KLF2 mutation is the most frequent somatic change in splenic marginal zone lymphoma and identifies a subset with distinct genotype

Alexandra Clipson,[1]* Ming Wang,[1]* Laurence de Leval,[2] Margaret Ashton-Key,[3] Andrew Wotherspoon,[4] George Vassiliou,[5,6] Niccolo Bolli,[5,6] Carolyn Grove,[5] Sarah Moody,[1] Leire Escudero Ibarz,[1] Gunes Gundem,[5] Kim Brugger,[7] Xuemin Xue,[1] Ella Mi,[1] Anthony Bench,[6] Mike Scott,[6] Hongxiang Liu,[8] George Follows,[6] Eloy F. Robles,[9] Jose Angel Martinez Climent,[9] David Oscier,[10] A James Watkins,[1,6] Ming-Qing Du[1]

[1]Division of Molecular Histopathology, Department of Pathology, University of Cambridge, UK;
[2]Institute of Pathology, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland;
[3]Department of Cellular Pathology, Southampton University Hospitals National Health Service Trust, Southampton, UK;
[4]Department of Histopathology, Royal Marsden Hospital, London, UK;
[5]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK;
[6]Department of Haematology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK;
[7]Department of Molecular Genetics**,** Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK;
[8]Molecular Malignancy Laboratory, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK;
[9]Division of Oncology, Center for Applied Medical Research CIMA, University of Navarra, Pamplona, Spain;
[10]Department of Haematology, Royal Bournemouth Hospital, Bournemouth, United Kingdom

* These authors contributed equally to this study.

Running title: *KLF2* mutation and associated genotype in SMZL

Key words: SMZL, *KLF2* mutation, *IGHV1-2*, 7q deletion, NF-κB

Corresponding author:

Professor Ming-Qing Du,

Division of Molecular Histopathology,

Department of Pathology,

University of Cambridge,

Level 3 Lab Block, Box 231,

Addenbrooke's Hospital,

Hills Road, Cambridge, CB2 2QQ, UK.

Email: mqd20@cam.ac.uk

Tel.: 00 44 (0)1223 767092

Fax: 00 44 (0)1223 586670

Conflict of interest: The authors declare no conflict of interest.

Sources of Research support: The research was supported by grants from Leukaemia & Lymphoma Research, U.K., Addenbrooke's Charitable Trust. SM is a PhD student supported by MRC and Department of Pathology, University of Cambridge. LEI is a PhD student supported by the Pathological Society of UK & Ireland. NB is a fellow of the European Hematology Association and was supported by a starter grant from the Academy of Medical Sciences.

ABSTRACT

To characterise the genetics of splenic marginal zone lymphoma (SMZL), we performed whole exome sequencing of 16 cases and identified novel recurrent inactivating mutations in *KLF2*, a gene whose deficiency was previously shown to cause splenic marginal zone hyperplasia in mice. *KLF2* mutation was found in 40 (42%) of 96 SMZLs, but rarely in other B-cell lymphomas. The majority of *KLF2* mutations were frameshift indels or nonsense changes, with missense mutations clustered in the C-terminal zinc finger domains. Functional assays showed that these mutations inactivated the ability of KLF2 to suppress NF-κB activation by TLR, BCR, BAFFR and TNFR signalling. Further extensive investigations revealed common and distinct genetic changes between SMZL with and without *KLF2* mutation. *IGHV1-2* rearrangement and 7q deletion were primarily seen in SMZL with *KLF2* mutation, while *MYD88* and *TP53* mutations were nearly exclusively found in those without *KLF2* mutation. *NOTCH2*, *TRAF3*, *TNFAIP3* and *CARD11* mutations were observed in SMZL both with and without *KLF2* mutation. Taken together, *KLF2* mutation is the most common genetic change in SMZL and identifies a subset with a distinct genotype characterised by multi-genetic changes. These different genetic changes may deregulate various signalling pathways and generate cooperative oncogenic properties, thereby contributing to lymphomagenesis.

INTRODUCTION

Splenic marginal zone lymphoma (SMZL), a low grade B-cell lymphoma, is difficult to diagnose accurately due to a lack of specific histological, immunophenotypic and genetic markers.[1] Patients with SMZL present with a highly variable clinical course with the majority showing a median survival of 10 years, ~25% of cases die of the disease within 5 years and a further ~5% of cases display high-grade transformation.[2,3] Despite the advances in treatment of other lymphomas, the survival of patients with SMZL has not been improved over the last decade.[4] These dilemmas in diagnosis and clinical management are largely due to poor understanding of its genetics and molecular mechanism.

There is mounting evidence suggesting a role for antigenic stimulation in the pathogenesis of SMZL. Approximately 20% of patients with SMZL present with autoimmune phenomena. A small proportion of cases are associated with HCV infection and can be effectively treated by antiviral therapy.[5] Importantly, >30% of SMZL has biased usage of *IG* heavy chain variable gene, *IGHV1-2.*[6,7] Most of the *IGHV1-2* rearrangements are characterised by minimal somatic mutations and longer complementarity determining region-3 (CDR3) sequence with common motifs, suggesting a possible selection by superantigens.[7] Together, these findings indicate a critical role of active BCR signalling in the pathogenesis of SMZL.

SMZL lacks recurrent chromosome translocations. Approximately 30% of SMZLs show hemizygous 7q deletion, which is also seen frequently in splenic B-cell lymphoma/leukaemia unclassifiable (SBCLU), but rarely in other lymphoma subtypes.[8-10] The gene(s) targeted by the 7q deletion remain obscure despite the combined investigation of genomic and transcriptomic profiles and mutation analysis of a number of candidate genes.[11,12]

Recent studies by whole exome sequencing (WES) identified a plethora of somatic mutations in SMZL.[13-16] These studies together with candidate gene sequencing showed a diverse spectrum of mutations in the NOTCH, NF-κB, BCR and TLR pathways, and in histone modifiers and transcriptional regulators.[13-18] Most of these mutations were found in <10% of cases, with only *NOTCH2* mutations occurring more frequently, but variably among different studies (6.5-25%).[13-16] Importantly, a significant proportion of SMZL lack any of these candidate pathogenic mutations. As the number of cases investigated by WES in each of the above studies was small (6-15 cases),[13-16] the mutation landscape in SMZL is likely not yet fully characterised. In this study, we identified further novel recurrent mutations in SMZL by WES and showed that *KLF2* (Kruppel-like factor 2) was inactivated by mutations in 42% of SMZL. Mechanistically, *KLF2* mutations abrogated KLF2-mediated suppression of NF-κB activation by TLR, BCR, BAFFR and TNFR signaling. *KLF2* mutation is significantly associated with 7q deletion, *IGHV1-2* usage, *NOTCH2*, *TNFAIP3* and *TRAF3* mutation, and may potentially cooperate with these genetic changes in oncogenesis.

MATERIALS AND METHODS

Patients samples: Tumour DNA was extracted from 105 cases of SMZL (77 from fresh frozen [FF] lymphoma tissues, 3 from leukemic peripheral blood samples, 25 from formalin-fixed paraffin-embedded [FFPE] lymphoma tissue), SBCLU (n=3), chronic lymphocytic leukaemia (CLL, n=39, all from bone marrow aspirate), hairy cell leukaemia (HCL, n=30, all from bone marrow aspirate), extranodal marginal zone lymphoma of mucosa-associated lymphoma tissue (MALT lymphoma, n=47, all from FFPE diagnostic tissue biopsies), mantle cell lymphoma (MCL, 7 from FF and 4 from FFPE diagnostic tissue biopsies), follicular lymphoma (FL, 5 from FF and 6 from FFPE diagnostic tissue biopsies), and diffuse large B-cell lymphoma (DLBCL, n=28, all FFPE diagnostic tissue biopsies). The lymphoma diagnosis was made according to the 2008 WHO classification of tumours of haematopoietic and lymphoid tissues. Where indicated, germline DNA was prepared from non-

neoplastic cells of tissues. The DNA quality was assessed by PCR of variably sized genomic fragments,[19] and samples with successful amplifications of >300bp were used for genetic analyses. Partial data on 7q deletion and *IGHV* usage in SMZL were available from previous studies,[10-12] with all other genetic data collected in this study. Local ethical guidelines were followed for the use of archival tissues for research with the approval of the ethics committees of the involved institutions.

