

| Kinase psite | Name | J | P |
|---|---|---|---|
| CARD11 pS439 | caspase recruitment domain family, member 11 | 99 | 99 |
| CDK14 pS95 | cyclin-dependent kinase 14 | 11.4 | 21 |
| LYN pS13 | LYN kinase | X | 72 |
| MAP3K2 pS153 | mitogen-activated protein kinase kinase kinase 2 | 314 | 314 |
| NADK pS48 | NAD kinase | 37 | 37 |
| PANK2 pS168 | pantothenate kinase 2 | 115 | 112 |
| PGK1 pS203 | phosphoglycerate kinase 1 | 405 | 405 |
| PRKAB2 pS108 | protein kinase, AMP-activated | 77 | 77 |
| SGK223 pS694 | homolog of rat pragma of Rnd2 | X | 63 |
| SIK3 pS493 | SIK family kinase 3 | 9.4 | 9.4 |
| CSF1R pY561/T567/Y571 | colony stimulating factor 1 receptor | 53 | X |
| PFKL pS775 | phosphofructokinase, liver | 8.1 | X |

| Phosphatase psite | Name | J | P |
|---|---|---|---|
| INPP5F pS907 | inositol polyphosphate-5-phosphatase F | 274 | X |
| LPIN2 pS243 | lipin 2 | 248 | 248 |
| PTPN12 pS435 | protein tyrosine phosphatase, non-receptor type 12 | 5.9 | 6.6 |
| PTPN12 pS449 | protein tyrosine phosphatase, non-receptor type 12 | 160 | 90 |
| PTPN22 pS414 | protein tyrosine phosphatase, non-receptor type 22 | 303 | 280 |
| PTPN6 pY536 | protein tyrosine phosphatase, non-receptor type 6 (SHP-1) | 167 | X |

IPA analysis of the P dataset proteins suggests the phosphoproteins most upregulated in CLL cells compared to tonsil are enriched for components of tight junction signalling, RAN (GTP-binding protein involved in nucleocytoplasmic transport) signalling, ATM signalling, glucose metabolism and many other pathways (table 7.9). The J dataset comparison was similarly enriched for glucose metabolic pathways, but also cytoskeletal signalling and other pathways. KEGG analysis also suggested glucose metabolism pathways as important in CLL *vs* Tonsil comparisons.

**Table 7.9: bioinformatic analysis of differentially phosphorylated proteins, CLL *vs* Tonsil.**
**Shaded = p-value <0.05**
**Ratio represents proteins in list/total proteins in pathway.**
**P=primary dataset, J=Joint**

**A: IPA pathway analysis**

| Ingenuity Canonical Pathways | Ratio P | p value P | Ratio J | p value J |
|---|---|---|---|---|
| Tight Junction Signalling | 0.03 | 0.0076 | 0.02 | 0.1259 |
| RAN Signalling | 0.11 | 0.0079 | 0.06 | 0.1294 |
| Acetyl-CoA Biosynthesis III (from Citrate) | X | X | 1.00 | 0.0087 |
| ATM Signalling | 0.05 | 0.0138 | 0.02 | 0.3999 |
| Glycolysis I | 0.04 | 0.1722 | 0.09 | 0.0151 |
| Caveolar-mediated Endocytosis Signalling | 0.04 | 0.0229 | 0.04 | 0.0234 |
| Actin Cytoskeleton Signalling | 0.02 | 0.0955 | 0.02 | 0.0309 |
| Pentose Phosphate Pathway (Oxidative Branch) | 0.25 | 0.0339 | 0.25 | 0.0339 |
| Branched-chain α-keto acid Dehydrogenase Complex | X | X | 0.25 | 0.0339 |
| FAK Signalling | 0.03 | 0.0347 | 0.02 | 0.1633 |
| Protein Kinase A Signalling | 0.01 | 0.1959 | 0.02 | 0.0363 |
| Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes | 0.03 | 0.0407 | 0.01 | 0.5383 |
| GDP-glucose Biosynthesis | 0.17 | 0.0417 | 0.17 | 0.0427 |
| 3-phosphoinositide Degradation | X | X | 0.04 | 0.0468 |
| Glucose and Glucose-1-phosphate Degradation | 0.14 | 0.0501 | 0.14 | 0.0501 |
| Paxillin Signalling | 0.03 | 0.0513 | 0.03 | 0.0525 |
| Sertoli Cell-Sertoli Cell Junction Signalling | 0.02 | 0.0513 | 0.01 | 0.4140 |
| Systemic Lupus Erythematosus Signalling | 0.02 | 0.1754 | 0.02 | 0.0575 |
| Acetyl-CoA Biosynthesis I (Pyruvate Dehydrogenase Complex) | 0.14 | 0.0589 | X | X |
| **B Cell Receptor Signalling** | **0.02** | **0.1349** | **0.02** | **0.8431** |

**B: KEGG pathway analysis**

| KEGG pathways | Ratio P | p value P | Ratio J | p value J |
|---|---|---|---|---|
| Pentose phosphate pathway | X | X | 0.12 | 0.011 |
| Glycolysis / Gluconeogenesis | 0.05 | 0.057 | 0.05 | 0.057 |

**C:IPA biological functions enriched in differentially phosphorylated proteins**

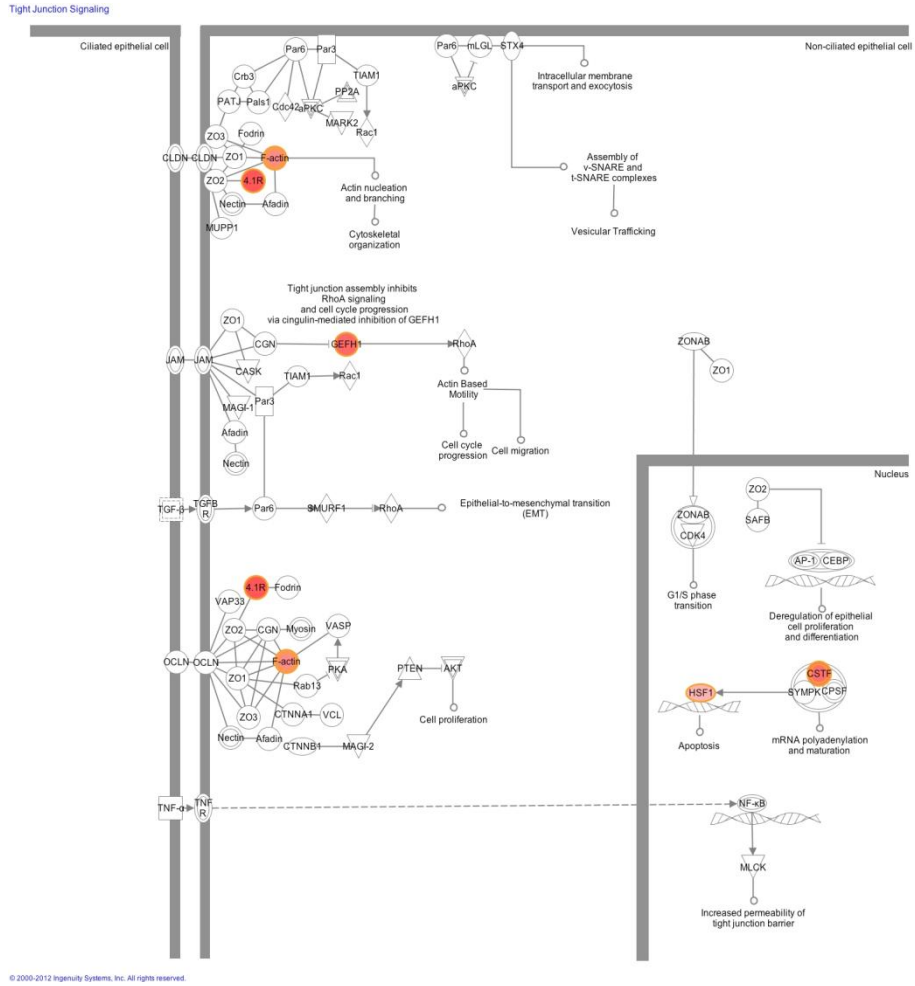
| Category | p-value P | p-value J |
|---|---|---|
| DNA Replication, Recombination, and Repair | 0.0085 | 0.0004 |
| Cell Morphology | 0.0005 | 0.0005 |
| Cardiovascular System Development and Function | 0.0005 | 0.0005 |
| Haematological Disease | 0.0005 | 0.0024 |
| Cancer | 0.0005 | 0.0024 |
| Nutritional Disease | 0.0011 | 0.0011 |
| Connective Tissue Disorders | 0.0011 | 0.0011 |
| Skeletal and Muscular Disorders | 0.0011 | 0.0011 |
| Metabolic Disease | 0.0011 | 0.0011 |
| RNA Post-Transcriptional Modification | 0.0029 | 0.0171 |
| Embryonic Development | 0.0031 | 0.0032 |
| Renal and Urological System Development and Function | 0.0031 | 0.0032 |
| Cellular Assembly and Organization | 0.0039 | 0.0032 |
| Cell Cycle | 0.0085 | 0.0032 |
| Hair and Skin Development and Function | 0.0038 | 0.0038 |
| Cellular Compromise | 0.0085 | 0.0038 |
| Gene Expression | 0.0039 | 0.0256 |
| Cellular Function and Maintenance | 0.0039 | 0.0086 |
| Cardiovascular Disease | 0.0053 | 0.0054 |
| Neurological Disease | 0.0053 | 0.0054 |
| Dermatological Diseases and Conditions | 0.0080 | 0.0086 |
| Immunological Disease | 0.0080 | 0.0086 |
| Tissue Development | 0.0085 | 0.0171 |
| Tissue Morphology | 0.0085 | 0.0171 |
| Molecular Transport | 0.0085 | 0.0171 |
| Connective Tissue Development and Function | 0.0085 | 0.0086 |
| Drug Metabolism | 0.0085 | 0.0086 |
| Tumour Morphology | 0.0085 | 0.0086 |
| Developmental Disorder | 0.0085 | 0.0086 |

**D: KEGG biological process enriched in differentially phosphorylated proteins**

| Term | P-Value P | P-Value J |
|---|---|---|
| haemopoiesis | 0.0001 | 0.0025 |
| haemopoietic or lymphoid organ development | 0.0002 | 0.0043 |
| immune system development | 0.0003 | 0.0060 |
| cytoskeleton organization | 0.0006 | 0.0211 |
| glucose catabolic process | 0.0011 | 0.0106 |
| lymphocyte differentiation | 0.0013 | 0.0013 |
| regulation of cellular response to stress | 0.0013 | 0.0087 |
| T cell differentiation | 0.0144 | 0.0017 |
| cell activation | 0.0019 | 0.0253 |
| hexose catabolic process | 0.0021 | 0.0169 |
| regulation of stress-activated protein kinase signalling pathway | 0.0021 | 0.0169 |
| DNA packaging | X | 0.0022 |
| monosaccharide catabolic process | 0.0023 | 0.0182 |
| leukocyte activation | 0.0029 | 0.0416 |
| regulation of protein kinase cascade | 0.0034 | 0.0135 |
| leukocyte differentiation | 0.0037 | 0.0037 |
| alcohol catabolic process | 0.0037 | 0.0257 |
| cellular carbohydrate catabolic process | 0.0044 | 0.0291 |
| chromosome organization | 0.0863 | 0.0047 |
| lymphocyte activation | 0.0047 | 0.0201 |
| actin cytoskeleton organization | 0.0087 | X |
| negative regulation of transcription | 0.0099 | X |
| regulation of MAPKKK cascade | 0.0106 | 0.0106 |
| carbohydrate catabolic process | 0.0106 | 0.0541 |
| RNA processing | 0.0106 | 0.0652 |
| induction of apoptosis by extracellular signals | 0.0578 | 0.0116 |
| regulation of protein amino acid phosphorylation | 0.0116 | X |
| actin filament-based process | 0.0117 | X |
| regulation of JNK cascade | 0.0144 | 0.0144 |
| microtubule-based process | 0.0146 | 0.0487 |

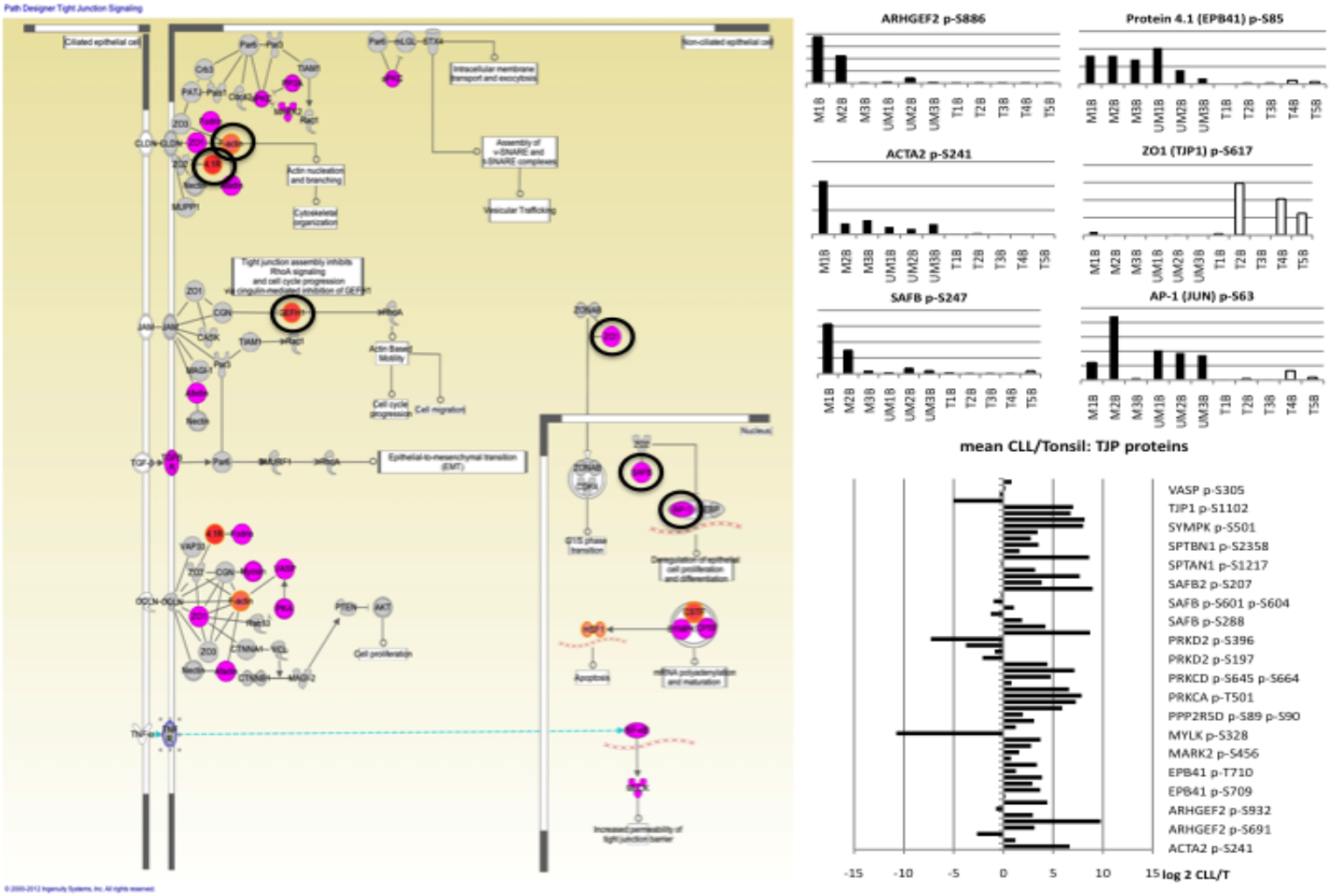
The Tight Junction pathway was the most significantly enriched pathway (p=0.0076) in the P dataset, yet the corresponding p-value for the J dataset was 0.13. Furthermore, as is often the case with gene set enrichment, a small number (5/157, yielding a ratio of 0.03) of genes from a pathway is enriched. In order to understand the Tight Junction Pathway components in our dataset further, we visualised the pathways using IPA. Figure 7.28 illustrates the 5 upregulated genes in the pathway:

**Figure 7.28: Tight Junction Pathway from Ingenuity Pathway Analysis**
**Red nodes represented significantly upregulated phosphoproteins (CLL v T) in the P dataset.**




Since the analysis used only highlighted phosphosites that were significantly different in all CLL (n=6) compared to all Tonsil (n=5) baseline samples, we attempted to visualise whether alterations in these proteins also occurred in a proportion of samples. We also sought to consider the same pathway in the J dataset. To this end, the overlapping proteins present in both our datasets and the IPA pathway were considered, and the expression levels in each sample visualised. Figure 7.29 is an illustration of this concept. It can be seen that the majority of tight junction pathway proteins had at least one phosphosite detected in our dataset, and that the majority of these had a greater than two fold upregulation (68%) or downregulation (12%) when CLL and Tonsil samples are compared. This illustrates the value of closer interrogation of pathways derived from bioinformatic tools, where use of data from stringent methods may underestimate the representation of pathway components.

**Figure 7.29: The IPA Tight Junction (TJ) pathway.**
**Proteins represented by phosphosites overexpressed in CLL v Tonsil are highlighted in red. In pink are those proteins that have at least one phosphosite detected. The 6 circled are representative examples showing the relative quantifications in each sample shown at the upper right. Below is the mean $log_2$ fold change CLL/Tonsil for each TJ component.**



### *7.3.9 Kinases upstream of CLL vs Tonsil differentially regulated phosphosites*

The CLL *vs* Tonsil differentially expressed phosphosites were considered in terms of known upstream kinases (Table 7.10). The PhosphSitePlus[549], PhosphoELM[550] and PhosphoPoint[551] databases were used for identifications. The fold change of a phosphosite was used to infer the different kinase activities. Where more than one substrate was identified, kinase activity was derived by using the geometric mean of fold changes for each substrate. Some had more than one upstream kinase.

8% of the phosphosites had an identified upstream kinase. There is substantial overlap between the two analyses, with multiple cAMP-dependent protein kinase alpha (PKACA), protein kinase C alpha (PKCA) and p21-activated kinases 1 and 2 (PAK1/2) substrates increased, indicating elevated activities of these kinases in CLL compared to Tonsil. Some phosphosite substrates have more than one upstream kinase capable of phosphorylating a particular site (substrate promiscuity) For example, the phosphosite WASP pY291 (CLL/T fold 107.1) could have been phosphorylated by multiple kinases (ABL, ACK, HCK, LCK, LYN, BTK and CK2).

**Table 7.10: Kinases of phosphosites significantly different between CLL & Tonsil.**
**The kinase and number of phosphosite substrates present are shown. The geometric mean CLL/T fold change of those substrates was used to derive a fold change of kinase activity. Primary (P) and Joined (J) datasets were analysed**
**P: 15/184 phosphosite substrates had at least one upstream kinase identified**
**J: 13/169 phosphosite substrates had at least one upstream kinase identified**

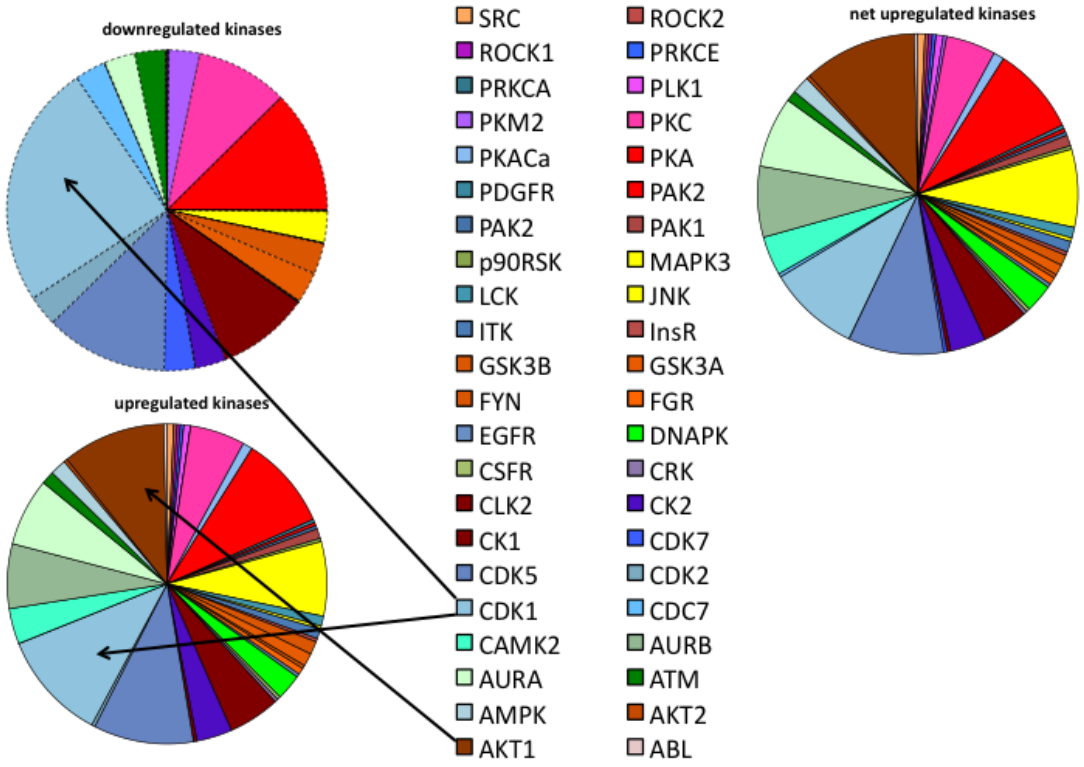
| kinase | substrate P | fold kinase activity P | substrate J | fold kinase activity J |
|---|---|---|---|---|
| PKACA | 4 | 65.1 | 5 | 45.1 |
| PKCA | 4 | 58.2 | 2 | 34.2 |
| PAK1 | 4 | 41.7 | 3 | 52.6 |
| PAK2 | 3 | 83.4 | 2 | 32.1 |
| p90RSK | 2 | 40.0 | 1 | 32.1 |
| CAMK2 | 1 | 489.1 | 0 | X |
| ABL | 1 | 107.1 | 1 | 167.5 |
| ACK | 1 | 107.1 | 0 | X |
| HCK | 1 | 107.1 | 0 | X |
| LCK | 1 | 107.1 | 1 | 167.5 |
| LYN | 1 | 107.1 | 0 | X |
| BTK | 1 | 107.1 | 0 | X |
| CK2 | 1 | 107.1 | 0 | X |
| HIPK2 | 1 | 105.7 | 0 | X |
| CDK1 | 1 | 63.8 | 0 | X |
| JNK | 1 | 20.7 | 1 | 20.7 |
| MAPK8/9/10 | 1 | 20.7 | 1 | 20.7 |
| PDHK1 | 1 | 15.7 | 0 | X |
| PDHK2 | 1 | 15.7 | 0 | X |
| PDHK3 | 1 | 15.7 | 0 | X |
| PDHK4 | 1 | 15.7 | 0 | X |
| PDPK1 | 1 | 15.7 | 0 | X |
| ROCK1/2 | 1 | 5.2 | 1 | 5.2 |
| AURKB | 1 | 5.2 | 1 | 5.2 |
| GSK3B | 1 | 0.01 | 0 | X |
| INSR | 0 | X | 1 | 167.5 |
| SRC | 0 | X | 1 | 167.5 |
| CSF1R | 0 | X | 1 | 52.9 |
| BCKDK | 0 | X | 1 | 19.8 |

### *7.3.10 Prediction of upstream Kinases*

The Scansite algorithm predicts upstream kinases based on amino acid residue motifs, and has been used to predict upstream kinases in a number of publications[523,552,553]. To assess the utility of this method with our dataset, upstream kinases were predicted for the 169 differentially regulated phosphosites in the J datasets. There were two groups within the dataset: Phosphosites decreased in CLL relative to tonsil (17% of those significantly differentially expressed) and those increased. The kinase activity is implied by the fold difference. There are therefore two groups of predicted kinases: those with higher activity in tonsil relative to CLL, and those with higher activity level in CLL.

Each phosphosite substrate had 0-10 predicted upstream kinases, and each predicted kinase had a mean of 7.7 substrates. This partly reflects the 'promiscuity' of kinases and phosphosites, and probably also the inaccuracy of Scansite as a kinase-substrate prediction tool. A number of kinases have predicted substrates that increased in tonsil *vs* CLL, and some that are decreased, which leads to paradoxical and opposing implied kinase activities. There are several possible explanations for this. Firstly, we have used an averaged ratio that does not fully take into account differences between samples, though this is mitigated by the statistical method used. The co-localisation of a kinase and its substrate is an important mechanism of regulation, and it may be that within normal B-cells, a highly active kinase may be reflected by a highly phosphorylated substrate, whereas the same active kinase in CLL cells may be localised in a different cellular location to its substrate, or that the substrate is at lower abundance, resulting in an apparent lower kinase activity. To a degree, the promiscuity of kinases/substrates and phosphatases may also account for these differences. Finally, the prediction algorithm of Scansite is likely to be inaccurate in many cases, and demonstrates the limitations of kinase-substrate predication based on motif recognition alone. To attempt to derive an overall impression of 'net' kinase activity, the number of substrates that are downregulated is subtracted from those upregulated, but this is a crude method and care must be taken that this should not be over interpreted. See figure 7.30

**Figure 7.30: Predicted upstream kinases down- or upregulated in CLL compared to tonsil.**
**The number of substrates per predicted kinase is reflected in the proportion of the pie chart. The net upregulated kinases is derived by subtracting the number of kinase substrates downregulated from those upregulated**



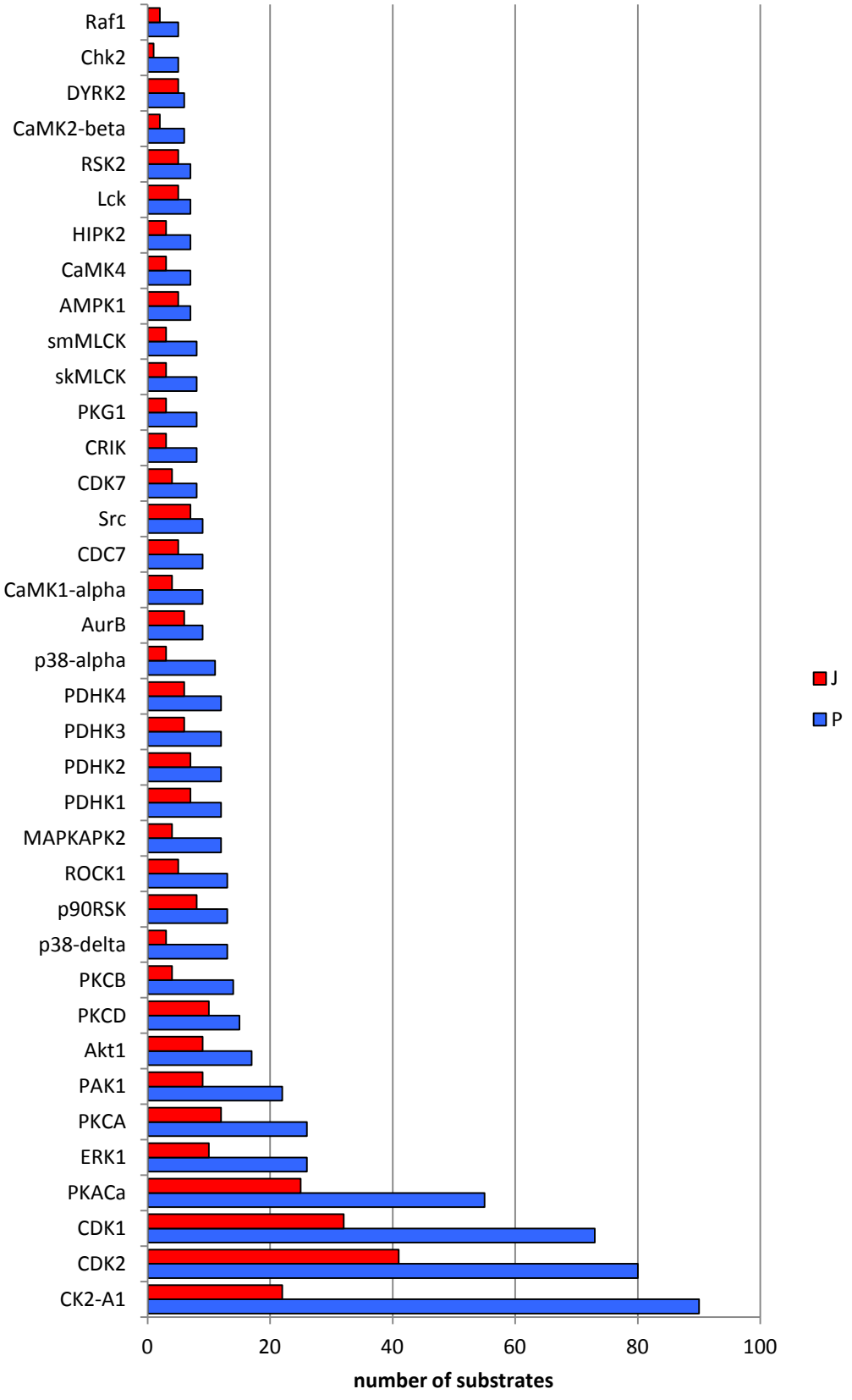

As an illustration of the limits of the method, there appears to be a number of substrates of CDK1 paradoxically upregulated in both CLL *vs* Tonsil and Tonsil *vs* CLL. This is also true of many of the other kinases. However, some kinases such as AKT1 appear to be consistently upregulated in CLL *vs* Tonsil.

### *7.3.11 Alternative Strategies for considering kinase activity*

The above analyses are dependent on a list of phosphoproteins derived from comparing leukaemia and tonsil samples, and are therefore critically dependent on whether tonsils are the most appropriate control. The analysis method is analogous to that used for gene expression data. Unlike gene expression data, which is derived from comparison of control and disease mRNA levels hybridised to arrays, the phosphoproteomic data produce expression values of phosphoproteins that are independent of control comparisons. If CLL samples alone had been used, the data will still be valid in terms of expression of phosphoproteins, and by extension their upstream kinases. This therefore permits a consideration of kinases independent of the use of tonsil controls, and suggests alternative analysis methods. The estimation of kinase activity using all phosphosite data may be more powerful than consideration of only the limited number of CLL *vs* Tonsil differentially expressed phosphosites.

The upstream kinases for all P dataset phosphoproteins were ascertained. 433/5088 (8.5%) of features had an identified kinase. The most frequent identified kinase was Casein kinase 2 (CK2-A1) with 90/5088 substrates. Numerous substrates of CDK1, CDK2, PKACa, PKCA, ERK1 and PAK1 were also present. In the J dataset, 208/3179 (6.5%) of phosphosites had an identified upstream kinase, with CK2 associated with 41/3179 phosphosites. These data are shown in fig 7.31

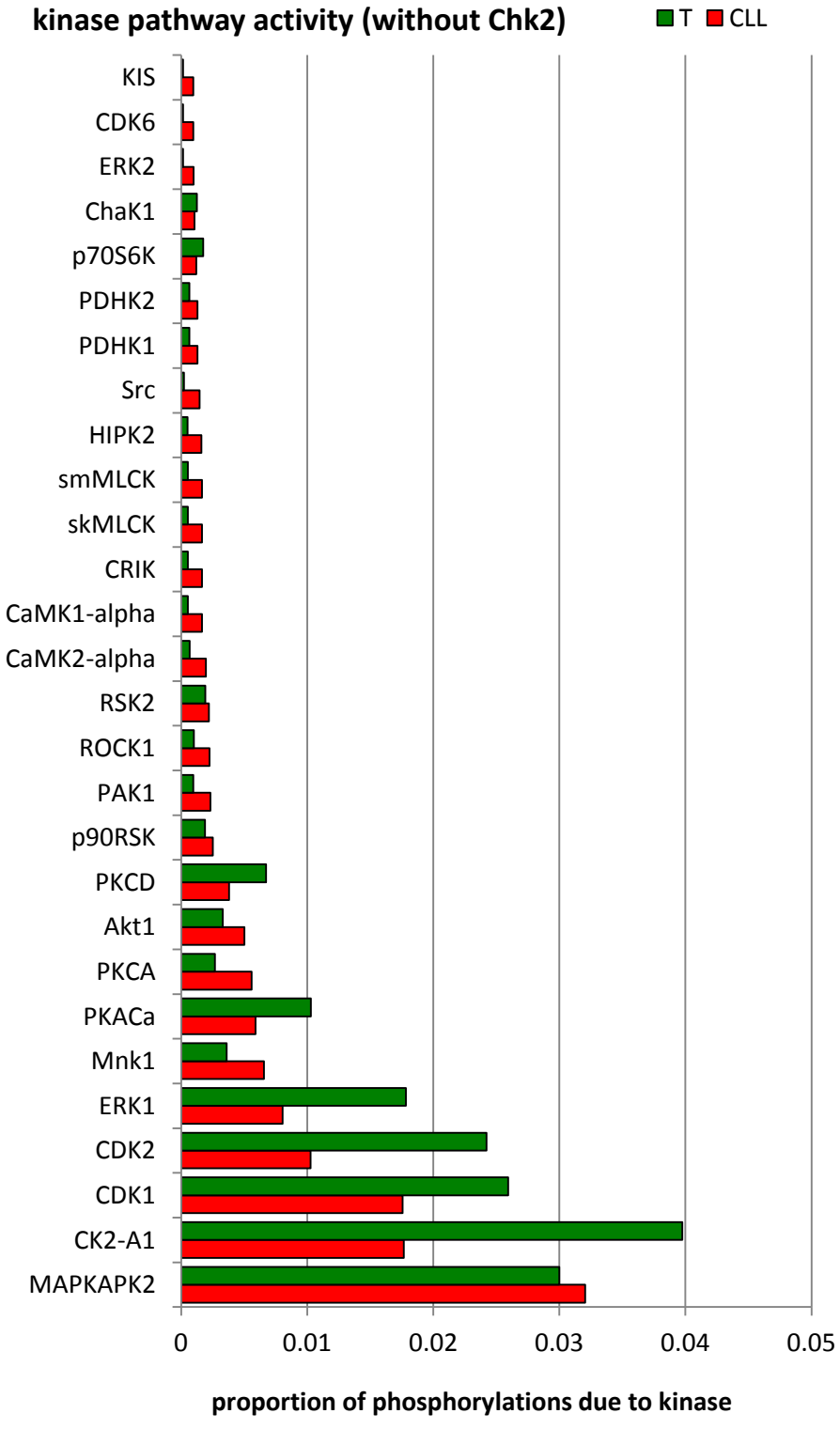
**Figure 7.31: upstream kinases for P & J dataset phosphosites.**
**The number of substrates for each kinase is shown. Kinases with at least 5 substrates shown.**



Note that one of the reasons for a large number of substrates for a given kinase is a bias derived from the literature, i.e. CK2 is well known to phosphorylate numerous substrates, and these have been reasonably well characterised, possibly creating an inflated estimation of its activity in experiments, though it is known to be a ubiquitously expressed kinase so might be expected to have high activity.

To derive a better measure of kinase activity, the phosphosite expression values for each kinase were summed, providing a number representing the proportion of total phosphosites detected in a sample that can be attributed to the action of a particular kinase. This is independent of J and P dataset differences. This analysis was performed for the mean expression values of CLL baseline samples and tonsil baseline samples. For example, the most active kinase derived from this method is Chk2, with 17% of the phosphosite expression within the CLL samples due to the Chk2-dependent phosphorlylation of Histone 1.2, the most abundant phosphoprotein within the dataset. Note that this is a product of both kinase activity and substrate abundance. Because this measure is not of 'pure' kinase activity, but of the pathway that involves Chk2 phosphorylating Histone 1.2, this derived measure is given the name of 'kinase pathway activity'. The top 30 kinase pathway activities (excluding Chk2) are shown in figure 7.32.

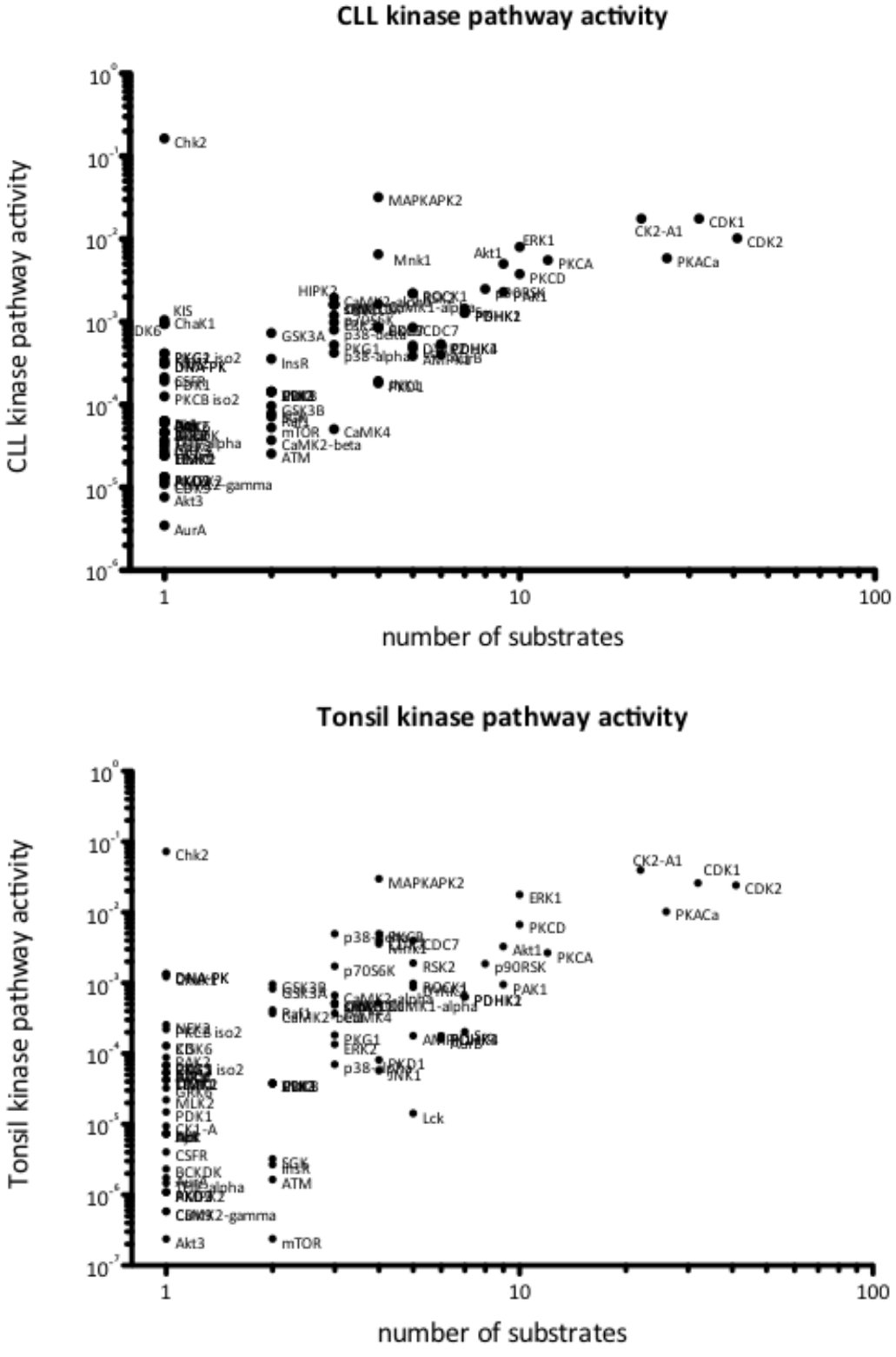
**Figure 7.32: Kinase pathway activity.**
**The mean value of all CLL & tonsil (T) baseline samples was derived for each phosphosite. The values for each phosphosite attributable to a particular kinase were summed. This provides a measure of the proportion of phosphorylations within a sample due to a particular kinase. The top 30 kinase pathway activities are shown. The data for Chk2 is not shown, as these dominate the data, with the proportion of phosphorylations due to Chk2 at 0.17 for CLL, 0.07 for T, making visualisation of the other kinases difficult.**



It might be expected that the more substrates available for measurement, the greater the confidence in this derived value for kinase pathway activity. In theory, the combination of the number of substrates with the kinase pathway activity may permit formal statistical comparisons to be made, in order to derive quantitative measures of confidence regarding the kinase pathway actvity. This is rather complex and beyond the scope of this thesis, and Dr P Cutillas and colleagues are working on such methods. Nevetheless, the principle is illustrated that kinase activity can be inferred from consideration of all phosphosites simultaneously, and independent of comparisons with controls.

Figure 7.33 shows a plot of kinase pathway activity against the number of substrates. This gives an overall impression of kinase activity within the CLL and Tonsil samples. It can be seen that CDK1/2, CK2-A1 and PKACa have numerous substrates detected, as well as high kinase activity, as might be expected. Chk2 has a high kinase activity, but represented by one highly abundant substrate.

**Figure 7.33: Kinase pathway activity *vs* number of substrates**
**Proportion of phosphosites in sample attributable to the action of a particular kinase *vs* number of substrates for that kinase. Shown are the data for the mean of all CLL and Tonsil baseline samples.**




Clearly, it appears that CLL and tonsil kinase pathway activity patterns are similar. This is unsurprising, as it would be expected that the 'core' B-cell phosphoproteome would

be common betweeen CLL and Tonsil, sharing many substrates and kinases. To highlight differences, a return to a comparison of CLL and Tonsil is required, and this is illustated in Figure 7.34 and 7.35

**Figure 7.34: CLL/Tonsil Kinase pathway activity *vs* number of substrates.**
**Proportion of phosphosites in sample attributable to the action of a particular kinase *vs* number of substrates for that kinase. The kinase pathway activity of CLL is divided by that for tonsil.**

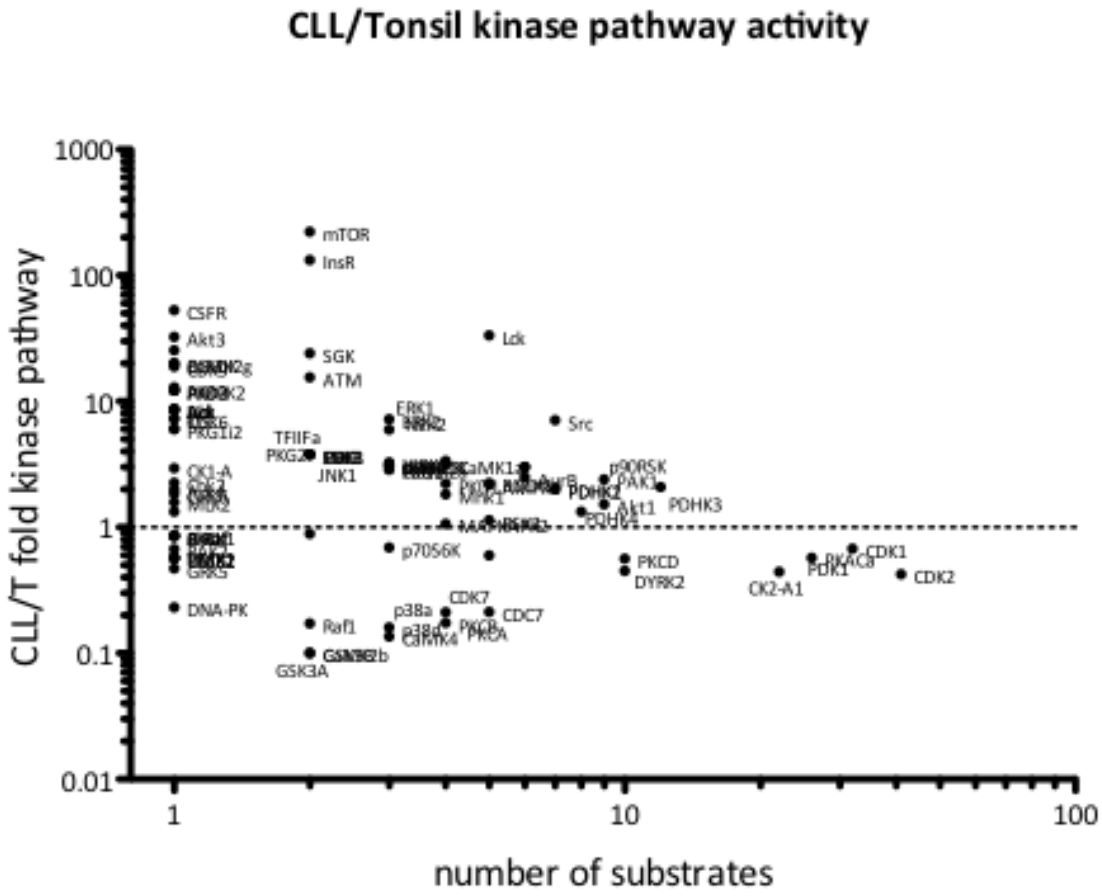
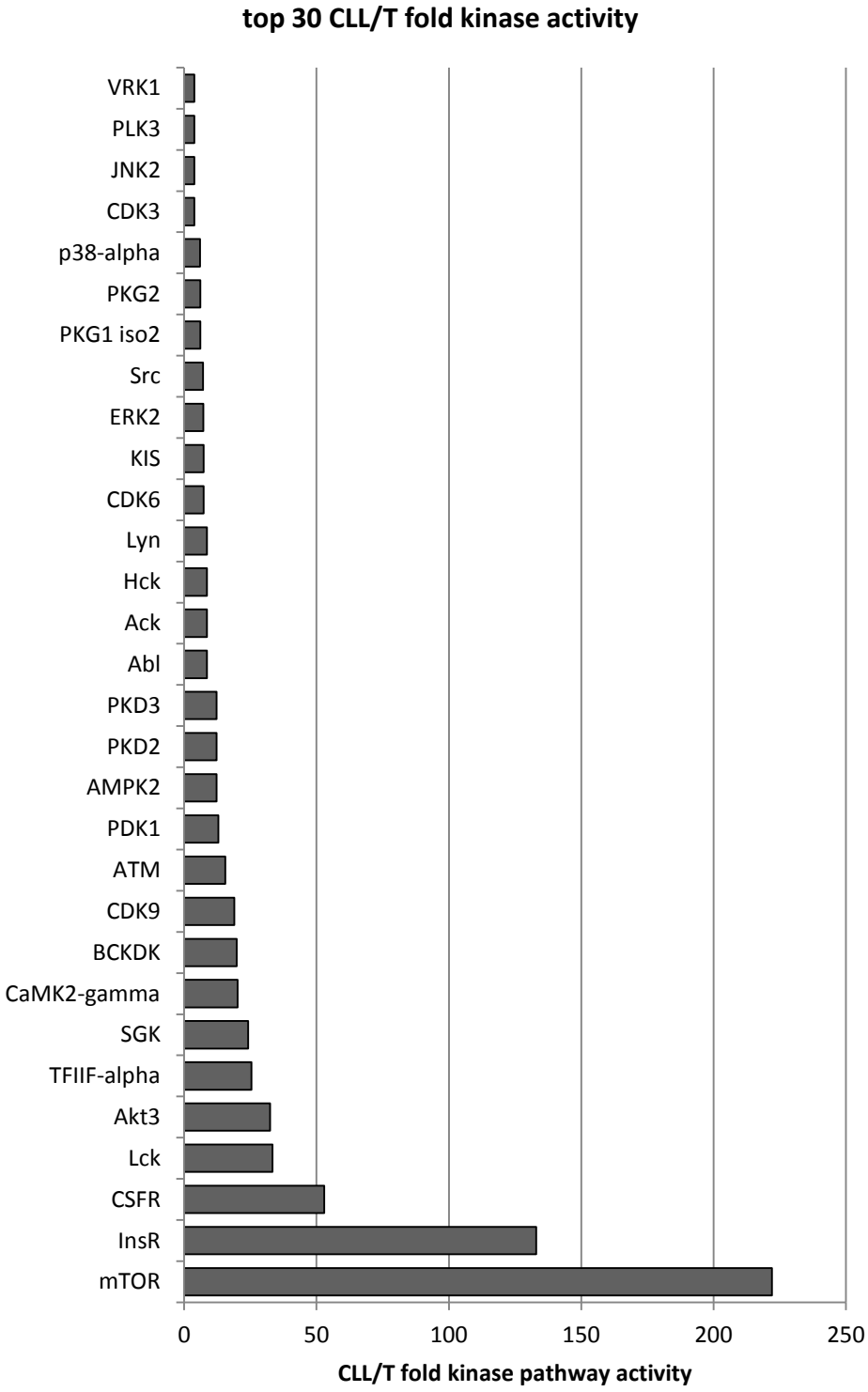
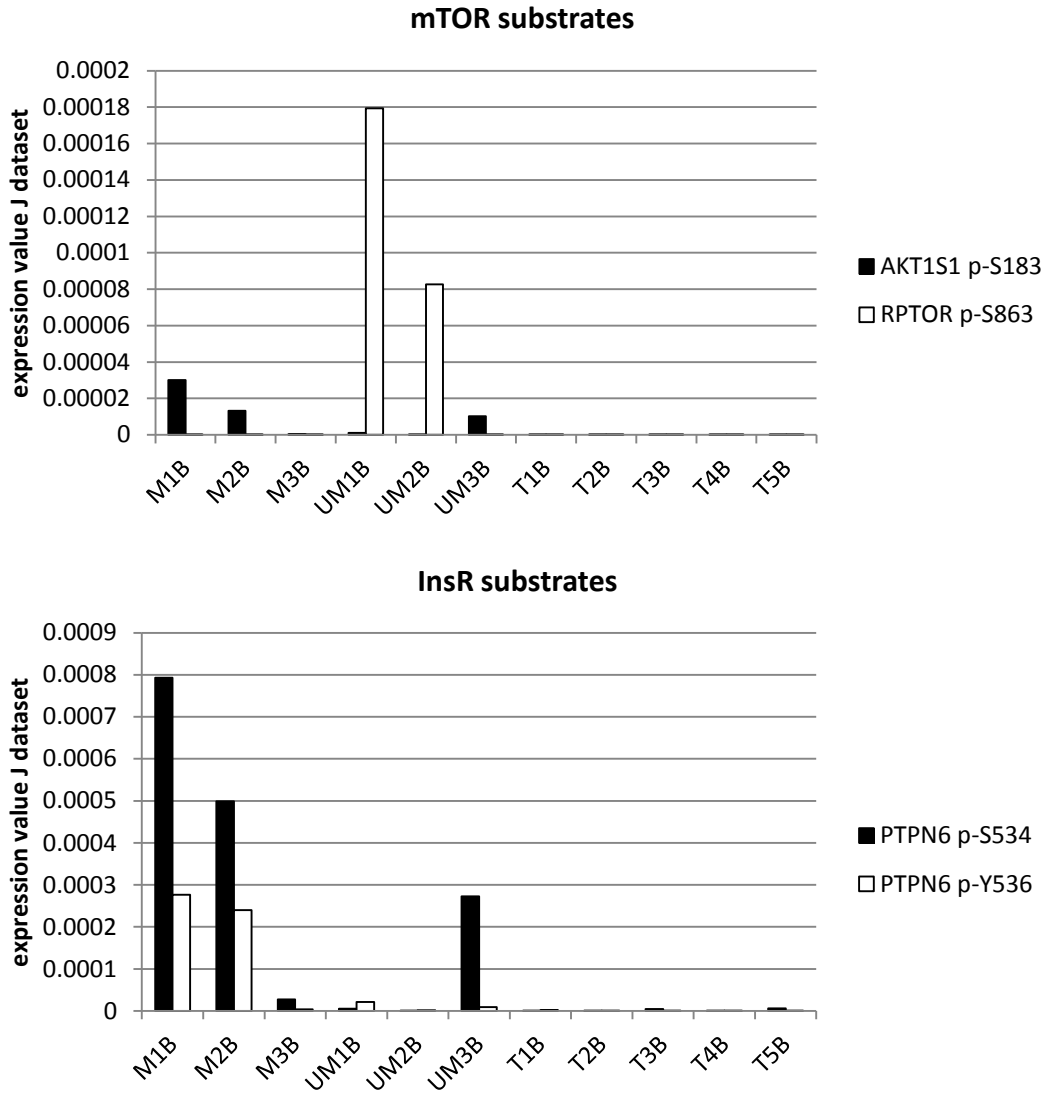


**Figure 7.35: CLL/Tonsil Kinase pathway activity, top 30 kinases**
**Proportion of phosphosites in sample attributable to the action of a particular kinase. The kinase pathway activity of CLL is divided by that for tonsil. The top 30 kinases are shown**



It can be seen that mTOR and the insulin receptor (InsR) kinase pathway activities are highest, yet each is represented by only 2 substrates. By consideration of expression values of these 2 substrates within each sample, (figure 7.36) it can be seen that mTOR and InsR activity are generally higher, but that not every CLL sample exhibits a high phosphosite level, explaining why these kinases are not highlighted by the previous statistical comparison of CLL *vs* Tonsil (see section 7.3.7)

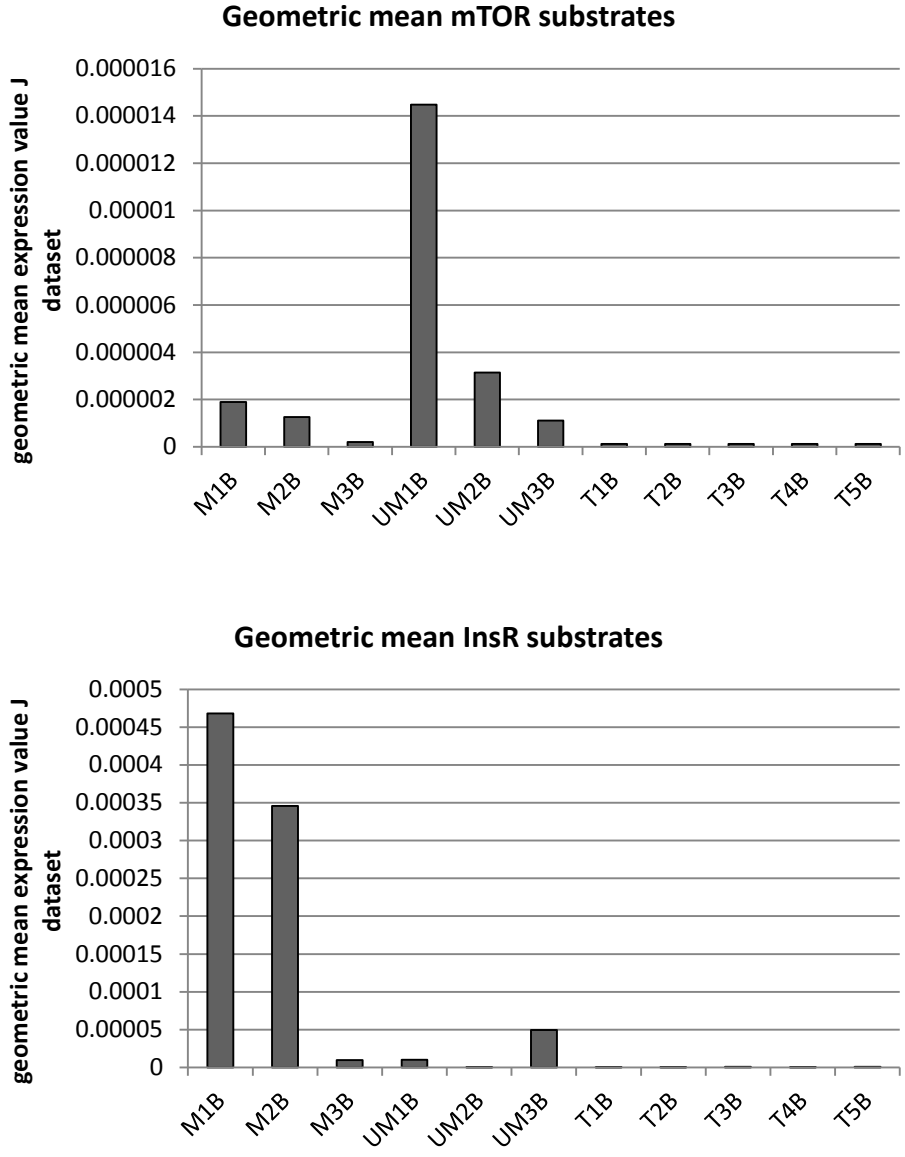
**Figure 7.36: Expression values of mTOR and InsR substrates within J dataset for each baseline sample**




It can also be seen that the highly expressed substrates contribute more to the derived kinase pathway activity measure, as this is derived from the sum of each phosphosite. To derive a measure of kinase activity that is less dependent on the substrate expression level, the geometric mean of each substrate in each sample was derived

(figure 7.37). This may give a better estimate of the kinase activity independent of the level of substrate. Further work in this kind of analysis is required, in order to add statistical robustness to these measures, but this demonstrates the potential for estimating kinase activity based on phosphoproteomic data.

