

Novel *GPR34* and *CCR6* mutation and distinct genetic profiles in MALT lymphomas of different sites

Sarah Moody,¹ Joe Sneath Thompson,¹ Shih-Sung Chuang,² Hongxiang Liu,³ Markus Raderer,⁴ George Vassiliou,⁵ Iwona Wlodarska,⁶ Fangtian Wu,¹ Sergio Cogliatti,⁷ Alistair Robson,⁸ Margaret Ashton-Key,⁹ Yingwen Bi,¹⁰ John Goodlad,¹¹ Ming-Qing Du^{1,3,12}

¹Division of Cellular and Molecular Pathology, Department of Pathology, University of Cambridge, Cambridge, UK;

²Department of Pathology, Chi-Mei Medical Centre, Tainan, Taiwan;

³Molecular Malignancy Laboratory, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK;

⁴Department of Medicine I, Clinical Division of Oncology, Medical University of Vienna, Austria;

⁵The Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom

⁶Center for Human Genetics, KU Leuven, Belgium;

⁷Institute of Pathology, State Hospital St. Gallen, St. Gallen, Switzerland;

⁸Department of Dermatopathology, St John's Institute of Dermatology, London, UK;

⁹Department of Cellular Pathology, Southampton University Hospitals National Health Service Trust, Southampton, UK;

¹⁰Department of Pathology, Eye & ENT Hospital, Fudan University, Shanghai, PR China;

¹¹Department of Pathology, Western General Hospital, NHS Lothian University Hospitals Trust, Edinburgh, UK;

¹²Department of Histopathology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK;

Author contributions: Experimental design, data collection and analysis: SM, HL, FW; Illumina sequencing analysis and variant calling: JST, GV; Case contribution: SSC, MR, IW, SC, AR, MAK, YB, JG; Manuscript writing and preparation: MQD, SM; Research funding, study design and coordination: MQD. All authors commented on the manuscript and approve its submission for publication. The authors declare no conflict of interest.

Running title: Novel GPCR mutation in MALT lymphoma

Key words: GPR34, CCR6, mutation, IGHV usage, MALT lymphoma

Correspondence to

Professor Ming-Qing Du,

Division of Cellular and Molecular Pathology,

Department of Pathology, University of Cambridge

Box 231, Level 3, Lab Block, Addenbrooke's Hospital,

Hills Road, Cambridge, CB2 2QQ, United Kingdom

Tel: +44 (0)1223 767092; Fax: +44 (0)1223 586670; Email: mqd20@cam.ac.uk

Word Count: Abstract: 248; the main manuscript text: 3610

Manuscript figure and tables: 4 Figures, 5 supplementary figures, 8 supplementary tables

Acknowledgements: The authors would like to thank David Withers, Cambridge Genomics Services and Graeme Clark (Medical Genetics) for technical assistance with sequencing. The research was supported by grants from Bloodwise (13006, 15002, 15019) UK, and Kay Kendal Leukaemia Fund (KKL582), UK. SM was initially supported by a PhD studentship from Medical Research Council, Department of Pathology, University of Cambridge and Addenbrooke's Charitable Trust.

ABSTRACT

MALT lymphoma originates from a background of diverse chronic inflammatory disorders at various anatomic sites. The genetics underlying its development, particularly in those associated with autoimmune disorders, is poorly characterised. By whole exome sequencing of 21 cases of MALT lymphomas of the salivary gland and thyroid, we have identified recurrent somatic mutations in 2 G-protein coupled receptors (*GPR34* and *CCR6*) not previously reported in human malignancies, 3 genes (*PIK3CD*, *TET2*, *TNFRSF14*) not previously implicated in MALT lymphoma, and a further 2 genes (*TBL1XR1*, *NOTCH1*) recently described in MALT lymphoma. The majority of mutations in *GPR34* and *CCR6* were nonsense and frameshift changes clustered in the C-terminal cytoplasmic tail, and would result in truncated proteins that lack the phosphorylation motif important for β -arrestin mediated receptor desensitization and internalisation. Screening of these newly identified mutations, together with previously defined genetic changes, revealed distinct mutation profiles in MALT lymphoma of various sites, with those of salivary gland characterised by frequent *TBL1XR1* and *GPR34* mutations, thyroid by frequent *TET2*, *TNFRSF14* and *PIK3CD* mutations, and ocular adnexa by frequent *TNFAIP3* mutation. Interestingly, in MALT lymphoma of the salivary gland, there was a significant positive association between *TBL1XR1* mutation and *GPR34* mutation/translocation ($P=0.0002$). In those of ocular adnexa, *TBL1XR1* mutation was mutually exclusive from *TNFAIP3* mutation ($P=0.049$), but significantly associated with IGHV3-23 usage ($P=0.03$) and *PIK3CD* mutation ($P=0.009$). These findings unravel novel insights into the molecular mechanisms of MALT lymphoma and provide further evidence for potential oncogenic cooperation between receptor signalling and genetic changes.