Exome sequencing and somatic variant calling: These were carried out by the Wellcome Trust Sanger Institute. High molecular weight (HMW) tumour DNA samples from 16 cases of SMZL and matched germline DNA samples from 3 of these cases (non-neoplastic FFPE tissues in 1, buccal swap or non-involved peripheral blood sample in 2) were used to generate genomic libraries with the Illumina Paired End Sample Prep Kit (Table S1). Enrichment was performed using the Agilent SureSelect Human All Exon 50Mb kit.[20-22] Each exome was sequenced using a 75bp paired-end protocol on an Illumina HiSeq platform. Sequencing reads were aligned to the hg19 reference genome using the BWA algorithm on default settings.

Novel variants were called by comparison of tumour and germline sequence reads. CaVEMan (Cancer Variants through Expectation Maximisation) was used to call single nucleotide substitution,[21] while Pindel was used to call insertions and deletions.[23] Post processing filters were applied to increase the specificity of the output, remove variants reported in poor quality sequences and remove known SNPs in databases and unmatched normals from this study and the 10,000 genomes project.

Somatic variant validation by PCR and Sanger sequencing: Where indicated, novel variants identified by WES and their potential somatic origin were first confirmed by PCR and Sanger sequencing. Depending on the nature of gene sequences, different approaches were employed for mutation screening. Mutations in *KLF2* were screened by PCR and Sanger sequencing as the gene

has a high GC content (Table S2). In each case, sequence change was confirmed by at least two independent PCR and sequencing experiments. The somatic mutation was ascertained by excluding germline changes through SNP database search and analysis of germline DNA samples where possible.

Somatic variant validation by Fluidigm Access Array PCR and Illumina MiSeq sequencing: Mutations in *NOTCH2*, *TNFAIP3, TRAF3*, *MYD88, IKBKB, CARD11, BCL10, CD79A, CD79B* and *TP53* were screened by massive parallel Fluidigm Access Array PCR and Illumina MiSeq sequencing using our established protocol from a parallel investigation (manuscript in preparation). The in house variant calling algorithm was developed and optimised against a large number of various known somatic mutations by Sanger sequencing. Please refer to Supplementary Methods for experimental details and variant calling algorithms (Table S3). Each sample was investigated in duplicate to eliminate any potential false positives. Any novel variants seen in both replicates of the same sample were further ascertained by an independent Fluidigm PCR and MiSeq sequencing or Sanger sequencing. Where indicated, their somatic nature was confirmed by PCR and Sanger sequencing of the paired non-tumour DNA sample, or by search of COSMIC somatic mutation database.

NF-κB reporter assay: The full-length coding sequence and various truncated forms of *KLF2* were amplified from pCMV6-AC-GFP (OriGene, USA) by PCR and cloned into the pIRES-puro2-HA vector at the *EcoRI* and *BamHI* sites. The *KLF2* mutant containing a single point mutation was generated from the wild type using the QuickChange Site-directed mutagenesis kit (Stratagene, USA). PCR and sequencing were performed to verify the *KLF2* sequence and reading frame. The effect of KLF2 and its mutants on suppression of NF-κB activation by various stimuli (TNFα, BAFF, mutant MYD88 or mutant CARD11) was investigated where appropriate in both HEK293T and OCI-LY19 human B-cell lymphoma cell lines using a Dual-Luciferase reporter assay (Promega, UK).[24,25] Please refer to Supplementary Methods for experimental details.

Analysis of rearranged *IGH* genes by PCR and Sanger sequencing: The rearranged *IGH* genes were amplified using BIOMED-2 FR1 and consensus JH primer sets.[26] The PCR product was purified and sequenced as previously described.[10] The VH sequence was identified using the IMGT/V-QUEST database (http://www.imgt.org). Cases harbouring < 97% homology were considered significantly mutated, while those with 97-99.9% identity were regarded minimally mutated.[7]

Statistical analyses: The student's t-test was performed using GraphPad Prism version 5.00 software (GraphPad Software, San Diego, USA). The correlation among categorical variables was evaluated by Fisher's exact probability test. Overall survival (OS) was measured from the date of diagnosis to death from any cause. Probabilities of OS were calculated by the Kaplan–Meier method, and the comparison between subgroups was performed via the log-rank test. Kaplan-Meier analysis, log-rank test and Fisher's exact test were carried out using SPSS, version 13.

RESULTS

Identification of *KLF2* mutation by WES

WES was successful for all 16 tumour and 3 matched germline DNA samples (Table S4). Based on the 3 cases with matched germline DNA, a total of 174 variants in 163 genes (average 58/case; range 45-82/case) were seen, with variants in 135 genes being novel, not reported previously in SMZL (Table S5). The number of variants in the remaining cases was much higher due to a lack of WES data from matched germline DNA, thus preventing the filter of all SNPs. Nonetheless, a total of 223 variants were observed in 159 genes known to be mutated in SMZL by previous studies, including those described in the NOTCH2 signalling pathway (*NOTCH2*, *NOTCH4*, *SPEN*), NF-κB pathway (*TNFAIP3*, *TRAF3*, *BIRC3*), BCR pathway (*CARD11*) and TLR pathway (*MYD88*) (Figure S1).[13-16]

Comparative analyses of the exome sequencing data from the 4 published WES studies also revealed little overlap among the variants identified in these studies (Figure S2).[13-16] Together, these findings suggest the presence of a remarkable heterogeneity or incomplete discovery of the somatic mutation profile in SMZL, or both.

Among the variants not reported previously, there were several recurrent changes not seen in the matched control DNA. Based on the frequency, possible functional impact and a comprehensive literature search, we identified *KLF2* mutation as a potentially significant genetic abnormality in SMZL (Figure S1). Among the 16 SMZLs investigated by WES in this study, *KLF2* mutation was seen in 5 cases, all being deleterious changes (frameshift insertion/deletion in 2, nonsense mutation in 2, and substitution change at an essential splice site in 1 case). Further PCR and Sanger sequencing confirmed these mutations, and their somatic origin in all 5 cases. In support of the pathogenic importance of these mutations in SMZL, Klf2 deficient mice were previously shown to have a marked increase in marginal zone B-cells and splenic marginal zone hyperplasia.[27-29]

KLF2 is frequently targeted by mutation in SMZL.

Next, we investigated *KLF2* mutation in 96 cases of SMZL including 13 of the 16 cases investigated by WES, and 7 other B-cell lymphoma entities to determine its frequencies and mutation spectrum. As the *KLF2* gene has a high GC content that may have accounted for the failure of detection of its mutation by previous exome sequencing studies,[13-16] PCR and Sanger sequencing were used for mutation screening. A total of 47 *KLF2* mutations were seen in 40 (42%) of the 96 cases of SMZL with double mutations in 7 cases, and their somatic nature was confirmed in each of the 8 cases (including 3 indels and 7 substitutions), for which non-tumour DNA was available (Figure 1A, Figure S3, Table S6). Of note, Sanger sequencing identified 4 additional mutations in 3 cases missed by WES, which included 1 frameshift deletion, 1 in-frame deletion and 2 missense substitutions. In contrast,

KLF2 mutation was not or rarely seen in SBCLU (0/3), CLL (0/39), HCL (3/30), FL (1/11), MCL (1/11), MALT lymphoma (2/47) and DLBCL (0/28) (Figure 1B).

Among the 47 *KLF2* mutations identified in SMZL, 27 were frameshift insertions / deletions, nonsense mutations, or substitutions affecting the essential splice site, thus resulting in a potentially truncated protein product (Figure 1A). Importantly, a high proportion of these deleterious mutations were localised toward the N-terminal activation domain (AD) and middle inhibitory domain (ID). The remaining 20 mutations seen in SMZL were 19 missense substitutions and 1 in-frame deletion, with 12 clustered in the C-terminal zinc finger (ZF) 1, mainly at conserved amino acid residues (Figure 1A). Fourteen missense mutations were predicted to be damaging by the PolyPhen-2 program,[30] and a further 2 missense mutations abolished the stop codon with potential extension of a further 62 amino acids. Among the 7 *KLF2* mutations found in other lymphoma entities, 2 were frameshift deletions in the ID and ZF3 domains respectively, and 5 were missense mutations with only 1 in the C-terminal ZF1 domain (Figure 1A).