**Figure 7.37 mTOR and InsR kinase activity for each sample.**
**The geometric mean of expression values for each substrate was derived for each sample, to yield a measure of kinase activity less dependent on absolute substrate expression**




Whilst a potentially interesting method, this analysis is still partly dependent on the abundance of substrates, on the comparisons between CLL and Tonsil, and as yet lacks the statistical rigour to establish a measure of confidence in the quantifications of kinase activity. This is discussed further.

## 7.4 Results: BCR signalling

### *7.4.1 Introduction*

Our hypothesis is that BCR signalling in CLL cells differs from that in healthy B-cells. We therefore seek to demonstrate differentially phosphorylated proteins between baseline and the IgM & IgD stimuli. Clustering analysis indicates that samples cluster according to patient rather than stimulus, so differences between patients are likely to be greater than between baseline and stimulated conditions in the same patient.

BCR signalling was assessed using two analysis methods. In both methods, the fold change in phosphosites was calculated by dividing the mean of M or D replicates by the mean of B replicates.

**1. Paired t-tests comparing means of replicates**

Each replicate is used as a technical replicate and the mean derived. The means of stimulated samples are then compared to baseline means. This should reveal common changes shared by patients.

$B_1,B_2,B_3,B_x$ *vs* $M_1,M_2,M_x$ etc

**2. Unpaired t-tests comparing stimulated replicates to baseline replicates within each sample.**

Each patient is treated as a different experiment, statistical comparison is made between the replicates within each patient

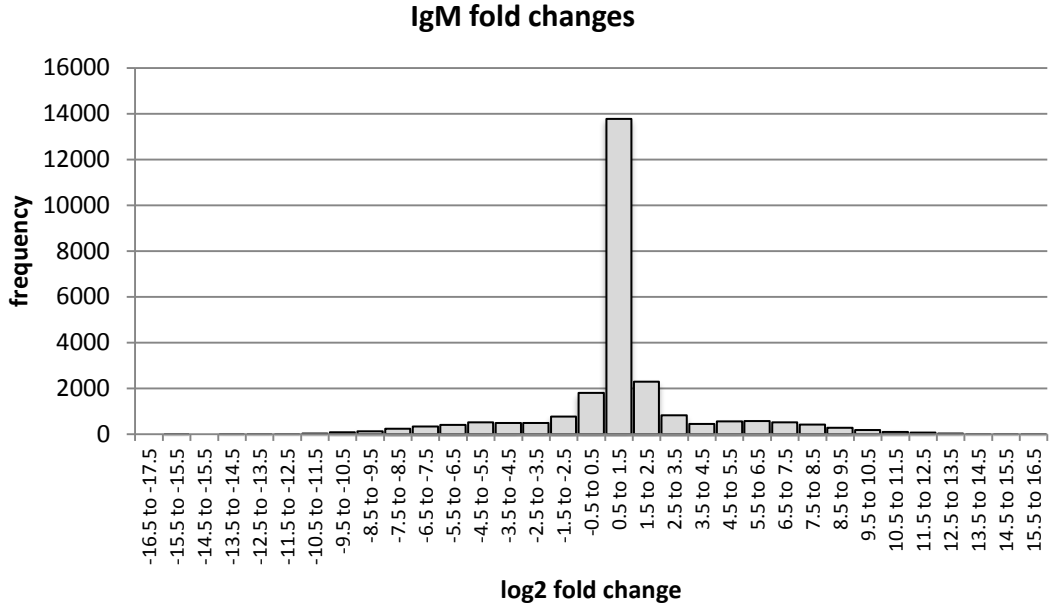
$B_{1a},B_{1b},B_{1c}$ *vs* $M_{1a},M_{1b},M_{1c}$ etc

It is envisaged that due to biological variation between patients, the results of stimulation may well appear very different in each patient, and this is indeed the case.

### *7.4.2 Paired t-tests comparing means of replicates to describe BCR signalling*

Mean $\log_2$ values were compared between the 5 tonsil baseline samples and their stimuli. When log fold changes are displayed in a histogram (Fig 7.38) it can be seen that the majority (74%) are less than two-fold.

**Figure 7.38 Log distribution of fold changes after IgM stimulation, P dataset.**




Fold changes appeared normally distributed, and so paired t-tests were felt appropriate and used to analyse differences between the stimuli and baseline. The differences between baseline and IgM/IgD mean (of triplicates) phosphosite quantifications were considered for each of the 5 tonsil and 6 CLL samples. This method should detect phosphosites consistently up/downregulated in all CLL or tonsil samples after stimulation. Table 7.12 shows that the number of significantly altered phosphosites was low, and in fact fell to zero when more stringent p-values using Benjamini-Hochberg corrections were used:

**Table 7.12: Number of differentially regulated phosphosites at different p values for Tonsil and CLL samples, P dataset.**

| Uncorrected P-value | Tonsil | | CLL | |
|---|---|---|---|---|
| | IgM | IgD | IgM | IgD |
| p<0.05 | 41 | 63 | 118 | 180 |
| p<0.01 | 8 | 7 | 11 | 22 |
| p<0.001 | 0 | 1 | 0 | 1 |

The most significant tonsil IgD-induced phosphosite at uncorrected p=0.0006 was Myosin regulatory light polypeptide 9 pS20, at 18-fold increase. It is involved in actin polymerization, so is a plausible candidate for involvement in BCR signalling. The most significantly altered after IgM signalling in tonsil (p=0.00157) was Actin-related protein 2/3 complex, another cytoskeletal protein The most significantly altered after IgM

signalling in CLL (p=0.00176) was Microtubule-associated protein 4, a cytoskeletal protein. The most significantly altered after IgD signalling in CLL (p=0.000186) was Phosphoglucomutase-1, a glycolytic pathway enzyme. Similarly, these and other cytoskeletal components were amongst the most significantly altered phosphoproteins in the J dataset. However, Benjamini-Hochberg correction with FDR at 0.05 and 0.1 did not result in any significant changes across all tonsil and CLL samples, for both P and J datasets. It seems that there are no phosphosites altered after BCR stimulation in all cases, using a stringent analysis method.

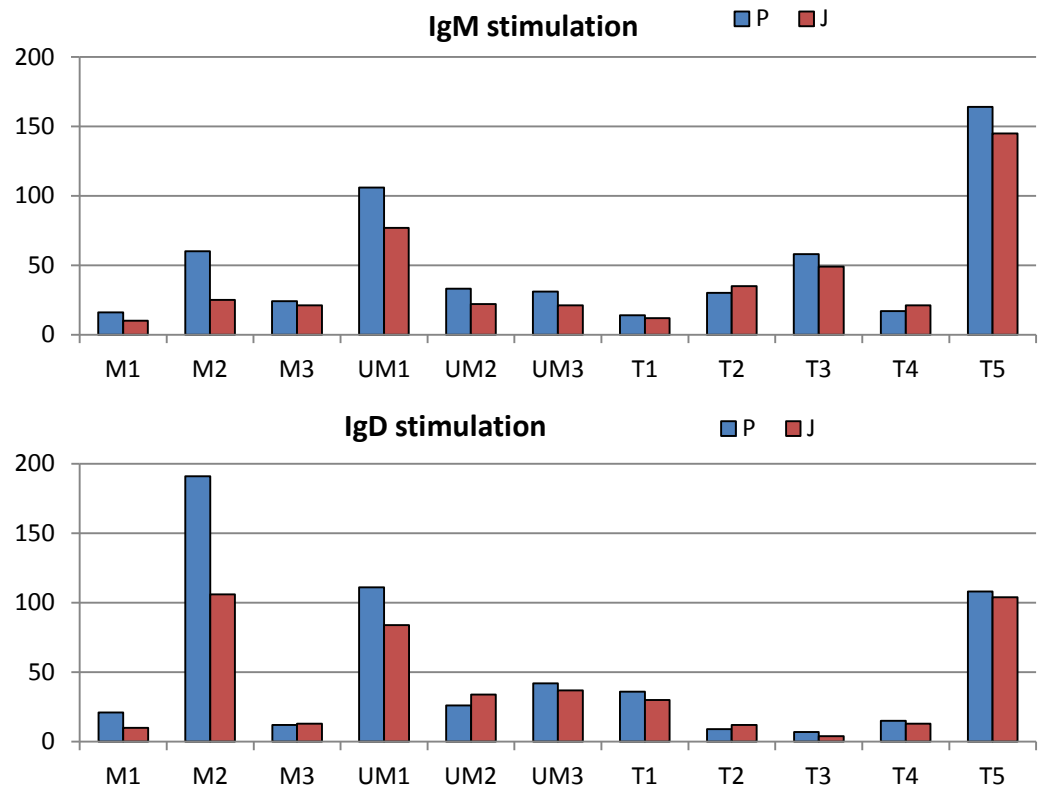
These results highlight the heterogeneity between patients. Some patients exhibited marked (up to $2^{16}$-fold) changes upon stimulation, but because they occurred in only 1 or 2 patients, they are not considered significant in this analysis. Possible explanations for these results are:

1. The phosphoproteomic techniques are either invalid or insensitive to the most important phosphosite changes with BCR stimulation, and this may be swamped by technical noise.
2. The analysis method is inappropriate
3. The heterogeneity between patients is greater than anticipated and any commonalities are 'swamped'
4. Any combination of the above

### *7.4.3 Using unpaired t-tests to compare each technical replicate*

We then made comparisons within each patient. Each stimulated condition triplicate was compared in an unpaired t-test to the baseline condition triplicate, with Benjamini-Hochberg correction at FDR of 0.05. This indicated whether a given phosphosite was altered significantly by stimulation in an individual patient. Different samples had differing numbers of phosphosites altered by IgM and IgD stimulation (see Fig 7.39). Overall 0.1-4.6% of phosphosites were altered by BCR stimulation. Figure 7.18 in section 7.3.4 provides examples of supervised clusters that illustrate examples of BCR signalling in patients.

**Figure 7.39: Number of phosphosites significantly altered by BCR stimulation in P and J datasets**




The overlaps between IgD and IgM within each sample were also compared (shown for P dataset in Table 7.13).

**Table 7.13: Overlapping phosphosites altered by both IgM and IgD stimulation, and those altered by each stimulus alone, P dataset**

| | M1 | M2 | M3 | UM1 | UM2 | UM3 | T1 | T2 | T3 | T4 | T5 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Overlap IgM and IgD** | 0 | 18 | 2 | 44 | 7 | 12 | 3 | 2 | 4 | 1 | 54 |
| **IgM only** | 16 | 42 | 22 | 62 | 26 | 19 | 11 | 28 | 54 | 16 | 110 |
| **IgD only** | 21 | 173 | 10 | 67 | 19 | 30 | 33 | 7 | 3 | 14 | 54 |

These data are combined to provide the proportions altered by both IgD and IgM or each stimulus alone (table 7.14), indicating that overall, 14-16% of the phosphosites are regulated by both IgM and IgD stimulation, with the remaining proportion unique to IgM or IgD stimulation. It also perhaps may be interpreted as IgD having a greater effect in CLL, compared to IgM, with the reverse in tonsil (see table 7.14). Of note, there was no correlation between the level of calcium flux induced by IgD or IgM and the number of altered phosphosites after the same stimulus.

**Table 7.14 percentages of all altered phosphosites in each group, overlaps. Overlapping phosphosites altered by both IgM and IgD stimulation, and those altered by each stimulus alone, P and J datasets combined.**

| | CLL P | CLL J | Tonsil P | Tonsil J |
|---|---|---|---|---|
| **Overlap IgM and IgD** | 16.2 | 14.5 | 16.5 | 15.3 |
| **IgM only** | 25.0 | 24.1 | 43.1 | 45.3 |
| **IgD only** | 42.5 | 46.8 | 23.9 | 24.0 |

Overlaps between patients were also considered, i.e. phosphosites that are consistently altered by stimulation in more than one patient (Table 7.15). Overlaps were limited, with a differentially regulated phosphosite after IgM stimulation in only three of the eleven patients (P dataset), and when tonsil and CLL are considered separately, there were only 5 phosphosites altered in at least 2 CLL patients, and 15 in tonsil:

**Table 7.15: Number of individual phosphosites significantly altered by IgM stimulation, overlaps P dataset.**
**Considering presence of the same phosphosite in more than one sample (overlap).**

| Number of samples with overlap | IgM ALL | IgM CLL | IgM T |
|---|---|---|---|
| **3** | 3 | 0 | 0 |
| **2** | 36 | 5 | 15 |
| **1** | 472 | 260 | 253 |
| **0** | 4577 | 4823 | 4820 |

The 3 differentially regulated phosphosites (and their putative functions) were:

- Pro-interleukin-16 pS844 (pleiotropic cytokine)
- Serine/threonine-protein kinase B-raf (MAP kinase/ERK signalling)
- ATP-dependent RNA helicase DHX8 pS460 (RNA processing)

Similarly, For the J dataset (Tables 7.16 & 7.17):

**Table 7.16: Number of individual phosphosites significantly altered by IgM stimulation, overlaps J dataset.**
**Considering presence of the same phosphosite in more than one sample (overlap).**

| Number of samples with overlap | IgM ALL | IgM CLL | IgM T |
|---|---|---|---|
| **3** | 7 | 0 | 2 |
| **2** | 31 | 6 | 14 |
| **1** | 355 | 164 | 228 |
| **0** | 2786 | 3009 | 2935 |

**Table 7.17: 7 IgM-induced phosphosites in at least 3 samples (J dataset)**

| Phosphoprotein overlap in 3 samples | Function |
|---|---|
| Protein phosphatase 1 regulatory subunit 12A | Regulator of protein phosphatase 1C. Mediates binding to myosin |
| Protein LAP2 | Adapter for the receptor ERBB2 |
| Serine/threonine-protein kinase B-raf | Transduction of mitogenic signals |
| ATP-dependent RNA helicase DHX8 | Facilitates nuclear export of spliced mRNA |
| Striatin-3 | May function as scaffolding or signalling protein |
| Cyclic AMP-responsive element-binding protein 1 | transcription factor |
| Translocation protein SEC62 | Required for preprotein translocation |

Similarly, for IgD, there was only one phosphosite (CD2 associated protein pS80) altered by IgD stimulation in 3 Tonsil patients (table 7.18).

**Table 7.18: Number of individual phosphosites significantly altered by IgD stimulation, overlaps P dataset.**
**Considering presence of the same phosphosite in more than one sample (overlap).**

| Number of samples with overlap | IgD ALL | IgD CLL | IgD T |
|---|---|---|---|
| **3** | 1 | 0 | 1 |
| **2** | 31 | 15 | 1 |
| **1** | 513 | 373 | 170 |
| **0** | 4543 | 4700 | 4916 |

And for J dataset, the same phosphosite (CD2 associated protein pS80) was altered in 3 Tonsil patients (Table 7.19). CD2 associated protein is a scaffolding molecule involved in regulation of the cytoskeleton. It has recently been associated with proximal BCR signalling events[554].

**Table 7.19: Number of individual phosphosites significantly altered by IgD stimulation, overlaps J dataset.**
**Considering presence of the same phosphosite in more than one sample (overlap).**

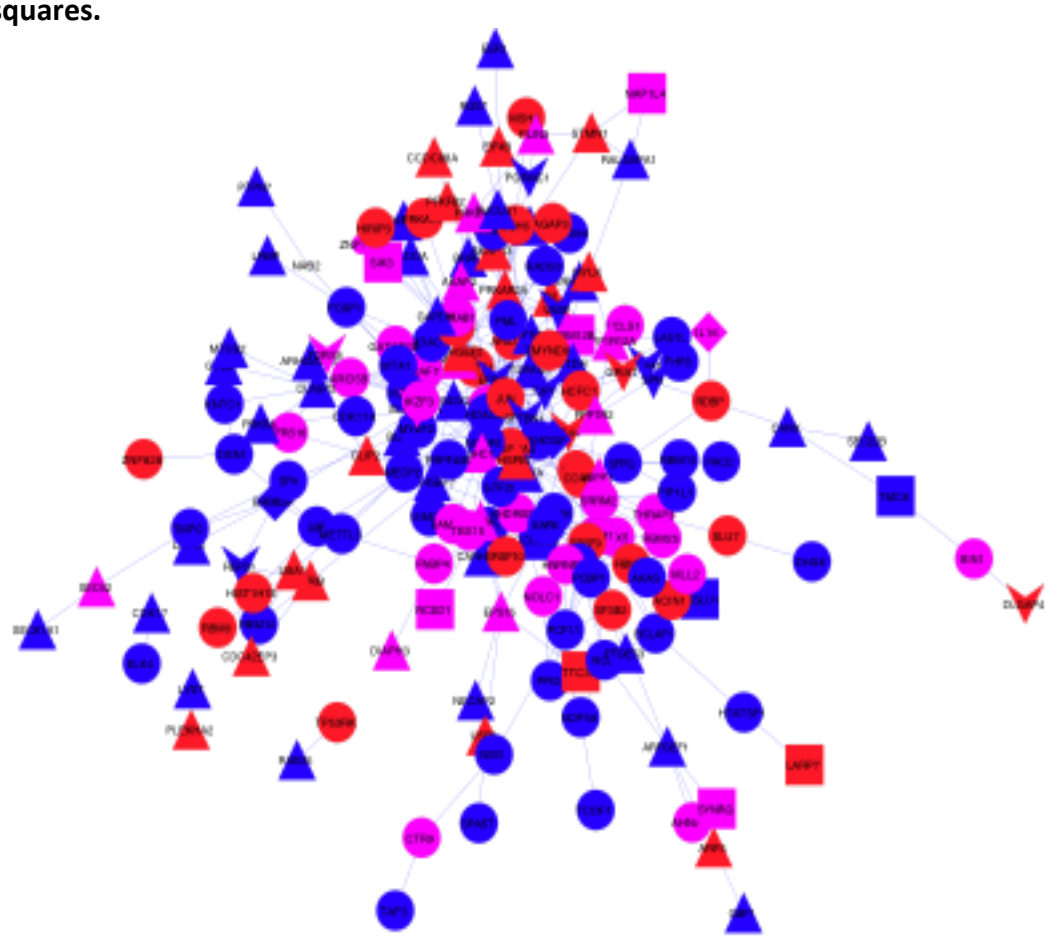
| Number of samples with overlap | IgD ALL | IgD CLL | IgD T |
|---|---|---|---|
| **3** | 1 | 0 | 1 |
| **2** | 37 | 14 | 2 |
| **1** | 370 | 256 | 156 |
| **0** | 2771 | 2909 | 3020 |

The percentage of phosphotyrosine phosphosites in the BCR altered lists was 1-1.5%, similar to the proportion for all detected phosphosites, indicating no enrichment of phosphotyrosine in BCR signalling.

This limited overlap is a reflection of the results of the analysis in section 7.3.13, demonstrating the absence of consistent phosphosite changes across patients. In an alternative strategy to search for generalities, the differentially regulated phosphosites after IgM or IgD stimulation in each CLL and Tonsil sample were combined into one list for further bioinformatic analysis.

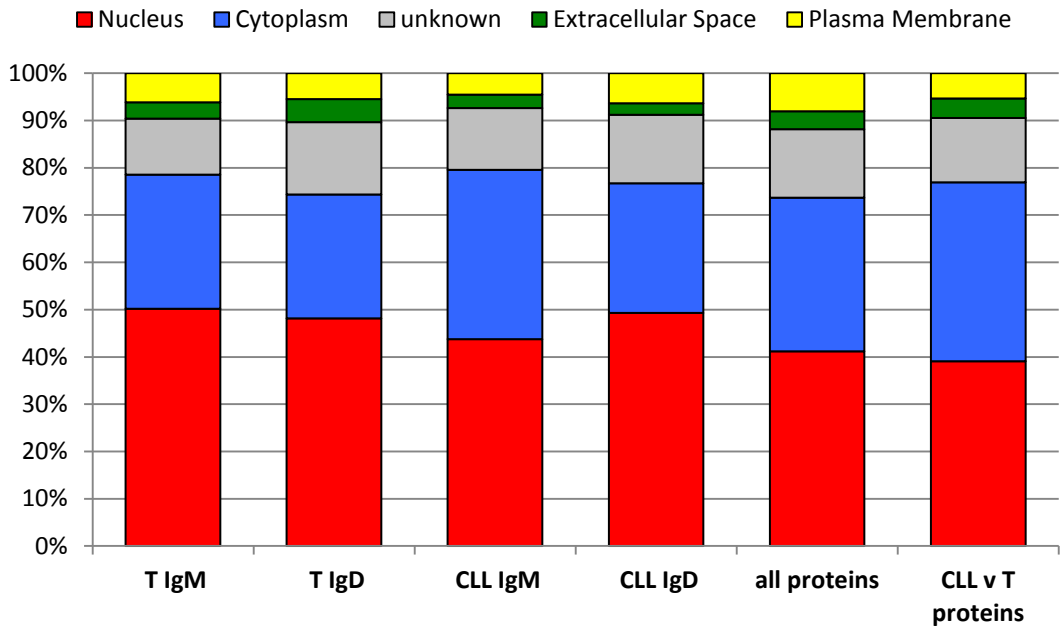
To estimate the degree to which correlation is also caused by direct protein-protein interactions, curated databases were considered. The Search Tool for the Retrieval of Interacting Genes (STRING) database provides access to access to both experimental as well as predicted protein-protein interaction information**.** By inputting protein identifications, known and predicted relationships between proteins can be visualised. The network data derived is then imported for visualisation in Cytoscape software (Figure 7.40)[555]. Evidence of interaction is derived from a combination of various published data. This includes proximity in genomes, co-expression in tissues, and evidence of direct binding from protein-interaction experiments and text mining techniques (such as co-occurrence in abstracts). As an example analysis, the phosphoproteins in the tonsil J dataset stimulated by BCR signalling were viewed in terms of their protein-protein interactions using the STRING database[556]. There was no obvious separation into IgD-stimulated, IgM-stimulated or common induced phosphoproteins by this analysis (see Figure 7.40), and this analysis technique was not pursued further.

**Figure 7.40: STRING interactions for Tonsil BCR signalling, J dataset.**
**The phosphoproteins altered by BCR stimulation with IgD (red nodes), IgM (blue nodes) or both (magenta nodes) were linked by known protein-binding interactions using STRING. Cytoplasmic proteins are represented by triangles, nuclear by circles, extracellular by diamonds, plasma membrane by arrows, unknown localisation by squares.**



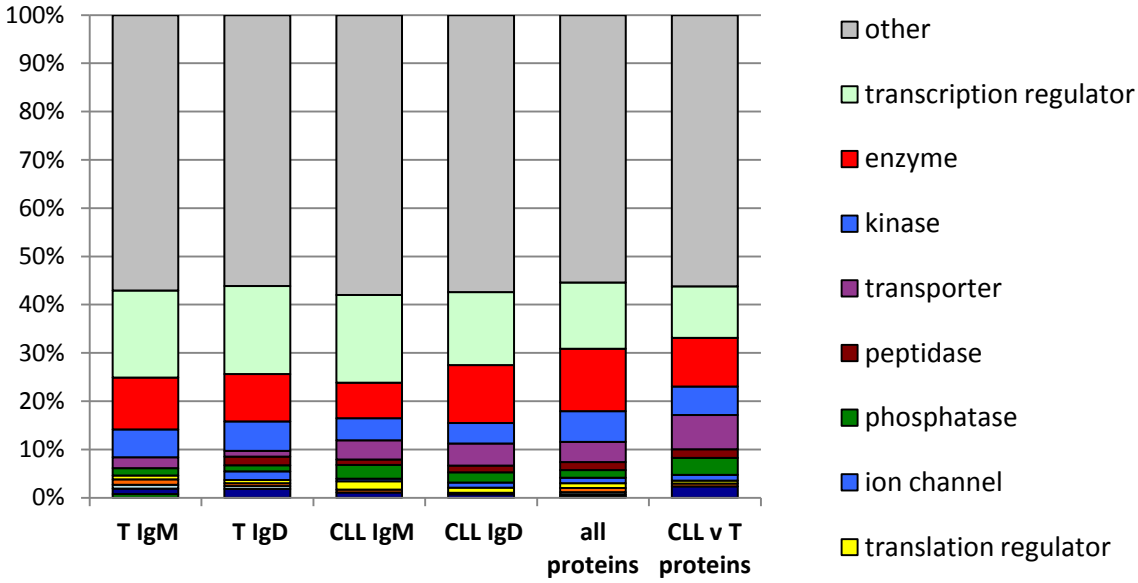
The cellular localisation of the proteins altered by BCR stimulation did not appear to be different from those in the entire dataset, or the CLL v Tonsil proteins (Fig 7.41):

**Figure 7.41. BCR-induced phosphoproteins by cellular localisation.**
**Shown are those altered by BCR stimulation for Tonsil (T) and CLL, J dataset and all proteins detected (all proteins) and those differing between CLL and Tonsil (CLL v T)**

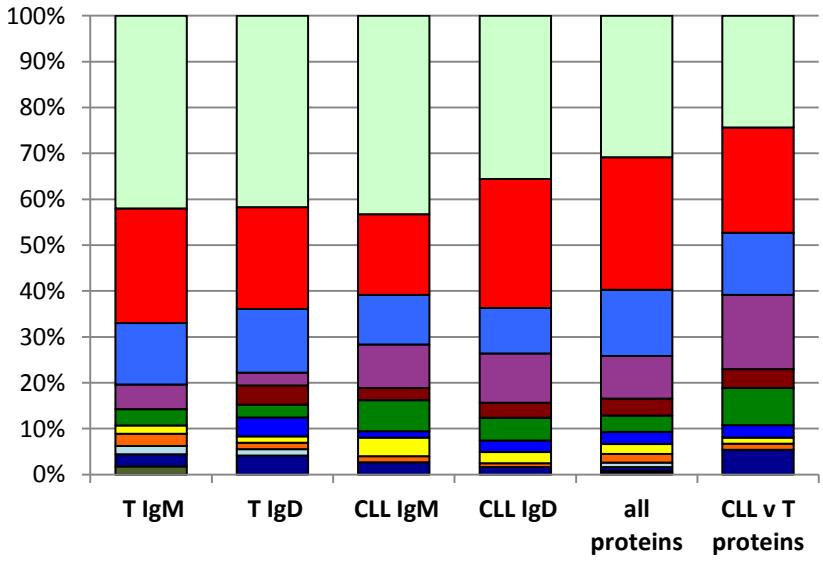



Similarly, the molecular function of the altered proteins was similar to those in the overall dataset, though there may be an increased proportion of transcriptional regulators after BCR stimulation (Fig 7.42)

**Figure 7.42. BCR-induced phosphoproteins by molecular function.**
**Shown are the proteins different after BCR stimulation for Tonsil (T) and CLL, J dataset, and the proteins in the entire dataset (all proteins) and those differing between CLL and Tonsil (CLL v T).**
**A: all functions**

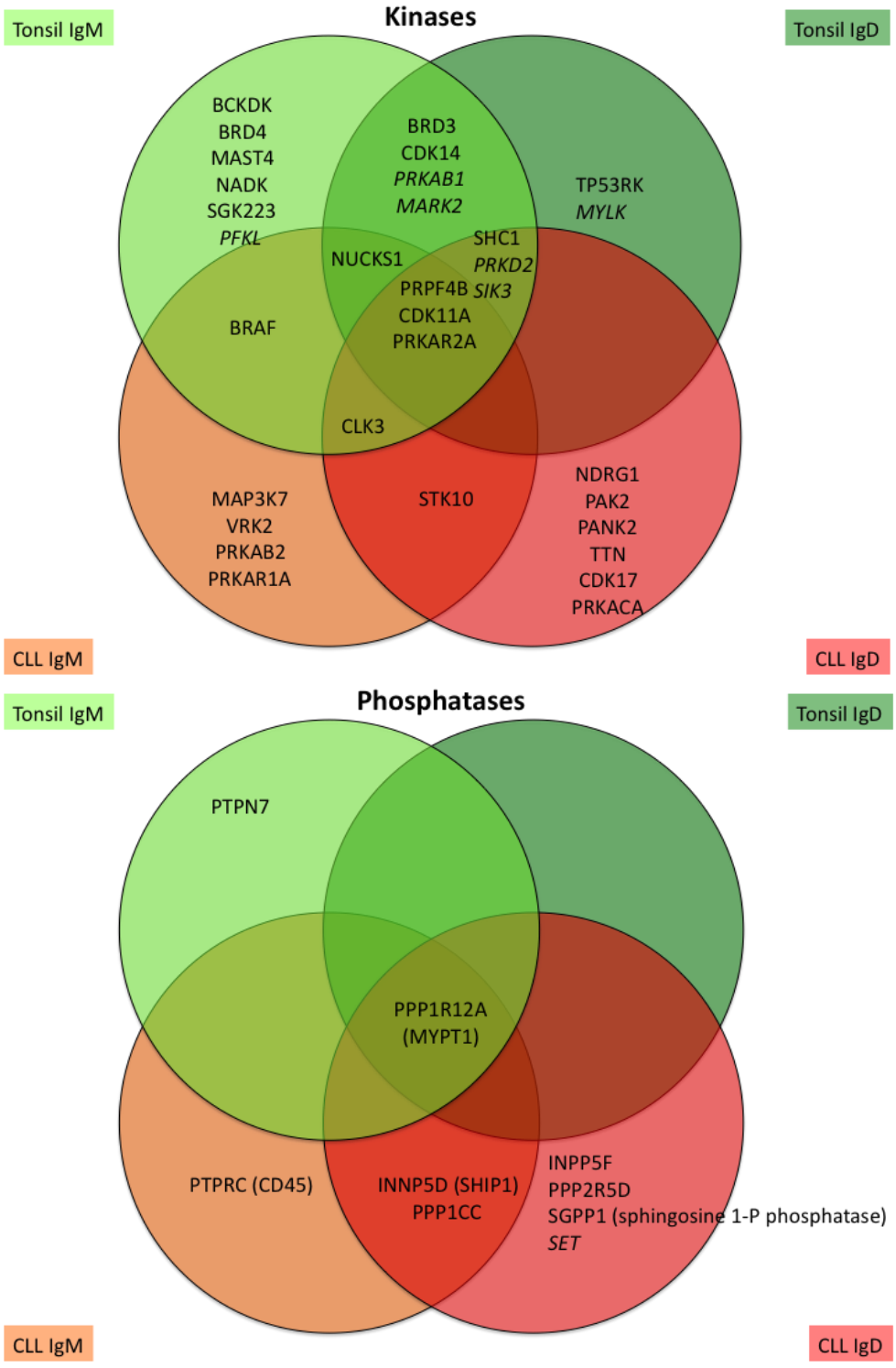


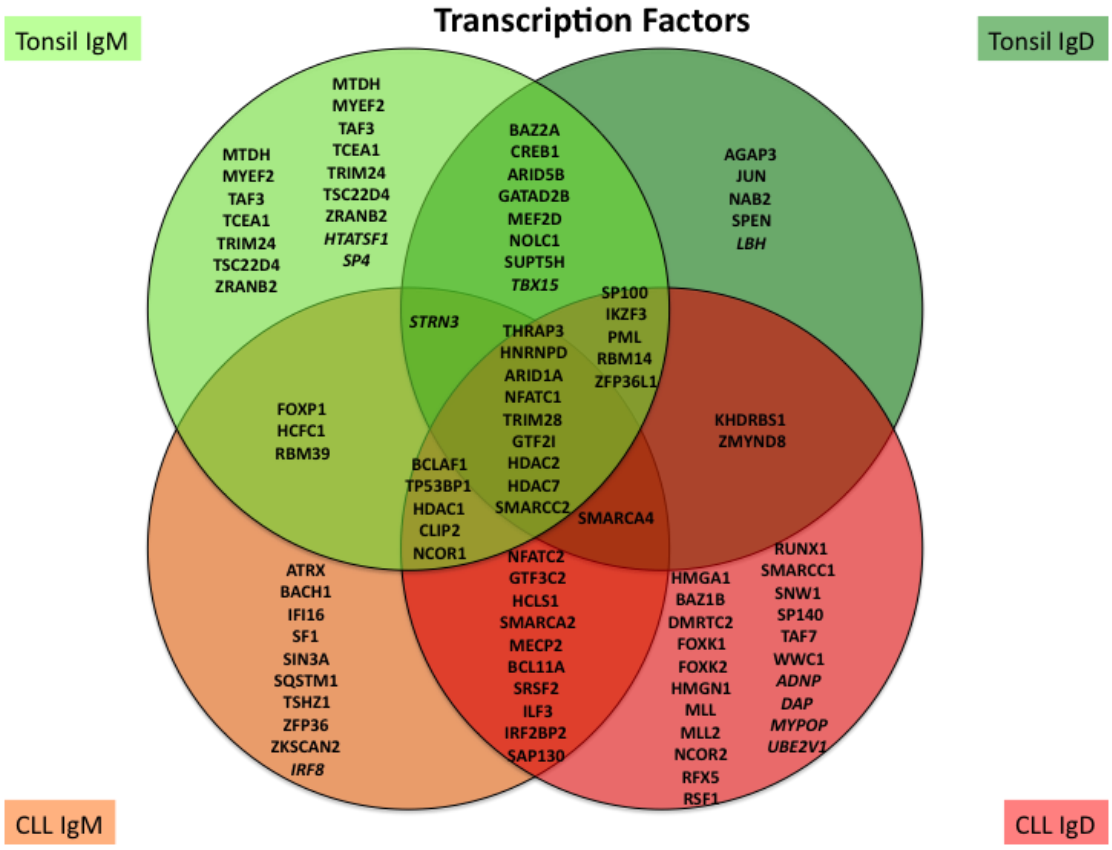

**B: As above, with the category 'other' removed**




Amongst the phosphosites with altered levels after BCR stimulation were several kinases, phosphatases and transcription factors. Figure 7.43 provides a summary of these according to stimulus.

**Figure 7.43 BCR-regulated phosphosites that are also kinases, phosphatases & transcription factors.**
**There was extensive overlap between the P and J datasets. Italicised kinases were present in J dataset only.**






The most enriched pathways after BCR stimulation are shown below (Table 7.20). The IPA analyses from CLL and Tonsil, and P and J datasets were compared. Also, each KEGG/IPA analysis was performed for each patient list to demonstrate the enrichment for pathway components within each patient, giving an impression of which patients are contributing to the enrichment. Note that if a component in one patient is combined with another pathway component in a different patient the combined list with 2 (or more) components may result in significant enrichment, even if enrichment is not considered as significant in an individual patient.

**Table 7.20: IPA and KEGG pathways enriched in proteins significantly altered by BCR stimulation.**
**The analysis was performed on combined lists of differentially regulated phosphoproteins after IgM or IgD stimulus. The same analysis was also performed on each separate patient lists. 'pts' indicates the number of patients that have a pathway identified when the analysis is run for each patient. Pathways enriched in ≥50% of patients in at least one dataset are shaded grey**

| Ingenuity Canonical Pathways IgM CLL | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| DNA Methylation & Transcriptional Repression Signalling | 0.13 | 0.0007 | 1 | 0.17 | 0.0001 | 3 |
| ERK/MAPK Signalling | 0.04 | 0.0011 | 4 | 0.04 | 0.0027 | 5 |
| Systemic Lupus Erythematosus Signalling | 0.03 | 0.0037 | 5 | 0.04 | 0.0018 | 5 |
| Phospholipase C Signalling | 0.02 | 0.0437 | 5 | 0.04 | 0.0025 | 5 |
| Hereditary Breast Cancer Signalling | 0.03 | 0.0741 | 3 | 0.05 | 0.0035 | 4 |
| Insulin Receptor Signalling | 0.02 | 0.0977 | 1 | 0.05 | 0.0063 | 5 |
| Role of BRCA1 in DNA Damage Response | 0.03 | 0.0977 | 3 | 0.06 | 0.0079 | 2 |
| Protein Kinase A Signalling | 0.02 | 0.0398 | 5 | 0.03 | 0.0081 | 6 |
| Tight Junction Signalling | 0.02 | 0.1306 | 4 | 0.04 | 0.0115 | 5 |
| Breast Cancer Regulation by Stathmin1 | 0.03 | 0.0245 | 3 | 0.04 | 0.0117 | 6 |
| iCOS-iCOSL Signalling in T Helper Cells | 0.04 | 0.0120 | 5 | 0.04 | 0.0437 | 4 |
| Acetyl-CoA Biosynthesis III (from Citrate) | X | X | 0 | 1.00 | 0.0132 | 1 |
| DNA DSB Repair by Homologous Recombination | 0.06 | 0.1159 | 1 | 0.13 | 0.0141 | 2 |
| Netrin Signalling | 0.04 | 0.0457 | 4 | 0.07 | 0.0145 | 4 |
| Telomere Extension by Telomerase | X | X | 0 | 0.13 | 0.0162 | 2 |
| Dopamine Receptor Signalling | 0.04 | 0.0282 | 1 | 0.05 | 0.0170 | 4 |
| Glucocorticoid Receptor Signalling | 0.02 | 0.0245 | 4 | 0.03 | 0.0195 | 4 |
| Estrogen Receptor Signalling | 0.03 | 0.0219 | 3 | 0.03 | 0.0759 | 3 |
| PI3K Signalling in B Lymphocytes | 0.03 | 0.0219 | 5 | 0.03 | 0.0759 | 4 |
| Calcium Signalling | 0.01 | 0.4315 | 4 | 0.03 | 0.0229 | 5 |
| Cardiac β-adrenergic Signalling | 0.03 | 0.0240 | 1 | 0.04 | 0.0245 | 4 |
| AMPK Signalling | 0.03 | 0.0251 | 4 | 0.04 | 0.0263 | 4 |
| IL-4 Signalling | 0.04 | 0.0251 | 4 | X | X | 3 |
| Spliceosomal Cycle | X | X | 0 | 0.14 | 0.0263 | 1 |
| Uridine-5'-phosphate Biosynthesis | X | X | 0 | 0.50 | 0.0263 | 1 |
| CDK5 Signalling | 0.03 | 0.0407 | 1 | 0.04 | 0.0275 | 4 |
| Nur77 Signalling in T Lymphocytes | 0.04 | 0.0741 | 3 | 0.05 | 0.0295 | 3 |
| ILK Signalling | 0.01 | 0.4699 | 1 | 0.03 | 0.0316 | 4 |
| Role of CHK Proteins in Cell Cycle Checkpoint Control | 0.02 | 0.3784 | 1 | 0.05 | 0.0339 | 2 |
| Telomerase Signalling | 0.02 | 0.2056 | 1 | 0.04 | 0.0380 | 3 |
| **B Cell Receptor Signalling** | **0.03** | **0.0417** | **5** | **0.03** | **0.0468** | 4 |
| ATM Signalling | 0.02 | 0.4055 | 1 | 0.05 | 0.0427 | 3 |
| Cell Cycle: G1/S Checkpoint Regulation | 0.02 | 0.4055 | 1 | 0.05 | 0.0427 | 3 |
| April Mediated Signalling | 0.05 | 0.0437 | 4 | 0.05 | 0.0891 | 3 |
| B Cell Activating Factor Signalling | 0.05 | 0.0479 | 4 | 0.05 | 0.0977 | 3 |

| KEGG pathway IgM CLL | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Spliceosome | 0.06 | 0.0024 | 2 | 0.03 | 0.0310 | 2 |
| Nucleotide excision repair | X | X | 0 | 0.09 | 0.0110 | 1 |
| **B cell receptor signalling pathway** | **0.04** | **0.0634** | **0** | **X** | **X** | **0** |

| Ingenuity Canonical Pathways IgM tonsil | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Phospholipase C Signalling | 0.04 | 0.0005 | 4 | 0.05 | 0.0000 | 5 |
| DNA Methylation & Transcriptional Repression Signalling | 0.13 | 0.0019 | 1 | 0.17 | 0.0001 | 4 |
| Telomerase Signalling | 0.04 | 0.0339 | 3 | 0.06 | 0.0008 | 4 |
| Sucrose Degradation V (Mammalian) | 0.14 | 0.0741 | 2 | 0.29 | 0.0019 | 3 |
| Tight Junction Signalling | 0.04 | 0.0023 | 3 | 0.03 | 0.0794 | 2 |
| Estrogen Receptor Signalling | 0.05 | 0.0045 | 2 | 0.05 | 0.0026 | 3 |
| Glycolysis I | 0.13 | 0.0026 | 5 | 0.09 | 0.0257 | 3 |
| Actin Nucleation by ARP-WASP Complex | 0.06 | 0.0048 | 3 | 0.06 | 0.0032 | 4 |
| Signalling by Rho Family GTPases | 0.02 | 0.0575 | 5 | 0.03 | 0.0032 | 5 |
| Role of NFAT in Cardiac Hypertrophy | 0.03 | 0.0708 | 1 | 0.04 | 0.0037 | 4 |
| ATM Signalling | 0.03 | 0.1722 | 2 | 0.07 | 0.0045 | 3 |
| Breast Cancer Regulation by Stathmin1 | 0.03 | 0.0324 | 3 | 0.04 | 0.0055 | 3 |
| Integrin Signalling | 0.03 | 0.0100 | 3 | 0.03 | 0.0055 | 4 |
| Protein Kinase A Signalling | 0.03 | 0.0170 | 2 | 0.03 | 0.0081 | 3 |
| Hereditary Breast Cancer Signalling | 0.03 | 0.1671 | 1 | 0.04 | 0.0087 | 4 |
| RhoGDI Signalling | 0.03 | 0.0537 | 4 | 0.03 | 0.0098 | 4 |
| Calcium Signalling | 0.02 | 0.1629 | 2 | 0.03 | 0.0120 | 4 |
| Diphthamide Biosynthesis | 0.50 | 0.0126 | 1 | X | X | 0 |
| Granzyme A Signalling | 0.12 | 0.0170 | 0 | 0.06 | 0.1671 | 1 |
| Thrombin Signalling | 0.03 | 0.0851 | 4 | 0.03 | 0.0186 | 4 |
| Nur77 Signalling in T Lymphocytes | 0.02 | 0.4797 | 1 | 0.05 | 0.0200 | 3 |
| Spliceosomal Cycle | X | X | 0 | 0.14 | 0.0224 | 1 |
| NADH Repair | 0.33 | 0.0251 | 1 | X | X | 0 |
| Calcium-induced T Lymphocyte Apoptosis | 0.02 | 0.5236 | 1 | 0.05 | 0.0282 | 3 |
| Cell Cycle: G1/S Checkpoint Regulation | 0.03 | 0.1722 | 1 | 0.05 | 0.0295 | 4 |
| Huntington's Disease Signalling | 0.02 | 0.1219 | 2 | 0.03 | 0.0302 | 4 |
| cAMP-mediated signalling | 0.02 | 0.1236 | 2 | 0.03 | 0.0309 | 2 |
| Gluconeogenesis I | 0.08 | 0.0316 | 3 | 0.04 | 0.2223 | 2 |
| PPARα/RXRα Activation | 0.03 | 0.0501 | 2 | 0.03 | 0.0331 | 2 |
| ERK5 Signalling | 0.03 | 0.1905 | 2 | 0.05 | 0.0347 | 2 |
| Corticotropin Releasing Hormone Signalling | 0.03 | 0.1611 | 2 | 0.03 | 0.0355 | 2 |
| Melatonin Signalling | 0.04 | 0.0537 | 1 | 0.04 | 0.0407 | 2 |
| Synaptic Long Term Potentiation | 0.03 | 0.0575 | 2 | 0.03 | 0.0407 | 3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **B Cell Receptor Signalling** | **0.02** | **0.2938** | 2 | **0.01** | **0.5070** | 2 |

| KEGG pathway IgM tonsil | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Spliceosome | 0.06 | 0.0013 | 1 | 0.08 | 0.0000 | 1 |
| Antigen processing and presentation | X | X | 0 | 0.05 | 0.0499 | 1 |
| Pathogenic Escherichia coli infection | 0.09 | 0.0022 | 1 | X | X | 0 |

| Ingenuity Canonical Pathways IgD CLL | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Telomerase Signalling | 0.05 | 0.0126 | 4 | 0.08 | 0.0003 | 4 |
| DNA Methylation & Transcriptional Repression Signalling | 0.09 | 0.0331 | 2 | 0.17 | 0.0004 | 4 |
| Estrogen Receptor Signalling | 0.05 | 0.0081 | 5 | 0.06 | 0.0015 | 5 |
| IL-4 Signalling | 0.05 | 0.0200 | 4 | 0.08 | 0.0018 | 3 |
| Insulin Receptor Signalling | 0.05 | 0.0098 | 2 | 0.06 | 0.0019 | 2 |
| RAR Activation | 0.03 | 0.0355 | 4 | 0.05 | 0.0112 | 3 |
| Diphthamide Biosynthesis | 0.50 | 0.0145 | 1 | 0.50 | 0.0178 | 1 |
| Hereditary Breast Cancer Signalling | 0.03 | 0.0759 | 3 | 0.05 | 0.0148 | 2 |
| ERK/MAPK Signalling | 0.04 | 0.0158 | 2 | 0.04 | 0.0162 | 3 |
| Glucocorticoid Receptor Signalling | 0.03 | 0.0316 | 4 | 0.04 | 0.0166 | 3 |
| Phospholipase C Signalling | 0.03 | 0.0398 | 4 | 0.04 | 0.0178 | 6 |
| CDK5 Signalling | 0.04 | 0.0372 | 2 | 0.06 | 0.0200 | 2 |
| Assembly of RNA Polymerase III Complex | 0.07 | 0.1718 | 1 | 0.14 | 0.0219 | 2 |
| Dopamine Receptor Signalling | 0.05 | 0.0229 | 2 | 0.05 | 0.0457 | 2 |
| AMPK Signalling | 0.04 | 0.0363 | 4 | 0.04 | 0.0263 | 3 |
| Dopamine-DARPP32 Feedback in cAMP Signalling | 0.04 | 0.0263 | 2 | 0.03 | 0.1510 | 2 |
| Calcium Signalling | 0.02 | 0.2208 | 4 | 0.04 | 0.0309 | 3 |
| Netrin Signalling | 0.04 | 0.1081 | 2 | 0.07 | 0.0324 | 2 |
| Granzyme A Signalling | 0.06 | 0.2070 | 1 | 0.18 | 0.0324 | 3 |
| Neuroprotective Role of THOP1 in Alzheimer's | 0.05 | 0.1130 | 2 | 0.07 | 0.0347 | 2 |
| Spliceosomal Cycle | X | X | 0 | 0.14 | 0.0355 | 2 |
| Uridine-5'-phosphate Biosynthesis | X | X | 0 | 0.50 | 0.0355 | 1 |
| FcγRIIB Signalling in B Lymphocytes | 0.04 | 0.1175 | 1 | 0.06 | 0.0363 | 2 |
| Role of NFAT in Cardiac Hypertrophy | 0.03 | 0.1069 | 3 | 0.04 | 0.0372 | 3 |
| Natural Killer Cell Signalling | 0.03 | 0.1928 | 2 | 0.05 | 0.0398 | 2 |
| Ins(1,4,5)$P_3$ Degradation | 0.06 | 0.2301 | 1 | 0.11 | 0.0407 | 1 |
| Renin-Angiotensin Signalling | 0.03 | 0.2037 | 3 | 0.04 | 0.0447 | 3 |
| 1D-myo-inositol Hexakisphosphate Biosynthesis | 0.05 | 0.2410 | 1 | 0.11 | 0.0447 | 1 |
| Ins(1,4,5)$P_3$ Biosynthesis | 0.05 | 0.2410 | 1 | 0.11 | 0.0447 | 1 |
| Cyclins and Cell Cycle Regulation | 0.03 | 0.1000 | 3 | 0.05 | 0.0490 | 2 |
| PDGF Signalling | 0.03 | 0.3048 | 1 | 0.05 | 0.0490 | 2 |
| Integrin Signalling | 0.03 | 0.0537 | 3 | 0.03 | 0.0513 | 4 |
| **B Cell Receptor Signalling** | **0.02** | **0.3631** | 2 | **0.04** | **0.0513** | 2 |

| KEGG pathway IgD CLL | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Spliceosome | 0.06 | 0.0024 | 2 | 0.10 | 0.0000 | 3 |
| Notch signalling pathway | 0.09 | 0.0146 | 2 | 0.11 | 0.0040 | 2 |
| RNA degradation | 0.09 | 0.0243 | 0 | X | X | 0 |
| Insulin signalling pathway | 0.04 | 0.0604 | 0 | 0.04 | 0.0394 | 0 |

| Ingenuity Canonical Pathways IgD tonsil | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Estrogen Receptor Signalling | 0.05 | 0.0006 | 2 | 0.05 | 0.0007 | 2 |
| Signalling by Rho Family GTPases | 0.03 | 0.0028 | 2 | 0.03 | 0.0029 | 3 |
| Phospholipase C Signalling | 0.02 | 0.0427 | 3 | 0.03 | 0.0032 | 4 |
| Cdc42 Signalling | 0.04 | 0.0047 | 2 | 0.02 | 0.0977 | 3 |
| Glucocorticoid Receptor Signalling | 0.03 | 0.0068 | 3 | 0.02 | 0.0245 | 3 |
| Granzyme A Signalling | 0.12 | 0.0083 | 0 | 0.06 | 0.1315 | 1 |
| SAPK/JNK Signalling | 0.03 | 0.0468 | 1 | 0.04 | 0.0089 | 3 |
| Protein Kinase A Signalling | 0.02 | 0.0135 | 3 | 0.02 | 0.0141 | 3 |
| Hypoxia Signalling in the Cardiovascular System | 0.03 | 0.0977 | 0 | 0.05 | 0.0162 | 1 |
| Spliceosomal Cycle | X | X | 0 | 0.14 | 0.0174 | 1 |
| Neurotrophin/TRK Signalling | 0.04 | 0.0204 | 0 | 0.04 | 0.0209 | 2 |
| ERK/MAPK Signalling | 0.03 | 0.0209 | 3 | 0.02 | 0.0741 | 3 |
| GDNF Family Ligand-Receptor Interactions | 0.04 | 0.0214 | 0 | 0.04 | 0.0219 | 2 |
| Cardiac β-adrenergic Signalling | 0.03 | 0.0234 | 3 | 0.03 | 0.0240 | 3 |
| Integrin Signalling | 0.02 | 0.0234 | 3 | 0.01 | 0.2254 | 2 |
| AMPK Signalling | 0.03 | 0.0245 | 3 | X | X | 1 |
| Regulation of Actin-based Motility by Rho | 0.03 | 0.0282 | 2 | 0.01 | 0.4887 | 2 |
| **B Cell Receptor Signalling** | **0.03** | **0.0407** | **1** | **0.03** | **0.0417** | 3 |
| April Mediated Signalling | 0.05 | 0.0427 | 1 | 0.05 | 0.0437 | 2 |
| PPAR Signalling | 0.02 | 0.1841 | 0 | 0.03 | 0.0447 | 2 |
| B Cell Activating Factor Signalling | 0.05 | 0.0468 | 1 | 0.05 | 0.0479 | 2 |
| PPARα/RXRα Activation | 0.02 | 0.0479 | 1 | 0.02 | 0.1589 | 2 |
| EIF2 Signalling | X | X | 0 | 0.02 | 0.0490 | 4 |

| KEGG pathway IgD Tonsil | Ratio J | P-value J | Pts J | Ratio P | P-value P | Pts P |
|---|---|---|---|---|---|---|
| Spliceosome | 0.07 | 0.000 | 1 | 0.10 | 0.000 | 2 |

Highlighted in Table 7.20 are the data indicating enrichment for components of the BCR signalling pathway. There are suggestions that our datasets do include these BCR pathway proteins. For IPA analysis, IgM CLL, IgD CLL (p=0.0513) and IgD Tonsil all seem to include BCR signalling components. The KEGG pathways analysis for IgM CLL has the BCR pathway enriched at p=0.06. More strikingly, the 'Spliceosome' KEGG pathway is significantly enriched in all IgM/IgD CLL/tonsil datasets, and the analogous 'Spliceosomal Cycle' appears in all IPA pathway lists. Also, Notch signalling is also enriched in IgD CLL KEGG analysis.

When the analyses are run for each list of phosphoproteins altered by BCR in each patient, it can be seen that the majority of pathways highlighted by combination of the

individual lists are also enriched in each patient. For example, protein kinase A signalling, highlighted in the IgM CLL analysis, is also significantly enriched when each of the 6 lists of IgM-stimulated phosphoproteins in the P dataset is considered. This suggests that whilst there is great patient-to-patient heterogeneity in the phosphoproteins stimulated by a particular stimulus, there is more unity when the pathways themselves are considered.

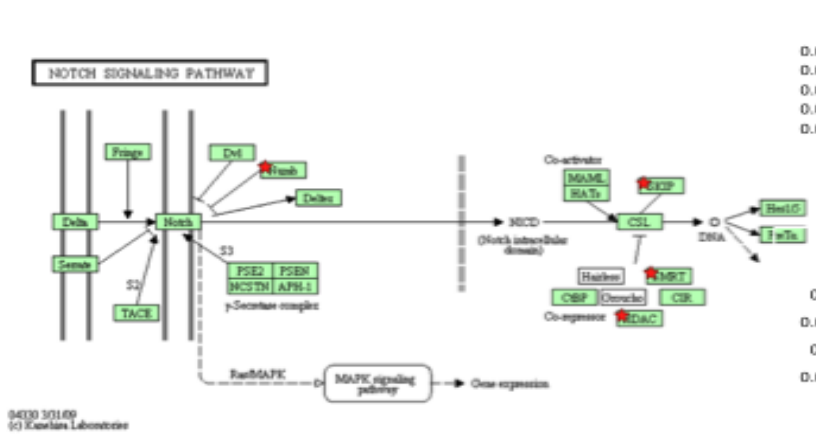
### *7.4.4 Notch & Spliceosome signalling is stimulated by BCR ligation*

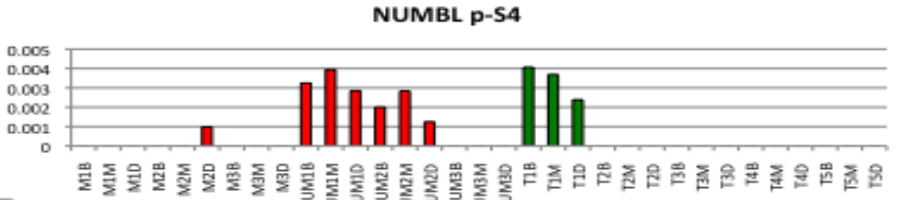
The IgD-stimulated CLL phosphosites were enriched for Notch signalling components: pS4 NUMBL, pS2269 NCOR2 (SMRT), pS224/232 SNW1 (SKIP) (phosphosite downstream of CDK2), pS410 HDAC1 (phosphosite downstream of CK2-A1) and pS421/423 HDAC2. The NOTCH pathway and the expression levels of the relevant phosphosites are shown in figure 7.44.

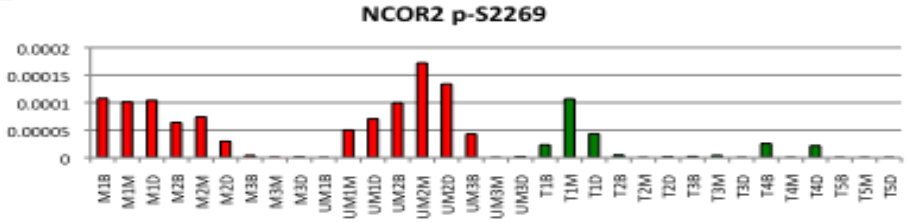
Altogether, 24 proteins in the KEGG spliceosome pathway were altered by at least one of the IgD or IgM stimuli. The KEGG pathway, and each of the altered phosphoproteins are shown in Figure 7.45. The alterations of expression by stimuli are complex, with the changes depending on sample, stimulus and phosphosite. Two examples (RBM25 and DDX42 are shown, but it is difficult to discern a pattern amongst all the altered phosphosites within the spliceosome components.