INTRODUCTION

Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue (MALT lymphoma) originates from acquired MALT at diverse anatomic sites. Accordingly, its development is closely associated with distinct inflammatory disorders. For example, MALT lymphomas of the stomach and ocular adnexa are associated with chronic infection by *Helicobacter pylori* and *Chlamydia psittaci* respectively, while those of the salivary gland and thyroid are associated with Sjögren syndrome and Hashimoto thyroiditis respectively.¹ In addition to its diverse aetiology, MALT lymphomas of various sites also show distinct biased usage of IGHV. For instance, up to a third of ocular adnexal MALT lymphomas have IGHV4-34 BCR that is most likely auto-reactive, binding to the carbohydrate I/i antigens, while more than 50% of salivary gland MALT lymphomas bear IGHV1-69 BCR with features of rheumatoid factors.²⁻⁹ These findings suggest that various chronic inflammatory disorders may generate immune responses, which preferentially expand B-cells with certain properties, such as autoreactive BCR, consequently leading to the development of MALT lymphoma.^{1,10}

There are a number of genetic changes described in MALT lymphomas, but many of these changes occur at considerable variable frequencies at different sites despite that they commonly affect the NF- κ B pathway.¹ For example, t(11;18)(q21;q21)/API2-MALT1 and to a lesser extent t(1;14)(p22;q32)/BCL10-IGH are frequently seen in MALT lymphoma of the lung and stomach, but rarely or not in those of the salivary gland and thyroid.¹¹⁻¹³ In contrast, *TNFAIP3* mutation and/or deletion, which encodes a NF- κ B negative regulator, is frequent in MALT lymphoma of the ocular adnexa, but occurs at low frequencies in those of other sites.¹⁴⁻¹⁶ The reason for such dramatic differences in the genetics of MALT lymphomas of different sites is unclear, and it remains to be investigated whether the occurrence of these genetic changes is influenced by the distinct background disorders or vice versa.

The development of MALT lymphoma is the result of oncogenic cooperation between immunological drive and acquired genetic changes as neither the above genetic changes nor the immunological drive alone is sufficient for malignant transformation. Nevertheless, the exact mechanisms of oncogenic cooperation in MALT lymphoma of various sites remain largely elusive. In a recent study, we found a significant association between *TNFAIP3* inactivation and biased IGHV4-34 usage in ocular adnexal MALT lymphoma, arguing for their cooperation in sustaining chronic BCR signalling, and thus NF- κ B activation.¹⁷ Interestingly, a very similar finding, namely a significant association between biased usage of IGHV1-2 and inactivating mutation of *KLF2* (a negative regulator of NF- κ B), was previously reported in splenic marginal zone lymphoma.¹⁸ Taken together, such co-operation between antigenic stimulation and genetic abnormalities may represent a common feature in marginal zone B-cell lymphoma. Nonetheless, this has not been fully investigated in marginal zone lymphoma of various sites, particularly in those of the salivary gland and thyroid, where very few genetic features have been identified. In the present study, we first investigated the mutation profile of 21 MALT lymphomas of the salivary gland and thyroid by whole exome sequencing (WES), then validated the recurrent mutations in MALT lymphoma of various sites by targeted sequencing and finally investigated their association with IGHV usage.