Functional characterisation of *KLF2* mutations

The nature and distribution of the *KLF2* mutations suggest that these genetic changes are likely to inactivate KLF2 function. We thus generated a series of *KLF2* expression constructs, representing various C-terminal truncated products that were lacking 1 or more ZF domains, and recurrent missense or in-frame deletion mutants that affected conserved amino acid residues in the ZF1 domain, and tested their impact on KLF2 function using *in vitro* reporter assays in both HEK293T and OCI-LY19 B-lymphoma cells. Previous studies have shown that KLF2 inhibits the transcriptional activity of NF-κB.[31] We therefore tested the ability of various KLF2 mutants to suppress NF-κB activation by TNFα, MYD88(S219C), CARD11(F130V) and BAFF, which were used to activate TNFR, TLR, canonical and non-canonical NF-κB signalling pathways respectively.

As expected, wild type KLF2 was highly potent in the suppression of NF-κB activation by different signalling pathways including stimulation by TNFα, MYD88(S219C), CARD11(F130V) and BAFF (Figure 2B). With the exception of KLF2-A291V mutant, all other 5 mutants including KLF2-C274Y and KLF2-ΔTY mutants showed a total or major loss in NF-κB suppression, although to varying extents depending on stimuli used to activate NF-κB. Interestingly, both cases with KLF2-A291V had a second mutation, one with P70S in the activation domain, the other with C274S, at which a C274Y change was shown to impair KLF2 function as described above (Table S6).

KLF2 mutation identifies a subset of SMZL with distinct genotypes

To further characterise the genetics of SMZL and understand their potential cooperation in lymphomagenesis, we comprehensively investigated somatic mutations in *NOTCH2*, *TNFAIP3*, *TRAF3*, *MYD88*, *CD79A*, *CD79B*, *CARD11*, *BCL10*, *IKBKB* and *TP53* (Figure 3, Table S6), 7q deletion and *IGHV* usage in the entire cohort of SMZL and correlated their changes with *KLF2* mutation using the Fisher's exact probability test. The analyses revealed several significant associations.

First, *KLF2* mutation identified a subset of SMZL with distinct genetic changes. The mutation was significantly associated with both 7q deletion (P=7.33×10^{-7}) and *IGHV1-2* usage (P=1.02×10^{-7}), seen in 77% cases with 7q deletion and 83% of those with *IGHV1-2* (Figure 4). As with previous studies,[6,7] the rearranged *IGHV1-2* was characterised by minimal somatic mutations (Figure 4A).

Second, mutations in the *NOTCH2*, *TRAF3*, *TNFAIP3* and *CARD11* genes were found in SMZL both with and without *KLF2* mutation, with *NOTCH2*, *TRAF3* and *TNFAIP3* mutations being significantly associated with *KLF2* mutations (Figure 4C, P=0.007, P=0.012 and P=0.015 respectively). In line with

the recent studies, [13,14,16] *NOTCH2* mutations were characterised by frameshift insertion/deletion, and nonsense mutations, which were clustered at the C-terminus and predicted to eliminate the C-terminal PEST, a domain critical for NOTCH2 proteasomal degradation (Figure 3). *TRAF3* and *TNFAIP3* mutations were featured by frameshift insertions/deletion and nonsense mutations, while *CARD11* mutations were typically activating changes reported elsewhere (Figure 3).[32,33]

Third, most SMZL without *KLF2* mutation showed heterogeneous usage of *IGHV* in their rearranged *IGH* genes, and the majority of these rearranged *IGHV* harboured high loads of somatic mutations. Interestingly, *MYD88* mutations were exclusively seen in cases without *KLF2* mutation (Figure 4A&C, *P*=0.021). *MYD88* mutations were typically those of activating changes, while *TP53* mutations were characteristic inactivating changes reported elsewhere.

Mutation in *CD79A* and *CD79B* (both seen in a single case), *BCL10* [2 cases, both mutations predicted a C-terminal truncated BCL10 with a potential gain of function[34]], and *IKBKB* (1 case) was found to be low in SMZL. Interestingly, these mutations were mutually exclusive from *CARD11* mutations.

Correlation among genetic abnormalities and clinicopathological parameters

Follow-up data were available for 60 cases of SMZL, ranging from 12 to 288 months (median = 55 months). Kaplan-Meier univariate analysis of the genetic and clinical variables showed that only *TP53* mutation was significantly associated with poor 5-year overall survival (*P* =0.002) (Table S7). However, the number of cases and death events were not sufficient for reliable multivariate analysis.

DISCUSSION

By WES and validation of the mutations identified, we have made several novel and significant discoveries in the present study. First, *KLF2* is frequently mutated in SMZL (42%), but not or rarely in other lymphomas; second, *KLF2* mutations are characterised by frameshift insertion/deletion, nonsense mutations, and a cluster of missense mutations in the ZF1 domain, which impair KLF2 function; third, there are distinct genetic changes according to *KLF2* mutation status. *IGHV1-2* rearrangement and 7q deletion are essentially seen in SMZL with *KLF2* mutation, while *MYD88* and *TP53* mutations are nearly exclusively seen in those without *KLF2* mutation. Mutations in *NOTCH2*, *TRAF3*, *TNFAIP3* and *CARD11* genes were found in SMZL both with and without *KLF2* mutation. These distinct mutation patterns indicate overlapping molecular mechanisms between SMZL with and without *KLF2* mutation, and also suggest the presence of different oncogenic cooperation between the two subgroups.

KLF2 mutation and its distinctively associated genetic changes

Among the diverse spectrum of mutations identified in SMZL, *KLF2* mutation (42%) is the most frequent genetic change, much higher than the recently identified *NOTCH2* mutation (6.5-25%).[13-16] The nature of *KLF2* mutations and our *in vitro* functional studies of *KLF2* mutants indicate that these mutations inactivate KLF2 function.

A pathogenic role of *KLF2* inactivating mutations in SMZL is strongly supported by recent findings, particularly those by studies of Klf2 knockout mice. Remarkably, B-cell specific Klf2-deficient mice show a dramatic increase of marginal zone B-cells.[27-29] Klf2 deficiency appears to promote follicular B-cells to gain a marginal zone like phenotype and migrate to the splenic marginal zone, but have little impact on their proliferation.[27-29] The molecular mechanism underlying the altered B-cell homeostasis and trafficking in Klf2-deficient mice is unclear although Klf2 most likely exerts such effects through transcriptional regulation of its target genes.

KLF2 is a member of the KLF family of transcription factors, and has been recently shown to be a negative regulator of inflammation and NF-κB activities.[35-38] KLF2 appears to regulate NF-κB activities by modulating recruitment of critical NF-κB coactivators.[38] Using an *in vitro* reporter assay, we showed that wild type KLF2 was a potent inhibitor of NF-κB activation by several signalling pathways including BCR (CARD11 mutant), TLR (MYD88 mutant), TNFR (TNFα) and BAFFR (BAFF). In contrast, KLF2 mutants had a total or major loss in suppression of NF-κB activation triggered by these signals. Given the importance of TLR, canonical and non-canonical NF-κB pathways in the development of marginal zone B-cells,[39,40] KLF2 inactivation by mutation may exert its oncogenic activities at least in part by deregulation of NF-κB activities triggered by these signals, leading to altered gene expression favouring B-cells homing to the marginal zone. However, *KLF2* inactivation alone is insufficient for malignant transformation, and requires cooperating genetic and cellular events in SMZL development.