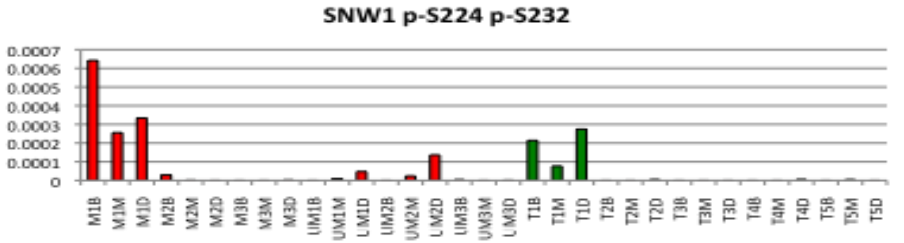
**Figure 7.44: NOTCH KEGG pathway**

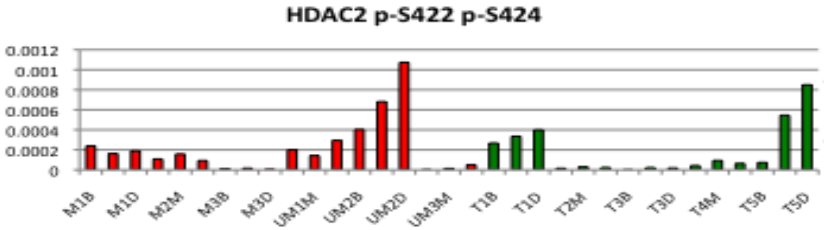
**Phosphoproteins in IgD CLL lists highlighted in red, and mean expression values for each condition (J dataset). Green=Tonsil, Red = CLL**



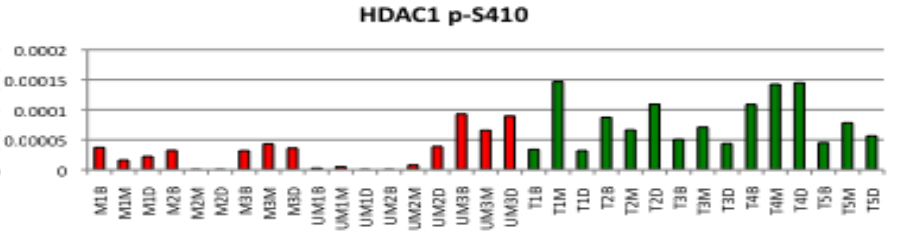
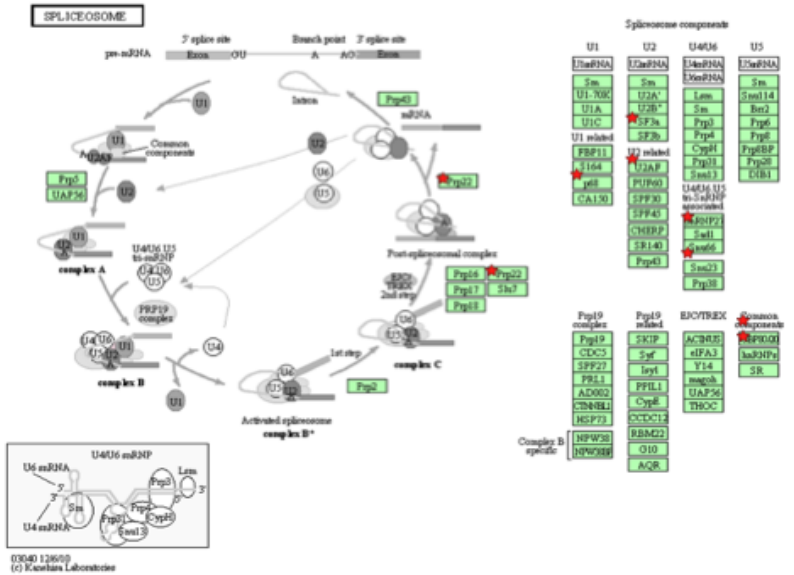
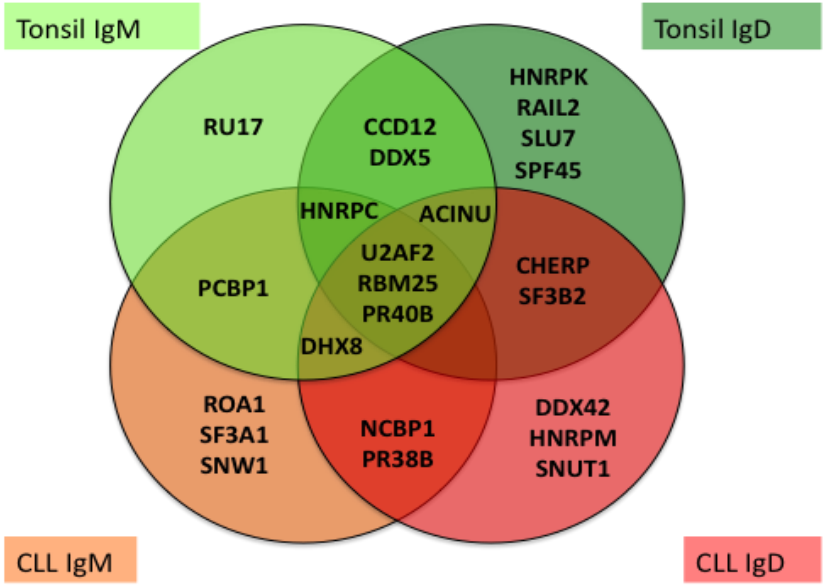
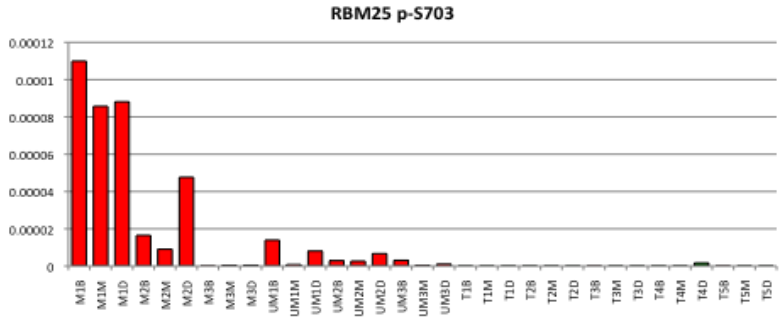
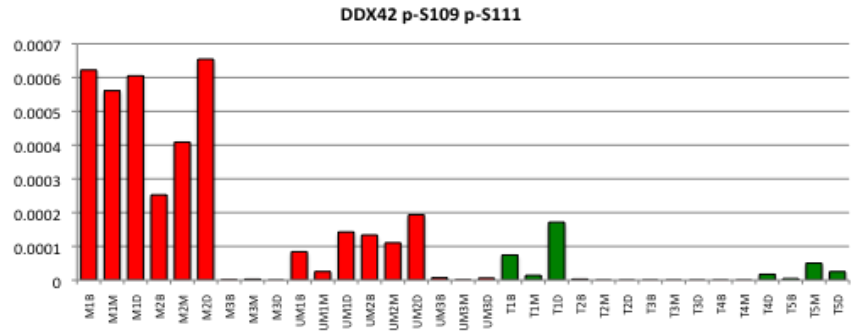



**Figure 7.45: Spliceosome KEGG pathway.**
**CLL IgM-stimulated phosphoproteins highlighted in red. To the right are the overlaps between the different stimuli and the spliceosome components altered by those stimuli, and below are two examples of expression levels in the different samples (J dataset)**



### *7.4.5 BCR pathways, the differences between IgM and IgD, CLL and tonsil.*

We sought to further clarify the overlaps and differences between IgD and IgM signalling, and between CLL and Tonsil. The lists of proteins altered by BCR signalling were compared. The characteristics of phosphoproteins altered by both IgD and IgM signalling, and those altered by only IgD or IgM signalling were compared. As has been shown (Table 7.21), there is a degree of overlap between IgM and IgD, but the majority of altered phosphosites are specific to each stimulus:

**Table 7.21 percentages of all altered phosphosites in each group, M & D overlap. Overlapping phosphosites altered by both IgM and stimulation, and those altered by each stimulus alone, P and J datasets shown separately.**

| | CLL P | CLL J | Tonsil P | Tonsil J |
|---|---|---|---|---|
| **Overlap IgM and IgD** | 16.2 | 14.5 | 16.5 | 15.3 |
| **IgM only** | 25.0 | 24.1 | 43.1 | 45.3 |
| **IgD only** | 42.5 | 46.8 | 23.9 | 24.0 |

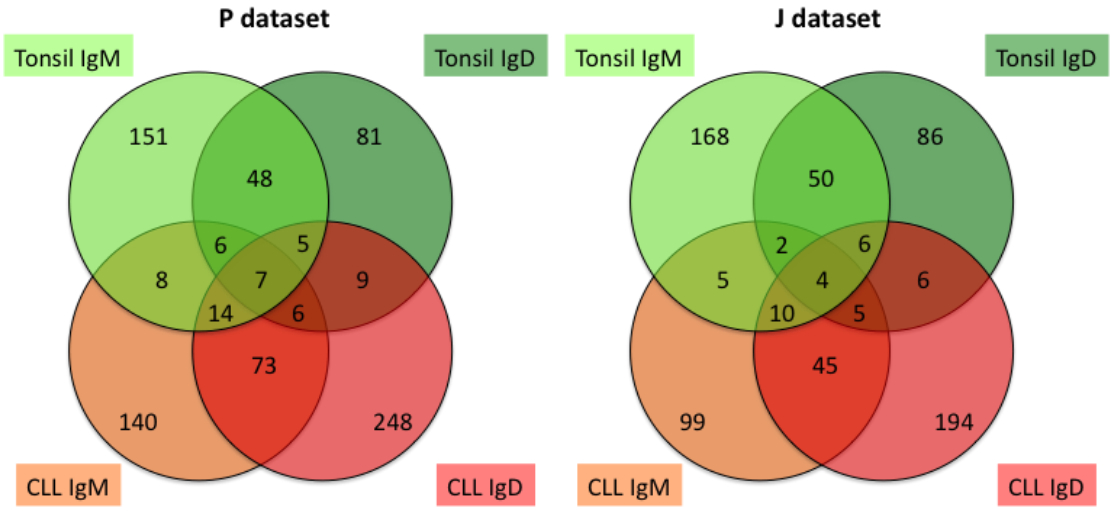
Similarly, there is limited overlap between CLL and Tonsil in terms of the phosphosites characterising IgM or IgD signalling (Table 7.22)

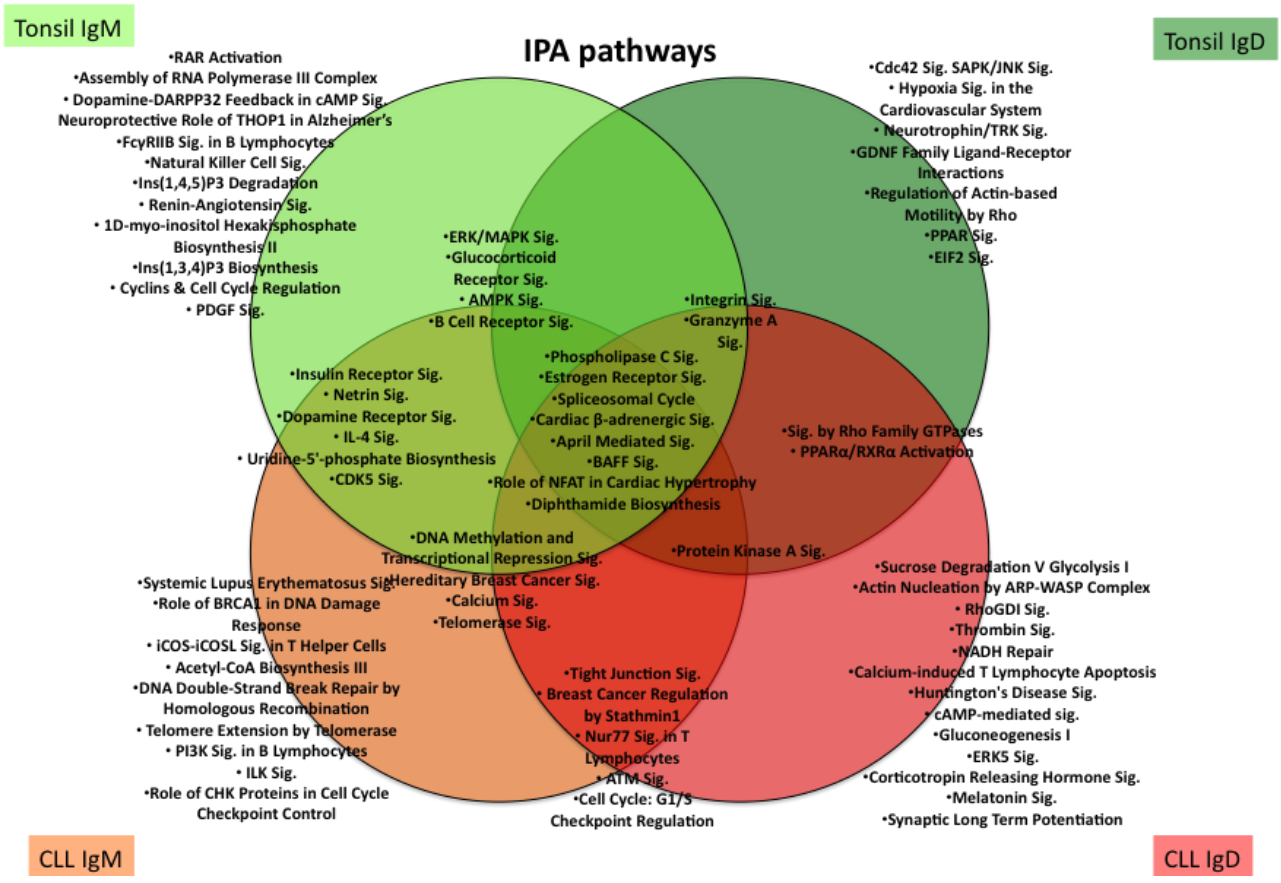
**Table 7.22 percentages of all altered phosphosites in each group, CLL & T overlap. Overlapping phosphosites altered by stimulation in both CLL and Tonsil, and those altered in each patient group alone, P and J datasets shown separately.**

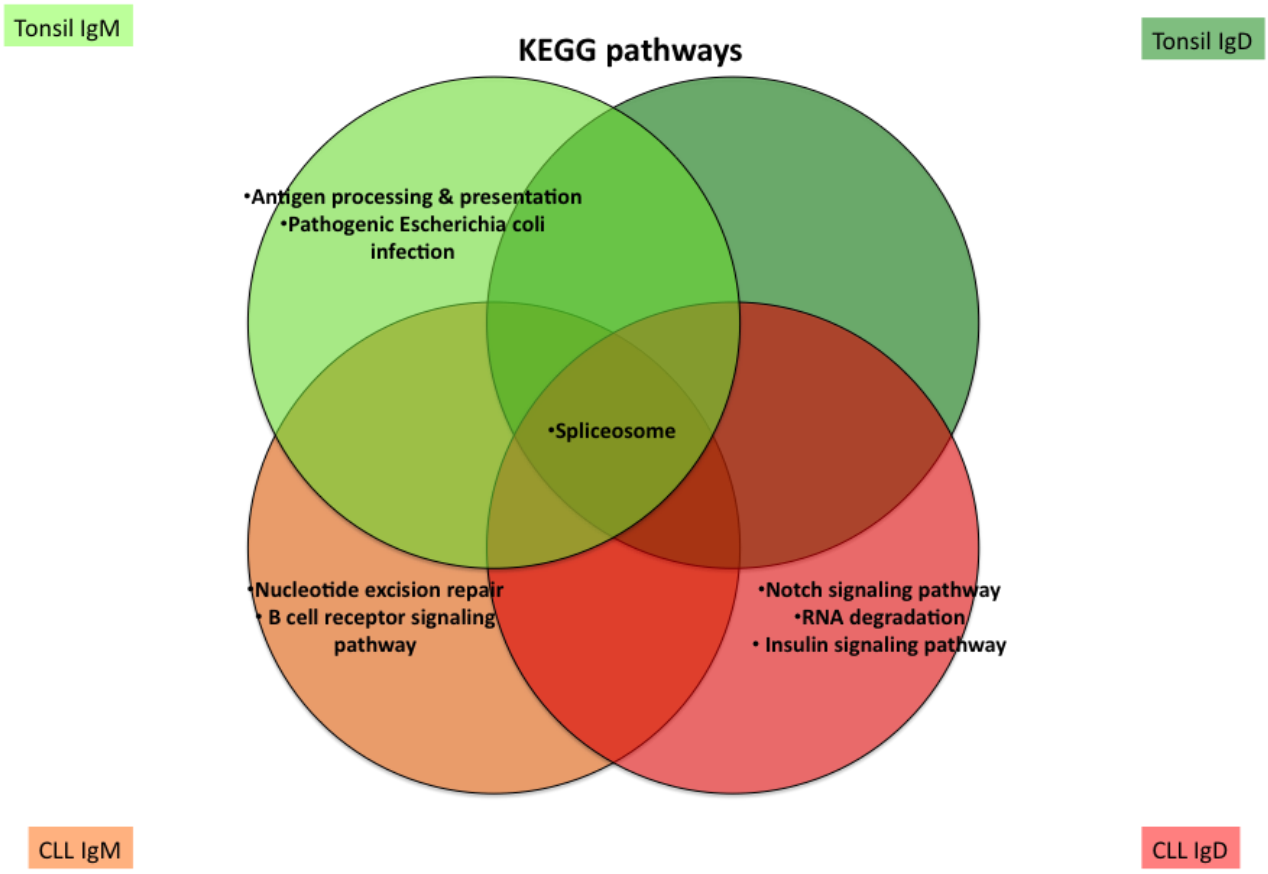
| | IgM P | IgM J | IgD P | IgD J |
|---|---|---|---|---|
| **Overlap CLL and T** | 7.1 | 5.1 | 5.2 | 4.9 |
| **CLL only** | 44.4 | 35.9 | 63.9 | 58.0 |
| **T only** | 41.4 | 54.0 | 25.8 | 32.2 |

Figure 7.46 illustrates the overlaps of unique phosphosites altered by stimulation in the CLL and Tonsil samples. For each list (IgM CLL, IgM Tonsil, IgD CLL, IgD Tonsil, see Table 7.20), the IPA and KEGG pathways were generated and overlaps considered in terms of the biological pathways that were enriched in either the P or J dataset. (i.e Phospholipase C Signalling appeared in all 4 lists, and so is placed at the overlap of the conditions. The pathway is not derived from analysis of the few phosphosites that overlapped in all conditions, though this analysis does suggest similar but fewer pathways, data not shown).

**Figure 7.46. The number of phosphosites altered by BCR stimulation considered in their overlaps between IgM & IgM, CLL and Tonsil.**
**Each of the 4 lists of IgM/IgD Below are the pathways containing the BCR stimulated phosphoproteins**



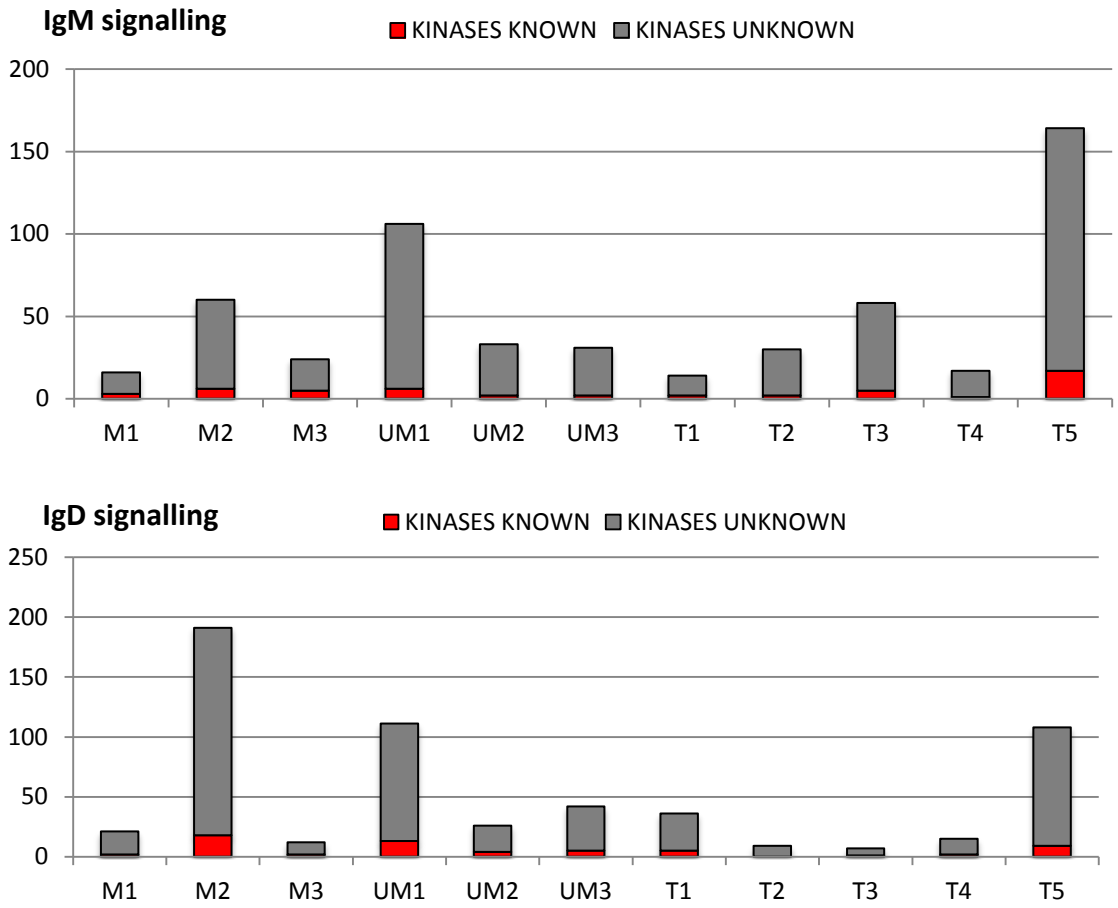






### *7.4.6 Kinases active after BCR ligation*

The BCR-stimulation regulated phosphosites were matched to upstream kinases, with a minority being represented. In the P dataset, 9% of phosphosites regulated after IgM had identified kinases (figure 7.47):

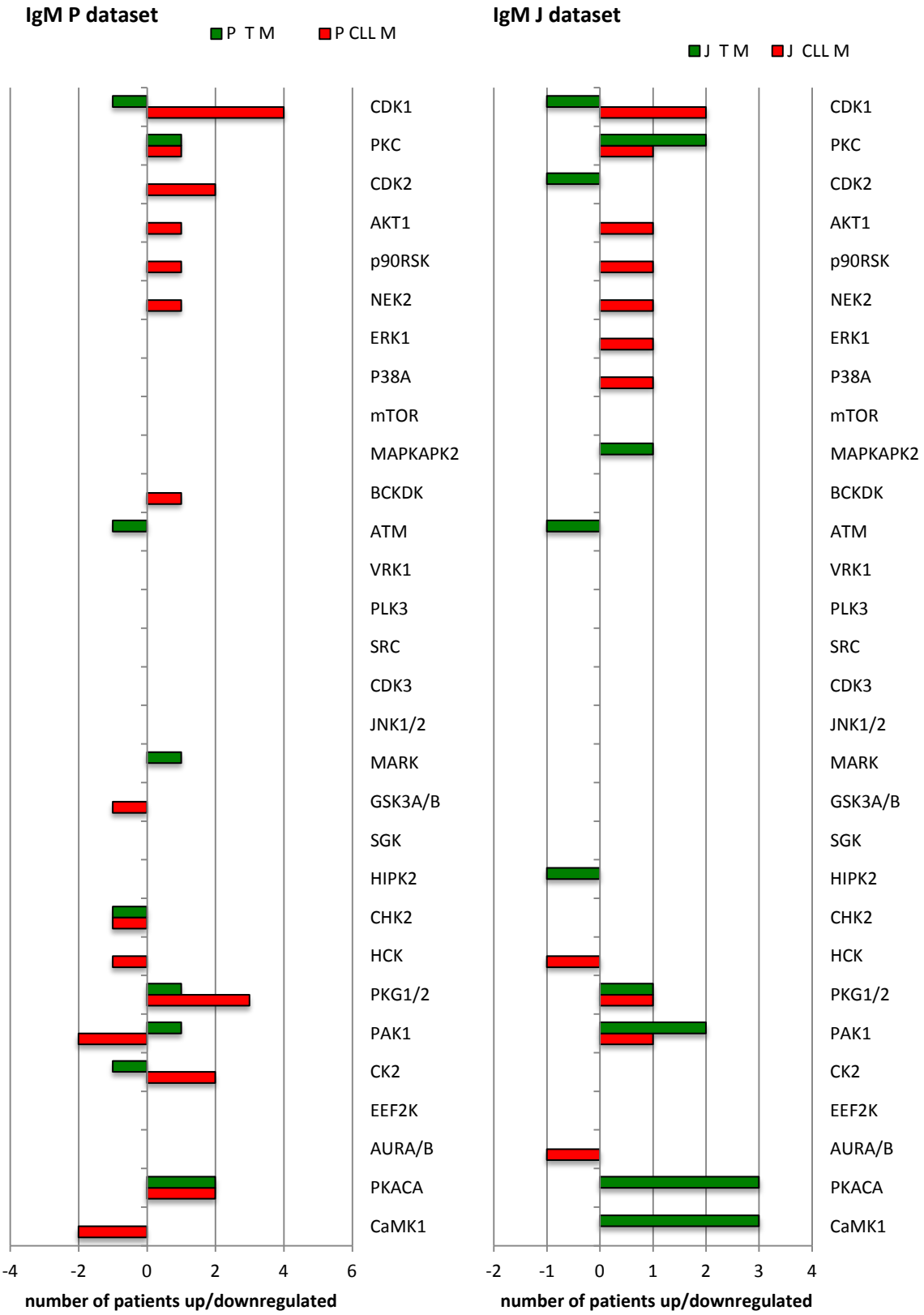
**Figure 7.47: Phosphosites with identified kinases (red) as proportion of all phosphosites for each stimulus, P dataset**




Overall, 7-11% of kinases were identified for the IgM & IgD induced phosphosites in the P and J datasets.

The identified kinases are shown below (fig 7.48). The kinase data are combined in Figure 7.49 to demonstrate the kinases and activated in CLL and Tonsil, by IgD and IgM signalling.

**Figure 7.48 kinases upstream of phosphosites differentially expressed after IgM and IgD stimulation.**
**The number of patients in which a kinase activity is up/down regulated is indicated.**



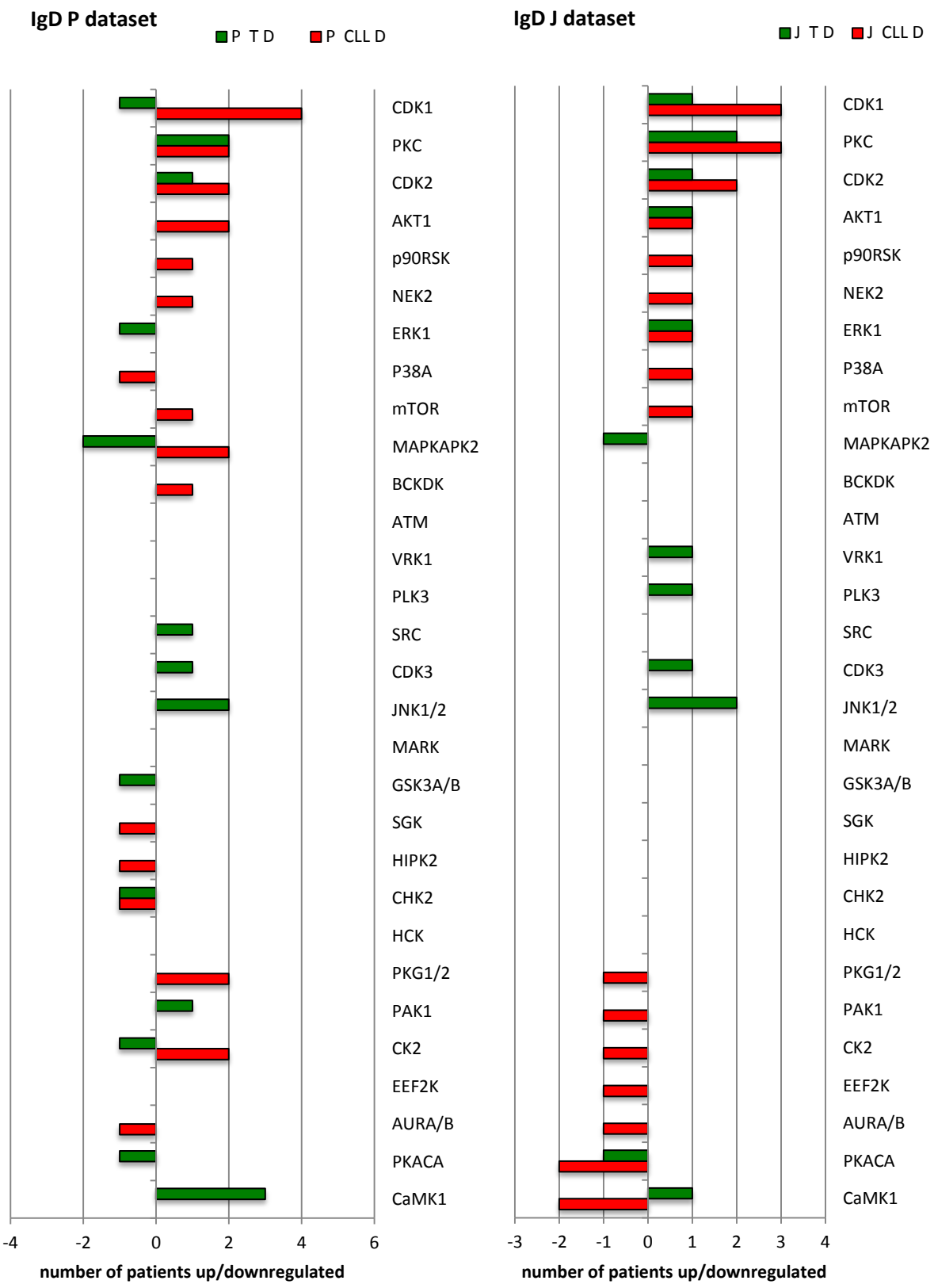
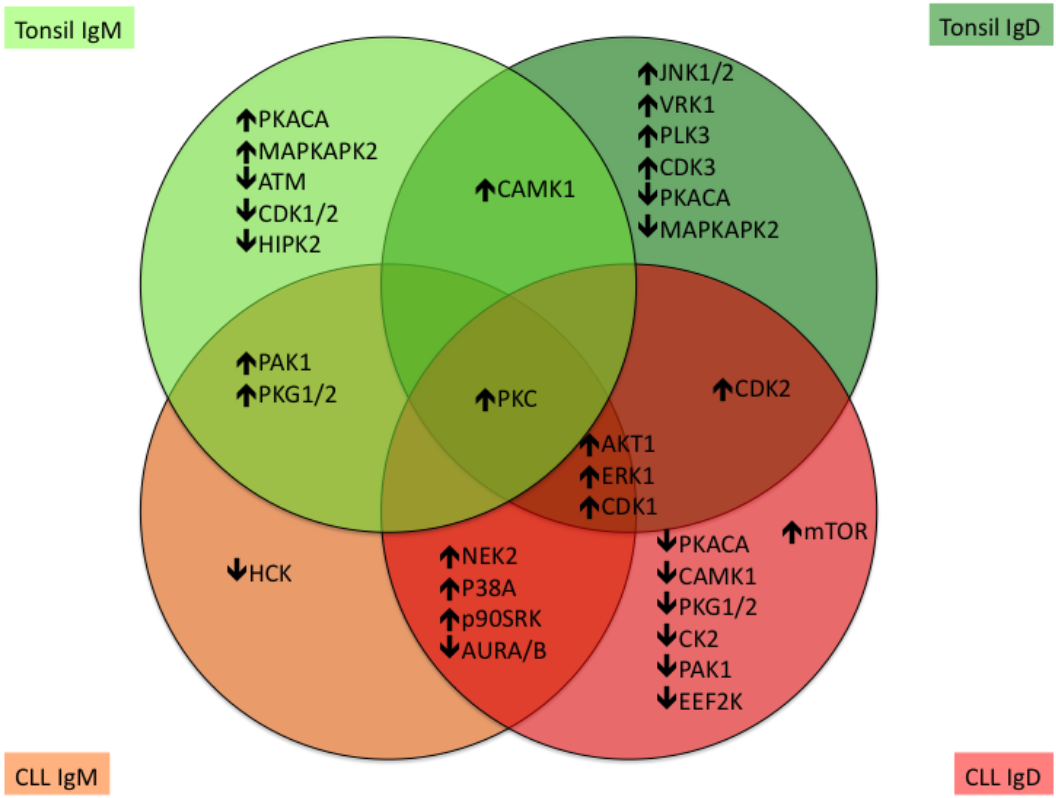

IgD P dataset
P T D
P CLL D
CDK1
PKC
CDK2
AKT1
p90RSK
NEK2
ERK1
P38A
mTOR
MAPKAPK2
BCKDK
ATM
VRK1
PLK3
SRC
CDK3
JNK1/2
MARK
GSK3A/B
SGK
HIPK2
CHK2
HCK
PKG1/2
PAK1
CK2
EEF2K
AURA/B
PKACA
CaMK1
-4
-2
0
2
4
6
number of patients up/downregulated
IgD J dataset
J T D
J CLL D
CDK1
PKC
CDK2
AKT1
p90RSK
NEK2
ERK1
P38A
mTOR
MAPKAPK2
BCKDK
ATM
VRK1
PLK3
SRC
CDK3
JNK1/2
MARK
GSK3A/B
SGK
HIPK2
CHK2
HCK
PKG1/2
PAK1
CK2
EEF2K
AURA/B
PKACA
CaMK1
-3
-2
-1
0
1
2
3
4
number of patients up/downregulated

**Figure 7.49: Kinases altered by BCR stimulation (in at least one patient)**

**P dataset**


**J dataset**



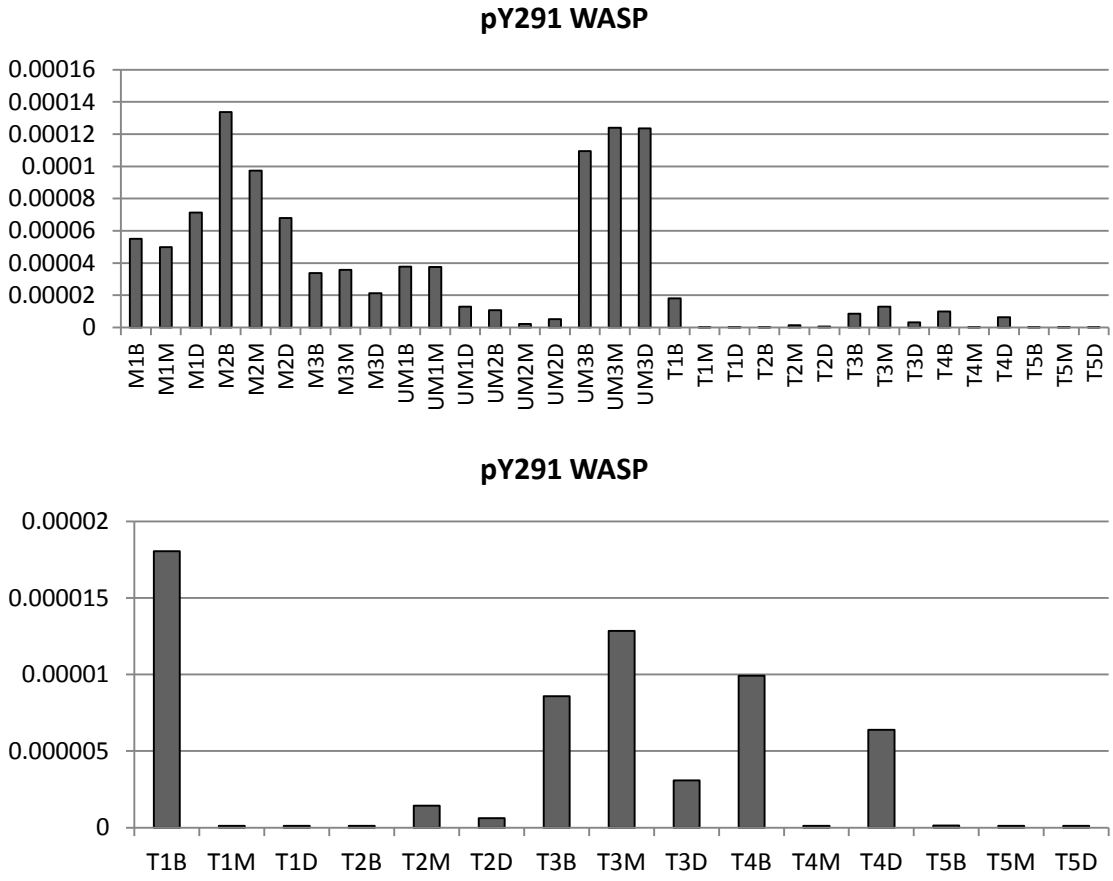
### *7.4.7 Substrates of kinases in the proximal B-Cell Receptor pathway are not generally detected, more distal kinases are upregulated by BCR stimulation*

The kinases generally associated with BCR signalling are SYK, BTK and LYN, and the analysis performed so far has not highlighted them. The generated data were examined for known substrates of these kinases.

**SYK:** The Phosphosite database lists 39 human phosphosites on 28 proteins as substrates of SYK. Whilst 8/28 proteins were in our dataset, none were known phosphosites of SYK.

**LYN:** Of 48 human phosphosite substrates of LYN, 9/35 proteins were in our dataset, with 1 known phosphosite substrate of LYN present: pS291 WASP (also downstream of ABL, ACK, LCK). This phosphosite was not significantly altered by BCR stimulation in any patient (though was significantly elevated in CLL *vs* Tonsil). The presumed LYN activity is shown in Figure 7.50

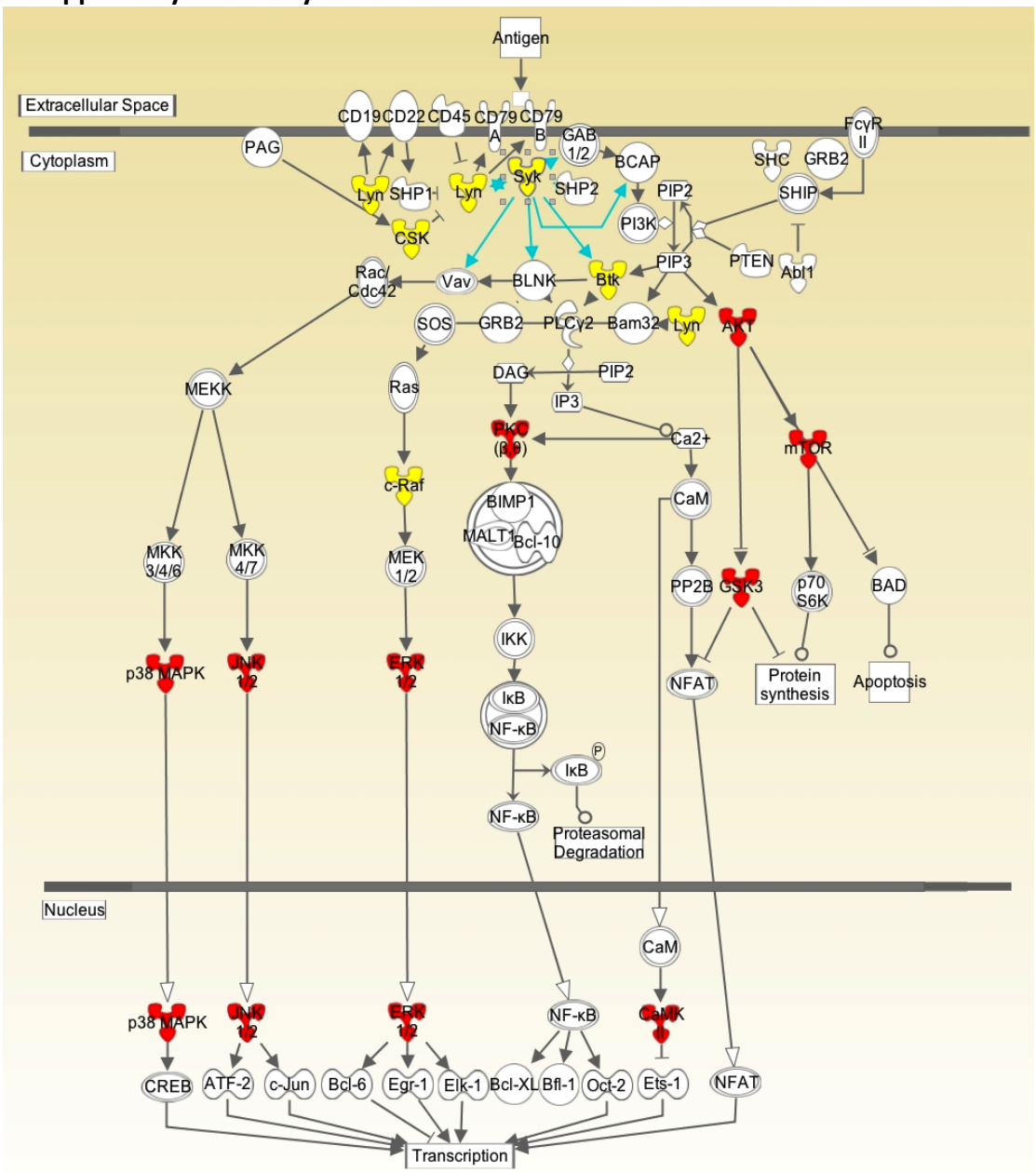
**Figure 7.50: Expression of pY291 WASP (and estimated LYN activity) in different samples.**
**The Tonsil samples are shown at higher scale below.**



**BTK:** Of the 13 phosphosites of BTK, 3/7 proteins were in our dataset, none of which had a known phosphosite. The phosphoELM database does consider pY291 WASP as a downstream substrate of BTK, with one referenced paper[557].

When the identified upstream kinases are mapped onto the canonical BCR pathway, it can be seen that whilst many of these proximal kinases are not seen to be upregulated by stimulation, many of those seen are downstream components of the canonical BCR signalling pathway (see fig 7.51)

**Figure 7.51 The B-Cell Receptor IPA canonical pathway.**
**Highlighted in red are the kinases upregulated in our datasets, in yellow are those not apparently altered by BCR stimulation in our dataset**



**Kinase prediction**

Due to the issues with the reliability of the Scansite algorithm and the labour-intensive nature of the analysis, limited kinase prediction based on BCR-stimulated phosphosites was performed, shown in figure 7.52 is the predicted kinases for one stimulus (chosen because there were limited numbers of phosphosites to analyse).

**Figure 7.52: M3M P dataset Scansite predicted kinases.**
**Each phosphosite significantly altered by IgM stimulus in samples M3 is shown with the fold change after stimulation, the known and predicted upstream kinases**

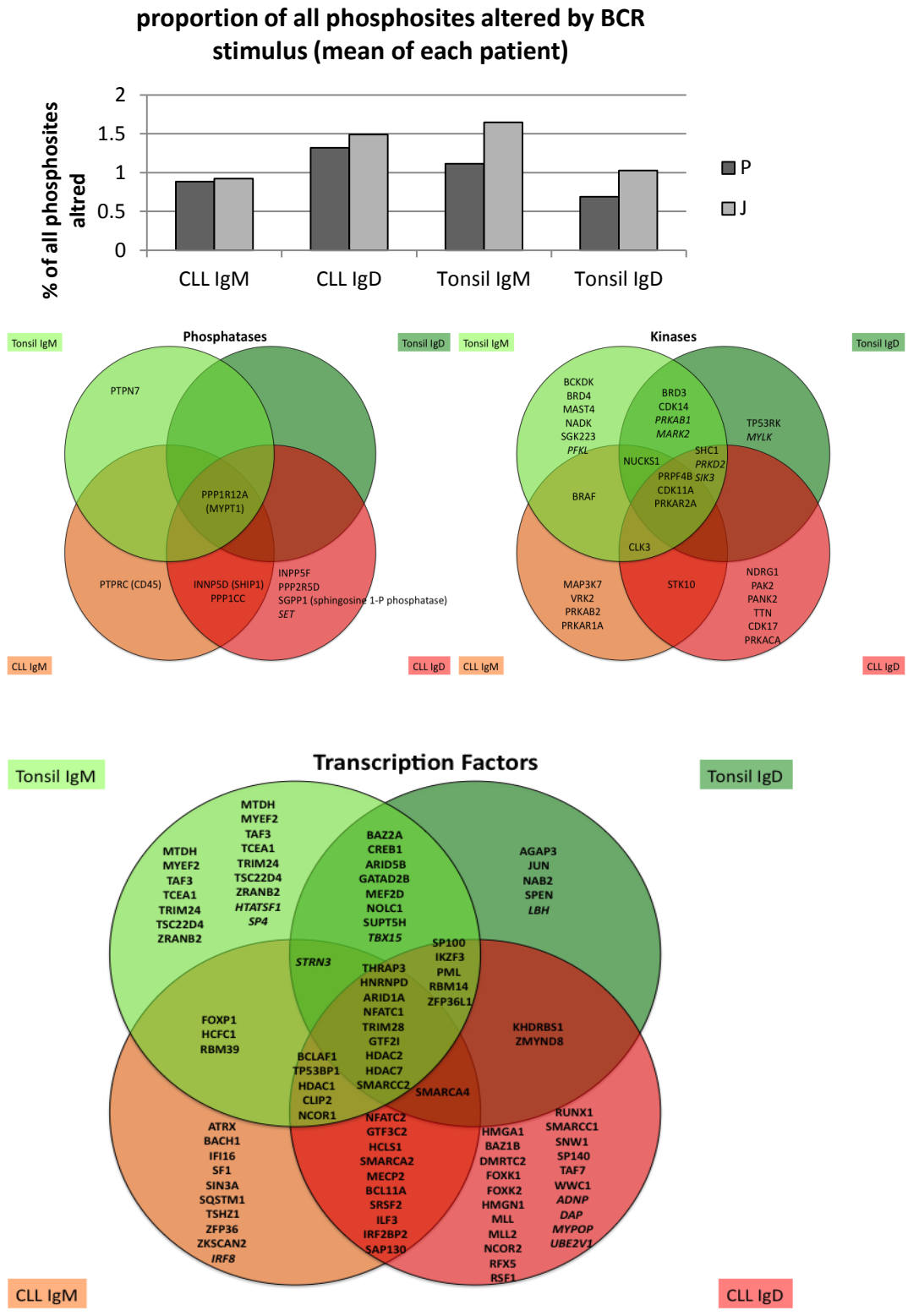
| Phosphosite | fold | Known kinase(s) | Predicted kinases (Scansite) |
|---|---|---|---|
| AHNAK P-S135 | 2390 | | AKT CAMK2 |
| METTL1 P-S27 | 784 | AKT1 P90RSK | CLK2 |
| PPP1R12A P-S299 | 59.8 | | CDK1 CDK5 PKA |
| PDCD4 P-S457 | 0.01 | AKT1 | CLK2K |
| FAM21A P-S333 | 4.06 | | AKT1 CK2 PKA |
| ARHGEF2 P-T695 | 61.9 | | CDK1 CDK5 |
| MAP3K7 P-S439 | 0.01 | | PKA AURB |
| PTPRC P-S1297 | 192 | | CDK1 CDC2K ERK1 P38MAPK |
| PLEKHA7 P-T1013 | 1.50 | | AKT1 PKA AURA |
| MAVS P-S222 | 0.05 | | CDK1 CDC2K CDK5K ERK1 GSK3A/B |
| GFRA1 P-Y420 P-S429 P-T432 | 0.01 | | GSK3A/B |
| CCDC86 P-S18 | 0.01 | | CDK1 CDK5 ERK1 |
| EIF3G P-T38 | 0.02 | | GSK3A/B |
| HNRNPM P-S618 | 76.8 | | DNAPK |
| TMPO P-S180 P-S184 | 1.59 | | CK2 |
| PTGES3 P-S113 I | 4.08 | CK2 | DNAPK |
| CAH4 P-S101 | 5.43 | | AURB |
| EIF2B5 P-S544 | 0.02 | GSK3A/B | CDK1 CDC2K CDK5 |
| MYL9 P-S20 | 0.00 | CAMK1A | AURB |
| LMO7 P-S246 | 48.9 | | AKT1 AMPK DNAPK CLK2K |
| STK10 P-S951 | 2.65 | | CDK1 |
| IL16 P-S844 | 3.35 | | AMPK AURB PKCE |

### *7.4.8 Summary: The effects of BCR signalling*

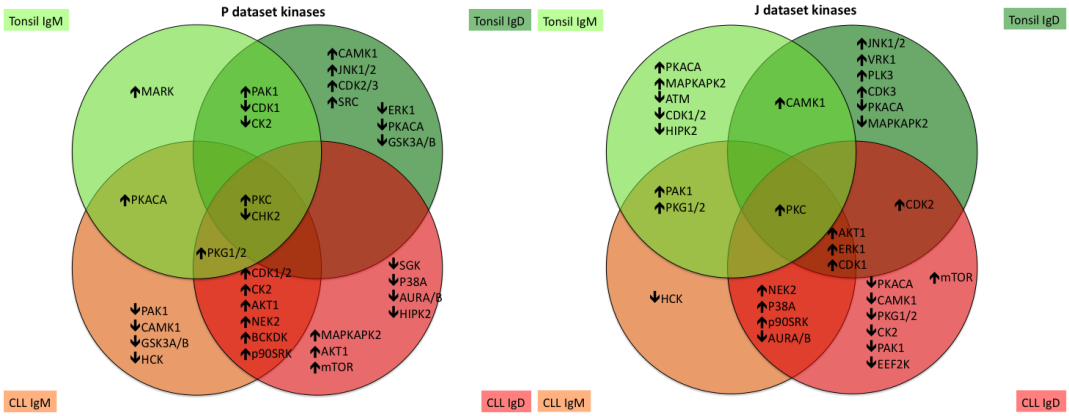
There is no simple means to summarise these data on events occurring after BCR ligation, or to highlight the differences between CLL and Tonsil, and between IgM and IgD. As an attempt, the summary images from above are combined into one image (Figure 7.53).

**Figure 7.53 BCR Signalling Summary.**

- **The mean proportions of all phosphosites altered by BCR stimuli, P and J datasets shown separately.**
- **BCR altered phosphorylations of kinases, phosphatases and transcription factors**
- **BCR altered IPA pathways**
- **BCR altered kinase activities**




Tonsil IgM
Tonsil IgD
IPA pathways
•RAR Activation
•Assembly of RNA Polymerase III Complex
• Dopamine-DARPP32 Feedback in cAMP Sig.
Neuroprotective Role of THOP1 in Alzheimer's
•FcγRIIB Sig. in B Lymphocytes
•Natural Killer Cell Sig.
•Ins(1,4,5)P3 Degradation
• Renin-Angiotensin Sig.
• 1D-myo-inositol Hexakisphosphate Biosynthesis II
•Ins(1,3,4)P3 Biosynthesis
• Cyclins & Cell Cycle Regulation
• PDGF Sig.
•Cdc42 Sig. SAPK/JNK Sig.
• Hypoxia Sig. in the Cardiovascular System
• Neurotrophin/TRK Sig.
•GDNF Family Ligand-Receptor Interactions
•Regulation of Actin-based Motility by Rho
•PPAR Sig.
•EIF2 Sig.
•ERK/MAPK Sig.
•Glucocorticoid Receptor Sig.
• AMPK Sig.
•B Cell Receptor Sig.
•Integrin Sig.
•Granzyme A Sig.
•Insulin Receptor Sig.
• Netrin Sig.
•Dopamine Receptor Sig.
• IL-4 Sig.
• Uridine-5'-phosphate Biosynthesis
•CDK5 Sig.
•Phospholipase C Sig.
•Estrogen Receptor Sig.
•Spliceosomal Cycle
•Cardiac β-adrenergic Sig.
•April Mediated Sig.
•BAFF Sig.
•Role of NFAT in Cardiac Hypertrophy
•Diphthamide Biosynthesis
•Sig. by Rho Family GTPases
• PPARα/RXRα Activation
•DNA Methylation and Transcriptional Repression Sig.
•Hereditary Breast Cancer Sig.
•Calcium Sig.
•Telomerase Sig.
•Protein Kinase A Sig.
•Systemic Lupus Erythematosus Sig.
•Role of BRCA1 in DNA Damage Response
• iCOS-iCOSL Sig. in T Helper Cells
• Acetyl-CoA Biosynthesis III
•DNA Double-Strand Break Repair by Homologous Recombination
• Telomere Extension by Telomerase
• PI3K Sig. in B Lymphocytes
• ILK Sig.
•Role of CHK Proteins in Cell Cycle Checkpoint Control
•Tight Junction Sig.
• Breast Cancer Regulation by Stathmin1
• Nur77 Sig. in T Lymphocytes
• ATM Sig.
•Cell Cycle: G1/S Checkpoint Regulation
•Sucrose Degradation V Glycolysis I
•Actin Nucleation by ARP-WASP Complex
• RhoGDI Sig.
•Thrombin Sig.
•NADH Repair
•Calcium-induced T Lymphocyte Apoptosis
•Huntington's Disease Sig.
• cAMP-mediated sig.
•Gluconeogenesis I
•ERK5 Sig.
•Corticotropin Releasing Hormone Sig.
•Melatonin Sig.
•Synaptic Long Term Potentiation
CLL IgM
CLL IgD


Tonsil IgM
P dataset kinases
↑MARK
↑PAK1
↓CDK1
↓CK2
↑CAMK1
↑JNK1/2
↑CDK2/3
↑SRC
↓ERK1
↓PKACA
↓GSK3A/B
↑PKACA
↑PKC
↓CHK2
↑PKG1/2
↑CDK1/2
↑CK2
↑AKT1
↑NEK2
↑BCKDK
↑p90SRK
↓SGK
↓P38A
↓AURA/B
↓HIPK2
↓PAK1
↓CAMK1
↓GSK3A/B
↓HCK
↑MAPKAPK2
↑AKT1
↑mTOR
CLL IgM
Tonsil IgD
Tonsil IgM
J dataset kinases
Tonsil IgD
↑PKACA
↑MAPKAPK2
↓ATM
↓CDK1/2
↓HIPK2
↑CAMK1
↑JNK1/2
↑VRK1
↑PLK3
↑CDK3
↓PKACA
↓MAPKAPK2
↑PAK1
↑PKG1/2
↑PKC
↑CDK2
↑AKT1
↑ERK1
↑CDK1
↓HCK
↑NEK2
↑P38A
↑p90SRK
↓AURA/B
↓PKACA
↓CAMK1
↓PKG1/2
↓CK2
↓PAK1
↓EEF2K
↑mTOR
CLL IgD
CLL IgM
CLL IgD

## 7.5 Results: Validation of phosphoproteomic quantification using western blotting

### *7.5.1 Introduction*

In order to validate the results of gene expression microarrays, quantitative real-time PCR is used to confirm and validate the relative quantifications of target genes. Similarly, quantification of individual phosphoproteins is desirable to validate the findings of phosphoproteomic data. In this situation, the analogous technique would be detection of phosphoproteins by western blotting[533,558]. Western blotting is often not considered to be a quantitative method, but a qualitative or semi-quantitative one, where relative quantifications of proteins detected by western blotting can be used to derive conclusions about relative quantifications in the tissue of interest. Beyond simple visual inspection of western blot, numerous authors have used quantification of proteins using densitometry[444], and this technique was used here.

Most known phosphoproteins do not have specific antibodies available for purchase, and this is true of the majority of the 3179 phosphosites in our analysis. Antibodies were sought for their specificity for phosphoproteins that had been significantly different between unstimulated and stimulated samples, or between CLL and tonsil, in order to minimise the variability between samples due to technical and biological noise.

### *7.5.2 Amounts of protein in lysates and available antibodies*

Each sample produced 0.2-2mg/mL of protein, with the tonsil samples (with a lower number of starting cells) producing at most 0.6 mg/mL. With 12 well gels, the maximum volume that can be loaded is 45 μL each well, setting an upper limit on the amount of protein that could be run. This led to some tonsil lysates being unusable, and when feasible, only 20μg of protein could be loaded into each well. The CLL samples, with higher concentrations, were used in a number of electrophoresis runs, but again, the amount of lysates was limited to 30μg/well. Furthermore, low protein concentrations are often poorly measured using the Bradford assays, and this led to unequal well loading in gels, as evaluated by GAPDH staining. Coupled with the inherent variability of western blotting, this made reliable quantification of phosphoproteins challenging.

To select appropriate antibodies for western blotting, the lists of phosphosites that were significantly differentially expressed between either CLL *vs* tonsil (220 in P & J datasets combined), or between the different stimuli (96 phosphosites altered by BCR stimulation in at least 2 samples in P & J datasets) were considered.