METHODS

Patient samples and DNA extraction

A total of 249 cases of MALT lymphoma were included in this study, originating from the ocular adnexa ($n=115$), salivary gland ($n=58$), stomach ($n=36$), thyroid ($n=13$), lung ($n=13$), and other sites ($n=14$). 179 of these cases were the subject of a previous study for somatic mutation in 17 genes.¹⁷ Genomic DNA was extracted from tumour rich areas (>40%) of formalin fixed paraffin embedded (FFPE) lymphoma biopsies in 217 cases, and where possible from non-neoplastic cells by

microdissection using the QIAamp DNA micro kit (Qiagen, Netherlands). Additionally, high molecular weight DNA was available in 32 cases of MALT lymphoma. DNA quality was assessed by PCR amplification of variably sized genomic fragments, with those amenable to PCR of ≥ 300 bp genomic fragment used for mutation analyses by targeted sequencing.¹⁹ Among the cases included in this study, clinical information such as site involvement and treatment details were available in 98 cases of ocular adnexal MALT lymphomas, and their correlation with genetic changes was examined.

Local ethical guidelines for the use of archival tissues for research with the approval of the ethics committees of the involved institutions.

Exome sequencing:

A total of 21 MALT lymphomas (14 from salivary gland, 7 from thyroid) together with 1 matched non-neoplastic biopsy were investigated by whole exome sequencing (Table S1). The initial 2 cases (1 thyroid and 1 salivary gland MALT lymphoma) were carried out on high molecular weight tumour DNA at the Eastern Sequence and Informatics Hub (EASIH). Genomic libraries were prepared using the Illumina Truseq DNA sample preparation v2 kit (Illumina, California, USA), captured with the Nimblegen SeqCap EZ Exome v3 (Roche, Germany) and sequenced on an Illumina HiSeq platform. The remaining 19 cases (6 thyroid and 13 salivary gland MALT lymphoma) together with 1 matched non-neoplastic sample were performed on FFPE tissue DNA at the Wellcome Trust Sanger Institute. Genomic libraries were generated using the Illumina Paired End Sample Prep Kit, enriched using the Agilent SureSelect Human All Exon 50Mb kit (Agilent, California, USA), and sequenced on an Illumina HiSeq platform using a 75bp paired end protocol. Sequencing reads were aligned to the hg19 reference genome using BWA with default settings, with single nucleotide variants called by CaVEMan (Cancer Variants through Expectation Maximisation), and insertions and deletions by PINDEL. Data were filtered against 300 unmatched normal controls available from the Cancer

Genome Project in addition to the 1000 genomes project and Ensembl variation databases to remove known single nucleotide polymorphisms (SNP).

Somatic variant validation by Fluidigm Access Array PCR and Illumina MiSeq Sequencing

Where indicated, the novel and shortlisted variants identified by WES were confirmed by PCR and Sanger Sequencing, and their somatic origin ascertained by PCR and Sanger sequencing analysis of DNA samples from microdissected non-neoplastic cells.

Mutations in *CCR6*, *FGFR3*, *FOXO1*, *GPR34*, *IKBKB*, *PIK3CD*, *NOTCH1*, *TBL1XR1*, *TET2* and *TNFRSF14* were then screened using the Fluidigm Access Array PCR and Illumina MiSeq protocol as previously described.¹⁹ Primer sequences and PCR conditions are shown in Table S2. In addition, mutation in a further 17 genes including BCR signalling (*CD79A*, *CD79B*, *CARD11*), NF- κ B (*TNFRSF11A*, *TNFAIP3*, *TRAF3*), TLR signalling (*MYD88*), plasma cell differentiation (*PRDM1*), histone modifiers (*CREBBP*, *EP300*, *EZH2*, *MLL2*, *MEF2B*, *KDM2B*), antigen presentation (*B2M*, *CD58*), and apoptosis (*TP53*) were similarly screened in the cases (n=70) that were not investigated previously.¹⁷ Briefly, DNA samples were sequenced in duplicate and analysed using an in-house variant calling protocol optimised against a large panel of known mutations.¹⁹ Variants that appeared in both replicates at $\geq 10\%$ AAF (alternative allele frequency) were further validated by inspection of their bam files, and then considered as true genetic changes.¹⁹ Where possible, variants were confirmed as somatic by PCR and Sanger sequencing of non-neoplastic DNA or assumed to be somatic if previously reported as somatic in the COSMIC database or other samples in the cohort.