The majority of SMZL with *KLF2* mutation have both 7q deletion and *IGHV1-2* rearrangement. The genes targeted by 7q deletion are unclear.[11,12] *IGHV1-2* usage is over-represented in SMZL, accounting for 30% of cases. Although the epitope recognised by *IGHV1-2* expressing BCR and its potential impact on clinicopathological presentation are unknown, the features of *IGHV1-2* rearrangements, including minimal somatic mutations and longer CDR3 sequence with common motifs, suggest a possible selection of T-cell independent marginal zone B-cells by superantigens,[7] thus a role of antigenic drive in the lymphomagenesis (Figure 5). Such active BCR signalling may cooperate with *KLF2* inactivation in SMZL development. Apart from *IGHV1-2* BCR stereotype, there are further genetic changes that potentially cooperate with KLF2 inactivation although these genetic changes occur in SMZL both with and without *KLF2* mutation.

Genetic changes common to SMZL both with and without *KLF2* mutation

NOTCH2, *TRAF3*, *CARD11* and *TNFAIP3* mutations were found in SMZL both with and without *KLF2* mutation. *NOTCH2*, *TRAF3* and *CARD11* mutations are most likely to enhance the NOTCH2, non-canonical NF-κB and BCR signalling respectively,[13,14,16,41,42] while *TNFAIP3* mutation may augment several molecular pathways including TNFR, TLR/IL1-R and BCR signalling.[43] All these innate signals are critical for the development of marginal zone B-cells although their precise role remains to be dissected. NOTCH2 signalling is critical for generation of marginal zone B-cells and their retention in the splenic marginal zone.[44-47] Active NOTCH2 signalling alone appears to have little impact on cell proliferation and survival, but sensitises B-cells to stimulation of surface TLR and CD40.[46,47] Non-canonical NF-κB signalling, typically triggered by stimulation of surface BAFFR and CD40, is also pivotal in development of marginal zone B-cells and formation of the splenic marginal zone.[48] Chronic active BCR signalling promotes cellular proliferation and survival, and TLR signalling may contribute to both the development and survival of marginal zone B-cells.[40] Marginal zone B cells express NOTCH2, BAFFR, CD40, and high levels of TLRs, while the splenic innate lymphoid cells (ILCs) express surface DLL1 (Notch ligand Delta-like 1), BAFF, CD40L and provide contact-dependent help to marginal zone B-cells by stimulation of the respective receptors in a cooperative manner.[39,40,49] Thus, *NOTCH2*, *TRAF3*, *CARD11* and *TNFAIP3* mutation may cooperate with the aforementioned surface receptor stimulation and cause constitutive activation of the corresponding signalling pathway.

The respective signalling enhanced by *NOTCH2*, *TRAF3*, *TNFAIP3* and *CARD11* mutation likely complements the molecular mechanism deregulated by *KLF2* mutation, *IGHV1-2* expressing BCR and genes targeted by 7q deletion, thus cooperating in SMZL development (Figure 5). However, the genetic events that cooperate with these mutations in SMZL without *KLF2* mutation are unclear.

Genetic changes preferentially associated with SMZL without *KLF2* mutation

There are also several interesting features in SMZL without *KLF2* mutation, including infrequent 7q deletion and *IGHV1-2* rearrangement, and nearly exclusive association with *MYD88* and *TP53* mutations (Figure 5). Among SMZL without *KLF2* mutation, there is a heterogeneous usage of *IGHV* and the majority of these rearranged *IGHV* genes show high levels of somatic mutations, suggesting origin from T-cell dependent marginal zone B-cells.[50]

The *MYD88* mutations seen in SMZL are typically those of gain-of-function change, capable of spontaneously assembling a signalling complex to activate NF-κB, STAT3 and AP1 transcription factors.[51] *MYD88* activation by mutation may lead to biological consequences similar to that by TLR activation implicated by KLF2 and/or TNFAIP3 inactivation.

The *TP53* mutations seen in SMZL were typically those reported elsewhere, and these mutations likely inactivate TP53 function. *TP53* mutation in SMZL, like in other lymphomas, is likely to be a secondary genetic event. In line with this, TP53 inactivation is associated with progression and poor prognosis in SMZL.[12,52,53]

In summary, there are common and distinct genetic changes between SMZL with and without *KLF2* mutation and these different genetic changes most likely deregulate several signalling pathways important for the generation of marginal zone B-cells, their migration and retention in the splenic marginal zone. Each of these genetic changes may have a predominant impact on a particular biological process and contribute to the lymphoma development through oncogenic cooperation with other concurrent changes (Figure 5). It is pertinent to tentatively speculate that 1) KLF2 inactivation may deregulate gene expression through the modulation of NF-κB activities and other unknown mechanisms, thereby promoting B-cells homing to the splenic marginal zone; 2) NOTCH2 activation, TRAF3 and TNFAIP3 inactivation and MYD88 activation by mutations may contribute to

the generation of marginal zone B-cells and their retention in the splenic marginal zone by augmenting the NOTCH2, non-canonical NF-κB pathway and TLR signalling respectively; 3) *IGHV1-2* rearrangement and CARD11 activation by mutations may lead to chronic active BCR signalling, consequently enhancing cell proliferation and survival. A simultaneous deregulation of the above signalling pathways in SMZL with *KLF2* mutation may generate complementary properties in oncogenic cooperation, leading to lymphoma development.

Acknowledgements: The authors would like to thank David Withers for his help in DNA sequencing, and Dr Yuanxue Huang, Howard Martin, Antje Schulze Selting and Robbie Zhao for technical assistance. We would also like to thank NIHR Cambridge Comprehensive Biomedical Research Centre Tissue and Blood Biobank for providing some of the lymphoma samples used in this study.

Author contributions: Data collection and analysis: AC, MW, SM, LEI, EM, HL, EFR, GV, JAMC, DO; Sequence analysis: AC, NB, CG, GG, KB, XX; Case contribution: LdL, MAK, AW, GV, AB, MS, GF, JAMC, DO; Manuscript writing and preparation: MQD, AC, MW. Study design and coordination: MQD, AJW. All authors commented on the manuscript and approve its submission for publication. There is no conflict of interest to declare.

Supplementary information is available at Leukemia's website.

References:

1 Isaacson PG, Piris MA, Berger F, Swerdlow SH, Thieblemont C, Pittaluga S, Harris NL. Splenic B-cell marginal zone lymphoma. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, editors. WHO Classification of Tumous of Haematopoietic and Lymphoid Tissues. Lyon: International Agency for Research on Cancer, 2008: 185-187.

2 Arcaini L, Lazzarino M, Colombo N, Burcheri S, Boveri E, Paulli M et al. Splenic marginal zone lymphoma: a prognostic model for clinical use. *Blood* 2006; **107**: 4643-4649.

3 Montalban C, Abraira V, Arcaini L, Domingo-Domenech E, Guisado-Vasco P, Iannito E et al. Risk stratification for Splenic Marginal Zone Lymphoma based on haemoglobin concentration, platelet count, high lactate dehydrogenase level and extrahilar lymphadenopathy: development and validation on 593 cases. *Br J Haematol* 2012; **159**: 164-171.

4 Olszewski AJ, Castillo JJ. Survival of patients with marginal zone lymphoma: Analysis of the Surveillance, Epidemiology, and End Results database. *Cancer* 2013; **119**: 629-638.

5 Arcaini L, Bruno R. Hepatitis C virus infection and antiviral treatment in marginal zone lymphomas. *Curr Clin Pharmacol* 2010; **5**: 74-81.

6 Zibellini S, Capello D, Forconi F, Marcatili P, Rossi D, Rattotti S et al. Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica* 2010; **95**: 1792-1796.

7 Bikos V, Darzentas N, Hadzidimitriou A, Davis Z, Hockley S, Traverse-Glehen A et al. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. *Leukemia* 2012; **26**: 1638-1646.

8 Mateo M, Mollejo M, Villuendas R, Algara P, Sanchez-Beato M, Martinez P et al. 7q31-32 allelic loss is a frequent finding in splenic marginal zone lymphoma. *Am J Pathol* 1999; **154**: 1583-1589.