Of these approximately 300 differentially expressed phosphoproteins, commercially available antibodies were available for only 2, both in the lists of sites differentially expressed between CLL and tonsil: Anti-Filamin A pS2152 and anti –WASP pY291.
The search was then extended to the approximately 900 phosphosites that were differently expressed in at least one sample after BCR stimulation. After consideration of the 300 greatest fold changes between samples, a further 5 antibodies were found to be commercially available.
Anti-phosphosite antibodies, with the full list in table 7.23

**Table 7.23: Proteins and phosphosites of for commercially available antibodies used for western blotting**

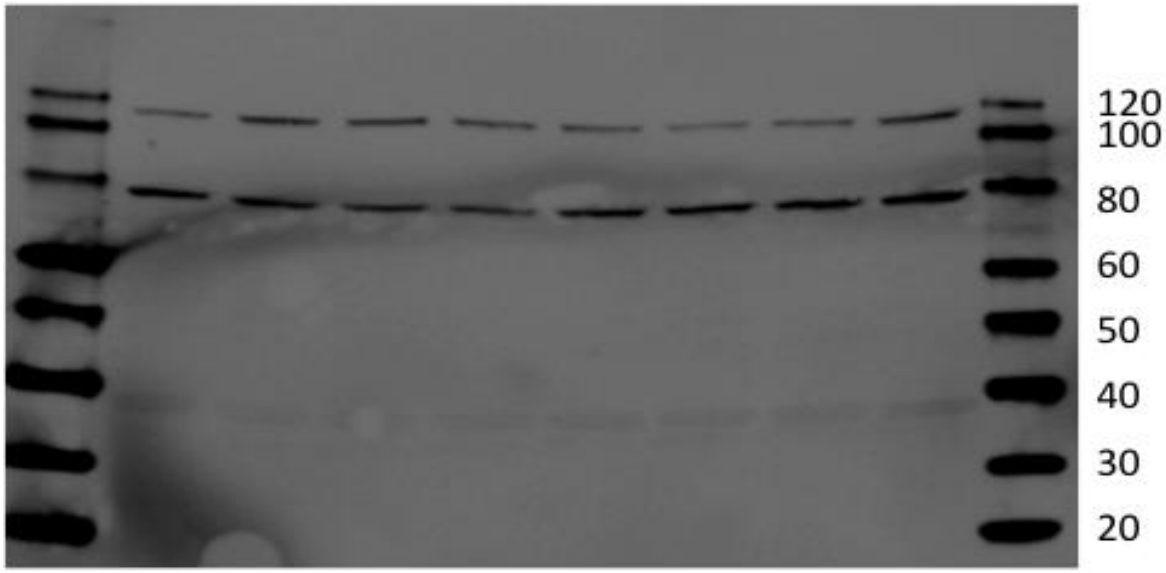
| Protein | Phosphosite |
|---|---|
| Filamin A | FLNA p-S2152 |
| Wiskott-Adrich syndrome protein | WAS p-Y291 |
| 40S ribosomal protein S6 | RS6 p-S235/S236 |
| Lymphocyte-specific protein 1 | LSP1 p-S204 |
| Myosin regulatory light polypeptide 9 | MYL9 pS20 |
| Myosin-9 | MYH9 pS1943 |
| Stathmin 1 | STMN1 p-S16 |

Of these, only FLNA and RS6 provided usable data. The other antibodies either produced non-specific staining, or no bands were seen. Briefly, a summary of the problems with the antibodies follows.

**LSP1 p-S204 (52kDa)**

The antibody used for pLSP1 produced two clear bands, neither at the predicted molecular weight of 52kDa (see figure 7.54).

**Figure 7.54 Gel electrophoresis of lysates from CRL cell line, probed with anti-phospho-LSP S204 antibody.**
**No band at 52kDa is seen. The higher molecular weight bands are clearly seen, and attributed to inappropriate specificity. On the right are shown the molecular weight ladders**




**WAS p-Y291 (60kDa)**

In cell lines and CLL, the anti-pWASP antibody showed prominent staining of the pervanadate treated cell line, but no appropriate bands were apparent in any samples (see fig 7.55). I interpret this as a capacity to bind phospho-tyrosine non-specifically.

**Figure 7.55: CLL lysates (blot 2a) probed for phospho-WASP.**
**The pervanadate-treated cell line (V) stained strongly positive in this and other blots, in the absence of a 60kDa band**

**MYL9 pS20 (18kDa)**

No bands seen in cell lines or primary samples.

**MYH9 pS1943 (230kDa)**

No bands seen in cell lines or primary samples.

**STMN1 p-S16 (20kDa)**

Whilst faint bands at 20kDa were seen in cell lines, no sample from primary cells displayed any staining.

I also attempted to probe for pBtk (T223) and pSyk (T525/526) in the primary samples, as a measure of BCR signalling. There was no apparent pBtk or pSyk or their equivalent total proteins in any of the primary samples, attributed to low amounts of protein.

Altogether, the limited amount and concentration of lysates resulted in less than ideal amounts of protein being used for electrophoresis, in particular with the tonsil samples. Producing good quality, evenly loaded western blots with small amounts of protein derived from primary cells was also challenging. The limited number of primary antibodies available for probing compounded these difficulties. Most antibodies purchased did not detect any phosphoproteins, probably due to the  to the low concentrations of protein, and a lack of efficacy of the antibody used. This resulted in probing being limited to using the pRS6 and pFLMN antibodies for validation. Both antibodies used have been validated by other publications, and the expression of pRS6 and pFLMN in the proteomic dataset was generally high, which may explain their success.

For each CLL sample comparison, electrophoresis was performed twice, with differing transfer times of 7 and 9 minutes, as optimisation using cell lines suggested that the high molecular weight filamin (280 kDa) appeared to transfer better at longer transfer times, whereas the short transfer time appeared to be more appropriate for the lower molecular weight proteins such as pRS6 (32 kDa). However, FLMN, RS6 and GAPDH (37 kDa) had a degree of successful western blotting with both transfer times, and therefore some comparisons of identical samples run on separate occasions could be made, albeit with different transfer techniques. Each gel was run with pervanadate

and untreated RL cell line lysates in order to produce positive controls. Pervanadate is a phosphatase inhibitor that generally induces non-specific increase in phosphorylation. Densitometry was performed, and correlations (both Pearson and Spearman r) made with quantifications derived from phosphoproteomic analysis. Low volumes of lysates with low concentrations from the tonsil samples meant that few comparisons could be made. Only one gel was run comparing CLL and tonsil, at the higher transfer time. Overall, for correlation of western blotting with proteomics data, sufficient lysates were available for only 5 electrophoresis runs using 12-well gels, not all high quality:

**1a (7 minute transfer) and 1b (9minute transfer**)

| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1B | UM1M | UM1D | UM2B | UM2M | UM2D | UM3B | UM3M | UM3D | | **V** | **C** |

**2a (7 minute transfer) and 2b (9minute transfer)**

| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M1B | M1M | M1D | M2B | M2M | M2D | M3B | M3M | M3D | | **V** | **C** |

**3 (incorporating tonsil samples, 9 minute transfer)**

| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1B | UM2B | UM3B | M1B | M2B | T1B | T2B | T3B | T4B | T5B | **V** | **C** |

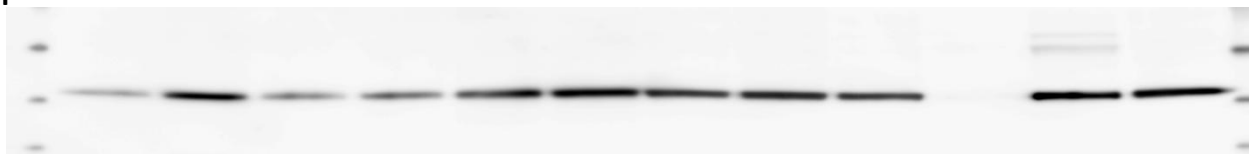
### *7.5.3 Results: Western blotting*

Densitometry was performed to derive quantifications for correlation. Those for GAPDH to compensate for loading variation divided the densitometry values for pRS6 and pFLMN. These were then correlated with phosphoproteomic quantifications, initially for those from the J dataset. For example, see figure 7.56:

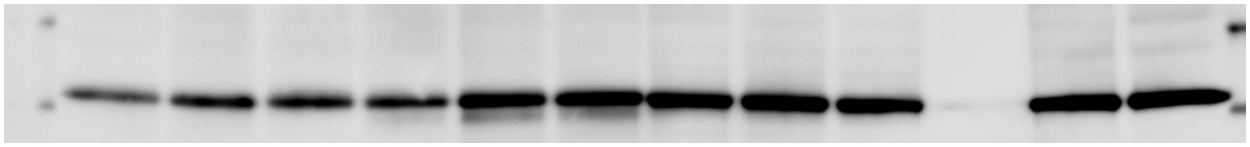
**Figure 7.56 Blot 1a (optimised transfer for RS6)-*anti-pRS6***
**The table shows the distribution of samples.**
**Empty box=empty well. V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pRS6, RS6 and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pRS6/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1B | UM1M | UM1D | UM2B | UM2M | UM2D | UM3B | UM3M | UM3D | | V | C |

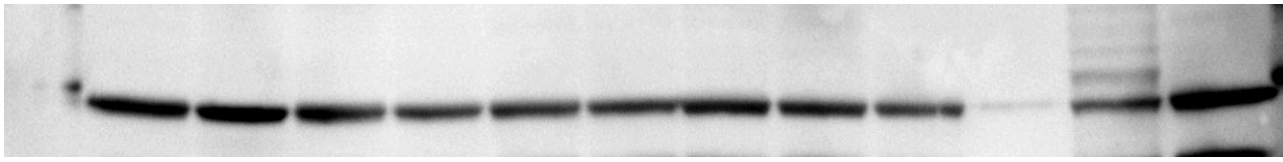
**pRS6**

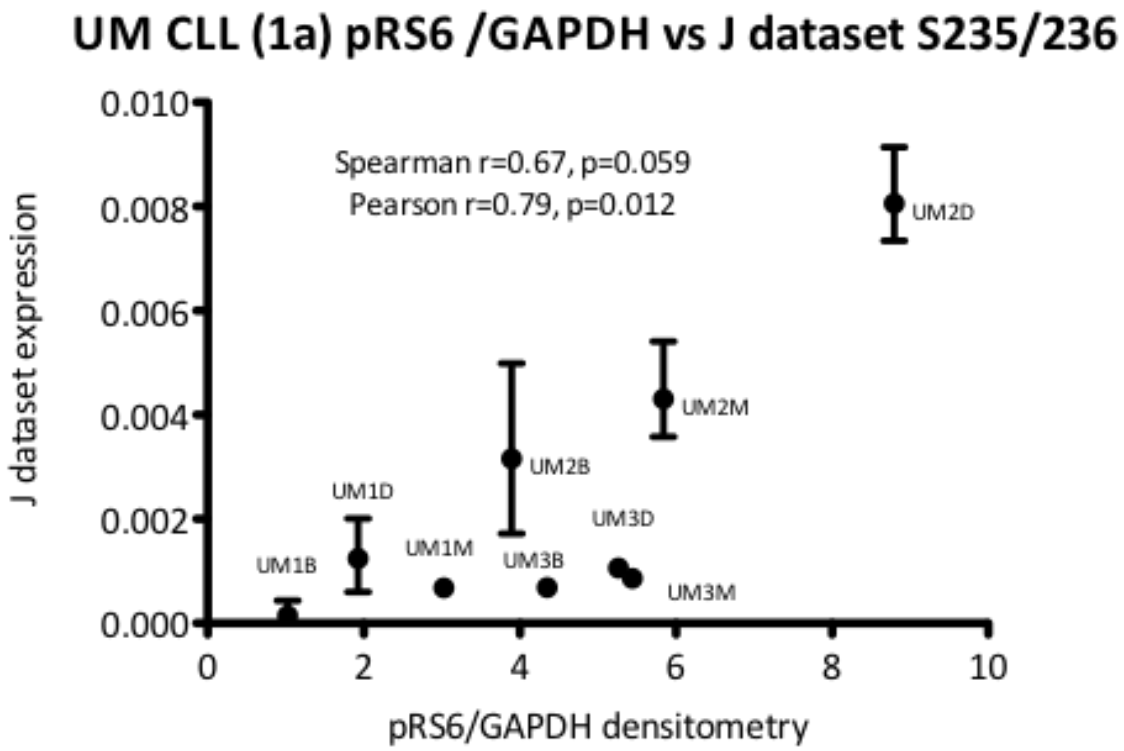


**total RS6**



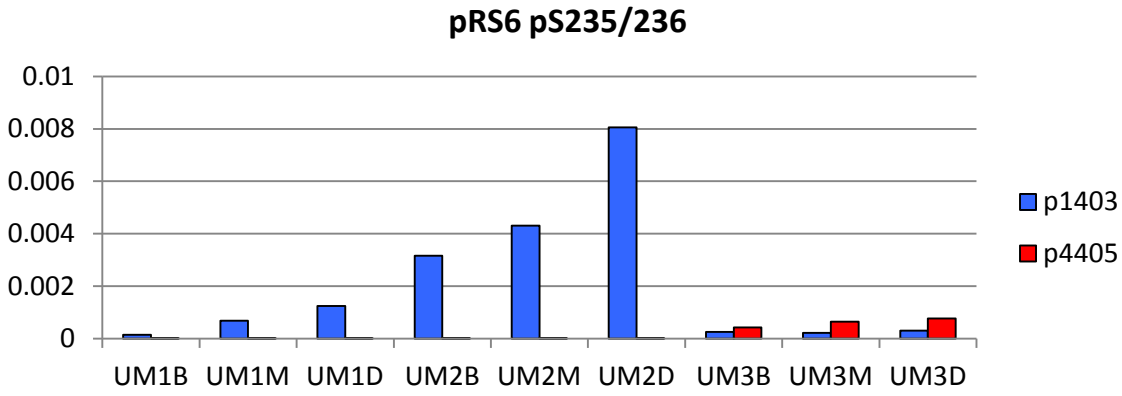
**GAPDH**





Therefore, there appears to be some correlation between western blotting densitometric determination of pRS6 expression and the J dataset means of triplicates. In order to determine the equivalent correlation between P dataset means and the western blots, correlations were also calculated for each feature representing pRS6. Two features, with the same peptide fragment, but differing in that p4405 was formylated, represented the pS235/236 phosphosite of RS6 (RL**S$^{p}$S$^{p}$**LR). The expression levels and correlations are represented in figure 7.57:

**Figure 7.57: Expression of two features in J dataset representing RS6 pS235/236. Shown are the expression levels of p1403 and p4405 (formylated version of p1403). Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of p1403 and p4405.**



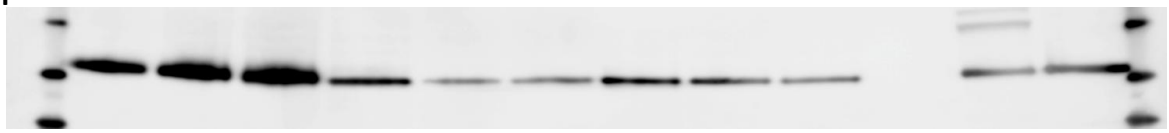

| | p1403 | p4405 | J dataset |
|---|---|---|---|
| Pearson r | 0.7221 | 0.2239 | 0.7883 |
| P value (two-tailed) | **0.0280** | 0.5625 | **0.0116** |
| Spearman r | 0.4833 | 0.2871 | 0.6667 |
| P value (two-tailed) | 0.1938 | 0.4630 | 0.0589 |

Thus, it can be seen that the more abundant feature (p1403) produces a higher (and significant) Pearson r, whilst the sum of the two features' expressions in the J dataset produces a higher Pearson and Spearman r (although p=0.059). The same analysis was used for subsequent blots. See Figures 7.58 & 7.59

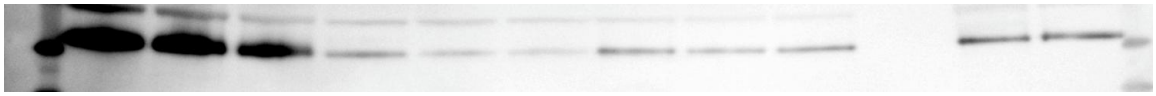
**Figure 7.58 Blot 2a (optimised transfer for RS6): *anti-pRS6***
**The table shows the distribution of samples.**
**Empty box=empty well. V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pRS6, RS6 and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pRS6/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M1B | M1M | M1D | M2B | M2M | M2D | M3B | M3M | M3D | | V | C |

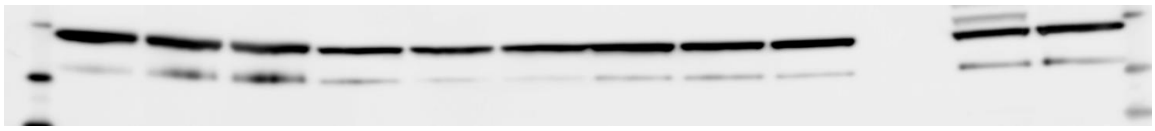
**pRS6**



**total RS6**



**GAPDH**



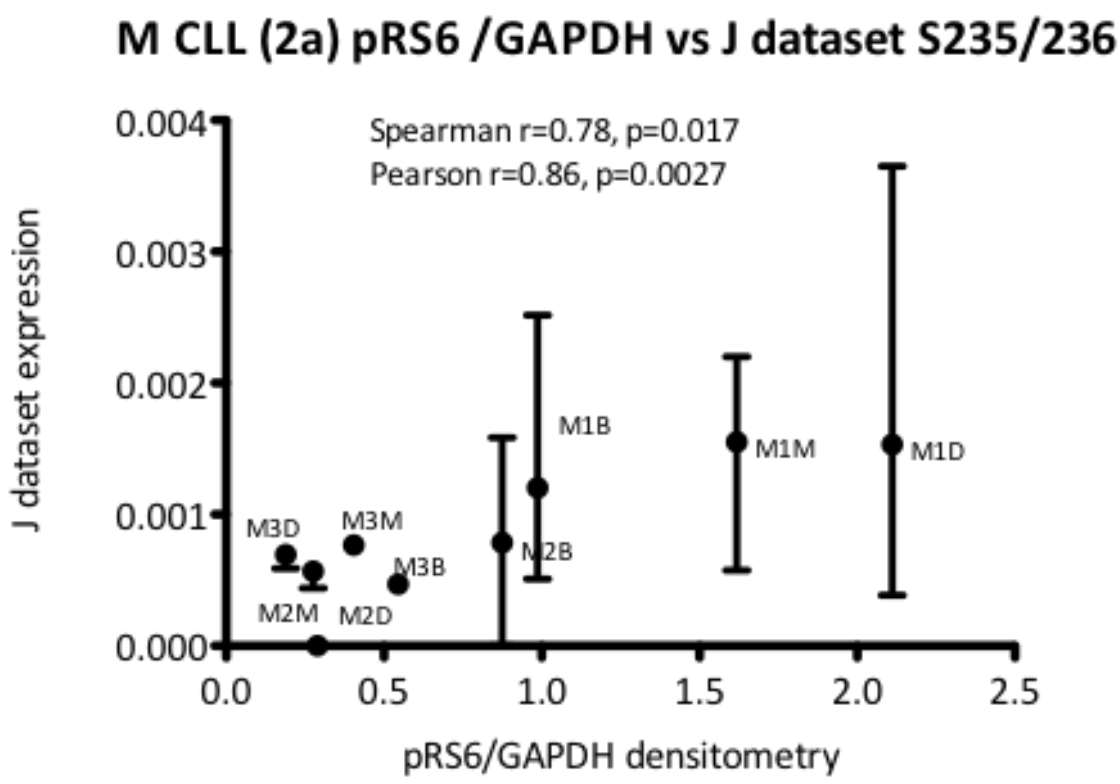
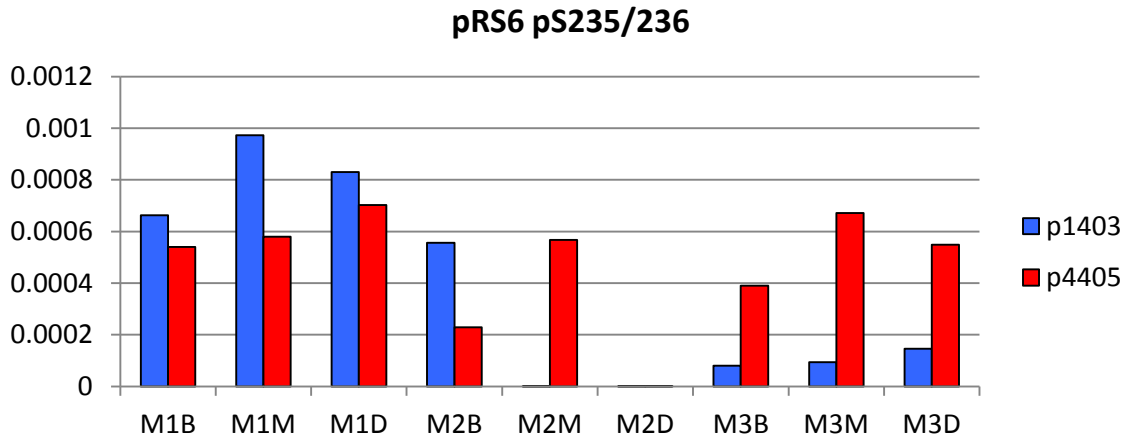


**Figure 7.59: Expression of two features in J dataset representing RS6 pS235/236. Shown are the expression levels of p1403 and p4405 (formylated version of p1403). Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of p1403 and p4405.**




| | p1403 | p4405 | J dataset |
|---|---|---|---|
| Pearson r | 0.9117 | 0.3881 | 0.8628 |
| P value (two-tailed) | **0.0006** | 0.3020 | **0.0027** |
| Spearman r | 0.7950 | 0.3167 | 0.7833 |
| P value (two-tailed) | **0.0138** | 0.4101 | **0.0172** |

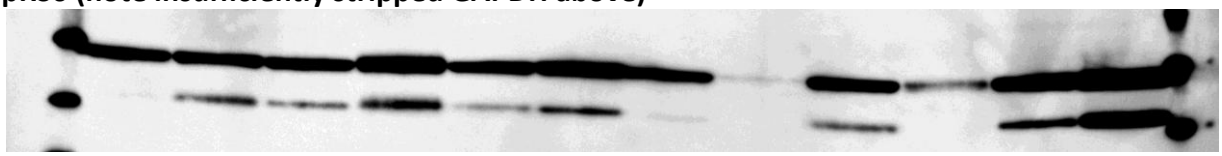
Again, p1403 demonstrates the higher correlation, though in this group of samples, p4405 has a higher level of expression in certain samples.

The same was performed for comparisons incorporating tonsil baselines (blot 3, Figure 7.60), though there was uneven loading.

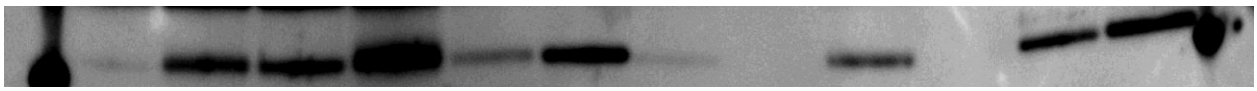
**Figure 7.60 Blot 3 (optimised transfer for FLMN): *anti-pRS6***
**The table shows the distribution of samples.**
**V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pRS6, RS6 and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pRS6/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data. Because of the poor loading of T3B and T5B, correlations are also shown when these samples are removed from the analysis**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1B | UM2B | UM3B | M1B | M2B | T1B | T2B | T3B | T4B | T5B | **V** | **C** |

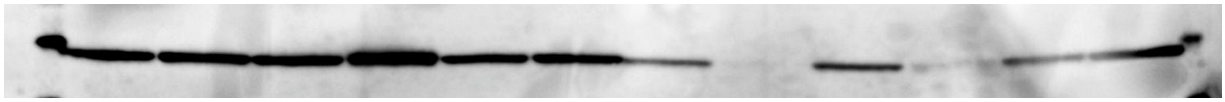
**pRS6 (note insufficiently stripped GAPDH above)**



**total RS6**



**GAPDH**



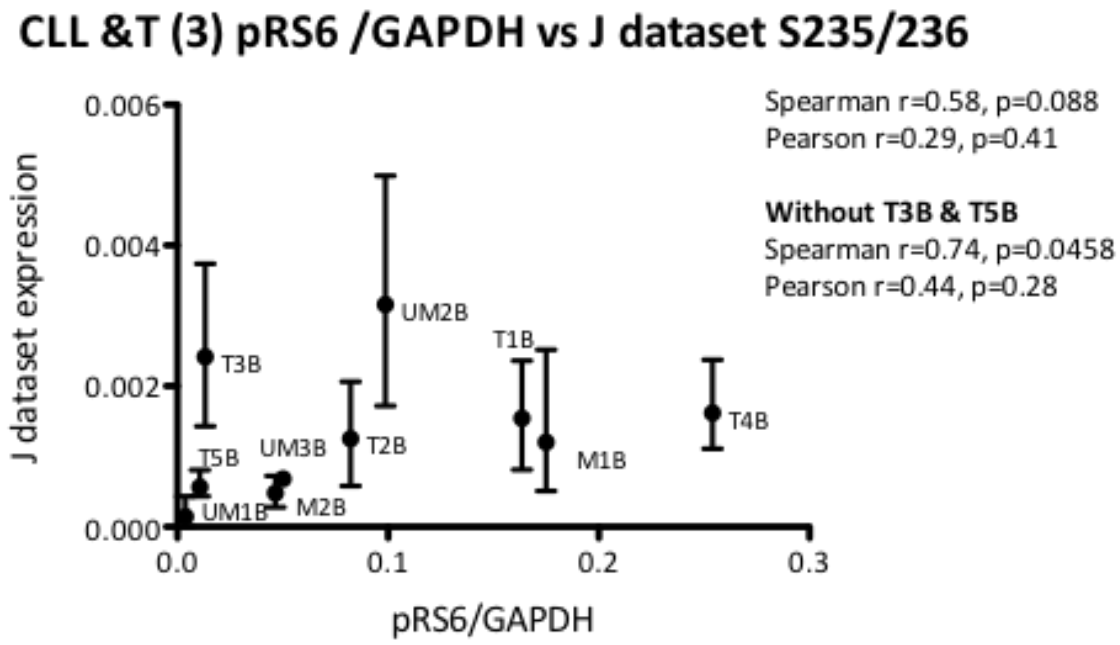
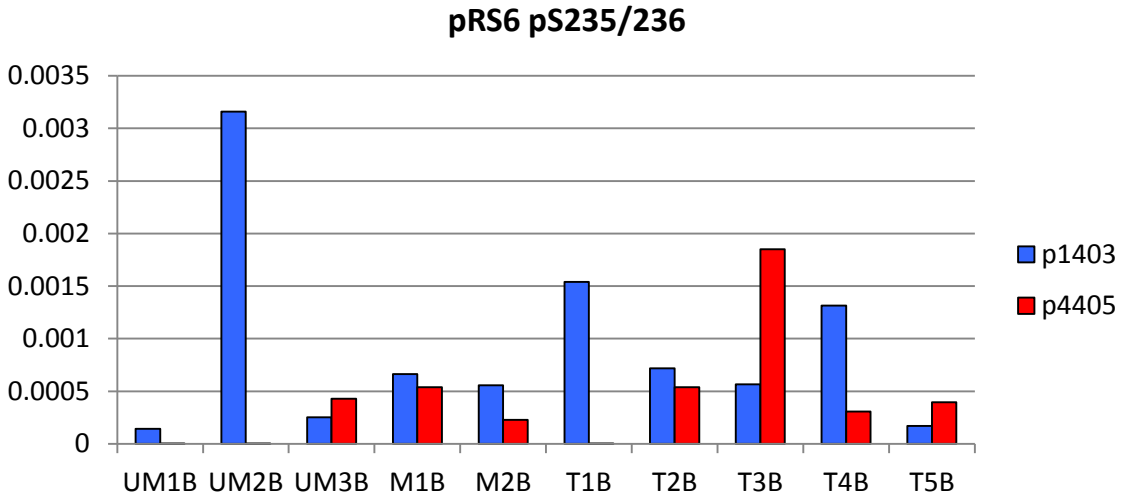


**Figure 7.61: Expression of two features in J dataset representing RS6 pS235/236. Shown are the expression levels of p1403 and p4405 (formylated version of p1403). Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of p1403 and p4405. Because of the poor loading of T3B and T5B, correlations are also shown when these samples are removed from the analysis**




| | p1403 | p4405 | J dataset |
|---|---|---|---|
| Pearson r | 0.4443 | -0.2633 | 0.2909 |
| P value (two-tailed) | 0.1983 | 0.4624 | 0.4148 |
| Spearman r | 0.8061 | 0.01841 | 0.5758 |
| P value (two-tailed) | **0.0072** | 0.9730 | 0.0883 |

| **without T3B and T5B** | **p1403** | **p4405** | **J dataset** |
|---|---|---|---|
| Pearson r | 0.3532 | 0.1880 | 0.4360 |
| P value (two-tailed) | 0.3908 | 0.6557 | 0.2802 |
| Spearman r | 0.6905 | 0.2928 | 0.7381 |
| P value (two-tailed) | 0.0694 | 0.4618 | **0.0458** |

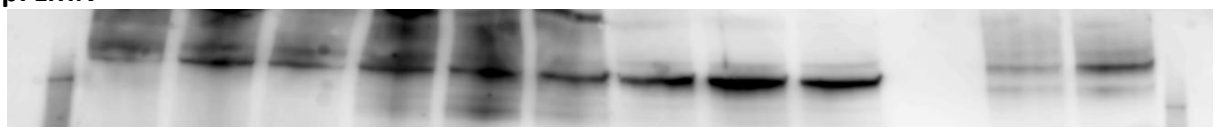
Again, p1403 has the greater correlation with densitometry, though this is actually reduced by removal of the poor-quality T3B and T5B densitometry values, whilst this correction results in somewhat improved correlation in the J dataset.

The same approach was adopted when probing for phospho-FLMN on the blots optimised for the transfer at high molecular weights (see Figures 7.62-7.67 ). Of note, phospho-FLMN was selected because of its significantly higher expression levels in all CLL samples relative to all tonsil samples, which is reflected in the western blots.

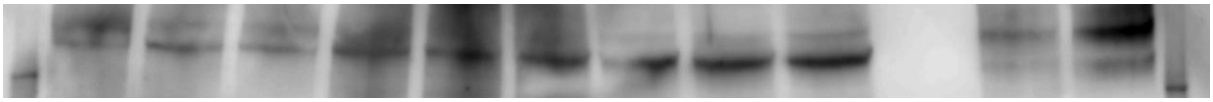
**Figure 7.62 Blot 1b (optimised transfer for FLMN): *anti-pFLMN***
**The table shows the distribution of samples.**
**V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pFLMN, FLMN and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pFLMN/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1 B | UM1 M | UM1 D | UM2 B | UM2 M | UM2 D | UM3 B | UM3 M | UM3 D | | V | C |

**pFLMN**



**total FLMN**



**GAPDH**

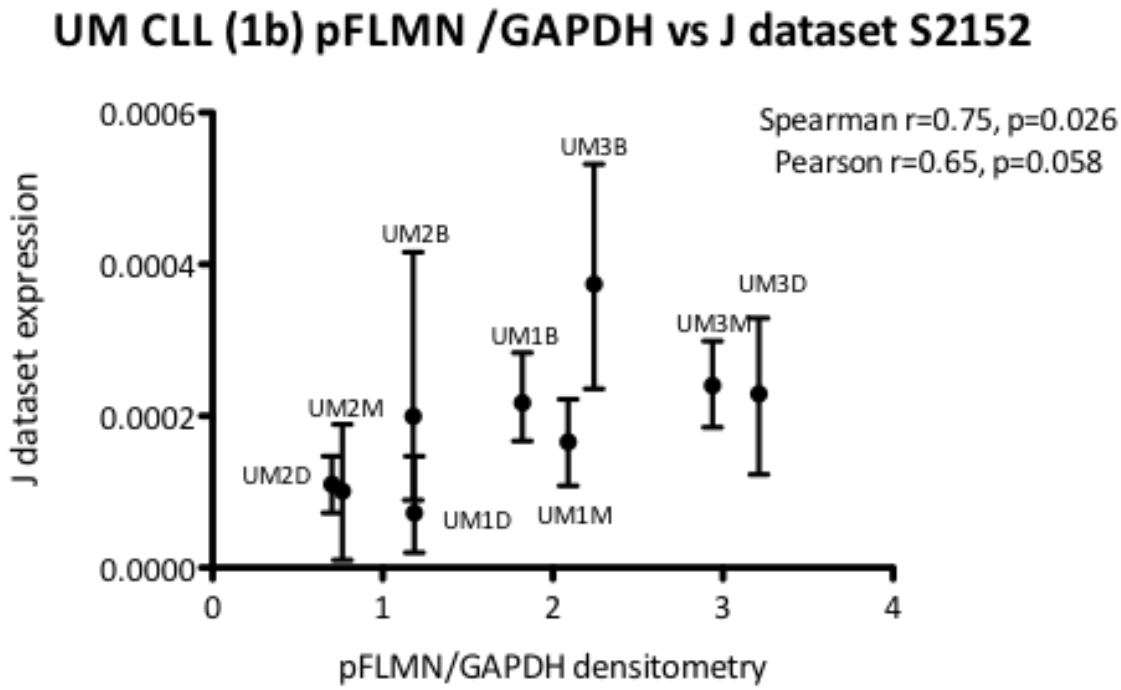
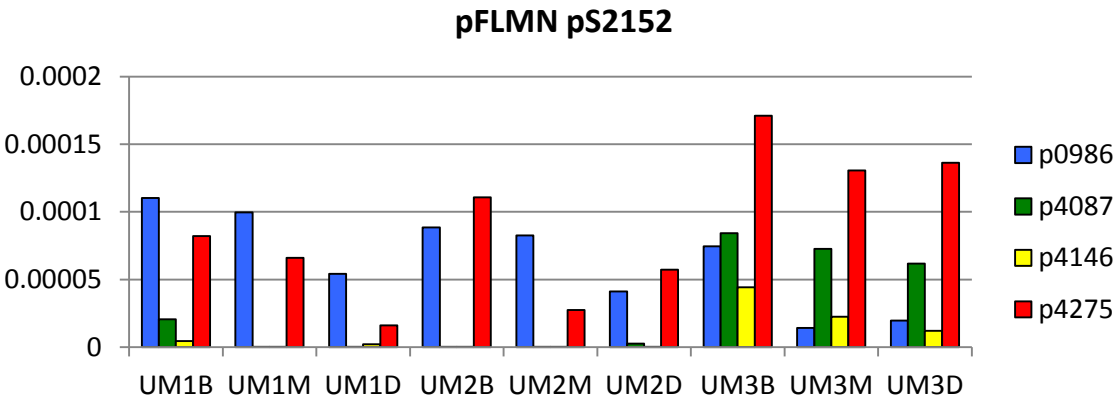


**Figure 7.63: Expression of 4 features in J dataset representing FLMN pS2152. Shown are the expression levels of p0986, p4087, p4146. p4275. Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of all 4 features.**



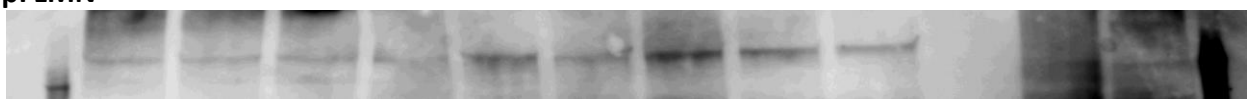

| | p0986 | p4087 | p4146 | p4275 | J dataset |
|---|---|---|---|---|---|
| Pearson r | -0.4356 | 0.7985 | 0.5600 | 0.7239 | 0.6494 |
| P value (two-tailed) | 0.2412 | **0.0099** | 0.1169 | **0.0274** | 0.0584 |
| Spearman r | -0.3667 | 0.6963 | 0.7833 | 0.7500 | 0.7500 |
| P value (two-tailed) | 0.3363 | **0.0433** | **0.0172** | **0.0255** | **0.0255** |

Figure 7.63 suggests that 3 of 4 features are correlated to some extent with the densitometry values, and in some cases, this relationship is weakened by considering their sum (J dataset). Also, the poorly correlated p0986 feature is not the least abundant, suggesting that poor correlation is not due to low abundance features.

**Figure 7.64 Blot 2b (optimised transfer for FLMN): *anti-pFLMN***
**The table shows the distribution of samples.**
**V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pFLMN, FLMN and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pFLMN/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M1C | M1M | M1D | M2C | M2M | M2D | M3C | M3M | M3D | | **V** | **C** |

**pFLMN**



**total FLMN**

**GAPDH**

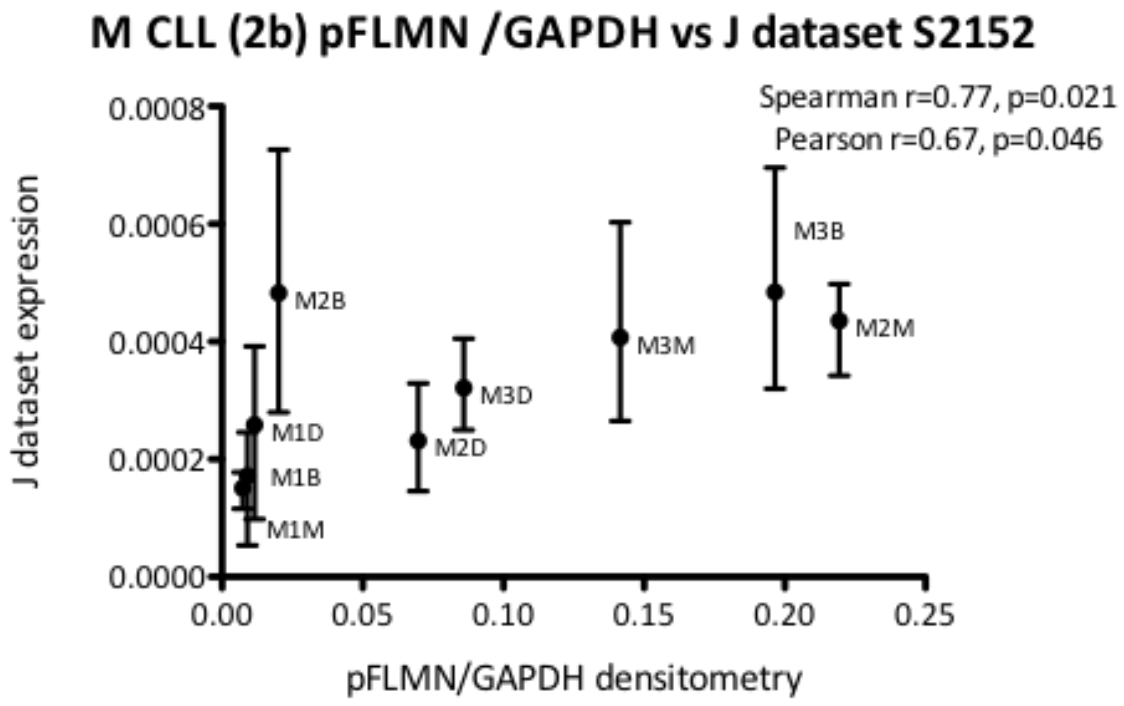
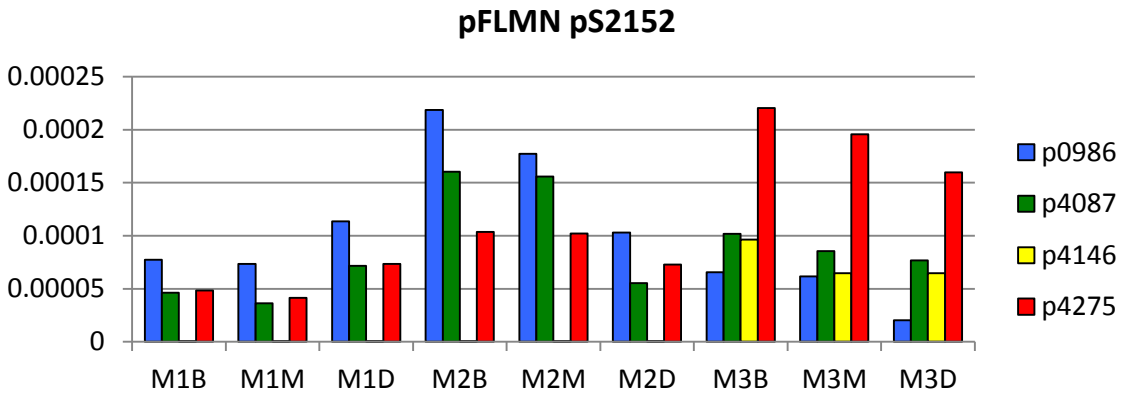


**Figure 7.65: Expression of 4 features in J dataset representing FLMN pS2152. Shown are the expression levels of p0986, P4087, p4146. p4275. Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of all 4 features.**



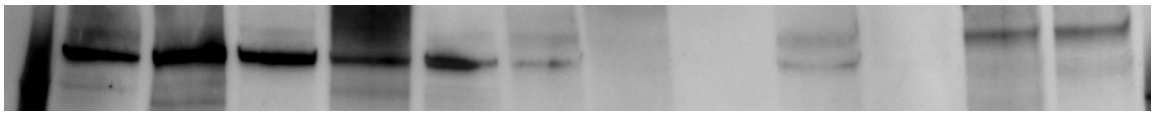

| | **p0986** | **p4087** | **p4146** | **p4275** | **J dataset** |
|---|---|---|---|---|---|
| Pearson r | -0.01445 | 0.4958 | 0.5655 | 0.6959 | 0.6742 |
| P value (two-tailed) | 0.9706 | 0.1747 | 0.1125 | **0.0373** | **0.0464** |
| Spearman r | -0.1333 | 0.7167 | 0.5743 | 0.7667 | 0.7667 |
| P value (two-tailed) | 0.7435 | **0.0369** | 0.1080 | **0.0214** | **0.0214** |

**Figure 7.66 Blot 3 (optimised transfer for FLMN): *anti-pFLMN***
**The table shows the distribution of samples.**
**V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pFLMN, FLMN and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pFLMN/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data. Because of the poor loading of T3B and T5B, correlations are also shown when these samples are removed from the analysis.**

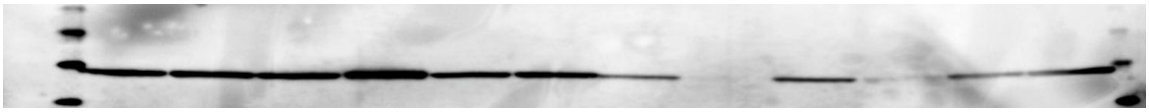
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1 B | UM2 B | UM3 B | M1B | M2B | T1B | T2B | T3B | T4B | T5B | V | C |

**pFLMN**

**total FLMN**



**GAPDH**



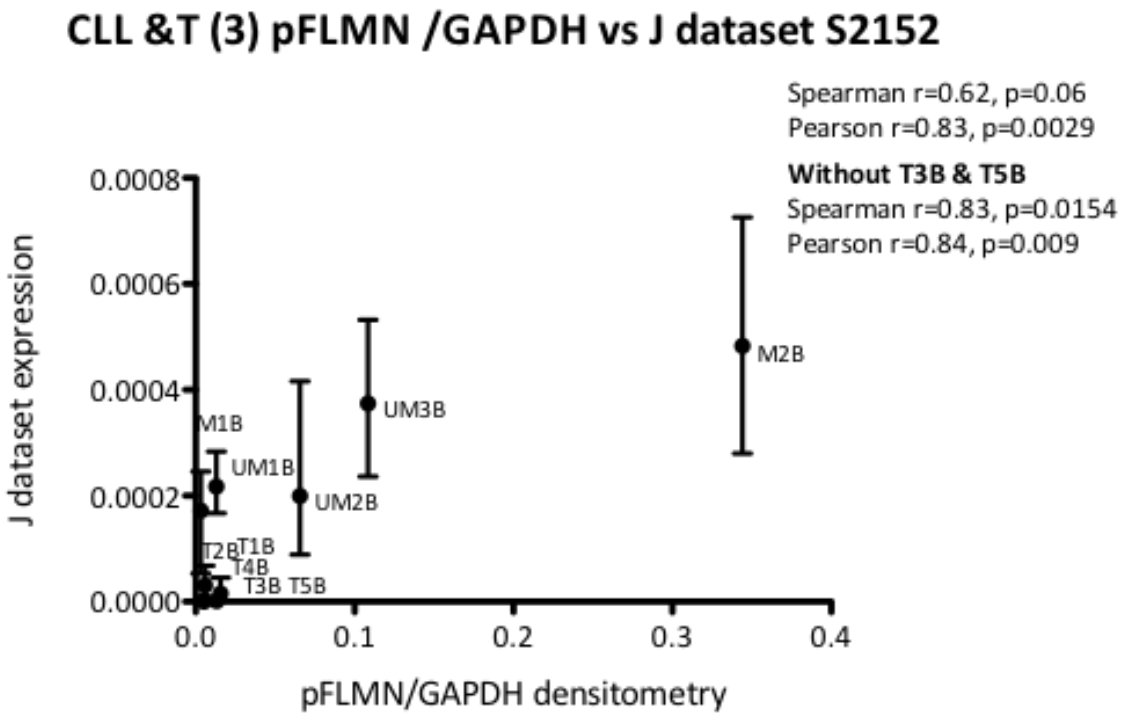
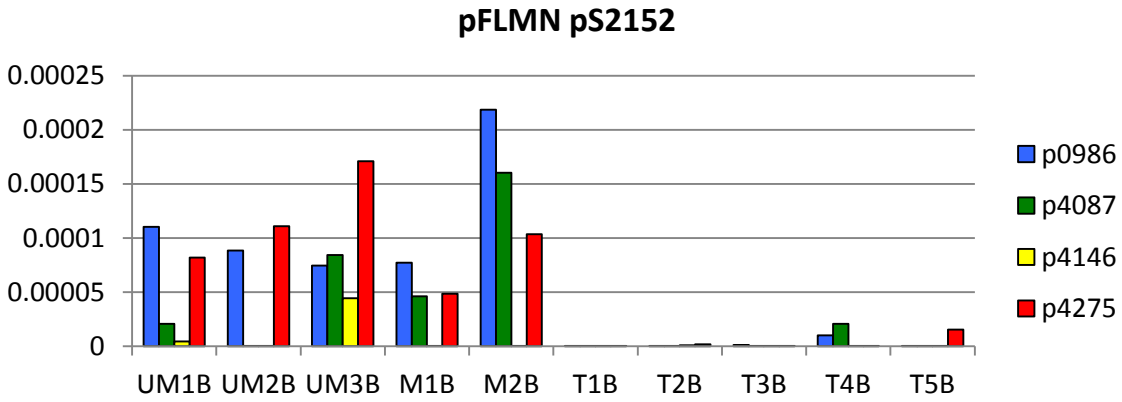


**Figure 7.67: Expression of 4 features in J dataset representing FLMN pS2152.**
**Shown are the expression levels of p0986, p4087, p4146, p4275. Below are shown the correlation coefficients for the comparisons with densitometry derived by western blot. The J dataset represents the sum of all 4 features.**
**Because of the poor loading of T3B and T5B, correlations are also shown when these samples are removed from the analysis**




| | p0986 | p4087 | p4146 | p4275 | J dataset |
|---|---|---|---|---|---|
| Pearson r | 0.8401 | 0.9095 | 0.1500 | 0.5524 | 0.8314 |
| P value (two-tailed) | **0.0023** | **0.0003** | 0.6792 | 0.0977 | **0.0029** |
| Spearman r | 0.4356 | 0.3039 | 0.1268 | 0.6319 | 0.6242 |
| P value (two-tailed) | 0.2044 | 0.3869 | 0.7330 | 0.0544 | 0.0603 |

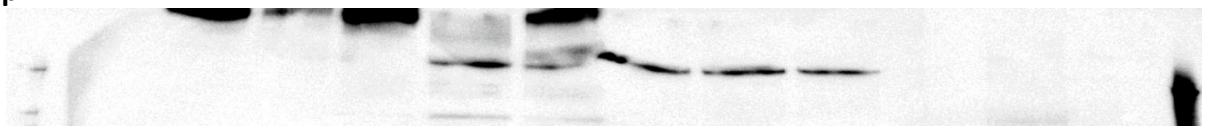
| without T3B and T5B | p0986 | p4087 | p4146 | p4275 | J dataset |
|---|---|---|---|---|---|
| Pearson r | 0.8466 | 0.9098 | 0.1140 | 0.5214 | 0.8382 |
| P value (two-tailed) | **0.0080** | **0.0017** | 0.7881 | 0.1851 | **0.0093** |
| Spearman r | 0.6108 | 0.4880 | 0.2728 | 0.7425 | 0.8333 |
| P value (two-tailed) | 0.1150 | 0.2162 | 0.5008 | **0.0458** | **0.0154** |

To demonstrate the differences caused by differing transfer times, further comparisons were made with the blots transferred at times not optimised for the phosphoproteins being probed (i.e. probing for pRS6 of blots optimised for FLMN transfer, and *vice versa*). These are shown in Figures 7.68-7.70

**Figure 7.68 Blot 1a (optimised transfer for RS6): *anti-pFLMN***
**The table shows the distribution of samples.**
**Empty box=empty well. V=pervanadate-treated cell line, C=untreated cell line.**
**Below is shown the probes for pFLMN, FLMN and GAPDH.**
**Below this, the figure shows the correlation between the means for each J dataset triplicate and the pFLMN/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

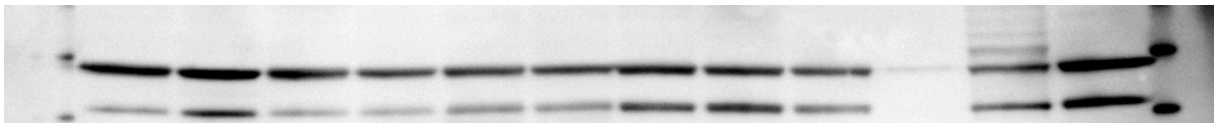
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1 B | UM1 M | UM1 D | UM2 B | UM2 M | UM2 D | UM3 B | UM3 M | UM3 D | | V | C |

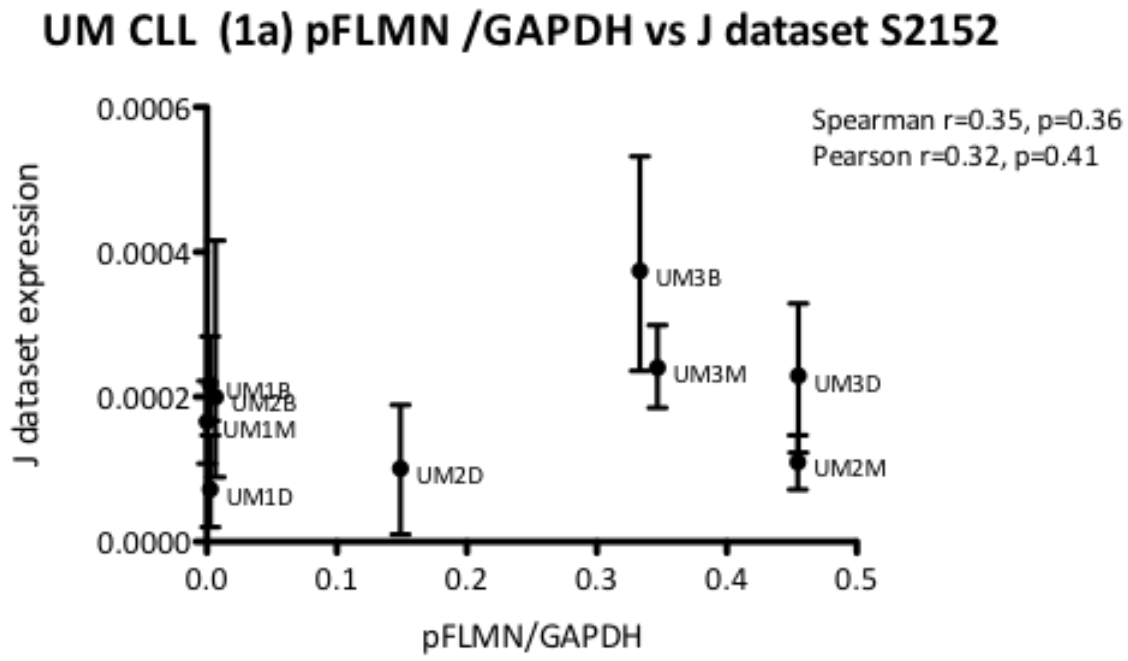
**pFLMN**



**total FLMN**

**GAPDH**



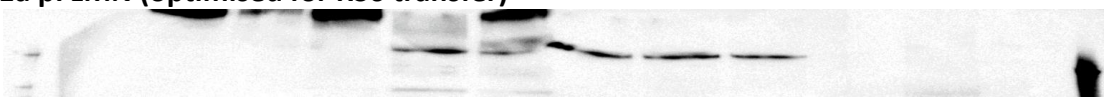



The comparisons between the two transfer times show poor correlation between them (figure 7.69):

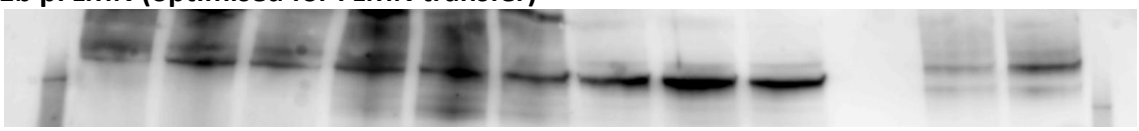
**Figure 7.69: Comparison of blots produced by different transfer time techniques, 1a vs 1b**

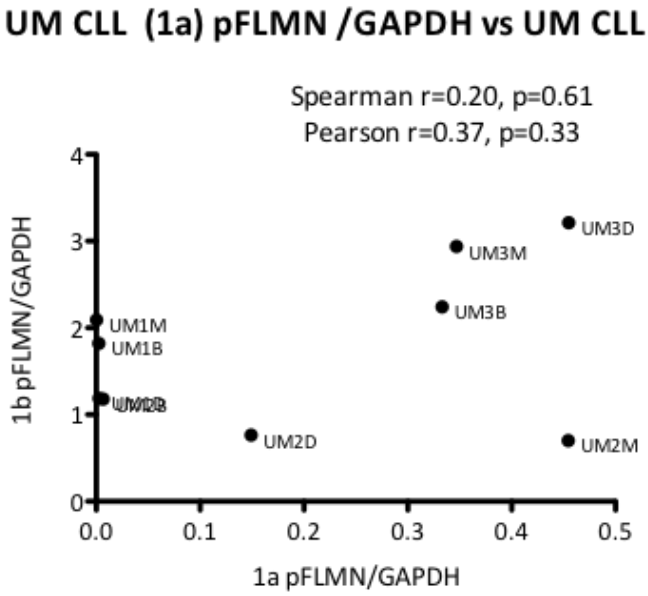
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1 B | UM1 M | UM1 D | UM2 B | UM2 M | UM2 D | UM3 B | UM3 M | UM3 D | | V | C |

**1a pFLMN (optimised for RS6 transfer)**



**1b pFLMN (optimised for FLMN transfer)**



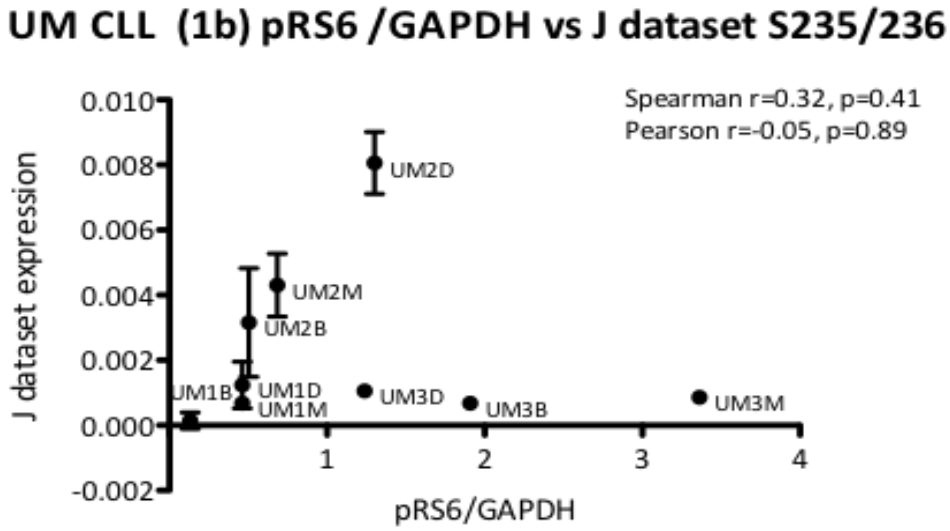



The same comparison was made with RS6. There was again poor correlation of pRS6 with the proteomics data, though there was a degree of correlation between the pRS6 expressions when the blots were compared (figure 7.70)

**7.70 Blot 1b (optimised transfer for FLMN): *anti-pRS6***

**The figure shows the correlation between the means for each J dataset triplicate and the pRS6/GAPDH ratio determined by densitometry. The error bars show the range of each replicate in the proteomic data.**

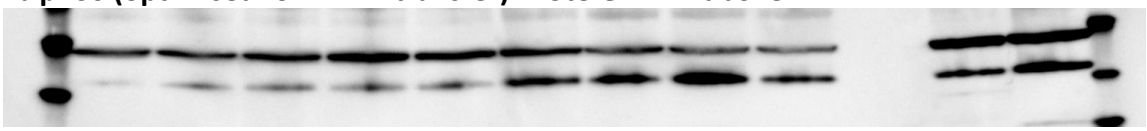
**Below is the comparison of blots produced by different transfer time techniques, 1a vs 1b**

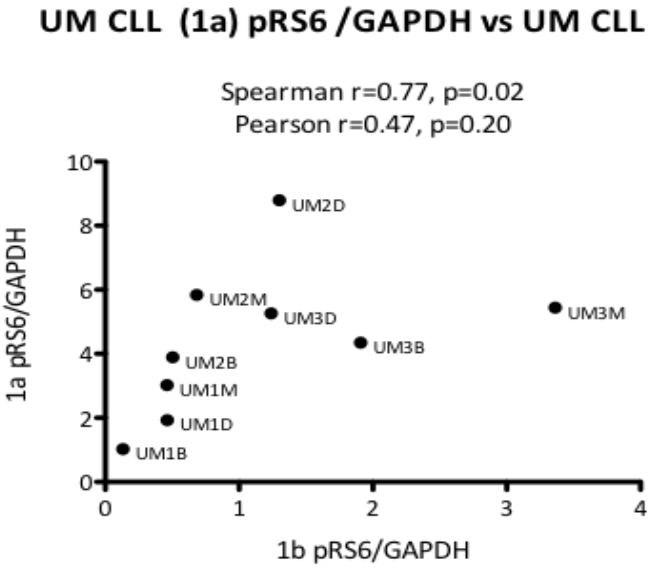



| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UM1 B | UM1 M | UM1 D | UM2 B | UM2 M | UM2 D | UM3 B | UM3 M | UM3 D | | V | C |

**1a pRS6 (optimised for RS6 transfer)**

**1b pRS6 (optimised for FLMN transfer). Note GAPDH above.**





To summarise, a limited number of phosphosites of interest had available antibodies, and only 2 of 7 of the purchased antibodies appeared to detect the relevant phosphoproteins. For these 2 antibodies, there was a degree of correlation between densitometry and proteomic expression values, but this was highly dependent on the quality of the western blot and the precise transfer method used. Table 7.24 summarises these data.