Fluorescent In Situ Hybridization (FISH) analysis

This was performed on FFPE tissue sections with Vysis LSI IGH Dual Colour Break Apart Probe (Abbott Molecular, Illinois, USA) and an in house GPR34 Dual Colour Break Apart probe as previously described.^{16,20} To generate the in house GPR34 break-apart probe, BAC clones RP11-643E21 and RP11-524P6 centromeric to *GPR34*, and RP11-360E17 and CTD-3202J9 telomeric to *GPR34* were amplified using the Templphi kit (GE Healthcare, Illinois, USA) and then labelled with SpectrumOrange and SpectrumGreen respectively using nick translation with random priming (Abbott Molecular, Illinois, USA).^{16,20}

Sequencing analysis of the rearranged immunoglobulin heavy chain genes

The sequence of clonally rearranged immunoglobulin heavy chain genes was available in 94 cases from a previous study.¹⁷ Additional PCR and sequencing of the rearranged immunoglobulin heavy chain genes were performed for the cases investigated by whole exome sequencing using the BIOMED-2 VH FR1-JH or FR2-JH primers as described previously.¹⁷ Sequences were analysed using IMGT/V-Quest software (www.imgt.org/IMGT_vquest/vquest), and successfully annotated in a total of 101 cases (Table S3).

Statistical analyses: Fishers exact test was used to test for associations between categorical variables.

RESULTS

Novel mutations in MALT lymphomas identified by WES

WES was successful in a total of 21 MALT lymphomas (14 from salivary gland, 7 from thyroid) together with 1 matched non-neoplastic DNA (Table S4). After filtering known SNPs, synonymous changes and benign variants by Polyphen2, 72 variants were seen in the case with matched germline

DNA, while an average of 111 variants (range 46-264/case) were observed in the remaining cases without matched germline DNA (Figure S1). Based on their frequencies, potential functional impact, involvement in cancer and lymphocyte biology and hits against relevant GO terms, novel variants in 10 genes were selected for validation and their somatic origin was confirmed by Sanger sequencing of corresponding non-neoplastic DNA (Table S5). These shortlisted genes included 2 G protein-coupled receptors (*GPR34*, *CCR6*) not yet reported as a mutation target in human malignancies, 6 (*FGFR3*, *FOXO1*, *IKBKB*, *PIK3CD*, *TET2*, *TNFRSF14*) not yet implicated in MALT lymphoma, and 2 (*TBL1XR1*, *NOTCH1*) recently described in MALT lymphoma (Table S4, Figure S2).²¹ All of these mutations identified by WES were potentially pathogenic (please see section below). To further investigate their mutation frequency and characteristics in MALT lymphoma, we screened a large cohort from various anatomic sites for mutations by Fluidigm Access Array PCR and Illumina MiSeq sequencing.

Variable involvement of newly identified mutations in MALT lymphoma of different sites

A total of 249 cases of MALT lymphoma from the ocular adnexa ($n=115$), salivary gland ($n=58$), stomach ($n=36$), thyroid ($n=13$), lung ($n=13$), and other sites ($n=14$) were investigated for somatic mutations in 27 genes with the 10 genes mentioned above investigated exclusively by the present study, and 17 genes studied by previous ($n=149$) and present ($n=70$) studies respectively (Table S6 & 7).¹⁷ Recurrent genetic changes are shown in Figure 1 (Figure S3). Interestingly, the findings showed remarkable variations in their involvement among MALT lymphoma of different sites.

GPR34 mutation was almost exclusively found in MALT lymphoma of salivary gland (9/56=16%), with exception of a single case from ocular adnexa (Figure 1). The majority of *GPR34* mutations were nonsense changes (6) and frameshift indel (1) that were clustered in the C-terminal regulatory

regions, resulting in truncated proteins (Figure 2). The remaining 3 mutations were missense changes, namely R84H, D151A and Y327N.

GPR34 is a G-protein coupled receptor (GPCR) and is deregulated by t(X:14)(p11.4;q32) in MALT lymphoma of salivary gland and lung with a history of Sjögren syndrome.^{20,22,23} In view of this finding, we screened salivary gland MALT lymphoma for t(X;14)(p11;q32) by interphase FISH, and identified the translocation in 2/58 (3%) cases (Figure S4). Interestingly, *GPR34* mutation and translocation were mutually exclusive, together seen in 11/58 (19%) cases of salivary gland MALT lymphoma (Figure 1, Figure S4).