9 Gruszka-Westwood AM, Hamoudi R, Osborne L, Matutes E, Catovsky D. Deletion mapping on the long arm of chromosome 7 in splenic lymphoma with villous lymphocytes. *Genes Chromosomes Cancer* 2003; **36**: 57-69.

10 Watkins AJ, Huang Y, Ye H, Chanudet E, Johnson N, Hamoudi R et al. Splenic marginal zone lymphoma: characterization of 7q deletion and its value in diagnosis. *J Pathol* 2010; **220**: 461-474.

11 Watkins AJ, Hamoudi RA, Zeng N, Yan Q, Huang Y, Liu H et al. An integrated genomic and expression analysis of 7q deletion in splenic marginal zone lymphoma. *PLoS One* 2012; **7**: e44997.

12 Fresquet V, Robles EF, Parker A, Martinez-Useros J, Mena M, Malumbres R et al. High-throughput sequencing analysis of the chromosome 7q32 deletion reveals IRF5 as a potential tumour suppressor in splenic marginal-zone lymphoma. *Br J Haematol* 2012; **158**: 712-726.

13 Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med* 2012; **209**: 1537-1551.

14 Kiel MJ, Velusamy T, Betz BL, Zhao L, Weigelin HG, Chiang MY et al. Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma. *J Exp Med* 2012; **209**: 1553-1565.

15 Martinez N, Almaraz C, Vaque JP, Varela I, Derdak S, Beltran S et al. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. *Leukemia* 2014; **28**: 1334-1340.

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

17 Yan Q, Huang Y, Watkins AJ, Kocialkowski S, Zeng N, Hamoudi RA et al. BCR and TLR signaling pathways are recurrently targeted by genetic changes in splenic marginal zone lymphomas. *Haematologica* 2012; **97**: 595-598.

18 Rossi D, Deaglio S, Dominguez-Sola D, Rasi S, Vaisitti T, Agostinelli C et al. Alteration of BIRC3 and multiple other NF-{kappa}B pathway genes in splenic marginal zone lymphoma. *Blood* 2011.

19 Liu H, Bench AJ, Bacon CM, Payne K, Huang Y, Scott MA et al. A practical strategy for the routine use of BIOMED-2 PCR assays for detection of B- and T-cell clonality in diagnostic haematopathology. *Br J Haematol* 2007; **138**: 31-43.

20 Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* 2011; **365**: 1384-1395.

21 Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* 2011; **469**: 539-542.

22 Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 2012; **486**: 400-404.

493 23 Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect
494 break points of large deletions and medium sized insertions from paired-end short reads.
495 *Bioinformatics* 2009; **25**: 2865-2871.

496 24 Bi Y, Zeng N, Chanudet E, Huang Y, Hamoudi RA, Liu H et al. A20 inactivation in ocular adnexal
497 MALT lymphoma. *Haematologica* 2012; **97**: 926-930.

498 25 Yan Q, Wang M, Moody S, Xue X, Huang Y, Bi Y et al. Distinct involvement of NF-kappaB
499 regulators by somatic mutation in ocular adnexal malt lymphoma. *Br J Haematol* 2013; **160**:
500 851-854.

501 26 van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL et al. Design
502 and standardization of PCR primers and protocols for detection of clonal immunoglobulin and
503 T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2
504 Concerted Action BMH4-CT98-3936. *Leukemia* 2003; **17**: 2257-2317.

505 27 Hoek KL, Gordy LE, Collins PL, Parekh VV, Aune TM, Joyce S et al. Follicular B cell trafficking
506 within the spleen actively restricts humoral immune responses. *Immunity* 2010; **33**: 254-265.

507 28 Winkelmann R, Sandrock L, Porstner M, Roth E, Mathews M, Hobeika E et al. B cell
508 homeostasis and plasma cell homing controlled by Kruppel-like factor 2. *Proc Natl Acad Sci U S*
509 *A* 2011; **108**: 710-715.

510 29 Hart GT, Wang X, Hogquist KA, Jameson SC. Kruppel-like factor 2 (KLF2) regulates B-cell
511 reactivity, subset differentiation, and trafficking molecule expression. *Proc Natl Acad Sci U S A*
512 2011; **108**: 716-721.

513 30 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P et al. A method and
514 server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248-249.

31 Das H, Kumar A, Lin Z, Patino WD, Hwang PM, Feinberg MW et al. Kruppel-like factor 2 (KLF2) regulates proinflammatory activation of monocytes. *Proc Natl Acad Sci U S A* 2006; **103**: 6653-6658.

32 Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* 2008; **319**: 1676-1679.

33 Dong G, Chanudet E, Zeng N, Appert A, Chen YW, Au WY et al. A20, ABIN-1/2 and CARD11 mutations and their prognostic value in gastrointestinal diffuse large B-cell lymphoma. *Clin Cancer Res* 2011.

34 Du MQ, Peng H, Liu H, Hamoudi RA, Diss TC, Willis TG et al. BCL10 gene mutation in lymphoma. *Blood* 2000; **95**: 3885-3890.

35 Mahabeleshwar GH, Kawanami D, Sharma N, Takami Y, Zhou G, Shi H et al. The myeloid transcription factor KLF2 regulates the host response to polymicrobial infection and endotoxic shock. *Immunity* 2011; **34**: 715-728.

36 Heo KS, Chang E, Takei Y, Le NT, Woo CH, Sullivan MA et al. Phosphorylation of protein inhibitor of activated STAT1 (PIAS1) by MAPK-activated protein kinase-2 inhibits endothelial inflammation via increasing both PIAS1 transrepression and SUMO E3 ligase activity. *Arterioscler Thromb Vasc Biol* 2013; **33**: 321-329.

37 Zahlten J, Steinicke R, Opitz B, Eitel J, N'guessan PD, Vinzing M et al. TLR2- and nucleotide-binding oligomerization domain 2-dependent Krüppel-like factor 2 expression down-regulates NF-kappa B-related gene expression. *J Immunol* 2010; **185**: 597-604.

38 Nayak L, Goduni L, Takami Y, Sharma N, Kapil P, Jain MK et al. Kruppel-like factor 2 is a transcriptional regulator of chronic and acute inflammation. *Am J Pathol* 2013; **182**: 1696-1704.

537 39 Pillai S, Cariappa A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev*
538 *Immunol* 2009; **9**: 767-777.

539 40 Cerutti A, Cols M, Puga I. Marginal zone B cells: virtues of innate-like antibody-producing
540 lymphocytes. *Nat Rev Immunol* 2013; **13**: 118-132.

541 41 Sun SC. The noncanonical NF-kappaB pathway. *Immunol Rev* 2012; **246**: 125-140.

542 42 Niemann CU, Wiestner A. B-cell receptor signaling as a driver of lymphoma development and
543 evolution. *Semin Cancer Biol* 2013; **23**: 410-421.

544 43 Catrysse L, Vereecke L, Beyaert R, Van Loo G. A20 in inflammation and autoimmunity. *Trends*
545 *Immunol* 2014; **35**: 22-31.

546 44 Tanigaki K, Han H, Yamamoto N, Tashiro K, Ikegawa M, Kuroda K et al. Notch-RBP-J signaling is
547 involved in cell fate determination of marginal zone B cells. *Nat Immunol* 2002; **3**: 443-450.

548 45 Saito T, Chiba S, Ichikawa M, Kunisato A, Asai T, Shimizu K et al. Notch2 is preferentially
549 expressed in mature B cells and indispensable for marginal zone B lineage development.
550 *Immunity* 2003; **18**: 675-685.

551 46 Hampel F, Ehrenberg S, Hojer C, Draeseke A, Marschall-Schroter G, Kuhn R et al. CD19-
552 independent instruction of murine marginal zone B-cell development by constitutive Notch2
553 signaling. *Blood* 2011; **118**: 6321-6331.

554 47 Simonetti G, Carette A, Silva K, Wang H, De Silva NS, Heise N et al. IRF4 controls the positioning
555 of mature B cells in the lymphoid microenvironments by regulating NOTCH2 expression and
556 activity. *J Exp Med* 2013; **210**: 2887-2902.