**Table 7.24 Summary of correlation coefficients and related p values for each comparison.**
**The western blot densitometry was correlated to the J dataset proteomic expression values.**
**Shaded= optimised transfer for that protein. Bold=p<0.05**

| | Western blot | 1a | 1b | 2a | 2b | 3 | 3 (without outliers) |
|---|---|---|---|---|---|---|---|
| **pRS6** | **Spearman r** | 0.67 | 0.32 | **0.78** | 0.35 | 0.59 | **0.74** |
| | **p value** | 0.059 | 0.41 | **0.017** | 0.36 | 0.088 | **0.0458** |
| | **Pearson r** | **0.79** | -0.05 | **0.86** | 0.51 | 0.29 | 0.44 |
| | **p value** | **0.012** | 0.89 | **0.0027** | 0.16 | 0.41 | 0.28 |
| **pFLMN** | **Spearman r** | 0.35 | **0.75** | 0.33 | **0.77** | 0.62 | **0.83** |
| | **p value** | 0.36 | **0.026** | 0.39 | **0.021** | 0.06 | **0.0029** |
| | **Pearson r** | 0.32 | 0.65 | 0.50 | **0.67** | **0.083** | **0.84** |
| | **p value** | 0.41 | 0.058 | 0.18 | **0.046** | **0.0029** | **0.009** |

Therefore, the blots with transfer times optimised for the phosphoprotein being probed for have reasonable correlations with the proteomics data. If one considers Spearman correlations, then 2 of 3 (excluding blot 3 outliers) have significant correlations for pRS6, with the non-significant p value being 0.059. For pFLMN, 3 of 3 blots show a significant Spearman correlation. Some of the data suggest a better correlation with the individual P dataset features, some with the J dataset sum of features. In general, I have taken these data as providing a reasonable but highly limited validation of the proteomic data.

### *7.5.4 Much of the variation in phosphoprotein expression is secondary to variation in total protein expression*

From inspection of the western blots, it is clear that high phosphoprotein expression is often related to a high total protein expression. This was formally assed by correlation of the densitometry values of phosphoprotein to the total protein (Table 7.25). Many of these correlations were significant, indicating that in most cases, the level of total protein largely determines the level of phosphoprotein.

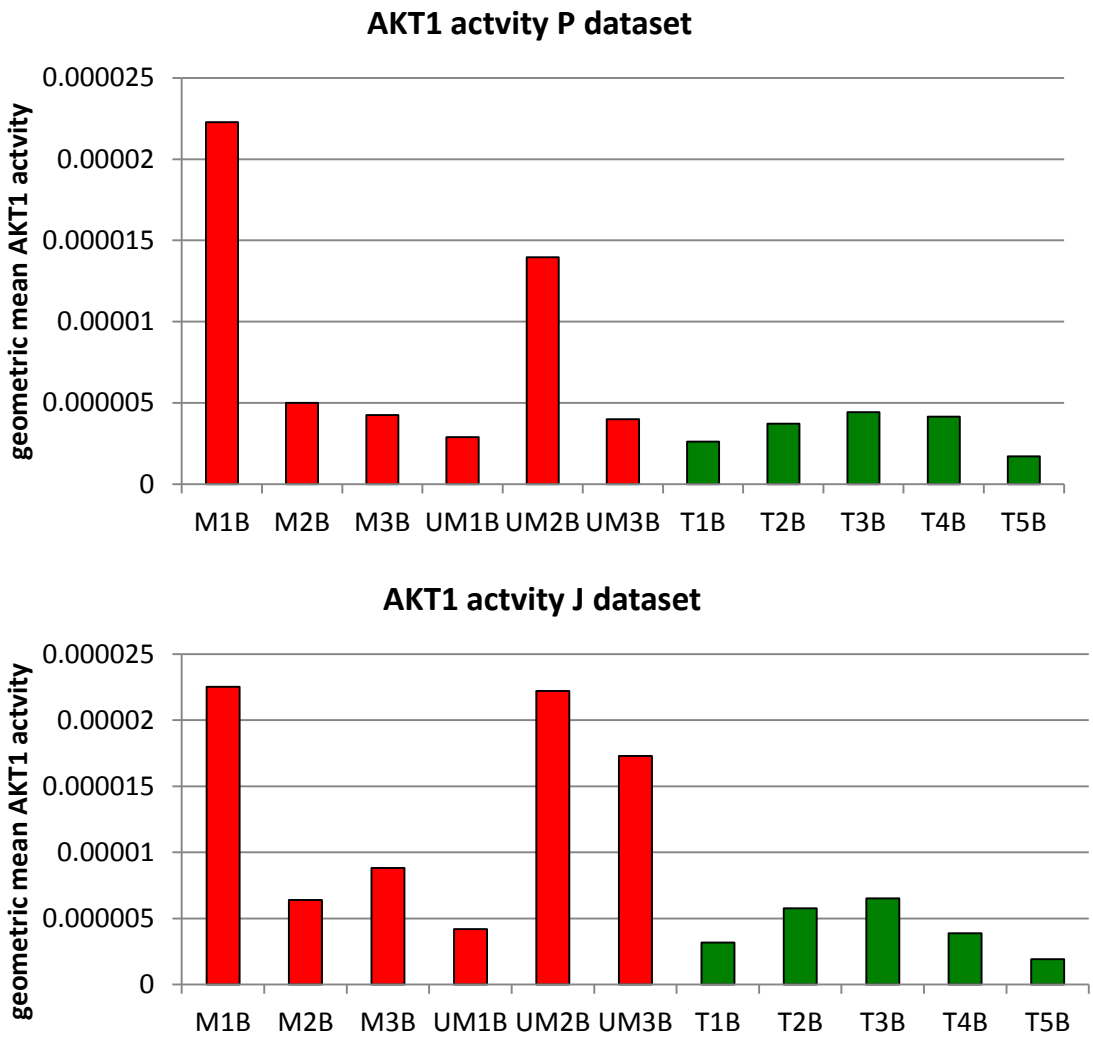
**Table 7.25 Summary of correlation coefficients and related p values for each comparison phosphoprotein *vs* total protein.**
**Shaded= optimised transfer for that protein. Bold=p<0.05. X= total not probed**

| | Western blot | 1a | 1b | 2a | 2b | 3 |
|---|---|---|---|---|---|---|
| **pRS6** | **Spearman r** | 0.62 | X | **0.78** | X | **0.96** |
| | **p value** | 0.086 | X | **0.017** | X | **0.000** |
| | **Pearson r** | **0.73** | X | **0.81** | X | **0.96** |
| | **p value** | **0.027** | X | **0.009** | X | **0.000** |
| **pFLMN** | **Spearman r** | 0.6 | **0.73** | 0.57 | **0.84** | **0.9** |
| | **p value** | 0.1 | **0.03** | 0.12 | **0.0046** | **0.001** |
| | **Pearson r** | 0.59 | **0.8** | 0.62 | **0.83** | 0.48 |
| | **p value** | 0.1 | **0.009** | 0.08 | **0.008** | 0.16 |

### 7.6 Correlation of AKT kinase activity with sensitivity to PI3K inhibition

The activity of PI3K was inferred from the expression of AKT substrates in the samples. 17 AKT substrates were identified in the P dataset, equivalent to 8 unique phosphosites. The geometric mean of all of these was taken to infer the AKT activity (see figure 7.71). Because of the differing methods for calculating J and P dataset AKT1 activity, some differences in predicted kinase activity are seen.

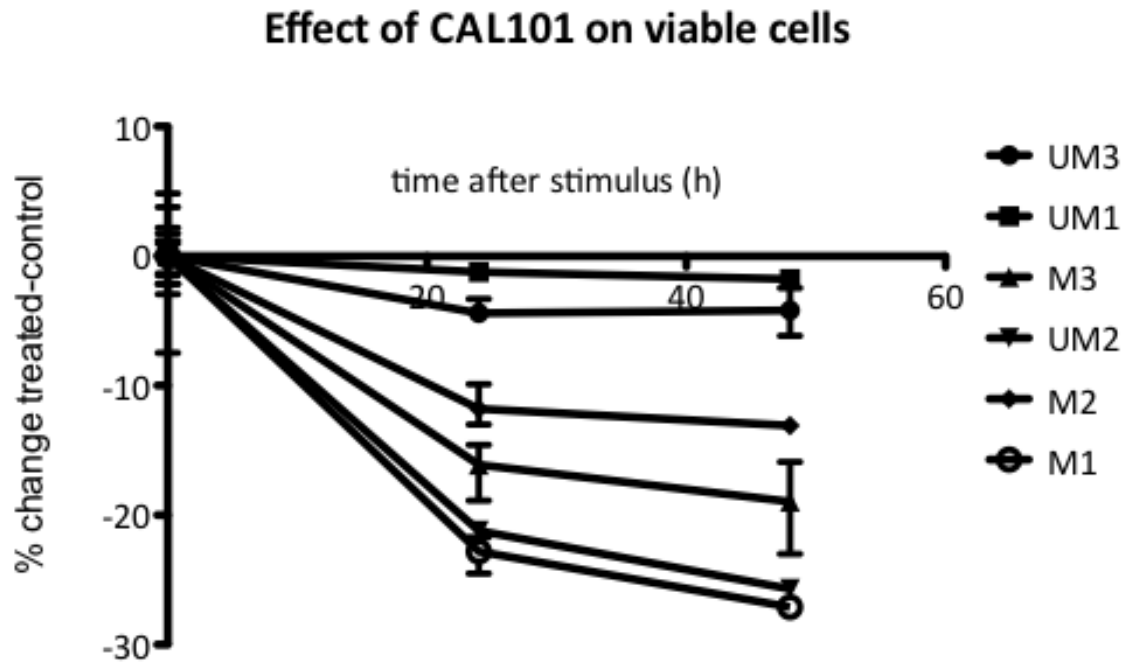
**Figure 7.71: Predicted AKT1 activity in each baseline sample.**
**The geometric mean of each substrate (n=17) expression value was combined to yield a prediction of the kinase activity in each sample. The same is shown below for the combined (n=8) J dataset. See Section 7.3.11 for details.**




CAL101/GS1101 is a PI3K delta-selective inhibitor that shows clinical activity in patients with CLL and is known to induce apoptosis in CLL cells *in vitro*[358]. The effect of CAL101/GS1101 at a concentration of 10 μM (similar to published data, and a

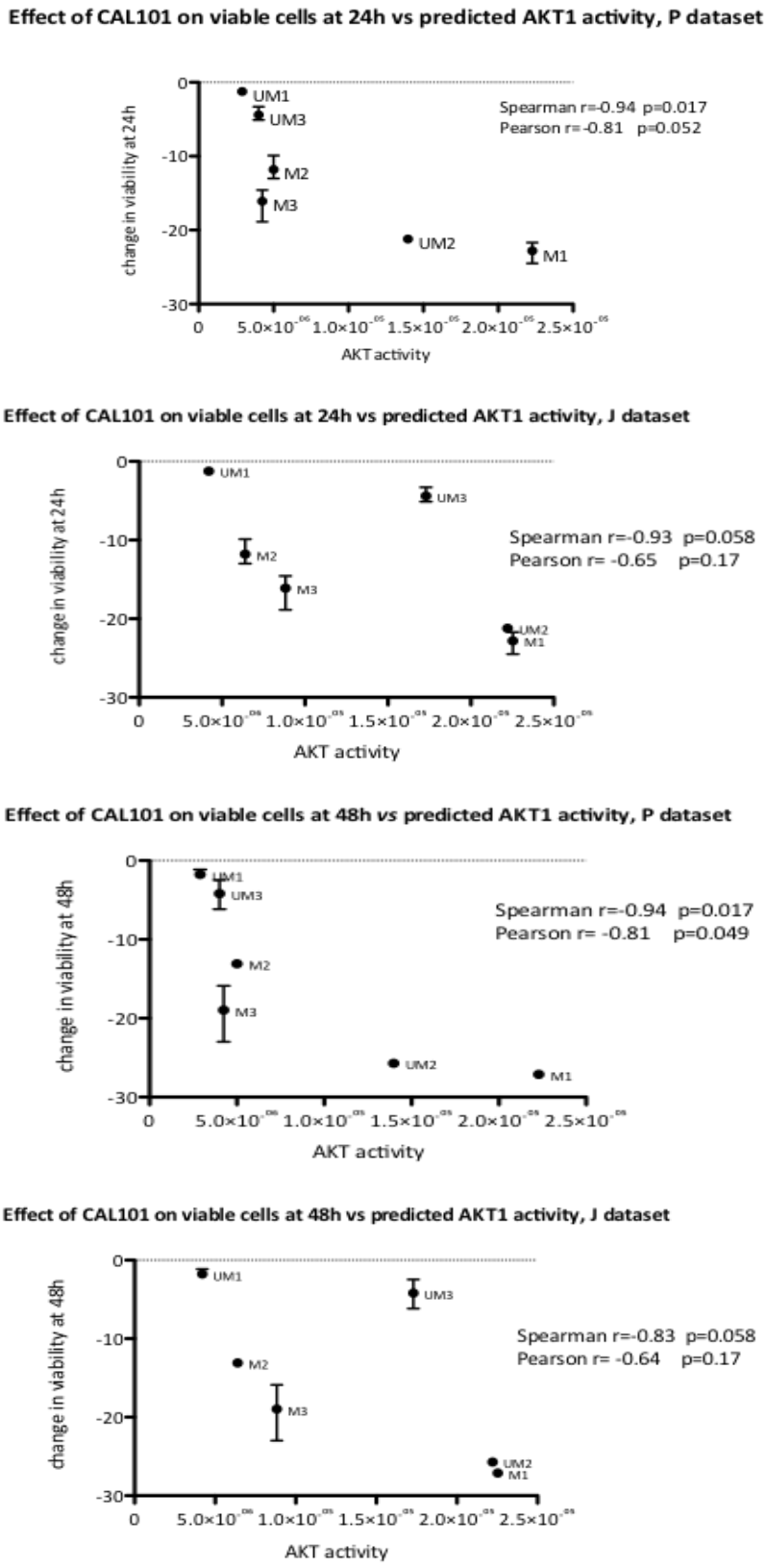
concentration that is achievable in patients) was added to CLL cells (in triplicate), and the viable cells at 24 and 48 h were evaluated (fig 7.72):

**Figure 7.72: Effect of CAL101/GS101 on CLL cell viability at 24 and 48 h.**
**The percentage of Annexin V/propidium iodide negative (viable) cells is subtracted from that with vehicle control. Error bars indicate range of triplicates.**




The degree of cell death at 24 and 48 hours was correlated with predicted AKT1 activity (figure 7.73):

**Figure 7.73: Correlation of predicted AKT1 activity with CAL101/GS101-induced apoptosis at 24 and 48 h.**




Unfortunately, probing for phosphoAkt S743 and total Akt by western blot detected no bands.

## 7.7 Discussion

### *7.7.1 Validation methods*

Phosphoproteomics is a relatively recent innovation and all techniques, particularly novel ones, demand validation by complementary methods. However, there are several challenges to the use of western blotting as a method of protein quantification, even with the most skilled practitioners. One argument is that western blot detection is a qualitative method and that visualising the presence or absence of a protein is the only appropriate use of the technique, however many publications routinely suggest differences in expression based on the size and density of protein bands by visual inspection alone. A natural extension of the subjective judgement of band intensity is the use of image software to attempt quantification by densitometry[444]. Densitometry involves image acquisition, selection of bands and computer-aided determination of the optical density (OD). When densitometry is used to determine protein quantification and correlated with complementary methods (such as ELISA), differing techniques can lead to widely varying statistical measures of correlation. When performed using certain guidelines, densitometry can be reasonably correlated with known protein concentrations[444]. There are a few general principles that are recommended: Image acquisition should be with charge-coupled camera devices rather than office scanners, and background correction should be applied. Intuition may suggest that the entire band should be included, but studies suggest that one third the diameter of the lane width provides a better correlation[444].

Although Densitometry is widely used, most publications are not explicit about the techniques or statistical analyses performed. One recent review found that only 1 in 100 publications using densitometry analysis provided sufficient information to adequately assess the validity of the methods used[444]. Beyond this, the inherent user-dependent variability in western blotting produces further challenges to the use of the method, with poor quality blots resulting from unequal loading and other technical issues. It is not clear which correlation statistic is ideal for correlation with proteomic data. Spearman rank correlation is probably the more appropriate as the data are not expected to be normally distributed, but the continuous quantitative nature of the

data is partially lost by this method, reducing the power, so some value might be expected to be gained from consideration of the Pearson r. The use of small numbers of samples per western blot also reduces the power of these statistical calculations. Combination of data from different gels is fraught with difficulties, and was considered inappropriate. A further challenge was provided by the fact that western blotting was performed on one sample, but the comparison would be with the proteomics triplicates. If more cells had been available, a more robust analysis might have considered triplicates for western blotting. With these caveats in mind, and in the absence of an obvious alternative I chose to validate the proteomic quantitative data using western blotting and densitometry.

Western blotting for phosphoproteins rather than total protein adds further challenges. Phosphoproteins are present at lower levels than the equivalent total protein, therefore ideally western blots should use as much protein as possible to maximise sensitivity (ideally 40-60 μg per well). In our experiments, the pellets produced in parallel to the proteomics that were used for western blotting only yielded sufficient protein to produce wells loaded with 20-30 μg of protein, particularly in the tonsil samples. This may explain the lack of detection of bands when probing for pBtk, pSyk, pMYL9 and pMYH9. Successful probing was apparent with pFLMN and pRS6, phosphoproteins at high abundance within our samples, as judged by their relative expression in the proteomics data. The use of a convenient yet possibly relatively insensitive method of transfer (semi-dry iBlot) may have amplified these problems.

Furthermore, it is recognised that many commercially available antibodies may not be as validated as is often claimed by manufacturers. Antibodies supplied may not bind the stated protein at all or may bind to proteins at completely different molecular weights. For instance, our anti-pLSP antibody clearly highlighted two bands at a completely different molecular weight than predicted, and the pWASP antibody strongly bound numerous proteins in pervanadate stimulated cell lines, indicating that it is a non-specific phosphotyrosine antibody. Non-specificity is a greater problem with phosphoprotein antibodies. Even with antibodies that have previously been found to

work well, a different batch from the same manufacturer may fail. Beyond this, differential glycosylation, the use of differing blocking reagents and many other factors may lead to inadequate optimisation.

Two antibodies seemed to work reasonably well: Anti-pFLMN and anti-pRS6. The high molecular weight (280kDa) FLMN may be less reliable because of issues with poor migration during electrophoresis, and poor transfer. Transfer time was optimised to facilitate FLMN transfer, but this possibly impacted on the transfer of lower molecular weight proteins such as RS6, or indeed GAPDH. It is difficult to estimate the variability introduced by these factors, but it was observed that greater correlations with the proteomic data were observed with the optimal transfer method. Of course, this argument risks circularity, as validity of western blotting is then confirmed in reference to the proteomic data.

Ideally, further validation of the proteomic dataset would be desirable. One way to improve sensitivity might be to use higher cell numbers to yield more protein for western blotting, and to use possibly more sensitive transfer methods. The use of replicate experiments and western blotting might also lend greater statistical reliability, though large numbers of primary cells would be necessary. Further optimisation of primary antibodies could then be considered. Even then, the reliability of commercially available antibodies must be questioned. The greatest challenge to the use of western blotting as a validation method is the lack of availability of primary antibodies for the vast majority of identified phosphosites. The number of discovered phosphosites continues to rise, and studies using phosphoproteomics are increasing in power. The availability of truly validated antibodies for use in western blotting cannot keep pace with advances in phosphoproteomics.

A label-free technique for proteomic quantifications was used in this study. The traditional approach to quantification in mass-spectrometry-based proteomics involves the use of labelling, but these approaches are very difficult when primary cells rather than cell lines are used. Label-free proteomics is gaining acceptance as a valid, accurate and reliable method of quantification, with the increasing sensitivity of

techniques allowing confident quantifications when multiple features are detected[439,440]. The techniques for analysing phosphoproteomic data are still in development, and to an extent must be modified according to the experimental design.

Somewhat ironically, a well-recognised method for identifying bands detected by protein electrophoresis is the use of mass-spectrometry-based proteomics. Discoveries in small-scale experiments are often assumed to be of better quality than those of high throughput experiments, but many have argued that the opposite is true: high-throughput experiments require extensive standardization and calibration, while a small-scale experiment is performed differently each time[430,553,559,560]. Furthermore, whilst small-scale experiments focus on what the experimenters consider interesting, a claim made for high-throughput methods is that they provide a view of the proteome and its interactions without presumptions as to what alterations are significant. Phosphoproteomics by MS may represent an unbiased approach capable of monitoring cellular phosphorylation events in the absence of *a priori* knowledge. This might lead to a greater confidence in the interaction and signalling networks produced by high-throughput techniques as compared to a consideration of the curated literature.

### *7.7.2 Noise in data and analysis methods*

The data produced by high throughput methods of mass spectrometry has a resemblance to gene expression data in terms of scale. The use of analysis methods well established in gene expression studies therefore helped with many aspects of this work. However, phosphoproteomic data is not gene expression data, and many analysis methods had to be tailored to the experiments performed and the nature of the data. One weakness of my analysis is a reliance on the collaboration with colleagues in terms of running the HPLC and mass spectrometer, as well as the interpretation of the raw mass spectrometry data to produce phosphopeptide identifications and quantifications. Our collaborators' extensive experience and publication record in this area nonetheless are reassuring[439,440,561,562], and collaboration is essential for the use of these kinds of novel techniques.

The remarkable increase in the power and sensitivity of proteomics techniques has led to an interesting problem in this analysis. Each phosphoprotein produces a number of different fragments during digestion and processing that can be identified as features by mass spectrometry. Some proteins produced >10 features per phosphosite. Some of these will be misidentifications, but it would be expected that most are true identifications. Even when peptide fragments are linked to their parent proteins with great confidence, there is a phenomenon whereby some fragment abundances appear to correlate with the parent protein abundance, whereas other fragment quantifications do not[440]. These 'non-proteotypic' peptides whose quantifications do not appear to reflect parent protein abundance are a problematic source of error, and ideally should be excluded from analyses. However, there is no *a priori* means of identifying non-proteotypic peptides apart from the basis of prior experience. Methods are being developed to deal with this issue, but cannot be applied to this study. Despite this, our data suggests that most fragments/features derived from a parent protein do correlate with each other within a sample, indicating that non-proteotypic peptides are infrequent and at generally lower abundance, at least for the two most abundant phosphoproteins.

To deal with this issue when performing comparisons between samples, two analysis datasets were used: the P dataset using individual features, and the J dataset using the sum of all features identifying a particular phosphosite. It was hoped that the J dataset would compensate for some of the issues regarding non-proteotypic peptides. As might be expected, these two different analysis methods produced overlapping, but not identical results.

The limited validation using western blotting did not clearly suggest that the use of the P dataset was superior to the J dataset or *vice versa*. The western blotting data are also consistent with the assumption that a minority of the features are non-proteotypic. In some cases, the more abundant feature had a greater correlation with densitometric plots. This might imply that consideration of individual features, as in the P dataset, would be most appropriate. However, this was not always true, and in some cases,

greater measures of correlation were produced for the J dataset. A theoretical argument in support of use of the J dataset is that it will compensate for the presence of non-proteotypic or misidentified features to an extent. The limited amount of validation does not permit a confident decision on which dataset to employ, and consideration of both analyses may be prudent. Caution in interpretation is clearly necessary, and highlights the requirement for better validation methods.

There is a large amount of 'noise' in the data when replicates are considered. The mean coefficient of variation between replicates was 52-53%; lower values would be desirable. An alternative measure of replicate reliability is the Pearson r correlation coefficient, which was 0.82 for the P dataset and 0.91 for the J dataset. Even so, this contains some replicates with a Pearson r as low as 0.4. To an extent, zero normalisation can compensate for some of his noise, but the variability must be taken into consideration. The clustering methods used differentiated between CLL and tonsil reasonably well, and between each patient, but again, there was much inter-replicate variability. Some analysis approaches exclude obvious outliers in order to reduce these measures of noise[440], but this may risk increasing the rate of inappropriate rejection of the null hypothesis, leading to false positives.

Some of this noise will be due to technical variability. In high-throughput mass spectrometry workflows, phosphorylation sites may be lost at several stages: some phosphoproteins are relatively difficult to purify, kinase and phosphatase activity may still be ongoing to different degrees in the lysates and enrichment for phosphopeptides favours certain amino acid compositions. In addition, even the most advanced LC-MS techniques do not reveal all phosphosites in a single biological sample: repeating the LC-MS run may reveal up to 50% more phosphosites when a duplicate sample is analysed[563]. Peaks in mass spectrometric data are selected stochastically based on peak size, and undersampling of complex mixtures results in problems with absent quantifications of peptides. Missing peptide quantifications rather than inappropriate misidentifications therefore often cause high inter-replicate variation.

Biological variation will also account for some of these differences. Each sample was rested for 2-3 hours and stimulated separately, and so may be considered biological replicates (same source), though for purpose of analysis their mean values are often used as simple technical replicates. One of the consequences of interpretation of high-throughput data is a greater appreciation of the true level of biological 'noise' that is present in even carefully controlled conditions. Intraclonal diversity is becoming very apparent in cancer biology, and accounts for some of this variability within populations. However, even when individual cells within clonal cell lines are monitored in terms of their responses to drugs, a high level of cell to cell variability is present, demonstrating that even 'identical' cells may respond differently to the same stimulus[564]. Individual cell fate is often dependent on stochastic process that manifest as predictable changes when millions of cells are considered at the population level[565].

Different researchers when analysing gene expression data favour various methods used for gene expression analysis. Many claim that fold changes alone are more 'biologically relevant' than simple statistical significance, and in fact show more reliability and correlation with other methods. Also, statistical significance becomes a problematic measure when multiplicity of hypotheses are considered: using a p-value $<0.05$ for our 5088 phosphosites would be expected to result in 254 false positives in a random dataset, a p value$<0.0002$ might be more appropriate to rule this out. This straightforward Bonferroni correction (multiplying the p-value by 5,088) is generally considered too stringent, and so False Discovery Rates (the expected proportion of false positives in a generated list of differentially regulated phosphosites) are calculated using the Benjamini-Hochberg method[538,539]. There is no consensus for analysing these kinds of data as yet, and so the overlapping results of different methods will need to be considered.

I adopted a stringent statistical approach to analysing the differences between CLL and Tonsil phosphosites, and the results of BCR stimulation. This will compensate for the large amount of 'noise' that is exhibited by inter-replicate variability; a phosphosite emerging from this noise is likely to be truly differentially expressed. It will also result in the rejection of several interesting differentially expressed phosphosites, but these

forms of high throughput data require a pragmatic approach to filter large volumes of information. In order to reach significance, a particular phosphosite generally required a high fold change to be present in all CLL samples as compared to all Tonsil samples. A phosphosite highly upregulated in 5/6 CLL samples will therefore generally be ignored using our approach. Alternative analysis methods may be able to deal with this aspect, but the most appropriate means of increasing statistical power would be an increase in the number of samples used. Overall, the statistical methods used for analysing large datasets are filters for data to enable hypotheses to be generated and tested. The threshold of the filter is set by convention and by the requirements of the researcher.

The high variability may account for the limited number of phosphosite alterations after BCR stimulus; these were few when compared to the differences between patients. This reduced the power to detect commonalities of BCR signalling shared across patient samples. This will reflect true differences between BCR signalling in one patient as compared to another. To an extent, the lack of shared BCR-stimulated proteins between patients is not seen when BCR stimulus is considered in terms of pathways (see below). It may be that our analysis is insufficiently sensitive to detect important phosphosite changes that occur in low-abundance phosphoproteins, or at low fold changes. It might be expected that BCR stimulus would result in the upregulation of substrates of SYK, LYN and BTK. However, since these substrates are not detected, no comment can be made about changes in activity of the upstream kinase. This is probably because these substrates are present at low abundance. When the kinases that appear to be modulated by BCR signalling are considered, they are present in distal parts of the BCR signalling pathway. A well-known aspect of signalling cascades is the amplification that occurs as signalling proceeds from the surface to the interior. A few activated proximal BCR component molecules will result in many more distal component phosphorylations, which may be the reason for the detection of distal components. Techniques are constantly improving, but it is likely that there will be a 'bias' for higher abundance phosphoprotein detection. Improvements in MS instrumentation and phosphopeptide enrichments techniques will increase sensitivity in the future. This study identified over 5,000 phosphopeptides, which provides a level of detection power that is surpassed by few other studies.

In all detected phosphoproteins, there was an expected low frequency (1-2%) of phosphotyrosine-containing proteins, but this proportion did not increase when lists of phosphoproteins differing between CLL and tonsil, of after BCR stimulation were considered. This suggests that serine-threonine phosphorylations may be as important functionally as tyrosine phosphorylations, which is consistent with some published data[533]. A common view is that phosphotyrosine signalling may be more critical to physiology and disease[428], but this may simply reflect the fact that enrichment techniques have generally favoured phosphotyrosines, and so they have been studied more frequently. Focussing on phosphotyrosine signalling in biological samples would require alternative enrichment techniques to those used in this study, and increasing numbers of researchers are performing $TiO_2$-enrichment and anti-phosphotyrosine antibody based methods in parallel to achieve this. This has the disadvantage of requiring even larger amounts of starting protein. Further refinements to increase sensitivity include the use of digestion techniques other than trypsinisation to widen the spectrum of peptides produced[559].

Finally, once unselected approaches have identified changes in phosphoenriched samples, repeating mass-spectrometry runs may permit tracking the levels of particular proteins in techniques such as in selected/multiple reaction monitoring[566], in essence following up unbiased approaches with targeted ones to answer more specific questions.

### *7.7.3 Choice of 'healthy' B-cell controls*

The cellular origin of CLL is much-debated[109,110]. CLL is comprised of two different subsets with and without *IGHV* gene somatic hypermutation (SHM), so the cell of origin is less obvious than with other lymphomas. There are several themes that are prominent in this debate - the fact that gene expression phenotypes suggest a memory/antigen-experienced origin, the expression of CD5, the presence or absence of SHM and the uncertainties surrounding the precise definitions of normal human B-cell subsets. Candidate normal B-cell subsets suggested to give rise to MBL and

subsequently CLL include human B1 cells, marginal zone cells, memory cells and transitional B-cells[109,110].

Whether the identification of the normal counterpart/cell of origin will truly influence our understanding or treatment of the disease is an outstanding question. It may well be that there are several different counterparts that are 'funnelled' to the same endpoint. There is evidence that both M-CLL and U-CLL derive from antigen-experienced B-cells. Whether they derive from a single subset of B-cells (naïve, marginal zone or $CD5^+$) is less clear[109,110]. The use of transformed B-cell lines as controls has many problematic implications, and it is preferable to use healthy B-cells where possible, though cell lines are useful for certain studies and for optimisation of techniques.

Other investigators have used various approaches to healthy B-cell controls in CLL. In previous studies from our own laboratory examining gene expression changes after BCR ligation, total B-cells from healthy donor peripheral blood were negatively selected and stimulated.[423] This decision was based upon the results of the gene expression studies of CLL demonstrating that memory cells most closely resembled CLL cells[219,343]. However tonsil gene expression profiles were more similar to CLL than $CD5^+$ cord blood B-cells. In this study, tonsils were taken from children. The median age of patients with CLL is 72, and so the controls are not age-matched. Tonsils are also not necessarily 'healthy', exposed as they are to bacteria and other aspects of the outside world, and tonsil removal may be performed for recurrent infections, though generally not during an ongoing infection.

The methods of phosphoproteomics dictated a pragmatic approach to achieving sufficient numbers of healthy B-cells as controls. The experiments outlined here require more than $10^7$ cells per experimental condition. This is not possible if cells are separated into different subsets. Also, there are issues with positive selection using magnetic bead or flow cytometry-assisted cell separation, as these techniques will often require ligation of the BCR (for instance sIgD-sorted cells to obtain naïve populations), or of molecules known to be involved in BCR signalling (such as CD19).

Therefore, a negative selection method is desirable, further reducing potential cell numbers in each sorted subset.

In order to provide sufficient cells, total B-cells from tonsils were used, which provide a mixture of naïve and memory B-cells. BCR signalling in this mixture of healthy B-cells is similar to CLL in terms of calcium flux after IgD and IgM stimulation. One of the primary aims of this study was examining differences between IgD and IgM, and this can only be achieved with the use of naïve B-cells, memory B-cells would have been inappropriate for this aspect. The choice of control is most important when analyses directly compare tonsil and CLL. Alternative analysis strategies that do not consider CLL versus Tonsil comparisons may be valid methods for inferring upstream kinase activities. Since the analysis of BCR signalling compares within patients, the issues with controls are less relevant in this consideration.

In the future, comparisons of malignant cell signalling pathways to the B-cell subsets that are the putative normal counterparts will hopefully be made. Currently, the normal counterpart(s) to the CLL cell have not been identified, though there are some likely candidates. In terms of comparing IgM and IgD signalling, the tonsillar B-cells are probably the most appropriate controls.

### *7.7.4 Challenges in analysing phosphoproteomic data*

A global view of cellular phosphoproteomes is hampered by the size and complexity of the proteome: there are tens of thousands of proteins in any one cell, and hundreds of thousands of potential interactions. Several databases of protein interactions and post-translational modifications have been curated from the literature and from the results of high-throughput experiments, and are continuously updated. These include MINT, networKIN, STRING, phosphoELM, PhosphoPoint, PhosphoSitePlus and Phosida[549,550,556,567-569]. The overlap between the databases is sometimes small, partly due to the differing aims of the databases. Different experimental protocols have differing preferences for detecting particular modifications. Most of the variation is probably derived from the different model tissues and experimental conditions employed. The gaps in knowledge are highlighted by the fact that estimates of the true

size of protein interaction networks exceed the number of interactions stored in these databases[430,560].

In protein signalling networks, nodes correspond to individual post-translational modifications (PTMs), with the connecting edges indicating how nodes interact with each other. The most common and important PTMs considered are those of phosphorylation on tyrosine (Y), serine (S) and threonine (T) residues. The node will represent a quantitative variable, such as concentration of a particular protein's phosphorylation at a particular residue: a 'phosphosite'. Graphical representations of these networks do not resemble the usual linear flow diagrams hitherto employed to understand cell signalling, but complex networks, with known pathways distributed across multiple nodes, and many pathways sharing nodes[563]. One of the puzzling aspects of bioinformatic pathway analyses such as IPA is that the same molecule is often labelled as being a component of multiple different pathways. This may create difficulties in deciding which pathway is important in a particular cancer cell or stimulus, but it also perhaps highlights the futility of attempting to conceptually confine particular molecules to one or two pathways, as in truth they act as multiply connected nodes in complex signalling networks.

Consideration of computational network models of interactions has led to the finding that most interactions follow a power-law, indicating that these are 'scale free' networks[421]. This implies that a majority of proteins interact with only a few partners, whilst a minority have multiple partners. These occasional protein 'hubs' are key members of the network, with removal of these highly connected members producing a greater perturbation of the entire network topology than if one of the less-connected nodes is removed. This model is supported by the finding of a correlation between the connectivity of protein nodes and lethality when these highly connected nodes are knocked out in yeast experiments[421].

In line with this, phosphoproteins in yeast and mammals follow a power-law distribution with respect to the number of phosphorylation sites[570]. This reflects the fact that most proteins have few phosphorylation sites, whilst a few have multiple

sites. Phosphorylation/dephosphorylation has often been assumed to act as a fundamental on/off switch for protein function. Even this overly simplistic picture is complicated by the fact that most proteins in eukaryotic cells harbour multiple phosphorylation sites[571]. Because this increases the possibilities for regulating protein function considerably, multisite phosphorylation may provide a method for fine-tuning responses and enable combinatorial regulation of responses[571]. Proteins with more phosphorylation sites tend to accumulate even more sites during evolution[523]. This supports the model whereby these highly phosphorylated proteins are densely connected and important members of the signalling networks.

Comparison of phosphoproteomes from different species has revealed significant overlaps in signalling pathways, and the set of proteins with phosphorylation sites identified in different species of eukaryotes is enriched for disease-associated genes[420], implying important physiological roles. Phosphorylation sites are evolutionarily conserved, relative to their unphosphorylated counterparts, suggesting their important functional role. Whilst there is extensive species overlap in phosphosites that play important roles, consideration of all phosphosites implies that the overlap between species is small. The small overlap is not only due to real differential phosphorylation, but also in limitations of experimental techniques[563]. This makes it difficult to estimate the full size of the phosphoproteome (all phosphosites present within one tissue or species) as there is no 'gold standard' to compare against. Nevertheless, estimates can be made. There are currently approximately 25,000 human phosphosites within databases[568]. Estimates of the full phosphoproteome range up to 200,000 phosphosites within the total 1.7 million S, T and Y residues in the human proteome[420,563]. The minority of these that form the tyrosine phosphoproteome appears to be more comprehensively mapped, either because tyrosine phosphorylation is more important functionally, or because the enrichment strategies have greater efficiency for phosphotyrosine. Clearly, more phosphosites are yet to be discovered.

Although some proteins are constitutively phosphorylated or fluctuate regularly throughout the cell cycle, some are only phosphorylated under specific conditions and

in specific tissues. Therefore, comprehensive understanding of a single phosphosite requires multiple experiments with a range of different conditions[433]. Whilst the number of identified phosphosites continues to increase at an alarming rate, it is likely that most of these phosphosites will only be significant in particular tissues and experimental conditions, and that the phenotypic differences between tissues and stimuli will largely be determined by a limited subset of the total known phosphoproteome.

The generation of large amounts of data from phosphoproteomic techniques leads to difficulty with analysis and interpretation. New phosphosites are being discovered at a rate far exceeding the growth in understanding of how a phosphosite affects protein function. Whilst upregulation of expression of a particular gene might reasonably be expected to result in increase protein expression and activity in many cases, an alteration in the stoichiometry of a phosphosite will generally not translate to a simple increase or decrease in activity.

It is generally assumed that phosphorylation of a protein will alter its function in some way. Classically, phosphorylation increases or decreases kinase activity. Probably more important is phosphorylation-induced modulation of protein-protein binding to induce complex formation or alter localisation. Evolutionary pressures exert their effects on protein phosphorylation, and natural selection might be expected to eliminate non-functional phosphorylations (a waste of ATP?). Nonetheless, it has been suggested that many phosphosites may not convey a functional alteration on a protein, although absence of evidence is not absence of evidence, and it is difficult to prove that a particular phosphosite labelled as non-functional will have no effect under all conditions and tissues[572]. It must also be remembered that a decrease in a protein's level of phosphorylation may be as likely to confer functionality as an increase, in contrast to gene expression data where an increase in RNA levels is presumed to reflect an increase in the functionality of the resultant protein. In this study, no assumptions regarding a phosphosite's function have been made, decreased phosphorylation presumed to be as relevant to protein function as an increase.

### *7.7.5 Analysing phosphoproteomic data in terms of biological pathways*

High throughput techniques such as gene expression analysis and mass-spectrometry-based proteomics are producing increasing volumes of data. The main driver has been improvements in technology, but it is also driven by a realisation that biologists must embrace complexity in order to understand the behaviour of cells, tissues and organisms at the molecular level. Whilst these complex datasets may be felt to accurately reflect cellular processes, the human brain is ill suited to deriving meaning from this kind of data directly. To an extent, these advances in experimental methods have also driven advances in bioinformatic techniques to analyse these unwieldy amounts of data. To overcome this challenge, curated databases have that reflect known pathways and disease processes can be interrogated. Two widely used pathway databases are Ingenuity Pathway Analysis (IPA) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG). Despite the wide use of Ingenuity (curated by a commercial company) and KEGG (curated by a single laboratory), the databases, algorithms and output formats are surprisingly different and non-overlapping[573]. In this study they were both used in an attempt to illustrate how they may be applied to phosphoproteomic data, as well as to gain insights in B-cell biology and BCR signalling. Other groups have applied these methods to phosphoproteomic data with a degree of success[553,574]. These methods can begin to make sense of the large amounts of data generated and lend plausibility to the techniques used. However, these tools were generated for analysing expression of the 23,000 known genes in the genome, as compared to the >100,000 largely unknown phosphosites in the phosphoproteome.

IPA and KEGG Databases were used to derive pathways at three levels of interpretation. Firstly, a consideration of all detected phosphoproteins in the CLL and Tonsil samples was made. The most abundant phosphoprotein was Histone 1.2, followed by various RNA binding proteins, cytoskeletal components and transcription factors. Strikingly, the top 3 IPA pathways enriched amongst the detected proteins were ‘ERK/MAPK Signalling’, ‘DNA Methylation/Transcriptional Repression Signalling’ and ‘B-Cell Receptor Signalling’. The top three KEGG pathways were ‘Spliceosome’, **‘**B-Cell Receptor Signalling’ and ‘Ribosome’. A number of interpretations are possible. The abundance of RNA binding proteins and transcription factors within

phosphoproteomes has been noted by others[523,534], and may explain enrichment for the 'Spliceosome', 'DNA Methylation/ Transcriptional Repression Signalling' and 'Ribosome' pathways. The presence of the BCR signalling, ERK and other identified pathways would be expected in B-cell tissues. The enrichment of BCR signalling components may be taken as evidence that these proteins are highly expressed because of the tissue origin, rather than activity of the BCR pathway itself, though this is also very possible. One study of mantle cell lines found constitutive phosphorylation of multiple BCR components, and the authors interpreted this as evidence of constitutive BCR signalling[575]. Further conclusions regarding pathway activity in resting cells are difficult to make with our data. Comparison with other tissues and phosphoproteomes of B-cells from other subsets may be helpful.

At the next interpretation level, the phosphoproteins differentially expressed between CLL and Tonsil were considered. One intriguing finding was that there appeared to be more phosphoproteins identified in CLL than in tonsil (see figure 7.23). One interpretation is that CLL has higher overall phosphoprotein levels in a given amount of total protein (all samples were enriched from 0.5mg of protein) when compared to tonsil. This is not inconceivable, and one study in CLL compared total phosphotyrosine levels on western blot with tonsil B-cells, and found markedly higher levels in CLL[330]. I am unaware of similar studies in phosphoproteomics. A more plausible, yet not mutually exclusive explanation is that CLL has a greater variety of phosphorylations, at least detected using our methods, supported by the finding that more phosphosites had a zero value in tonsil when compared to CLL. This would manifest as generally higher CLL/Tonsil fold changes.

IPA analysis suggests that the phosphoproteins most upregulated in CLL cells compared to tonsil are enriched for components of tight junction signalling (unreported in literature), RAN (GTP-binding protein involved in nucleocytoplasmic transport) signalling[576], ATM (recognised association with CLL because of 11q23 deletions) signalling[30], glucose metabolism, cytoskeletal signalling and many other pathways. Abnormalities of the cytoskeleton in CLL have been recognised for some time[577], and Professor Caligaris-Cappio's group suggest that the importance of the

prognostic marker Hematopoietic cell–specific Lyn substrate-1 (HS1/HCLS1) in CLL lies in its role in regulating the cytoskeleton and cell trafficking[578]. Multiple phosphorylated forms of HS1 were detected in our data, but whilst BCR stimulation did result in altered HS1 phosphorylation, it was not differently expressed between CLL and Tonsil. An associated protein of unknown function, HCLS1-binding protein 3 was also present, and was differentially phosphorylated between CLL and Tonsil. IPA and KEGG analyses also suggested glucose metabolism pathways as important in CLL *vs* Tonsil comparisons. Abnormalities in CLL glucose metabolism and glycolysis have been reported by many studies since an original report in 1969[579], but there are few publications linking glucose metabolism to CLL in more recent years, though a recent study suggests that the pentose phosphate pathway may be important in CLL and a number of other B-cell malignancies[580], and some studies link CLL gene expression to more general abnormalities in 'cell metabolism'[581].

Aside from IPA pathway analysis, the IPA and KEGG 'biological functions' suggested enrichment for various processes such as 'Haematological Disease', 'Cancer', 'Haemopoiesis' and 'Immune system development' (see Table 7.9). Whilst not particularly illuminating in themselves, they at least suggest that the use of these bioinformatic techniques in combination with our experimental data is producing plausible results.

I engaged in exploratory analyses with closer focus on the Tight Junction Pathway, suggesting more pathway component involvement than initially highlighted. TJ pathway components were also altered by BCR stimulation in CLL. A pathway classically associated with epithelium rather than B-cells seems implausible in this context, but the TJ pathway is known to link transmembrane signalling to the cytoskeleton more generally, and the only other study examining BCR signalling in a cell line found several TJP components altered after IgM ligation[534].

Further analysis of this kind may be instructive in exploration of the pathogenesis of CLL, but there is a critical dependence on the use of Tonsil as a control comparison with these and other analyses, and caution is therefore required. It may be instructive

to examine co-expression patterns in the entire phosphoproteome in an analogous fashion, but this is beyond the scope of this study.

The principle aim of this study was an exploration of BCR signalling in CLL at the next level of interpretive analysis. Statistical comparisons using clustering techniques and paired t-tests suggested that there were no consistent phosphoprotein differences after BCR stimulation in all CLL or tonsil samples. Whilst this may seem discouraging, it is not entirely surprising. The limited sensitivity and large amounts of technical noise have already been discussed. It appears that the differences between patients are much greater than those after BCR stimulation. Each patient seemed to have a relatively unique alteration in phosphoproteome after BCR stimulation. We found that each patient had 0.1-4.6% of phosphosites altered by BCR stimulation, and when considered collectively, 19% of all unique phosphosites were altered in at least one sample after IgD or IgM stimulation. It is difficult to make valid comparisons with published data. Most studies of stimulus-induced phosphoproteome changes use quantitative labelling techniques without use of replicates and equivalent statistics, analyses most often used cell lines and often pooled lysates from different samples when primary samples were studied. Estimates of the proportion of the detected phosphoproteome that changes after stimulation varied from 3-30%, with the higher percentage representing a pooling of data from multiple cell line experiments[533,558,582-584]. Our data showing at most 3.7% of all phosphosites altered per sample, or 17% (P dataset) to 21% (J dataset) when all phosphosites are considered is generally consistent with published data.

There was a disappointing lack of overlap of BCR-induced phosphoproteins across all patients. The handful of overlaps included pro-IL16 (a cytokine released by lymphocytes that binds CD4, thereby attracting T-cells[585]), B-RAF (known to be involved in ERK signalling, BCR signalling[368], and known to be mutated in some cases of CLL[416]), DHX8 (involved in the spliceosome[585]), CREB1 (a transcription factor known to be phosphorylated after BCR ligation[586]) and CD2AP (recently suggested to involved in proximal events in BCR signalling, linking the cytoskeleton to BLNK[554]). Interestingly, CD2AP was altered by IgD signalling in 3 of the 5 tonsil patients in both the J and P

datasets, but not by IgM ligation, so may be a candidate component of an IgD-specific BCR signalling pathway. Therefore, whilst there are limited overlaps, these differentially phosphorylated proteins are all plausible candidates for involvement in the BCR signalling cascade.

The phosphoproteins altered by BCR stimulation contained a higher proportion of transcription factors (TFs) than all of the phosphoproteins detected, suggesting TFs were altered by BCR signalling, as is well-recognised[122]. The canonical IPA pathway terminates with the TFs CREB1, ATF2, JUN, BCL6, EGR1, ELK1, OCT2, ETS1, and NFAT. Of these, CREB1, JUN and NFAT were all differentially phosphorylated by BCR ligation in this study.

The pathways altered in each patient were considered in two approaches. Firstly, the lists of differentially phosphorylated from each patient were combined into lists to give a general view of IgM or IgD signalling in CLL or tonsil samples. These were then used in bioinformatic analyses. Using the results of these, the individual lists from each patient were then considered in a 'look back' manner. The value of this approach was illustrated by the findings overlapping pathways activated by BCR ligation shared across patients, even when the individual pathway components themselves were not shared across patients. For instance, the 'Phospholipase C signalling' IPA pathway (known to be a central component within the BCR signalling cascade, leading to calcium flux) was significantly enriched in the IgM-stimulated tonsil phosphoprotein list, and also in all 5 separate patient lists. Therefore, I suggest that whilst each patient may appear to undergo BCR signalling with a unique pattern of phosphosite changes, a consideration at the level of biological pathways suggests greater unity than is initially appreciated.

One group has used phosphoproteomic analysis to predict the migration behaviour of ErbB2-overexpressing mammary epithelial cells in response to ligands[587] or pharmacological inhibition of key kinases[588]. Partial least-squares regression analysis produced a quantitative combination of a limited number of phosphorylation sites (5-10) on a few key proteins (a 'network gauge', or 'rheostat') that could successfully

interpret the effects of kinase inhibitors on cell migratory behaviour, or predict response to ligand stimulation. No individual phosphorylation site could predict the behaviour alone, suggesting that models attempting to predict cell behaviour will need to include markers of multiple pathway activity, and that attempting to predict cellular behaviour on the basis of individual pathway components will be futile[411].

The IPA and KEGG pathways highlighted after BCR stimulation contain some expected pathways (such as BCR signalling, ERK/MAPK signalling, phospholipase C signalling), some plausible ones (PI3K signalling in B lymphocytes, BAFF signalling, regulation of actin motility by Rho, Systemic Lupus Erythematosus Signalling) and some surprising pathways (Tight Junction Signalling, Insulin Receptor Signalling, Spliceosomal Cycle). To an extent this can be explained by a deeper consideration of the IPA/KEGG results, and a degree of caution about over interpreting results of these analyses. Usually only a handful of proteins within a pathway are highlighted by a particular analysis, though this is often all that is required to achieve a significant enrichment to indicate the involvement of a pathway. As discussed, each protein is expected to be a component of multiple, interlinked pathways in a complex network. This conception of a particular pathway component might lead to rejection of the entire idea of discrete pathways as units of analysis. Alternatively (but not exclusively), it points to the increasing realisation that one molecule may truly be an important component of multiple pathways. For example SNW1 (SKIP) is well recognised to be involved in both the spliceosome and NOTCH signalling pathways.

This draws attention to a striking finding of the analysis: the recurrent finding of multiple components of the spliceosome pathway as differentially phosphorylated after BCR stimulus in both CLL and Tonsil samples. This may partly reflect the high proportion and abundance of RNA-binding proteins within the entire dataset, as 'Spliceosome' is also highlighted when considering all detected proteins (though not in the CLL *vs* Tonsil comparisons). The patterns of spliceosome component phosphorylation after BCR stimulus are complex, different in Tonsil and CLL, and after IgD or IgM stimulus. This creates a huge challenge in deriving a simple model for how BCR stimulation may be linked to alterations in the spliceosome complex. BCR

stimulation is known to have pleiotropic effects in all B-cells, and it is certainly plausible that it may cause change in the function of this complex, but published data has not called attention to such a link, and I suggest that further investigation of a possible relationship between BCR signalling and the spliceosome may be fruitful in both CLL and general B-cell physiology. Given the recent findings of recurrent mutations in SF3B1 in CLL[44], this may prove an interesting link.

The other striking pathway derived from this analysis is the IgD-induced change in phosphorylation of several NOTCH signalling components (as judged by KEGG pathway analysis) in CLL. Given that the other recurrent mutation recently described at high frequency in CLL is in NOTCH1[40], this again suggests an as yet unpublished link between BCR signalling and NOTCH signalling in CLL which may well be worth exploring.

Overall, many intriguing pathways were highlighted, each of which might warrant deeper consideration and validation.

### *7.7.6 Phosphatases and anergy*

In humans, the genome contains 518 protein kinases and ~140 phosphatases[431,432]. There are a number of diseases that result from mutations in particular protein kinases and phosphatases[430]. With most proteins having multiple phosphosites and most kinases and phosphatases numerous substrates, phosphorylation cannot be understood as simple pair-wise interactions in linear pathways. Phosphorylation networks are defined by the promiscuous relationships of kinases, phosphatases and substrates. Furthermore, analysis of isolated kinase-substrate relationships in a simplified *in vitro* assay fails to capture the regulatory molecules, subcellular localisation and other influences that operate *in vivo*[433]. Phosphoproteomic analysis of lysates from cells can begin to capture some of this context.

This study detected several phosphoproteins that are also kinases and phosphatases, and some are differentially phosphorylated between CLL and Tonsil, and by BCR stimulation. The phosphatase PTPN6 (SHP-1) pY536 phosphorylation was significantly

higher in CLL *vs* Tonsil samples (168-fold in J dataset, p=0.048). This phosphorylation increases phosphatase activity and is downstream of InsR, Lck & Src kinases (and possibly Abl and Lyn)[517]. SHP-1 is well-known to act as a negative regulator of BCR signalling and to be an important mediator of B-cell anergy[234], and to exhibit abnormalities in CLL[589]. Similarly, differential phosphorylation of the phosphatase regulator SET was observed after IgD stimulation in CLL, and is known to be overexpressed in CLL and lymphoma[590]. The SHIP1 (stimulated by IgD and IgM ligation) and CD45 (IgM stimulated) phosphatases were altered in the CLL samples, and these are also important regulators of BCR signalling and anergy[234,591]. Interrogating the detected proteins for other molecules involved in anergy[234], we did not detect DOK, CD72, PIR-b or CBL. STIM1 (pS257, phosphosite of unknown function) was detected but not different between CLL *vs* Tonsil or BCR stimulated conditions. The transcriptional regulator NFAT1 is critical to anergy maintenance in B-cells[234], and is known to be active in CLL cells, a finding which led to the initial analogy of CLL cells with anergic B-cells[115]. Multiple phosphorylations of NFAT1 (of unknown function and kinase) were detected in this study, several were differentially expressed between CLL and Tonsil, and were significantly altered by IgD and IgM stimulation in CLL. Collectively, these data are suggestive of involvement of the known molecular mediators of anergy in CLL, and in CLL BCR signalling. There are some differences between IgM and IgD-induced phosphatase/NFAT phosphorylation, but also overlaps. Because the function of the individual phosphosites is mostly unknown, it is difficult to claim that this is direct evidence of anergic BCR signalling in CLL, but it is certainly consistent with this model. The choice of CLL samples able to undergo IgM calcium flux may also have confused the picture, as we are examining signalling in cases where anergy might be expected to be less relevant (discussed further in Chapter 8).

Despite similar calcium flux responses to IgM and IgD in all patients, there are suggestions that greater numbers of phosphosites were altered after IgD ligation in CLL, when compared to IgM-induced phosphosites, and that the reverse is true in Tonsil. There was no correlation of the level of calcium flux and the number of phosphosites altered in each patient. If we consider the number of altered phosphosites as a measure of the level of BCR signal strength (a rather large

assumption), then it appears that IgM signalling in CLL is 'weaker' than that of IgD, again consistent with an anergic model. The converse, with IgM signalling in tonsil being 'stronger' is consistent with the classical model of IgM signalling being more important in normal physiology. This may be an over interpretation of the data.

Finally, the phosphatase PTPN22 was also differentially phosphorylated in CLL (303-fold overexpressed relative to tonsil, unknown function/kinase of phosphosite), and several recent publications have focussed on its higher expression in CLL, with one suggesting that it blocks pro-apoptotic BCR signals by selectively activating the pro-survival AKT pathway after IgM ligation[388].

### *7.7.7 Kinases*

Perhaps the greatest potential of phosphoproteomic data is in exploring the activities of kinases. Several Kinases were themselves phosphoproteins in the CLL *vs* Tonsil list (Table 7.8). Increased phosphorylation of CARD11 and LYN (both involved in BCR signalling) kinases were seen in CLL *vs* Tonsil. Several unexpected kinases such as an activating phosphorylation of the Colony-Stimulating Factor Receptor Kinase (CSF1R) were also found, and a search of the literature suggests that CSF1R has elevated expression in CLL[592]. The approach looking at 'kinase pathway activity' (Figure 7.36) also suggested high activity of the CSF1R. Several kinases involved in glucose metabolism (PFKL, PGK1, NADK) were also altered. Most of these phosphorylations have unknown effects on the kinase activity. Only one (Lyn) phosphorylation was associated with evidence of altered kinase activity as judged by its effect on downstream substrates.