TBL1XR1 mutation was most frequent in MALT lymphoma of salivary gland (14/58=24%), occurring less frequently in MALT lymphomas of the stomach (3/36=8%), lung (1/13=8%), ocular adnexa (7/115=6%), skin (1/8 =13%) and in a single case from tonsil, but not in the thyroid (Figure 1). A total of 30 mutations were found in 27 cases, with 3 cases affected by two mutations. The majority of these mutations (20 missense changes and 5 inframe deletions) were within the WD40 domains and frequently affected regions or residues of structural importance, which are likely critical for interaction with NCoR (Figure 3). The remaining 5 mutations included 3 substitution changes at exon 11 splice site, 1 frameshift deletion and 1 nonsense change, most likely resulting in truncated proteins with potential increased binding to UbcH5.²⁴

TET2 mutation was found in the majority of MALT lymphomas of the thyroid (8/13=62%), and also recurrently seen in those of salivary gland (5/58=9%), ocular adnexa (5/115=4%), stomach (3/36=8%) and lung (1/13=8%) (Figure 1). A total of 31 *TET2* mutations were detected in 23 cases, with 8 cases showing multiple mutations (Figure 2). Interestingly, 7/8 of the cases with multiple *TET2* mutations were MALT lymphomas of the thyroid. The vast majority of these mutations were deleterious changes including 11 frameshift indels, 4 nonsense, 1 splice site and 15 missense mutations. Of

these, the majority of frameshift (8/11) and nonsense (3/4) mutations were found in MALT lymphomas of the thyroid.

To our surprise, *TNFRSF14* mutation was highly frequent in MALT lymphoma of thyroid (6/13=46%), albeit infrequent in those of other sites (Figure 1). A total of 21 *TNFRSF14* mutations were detected in 18 cases, and all mutations were deleterious changes comprising 6 nonsense, 2 frameshift indel, 1 inframe deletion, 11 missense changes including 2 in the first codon thus abolishing the translation start site, and a single splice site mutation (Figure 2). As *TNFRSF14* mutation is common in paediatric-type follicular lymphoma and the mutations identified are of similar nature,²⁵ we carefully reviewed the histopathology and immunophenotype of our cases with *TNFRSF14* mutations, where necessary performed further immunohistochemistry for CD10 and BCL6, and also FISH for *BCL2* and *BCL6* translocation. These additional analyses confirmed the original diagnosis of MALT lymphoma in these cases.

PIK3CD was another gene frequently mutated in MALT lymphoma of thyroid (3/13=23%), followed by salivary gland (5/58=9%), but rarely in those of other sites (Figure 1). A total of 11 *PIK3CD* mutations were identified in 11 MALT lymphomas, and the vast majority of these mutations were missense changes including N334K and E1021K, previously reported in patients with activated PI3K δ syndrome (hyper IgM syndrome) and in DLBCL (Figure 2).²⁶⁻²⁹

Mutation in the remaining genes was relevantly infrequent, without any apparent bias among the sites examined. Mutation in *CCR6* was recurrent, and seen in a total of 7 MALT lymphomas, with 7 mutations including 5 deleterious changes (3 nonsense, 3 frameshift indels) and 2 missense changes, largely clustered in the C-terminal regulatory region (Figures 1&2). Interestingly, mutations in *GPR34* and *CCR6*, both members of the GPCR family, were mutually exclusive.

Distinct association of genetic changes in MALT lymphoma of various sites

Correlation analysis revealed several distinct associations among genetic changes in MALT lymphoma of different sites. In salivary gland MALT lymphoma, there was a significant association between *TBL1XR1* mutation and *GPR34* mutation/translocation ($P=0.0002$), with *TBL1XR1* mutation seen in 8/11 (73%) cases with *GPR34* mutation/translocation, but only in 6/47 (13%) cases without *GPR34* genetic abnormalities (Figure 4, Figure S5). Both *GPR34* and *TBL1XR1* genetic changes were more frequently seen in the cases with non-IGHV1-69 than those with IGHV1-69 rearrangement (29% vs 9%; 57% vs 18% respectively) although this does not reach statistical significance.