557 48 Guo F, Weih D, Meier E, Weih F. Constitutive alternative NF-kappaB signaling promotes
558 marginal zone B-cell development but disrupts the marginal sinus and induces HEV-like
559 structures in the spleen. *Blood* 2007; **110**: 2381-2389.

560 49 Magri G, Miyajima M, Bascones S, Mortha A, Puga I, Cassis L et al. Innate lymphoid cells
561 integrate stromal and immunological signals to enhance antibody production by splenic
562 marginal zone B cells. *Nat Immunol* 2014; **15**: 354-364.

563 50 Pillai S, Cariappa A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev*
564 *Immunol* 2009; **9**: 767-777.

565 51 Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH et al. Oncogenically active MYD88
566 mutations in human lymphoma. *Nature* 2011; **470**: 115-119.

567 52 Gruszka-Westwood AM, Hamoudi RA, Matutes E, Tuset E, Catovsky D. p53 abnormalities in
568 splenic lymphoma with villous lymphocytes. *Blood* 2001; **97**: 3552-3558.

569 53 Salido M, Baro C, Oscier D, Stamatopoulos K, Dierlamm J, Matutes E et al. Cytogenetic
570 aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell
571 lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. *Blood* 2010; **116**:
572 1479-1488.

573 **FIGURE LEGENDS:**

574 **Figure 1:** Nature and incidence of *KLF2* mutations in SMZL and other B-cell lymphomas.

575

576 A) Nature and distribution of *KLF2* mutations in lymphoma. The majority of *KLF2* mutations seen in
577 SMZL are frameshift deletion / insertion or nonsense mutations. Missense mutations are largely
578 clustered in ZF1 (zinc finger), particularly at conserved amino acid residues. Mutations were
579 confirmed by two independent PCR and sequencing experiments. Where possible the somatic
580 nature of the mutation identified was confirmed by PCR and sequencing analysis of the paired non-
581 tumour DNA and indicated by red symbols. Mutations identical to those confirmed to be somatic
582 are highlighted in blue. Concurrent mutations seen in the same cases are indicated by their case
583 number.

584

585 B) Frequencies of *KLF2* mutations in SMZL and various other B-cell lymphomas.

586

587 ZF: zinc finger; NLS: putative nuclear localisation signal.

588

589 **Figure 2:** Functional characterisation of *KLF2* mutations.

590

591 A) The representative *KLF2* mutants investigated by *in vitro* reporter assays, which include three
592 truncation, two recurrent missense and one in-frame deletion mutants. The missense change and
593 in-frame deletion affect the conserved amino acid residues.

594

595 B) NF-κB reporter assay shows that wild type KLF2 is a potent inhibitor of NF-κB activation by TNFα,
596 BAFF, MYD88 and CARD11 mutants in both HEK293T and OCI-LY19 B-lymphoma cells. With the
597 exception of KLF2-A291V mutant, all other 5 mutants including KLF2-C274Y and KLF2-ΔTY mutants
598 showed a total or major loss in NF-κB suppression. Interestingly, both the cases with KLF2-A291V

had a second mutation, one with P70S in the activation domain, the other with C274S. The data is from at least three independent experiments and presented as a mean ± standard deviation, and the difference between KLF2 and its mutants is analysed by the Student t-test. As indicated by an arrowhead, the KLF2-AD-ZF1 mutant consistently shows an additional band, ~7 kDa larger than the expected size (32 kDa). We have performed a series of experiments to confirm the correct sequence and also rule out any cross-contamination of an additional clone. The band shift is most likely caused by post-translational modifications, but its nature remains to be established. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 3: Nature and distribution of mutations in *NOTCH2*, *TNFAIP3*, *TRAF3*, *CARD11*, *MYD88* and *TP53* in SMZL. *NOTCH2* mutations are characterised by frameshift insertion/deletion, and nonsense mutations that are clustered at the C-terminus and predicted to eliminate the C-terminal PEST, a domain critical for NOTCH2 proteasomal degradation. *TNFAIP3* and *TRAF3* mutations are featured by frameshift insertion/deletion and nonsense mutations, which inactivate their protein functions. *CARD11* and *MYD88* mutations are typically those of activating changes reported elsewhere. *TP53* mutations are also characteristic inactivating changes extensively reported in the literature. The mutations identified by Fluidigm PCR and MiSeq sequencing are confirmed either by an independent Fluidigm PCR and MiSeq sequencing or Sanger sequencing. Where possible, the somatic nature of mutation was determined: those confirmed by PCR and sequencing of the paired non-tumour DNA are shown by red symbols, while those identified in a search of the COSMIC somatic mutation database are shown by blue symbols. Concurrent mutations seen in the same cases are indicated by their case number.

LNR: LIN-12/NOTCH repeats; HD: heterodimerisation; TM: transmembrane; RAM: regulation of amino acid metabolism; TAD: transactivation domain; PEST: Proline, glutamic acid, serine and

625 threonine rich domain; ZF: zinc finger; CARD: caspase recruitment domain; PDZ: (PSD95, DLG and
626 ZO1 homology) domain; SH3: Src homology motif; GUK: guanylate kinase domain
627
628 **Figure 4:** Correlation of *KLF2* mutation with other genetic changes in SMZL.

629 A) Heatmap shows *KLF2* mutation and other genetic changes in 101 cases of SMZL. Rows
630 correspond to genetic change, while columns indicate individual cases. Positive genetic changes are
631 shown in green. Genes included in the BCR pathway mutations are *CARD11*, *BCL10*, *CD79A* and
632 *CD79B*.

633 B) Frequencies of *KLF2* mutation and other genetic changes in SMZL;

634 C) Correlation among *KLF2* mutation and other genetic changes in SMZL. *KLF2* mutation is
635 significantly and positively associated with 7q deletion, *IGHV1-2* usage, *NOTCH2*, *TRAF3* and *TNFAIP3*
636 mutations, but negatively correlated with *MYD88* mutation. BCR: B-cell receptor; Mut: mutation;
637 Del: deletion.

638

639 **Figure 5:** A summary of the proposed molecular mechanism of SMZL.

640 The majority of SMZL with *KLF2* mutation have the rearranged *IGHV1-2* that carries minimal levels of
641 somatic mutations, suggesting derivation of these lymphoma cells from T-cell independent marginal
642 zone B-cells. The biased usage of *IGHV1-2* indicates possible antigenic drive by superantigen, hence
643 chronic BCR signalling. *KLF2* inactivation by mutation may facilitate marginal zone B-cell
644 differentiation and their homing to the splenic marginal zone. 7q deletion is predominately seen in
645 cases with *KLF2* mutation and its role in the lymphoma pathogenesis is unknown.

646 *CARD11*, *NOTCH2*, *TRAF3* and *TNFAIP3* mutations are found in cases with and without *KLF2* mutation.
647 Mutations in *CARD11* and others (*CD79A/B*, *BCL10*) may lead to active BCR signalling, thereby

promoting cell proliferation and survival. Activation of NOTCH2, BAFFR/CD40 and TLR signalling by *NOTCH2*, *TRAF3* and *TNFAIP3* mutation may primarily contribute to marginal zone B-cell generation and their retention in the splenic marginal zone.

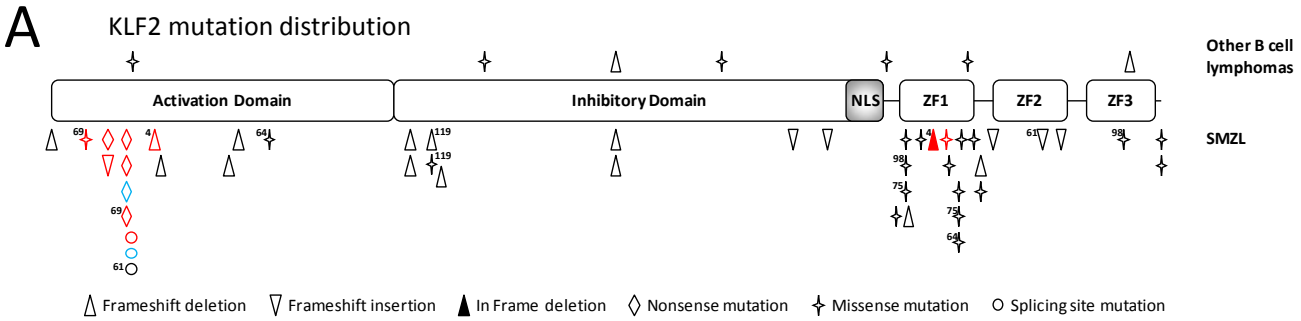
The majority of SMZL without *KLF2* mutation have heterogeneous usage of *IGHV* that carries high loads of somatic mutation, suggesting origination of these lymphoma cells from T-cell dependent marginal zone B-cells. *MYD88* and *TP53* mutations are nearly exclusively seen in cases without *KLF2* mutation. *MYD88* mutation most likely causes constitutive TLR signalling, while *TP53* mutation inactivates its tumour suppressor function and may promote disease progression and high-grade transformation.