There were also numerous BCR-induced kinase phosphorylations (Figure 7.47), some of which were also observed to have differing downstream effects when their kinase activities were considered (figure 7.52), for instance with MARK and PKACA, but the overlaps were limited. BRAF was induced by IgM ligation in both CLL and Tonsil, and has been identified as one of the targets of infrequent somatic mutations in CLL[416]. The same study found mutations in NEK1, which has overlapping substrates with NEK2 (upregulated by IgM and IgD stimulation in this study).

One study has used array methods to examine upregulation of gene expression of kinases in CLL[593]. Amongst many others, they found upregulated LCK, LYN, MAP3K2, MAPK8/9 and STK10 kinases, which have evidence of increased activity in CLL *vs* Tonsil, or after BCR ligation in CLL in this study. Similarly, others have found elevated CK2 [594] and Aurora Kinase[595] activity in CLL and other malignancies. The elevated expression of constitutive activation of LYN in CLL that was suggested in the CLL vs Tonsil data has been recognised for some time[330].

The kinase pathway activity method suggested that mTOR pathway activity may be elevated in CLL in this study. The known elevated mTOR activity in CLL is already leading to clinical trials of mTOR inhibitors[596]. A superficially less plausible suggestion for increased kinase activity in CLL is that of the insulin receptor (InsR) kinase, but a recent study has indeed found elevated levels of InsR and functional significance in CLL[597]. One might even speculatively link elevated InsR kinase activity to the numerous suggestions in this study of glucose metabolic abnormalities.

The 'kinase pathway activity' approach raises important points about the relationship between kinases and their substrates, and the relationship between phosphorylated proteins and the total amount of protein. Generally, a particular phosphoprotein is conceived as forming a minor component of the total protein. However, there is a clear correlation between the total amount of a protein and the amount of phosphoprotein. Several authors imply that phosphorylation only occurs on signalling proteins at low abundance, but phosphoproteomic datasets are not overrepresented in terms of low abundance proteins, and the profile of abundance distribution is similar for the proteome and the phosphoproteome[523]. One of the major determinants of the amount of a phosphoprotein is the amount of the parent protein, as is seen when the total protein levels are detected with western blotting. This makes the interpretation of phosphoprotein levels as simple readouts of kinase activity somewhat less straightforward than is often presented.

The levels of a particular phosphoprotein are a product of the abundance of the parent protein substrate and the activity of the responsible kinases (and phosphatases). If no

substrate is present then even a kinase that is highly active will be unable to manifest activity in terms of that substrate. A higher level of phosphoprotein may therefore be partly due to a higher level of the parent protein. Conversely, if the kinase is absent, then high levels of parent proteins do not translate into high levels of phosphoproteins. A highly abundant phosphoprotein is therefore not a simple measure of the kinase activity, though it is clearly related to this. A more appropriate conception might be that it reflects the activity of particular pathways that are defined by the kinase-substrate pairing. Regulation of the pathway can be via both alteration of kinase activities and substrate levels, and neoplastic changes may affect both of these pathway components. It might be expected that kinase activity modulation is a more rapid means for regulating signalling, and that changes in substrate abundance will vary over longer timescales, but protein turnover can be quite rapid, particularly in terms of degradation. These issues may be less relevant when considering the changes seen after 5 minutes of BCR stimulation, but will hugely affect considerations of differences between healthy and tonsil phosphoproteomes.

In this study, I have partially used the 'kinase pathway activity' approach to acknowledge this relationship between kinase activity and substrate levels and to attempt to move away from a dependence on comparison with tonsil controls. I attempted to derive kinase activity values less dependent on substrate levels by using the geometric mean of phosphosite expressions. This is inspired by the work of Dr Pedro Cutillas and others, who use the term 'net kinase' activity to describe the number of times a kinase has phosphorylated its substrates before measurement, as opposed to 'specific' kinase activity, which describes the inherent ability of a kinase to phosphorylate a substrate[598], which is rather more difficult to measure. I have often implied that phosphosite quantification implies kinase activity, but it is clearly not a straightforward relationship It could be argued that conceptions of kinase activity independent of substrate is relatively meaningless in some ways, that the net activity of a pathway involving a kinase is more relevant for most applications.

The net activity of a kinase can be inferred from the abundance of its phosphorylated substrates. If several phosphorylated substrates of a kinase are present and at high

abundance then we might have greater confidence regarding high kinase activity than if only one substrate is providing this readout, particular in the face of substrate and kinase promiscuity. In principle, this confidence can be translated into a p-value that relates to the number of substrates being measured. If this approach is taken, then net kinase activities could be inferred from consideration of all the known substrates of all known kinases in a particular biological sample, in a statistically robust manner. This would be a much more powerful method of assessing kinase activity from a phosphoproteome, as compared to the methods used in this study, which considered lists of differentially phosphorylated proteins generated using statistics used for analysing gene expression data. Dr Cutillas is pioneering such methods.

The greatest challenge to progress in this field is the lack of well-defined kinase-substrate relationships. Whilst >170,000 different phosphosites have been characterised, the majority of these do not have a defined upstream kinase. An estimate of the size of this challenge in 2004 was that 35% of phosphosites then known had a clearly identified kinase[599]. The same database in 2011 suggested that this proportion had decreased to 12%, due to the rapid increase in identification of novel phosphosites without an accompanying increase in understanding kinase-substrate interactions[550]. In a high proportion of cases (~60%) the upstream kinase can be predicted by considering the preference of particular kinases for serine, threonine and tyrosine residues in the context of amino acid motifs, particularly when contextual data such as cellular localisation are incorporated[600]. As more confirmatory data accumulate, this proportion is likely to rise. However, there is still great scepticism as to the accuracy of such prediction tools, and validation by other methods is still felt to be necessary. The widely used Scansite algorithm based on motifs alone cannot claim great accuracy[552], and the limitations of its use as a prediction tool are highlighted in this study. To fully unleash the potential of phosphoproteomic data, a huge collaborative effort to establish kinase-substrate relationships is required, a 'human phosphoproteome' project analogous to that which sequenced the entire human genome.

In cosmology, Dark matter is hypothesized to account for 84% of the matter in the universe. It does not absorb light or any electromagnetic radiation and so cannot be detected by telescopes. Its existence is inferred from its gravitational effects on conventional matter, radiation and the large scale structure of the universe, but the nature of dark matter is unknown[601]. Similarly, the thousands of human phosphosites are the result of kinases that we know must exist, but we do not know the details of the vast majority of kinase-phosphosite relationships, in fact our ignorance is increasing. Whilst the effects of the kinases can be inferred by alterations in thousands of phosphosites, kinases still constitute the dark matter of intracellular signalling cascades.

### *7.7.8 Translational aspects of this study*

Drugs that inhibit protein kinases or phosphatases are being developed to target cancer and other diseases. The overexpression of growth factor receptor tyrosine kinases or their mutation to constitutively active forms is important in the pathogenesis of several cancers. An example of a particularly successful targeting of a constitutively active kinase is imatinib, which is a potent inhibitor of the Abelson tyrosine kinase (abl) that is abnormal in chronic myeloid leukaemia (CML). The use of imatinib has revolutionised the treatment of this disease[434].

Growth factors such as epidermal growth factor (EGF) activate the classical mitogen-activated protein kinase (MAPK) cascade to cause cell proliferation. The uncontrolled activation of this pathway can cause cancer[437]. One of the best-characterized phosphoproteomes is that downstream of the Epidermal Growth Factor Receptor (EGFR)[435,436]. Time-course studies have detailed the multiple phosphorylations that occur after EGF ligation. Increasing levels of particular EGFR mutants shifts signalling away from ERK and STAT3 towards the PI3K pathway, and induces phosphorylation and transactivation of the hepatocyte growth factor MET. Consequently, combination of EGFR and MET inhibitors synergized to kill tumour cells in vitro, a rational combination based on analyses of phosphoproteomic signalling networks[435].

In this study, an estimation of the activity of the AKT1 kinase was derived for each CLL sample. This showed a significant correlation with the degree of apoptosis induction with PI3K inhibition. This is clearly a very limited experiment that should not be over interpreted, but it does begin to show how the technique of phosphoproteomics might be used as a means for biomarker prediction and the use of rational drug combinations in cancer therapy. This is discussed further in Chapter 8.

The particular combination of kinase and phosphatase activation determines the topology of intracellular signalling networks. The understanding of these networks in cancer cells will be necessary to understand how therapies that target tumour cells affect signalling, and how resistance to therapy arises by bypassing particular network elements. Inhibitors of kinases are promising in their efficacy and lack of toxicity, often due to their selective targeting of kinases. Phosphoproteomics will be instrumental in determining the modes of action of these kinases, as well as the off-target effects and mechanisms of resistance of these kinases. This understanding should lead to better treatments for patients.

### *7.7.9 Summary: mechanisms of IgM and IgD signalling in CLL and healthy B-cells*

I set out to determine the mechanisms of BCR signalling using a phosphoproteomic approach. It might seem straightforward to pose the questions 'how does BCR signalling differ between leukaemic and healthy B-cells?' or 'How do IgM and IgD signalling differ?' However, the nature of how we 'measure' BCR signalling depends on the tools that are available for examining events after a BCR binds a ligand, and on the models used to interpret and extend those data. BCR signalling can be defined by abrupt rises in cytosolic calcium ions, by phosphorylation of particular proteins (that have been previously shown to be recurrently phosphorylated by BCR ligation), or by alteration in B-cell fate that is an appropriate deployment of the immune system against a harmful or harmless antigen.

There are several other ways of 'measuring' BCR signalling. One argument is that the phosphoproteomic methods used in this study would suggest that any alterations in phosphorylation seen after BCR stimulus would define BCR signalling, regardless of assumptions about known pathways. However, to be valid, recurrent changes in a variety of experimental settings would be required before novel findings are accepted as defining a canonical pathway. Whilst this study provided huge amounts of data, with an accompanying large amount of analysis, it was still performed on a limited number of primary samples, and more insights will be provided by extension to more patient samples, in different conditions depending on the question to be asked.

There is an increasing realisation that each patient's cancer may be almost unique in terms of the repertoire of mutations, compounded by the issue of tumour heterogeneity and evolution. This may also be true for signalling networks defined by phosphoproteins, if our data are valid. To gain insights into the characteristics of BCR signalling, changes in phosphosites were sought that were consistent across different patient samples. Whilst BCR stimulation occasionally led a significant change in 2 or at most 3 samples, this was the limit of the consistency. It seemed BCR signalling was characterised by a near unique pattern of phosphosite changes. To an extent this is due to lack of sensitivity and noisy data.

When considered using a pathway-centric approach, greater overlaps between patients were seen. This concept is mirrored in some suggested interpretation of the results of whole genome sequencing efforts in CLL and other tumours[44]. Multiple infrequent mutations may be more usefully considered as components of a few pathways that are recurrently abnormal in a particular form of cancer. Perhaps then, recurrent changes in the same phosphoproteins might not be the unit of consideration when describing a signalling cascade. In truth, there may be multiple levels at which it is valid to describe signalling pathways.

It might be desirable to have a visual summary of a pathway that sums up my conception of BCR signalling, and to an extent I have attempted this in Figure 7.53. This is an attempt to highlight how BCR signalling varies between CLL and Tonsil, between IgM and IgD. A more faithful depiction of BCR signalling would probably look like a complex network of nodes representing proteins linked by connections representing their relationships. Each node would have several attributes describing the protein's function, which phosphosites are present, cellular localisation, and so on. Some of these nodes would also be kinases or phosphatases that alter the connections between other nodes, adding further complexity. The network may extend into three or even more dimensions. Furthermore, because of the current lack of scientific knowledge regarding the relationships between kinases and substrates and other protein-protein interactions, the majority of this network would be full of holes. It would be as hard to comprehend as it would to construct. The true value of these kinds of model will be the insights provided, the predictions that can be tested, and the translations into therapy for patients.

# 8. Discussion

## 8.1 Introduction

This thesis set out to determine the role of the IgM and IgD isotypes of the B Cell Receptor in Chronic Lymphocytic Leukaemia. CLL, like other malignancies, is dependent on the microenvironment for pro-survival signals, and many lines of evidence suggest that antigen acting via the BCR may be a critical microenvironmental ligand[49]. The restricted repertoire of *IGHV* genes, stereotyped CDRs and autoantigenic specificities all indicate that the antigen specificity of CLL B-cells has been influential at some point in transformation. These findings also highlight the degree to which 'healthy' B-cells exhibit autoreactivity, and that these autoreactive cells may play important roles in physiology[200].

A recent study has led to the suggestion that in a number of cases, the binding specificity of the CLL BCR is reactive for part of the framework region of BCRs on the same cell, that the antigen stimulating the BCR is the BCR itself[216], although, a very limited repertoire of patient samples was studied, so this finding may not extend to all cases of CLL. If, however, this model does apply to all CLL cases, it is difficult to reconcile with many other studies of the BCR in CLL. For example, other autoantigens and microbial antigens have been shown to bind CLL BCRs, and would presumably also exert an effect[185]. IgM signalling anergy is present in many CLL cases, and whilst this model may explain anergy, the reversal of anergy on *in vitro* incubation is not explained, as the antigen cannot be removed[116]. If BCR signalling occurs, as suggested, in the LN microenvironment[53], how would this fit with a model where the antigen is not localised in a tissue compartment but on the same cell? Perhaps other microenvironmental factors known to be important then come into play[22]. If BCR ligation is constantly occurring, why would *in vitro* BCR ligation have any effects in CLL in terms of apoptosis and other changes? Perhaps all studies examining in vitro BCR stimulus are using a supraphysiologic stimulus with crosslinking, and that only a limited amount of BCR stimulation is required for survival in vivo. Clearly, the exact model and stimulus conditions have a huge influence on the results of BCR stimulation in CLL[320].

Whilst several *in vitro* studies suggest the importance of BCR ligation via its IgM isotype in CLL, there has been relatively little exploration of the role of IgD.

My aims were to:

1. Establish and confirm the patterns of expression of the IgM and IgD isotypes on the surface of CLL cells in peripheral blood, bone marrow and lymph nodes.
2. Clarify the outcomes of *in vitro* BCR ligation of the IgM and IgD isotypes in CLL.
3. Explore the mechanisms of BCR signalling in CLL and tonsil B-cells using a phosphoproteomic approach, and compare IgM and IgD signalling.

**8.2 Immunoglobulin Expression in CLL**

Surface immunoglobulin expression in CLL is a pre-requisite for diagnosis, and is often described as 'dim' or 'weak', suggesting low expression compared to healthy B-cells[5]. It is not clear whether 'dim' applies to both IgM and IgD. Laboratories often only measure surface IgM, but it is well known that IgD is commonly co-expressed with IgM, recapitulating the immunophenotype of naïve B-cells[449]. When immunophenotyping methods are considered in depth, straightforward interpretations of the levels of IgD or IgM are elusive. Expression is a function of the percentage of cells expressing an isotype, and the levels per cell, even before issues regarding binding affinity and fluorescence are considered. Comparing IgM and IgD, or peripheral blood flow cytometry to lymph node immunohistochemistry has many caveats. Nonetheless, results can be interpreted if a sufficiently critical approach to the data is taken.

In peripheral blood CLL cells, IgD and IgM expression seem to be continuous variables[449]. They can both be detected on cells in the majority of patients. In this study, IgD expression was more common than IgM expression, with dual positivity forming the largest group. IgD expression was also generally higher than IgM expression. The presence or absence of IgM or IgD as defined using thresholds did not appear to correlate with prognosis, suggesting that variation of BCR surface levels in peripheral blood do not obviously influence disease biology. It may be that BCR expression alone supports survival, but that the level does not influence disease

aggressiveness. The BCR isotype expression in cases defined as $IgM^{-}IgD^{-}$ would need exploration in greater detail to clarify this. To an extent IgG expression may provide the pro-survival signals in those cases apparently $IgM^{-}IgD^{-}$. IgM and IgD levels do not correlate with each other, suggesting that they have differing modes of regulation and differing functional roles.

When the lymph node microenvironment is considered, albeit in a limited number of patients, it appears that IgD expression in LN is higher than that of IgM and expression of the two isotypes is not correlated, again suggesting differential regulation and function. Whilst IgD levels in the LN reflect those in PB, IgM levels in the two compartments differ, with apparently lower levels of IgM levels in LN than in PB in the same patient. It should be remembered also that lymph node biopsies are not typical of the general population of patients with CLL and that some patient selection may be involved in examining this group of patients in whom both tissues are available. However, it does suggest that IgM and IgD may play differing roles in the LN compartment. I suggest a model where IgD expression is generally unchanged as CLL cells traffic between the PB and LN compartment, whereas IgM expression may be downregulated in LN, in line with models of B-cell anergy. In line with other results[481], no consistent differences in BCR expression between pseudofollicular and interfollicular areas were noted, implying that CLL cells with higher proliferation do not appear to modulate their BCR expression within the LN. Perhaps factors such as T-cell help become more important.

It is easier to study cells isolated from PB, but important events in leukaemogenesis probably occur in the LN microenvironment[22,53]. Various studies have modelled the microenvironment using cytokines and stromal cells *in vitro*, and suggest that these effects are better mimics of in vivo events than culture in medium alone[22,380,602]. Perhaps immobilised BCR stimulus also simulates the mode of antigen encounter seen *in vivo* better than a soluble stimulus[320]. Whereas it is difficult to judge how one experimental condition may be 'more physiological' than another, it would seem a logical step to combine different putative microenvironmental stimuli that have been shown to have effects in isolation in an attempt to simulate the full complexity of the

microenvironment. It would remain difficult to avoid supraphysiological stimuli in such relatively crude *in vitro* models.

The recent study by Herishanu et al suggests that gene expression differences between the CLL LN and PB compartments imply BCR signalling[53]. They mimicked LN-upregulated gene expression patterns by IgM stimulus of PB cells. A similar strategy could be attempted using phosphoproteomics. Paired lymph node and peripheral blood CLL phosphoproteomes could be compared. If different, then stimuli such as BCR ligation, CXCR4 exposure or culture with stromal cells could be attempted, followed by phosphoproteome analysis. If BCR stimulus resulted in a phosphoproteome more akin to LN CLL, then this might be evidence for its importance in the microenvironment. An attempt to combine the different stimuli might result in phosphoproteomes that increasingly mimic that seen in the microenvironment, the conditions recapitulating the signalling pathways seen in the microenvironment might then be declared the 'most physiological'. These conditions could then be used to evaluate drug sensitivity. The design and performance of these experiments would be highly challenging, and it would be surprising if *in vivo* states could be precisely mimicked *in vitro*.

Regardless of whether IgM levels are truly modulated in the fashion suggested by an anergic model, my data suggest that IgD expression is more common and/or higher than IgM expression in CLL. If BCR plays a role in CLL, then it follows that IgD may play a more important role than IgM in this disease. However, the same data can be used to support an opposite argument: the downregulation of IgM in CLL and anergy may imply that signals received via the IgM isotype are more 'potent' than via IgD and so need to be reduced to prevent autoimmunity or apoptosis. This model, though plausible, has the disadvantage of greater complexity. It is likely that both IgM and IgD play a role in normal B-cell physiology, as well as in disease. If CLL is binding antigen, it will do so via both IgM and IgD. The retention of IgM and IgD in the genome since the immunoglobulin genes arose 500 million years ago is a potent argument for the importance of both isotypes[255]. It is has been suggested that one of the functions of IgD may be in the maintenance of autoreactive pool of naïve B-cells[94]. CLL would be missing a trick if it did not co-opt IgD-induced survival signals for its own purposes. In

the context of published research, this study suggests that whilst there are many functional similarities between IgM and IgD, there are also important differences. Any study claiming to investigate BCR signalling in CLL will be required to examine IgD as well as IgM signalling.

### 8.3 Outcomes of BCR ligation in vitro

Defining or measuring BCR signalling is not as straightforward as is often implied. The definitions generally encompass early events such as calcium flux and/or phosphorylation or later cellular outcomes such as cell proliferation and death[383]. Interpreting these outcomes implies a teleological view of the function of the B-cell: 'successful' BCR signalling should result in activation and proliferation such that an offending pathogen can be eliminated. However, given that most antigens encountered by B-cells turn out to be autoantigens, it can be argued that 'successful' BCR signalling should in many cases manifest as apoptosis or anergy. When *in vitro* stimulation of the BCR can result in both activation and apoptosis as outcomes of 'successful' BCR signalling, the picture can clearly become quite complex and studies of BCR signalling become difficult to compare and interpret. It is clear that other contexts such as secondary signals derived from T-cells or inflammatory mediators play an important role, as does the developmental stage of the B-cell and nature of the antigen[122].

There is evidence that antigen is critical at some stage in CLL development, but it is not clear when. It may be that CLL arises from transformation of a normal B-cell subset that is enriched for autoreactive cells. There is circumstantial evidence for ongoing antigen encounter in CLL. The phenotype and gene expression profile of CLL cells most closely resembles that of antigen-experienced memory B-cells[219]. There is some evidence for ongoing intraclonal diversification in terms of SHM and CSR, events that imply ongoing antigen selection processes involving AID[494]. Ongoing BCR stimulus in itself may be mutagenic, as AID is known to have many non-V gene targets that can result in lymphomagenesis[99]. The finding in this and other studies that BCR ligation *in vitro* can promote CLL survival suggests that antigen may be utilised by CLL cells *in vivo*.

The phenotypic and functional resemblance of CLL cells to anergic B-cells is also evidence that antigen engagement may be occurring *in vivo,* and the reversal of anergy after incubation *in vitro* is consistent with the model whereby the anergic state is dependent on continuous antigen exposure[115,116,234]. Teleologically, it might be argued that it would be in the cells' best interests to use antigen as an opportunity to trigger activation. The absence of clinically relevant autoantibody production by CLL cells might be taken as evidence that such activation does not occur. It may be that the CLL cells simply retain enough of the behaviour of normal B-cells that include anergy in the face of chronic autoantigen binding. Proliferation despite anergy may be permitted because of defects in apoptosis that are known to exist in CLL[603], or by other mechanisms and one such mechanism might be aberrant BCR signalling. The pattern of intracellular pathways activated after BCR ligation is a critical determinant of the outcome of antigen binding[122]. This study suggests that BCR signalling in CLL cells differs in multiple ways from that seen in non-malignant B-cells. It might be suggested that aberrant BCR signalling pathways are responsible for the proliferation of CLL cells. In normal physiology, B-cells encountering autoantigen in this context should undergo anergy and fail to survive and proliferate. The BCR signalling pathway may be corrupted in CLL, permitting inappropriate use of autoantigen as a pro-survival ligand.

An anergic model of CLL BCR signalling may permit further explanations and predictions. The time course of anergy induction and reversal seems to vary from a few minutes to 48h, depending on the model and antigen used[234]. In this study, all cells were 'rested' for 2-3 hours after thawing, as this was presumed to allow time for recovery of physiological functions, but some reversal may occur during this time. This study suggests that most of the reversal of anergy occurs after this, over the subsequent 24h, and Mockridge et al suggest that this can increase further over 48-72h[116]. If, as suggested, the antigen is the BCR itself, or an autoantigen expressed by CLL cells themselves, then reversal would not be expected to occur. Those advocating the model of CLL as anergic cells suggest that PB cells have left a LN microenvironment rich in antigen, and that a degree of reversal may occur in PB[116]. Some support for this is my finding of lower IgM expression in LN compared to PB. Presumably those cells exiting the LN most recently have lower levels of IgM and/or anergic signalling. Over

time, IgM expression and signalling would recover and as the cell re-enters the LN it would again derive survival signals. This links anergy to a model involving cell trafficking, and it would be interesting to examine co-expression of IgM and molecules involved in trafficking, or use *in vivo* labelling techniques to examine cell trafficking.

In the phosphoproteomic explorations of BCR signalling mechanisms, CLL cells that had capacity for IgM calcium flux were selected to ensure the comparability of IgM and IgD signalling. This ignores any signalling that may occur via IgM in the absence of Ca flux, which may be very relevant, but difficult to study. It may be that the conditions used measure a partially anergic IgM pathway. This hypothesis is consistent with the finding of numerous molecules involved in B-cell anergy in the CLL subset, providing further evidence for an anergic model of CLL. But this raises further questions regarding the definition of anergy. Can anergic signalling occur in the presence of apparently robust calcium flux? Are there degrees of anergy? What are the apoptotic responses of CLL cells that have undergone reversal of anergy? Are the high levels of phosphatase activity characterising CLL BCR signalling reversible in the same way as calcium flux, or are they fixed attributes of transformation? Further studies of the dynamics of anergy will be required to answer these questions. These may help answer certain questions raised regarding the recent *Nature* paper describing the BCR itself as the putative antigen in CLL[216]. Because antigen ligation of the BCR is only one amongst many survival signals derived from the microenvironment, many other factors will also need to be considered.

### 8.4 Phosphoproteomics and signalling networks

*"... the totality is not, as it were, a mere heap, but the whole is something besides the parts ..."*[604]

This is the first time that BCR signalling has been reported in primary B-cells using a phosphoproteomic approach. In fact, phosphoproteomics has seldom been used on primary tumour samples, presumably due to difficulties with cell number and labelling. The number of phosphopeptides detected is amongst the highest using these kinds of techniques. Despite the increasing power of MS-based phosphoproteomics, increased

sensitivity is desirable. Proximal events in signalling pathways occur at low stoichiometry, and low abundance phosphoproteins will not be detected. There are also issues with large amounts of noise and fundamental issues regarding quantification of non-proteotypic peptides.

I have suggested that multiple kinases may be more active in CLL compared to healthy B-cells, and that the kinases active after BCR ligation also differ from that seen in healthy cells. The frustrating inability to link kinases to most of their substrates limits the power of this study to identify the full extent of signalling networks. Further increase in the power of phosphoproteomic approaches is anticipated with novel analysis methods, but these will still be limited by ignorance of the substrate specificities of most kinases. To an extent this ignorance affects any study of intracellular signalling pathways, the high-throughput techniques simply make the scale of our ignorance more apparent. Extensive efforts to better characterise the human phosphoproteome are needed.

Recent genomic studies of CLL have discovered numerous but infrequent somatic mutations[40,44]. In searching for unifying principles, some have suggested that a pathway-centric approach can be useful, whereby multiple seemingly disparate mutations may perturb a common pathway[48]. To an extent, an analogous approach has proved useful in this study; seemingly disparate changes in phosphoproteins after BCR stimulation can be seen to lie in the same pathway. Often these pathways are what one might expect, most obviously the BCR pathway itself. The purpose of these studies is to generate novel hypotheses regarding CLL pathophysiology, and the suggestions of links between BCR signalling and the spliceosome or Notch signalling pathways are intriguing and require verification.

Attempts to describe intracellular signalling better typically result in complex networks of proteins that are not confined to any one classically conceived pathway[605]. The teleological purpose of BCR signalling is to alter the behaviour of the B-cell, and this will result in the BCR pathway feeding into the diverse pathways involved in metabolism, cell cycling and antibody synthesis. Beyond this network formed of linking

simpler linear pathways, intracellular signalling is probably better conceived as being distributed across multiple nodes[606]. The cell could best be described as analysing inputs in a massively parallel fashion, integrating signals and producing multiple outputs, much as a computer. The most important word in the intracellular signalling dictionary is phosphorylation.

It has been repeatedly suggested that phosphoproteomics may provide an unbiased approach to the elucidation of intracellular signalling pathways in the absence of *a priori* assumptions about which pathways are important[426,553]. Whilst this may avoid prejudices based on the scientific paradigm of the researchers, it substitutes them with the assumption that phosphorylation is the most important determinant of signalling. We have seen that interpretation of the phosphoproteome is critically dependent on the total proteome, and inferring kinase activity is not straightforward. Increasing focus is placed on other post-translational modifications (PTMs) such as glycosylation, acetylation, sulfation and ubiquitylation, which may be very important but have been relatively neglected[544]. Phosphorylation may still turn out to be the dominant PTM in signalling networks, but other modifications will find their place in the vocabulary of cells. New discoveries are partly driven by technological innovations that permit detection of these molecular changes. Validation, analysis and bioinformatic methods are still lagging behind detection methods[553].

This degree of complexity can seem overwhelmingly intractable. The emergence of complex phenomena from deterministic simple interactions is much discussed within the context of complexity theory, which is rooted in chaos theory[607]. The interactions of hundreds of atoms within a molecule are difficult to model, yet we derive chemistry as an emergent property of physics. Organisms are complex in themselves yet the study of ecology where organisms are the unit of consideration is possible. Despite great complexity and non-linearity, emergent properties will arise from the large-scale properties of protein networks that can make predictions about cell biology. In fact networks are more robust in the face of perturbations than linear pathways[605]. Removal of nodes in a network (for example by mutation) may not be as catastrophic as removal of a node in a linear pathway. However, it is known that removal of

critically connected nodes (such as P53) can result in widespread changes throughout the network[426]. Certain nodes are more important than others, but their importance lies in their connectedness. High-throughput techniques are demanding new theoretical models. Researchers have been promising profound insights into biology for some time, yet there is much work to be done before we see findings that can be translated into improvements in human health.

**8.5 Translational aspects of this study**

Despite remarkable advances in the treatment of CLL, there is still an unmet need in the treatment of this disease, as the standard of care still exhibits significant toxicity in a group of patients less suited to tolerate this. New therapies showing great efficacy with reduced toxicity are often stated to exert their effects via inhibition of BCR signalling, but their effects may depend partly on interference with other important pathways involving the microenvironment[360]. The new small molecule inhibitors of BTK, PI3K and SYK all have a class effect whereby treatment leads to a pronounced lymphocytosis, as lymph nodes shrink. This mobilisation of CLL cells to exit the LN is felt to be an important component of their efficacy, and have led to the need to adjust response criteria[608]. It seems that these drugs interfere with some aspect of the CLL cell localisation to their microenvironment. BCR binding antigen may be this localisation signal, but these drugs interfere with multiple pathways, including CXCR4 signalling, which may be just as or even more important in this context[360]. Whatever the mechanism, the efficacy of these new agents is heartening.

This study produced estimations of AKT kinase activity and correlated them with sensitivity to PI3K inhibition. If there had been no correlation, one of the explanations would have been the bypassing of the PI3K pathway by parallel pathway activation, a mechanism of drug resistance. Clearly, there are much simpler methods to predict drug sensitivity, but the power of the phosphoproteomic mechanism may be to enable prediction of rational combinations to target multiple pathways that may be active in cancer cells[606]. One study has suggested links between phosphoprotein profiles and responses to tyrosine kinase inhibitors in chronic myeloid leukaemia[609]. Another study successfully predicted a rational combination of EGFR inhibitors and MET inhibitors

that showed synergy in killing cancer cell lines[610]. Studies need not be confined to *in vitro* work, it would be interesting to study CLL samples from patients being treated with kinase inhibitors, particularly in the context of clinical drug resistance or relapse. It is conceivable that each patient will have unique patterns of pathway activation and sensitivity to inhibition, so this approach can appeal to notions of personalised therapy or biomarker prediction[48].

One disadvantage of the phosphoproteomic approach is the need for large amounts of protein from large numbers of cells ($<10^7$ per sample). This limits its applicability with primary samples, but haematological cancers may be amenable to this approach. Consideration of bulk populations also ignores issues such as intratumoral heterogeneity, and it is increasingly recognised that a major challenge to successful therapy is intraclonal diversity[412]. Great excitement has been produced by the use of small molecule kinases inhibitors in cancer therapy, and this is true of CLL. The early success of these drugs requires longer follow up, and it is expected that development of resistance is likely. Combination therapy is likely to become the paradigm, and if these combinations are rationally based on pre-clinical studies they might be more successful[48]. A deeper understanding of leukaemia biology will surely aid the development of new treatments for patients.

# 9. References


1 Kristinsson, S. Y. *et al.* Improved survival in chronic lymphocytic leukemia in the past decade: a population-based study including 11,179 patients diagnosed between 1973-2003 in Sweden. *Haematologica* **94**, 1259-1265, doi:94/9/1259 [pii]
10.3324/haematol.2009.007849 (2009).

2 Dores, G. M. *et al.* Chronic lymphocytic leukaemia and small lymphocytic lymphoma: overview of the descriptive epidemiology. *Br J Haematol* **139**, 809-819, doi:BJH6856 [pii]
10.1111/j.1365-2141.2007.06856.x (2007).

3 Damle, R. N. *et al.* Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94**, 1840-1847 (1999).

4 Hamblin, T. J. *et al.* CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. *Blood* **99**, 1023-1029 (2002).

5 Hallek, M. *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* **111**, 5446-5456, doi:blood-2007-06-093906 [pii]
10.1182/blood-2007-06-093906 (2008).

6 Landgren, O. *et al.* B-cell clones as early markers for chronic lymphocytic leukemia. *N Engl J Med* **360**, 659-667, doi:360/7/659 [pii]
10.1056/NEJMoa0806122 (2009).

7 Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87**, 159-170, doi:S0092-8674(00)81333-1 [pii] (1996).

8 Almeida, J. *et al.* CLL-like B-lymphocytes are systematically present at very low numbers in peripheral blood of healthy adults. *Leukemia*, doi:leu2010305 [pii]
10.1038/leu.2010.305 (2011).

9 Dunn-Walters, D. K. & Ademokun, A. A. B cell repertoire and ageing. *Curr Opin Immunol* **22**, 514-520, doi:S0952-7915(10)00073-7 [pii]
10.1016/j.coi.2010.04.009 (2010).

10 Lanasa, M. C. *et al.* Single-cell analysis reveals oligoclonality among 'low-count' monoclonal B-cell lymphocytosis. *Leukemia* **24**, 133-140, doi:leu2009192 [pii]
10.1038/leu.2009.192 (2010).

11 Fazi, C. *et al.* General population low-count CLL-like MBL persist over time without clinical progression, though carrying the same cytogenetic abnormalities of CLL. *Blood*, doi:blood-2011-05-357251 [pii]
10.1182/blood-2011-05-357251 (2011).

12 Farace, F. *et al.* T cell repertoire in patients with B chronic lymphocytic leukemia. Evidence for multiple in vivo T cell clonal expansions. *J Immunol* **153**, 4281-4290 (1994).

13 Kikushige, Y. *et al.* Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell* **20**, 246-259, doi:S1535-6108(11)00259-5 [pii]
10.1016/j.ccr.2011.06.029 (2011).
14 Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials. CLL Trialists' Collaborative Group. *J Natl Cancer Inst* **91**, 861-868 (1999).
15 Hallek, M. *et al.* Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* **376**, 1164-1174, doi:S0140-6736(10)61381-5 [pii]
10.1016/S0140-6736(10)61381-5 (2010).
16 Butler, T. & Gribben, J. G. Biologic and clinical significance of molecular profiling in Chronic Lymphocytic Leukemia. *Blood Rev* **24**, 135-141, doi:S0268-960X(10)00014-7 [pii]
10.1016/j.blre.2010.03.004 (2010).
17 Castillo, J. J., Furman, M. & Winer, E. S. CAL-101: a phosphatidylinositol-3-kinase p110-delta inhibitor for the treatment of lymphoid malignancies. *Expert Opin Investig Drugs*, doi:10.1517/13543784.2012.640318 (2011).
18 Chen, C. I. *et al.* Single-agent lenalidomide in the treatment of previously untreated chronic lymphocytic leukemia. *J Clin Oncol* **29**, 1175-1181, doi:JCO.2010.29.8133 [pii]
10.1200/JCO.2010.29.8133 (2011).
19 Reeder, C. B. & Ansell, S. M. Novel therapeutic agents for B-cell lymphoma: developing rational combinations. *Blood* **117**, 1453-1462, doi:blood-2010-06-255067 [pii]
10.1182/blood-2010-06-255067 (2011).
20 Lamanna, N. Treatment of Older Patients with Chronic Lymphocytic Leukemia. *Curr Hematol Malig Rep*, doi:10.1007/s11899-011-0111-0 (2012).
21 Bissell, M. J. & Hines, W. C. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med* **17**, 320-329, doi:nm.2328 [pii]
10.1038/nm.2328 (2011).
22 Burger, J. A. Nurture versus Nature: The Microenvironment in Chronic Lymphocytic Leukemia. *Hematology Am Soc Hematol Educ Program* **2011**, 96-103, doi:2011/1/96 [pii]
10.1182/asheducation-2011.1.96 (2011).
23 Ramsay, A. G. & Gribben, J. G. Immune dysfunction in chronic lymphocytic leukemia T cells and lenalidomide as an immunomodulatory drug. *Haematologica* **94**, 1198-1202, doi:94/9/1198 [pii]
10.3324/haematol.2009.009274 (2009).
24 JN, M. Ulcere. *Dictionnaire de Medicine* **21**, 31-50 (1828).
25 N, T. & GT, P. The development of cancer in burns scars. *Surg Gynecol Obstet* **51**, 749-782 (1930).
26 Dohner, H. *et al.* Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* **343**, 1910-1916, doi:MJBA-432602 [pii] (2000).

27 Krober, A. *et al.* V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* **100**, 1410-1416 (2002).
28 Calin, G. A. *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **99**, 15524-15529, doi:10.1073/pnas.242606799
242606799 [pii] (2002).
29 Lia, M. *et al.* Functional dissection of the chromosome 13q14 tumor-suppressor locus using transgenic mouse lines. *Blood* **119**, 2981-2990, doi:blood-2011-09-381814 [pii]
10.1182/blood-2011-09-381814 (2012).
30 Rossi, D. & Gaidano, G. ATM and chronic lymphocytic leukemia: mutations, and not only deletions, matter. *Haematologica* **97**, 5-8, doi:haematol.2011.057109 [pii]
10.3324/haematol.2011.057109 (2012).
31 Zenz, T., Mertens, D., Dohner, H. & Stilgenbauer, S. Importance of genetics in chronic lymphocytic leukemia. *Blood Rev* **25**, 131-137, doi:S0268-960X(11)00020-8 [pii]
10.1016/j.blre.2011.02.002 (2011).
32 Gonzalez, D. *et al.* Mutational Status of the TP53 Gene As a Predictor of Response and Survival in Patients With Chronic Lymphocytic Leukemia: Results From the LRF CLL4 Trial. *J Clin Oncol* **29**, 2223-2229, doi:JCO.2010.32.0838 [pii]
10.1200/JCO.2010.32.0838 (2011).
33 Pekova, S. *et al.* A comprehensive study of TP53 mutations in chronic lymphocytic leukemia: Analysis of 1287 diagnostic and 1148 follow-up CLL samples. *Leuk Res* **35**, 889-898, doi:S0145-2126(10)00612-0 [pii]
10.1016/j.leukres.2010.12.016 (2011).
34 Trbusek, M., Malcikova, J. & Mayer, J. Selection of new TP53 mutations by therapy in chronic lymphocytic leukemia. *Leuk Res* **35**, 981-982, doi:S0145-2126(11)00067-1 [pii]
10.1016/j.leukres.2011.02.005 (2011).
35 Trbusek, M. *et al.* Missense mutations located in structural p53 DNA-binding motifs are associated with extremely poor survival in chronic lymphocytic leukemia. *J Clin Oncol* **29**, 2703-2708, doi:JCO.2011.34.7872 [pii]
10.1200/JCO.2011.34.7872 (2011).
36 Zainuddin, N. *et al.* TP53 Mutations are infrequent in newly diagnosed chronic lymphocytic leukemia. *Leuk Res* **35**, 272-274, doi:S0145-2126(10)00428-5 [pii]
10.1016/j.leukres.2010.08.023 (2011).
37 de Viron, E., Michaux, L., Put, N., Bontemps, F. & van den Neste, E. Present status and perspectives in functional analysis of p53 in chronic lymphocytic leukemia. *Leuk Lymphoma*, doi:10.3109/10428194.2012.660630 (2012).
38 Pospisilova, S. *et al.* ERIC recommendations on TP53 mutation analysis in Chronic Lymphocytic Leukemia. *Leukemia*, doi:leu201225 [pii]
10.1038/leu.2012.25 (2012).
39 Fabbri, G. *et al.* Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med*, doi:jem.20110921 [pii]
10.1084/jem.20110921 (2011).

40 Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*, doi:nature10113 [pii]
10.1038/nature10113 (2011).
41 Rossi, D. *et al.* Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*, doi:blood-2011-09-379966 [pii]
10.1182/blood-2011-09-379966 (2011).
42 Rasi, S. *et al.* Analysis of NOTCH1 mutations in monoclonal B-cell lymphocytosis. *Haematologica* **97**, 153-154, doi:haematol.2011.053090 [pii]
10.3324/haematol.2011.053090 (2012).
43 Rasi, S. *et al.* Analysis of NOTCH1 mutations in monoclonal B cell lymphocytosis. *Haematologica*, doi:haematol.2011.053090 [pii]
10.3324/haematol.2011.053090 (2011).
44 Wang, L. *et al.* SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* **365**, 2497-2506, doi:10.1056/NEJMoa1109016 (2011).
45 Rossi, D. *et al.* Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood*, doi:blood-2011-08-373159 [pii]
10.1182/blood-2011-08-373159 (2011).
46 Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* **44**, 47-52, doi:ng.1032 [pii]
10.1038/ng.1032 (2011).
47 Fabbri, G. *et al.* Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med* **208**, 1389-1401, doi:jem.20110921 [pii]
10.1084/jem.20110921 (2011).
48 Sledge, G. W., Jr. The challenge and promise of the genomic era. *J Clin Oncol* **30**, 203-209, doi:JCO.2011.39.0088 [pii]
10.1200/JCO.2011.39.0088 (2012).
49 Stevenson, F. K. & Caligaris-Cappio, F. Chronic lymphocytic leukemia: revelations from the B-cell receptor. *Blood* **103**, 4389-4395, doi:10.1182/blood-2003-12-4312
2003-12-4312 [pii] (2004).
50 Chilosi, M. *et al.* Immunohistochemical demonstration of follicular dendritic cells in bone marrow involvement of B-cell chronic lymphocytic leukemia. *Cancer* **56**, 328-332 (1985).
51 Schmid, C. & Isaacson, P. G. Proliferation centres in B-cell malignant lymphoma, lymphocytic (B-CLL): an immunophenotypic study. *Histopathology* **24**, 445-451 (1994).
52 Ghia, P. *et al.* Differential effects on CLL cell survival exerted by different microenvironmental elements. *Curr Top Microbiol Immunol* **294**, 135-145 (2005).
53 Herishanu, Y. *et al.* The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* **117**, 563-574, doi:blood-2010-05-284984 [pii]
10.1182/blood-2010-05-284984 (2011).

54 Boelens, J., Philippe, J. & Offner, F. B-CLL cells from lymph nodes express higher ZAP-70 levels than B-CLL cells from peripheral blood. *Leuk Res* **31**, 719-720, doi:S0145-2126(06)00209-8 [pii]
10.1016/j.leukres.2006.05.024 (2007).
55 Caligaris-Cappio, F. & Ghia, P. Novel insights in chronic lymphocytic leukemia: are we getting closer to understanding the pathogenesis of the disease? *J Clin Oncol* **26**, 4497-4503, doi:JCO.2007.15.4393 [pii]
10.1200/JCO.2007.15.4393 (2008).
56 Lagneaux, L., Delforge, A., Bron, D., De Bruyn, C. & Stryckmans, P. Chronic lymphocytic leukemic B cells but not normal B cells are rescued from apoptosis by contact with normal bone marrow stromal cells. *Blood* **91**, 2387-2396 (1998).
57 Ysebaert, L. & Fournie, J. J. Genomic and phenotypic characterization of nurse-like cells that promote drug resistance in chronic lymphocytic leukemia. *Leuk Lymphoma* **52**, 1404-1406, doi:10.3109/10428194.2011.568078 (2011).
58 Hayden, R. E., Pratt, G., Roberts, C., Drayson, M. T. & Bunce, C. M. Treatment of chronic lymphocytic leukemia requires targeting of the protective lymph node environment with novel therapeutic approaches. *Leuk Lymphoma*, doi:10.3109/10428194.2011.610014 (2011).
59 Kurtova, A. V. *et al.* Diverse marrow stromal cells protect CLL cells from spontaneous and drug-induced apoptosis: development of a reliable and reproducible system to assess stromal cell adhesion-mediated drug resistance. *Blood* **114**, 4441-4450, doi:blood-2009-07-233718 [pii]
10.1182/blood-2009-07-233718 (2009).
60 Schrottner, P., Leick, M. & Burger, M. The role of chemokines in B cell chronic lymphocytic leukaemia: pathophysiological aspects and clinical impact. *Ann Hematol* **89**, 437-446, doi:10.1007/s00277-009-0876-6 (2010).
61 Scielzo, C. *et al.* The functional in vitro response to CD40 ligation reflects a different clinical outcome in patients with chronic lymphocytic leukemia. *Leukemia* **25**, 1794, doi:leu2011191 [pii]
10.1038/leu.2011.191 (2011).
62 Bernal, A. *et al.* Survival of leukemic B cells promoted by engagement of the antigen receptor. *Blood* **98**, 3050-3057 (2001).
63 Endo, T. *et al.* BAFF and APRIL support chronic lymphocytic leukemia B-cell survival through activation of the canonical NF-kappaB pathway. *Blood* **109**, 703-710, doi:blood-2006-06-027755 [pii]
10.1182/blood-2006-06-027755 (2007).
64 Sivina, M. *et al.* CCL3 (MIP-1alpha) plasma levels and the risk for disease progression in chronic lymphocytic leukemia. *Blood* **117**, 1662-1669, doi:blood-2010-09-307249 [pii]
10.1182/blood-2010-09-307249 (2011).
65 Malavasi, F. *et al.* CD38 and chronic lymphocytic leukemia: a decade later. *Blood*, doi:blood-2011-06-275610 [pii]
10.1182/blood-2011-06-275610 (2011).
66 Burger, J. A. & Peled, A. CXCR4 antagonists: targeting the microenvironment in leukemia and other cancers. *Leukemia* **23**, 43-52, doi:leu2008299 [pii]
10.1038/leu.2008.299 (2009).

67 Rossi, D. *et al.* CD49d expression is an independent risk factor of progressive disease in early stage chronic lymphocytic leukemia. *Haematologica* **93**, 1575-1579, doi:haematol.13103 [pii]
10.3324/haematol.13103 (2008).
68 Ghia, P. *et al.* Chronic lymphocytic leukemia B cells are endowed with the capacity to attract CD4+, CD40L+ T cells by producing CCL22. *Eur J Immunol* **32**, 1403-1413, doi:10.1002/1521-4141(200205)32:5<1403::AID-IMMU1403>3.0.CO;2-Y (2002).
69 Burger, J. A. *et al.* High-level expression of the T-cell chemokines CCL3 and CCL4 by chronic lymphocytic leukemia B cells in nurselike cell cocultures and after BCR stimulation. *Blood* **113**, 3050-3058, doi:blood-2008-07-170415 [pii]
10.1182/blood-2008-07-170415 (2009).
70 Granziero, L. *et al.* Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. *Blood* **97**, 2777-2783 (2001).
71 Hewamana, S. *et al.* The NF-kappaB subunit Rel A is associated with in vitro survival and clinical disease progression in chronic lymphocytic leukemia and represents a promising therapeutic target. *Blood* **111**, 4681-4689, doi:blood-2007-11-125278 [pii]
10.1182/blood-2007-11-125278 (2008).
72 Ramsay, A. G. *et al.* Chronic lymphocytic leukemia T cells show impaired immunological synapse formation that can be reversed with an immunomodulating drug. *J Clin Invest* **118**, 2427-2437, doi:10.1172/JCI35017 (2008).
73 Riches, J. C., Ramsay, A. G. & Gribben, J. G. T-cell function in chronic lymphocytic leukaemia. *Semin Cancer Biol* **20**, 431-438, doi:S1044-579X(10)00084-2 [pii]
10.1016/j.semcancer.2010.09.006 (2010).
74 Piper, K. P. *et al.* Chronic lymphocytic leukaemia cells drive the global CD4(+) T cell repertoire towards a regulatory phenotype and leads to the accumulation of CD4(+) forkhead box P3(+) T cells. *Clin Exp Immunol* **166**, 154-163, doi:10.1111/j.1365-2249.2011.04466.x (2011).
75 Riches, J. C., Ramsay, A. G. & Gribben, J. G. Immune Reconstitution in Chronic Lymphocytic Leukemia. *Curr Hematol Malig Rep*, doi:10.1007/s11899-011-0106-x (2012).
76 Gorgun, G., Holderried, T. A., Zahrieh, D., Neuberg, D. & Gribben, J. G. Chronic lymphocytic leukemia cells induce changes in gene expression of CD4 and CD8 T cells. *J Clin Invest* **115**, 1797-1805, doi:10.1172/JCI24176 (2005).
77 Lee, B. N. *et al.* Treatment with lenalidomide modulates T-cell immunophenotype and cytokine production in patients with chronic lymphocytic leukemia. *Cancer* **117**, 3999-4008, doi:10.1002/cncr.25983 (2011).
78 Ferrajoli, A. *et al.* Lenalidomide induces complete and partial remissions in patients with relapsed and refractory chronic lymphocytic leukemia. *Blood* **111**, 5291-5297, doi:blood-2007-12-130120 [pii]
10.1182/blood-2007-12-130120 (2008).

79 Burger, J. A. & Hoellenriegel, J. Phosphoinositide 3'-kinase delta: turning off BCR signaling in Chronic Lymphocytic Leukemia. *Oncotarget* **2**, 737-738, doi:341 [pii] (2011).
80 Flinn, I. W. B-cell receptor inhibitors in chronic lymphocytic leukemia. *Clin Adv Hematol Oncol* **9**, 605-606 (2011).
81 Kuppers, R., Klein, U., Hansmann, M. L. & Rajewsky, K. Cellular origin of human B-cell lymphomas. *N Engl J Med* **341**, 1520-1529 (1999).
82 Welner, R. S., Pelayo, R. & Kincade, P. W. Evolving views on the genealogy of B cells. *Nat Rev Immunol* **8**, 95-106, doi:nri2234 [pii]
10.1038/nri2234 (2008).
83 Swerdlow, S. H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H., Thiele, J., Vardiman, J.W & IARC. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition*. (2008).
84 Perez-Andres, M. *et al.* Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. *Cytometry B Clin Cytom* **78 Suppl 1**, S47-60, doi:10.1002/cyto.b.20547 (2010).
85 Hombach, J., Tsubata, T., Leclercq, L., Stappert, H. & Reth, M. Molecular components of the B-cell antigen receptor complex of the IgM class. *Nature* **343**, 760-762, doi:10.1038/343760a0 (1990).
86 Black, C. A. A brief history of the discovery of the immunoglobulins and the origin of the modern immunoglobulin nomenclature. *Immunol Cell Biol* **75**, 65-68, doi:10.1038/icb.1997.10 (1997).
87 Carsetti, R., Rosado, M. M. & Wardmann, H. Peripheral development of B cells in mouse and man. *Immunol Rev* **197**, 179-191, doi:109 [pii] (2004).
88 LeBien, T. W. & Tedder, T. F. B lymphocytes: how they develop and function. *Blood* **112**, 1570-1580, doi:112/5/1570 [pii]
10.1182/blood-2008-02-078071 (2008).
89 Schlissel, M. S. Regulating antigen-receptor gene assembly. *Nat Rev Immunol* **3**, 890-899, doi:10.1038/nri1225
nri1225 [pii] (2003).
90 Ghia, P., Rosenquist, R. & Davi, F. *Immunoglbulin Gene Analysis in Chronic Lymphocytic Leukemia*. (Wolters Kluwer, 2009).
91 Nemazee, D. Receptor editing in lymphocyte development and central tolerance. *Nat Rev Immunol* **6**, 728-740, doi:nri1939 [pii]
10.1038/nri1939 (2006).
92 Chen, C., Prak, E. L. & Weigert, M. Editing disease-associated autoantibodies. *Immunity* **6**, 97-105, doi:S1074-7613(00)80673-1 [pii] (1997).
93 Kohler, F. *et al.* Autoreactive B cell receptors mimic autonomous pre-B cell receptor signaling and induce proliferation of early B cells. *Immunity* **29**, 912-921, doi:S1074-7613(08)00505-0 [pii]
10.1016/j.immuni.2008.10.013 (2008).
94 Zikherman, J., Parameswaran, R. & Weiss, A. Endogenous antigen tunes the responsiveness of naive B cells but not T cells. *Nature*, doi:nature11311 [pii]
10.1038/nature11311 (2012).
95 Allen, C. D., Okada, T. & Cyster, J. G. Germinal-center organization and cellular dynamics. *Immunity* **27**, 190-202, doi:S1074-7613(07)00370-6 [pii]
10.1016/j.immuni.2007.07.009 (2007).

96 Liu, Y. J. *et al.* Mechanism of antigen-driven selection in germinal centres. *Nature* **342**, 929-931, doi:10.1038/342929a0 (1989).
97 Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annu Rev Immunol*, doi:10.1146/annurev-immunol-020711-075032 (2011).
98 Maddaly, R. *et al.* Receptors and signaling mechanisms for B-lymphocyte activation, proliferation and differentiation--insights from both in vivo and in vitro approaches. *FEBS Lett* **584**, 4883-4894, doi:S0014-5793(10)00678-2 [pii]
10.1016/j.febslet.2010.08.022 (2010).
99 Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol* **8**, 22-33, doi:nri2217 [pii]
10.1038/nri2217 (2008).
100 Mahowald, G. K., Baron, J. M. & Sleckman, B. P. Collateral damage from antigen receptor gene diversification. *Cell* **135**, 1009-1012, doi:S0092-8674(08)01499-2 [pii]
10.1016/j.cell.2008.11.024 (2008).
101 Stavnezer, J., Guikema, J. E. & Schrader, C. E. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* **26**, 261-292, doi:10.1146/annurev.immunol.26.021607.090248 (2008).
102 Weill, J. C., Weller, S. & Reynaud, C. A. Human marginal zone B cells. *Annu Rev Immunol* **27**, 267-285, doi:10.1146/annurev.immunol.021908.132607 (2009).
103 Forconi, F. *et al.* The normal IGHV1-69-derived B-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood* **115**, 71-77, doi:blood-2009-06-225813 [pii]
10.1182/blood-2009-06-225813 (2010).
104 Forconi, F. *et al.* The IGHV1-69/IGHJ3 recombinations of unmutated CLL are distinct from those of normal B cells. *Blood*, doi:blood-2011-08-375501 [pii]
10.1182/blood-2011-08-375501 (2012).
105 Visvader, J. E. Cells of origin in cancer. *Nature* **469**, 314-322, doi:nature09781 [pii]
10.1038/nature09781 (2011).
106 Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat Med* **17**, 313-319, doi:nm.2304 [pii]
10.1038/nm.2304 (2011).
107 Hasserjian, R. P. *et al.* Commentary on the WHO classification of tumors of lymphoid tissues (2008): "Gray zone" lymphomas overlapping with Burkitt lymphoma or classical Hodgkin lymphoma. *J Hematop*, doi:10.1007/s12308-009-0039-7 (2009).
108 Weller, S. *et al.* CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc Natl Acad Sci U S A* **98**, 1166-1170, doi:10.1073/pnas.98.3.1166
98/3/1166 [pii] (2001).
109 Caligaris-Cappio, F. & Ghia, P. The normal counterpart to the chronic lymphocytic leukemia B cell. *Best Pract Res Clin Haematol* **20**, 385-397, doi:S1521-6926(07)00021-7 [pii]
10.1016/j.beha.2007.02.005 (2007).