In ocular adnexal MALT lymphoma, *TBL1XR1* mutation was mutually exclusive from *TNFAIP3* mutation ($P=0.049$), but significantly associated with *PIK3CD* mutation ($P=0.009$), with *TBL1XR1* mutation seen in 2/3 (67%) cases with *PIK3CD* mutation, but in only 5/112 (4%) cases without *PIK3CD* mutation (Figure 4, Figure S5). *TBL1XR1* mutation was also mutually exclusive from IGHV4-34 rearrangement, but significantly associated with IGHV3-23 usage ($P=0.03$), with *TBL1XR1* mutation seen in 3/11 (27%) cases with IGHV3-23 rearrangement, but in only 2/52 (4%) cases with other rearrangements (Figure 4). *MYD88* mutation also appeared to be more frequent in *TBL1XR1* mutated cases (2/7=29%) compared to non-mutated cases (6/108= 6%), although this did not reach significance ($P=0.07$).

In gastric MALT lymphoma, t(11;18) API2-MALT1 was mutually exclusive from other genetic changes (Figure S5, $P=0.025$).

Correlation between genetic changes and clinicopathological parameters

This was carried out in 98 cases of ocular adnexal MALT lymphoma where clinicopathological data were available (Table S8). *TBL1XR1* mutation and IGHV3-23 usage were significantly associated with involvement of the conjunctiva ($P=0.002$ in each respectively), while *TNFRSF14* mutation was significantly associated with involvement of the orbit ($P=0.04$). *MYD88* and *TP53* mutations were much higher in cases involving both orbit and conjunctiva than those involving only a single site but only the former showing a statistical significance (4/23=17% vs 2/75=3%, $P=0.03$). With the exception of *TNFAIP3* deletion/mutation that was significantly associated with higher radiation dosages to achieve complete remission as reported previously,¹⁶, none of the other genetic changes showed any significant association with radiation dosages.

DISCUSSION

By whole exome sequencing of 21 cases of MALT lymphomas of the salivary gland and thyroid, we have identified several novel genetic changes in MALT lymphoma, including *GPR34* and *CCR6* mutations not yet previously reported in human malignancies. Screening these newly identified mutations revealed distinct mutation profiles in MALT lymphoma of various sites, and provide further evidence of potential oncogenic cooperation between receptor signalling and genetic changes.

***GPR34* and *CCR6* mutations in MALT lymphoma**

Both *GPR34* and *CCR6* are members of class A G protein-coupled receptor (GPCR) superfamily, which transduce extracellular stimulation into intracellular signalling through G protein and β -arrestin. By interacting with G protein, GPCR can activate diverse signalling pathways critical for cellular function. This receptor signalling is tightly regulated through a process known as desensitization that is mediated by GPCR phosphorylation, allowing its binding to β -arrestin which triggers receptor

internalisation, consequently dampening intracellular signalling.³⁰ A common phosphorylation code (motif) has been identified in the C-terminal cytoplasmic tail of most GPCRs, and its phosphorylation enables electrostatic interactions with β -arrestin.³¹

GPR34 contains 7 transmembrane domains, followed by a C-terminal cytoplasmic tail. The majority of *GPR34* mutations identified in MALT lymphoma are nonsense changes or frameshift indels that are clustered in its C-terminal region, resulting in truncated proteins. A phosphorylation code (motif) (S³⁵¹S³⁵³T³⁵⁶) is seen in the C-terminal cytoplasmic tail of GPR34 according to Zhou et al,³¹ and all the above nonsense and frameshift changes would eliminate or impair these phosphorylation sites, thus potentially deregulating the receptor desensitization process.

The remaining *GPR34* mutations are missense changes including R84H, D151A and Y327N. R84H affects the tri-basic motif (RKR) in the first intracellular loop, which is the key topogenic signal determining the orientation of the first transmembrane domain.³² D151A affects the highly conserved E/DRY motif and is predicted to cause the receptor constitutive activation.³³ Y327N is close to the interface of the seventh TM domain and the cytoplasmic tails, and is predicted to be highly damaging by PolyPhen-2, although its potential functional impact is unclear.

Similarly, *CCR6* also contains 7 transmembrane domains, followed by a C-terminal cytoplasmic tail, and the majority of *CCR6* mutations are nonsense or frameshift changes that are clustered in its C-terminal region. A phosphorylation code (motif) (S³⁵⁷T³⁶⁰T³⁶³) is present in the C-terminal cytoplasmic tail of *CCR6*,³¹ and all the nonsense and frameshift changes identified would eliminate these phosphorylation sites, potentially impairing the receptor desensitization process. The functional impact of the two *CCR6* missense mutations (R159S and Y352C) is unclear.