SUPPLEMENTARY MATERIALS:

Supplementary methods

List of Supplementary Figures

List of Supplementary tables



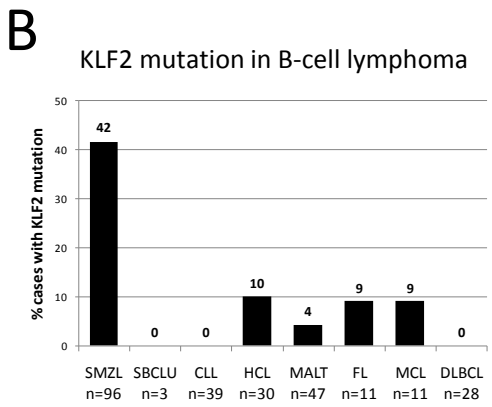


Figure 1: Nature and incidence of *KLF2* mutations in SMZL and other B-cell lymphomas.

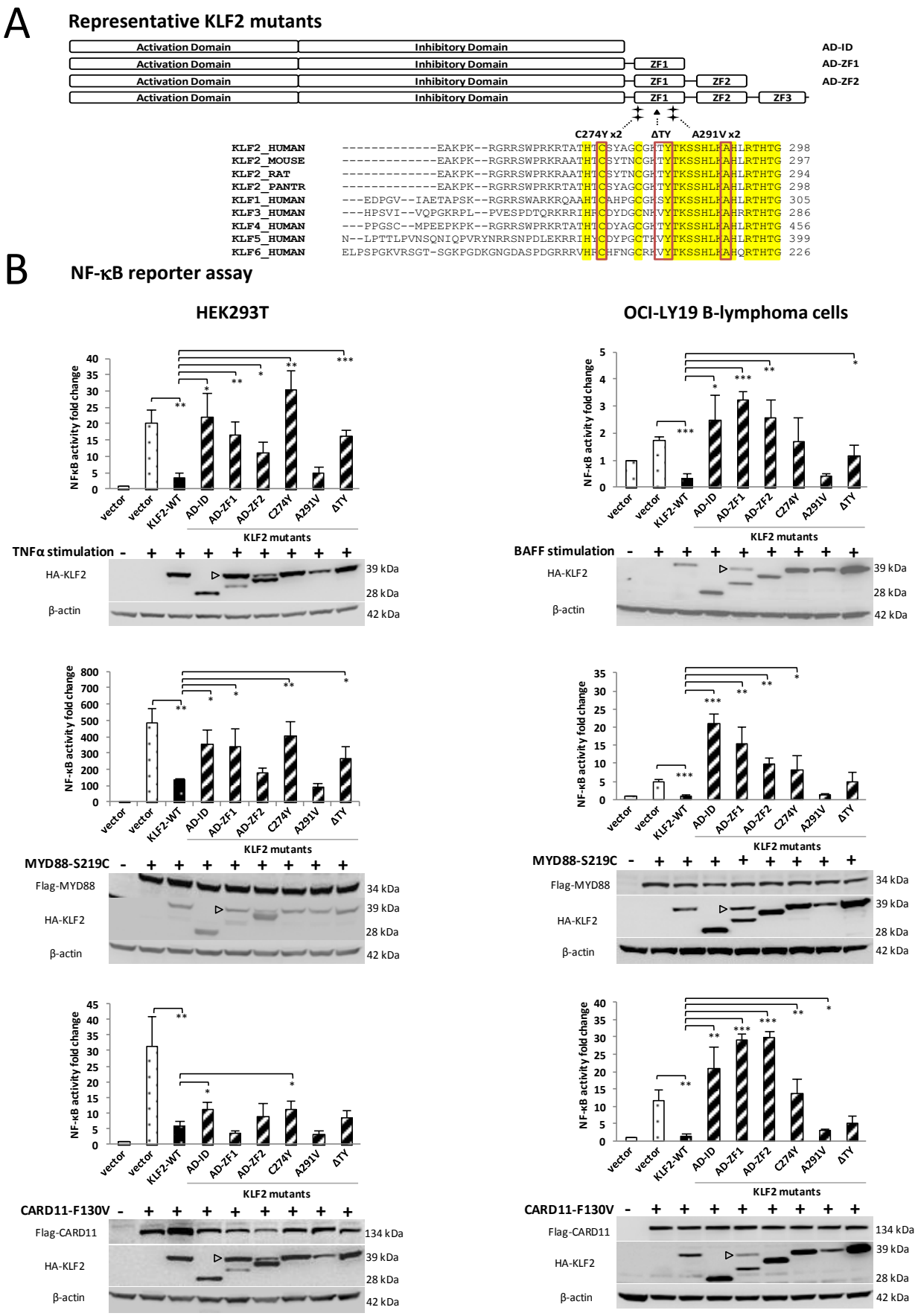


Figure 2: *KLF2* mutations and their functional characterisation.