110 Chiorazzi, N. & Ferrarini, M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood*, doi:blood-2010-07-155663 [pii]
10.1182/blood-2010-07-155663 (2010).
111 Darzentas, N. *et al.* A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia* **24**, 125-132, doi:leu2009186 [pii]
10.1038/leu.2009.186 (2010).
112 Bahler, D. W. & Levy, R. Clonal evolution of a follicular lymphoma: evidence for antigen selection. *Proc Natl Acad Sci U S A* **89**, 6770-6774 (1992).
113 Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**, 1848-1854 (1999).
114 Binder, M. *et al.* Stereotypical chronic lymphocytic leukemia B-cell receptors recognize survival promoting antigens on stromal cells. *PLoS ONE* **5**, e15992, doi:10.1371/journal.pone.0015992 (2010).
115 Muzio, M. *et al.* Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood* **112**, 188-195, doi:blood-2007-09-111344 [pii]
10.1182/blood-2007-09-111344 (2008).
116 Mockridge, C. I. *et al.* Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood* **109**, 4424-4431, doi:blood-2006-11-056648 [pii]
10.1182/blood-2006-11-056648 (2007).
117 Packham, G. & Stevenson, F. The role of the B-cell receptor in the pathogenesis of chronic lymphocytic leukaemia. *Semin Cancer Biol*, doi:S1044-579X(10)00076-3 [pii]
10.1016/j.semcancer.2010.08.004 (2010).
118 Lam, K. P., Kuhn, R. & Rajewsky, K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell* **90**, 1073-1083, doi:S0092-8674(00)80373-6 [pii] (1997).
119 Smith, S. H. & Reth, M. Perspectives on the nature of BCR-mediated survival signals. *Mol Cell* **14**, 696-697, doi:10.1016/j.molcel.2004.06.015
S1097276504003430 [pii] (2004).
120 Grande, S. M., Bannish, G., Fuentes-Panana, E. M., Katz, E. & Monroe, J. G. Tonic B-cell and viral ITAM signaling: context is everything. *Immunol Rev* **218**, 214-234, doi:IMR535 [pii]
10.1111/j.1600-065X.2007.00535.x (2007).
121 Delgado, P. *et al.* Essential function for the GTPase TC21 in homeostatic antigen receptor signaling. *Nat Immunol* **10**, 880-888, doi:ni.1749 [pii]
10.1038/ni.1749 (2009).
122 Kurosaki, T., Shinohara, H. & Baba, Y. B cell signaling and fate decision. *Annu Rev Immunol* **28**, 21-55, doi:10.1146/annurev.immunol.021908.132541 (2010).
123 Kuppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* **5**, 251-262, doi:nrc1589 [pii]
10.1038/nrc1589 (2005).

124 Caldwell, R. G., Wilson, J. B., Anderson, S. J. & Longnecker, R. Epstein-Barr virus LMP2A drives B cell development and survival in the absence of normal B cell receptor signals. *Immunity* **9**, 405-411, doi:S1074-7613(00)80623-8 [pii] (1998).
125 de Martel, C. *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol* **13**, 607-615, doi:S1470-2045(12)70137-7 [pii]
10.1016/S1470-2045(12)70137-7 (2012).
126 Saha, A. & Robertson, E. S. Epstein-Barr virus-associated B-cell lymphomas: pathogenesis and clinical outcomes. *Clin Cancer Res* **17**, 3056-3063, doi:1078-0432.CCR-10-2578 [pii]
10.1158/1078-0432.CCR-10-2578 (2011).
127 Carbone, A. & Gloghini, A. KSHV/HHV8-associated lymphomas. *Br J Haematol* **140**, 13-24, doi:BJH6879 [pii]
10.1111/j.1365-2141.2007.06879.x (2008).
128 Silvestri, F. *et al.* Prevalence of hepatitis C virus infection in patients with lymphoproliferative disorders. *Blood* **87**, 4296-4301 (1996).
129 Reitz, M. S., Jr. *et al.* Human T-cell leukemia/lymphoma virus: the retrovirus of adult T-cell leukemia/lymphoma. *J Infect Dis* **147**, 399-405 (1983).
130 Jelic, S. & Filipovic-Ljeskovic, I. Positive serology for Lyme disease borrelias in primary cutaneous B-cell lymphoma: a study in 22 patients; is it a fortuitous finding? *Hematol Oncol* **17**, 107-116, doi:10.1002/(SICI)1099-1069(199909)17:3<107::AID-HON644>3.0.CO;2-R [pii] (1999).
131 Lecuit, M. *et al.* Immunoproliferative small intestinal disease associated with Campylobacter jejuni. *N Engl J Med* **350**, 239-248, doi:10.1056/NEJMoa031887
350/3/239 [pii] (2004).
132 Ferreri, A. J. *et al.* Chlamydophila Psittaci Eradication With Doxycycline As First-Line Targeted Therapy for Ocular Adnexae Lymphoma: Final Results of an International Phase II Trial. *J Clin Oncol* **30**, 2988-2994, doi:JCO.2011.41.4466 [pii]
10.1200/JCO.2011.41.4466 (2012).
133 Parsonnet, J. & Isaacson, P. G. Bacterial infection and MALT lymphoma. *N Engl J Med* **350**, 213-215, doi:10.1056/NEJMp038200
350/3/213 [pii] (2004).
134 Wotherspoon, A. C. *et al.* Regression of primary low-grade B-cell gastric lymphoma of mucosa-associated lymphoid tissue type after eradication of Helicobacter pylori. *Lancet* **342**, 575-577, doi:0140-6736(93)91409-F [pii] (1993).
135 Sagaert, X., Van Cutsem, E., De Hertogh, G., Geboes, K. & Tousseyn, T. Gastric MALT lymphoma: a model of chronic inflammation-induced tumor development. *Nat Rev Gastroenterol Hepatol* **7**, 336-346, doi:nrgastro.2010.58 [pii]
10.1038/nrgastro.2010.58 (2010).
136 Ngo, V. N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115-119, doi:nature09671 [pii]
10.1038/nature09671 (2011).

137 Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-721, doi:nature07968 [pii]
10.1038/nature07968 (2009).
138 Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676-1679, doi:1153629 [pii]
10.1126/science.1153629 (2008).
139 Goodyear, C. S. & Silverman, G. J. Death by a B cell superantigen: In vivo VH-targeted apoptotic supraclonal B cell deletion by a Staphylococcal Toxin. *J Exp Med* **197**, 1125-1139, doi:10.1084/jem.20020552
jem.20020552 [pii] (2003).
140 Silverman, G. J. & Goodyear, C. S. Confounding B-cell defences: lessons from a staphylococcal superantigen. *Nat Rev Immunol* **6**, 465-475, doi:nri1853 [pii]
10.1038/nri1853 (2006).
141 Kipps, T. J. *et al.* Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **86**, 5913-5917 (1989).
142 Murray, F. *et al.* Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood* **111**, 1524-1533, doi:blood-2007-07-099564 [pii]
10.1182/blood-2007-07-099564 (2008).
143 Johnson, T. A., Rassenti, L. Z. & Kipps, T. J. Ig VH1 genes expressed in B cell chronic lymphocytic leukemia exhibit distinctive molecular features. *J Immunol* **158**, 235-246 (1997).
144 Zibellini, S. *et al.* Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica* **95**, 1792-1796, doi:haematol.2010.025437 [pii]
10.3324/haematol.2010.025437 (2010).
145 Thelander, E. F. & Rosenquist, R. Molecular genetic characterization reveals new subsets of mantle cell lymphoma. *Leuk Lymphoma* **49**, 1042-1049, doi:792731304 [pii]
10.1080/10428190801947559 (2008).
146 Miklos, J. A., Swerdlow, S. H. & Bahler, D. W. Salivary gland mucosa-associated lymphoid tissue lymphoma immunoglobulin V(H) genes show frequent use of V1-69 with distinctive CDR3 features. *Blood* **95**, 3878-3884 (2000).
147 Starkebaum, G. & Sasso, E. H. Hepatitis C and B cells: induction of autoimmunity and lymphoproliferation may reflect chronic stimulation through cell-surface receptors. *J Rheumatol* **31**, 416-418, doi:0315162X-31-416 [pii] (2004).
148 Chan, C. H., Hadlock, K. G., Foung, S. K. & Levy, S. V(H)1-69 gene is preferentially used by hepatitis C virus-associated B cell lymphomas and by normal B cells responding to the E2 viral antigen. *Blood* **97**, 1023-1026 (2001).
149 Ferri, C. *et al.* Etiopathogenetic role of hepatitis C virus in mixed cryoglobulinemia, chronic liver diseases and lymphomas. *Clin Exp Rheumatol* **13 Suppl 13**, S135-140 (1995).

150 Widhopf, G. F., 2nd *et al.* Nonstochastic pairing of immunoglobulin heavy and light chains expressed by chronic lymphocytic leukemia B cells is predicated on the heavy chain CDR3. *Blood* **111**, 3137-3144, doi:blood-2007-02-073130 [pii]
10.1182/blood-2007-02-073130 (2008).
151 De Re, V. *et al.* HCV-NS3 and IgG-Fc crossreactive IgM in patients with type II mixed cryoglobulinemia and B-cell clonal proliferations. *Leukemia* **20**, 1145-1154, doi:2404201 [pii]
10.1038/sj.leu.2404201 (2006).
152 Bhat, N. M., Bieber, M. M., Chapman, C. J., Stevenson, F. K. & Teng, N. N. Human antilipid A monoclonal antibodies bind to human B cells and the i antigen on cord red blood cells. *J Immunol* **151**, 5011-5021 (1993).
153 Silberstein, L. E., George, A., Durdik, J. M. & Kipps, T. J. The V4-34 encoded anti-i autoantibodies recognize a large subset of human and mouse B-cells. *Blood Cells Mol Dis* **22**, 126-138, doi:S1079-9796(96)90020-0 [pii]
10.1006/bcmd.1996.0020 (1996).
154 Potter, K. N., Hobby, P., Klijn, S., Stevenson, F. K. & Sutton, B. J. Evidence for involvement of a hydrophobic patch in framework region 1 of human V4-34-encoded Igs in recognition of the red blood cell I antigen. *J Immunol* **169**, 3777-3782 (2002).
155 Pugh-Bernard, A. E. *et al.* Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. *J Clin Invest* **108**, 1061-1070, doi:10.1172/JCI12462 (2001).
156 Pascual, V. *et al.* VH restriction among human cold agglutinins. The VH4-21 gene segment is required to encode anti-I and anti-i specificities. *J Immunol* **149**, 2337-2344 (1992).
157 Montesinos-Rongen, M. *et al.* Primary central nervous system lymphomas are derived from germinal-center B cells and show a preferential usage of the V4-34 gene segment. *Am J Pathol* **155**, 2077-2086 (1999).
158 Thompsett, A. R., Ellison, D. W., Stevenson, F. K. & Zhu, D. V(H) gene sequences from primary central nervous system lymphomas indicate derivation from highly mutated germinal center B cells with ongoing mutational activity. *Blood* **94**, 1738-1746 (1999).
159 Bhat, N. M. *et al.* B cell lymphoproliferative disorders and VH4-34 gene encoded antibodies. *Hum Antibodies* **13**, 63-68 (2004).
160 Chapman, C. J. *et al.* Autoanti-red cell antibodies synthesized by patients with infectious mononucleosis utilize the VH4-21 gene segment. *J Immunol* **151**, 1051-1061 (1993).
161 Kostareli, E. *et al.* Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Leukemia*, doi:leu2008379 [pii]
10.1038/leu.2008.379 (2009).
162 Klein, E. & Nagy, N. Restricted expression of EBV encoded proteins in in vitro infected CLL cells. *Semin Cancer Biol* **20**, 410-415, doi:S1044-579X(10)00097-0 [pii]
10.1016/j.semcancer.2010.10.013 (2010).

163 Isenberg, D., Spellerberg, M., Williams, W., Griffiths, M. & Stevenson, F. Identification of the 9G4 idiotope in systemic lupus erythematosus. *Br J Rheumatol* **32**, 876-882 (1993).
164 Mockridge, C. I. *et al.* Common patterns of B cell perturbation and expanded V4-34 immunoglobulin gene usage in autoimmunity and infection. *Autoimmunity* **37**, 9-15 (2004).
165 Cappione, A. J., Pugh-Bernard, A. E., Anolik, J. H. & Sanz, I. Lupus IgG VH4.34 antibodies bind to a 220-kDa glycoform of CD45/B220 on the surface of human B lymphocytes. *J Immunol* **172**, 4298-4307 (2004).
166 Fais, F. *et al.* Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest* **102**, 1515-1525, doi:10.1172/JCI3009 (1998).
167 Herve, M. *et al.* Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J Clin Invest* **115**, 1636-1643, doi:10.1172/JCI24387 (2005).
168 Efremov, D. G. *et al.* Restricted immunoglobulin VH region repertoire in chronic lymphocytic leukemia patients with autoimmune hemolytic anemia. *Blood* **87**, 3869-3876 (1996).
169 Tobin, G. *et al.* Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. *Blood* **104**, 2879-2885, doi:10.1182/blood-2004-01-0132
2004-01-0132 [pii] (2004).
170 Dal-Bo, M. *et al.* B-cell receptor, clinical course and prognosis in chronic lymphocytic leukaemia: the growing saga of the IGHV3 subgroup gene usage. *Br J Haematol*, doi:10.1111/j.1365-2141.2010.08440.x (2011).
171 Ghiotto, F. *et al.* Mutation pattern of paired immunoglobulin heavy and light variable domains in chronic lymphocytic leukemia B-cells. *Mol Med*, doi:molmed.2011.00104 [pii]
10.2119/molmed.2011.00104 (2011).
172 Maura, F. *et al.* Relevance of stereotyped B-cell receptors in the context of the molecular, cytogenetic and clinical features of chronic lymphocytic leukemia. *PLoS One* **6**, e24313, doi:10.1371/journal.pone.0024313
PONE-D-11-09072 [pii] (2011).
173 Tsakou, E. *et al.* Partial Versus Productive IGH Locus Rearrangements in Chronic Lymphocytic Leukemia: Implications for B-Cell Receptor Stereotypy. *Mol Med*, doi:molmed.2011.00216 [pii]
10.2119/molmed.2011.00216 (2011).
174 Messmer, B. T. *et al.* Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med* **200**, 519-525, doi:10.1084/jem.20040544
jem.20040544 [pii] (2004).
175 Bomben, R. *et al.* Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicentre study. *Br J Haematol* **144**, 492-506, doi:BJH7469 [pii]
10.1111/j.1365-2141.2008.07469.x (2009).

176 Rossi, D. *et al.* Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to richter syndrome. *Clin Cancer Res* **15**, 4415-4422, doi:1078-0432.CCR-08-3266 [pii]
10.1158/1078-0432.CCR-08-3266 (2009).
177 Athanasiadou, A. *et al.* Recurrent cytogenetic findings in subsets of patients with chronic lymphocytic leukemia expressing IgG-switched stereotyped immunoglobulins. *Haematologica* **93**, 473-474, doi:93/3/473 [pii]
10.3324/haematol.11872 (2008).
178 Balatti, V. *et al.* NOTCH1 mutations in CLL associated with trisomy 12. *Blood*, doi:blood-2011-10-386144 [pii]
10.1182/blood-2011-10-386144 (2011).
179 Bomben, R. *et al.* Expression of mutated IGHV3-23 genes in chronic lymphocytic leukemia identifies a disease subset with peculiar clinical and biological features. *Clin Cancer Res* **16**, 620-628, doi:1078-0432.CCR-09-1638 [pii]
10.1158/1078-0432.CCR-09-1638 (2010).
180 Bende, R. J. *et al.* Among B cell non-Hodgkin's lymphomas, MALT lymphomas express a unique antibody repertoire with frequent rheumatoid factor reactivity. *J Exp Med* **201**, 1229-1241, doi:jem.20050068 [pii]
10.1084/jem.20050068 (2005).
181 Ghiotto, F. *et al.* Remarkably similar antigen receptors among a subset of patients with chronic lymphocytic leukemia. *J Clin Invest* **113**, 1008-1016, doi:10.1172/JCI19399 (2004).
182 Stamatopoulos, K. *et al.* Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood* **109**, 259-270, doi:blood-2006-03-012948 [pii]
10.1182/blood-2006-03-012948 (2007).
183 Potter, K. N. *et al.* Features of the overexpressed V1-69 genes in the unmutated subset of chronic lymphocytic leukemia are distinct from those in the healthy elderly repertoire. *Blood* **101**, 3082-3084, doi:10.1182/blood-2002-08-2432
2002-08-2432 [pii] (2003).
184 Potter, K. N. *et al.* Structural and functional features of the B-cell receptor in IgG-positive chronic lymphocytic leukemia. *Clin Cancer Res* **12**, 1672-1679, doi:12/6/1672 [pii]
10.1158/1078-0432.CCR-05-2164 (2006).
185 Rosen, A., Murray, F., Evaldsson, C. & Rosenquist, R. Antigens in chronic lymphocytic leukemia--implications for cell origin and leukemogenesis. *Semin Cancer Biol* **20**, 400-409, doi:S1044-579X(10)00082-9 [pii]
10.1016/j.semcancer.2010.09.004 (2010).
186 Efremov, D. G., Gobessi, S. & Longo, P. G. Signaling pathways activated by antigen-receptor engagement in chronic lymphocytic leukemia B-cells. *Autoimmun Rev* **7**, 102-108, doi:S1568-9972(07)00058-4 [pii]
10.1016/j.autrev.2007.02.021 (2007).
187 Mackay, I. R. & Rose, N. R. Autoimmunity and lymphoma: tribulations of B cells. *Nat Immunol* **2**, 793-795, doi:10.1038/ni0901-793
ni0901-793 [pii] (2001).
188 Kassan, S. S. *et al.* Increased risk of lymphoma in sicca syndrome. *Ann Intern Med* **89**, 888-892 (1978).

189 Lindsay, S. & Dailey, M. E. Malignant lymphoma of the thyroid gland and its relation to Hashimoto disease: a clinical and pathologic study of 8 patients. *J Clin Endocrinol Metab* **15**, 1332-1353 (1955).
190 Caligaris-Cappio, F. Autoimmune disorders and lymphoma. *Ann Oncol* **19 Suppl 4**, iv31-34, doi:mdn190 [pii]
10.1093/annonc/mdn190 (2008).
191 Goldin, L. R. & Landgren, O. Autoimmunity and lymphomagenesis. *Int J Cancer* **124**, 1497-1502, doi:10.1002/ijc.24141 (2009).
192 Landgren, O. *et al.* Autoimmunity and susceptibility to Hodgkin lymphoma: a population-based case-control study in Scandinavia. *J Natl Cancer Inst* **98**, 1321-1330, doi:98/18/1321 [pii]
10.1093/jnci/djj361 (2006).
193 Landgren, O. *et al.* Patterns of autoimmunity and subsequent chronic lymphocytic leukemia in Nordic countries. *Blood* **108**, 292-296, doi:2005-11-4620 [pii]
10.1182/blood-2005-11-4620 (2006).
194 Zintzaras, E., Voulgarelis, M. & Moutsopoulos, H. M. The risk of lymphoma development in autoimmune diseases: a meta-analysis. *Arch Intern Med* **165**, 2337-2344, doi:165/20/2337 [pii]
10.1001/archinte.165.20.2337 (2005).
195 Hodgson, K., Ferrer, G., Pereira, A., Moreno, C. & Montserrat, E. Autoimmune cytopenia in chronic lymphocytic leukaemia: diagnosis and treatment. *Br J Haematol* **154**, 14-22, doi:10.1111/j.1365-2141.2011.08707.x (2011).
196 Preud'homme, J. L. & Seligmann, M. Anti-human immunoglobulin G activity of membrane-bound monoclonal immunoglobulin M in lymphoproliferative disorders. *Proc Natl Acad Sci U S A* **69**, 2132-2135 (1972).
197 Kipps, T. J. & Carson, D. A. Autoantibodies in chronic lymphocytic leukemia and related systemic autoimmune diseases. *Blood* **81**, 2475-2487 (1993).
198 Broker, B. M. *et al.* Chronic lymphocytic leukemic (CLL) cells secrete multispecific autoantibodies. *J Autoimmun* **1**, 469-481 (1988).
199 Sthoeger, Z. M. *et al.* Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. *J Exp Med* **169**, 255-268 (1989).
200 Ehrenstein, M. R. & Notley, C. A. The importance of natural IgM: scavenger, protector and regulator. *Nat Rev Immunol* **10**, 778-786, doi:nri2849 [pii]
10.1038/nri2849 (2010).
201 Hansson, G. K. & Hermansson, A. The immune system in atherosclerosis. *Nat Immunol* **12**, 204-212, doi:ni.2001 [pii]
10.1038/ni.2001 (2011).
202 Dighiero, G. *et al.* Autoantibody activity of immunoglobulins isolated from B-cell follicular lymphomas. *Blood* **78**, 581-585 (1991).
203 Casciola-Rosen, L. A., Anhalt, G. & Rosen, A. Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. *J Exp Med* **179**, 1317-1330 (1994).
204 Shaw, P. X., Goodyear, C. S., Chang, M. K., Witztum, J. L. & Silverman, G. J. The autoreactivity of anti-phosphorylcholine antibodies for atherosclerosis-associated neo-antigens and apoptotic cells. *J Immunol* **170**, 6151-6157 (2003).

205 Chang, M. K. *et al.* Apoptotic cells with oxidation-specific epitopes are immunogenic and proinflammatory. *J Exp Med* **200**, 1359-1370, doi:jem.20031763 [pii]
10.1084/jem.20031763 (2004).
206 Chu, C. C. *et al.* Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood* **112**, 5122-5129, doi:blood-2008-06-162024 [pii]
10.1182/blood-2008-06-162024 (2008).
207 Chu, C. C. *et al.* Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: implications for patient outcome and cell of origin. *Blood* **115**, 3907-3915, doi:blood-2009-09-244251 [pii]
10.1182/blood-2009-09-244251 (2010).
208 Binder, M. *et al.* B-cell receptor epitope recognition correlates with the clinical course of chronic lymphocytic leukemia. *Cancer*, doi:10.1002/cncr.25755 (2010).
209 Lanemo Myhrinder, A. *et al.* A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood* **111**, 3838-3848, doi:blood-2007-11-125450 [pii]
10.1182/blood-2007-11-125450 (2008).
210 Vossenaar, E. R. *et al.* Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. *Arthritis Res Ther* **6**, R142-150, doi:10.1186/ar1149
ar1149 [pii] (2004).
211 Steininger, C. *et al.* Recombinant antibodies encoded by IGHV1-69 react with pUL32, a phosphoprotein of cytomegalovirus and B-cell superantigen. *Blood*, doi:blood-2011-08-374058 [pii]
10.1182/blood-2011-08-374058 (2012).
212 Binder, C. J. *et al.* Pneumococcal vaccination decreases atherosclerotic lesion formation: molecular mimicry between Streptococcus pneumoniae and oxidized LDL. *Nat Med* **9**, 736-743, doi:10.1038/nm876
nm876 [pii] (2003).
213 Anderson, L. A., Landgren, O. & Engels, E. A. Common community acquired infections and subsequent risk of chronic lymphocytic leukaemia. *Br J Haematol*, doi:BJH7849 [pii]
10.1111/j.1365-2141.2009.07849.x (2009).
214 Moreira, J. *et al.* Infectious complications among individuals with clinical Monoclonal B-cell Lymphocytosis (MBL): a cohort study of newly diagnosed cases compared to controls. *Leukemia*, doi:leu2012187 [pii]
10.1038/leu.2012.187 (2012).
215 trials.gov, C. *CLL Empirical Antibiotic Regimen (CLEAR)*, <http://clinicaltrials.gov/ct2/show/NCT01279252> (2012).
216 Minden, M. D. *et al.* Chronic lymphocytic leukaemia is driven by antigen-independent cell-autonomous signalling. *Nature*, doi:nature11309 [pii]
10.1038/nature11309 (2012).
217 Borche, L., Lim, A., Binet, J. L. & Dighiero, G. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood* **76**, 562-569 (1990).

218 Damle, R. N. *et al.* B-cell chronic lymphocytic leukemia cells express a surface membrane phenotype of activated, antigen-experienced B lymphocytes. *Blood* **99**, 4087-4093 (2002).
219 Klein, U. *et al.* Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med* **194**, 1625-1638 (2001).
220 Sutton, L. A. *et al.* Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. *Blood* **114**, 4460-4468, doi:blood-2009-05-221309 [pii]
10.1182/blood-2009-05-221309 (2009).
221 Gurrieri, C. *et al.* Chronic lymphocytic leukemia B cells can undergo somatic hypermutation and intraclonal immunoglobulin V(H)DJ(H) gene diversification. *J Exp Med* **196**, 629-639 (2002).
222 Marantidou, F. *et al.* Activation-induced cytidine deaminase splicing patterns in chronic lymphocytic leukemia. *Blood Cells Mol Dis* **44**, 262-267, doi:S1079-9796(09)00243-5 [pii]
10.1016/j.bcmd.2009.12.005 (2010).
223 Green, A. Vol. 6 (2012).
224 Moreno, C. *et al.* Autoimmune cytopenia in chronic lymphocytic leukemia: prevalence, clinical associations, and prognostic significance. *Blood* **116**, 4771-4776, doi:blood-2010-05-286500 [pii]
10.1182/blood-2010-05-286500 (2010).
225 Kay, N. E., Johnson, J. D., Stanek, R. & Douglas, S. D. T-cell subpopulations in chronic lymphocytic leukemia: abnormalities in distribtuion and in in vitro receptor maturation. *Blood* **54**, 540-544 (1979).
226 Bojarska-Junak, A. *et al.* BAFF and APRIL expression in B-cell chronic lymphocytic leukemia: correlation with biological and clinical features. *Leuk Res* **33**, 1319-1327, doi:S0145-2126(09)00158-1 [pii]
10.1016/j.leukres.2009.03.030 (2009).
227 Molica, S. *et al.* Baff serum level predicts time to first treatment in early chronic lymphocytic leukemia. *Eur J Haematol*, doi:EJH1482 [pii]
10.1111/j.1600-0609.2010.01482.x (2010).
228 Pham, L. V. & Ford, R. J. The role of BAFF-R dysregulation in B-lymphoid lineage malignancies. *Cell Cycle* **10**, 189-190, doi:14570 [pii] (2011).
229 Mackay, F. *et al.* Mice transgenic for BAFF develop lymphocytic disorders along with autoimmune manifestations. *J Exp Med* **190**, 1697-1710 (1999).
230 Lesley, R. *et al.* Reduced competitiveness of autoantigen-engaged B cells due to increased dependence on BAFF. *Immunity* **20**, 441-453, doi:S1074761304000792 [pii] (2004).
231 Andersen-Ranberg, K., M, H. O.-M., Wiik, A., Jeune, B. & Hegedus, L. High prevalence of autoantibodies among Danish centenarians. *Clin Exp Immunol* **138**, 158-163, doi:10.1111/j.1365-2249.2004.02575.x
CEI2575 [pii] (2004).
232 Cambier, J. C., Gauld, S. B., Merrell, K. T. & Vilen, B. J. B-cell anergy: from transgenic models to naturally occurring anergic B cells? *Nat Rev Immunol* **7**, 633-643, doi:nri2133 [pii]

10.1038/nri2133 (2007).
233 Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374-1377, doi:10.1126/science.1086907
1086907 [pii] (2003).
234 Yarkoni, Y., Getahun, A. & Cambier, J. C. Molecular underpinning of B-cell anergy. *Immunol Rev* **237**, 249-263, doi:IMR936 [pii]
10.1111/j.1600-065X.2010.00936.x (2010).
235 Goodnow, C. C. *et al.* Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* **334**, 676-682, doi:10.1038/334676a0 (1988).
236 Goodnow, C. C., Crosbie, J., Jorgensen, H., Brink, R. A. & Basten, A. Induction of self-tolerance in mature peripheral B lymphocytes. *Nature* **342**, 385-391, doi:10.1038/342385a0 (1989).
237 Hartley, S. B. *et al.* Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature* **353**, 765-769, doi:10.1038/353765a0 (1991).
238 Goodnow, C. C., Brink, R. & Adams, E. Breakdown of self-tolerance in anergic B lymphocytes. *Nature* **352**, 532-536, doi:10.1038/352532a0 (1991).
239 Getahun, A., O'Neill, S. K. & Cambier, J. C. Establishing anergy as a bona fide in vivo mechanism of B cell tolerance. *J Immunol* **183**, 5439-5441, doi:183/9/5439 [pii]
10.4049/jimmunol.0990088 (2009).
240 Merrell, K. T. *et al.* Identification of anergic B cells within a wild-type repertoire. *Immunity* **25**, 953-962, doi:S1074-7613(06)00524-3 [pii]
10.1016/j.immuni.2006.10.017 (2006).
241 Duty, J. A. *et al.* Functional anergy in a subpopulation of naive B cells from healthy humans that express autoreactive immunoglobulin receptors. *J Exp Med* **206**, 139-151, doi:jem.20080611 [pii]
10.1084/jem.20080611 (2009).
242 Tough, D. F. & Sprent, J. Lifespan of lymphocytes. *Immunol Res* **14**, 1-12 (1995).
243 Fulcher, D. A. & Basten, A. Reduced life span of anergic self-reactive B cells in a double-transgenic model. *J Exp Med* **179**, 125-134 (1994).
244 Mackay, F., Schneider, P., Rennert, P. & Browning, J. BAFF AND APRIL: a tutorial on B cell survival. *Annu Rev Immunol* **21**, 231-264, doi:10.1146/annurev.immunol.21.120601.141152
120601.141152 [pii] (2003).
245 Cristina, T. *et al.* April (a Proliferation-Inducing Ligand) Serum Levels Predict Time to First Treatment in Patients Affected by B-Cell Chronic Lymphocytic Leukemia. *Eur J Haematol*, doi:10.1111/j.1600-0609.2011.01650.x (2011).
246 Oliver, P. M., Vass, T., Kappler, J. & Marrack, P. Loss of the proapoptotic protein, Bim, breaks B cell anergy. *J Exp Med* **203**, 731-741, doi:jem.20051407 [pii]
10.1084/jem.20051407 (2006).
247 Rathmell, J. C. *et al.* CD95 (Fas)-dependent elimination of self-reactive B cells upon interaction with CD4+ T cells. *Nature* **376**, 181-184, doi:10.1038/376181a0 (1995).

248 Buggins, A. G. & Pepper, C. J. The role of Bcl-2 family proteins in chronic lymphocytic leukaemia. *Leuk Res* **34**, 837-842, doi:S0145-2126(10)00152-9 [pii]
10.1016/j.leukres.2010.03.011 (2010).
249 Stevenson, F. K., Krysov, S., Davies, A. J., Steele, A. J. & Packham, G. B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **118**, 4313-4320, doi:blood-2011-06-338855 [pii]
10.1182/blood-2011-06-338855 (2011).
250 Batista, F. D. & Harwood, N. E. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol* **9**, 15-27, doi:nri2454 [pii]
10.1038/nri2454 (2009).
251 Colombo, M. *et al.* Intraclonal cell expansion and selection driven by B cell receptor in chronic lymphocytic leukemia. *Mol Med* **17**, 834-839, doi:molmed.2011.00047 [pii]
10.2119/molmed.2011.00047 (2011).
252 Ligler, F. S., Kettman, J. R., Smith, R. G. & Frenkel, E. P. Immunoglobulin phenotype on B cells correlates with clinical stage of chronic lymphocytic leukemia. *Blood* **62**, 256-263 (1983).
253 Rudders, R. A. B lymphocyte subpopulations in chronic lymphocytic leukemia. *Blood* **47**, 229-235 (1976).
254 Geisberger, R., Lamers, M. & Achatz, G. The riddle of the dual expression of IgM and IgD. *Immunology* **118**, 429-437, doi:IMM2386 [pii]
10.1111/j.1365-2567.2006.02386.x (2006).
255 Chen, K. & Cerutti, A. New insights into the enigma of immunoglobulin D. *Immunol Rev* **237**, 160-179, doi:IMR929 [pii]
10.1111/j.1600-065X.2010.00929.x (2010).
256 Litman, G. W., Anderson, M. K. & Rast, J. P. Evolution of antigen binding receptors. *Annu Rev Immunol* **17**, 109-147, doi:10.1146/annurev.immunol.17.1.109 (1999).
257 Magor, B. G., Ross, D. A., Pilstrom, L. & Warr, G. W. Transcriptional enhancers and the evolution of the IgH locus. *Immunol Today* **20**, 13-17, doi:S0167569998013802 [pii] (1999).
258 Rowe, D. S. & Fahey, J. L. A New Class of Human Immunoglobulins. I. A Unique Myeloma Protein. *J Exp Med* **121**, 171-184 (1965).
259 Chen, K. & Cerutti, A. The function and regulation of immunoglobulin D. *Curr Opin Immunol* **23**, 345-352, doi:S0952-7915(11)00007-0 [pii]
10.1016/j.coi.2011.01.006 (2011).
260 Kantor, G. L., Van Herle, A. J. & Barnett, E. V. Auto-antibodies of the IgD class. *Clin Exp Immunol* **6**, 951-962 (1970).
261 Chen, K. *et al.* Immunoglobulin D enhances immune surveillance by activating antimicrobial, proinflammatory and B cell-stimulating programs in basophils. *Nat Immunol* **10**, 889-898, doi:ni.1748 [pii]
10.1038/ni.1748 (2009).
262 Spiegelberg, H. L. The structure and biology of human IgD. *Immunol Rev* **37**, 3-24 (1977).
263 Loset, G. A., Roux, K. H., Zhu, P., Michaelsen, T. E. & Sandlie, I. Differential segmental flexibility and reach dictate the antigen binding mode of chimeric

IgD and IgM: implications for the function of the B cell receptor. *J Immunol* **172**, 2925-2934 (2004).

264 Wienands, J., Hombach, J., Radbruch, A., Riesterer, C. & Reth, M. Molecular components of the B cell antigen receptor complex of class IgD differ partly from those of IgM. *EMBO J* **9**, 449-455 (1990).

265 Wu, Y., Pun, C. & Hozumi, N. Roles of calnexin and Ig-alpha beta interactions with membrane Igs in the surface expression of the B cell antigen receptor of the IgM and IgD classes. *J Immunol* **158**, 2762-2770 (1997).

266 Terashima, M. *et al.* The IgM antigen receptor of B lymphocytes is associated with prohibitin and a prohibitin-related protein. *EMBO J* **13**, 3782-3792 (1994).

267 Adachi, T. *et al.* The specificity of association of the IgD molecule with the accessory proteins BAP31/BAP29 lies in the IgD transmembrane sequence. *EMBO J* **15**, 1534-1541 (1996).

268 Wienands, J. & Reth, M. The B cell antigen receptor of class IgD can be expressed on the cell surface in two different forms. *Eur J Immunol* **21**, 2373-2378, doi:10.1002/eji.1830211012 (1991).

269 Chaturvedi, A., Siddiqui, Z., Bayiroglu, F. & Rao, K. V. A GPI-linked isoform of the IgD receptor regulates resting B cell activation. *Nat Immunol* **3**, 951-957, doi:10.1038/ni839
ni839 [pii] (2002).

270 Yuan, D., Witte, P. L., Tan, J., Hawley, J. & Dang, T. Regulation of IgM and IgD heavy chain gene expression: effect of abrogation of intergenic transcriptional termination. *J Immunol* **157**, 2073-2081 (1996).

271 Kim, M., Qiu, P., Abuodeh, R., Chen, J. & Yuan, D. Differential regulation of transcription termination occurring at two different sites on the micro-delta gene complex. *Int Immunol* **11**, 813-824 (1999).

272 Ashfield, R. *et al.* MAZ-dependent termination between closely spaced human complement genes. *EMBO J* **13**, 5656-5667 (1994).

273 Havran, W. L., DiGiusto, D. L. & Cambier, J. C. mIgM:mIgD ratios on B cells: mean mIgD expression exceeds mIgM by 10-fold on most splenic B cells. *J Immunol* **132**, 1712-1716 (1984).

274 Yuan, D. Regulation of IgM and IgD synthesis in B lymphocytes. II. Translational and post-translational events. *J Immunol* **132**, 1566-1570 (1984).

275 Nicholson, I. C., Brisco, M. J. & Zola, H. Memory B lymphocytes in human tonsil do not express surface IgD. *J Immunol* **154**, 1105-1113 (1995).

276 Kluin, P. M. *et al.* IgD class switching: identification of a novel recombination site in neoplastic and normal B cells. *Eur J Immunol* **25**, 3504-3508, doi:10.1002/eji.1830251244 (1995).

277 Seifert, M. *et al.* A model for the development of human IgD-only B cells: Genotypic analyses suggest their generation in superantigen driven immune responses. *Mol Immunol* **46**, 630-639, doi:S0161-5890(08)00332-5 [pii]
10.1016/j.molimm.2008.07.032 (2009).

278 Zheng, N. Y. *et al.* Human immunoglobulin selection associated with class switch and possible tolerogenic origins for C delta class-switched B cells. *J Clin Invest* **113**, 1188-1201, doi:10.1172/JCI20255 (2004).

279 Koelsch, K. *et al.* Mature B cells class switched to IgD are autoreactive in healthy individuals. *J Clin Invest* **117**, 1558-1565, doi:10.1172/JCI27628 (2007).

280 Forsgren, A. & Grubb, A. O. Many bacterial species bind human IgD. *J Immunol* **122**, 1468-1472 (1979).
281 Ruan, M. R., Akkoyunlu, M., Grubb, A. & Forsgren, A. Protein D of Haemophilus influenzae. A novel bacterial surface protein with affinity for human IgD. *J Immunol* **145**, 3379-3384 (1990).
282 Forsgren, A. *et al.* Isolation and characterization of a novel IgD-binding protein from Moraxella catarrhalis. *J Immunol* **167**, 2112-2120 (2001).
283 Nguyen, T. G. *et al.* Anti-IgD antibody attenuates collagen-induced arthritis by selectively depleting mature B-cells and promoting immune tolerance. *J Autoimmun* **35**, 86-97, doi:S0896-8411(10)00035-1 [pii]
10.1016/j.jaut.2010.03.003 (2010).
284 Carsetti, R., Kohler, G. & Lamers, M. C. Transitional B cells are the target of negative selection in the B cell compartment. *J Exp Med* **181**, 2129-2140 (1995).
285 Su, T. T. & Rawlings, D. J. Transitional B lymphocyte subsets operate as distinct checkpoints in murine splenic B cell development. *J Immunol* **168**, 2101-2110 (2002).
286 Heltemes, L. M. & Manser, T. Level of B cell antigen receptor surface expression influences both positive and negative selection of B cells during primary development. *J Immunol* **169**, 1283-1292 (2002).
287 Preud'homme, J. L. *et al.* Structural and functional properties of membrane and secreted IgD. *Mol Immunol* **37**, 871-887, doi:S0161589001000062 [pii] (2000).
288 Roes, J. & Rajewsky, K. Immunoglobulin D (IgD)-deficient mice reveal an auxiliary receptor function for IgD in antigen-mediated recruitment of B cells. *J Exp Med* **177**, 45-55 (1993).
289 Nitschke, L., Kosco, M. H., Kohler, G. & Lamers, M. C. Immunoglobulin D-deficient mice can mount normal immune responses to thymus-independent and -dependent antigens. *Proc Natl Acad Sci U S A* **90**, 1887-1891 (1993).
290 Lutz, C. *et al.* IgD can largely substitute for loss of IgM function in B cells. *Nature* **393**, 797-801, doi:10.1038/31716 (1998).
291 Kitamura, D., Roes, J., Kuhn, R. & Rajewsky, K. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin mu chain gene. *Nature* **350**, 423-426, doi:10.1038/350423a0 (1991).
292 Brink, R. *et al.* Immunoglobulin M and D antigen receptors are both capable of mediating B lymphocyte activation, deletion, or anergy after interaction with specific antigen. *J Exp Med* **176**, 991-1005 (1992).
293 Cambier, J. C. & Monroe, J. G. B cell activation. V. Differentiation signaling of B cell membrane depolarization, increased I-A expression, G0 to G1 transition, and thymidine uptake by anti-IgM and anti-IgD antibodies. *J Immunol* **133**, 576-581 (1984).
294 Kim, K. M. & Reth, M. The B cell antigen receptor of class IgD induces a stronger and more prolonged protein tyrosine phosphorylation than that of class IgM. *J Exp Med* **181**, 1005-1014 (1995).
295 Mayumi, M. *et al.* Positive and negative signals transduced through surface immunoglobulins in human B cells. *J Allergy Clin Immunol* **94**, 612-619, doi:a57121 [pii] (1994).
296 Harwood, N. E. & Batista, F. D. Early events in B cell activation. *Annu Rev Immunol* **28**, 185-210, doi:10.1146/annurev-immunol-030409-101216 (2010).

297 Treanor, B. *et al.* The membrane skeleton controls diffusion dynamics and signaling through the B cell receptor. *Immunity* **32**, 187-199, doi:S1074-7613(10)00043-9 [pii]
10.1016/j.immuni.2009.12.005 (2010).
298 Depoil, D. *et al.* CD19 is essential for B cell activation by promoting B cell receptor-antigen microcluster formation in response to membrane-bound ligand. *Nat Immunol* **9**, 63-72, doi:ni1547 [pii]
10.1038/ni1547 (2008).
299 Vilen, B. J., Nakamura, T. & Cambier, J. C. Antigen-stimulated dissociation of BCR mIg from Ig-alpha/Ig-beta: implications for receptor desensitization. *Immunity* **10**, 239-248, doi:S1074-7613(00)80024-2 [pii] (1999).
300 Ales-Martinez, J. E., Warner, G. L. & Scott, D. W. Immunoglobulins D and M mediate signals that are qualitatively different in B cells with an immature phenotype. *Proc Natl Acad Sci U S A* **85**, 6919-6923 (1988).
301 Kim, K. M. & Reth, M. Signaling difference between class IgM and IgD antigen receptors. *Ann N Y Acad Sci* **766**, 81-88 (1995).
302 Mongini, P. K., Blessinger, C., Posnett, D. N. & Rudich, S. M. Membrane IgD and membrane IgM differ in capacity to transduce inhibitory signals within the same human B cell clonal populations. *J Immunol* **143**, 1565-1574 (1989).
303 Tisch, R., Roifman, C. M. & Hozumi, N. Functional differences between immunoglobulins M and D expressed on the surface of an immature B-cell line. *Proc Natl Acad Sci U S A* **85**, 6914-6918 (1988).
304 Matutes, E. *et al.* The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia* **8**, 1640-1645 (1994).
305 Lanham, S. *et al.* Differential signaling via surface IgM is associated with VH gene mutational status and CD38 expression in chronic lymphocytic leukemia. *Blood* **101**, 1087-1093, doi:10.1182/blood-2002-06-1822
2002-06-1822 [pii] (2003).
306 Zupo, S. *et al.* Apoptosis or plasma cell differentiation of CD38-positive B-chronic lymphocytic leukemia cells induced by cross-linking of surface IgM or IgD. *Blood* **95**, 1199-1206 (2000).
307 Zupo, S. *et al.* CD38 expression distinguishes two groups of B-cell chronic lymphocytic leukemias with different responses to anti-IgM antibodies and propensity to apoptosis. *Blood* **88**, 1365-1374 (1996).
308 Pierce, S. K. & Liu, W. The tipping points in the initiation of B cell signalling: how small changes make big differences. *Nat Rev Immunol* **10**, 767-777, doi:nri2853 [pii]
10.1038/nri2853 (2010).
309 Srinivasan, L. *et al.* PI3 kinase signals BCR-dependent mature B cell survival. *Cell* **139**, 573-586, doi:S0092-8674(09)01115-5 [pii]
10.1016/j.cell.2009.08.041 (2009).
310 Fruman, D. A., Satterthwaite, A. B. & Witte, O. N. Xid-like phenotypes: a B cell signalosome takes shape. *Immunity* **13**, 1-3, doi:S1074-7613(00)00002-9 [pii] (2000).
311 Harwood, N. E. & Batista, F. D. The cytoskeleton coordinates the early events of B-cell activation. *Cold Spring Harb Perspect Biol* **3**, doi:cshperspect.a002360 [pii]

10.1101/cshperspect.a002360 (2011).
312 Treanor, B., Depoil, D., Bruckbauer, A. & Batista, F. D. Dynamic cortical actin remodeling by ERM proteins controls BCR microcluster organization and integrity. *J Exp Med* **208**, 1055-1068, doi:jem.20101125 [pii]
10.1084/jem.20101125 (2011).
313 Cambier, J. C. & Getahun, A. B cell activation versus anergy; the antigen receptor as a molecular switch. *Immunol Lett* **128**, 6-7, doi:S0165-2478(09)00235-1 [pii]
10.1016/j.imlet.2009.09.006 (2010).
314 Keshvara, L. M., Isaacson, C., Harrison, M. L. & Geahlen, R. L. Syk activation and dissociation from the B-cell antigen receptor is mediated by phosphorylation of tyrosine 130. *J Biol Chem* **272**, 10377-10381 (1997).
315 Hibbs, M. L. *et al.* Multiple defects in the immune system of Lyn-deficient mice, culminating in autoimmune disease. *Cell* **83**, 301-311, doi:0092-8674(95)90171-X [pii] (1995).
316 Goodnow, C. C. *et al.* Clonal silencing of self-reactive B lymphocytes in a transgenic mouse model. *Cold Spring Harb Symp Quant Biol* **54 Pt 2**, 907-920 (1989).
317 Allsup, D. J. *et al.* B-cell receptor translocation to lipid rafts and associated signaling differ between prognostically important subgroups of chronic lymphocytic leukemia. *Cancer Res* **65**, 7328-7337, doi:65/16/7328 [pii]
10.1158/0008-5472.CAN-03-1563 (2005).
318 Chen, L. *et al.* ZAP-70 directly enhances IgM signaling in chronic lymphocytic leukemia. *Blood* **105**, 2036-2041, doi:2004-05-1715 [pii]
10.1182/blood-2004-05-1715 (2005).
319 Nedellec, S. *et al.* B cell response to surface IgM cross-linking identifies different prognostic groups of B-chronic lymphocytic leukemia patients. *J Immunol* **174**, 3749-3756, doi:174/6/3749 [pii] (2005).
320 Petlickovski, A. *et al.* Sustained signaling through the B-cell receptor induces Mcl-1 and promotes survival of chronic lymphocytic leukemia B cells. *Blood* **105**, 4820-4827, doi:2004-07-2669 [pii]
10.1182/blood-2004-07-2669 (2005).
321 Deglesne, P. A. *et al.* Survival response to B-cell receptor ligation is restricted to progressive chronic lymphocytic leukemia cells irrespective of Zap70 expression. *Cancer Res* **66**, 7158-7166, doi:66/14/7158 [pii]
10.1158/0008-5472.CAN-06-0085 (2006).
322 Guarini, A. *et al.* BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. *Blood* **112**, 782-792, doi:blood-2007-12-127688 [pii]
10.1182/blood-2007-12-127688 (2008).
323 Zomas, A. P. *et al.* Expression of the immunoglobulin-associated protein B29 in B cell disorders with the monoclonal antibody SN8 (CD79b). *Leukemia* **10**, 1966-1970 (1996).
324 Gordon, M. S. *et al.* Aberrant B cell receptor signaling from B29 (Igbeta, CD79b) gene mutations of chronic lymphocytic leukemia B cells. *Proc Natl Acad Sci U S A* **97**, 5504-5509, doi:10.1073/pnas.090087097

090087097 [pii] (2000).
325 Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88-92, doi:nature08638 [pii]
10.1038/nature08638 (2010).
326 Vig, M. & Kinet, J. P. Calcium signaling in immune cells. *Nat Immunol* **10**, 21-27, doi:ni.f.220 [pii]
10.1038/ni.f.220 (2009).
327 Michel, F. *et al.* Defective calcium response in B-chronic lymphocytic leukemia cells. Alteration of early protein tyrosine phosphorylation and of the mechanism responsible for cell calcium influx. *J Immunol* **150**, 3624-3633 (1993).
328 Hivroz, C., Grillot-Courvalin, C., Labaume, S., Miglierina, R. & Brouet, J. C. Cross-linking of membrane IgM on B CLL cells: dissociation between intracellular free Ca2+ mobilization and cell proliferation. *Eur J Immunol* **18**, 1811-1817, doi:10.1002/eji.1830181124 (1988).
329 Lankester, A. C. *et al.* Antigen receptor nonresponsiveness in chronic lymphocytic leukemia B cells. *Blood* **86**, 1090-1097 (1995).
330 Contri, A. *et al.* Chronic lymphocytic leukemia B cells contain anomalous Lyn tyrosine kinase, a putative contribution to defective apoptosis. *J Clin Invest* **115**, 369-378, doi:10.1172/JCI22094 (2005).
331 Hallaert, D. Y. *et al.* c-Abl kinase inhibitors overcome CD40-mediated drug resistance in CLL; Implications for therapeutic targeting of chemoresistant niches. *Blood*, doi:blood-2008-03-146704 [pii]
10.1182/blood-2008-03-146704 (2008).
332 Veldurthy, A. *et al.* The kinase inhibitor dasatinib induces apoptosis in chronic lymphocytic leukemia cells in vitro with preference for a subgroup of patients with unmutated IgVH genes. *Blood* **112**, 1443-1452, doi:blood-2007-11-123984 [pii]
10.1182/blood-2007-11-123984 (2008).
333 Ruhe, J. E. *et al.* Genetic alterations in the tyrosine kinase transcriptome of human cancer cell lines. *Cancer Res* **67**, 11368-11376, doi:67/23/11368 [pii]
10.1158/0008-5472.CAN-07-2703 (2007).
334 Philippen, A. *et al.* SYK carries no activating point mutations in patients with chronic lymphocytic leukaemia (CLL). *Br J Haematol*, doi:BJH8244 [pii]
10.1111/j.1365-2141.2010.08244.x (2010).
335 Baudot, A. D. *et al.* The tyrosine kinase Syk regulates the survival of chronic lymphocytic leukemia B cells through PKCdelta and proteasome-dependent regulation of Mcl-1 expression. *Oncogene* **28**, 3261-3273, doi:onc2009179 [pii]
10.1038/onc.2009.179 (2009).
336 Buchner, M. *et al.* Spleen tyrosine kinase is overexpressed and represents a potential therapeutic target in chronic lymphocytic leukemia. *Cancer Res* **69**, 5424-5432, doi:0008-5472.CAN-08-4252 [pii]
10.1158/0008-5472.CAN-08-4252 (2009).
337 Semichon, M., Merle-Beral, H., Lang, V. & Bismuth, G. Normal Syk protein level but abnormal tyrosine phosphorylation in B-CLL cells. *Leukemia* **11**, 1921-1928 (1997).

338 Guinamard, R. *et al.* B cell antigen receptor engagement inhibits stromal cell-derived factor (SDF)-1alpha chemotaxis and promotes protein kinase C (PKC)-induced internalization of CXCR4. *J Exp Med* **189**, 1461-1466 (1999).

339 Ganju, R. K., Brubaker, S. A., Chernock, R. D., Avraham, S. & Groopman, J. E. Beta-chemokine receptor CCR5 signals through SHP1, SHP2, and Syk. *J Biol Chem* **275**, 17263-17268, doi:10.1074/jbc.M000689200
M000689200 [pii] (2000).

340 Zarbock, A., Lowell, C. A. & Ley, K. Spleen tyrosine kinase Syk is necessary for E-selectin-induced alpha(L)beta(2) integrin-mediated rolling on intercellular adhesion molecule-1. *Immunity* **26**, 773-783, doi:S1074-7613(07)00283-X [pii]
10.1016/j.immuni.2007.04.011 (2007).

341 Quiroga, M. P. *et al.* B-cell antigen receptor signaling enhances chronic lymphocytic leukemia cell migration and survival: specific targeting with a novel spleen tyrosine kinase inhibitor, R406. *Blood* **114**, 1029-1037, doi:blood-2009-03-212837 [pii]
10.1182/blood-2009-03-212837 (2009).

342 Friedberg, J. W. *et al.* Inhibition of Syk with fostamatinib disodium has significant clinical activity in non-Hodgkin lymphoma and chronic lymphocytic leukemia. *Blood* **115**, 2578-2585, doi:blood-2009-08-236471 [pii]
10.1182/blood-2009-08-236471.

343 Rosenwald, A. *et al.* Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med* **194**, 1639-1647 (2001).

344 Wiestner, A. *et al.* ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood* **101**, 4944-4951, doi:10.1182/blood-2002-10-3306
2002-10-3306 [pii] (2003).

345 Palacios, E. H. & Weiss, A. Distinct roles for Syk and ZAP-70 during early thymocyte development. *J Exp Med* **204**, 1703-1715, doi:jem.20070405 [pii]
10.1084/jem.20070405 (2007).

346 Chen, L. *et al.* Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **100**, 4609-4614, doi:10.1182/blood-2002-06-1683
2002-06-1683 [pii] (2002).

347 Latour, S., Chow, L. M. & Veillette, A. Differential intrinsic enzymatic activity of Syk and Zap-70 protein-tyrosine kinases. *J Biol Chem* **271**, 22782-22790 (1996).

348 Gobessi, S. *et al.* ZAP-70 enhances B-cell-receptor signaling despite absent or inefficient tyrosine kinase activation in chronic lymphocytic leukemia and lymphoma B cells. *Blood* **109**, 2032-2039, doi:blood-2006-03-011759 [pii]
10.1182/blood-2006-03-011759 (2007).

349 Salim, K. *et al.* Distinct specificity in the recognition of phosphoinositides by the pleckstrin homology domains of dynamin and Bruton's tyrosine kinase. *EMBO J* **15**, 6241-6250 (1996).

350 Herman, S. E. *et al.* Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood* **117**, 6287-6296, doi:blood-2011-01-328484 [pii]

10.1182/blood-2011-01-328484 (2011).
351 Rooij, M. F. *et al.* The clinically active BTK inhibitor PCI-32765 targets B-cell receptor- and chemokine-controlled adhesion and migration in chronic lymphocytic leukemia. *Blood*, doi:blood-2011-11-390989 [pii]
10.1182/blood-2011-11-390989 (2012).
352 Pleiman, C. M., Hertz, W. M. & Cambier, J. C. Activation of phosphatidylinositol-3' kinase by Src-family kinase SH3 binding to the p85 subunit. *Science* **263**, 1609-1612 (1994).
353 Werner, M., Hobeika, E. & Jumaa, H. Role of PI3K in the generation and survival of B cells. *Immunol Rev* **237**, 55-71, doi:IMR934 [pii]
10.1111/j.1600-065X.2010.00934.x (2010).
354 Tedder, T. F., Zhou, L. J. & Engel, P. The CD19/CD21 signal transduction complex of B lymphocytes. *Immunol Today* **15**, 437-442 (1994).
355 Barragan, M. *et al.* Involvement of protein kinase C and phosphatidylinositol 3-kinase pathways in the survival of B-cell chronic lymphocytic leukemia cells. *Blood* **99**, 2969-2976 (2002).
356 Ringshausen, I. *et al.* Constitutively activated phosphatidylinositol-3 kinase (PI-3K) is involved in the defect of apoptosis in B-CLL: association with protein kinase Cdelta. *Blood* **100**, 3741-3748, doi:10.1182/blood-2002-02-0539
2002-02-0539 [pii] (2002).
357 Leupin, N. *et al.* Disparate expression of the PTEN gene: a novel finding in B-cell chronic lymphocytic leukaemia (B-CLL). *Br J Haematol* **121**, 97-100, doi:4227 [pii] (2003).
358 Lannutti, B. J. *et al.* CAL-101, a p110delta selective phosphatidylinositol-3-kinase inhibitor for the treatment of B-cell malignancies, inhibits PI3K signaling and cellular viability. *Blood* **117**, 591-594, doi:blood-2010-03-275305 [pii]
10.1182/blood-2010-03-275305 (2011).
359 Hoellenriegel, J. *et al.* The phosphoinositide 3'-kinase delta inhibitor, CAL-101, inhibits B-cell receptor signaling and chemokine networks in chronic lymphocytic leukemia. *Blood* **118**, 3603-3612, doi:blood-2011-05-352492 [pii]
10.1182/blood-2011-05-352492 (2011).
360 Woyach, J. A., Johnson, A. J. & Byrd, J. C. The B-cell receptor signaling pathway as a therapeutic target in CLL. *Blood*, doi:blood-2012-02-362624 [pii]
10.1182/blood-2012-02-362624 (2012).
361 Wendel, H. G. *et al.* Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature* **428**, 332-337, doi:10.1038/nature02369
nature02369 [pii] (2004).
362 Zhuang, J. *et al.* Akt is activated in chronic lymphocytic leukemia cells and delivers a pro-survival signal: the therapeutic potential of Akt inhibition. *Haematologica* **95**, 110-118, doi:haematol.2009.010272 [pii]
10.3324/haematol.2009.010272 (2009).
363 Cuni, S. *et al.* A sustained activation of PI3K/NF-kappaB pathway is critical for the survival of chronic lymphocytic leukemia B cells. *Leukemia* **18**, 1391-1400, doi:10.1038/sj.leu.2403398
2403398 [pii] (2004).