The nature and distribution of mutations in *GPR34* and *CCR6* identified in this study are very similar to those seen in *CXCR4* in Waldenström's macroglobulinaemia and WHIM syndrome, as well as those found in *CCR4* in adult T cell leukaemia/lymphoma, which have been shown to be gain-of-function changes.³⁴⁻³⁶ Despite their similar mutation pattern, *GPR34* and *CCR6* have distinct ligands, thus respond to different environmental cues, potentially explaining their differential involvement in MALT lymphoma of different sites. *GPR34* has been shown to be triggered by lysophosphatidylserine, while *CCR6* is activated by CCL20.³⁷⁻³⁹ Over-expression of wild type *GPR34* *in vitro* resulted in ERK, PI3K/AKT and PKC signalling, inducing AP1 and NF-κB-mediated gene transcription.^{20,40} It remains to be investigated whether the above *GPR34* and *CCR6* mutations impact similar downstream signalling pathways, and if they would enhance these intracellular signalling pathways through impaired receptor desensitization and internalisation or via independent mechanisms.

Distinct genetic profile in MALT lymphoma of various sites.

The present study provides further evidence showing distinct genetic profiles among MALT lymphoma of various sites (Figure 1). MALT lymphoma of the salivary gland features frequent mutation of *TBL1XR1* and *GPR34*, while those of the thyroid are characterised by frequent mutations in *TET2*, *TNFRSF14* and *PIK3CD*. MALT lymphoma of the ocular adnexa is noted for frequent *TNFAIP3* mutation. In contrast, MALT lymphomas of the stomach and lung are distinguished by a high prevalence of t(11;18)(q21;q21).

The present study also unravels several distinct associations among genetic changes in MALT lymphoma. In MALT lymphoma of the salivary gland, there is a significant positive association between *TBL1XR1* mutation and *GPR34* mutation/translocation. The effect of *TBL1XR1* mutations has not been fully characterised, but a previous report suggests that mutations enhance *TBL1XR1*

binding to NCoR and facilitate its degradation, consequently promoting NF- κ B and JUN target gene expression.⁴¹ Thus, *TBL1XR1* mutation may sustain *GPR34* mediated NF- κ B and AP1 activation via this mechanism.

Why *GPR34* mutation and translocation are preferentially associated with MALT lymphoma of the salivary gland is unclear. *GPR34* has been shown to be activated by lyso-phosphatidylserine, which could be generated by hydrolysis of the exposed membrane lipids in apoptotic cells through phospholipase A.^{37,38} Among MALT lymphoma of various sites, the lymphoepithelial lesions are most prominent in those of salivary gland. It remains to be investigated whether such lymphoepithelial lesions may facilitate the production of lyso-phosphatidylserine, the ligand for *GPR34*, thereby facilitating the expansion and localisation of B-cells carrying *GPR34* genetic changes surrounding the lymphoepithelial lesion.

In MALT lymphoma of the ocular adnexa, *TBL1XR1* mutation is mutually exclusive from *TNFAIP3* mutation and IGHV4-34 usage, but significantly associated with IGHV3-23 usage and *PIK3CD* mutation, albeit with limited numbers of cases involved. Like IGHV4-34 rearrangements, the IGHV3-23 rearrangement seen in ocular adnexal MALT lymphoma also encodes autoreactive BCR as shown by recombinant antibody studies.⁵ In this context, the significant association between *TBL1XR1* mutation and IGHV3-23 rearrangement is very much analogous to that between *TNFAIP3* mutation and IGHV4-34 rearrangement reported previously,¹⁷ hence expanding the evidence supporting oncogenic cooperation between antigenic drive and acquired genetic changes in MALT lymphoma. Similarly, the significant association between *TBL1XR1* and *PIK3CD* mutation may also signify oncogenic cooperation between the two genetic events.

In summary, our study reveals several novel genetic changes including *GPR34* and *CCR6* mutations not yet reported in human malignancies, and provides further evidence of distinct mutation profiles

in MALT lymphoma of various sites. These novel findings offer exciting prospects for further characterisation of the molecular pathogenesis of MALT lymphoma.