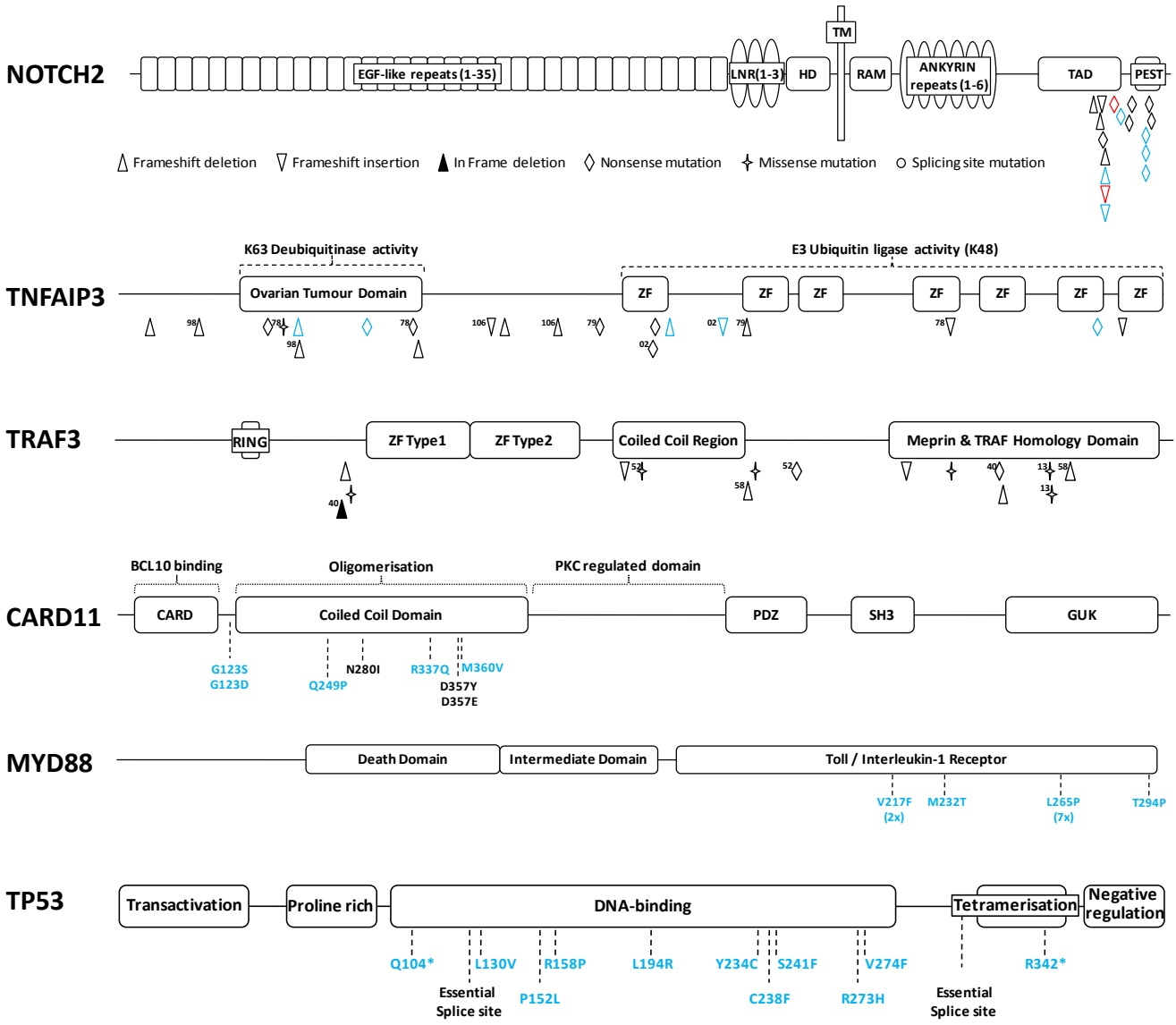
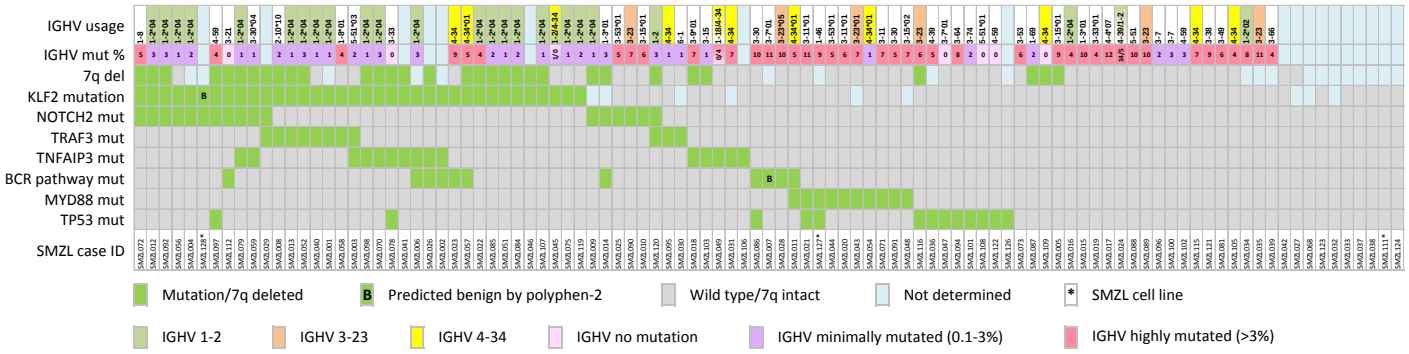
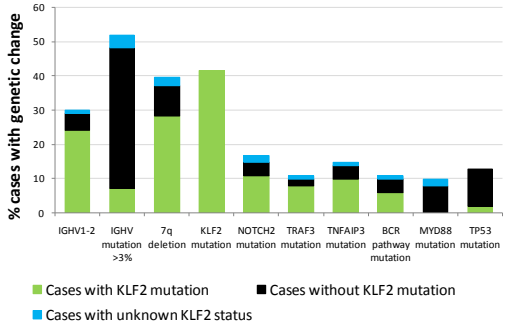


Figure 3: Nature and distribution of mutations in *NOTCH2*, *TNFAIP3*, *TRAF3*, *CARD11*, *MYD88* and *TP53* in SMZL.

A



B



C

	IGHV1-2	7q deletion	KLF2 mutation	NOTCH2 mutation	TRAF3 mutation	TNFAIP3 mutation	BCR pathway mutation	MYD88 mutation
7q del	P = 0.003							
KLF2 mut	P = 1.02×10^{-7}	P = 7.33×10^{-7}						
NOTCH2 mut	P = 0.134	P = 0.007	P = 0.007					
TRAF3 mut	P = 0.159	P = 0.505	P = 0.012	P = 1				
TNFAIP3 mut	P = 0.727	P = 0.028	P = 0.015	P = 1	P = 1.000			
BCR pathway mut	P = 0.266	P = 0.426	P = 0.147	P = 1	P = 0.596	P = 0.167		
MYD88 mut	P = 0.028	P = 0.010	P = 0.021	P = 0.205	P = 0.596	P = 0.351	P = 1	
TP53 mut	P = 0.015	P = 0.732	P = 0.071	P = 0.690	P = 0.350	P = 0.685	P = 0.352	P = 0.613

Red cell: positive correlation; black cell: negative correlation

Figure 4: Correlation of *KLF2* mutation with other genetic changes in SMZL.

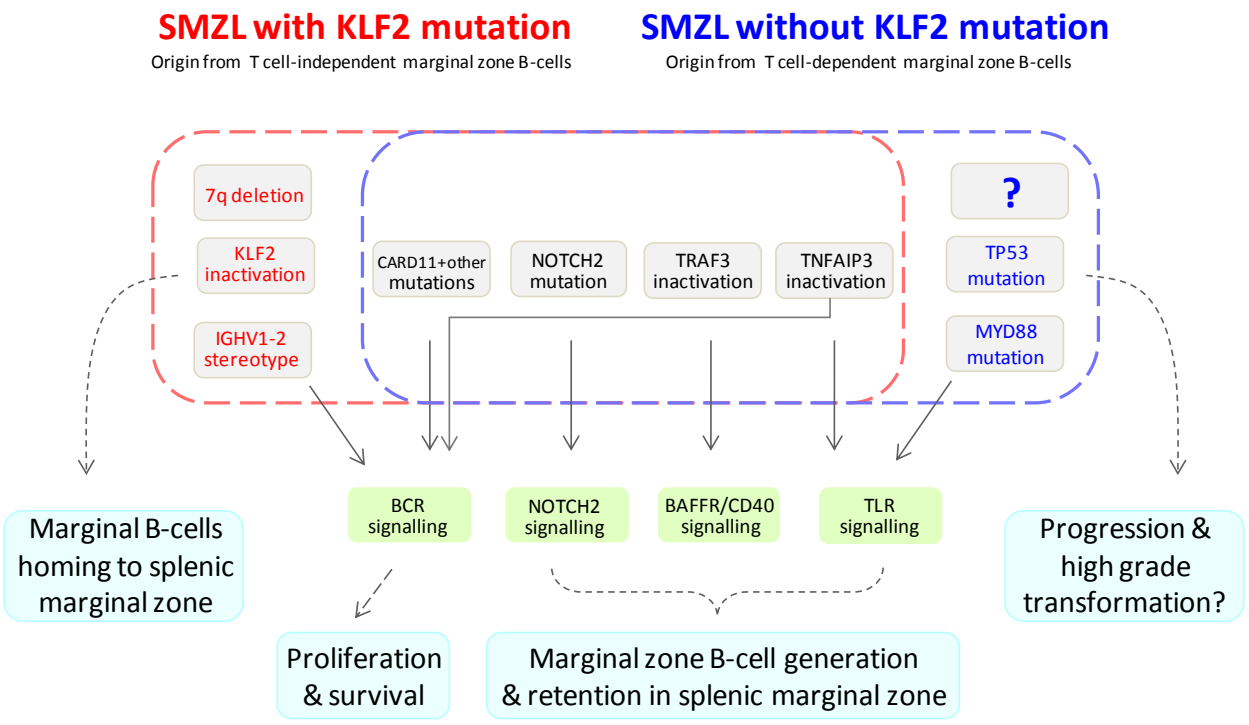


Figure 5: A summary of the proposed molecular mechanism of SMZL.