364 Abrams, S. T. *et al.* B-cell receptor signaling in chronic lymphocytic leukemia cells is regulated by overexpressed active protein kinase CbetaII. *Blood* **109**, 1193-1201, doi:blood-2006-03-012021 [pii]
10.1182/blood-2006-03-012021 (2007).
365 Su, T. T. *et al.* PKC-beta controls I kappa B kinase lipid raft recruitment and activation in response to BCR signaling. *Nat Immunol* **3**, 780-786, doi:10.1038/ni823
ni823 [pii] (2002).
366 zum Buschenfelde, C. M. *et al.* Recruitment of PKC-betaII to lipid rafts mediates apoptosis-resistance in chronic lymphocytic leukemia expressing ZAP-70. *Leukemia* **24**, 141-152, doi:leu2009216 [pii]
10.1038/leu.2009.216 (2010).
367 Holler, C. *et al.* PKCbeta is essential for the development of chronic lymphocytic leukemia in the TCL1 transgenic mouse model: validation of PKCbeta as a therapeutic target in chronic lymphocytic leukemia. *Blood* **113**, 2791-2794, doi:blood-2008-06-160713 [pii]
10.1182/blood-2008-06-160713 (2009).
368 Chung, E. & Kondo, M. Role of Ras/Raf/MEK/ERK signaling in physiological hematopoiesis and leukemia development. *Immunol Res*, doi:10.1007/s12026-010-8187-5 (2010).
369 Healy, J. I. *et al.* Different nuclear signals are activated by the B cell receptor during positive versus negative signaling. *Immunity* **6**, 419-428, doi:S1074-7613(00)80285-X [pii] (1997).
370 Kawauchi, K., Ogasawara, T. & Yasuyama, M. Activation of extracellular signal-regulated kinase through B-cell antigen receptor in B-cell chronic lymphocytic leukemia. *Int J Hematol* **75**, 508-513 (2002).
371 Marzo, I. *et al.* Farnesyltransferase inhibitor BMS-214662 induces apoptosis in B-cell chronic lymphocytic leukemia cells. *Leukemia* **18**, 1599-1604, doi:10.1038/sj.leu.2403469
2403469 [pii] (2004).
372 Longo, P. G. *et al.* The Akt/Mcl-1 pathway plays a prominent role in mediating antiapoptotic signals downstream of the B-cell receptor in chronic lymphocytic leukemia B cells. *Blood* **111**, 846-855, doi:blood-2007-05-089037 [pii]
10.1182/blood-2007-05-089037 (2008).
373 Karin, M. & Greten, F. R. NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol* **5**, 749-759, doi:nri1703 [pii]
10.1038/nri1703 (2005).
374 Furman, R. R., Asgary, Z., Mascarenhas, J. O., Liou, H. C. & Schattner, E. J. Modulation of NF-kappa B activity and apoptosis in chronic lymphocytic leukemia B cells. *J Immunol* **164**, 2200-2206, doi:ji_v164n4p2200 [pii] (2000).
375 Pickering, B. M. *et al.* Pharmacological inhibitors of NF-kappaB accelerate apoptosis in chronic lymphocytic leukaemia cells. *Oncogene* **26**, 1166-1177, doi:1209897 [pii]
10.1038/sj.onc.1209897 (2007).
376 Ougolkov, A. V., Bone, N. D., Fernandez-Zapico, M. E., Kay, N. E. & Billadeau, D. D. Inhibition of glycogen synthase kinase-3 activity leads to epigenetic silencing

of nuclear factor kappaB target genes and induction of apoptosis in chronic lymphocytic leukemia B cells. *Blood* **110**, 735-742, doi:blood-2006-12-060947 [pii]
10.1182/blood-2006-12-060947 (2007).
377 MacFarlane, M. *et al.* Mechanisms of resistance to TRAIL-induced apoptosis in primary B cell chronic lymphocytic leukaemia. *Oncogene* **21**, 6809-6818, doi:10.1038/sj.onc.1205853 (2002).
378 Edelmann, J. *et al.* Bone marrow fibroblasts induce expression of PI3K/NF-kappaB pathway genes and a pro-angiogenic phenotype in CLL cells. *Leuk Res* **32**, 1565-1572, doi:S0145-2126(08)00126-4 [pii]
10.1016/j.leukres.2008.03.003 (2008).
379 Nishio, M. *et al.* Nurselike cells express BAFF and APRIL, which can promote survival of chronic lymphocytic leukemia cells via a paracrine pathway distinct from that of SDF-1alpha. *Blood* **106**, 1012-1020, doi:2004-03-0889 [pii]
10.1182/blood-2004-03-0889 (2005).
380 Burger, J. A. *et al.* Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1. *Blood* **96**, 2655-2663 (2000).
381 Hertlein, E. & Byrd, J. C. Signalling to drug resistance in CLL. *Best Pract Res Clin Haematol* **23**, 121-131, doi:S1521-6926(10)00008-3 [pii]
10.1016/j.beha.2010.01.007 (2010).
382 Kumar, D. *et al.* Cellular phosphatases facilitate combinatorial processing of receptor-activated signals. *BMC Res Notes* **1**, 81, doi:1756-0500-1-81 [pii]
10.1186/1756-0500-1-81 (2008).
383 Dal Porto, J. M. *et al.* B cell antigen receptor signaling 101. *Mol Immunol* **41**, 599-613, doi:10.1016/j.molimm.2004.04.008
S0161589004001245 [pii] (2004).
384 Matutes, E., Wotherspoon, A. & Catovsky, D. Differential diagnosis in chronic lymphocytic leukaemia. *Best Pract Res Clin Haematol* **20**, 367-384, doi:S1521-6926(07)00022-9 [pii]
10.1016/j.beha.2007.03.001 (2007).
385 Gabelloni, M. L. *et al.* SHIP-1 protein level and phosphorylation status differs between CLL cells segregated by ZAP-70 expression. *Br J Haematol* **140**, 117-119, doi:BJH6891 [pii]
10.1111/j.1365-2141.2007.06891.x (2008).
386 Chen, L., Juszczynski, P., Takeyama, K., Aguiar, R. C. & Shipp, M. A. Protein tyrosine phosphatase receptor-type O truncated (PTPROt) regulates SYK phosphorylation, proximal B-cell-receptor signaling, and cellular proliferation. *Blood* **108**, 3428-3433, doi:blood-2006-03-013821 [pii]
10.1182/blood-2006-03-013821 (2006).
387 Juszczynski, P. *et al.* BCL6 modulates tonic BCR signaling in diffuse large B-cell lymphomas by repressing the SYK phosphatase, PTPROt. *Blood* **114**, 5315-5321, doi:blood-2009-02-204362 [pii]
10.1182/blood-2009-02-204362 (2009).
388 Negro, R. *et al.* Overexpression of the autoimmunity-associated phosphatase PTPN22 promotes survival of antigen-stimulated chronic lymphocytic leukemia

cells by selectively activating the AKT pathway. *Blood*, doi:blood-2012-01-403162 [pii]
10.1182/blood-2012-01-403162 (2012).
389 Menard, L. *et al.* The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. *J Clin Invest*, doi:45790 [pii]
10.1172/JCI45790 (2011).
390 Howard, M. *et al.* Formation and hydrolysis of cyclic ADP-ribose catalyzed by lymphocyte antigen CD38. *Science* **262**, 1056-1059 (1993).
391 Malavasi, F. *et al.* Human CD38: a glycoprotein in search of a function. *Immunol Today* **15**, 95-97 (1994).
392 Deaglio, S. *et al.* Human CD38 (ADP-ribosyl cyclase) is a counter-receptor of CD31, an Ig superfamily member. *J Immunol* **160**, 395-402 (1998).
393 Deaglio, S., Aydin, S., Vaisitti, T., Bergui, L. & Malavasi, F. CD38 at the junction between prognostic marker and therapeutic target. *Trends Mol Med* **14**, 210-218, doi:S1471-4914(08)00081-6 [pii]
10.1016/j.molmed.2008.02.005 (2008).
394 Elgueta, R. *et al.* Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunol Rev* **229**, 152-172, doi:IMR782 [pii]
10.1111/j.1600-065X.2009.00782.x (2009).
395 Willimott, S., Baou, M., Huf, S., Deaglio, S. & Wagner, S. D. Regulation of CD38 in proliferating chronic lymphocytic leukemia cells stimulated with CD154 and interleukin-4. *Haematologica* **92**, 1359-1366, doi:92/10/1359 [pii]
10.3324/haematol.11340 (2007).
396 Schattner, E. J. *et al.* Chronic lymphocytic leukemia B cells can express CD40 ligand and demonstrate T-cell type costimulatory capacity. *Blood* **91**, 2689-2697 (1998).
397 Pham, L. V., Tamayo, A. T., Yoshimura, L. C., Lin-Lee, Y. C. & Ford, R. J. Constitutive NF-kappaB and NFAT activation in aggressive B-cell lymphomas synergistically activates the CD154 gene and maintains lymphoma cell survival. *Blood* **106**, 3940-3947, doi:2005-03-1167 [pii]
10.1182/blood-2005-03-1167 (2005).
398 Romano, M. F. *et al.* Triggering of CD40 antigen inhibits fludarabine-induced apoptosis in B chronic lymphocytic leukemia cells. *Blood* **92**, 990-995 (1998).
399 Minuzzo, S. *et al.* CD40 activation of B-CLL cells is associated with augmented intracellular levels of CD79b and increased BCR expression in a subset of patients. *Leukemia* **19**, 1099-1101, doi:2403772 [pii]
10.1038/sj.leu.2403772 (2005).
400 Herling, M. *et al.* High TCL1 levels are a marker of B-cell receptor pathway responsiveness and adverse outcome in chronic lymphocytic leukemia. *Blood* **114**, 4675-4686, doi:blood-2009-03-208256 [pii]
10.1182/blood-2009-03-208256 (2009).
401 Mackus, W. J. *et al.* Prevention of B cell antigen receptor-induced apoptosis by ligation of CD40 occurs downstream of cell cycle regulation. *Int Immunol* **14**, 973-982 (2002).

402 Hayden, R. E., Pratt, G., Drayson, M. T. & Bunce, C. M. Lycorine sensitizes CD40 ligand-protected chronic lymphocytic leukemia cells to bezafibrate- and medroxyprogesterone acetate-induced apoptosis but dasatanib does not overcome reported CD40-mediated drug resistance. *Haematologica* **95**, 1889-1896, doi:haematol.2010.027821 [pii]
10.3324/haematol.2010.027821 (2010).
403 Farahani, M. *et al.* Autocrine VEGF mediates the antiapoptotic effect of CD154 on CLL cells. *Leukemia* **19**, 524-530, doi:2403631 [pii]
10.1038/sj.leu.2403631 (2005).
404 Jak, M., van Bochove, G. G., van Lier, R. A., Eldering, E. & van Oers, M. H. CD40 stimulation sensitizes CLL cells to rituximab-induced cell death. *Leukemia*, doi:leu201139 [pii]
10.1038/leu.2011.39 (2011).
405 Bleul, C. C., Schultze, J. L. & Springer, T. A. B lymphocyte chemotaxis regulated in association with microanatomic localization, differentiation state, and B cell receptor engagement. *J Exp Med* **187**, 753-762 (1998).
406 Buchner, M. *et al.* Spleen tyrosine kinase inhibition prevents chemokine- and integrin-mediated stromal protective effects in chronic lymphocytic leukemia. *Blood* **115**, 4497-4506, doi:blood-2009-07-233692 [pii]
10.1182/blood-2009-07-233692 (2010).
407 Monroe, J. G. & Keir, M. E. Bridging Toll-like- and B cell-receptor signaling: meet me at the autophagosome. *Immunity* **28**, 729-731, doi:S1074-7613(08)00243-4 [pii]
10.1016/j.immuni.2008.05.006 (2008).
408 Dufner, A. & Schamel, W. W. B cell antigen receptor-induced activation of an IRAK4-dependent signaling pathway revealed by a MALT1-IRAK4 double knockout mouse model. *Cell Commun Signal* **9**, 6, doi:1478-811X-9-6 [pii]
10.1186/1478-811X-9-6 (2011).
409 Bernasconi, N. L., Onai, N. & Lanzavecchia, A. A role for Toll-like receptors in acquired immunity: up-regulation of TLR9 by BCR triggering in naive B cells and constitutive expression in memory B cells. *Blood* **101**, 4500-4504, doi:10.1182/blood-2002-11-3569
2002-11-3569 [pii] (2003).
410 Himmelweit, F. *Collectd Papers of Paul Ehrlich*. 253 (Pergamon, 1960).
411 Kreeger, P. K. & Lauffenburger, D. A. Cancer systems biology: a network modeling perspective. *Carcinogenesis* **31**, 2-8, doi:bgp261 [pii]
10.1093/carcin/bgp261 (2010).
412 Yap, T. A., Gerlinger, M., Futreal, P. A., Pusztai, L. & Swanton, C. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med* **4**, 127ps110, doi:4/127/127ps10 [pii]
10.1126/scitranslmed.3003854 (2012).
413 Fox, E. J., Salk, J. J. & Loeb, L. A. Cancer genome sequencing--an interim analysis. *Cancer Res* **69**, 4948-4950, doi:0008-5472.CAN-09-1231 [pii]
10.1158/0008-5472.CAN-09-1231 (2009).
414 Domenech, E. *et al.* New mutations in chronic lymphocytic leukemia identified by target enrichment and deep sequencing. *PLoS One* **7**, e38158, doi:10.1371/journal.pone.0038158

PONE-D-12-00287 [pii] (2012).
415 Brown, J. R. *et al.* Systematic genomic screen for tyrosine kinase mutations in CLL. *Leukemia* **22**, 1966-1969, doi:leu2008222 [pii]
10.1038/leu.2008.222 (2008).
416 Zhang, X. *et al.* Sequence analysis of 515 kinase genes in chronic lymphocytic leukemia. *Leukemia* **25**, 1908-1910, doi:leu2011163 [pii]
10.1038/leu.2011.163 (2011).
417 Bild, A. H., Potti, A. & Nevins, J. R. Linking oncogenic pathways with therapeutic opportunities. *Nat Rev Cancer* **6**, 735-741, doi:nrc1976 [pii]
10.1038/nrc1976 (2006).
418 Kapoor, A. & Figlin, R. A. Targeted inhibition of mammalian target of rapamycin for the treatment of advanced renal cell carcinoma. *Cancer* **115**, 3618-3630, doi:10.1002/cncr.24409 (2009).
419 Pawson, T. & Warner, N. Oncogenic re-wiring of cellular signaling pathways. *Oncogene* **26**, 1268-1275, doi:1210255 [pii]
10.1038/sj.onc.1210255 (2007).
420 Tan, C. S. *et al.* Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci Signal* **2**, ra39, doi:2/81/ra39 [pii]
10.1126/scisignal.2000316 (2009).
421 Pieroni, E. *et al.* Protein networking: insights into global functional organization of proteomes. *Proteomics* **8**, 799-816, doi:10.1002/pmic.200700767 (2008).
422 Mani, K. M. *et al.* A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol Syst Biol* **4**, 169, doi:msb20082 [pii]
10.1038/msb.2008.2 (2008).
423 Vallat, L. D., Park, Y., Li, C. & Gribben, J. G. Temporal genetic program following B-cell receptor cross-linking: altered balance between proliferation and death in healthy and malignant B cells. *Blood* **109**, 3989-3997, doi:blood-2006-09-045377 [pii]
10.1182/blood-2006-09-045377 (2007).
424 Itadani, H., Mizuarai, S. & Kotani, H. Can systems biology understand pathway activation? Gene expression signatures as surrogate markers for understanding the complexity of pathway activation. *Curr Genomics* **9**, 349-360, doi:10.2174/138920208785133235 (2008).
425 Kolch, W. & Pitt, A. Functional proteomics to dissect tyrosine kinase signalling pathways in cancer. *Nat Rev Cancer* **10**, 618-629, doi:nrc2900 [pii]
10.1038/nrc2900 (2010).
426 Pflieger, D., Gonnet, F., de la Fuente van Bentem, S., Hirt, H. & de la Fuente, A. Linking the proteins--elucidation of proteome-scale networks using mass spectrometry. *Mass Spectrom Rev* **30**, 268-297, doi:10.1002/mas.20278 (2011).
427 Kozuka-Hata, H., Tasaki, S. & Oyama, M. Phosphoproteomics-based systems analysis of signal transduction networks. *Front Physiol* **2**, 113, doi:10.3389/fphys.2011.00113 (2011).
428 Nita-Lazar, A., Saito-Benz, H. & White, F. M. Quantitative phosphoproteomics by mass spectrometry: past, present, and future. *Proteomics* **8**, 4433-4443, doi:10.1002/pmic.200800231 (2008).

429 Hunter, T. & Sefton, B. M. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc Natl Acad Sci U S A* **77**, 1311-1315 (1980).
430 Lin, J., Xie, Z., Zhu, H. & Qian, J. Understanding protein phosphorylation on a systems level. *Brief Funct Genomics* **9**, 32-42, doi:elp045 [pii]
10.1093/bfgp/elp045 (2010).
431 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934, doi:10.1126/science.1075762
298/5600/1912 [pii] (2002).
432 Alonso, A. *et al.* Protein tyrosine phosphatases in the human genome. *Cell* **117**, 699-711, doi:10.1016/j.cell.2004.05.018
S0092867404005343 [pii] (2004).
433 Grimsrud, P. A., Swaney, D. L., Wenger, C. D., Beauchene, N. A. & Coon, J. J. Phosphoproteomics for the masses. *ACS Chem Biol* **5**, 105-119, doi:10.1021/cb900277e (2010).
434 Kantarjian, H. *et al.* Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *N Engl J Med* **346**, 645-652, doi:10.1056/NEJMoa011573
346/9/645 [pii] (2002).
435 Huang, P. H. *et al.* Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proc Natl Acad Sci U S A* **104**, 12867-12872, doi:0705158104 [pii]
10.1073/pnas.0705158104 (2007).
436 Blagoev, B., Ong, S. E., Kratchmarova, I. & Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol* **22**, 1139-1145, doi:10.1038/nbt1005
nbt1005 [pii] (2004).
437 Zhang, Y. *et al.* Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol Cell Proteomics* **4**, 1240-1250, doi:M500089-MCP200 [pii]
10.1074/mcp.M500089-MCP200 (2005).
438 Rabinovitch, P. S., June, C. H., Grossmann, A. & Ledbetter, J. A. Heterogeneity among T cells in intracellular free calcium responses after mitogen stimulation with PHA or anti-CD3. Simultaneous use of indo-1 and immunofluorescence with flow cytometry. *J Immunol* **137**, 952-961 (1986).
439 Montoya, A., Beltran, L., Casado, P., Rodriguez-Prados, J. C. & Cutillas, P. R. Characterization of a TiO enrichment method for label-free quantitative phosphoproteomics. *Methods* **54**, 370-378, doi:S1046-2023(11)00039-9 [pii]
10.1016/j.ymeth.2011.02.004 (2011).
440 Casado, P. & Cutillas, P. R. A self-validating quantitative mass spectrometry method for assessing the accuracy of high-content phosphoproteomic experiments. *Mol Cell Proteomics* **10**, M110 003079, doi:M110.003079 [pii]
10.1074/mcp.M110.003079 (2011).
441 Savitski, M. M. *et al.* Confident phosphorylation site localization using the Mascot Delta Score. *Mol Cell Proteomics* **10**, M110 003830, doi:M110.003830 [pii]

10.1074/mcp.M110.003830 (2011).
442 Cutillas, P. R. & Vanhaesebroeck, B. Quantitative profile of five murine core proteomes using label-free functional proteomics. *Mol Cell Proteomics* **6**, 1560-1573, doi:M700037-MCP200 [pii]
10.1074/mcp.M700037-MCP200 (2007).
443 Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354 (1979).
444 Gassmann, M., Grenacher, B., Rohde, B. & Vogel, J. Quantifying Western blots: pitfalls of densitometry. *Electrophoresis* **30**, 1845-1855, doi:10.1002/elps.200800720 (2009).
445 Hamblin, T. & Hough, D. Chronic lymphatic leukaemia: correlation of immunofluorescent characteristics and clinical features. *Br J Haematol* **36**, 359-365 (1977).
446 Baldini, L. *et al.* Prognostic significance of immunoglobulin phenotype in B cell chronic lymphocytic leukemia. *Blood* **65**, 340-344 (1985).
447 Kimby, E., Mellstedt, H., Bjorkholm, M. & Holm, G. Surface immunoglobulin pattern of the leukaemic cell population in chronic lymphocytic leukaemia (CLL) in relation to disease activity. *Hematol Oncol* **3**, 261-269 (1985).
448 Hamblin, T. J., Oscier, D. G., Stevens, J. R. & Smith, J. L. Long survival in B-CLL correlates with surface IgM kappa phenotype. *Br J Haematol* **66**, 21-26 (1987).
449 Geisler, C. H. *et al.* Prognostic importance of flow cytometric immunophenotyping of 540 consecutive patients with B-cell chronic lymphocytic leukemia. *Blood* **78**, 1795-1802 (1991).
450 Shen, P. U., Fuller, S. G., Rezuke, W. N., Sherburne, B. J. & DiGiuseppe, J. A. Laboratory, morphologic, and immunophenotypic correlates of surface immunoglobulin heavy chain isotype expression in B-cell chronic lymphocytic leukemia. *Am J Clin Pathol* **116**, 905-912, doi:10.1309/1TYF-VPM9-CQ2C-B0TF (2001).
451 Marantidou, F. *et al.* Activation-induced cytidine deaminase splicing patterns in chronic lymphocytic leukemia. *Blood Cells Mol Dis* **44**, 262-267, doi:S1079-9796(09)00243-5 [pii]
10.1016/j.bcmd.2009.12.005.
452 Dono, M., Hashimoto, S., Ferrarini, M. & Chiorazzi, N. In vivo isotype class switching in CD5+ chronic lymphocytic leukemia B cells. *Ann N Y Acad Sci* **764**, 478-481 (1995).
453 Efremov, D. G., Ivanovski, M., Batista, F. D., Pozzato, G. & Burrone, O. R. IgM-producing chronic lymphocytic leukemia cells undergo immunoglobulin isotype-switching without acquiring somatic mutations. *J Clin Invest* **98**, 290-298, doi:10.1172/JCI118792 (1996).
454 Matolcsy, A., Casali, P., Nador, R. G., Liu, Y. F. & Knowles, D. M. Molecular characterization of IgA- and/or IgG-switched chronic lymphocytic leukemia B cells. *Blood* **89**, 1732-1739 (1997).
455 Oppezzo, P. *et al.* Chronic lymphocytic leukemia B cells expressing AID display dissociation between class switch recombination and somatic hypermutation. *Blood* **101**, 4029-4032, doi:10.1182/blood-2002-10-3175
2002-10-3175 [pii] (2003).

456 Kumagai, K., Abo, T., Sekizawa, T. & Sasaki, M. Studies of surface immunoglobulins on human B lymphocytes. I. Dissociation of cell-bound immunoglobulins with acid pH or at 37 degrees C. *J Immunol* **115**, 982-987 (1975).
457 Wu, C. J., Karttunen, J. T., Chin, D. H., Sen, D. & Gilbert, W. Murine memory B cells are multi-isotype expressors. *Immunology* **72**, 48-55 (1991).
458 Dono, M. *et al.* Evidence for progenitors of chronic lymphocytic leukemia B cells that undergo intraclonal differentiation and diversification. *Blood* **87**, 1586-1594 (1996).
459 Cerutti, A. *et al.* Ongoing in vivo immunoglobulin class switch DNA recombination in chronic lymphocytic leukemia B cells. *J Immunol* **169**, 6594-6603 (2002).
460 Palacios, F. *et al.* High expression of AID and active class switch recombination might account for a more aggressive disease in unmutated CLL patients: link with an activated microenvironment in CLL disease. *Blood* **115**, 4488-4496, doi:blood-2009-12-257758 [pii]
10.1182/blood-2009-12-257758 (2010).
461 Oppezzo, P. *et al.* Do CLL B cells correspond to naive or memory B-lymphocytes? Evidence for an active Ig switch unrelated to phenotype expression and Ig mutational pattern in B-CLL cells. *Leukemia* **16**, 2438-2446, doi:10.1038/sj.leu.2402731 (2002).
462 Mann, H. & Whitney, D. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Annals of Mathematical Statistics* **18**, 50-60 (1947).
463 Wilcoxon, F. Individual Comparisons by ranking methods. *Biometrics Bulletin* **1**, 80-83 (1945).
464 Kaplan, E. & Meier, P. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assn.* **53**, 457-481 (1958).
465 Mantel, N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* **50**, 163-170 (1966).
466 Bhayat, F., Das-Gupta, E., Smith, C., McKeever, T. & Hubbard, R. The incidence of and mortality from leukaemias in the UK: a general population-based study. *BMC Cancer* **9**, 252, doi:1471-2407-9-252 [pii]
10.1186/1471-2407-9-252 (2009).
467 Tam, C. S. *et al.* De novo deletion 17p13.1 chronic lymphocytic leukemia shows significant clinical heterogeneity: the M. D. Anderson and Mayo Clinic experience. *Blood* **114**, 957-964, doi:blood-2009-03-210591 [pii]
10.1182/blood-2009-03-210591 (2009).
468 Chiorazzi, N. & Ferrarini, M. Evolving view of the in-vivo kinetics of chronic lymphocytic leukemia B cells. *Hematology Am Soc Hematol Educ Program*, 273-278, 512, doi:2006/1/273 [pii]
10.1182/asheducation-2006.1.273 (2006).
469 Ponzoni, M., Doglioni, C. & Caligaris-Cappio, F. Chronic lymphocytic leukemia: the pathologist's view of lymph node microenvironment. *Semin Diagn Pathol* **28**, 161-166 (2011).

470 Messmer, B. T. *et al.* In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. *J Clin Invest* **115**, 755-764, doi:10.1172/JCI23409 (2005).
471 Patten, P. E. *et al.* CD38 expression in chronic lymphocytic leukemia is regulated by the tumor microenvironment. *Blood* **111**, 5173-5181, doi:blood-2007-08-108605 [pii]
10.1182/blood-2007-08-108605 (2008).
472 Smit, L. A. *et al.* Differential Noxa/Mcl-1 balance in peripheral versus lymph node chronic lymphocytic leukemia cells correlates with survival capacity. *Blood* **109**, 1660-1668, doi:blood-2006-05-021683 [pii]
10.1182/blood-2006-05-021683 (2007).
473 Ben-Ezra, J. *et al.* Small lymphocytic lymphoma: a clinicopathologic analysis of 268 cases. *Blood* **73**, 579-587 (1989).
474 Gine, E. *et al.* Expanded and highly active proliferation centers identify a histological subtype of chronic lymphocytic leukemia ("accelerated" chronic lymphocytic leukemia) with aggressive clinical behavior. *Haematologica* **95**, 1526-1533, doi:haematol.2010.022277 [pii]
10.3324/haematol.2010.022277 (2010).
475 Morabito, F. *et al.* More on the determination of Ki-67 as a novel potential prognostic marker in B-cell chronic lymphocytic leukemia. *Leuk Res*, doi:S0145-2126(10)00379-6 [pii]
10.1016/j.leukres.2010.07.036 (2010).
476 Ciccone, M. *et al.* Proliferation centers in chronic lymphocytic leukemia: correlation with cytogenetic and clinicobiological features in consecutive patients analyzed on tissue microarrays. *Leukemia*, doi:leu2011247 [pii]
10.1038/leu.2011.247 (2011).
477 Balogh, Z. *et al.* High rate of neoplastic cells with genetic abnormalities in proliferation centers of chronic lymphocytic leukemia. *Leuk Lymphoma* **52**, 1080-1084, doi:10.3109/10428194.2011.555889 (2011).
478 Inamdar, K. V. & Bueso-Ramos, C. E. Pathology of chronic lymphocytic leukemia: an update. *Ann Diagn Pathol* **11**, 363-389, doi:S1092-9134(07)00125-6 [pii]
10.1016/j.anndiagpath.2007.08.002 (2007).
479 van Gent, R. *et al.* In vivo dynamics of stable chronic lymphocytic leukemia inversely correlate with somatic hypermutation levels and suggest no major leukemic turnover in bone marrow. *Cancer Res* **68**, 10137-10144, doi:68/24/10137 [pii]
10.1158/0008-5472.CAN-08-2325 (2008).
480 Rodriguez, A. *et al.* Variability in the degree of expression of phosphorylated IkappaBalpha in chronic lymphocytic leukemia cases with nodal involvement. *Clin Cancer Res* **10**, 6796-6806, doi:10/20/6796 [pii]
10.1158/1078-0432.CCR-04-0753 (2004).
481 Herreros, B. *et al.* Proliferation centers in chronic lymphocytic leukemia: the niche where NF-kappaB activation takes place. *Leukemia* **24**, 872-876, doi:leu2009285 [pii]
10.1038/leu.2009.285 (2010).

482 Calissano, C. *et al.* Intra-clonal complexity in chronic lymphocytic leukemia: fractions enriched in recently born/divided and older/quiescent cells. *Mol Med*, doi:molmed.2011.00360 [pii]
10.2119/molmed.2011.00360 (2011).
483 Kurosaki, T. Regulation of BCR signaling. *Mol Immunol* **48**, 1287-1291, doi:S0161-5890(10)00664-4 [pii]
10.1016/j.molimm.2010.12.007 (2011).
484 Chen, L. *et al.* ZAP-70 enhances IgM signaling independent of its kinase activity in chronic lymphocytic leukemia. *Blood* **111**, 2685-2692, doi:blood-2006-12-062265 [pii]
10.1182/blood-2006-12-062265 (2008).
485 Krysov, S. *et al.* Surface IgM of CLL cells displays unusual glycans indicative of engagement of antigen in vivo. *Blood* **115**, 4198-4205, doi:blood-2009-12-254847 [pii]
10.1182/blood-2009-12-254847 (2010).
486 Gauld, S. B., Benschop, R. J., Merrell, K. T. & Cambier, J. C. Maintenance of B cell anergy requires constant antigen receptor occupancy and signaling. *Nat Immunol* **6**, 1160-1167, doi:ni1256 [pii]
10.1038/ni1256 (2005).
487 Gobessi, S. *et al.* Inhibition of constitutive and BCR-induced Syk activation downregulates Mcl-1 and induces apoptosis in chronic lymphocytic leukemia B cells. *Leukemia* **23**, 686-697, doi:leu2008346 [pii]
10.1038/leu.2008.346 (2009).
488 Cragg, M. S. *et al.* The alternative transcript of CD79b is overexpressed in B-CLL and inhibits signaling for apoptosis. *Blood* **100**, 3068-3076, doi:10.1182/blood.V100.9.3068 (2002).
489 Minuzzo, S. *et al.* Heterogeneous intracellular expression of B-cell receptor components in B-cell chronic lymphocytic leukaemia (B-CLL) cells and effects of CD79b gene transfer on surface immunoglobulin levels in a B-CLL-derived cell line. *Br J Haematol* **130**, 878-889, doi:BJH5699 [pii]
10.1111/j.1365-2141.2005.05699.x (2005).
490 Rabinovitch, P. & June, C. in *Flow Cytometry* (ed M Ormerod) (Oxford University Press, 2005).
491 Andreeff, M., Darzynkiewicz, Z., Sharpless, T. K., Clarkson, B. D. & Melamed, M. R. Discrimination of human leukemia subtypes by flow cytometric analysis of cellular DNA and RNA. *Blood* **55**, 282-293 (1980).
492 Kardava, L. *et al.* The B lineage transcription factor E2A regulates apoptosis in chronic lymphocytic leukemia (CLL) cells. *Int Immunol*, doi:dxr027 [pii]
10.1093/intimm/dxr027 (2011).
493 Cao, X. *et al.* The expression of SOX11, cyclin D1, cyclin D2, and cyclin D3 in B-cell lymphocytic proliferative diseases. *Med Oncol*, doi:10.1007/s12032-011-9937-5 (2011).
494 Calissano, C. *et al.* In vivo intraclonal and interclonal kinetic heterogeneity in B-cell chronic lymphocytic leukemia. *Blood* **114**, 4832-4842, doi:blood-2009-05-219634 [pii]
10.1182/blood-2009-05-219634 (2009).

495 Lin, T. T. *et al.* Highly purified CD38 sub-populations show no evidence of preferential clonal evolution despite having increased proliferative activity when compared with CD38 sub-populations derived from the same chronic lymphocytic leukaemia patient. *Br J Haematol* **142**, 595-605, doi:BJH7236 [pii]
10.1111/j.1365-2141.2008.07236.x (2008).
496 Obermann, E. C. *et al.* Cell cycle phase distribution analysis in chronic lymphocytic leukaemia: a significant number of cells reside in early G1-phase. *J Clin Pathol* **60**, 794-797, doi:jcp.2006.040956 [pii]
10.1136/jcp.2006.040956 (2007).
497 Damle, R. N. *et al.* CD38 expression labels an activated subset within chronic lymphocytic leukemia clones enriched in proliferating B cells. *Blood* **110**, 3352-3359, doi:blood-2007-04-083832 [pii]
10.1182/blood-2007-04-083832 (2007).
498 Bennett, F. *et al.* B-cell chronic lymphocytic leukaemia cells show specific changes in membrane protein expression during different stages of cell cycle. *Br J Haematol* **139**, 600-604, doi:BJH6790 [pii]
10.1111/j.1365-2141.2007.06790.x (2007).
499 Kerr, J. F., Wyllie, A. H. & Currie, A. R. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**, 239-257 (1972).
500 Elmore, S. Apoptosis: a review of programmed cell death. *Toxicol Pathol* **35**, 495-516, doi:779478428 [pii]
10.1080/01926230701320337 (2007).
501 Chen, L. *et al.* SYK-dependent tonic B-cell receptor signaling is a rational treatment target in diffuse large B-cell lymphoma. *Blood* **111**, 2230-2237, doi:blood-2007-07-100115 [pii]
10.1182/blood-2007-07-100115 (2008).
502 Gururajan, M., Jennings, C. D. & Bondada, S. Cutting edge: constitutive B cell receptor signaling is critical for basal growth of B lymphoma. *J Immunol* **176**, 5715-5719, doi:176/10/5715 [pii] (2006).
503 Cutrona, G. *et al.* Clonal heterogeneity in chronic lymphocytic leukemia cells: superior response to surface IgM cross-linking in CD38, ZAP-70-positive cells. *Haematologica* **93**, 413-422, doi:haematol.11646 [pii]
10.3324/haematol.11646 (2008).
504 Diaz, C. & Schroit, A. J. Role of translocases in the generation of phosphatidylserine asymmetry. *J Membr Biol* **151**, 1-9 (1996).
505 Vermes, I., Haanen, C., Steffens-Nakken, H. & Reutelingsperger, C. A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. *J Immunol Methods* **184**, 39-51, doi:002217599500072I [pii] (1995).
506 Morabito, F. *et al.* Prognostic relevance of in vitro response to cell stimulation via surface IgD in binet stage a CLL. *Br J Haematol* **149**, 160-163, doi:BJH8032 [pii]
10.1111/j.1365-2141.2009.08032.x (2009).
507 Lang, J. *et al.* B cells are exquisitely sensitive to central tolerance and receptor editing induced by ultralow affinity, membrane-bound antigen. *J Exp Med* **184**, 1685-1697 (1996).

508 Mongini, P. K., Blessinger, C. A., Highet, P. F. & Inman, J. K. Membrane IgM-mediated signaling of human B cells. Effect of increased ligand binding site valency on the affinity and concentration requirements for inducing diverse stages of activation. *J Immunol* **148**, 3892-3901 (1992).
509 Chaturvedi, A., Martz, R., Dorward, D., Waisberg, M. & Pierce, S. K. Endocytosed BCRs sequentially regulate MAPK and Akt signaling pathways from intracellular compartments. *Nat Immunol* **12**, 1119-1126, doi:ni.2116 [pii]
10.1038/ni.2116 (2011).
510 Hou, P. *et al.* B cell antigen receptor signaling and internalization are mutually exclusive events. *PLoS Biol* **4**, e200, doi:05-PLBI-RA-1312R3 [pii]
10.1371/journal.pbio.0040200 (2006).
511 Carrasco, Y. R. & Batista, F. D. B cell recognition of membrane-bound antigen: an exquisite way of sensing ligands. *Curr Opin Immunol* **18**, 286-291, doi:S0952-7915(06)00061-6 [pii]
10.1016/j.coi.2006.03.013 (2006).
512 Brunswick, M. *et al.* Picogram quantities of anti-Ig antibodies coupled to dextran induce B cell proliferation. *J Immunol* **140**, 3364-3372 (1988).
513 Damle, R. N. *et al.* T-cell independent, B-cell receptor-mediated induction of telomerase activity differs among IGHV mutation-based subgroups of chronic lymphocytic leukemia patients. *Blood*, doi:blood-2012-02-409110 [pii]
10.1182/blood-2012-02-409110 (2012).
514 Brunswick, M. *et al.* Surface immunoglobulin-mediated B-cell activation in the absence of detectable elevations in intracellular ionized calcium: a model for T-cell-independent B-cell activation. *Proc Natl Acad Sci U S A* **86**, 6724-6728 (1989).
515 Vlad, A. *et al.* Down-regulation of CXCR4 and CD62L in chronic lymphocytic leukemia cells is triggered by B-cell receptor ligation and associated with progressive disease. *Cancer Res* **69**, 6387-6395, doi:0008-5472.CAN-08-4750 [pii]
10.1158/0008-5472.CAN-08-4750 (2009).
516 Kaveri, S. V., Silverman, G. J. & Bayry, J. Natural IgM in Immune Equilibrium and Harnessing Their Therapeutic Potential. *J Immunol* **188**, 939-945, doi:188/3/939 [pii]
10.4049/jimmunol.1102107 (2012).
517 PhosphoSite. *PhosphoSite*, <www.phosphosite.org> (2011).
518 Dengjel, J. *et al.* Quantitative proteomic assessment of very early cellular signaling events. *Nat Biotechnol* **25**, 566-568, doi:nbt1301 [pii]
10.1038/nbt1301 (2007).
519 Olsen, J. V. *et al.* Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635-648, doi:S0092-8674(06)01274-8 [pii]
10.1016/j.cell.2006.09.026 (2006).
520 Seet, B. T., Dikic, I., Zhou, M. M. & Pawson, T. Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol* **7**, 473-483, doi:nrm1960 [pii]
10.1038/nrm1960 (2006).
521 Morley, R. *A guide to methods in the biomedicaal sciences*. (springer, 2005).
522 Beltran, L. & Cutillas, P. R. Advances in phosphopeptide enrichment techniques for phosphoproteomics. *Amino Acids*, doi:10.1007/s00726-012-1288-9 (2012).

523 Mayya, V. & Han, D. K. Phosphoproteomics by mass spectrometry: insights, implications, applications and limitations. *Expert Rev Proteomics* **6**, 605-618, doi:10.1586/epr.09.84 (2009).
524 Rush, J. *et al.* Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol* **23**, 94-101, doi:nbt1046 [pii]
10.1038/nbt1046 (2005).
525 Ficarro, S. B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat Biotechnol* **20**, 301-305, doi:10.1038/nbt0302-301
nbt0302-301 [pii] (2002).
526 Ficarro, S. B. *et al.* Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun Mass Spectrom* **19**, 57-71, doi:10.1002/rcm.1746 (2005).
527 Cutillas, P. R. *et al.* Ultrasensitive and absolute quantification of the phosphoinositide 3-kinase/Akt signal transduction pathway by mass spectrometry. *Proc Natl Acad Sci U S A* **103**, 8959-8964, doi:0602101103 [pii]
10.1073/pnas.0602101103 (2006).
528 Walters, D. K. *et al.* Phosphoproteomic analysis of AML cell lines identifies leukemic oncogenes. *Leuk Res* **30**, 1097-1104, doi:S0145-2126(06)00002-6 [pii]
10.1016/j.leukres.2006.01.001 (2006).
529 Goss, V. L. *et al.* A common phosphotyrosine signature for the Bcr-Abl kinase. *Blood* **107**, 4888-4897, doi:2005-08-3399 [pii]
10.1182/blood-2005-08-3399 (2006).
530 Cecconi, D. *et al.* Signal transduction pathways of mantle cell lymphoma: a phosphoproteome-based study. *Proteomics* **8**, 4495-4506, doi:10.1002/pmic.200800080 (2008).
531 O'Hayre, M. *et al.* Elucidating the CXCL12/CXCR4 signaling network in chronic lymphocytic leukemia through phosphoproteomics analysis. *PLoS ONE* **5**, e11716, doi:10.1371/journal.pone.0011716 (2010).
532 Cao, L. *et al.* Quantitative time-resolved phosphoproteomic analysis of mast cell signaling. *J Immunol* **179**, 5864-5876, doi:179/9/5864 [pii] (2007).
533 Navarro, M. N., Goebel, J., Feijoo-Carnero, C., Morrice, N. & Cantrell, D. A. Phosphoproteomic analysis reveals an intrinsic pathway for the regulation of histone deacetylase 7 that controls the function of cytotoxic T lymphocytes. *Nat Immunol* **12**, 352-361, doi:ni.2008 [pii]
10.1038/ni.2008 (2011).
534 Matsumoto, M. *et al.* Large-scale proteomic analysis of tyrosine-phosphorylation induced by T-cell receptor or B-cell receptor activation reveals new signaling pathways. *Proteomics* **9**, 3549-3563, doi:10.1002/pmic.200900011 (2009).
535 Alcolea, M. P. & Cutillas, P. R. Quantification of protein kinase activities by LC-MS. *Methods Mol Biol* **658**, 325-337, doi:10.1007/978-1-60761-780-8_20 (2010).
536 Bourgon, R., Gentleman, R. & Huber, W. Independent filtering increases detection power for high-throughput experiments. *Proc Natl Acad Sci U S A* **107**, 9546-9551, doi:0914005107 [pii]

10.1073/pnas.0914005107 (2010).
537 Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**, 14863-14868 (1998).
538 Farcomeni, A. A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Stat Methods Med Res* **17**, 347-388, doi:0962280206079046 [pii]
10.1177/0962280206079046 (2008).
539 Jones, L. V. & Tukey, J. W. A sensible formulation of the significance test. *Psychol Methods* **5**, 411-414 (2000).
540 van Iterson, M., Boer, J. M. & Menezes, R. X. Filtering, FDR and power. *BMC Bioinformatics* **11**, 450, doi:1471-2105-11-450 [pii]
10.1186/1471-2105-11-450 (2010).
541 D, S. *Microarray Bioinformatics*. (Cambridge University Press, 2003).
542 Zhou, Y., Cras-Meneur, C., Ohsugi, M., Stormo, G. D. & Permutt, M. A. A global approach to identify differentially expressed genes in cDNA (two-color) microarray experiments. *Bioinformatics* **23**, 2073-2079, doi:btm292 [pii]
10.1093/bioinformatics/btm292 (2007).
543 McLachlin, D. T. & Chait, B. T. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr Opin Chem Biol* **5**, 591-602, doi:S1367-5931(00)00250-7 [pii] (2001).
544 Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat Biotechnol* **21**, 255-261, doi:10.1038/nbt0303-255
nbt0303-255 [pii] (2003).
545 Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D. & Wishart, D. S. MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. *Nucleic Acids Res* **40**, W127-133, doi:gks374 [pii]
10.1093/nar/gks374 (2012).
546 Saldanha, A. J. Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246-3248, doi:10.1093/bioinformatics/bth349
bth349 [pii] (2004).
547 Thomas, P. D. *et al.* PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* **31**, 334-341 (2003).
548 Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. Serial analysis of gene expression. *Science* **270**, 484-487 (1995).
549 Hornbeck, P. V. *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* **40**, D261-270, doi:gkr1122 [pii]
10.1093/nar/gkr1122 (2012).
550 Dinkel, H. *et al.* Phospho.ELM: a database of phosphorylation sites--update 2011. *Nucleic Acids Res* **39**, D261-267, doi:gkq1104 [pii]
10.1093/nar/gkq1104 (2011).
551 Yang, C. Y. *et al.* PhosphoPOINT: a comprehensive human kinase interactome and phospho-protein database. *Bioinformatics* **24**, i14-20, doi:btn297 [pii]
10.1093/bioinformatics/btn297 (2008).

552 Obenauer, J. C., Cantley, L. C. & Yaffe, M. B. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res* **31**, 3635-3641 (2003).
553 Schaab, C. in *Methods in Molecular Biology, Data Mining in Proteomics, Part 1* Vol. Volume 696 41-57 (Humana Press, 2011).
554 Oellerich, T. *et al.* The B-cell antigen receptor signals through a preformed transducer module of SLP65 and CIN85. *EMBO J* **30**, 3620-3634, doi:emboj2011251 [pii]
10.1038/emboj.2011.251 (2011).
555 Cline, M. S. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* **2**, 2366-2382, doi:nprot.2007.324 [pii]
10.1038/nprot.2007.324 (2007).
556 Szklarczyk, D. *et al.* The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **39**, D561-568, doi:gkq973 [pii]
10.1093/nar/gkq973 (2011).
557 Baba, Y. *et al.* Involvement of wiskott-aldrich syndrome protein in B-cell cytoplasmic tyrosine kinase pathway. *Blood* **93**, 2003-2012 (1999).
558 Ruperez, P., Gago-Martinez, A., Burlingame, A. L. & Oses-Prieto, J. A. Quantitative phosphoproteomic analysis reveals a role for serine and threonine kinases in the cytoskeletal reorganization in early T-cell receptor (TCR) activation in human primary T-cells. *Mol Cell Proteomics*, doi:M112.017863 [pii]
10.1074/mcp.M112.017863 (2012).
559 Preisinger, C., von Kriegsheim, A., Matallanas, D. & Kolch, W. Proteomics and phosphoproteomics for the mapping of cellular signalling networks. *Proteomics* **8**, 4402-4415, doi:10.1002/pmic.200800136 (2008).
560 Cirit, M. & Haugh, J. M. Quantitative models of signal transduction networks: How detailed should they be? *Commun Integr Biol* **4**, 353-356, doi:10.4161/cib.4.3.15149
1942-0889-4-3-30 [pii] (2011).
561 Alcolea, M. P., Kleiner, O. & Cutillas, P. R. Increased confidence in large-scale phosphoproteomics data by complementary mass spectrometric techniques and matching of phosphopeptide data sets. *J Proteome Res* **8**, 3808-3815, doi:10.1021/pr800955n (2009).
562 Cutillas, P. R. & Timms, J. F. Approaches and applications of quantitative LC-MS for proteomics and activitomics. *Methods Mol Biol* **658**, 3-17, doi:10.1007/978-1-60761-780-8_1 (2010).
563 Boekhorst, J. *et al.* Evaluating experimental bias and completeness in comparative phosphoproteomics analysis. *PLoS One* **6**, e23276, doi:10.1371/journal.pone.0023276
PONE-D-11-08687 [pii] (2011).
564 Cohen, A. A. *et al.* Dynamic proteomics of individual cancer cells in response to a drug. *Science* **322**, 1511-1516, doi:1160165 [pii]
10.1126/science.1160165 (2008).
565 Duffy, K. R. *et al.* Activation-induced B cell fates are selected by intracellular stochastic competition. *Science* **335**, 338-341, doi:science.1213230 [pii]
10.1126/science.1213230 (2012).

566 Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. & Park, B. K. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**, 1229-1239, doi:S1570-0232(08)00837-4 [pii]
10.1016/j.jchromb.2008.11.013 (2009).
567 Linding, R. *et al.* NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* **36**, D695-699, doi:gkm902 [pii]
10.1093/nar/gkm902 (2008).
568 Gnad, F., Gunawardena, J. & Mann, M. PHOSIDA 2011: the posttranslational modification database. *Nucleic Acids Res* **39**, D253-260, doi:gkq1159 [pii]
10.1093/nar/gkq1159 (2011).
569 Ceol, A. *et al.* MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res* **38**, D532-539, doi:gkp983 [pii]
10.1093/nar/gkp983 (2010).
570 Yachie, N., Saito, R., Sugahara, J., Tomita, M. & Ishihama, Y. In silico analysis of phosphoproteome data suggests a rich-get-richer process of phosphosite accumulation over evolution. *Mol Cell Proteomics* **8**, 1061-1071, doi:M800466-MCP200 [pii]
10.1074/mcp.M800466-MCP200 (2009).
571 Salazar, C. & Hofer, T. Multisite protein phosphorylation--from molecular mechanisms to kinetic models. *FEBS J* **276**, 3177-3198, doi:EJB7027 [pii]
10.1111/j.1742-4658.2009.07027.x (2009).
572 Lienhard, G. E. Non-functional phosphorylations? *Trends Biochem Sci* **33**, 351-352, doi:S0968-0004(08)00127-8 [pii]
10.1016/j.tibs.2008.05.004 (2008).
573 Soh, D., Dong, D., Guo, Y. & Wong, L. Consistency, comprehensiveness, and compatibility of pathway databases. *BMC Bioinformatics* **11**, 449, doi:1471-2105-11-449 [pii]
10.1186/1471-2105-11-449 (2010).
574 Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**, D109-114, doi:gkr988 [pii]
10.1093/nar/gkr988 (2012).
575 Pighi, C. *et al.* Phospho-proteomic analysis of mantle cell lymphoma cells suggests a pro-survival role of B-cell receptor signaling. *Cell Oncol (Dordr)* **34**, 141-153, doi:10.1007/s13402-011-0019-7 (2011).
576 Solomou, E. E. *et al.* 13q deletion in chronic lymphocytic leukemia: characterization of E4.5, a novel chromosome condensation regulator-like guanine nucleotide exchange factor. *Leuk Lymphoma* **44**, 1579-1585, doi:10.3109/10428190309178782 (2003).
577 Caligaris-Cappio, F. *et al.* Cytoskeleton organization is aberrantly rearranged in the cells of B chronic lymphocytic leukemia and hairy cell leukemia. *Blood* **67**, 233-239 (1986).
578 Scielzo, C. *et al.* HS1 has a central role in the trafficking and homing of leukemic B cells. *Blood* **116**, 3537-3546, doi:blood-2009-12-258814 [pii]
10.1182/blood-2009-12-258814 (2010).

579 Brody, J. I., Oski, F. A. & Singer, D. E. Impaired pentose phosphate shunt and decreased glycolytic activity in lymphocytes of chronic lymphocytic leukemia. Metabolic pathway. *Blood* **34**, 421-429 (1969).
580 McBrayer, S. K. *et al.* Integrative Gene Expression Profiling Reveals G6PD-Mediated Resistance to RNA-Directed Nucleoside Analogues in B-Cell Neoplasms. *PLoS One* **7**, e41455, doi:10.1371/journal.pone.0041455
PONE-D-11-08927 [pii] (2012).
581 Chuang, H. Y. *et al.* Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. *Blood*, doi:blood-2012-03-416461 [pii]
10.1182/blood-2012-03-416461 (2012).
582 Weintz, G. *et al.* The phosphoproteome of toll-like receptor-activated macrophages. *Mol Syst Biol* **6**, 371, doi:msb201029 [pii]
10.1038/msb.2010.29 (2010).
583 Brockmeyer, C. *et al.* T Cell Receptor (TCR)-induced Tyrosine Phosphorylation Dynamics Identifies THEMIS as a New TCR Signalosome Component. *J Biol Chem* **286**, 7535-7547, doi:M110.201236 [pii]
10.1074/jbc.M110.201236 (2011).
584 Weigand, S. *et al.* Global quantitative phosphoproteome analysis of human tumor xenografts treated with a CD44 antagonist. *Cancer Res*, doi:0008-5472.CAN-12-0136 [pii]
10.1158/0008-5472.CAN-12-0136 (2012).
585 UniProt. *UniprotKB*, <www.uniprot.org> (2012).
586 Yasuda, T. *et al.* Erk kinases link pre-B cell receptor signaling to transcriptional events required for early B cell expansion. *Immunity* **28**, 499-508, doi:S1074-7613(08)00111-8 [pii]
10.1016/j.immuni.2008.02.015 (2008).
587 Kumar, N., Wolf-Yadlin, A., White, F. M. & Lauffenburger, D. A. Modeling HER2 effects on cell behavior from mass spectrometry phosphotyrosine data. *PLoS Comput Biol* **3**, e4, doi:06-PLCB-RA-0339R2 [pii]
10.1371/journal.pcbi.0030004 (2007).
588 Kumar, N., Afeyan, R., Kim, H. D. & Lauffenburger, D. A. Multipathway model enables prediction of kinase inhibitor cross-talk effects on migration of Her2-overexpressing mammary epithelial cells. *Mol Pharmacol* **73**, 1668-1678, doi:mol.107.043794 [pii]
10.1124/mol.107.043794 (2008).
589 Tibaldi, E. *et al.* Lyn-mediated SHP-1 recruitment to CD5 contributes to resistance to apoptosis of B-cell chronic lymphocytic leukemia cells. *Leukemia* **25**, 1768-1781, doi:leu2011152 [pii]
10.1038/leu.2011.152 (2011).
590 Christensen, D. J. *et al.* SET oncoprotein overexpression in B-cell chronic lymphocytic leukemia and non-Hodgkin's lymphoma: a predictor of aggressive disease and new treatment target. *Blood*, doi:blood-2011-04-351072 [pii]
10.1182/blood-2011-04-351072 (2011).
591 Zikherman, J., Doan, K., Parameswaran, R., Raschke, W. & Weiss, A. Quantitative differences in CD45 expression unmask functions for CD45 in B-

cell development, tolerance, and survival. *Proc Natl Acad Sci U S A* **109**, E3-12, doi:1117374108 [pii]
10.1073/pnas.1117374108 (2012).
592 Tsuda, H. *et al.* Myelo-Monocytoid Immunophenotypes of B-Cell Chronic Lymphocytic-Leukemia Cells. *Internal Med* **32**, 533-539, doi:Doi 10.2169/Internalmedicine.32.533 (1993).
593 Tavolaro, S. *et al.* Gene expression profile of protein kinases reveals a distinctive signature in chronic lymphocytic leukemia and in vitro experiments support a role of second generation protein kinase inhibitors. *Leuk Res* **34**, 733-741, doi:S0145-2126(09)00527-X [pii]
10.1016/j.leukres.2009.11.005 (2010).
594 Piazza, F. *et al.* Protein kinase CK2 in hematologic malignancies: reliance on a pivotal cell survival regulator by oncogenic signaling pathways. *Leukemia* **26**, 1174-1179, doi:leu2011385 [pii]
10.1038/leu.2011.385 (2012).
595 de Paula Careta, F. *et al.* The Aurora A and B kinases are up-regulated in bone marrow-derived chronic lymphocytic leukemia cells and represent potential therapeutic targets. *Haematologica* **97**, 1246-1254, doi:haematol.2011.054668 [pii]
10.3324/haematol.2011.054668 (2012).
596 Ringshausen, I., Peschel, C. & Decker, T. Mammalian target of rapamycin (mTOR) inhibition in chronic lymphocytic B-cell leukemia: A new therapeutic option. *Leukemia Lymphoma* **46**, 11-19, doi:Doi 10.1080/10428190400005353 (2005).
597 Saiya-Cork, K. *et al.* A pathobiological role of the insulin receptor in chronic lymphocytic leukemia. *Clin Cancer Res* **17**, 2679-2692, doi:1078-0432.CCR-10-2058 [pii]
10.1158/1078-0432.CCR-10-2058 (2011).
598 Cutillas, P. R. & Jorgensen, C. Biological signalling activity measurements using mass spectrometry. *Biochemical Journal* **434**, 189-199, doi:Doi 10.1042/Bj20101974 (2011).
599 Diella, F. *et al.* Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* **5**, 79, doi:10.1186/1471-2105-5-79
1471-2105-5-79 [pii] (2004).
600 Linding, R. *et al.* Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415-1426, doi:S0092-8674(07)00727-1 [pii]
10.1016/j.cell.2007.05.052 (2007).
601 Trimble, V. Existence and Nature of Dark Matter in the Universe. *Annu Rev Astron Astr* **25**, 425-472, doi:Doi 10.1146/Annurev.Astro.25.1.425 (1987).
602 Balakrishnan, K. *et al.* Influence of bone marrow stromal microenvironment on forodesine-induced response in CLL primary cells. *Blood*, doi:blood-2009-10-246199 [pii]
10.1182/blood-2009-10-246199 (2010).
603 Wickremasinghe, R. G., Prentice, A. G. & Steele, A. J. Aberrantly activated anti-apoptotic signalling mechanisms in chronic lymphocytic leukaemia cells: clues

to the identification of novel therapeutic targets. *Br J Haematol* **153**, 545-556, doi:10.1111/j.1365-2141.2011.08676.x (2011).

604 Aristotle & (transl. Bostock, D. *Metaphysics : Books Z and H*. (Oxford : Clarendon, 355-323 BC).

605 Goh, W. W., Lee, Y. H., Chung, M. & Wong, L. How advancement in biological network analysis methods empowers proteomics. *Proteomics* **12**, 550-563, doi:10.1002/pmic.201100321 (2012).

606 Erler, J. T. & Linding, R. Network medicine strikes a blow against breast cancer. *Cell* **149**, 731-733, doi:S0092-8674(12)00511-9 [pii]
10.1016/j.cell.2012.04.014 (2012).

607 Gribbin, J. *Deep simplicity : chaos, complexity and the emergence of life*. (Penguin, 2005).

608 Cheson, B. D. *et al.* Novel targeted agents and the need to refine clinical end points in chronic lymphocytic leukemia. *J Clin Oncol* **30**, 2820-2822, doi:JCO.2012.43.3748 [pii]
10.1200/JCO.2012.43.3748 (2012).

609 Jalkanen, S. E. *et al.* Phosphoprotein profiling predicts response to tyrosine kinase inhibitor therapy in chronic myeloid leukemia patients. *Experimental Hematology* **40**, 705-714, doi:Doi 10.1016/J.Exphem.2012.05.010 (2012).

610 Guo, A. *et al.* Signaling networks assembled by oncogenic EGFR and c-Met. *Proc Natl Acad Sci U S A* **105**, 692-697, doi:0707270105 [pii]
10.1073/pnas.0707270105 (2008).