REFERENCES:

1. Du MQ. MALT lymphoma: A paradigm of NF-kappaB dysregulation. *Semin Cancer Biol.* 2016;3949-60.
2. van Maldegem F, Wormhoudt TA, Mulder MM, *et al.* Chlamydia psittaci-negative ocular adnexal marginal zone B-cell lymphomas have biased VH4-34 immunoglobulin gene expression and proliferate in a distinct inflammatory environment. *Leukemia.* 2012;26(7):1647-1653.
3. Mannami T, Yoshino T, Oshima K, *et al.* Clinical, histopathological, and immunogenetic analysis of ocular adnexal lymphoproliferative disorders: characterization of malt lymphoma and reactive lymphoid hyperplasia. *Mod Pathol.* 2001;14(7):641-649.
4. Bahler DW, Szankasi P, Kulkarni S, *et al.* Use of similar immunoglobulin VH gene segments by MALT lymphomas of the ocular adnexa. *Mod Pathol.* 2009;22(6):833-838.
5. Zhu D, Bhatt S, Lu X, *et al.* Chlamydia psittaci-negative ocular adnexal marginal zone lymphomas express self polyreactive B-cell receptors. *Leukemia.* 2015;29(7):1587-1599.
6. Dagklis A, Ponzoni M, Govi S, *et al.* Immunoglobulin gene repertoire in ocular adnexal lymphomas: hints on the nature of the antigenic stimulation. *Leukemia.* 2012;26(4):814-821.
7. Bende RJ, Aarts WM, Riedl RG, *et al.* Among B cell non-Hodgkin's lymphomas, MALT lymphomas express a unique antibody repertoire with frequent rheumatoid factor reactivity. *J Exp Med.* 2005;201(8):1229-1241.
8. Bahler DW, Miklos JA, Swerdlow SH. Ongoing Ig gene hypermutation in salivary gland mucosa-associated lymphoid tissue-type lymphomas. *Blood.* 1997;89(9):3335-3344.
9. Miklos JA, Swerdlow SH, Bahler DW. Salivary gland mucosa-associated lymphoid tissue lymphoma immunoglobulin V(H) genes show frequent use of V1-69 with distinctive CDR3 features. *Blood.* 2000;95(12):3878-3884.
10. Hamoudi RA, Appert A, Ye H, *et al.* Differential expression of NF-kappaB target genes in MALT lymphoma with and without chromosome translocation: insights into molecular mechanism. *Leukemia.* 2010;24(8):1487-1497.
11. Ye H, Liu H, Attygalle A, *et al.* Variable frequencies of t(11;18)(q21;q21) in MALT lymphomas of different sites: significant association with CagA strains of H pylori in gastric MALT lymphoma. *Blood.* 2003;102(3):1012-1018.
12. Ye H, Gong L, Liu H, *et al.* MALT lymphoma with t(14;18)(q32;q21)/IGH-MALT1 is characterized by strong cytoplasmic MALT1 and BCL10 expression. *J Pathol.* 2005;205(3):293-301.
13. Streubel B, Simonitsch-Klupp I, Mullauer L, *et al.* Variable frequencies of MALT lymphoma-associated genetic aberrations in MALT lymphomas of different sites. *Leukemia.* 2004;18(10):1722-1726.

14. Chanudet E, Ye H, Ferry J, *et al.* A20 deletion is associated with copy number gain at the TNFA/B/C locus and occurs preferentially in translocation-negative MALT lymphoma of the ocular adnexa and salivary glands. *J Pathol.* 2009;217(3):420-430.
15. Honma K, Tsuzuki S, Nakagawa M, *et al.* TNFAIP3/A20 functions as a novel tumor suppressor gene in several subtypes of non-Hodgkin lymphomas. *Blood.* 2009;114(12):2467-2475.
16. Bi Y, Zeng N, Chanudet E, *et al.* A20 inactivation in ocular adnexal MALT lymphoma. *Haematologica.* 2012;97(6):926-930.
17. Moody S, Escudero-Ibarz L, Wang M, *et al.* Significant association between TNFAIP3 inactivation and biased immunoglobulin heavy chain variable region 4-34 usage in mucosa-associated lymphoid tissue lymphoma. *J Pathol.* 2017;243(1):3-8.
18. Clipson A, Wang M, de Leval L, *et al.* KLF2 mutation is the most frequent somatic change in splenic marginal zone lymphoma and identifies a subset with distinct genotype. *Leukemia.* 2015;29(5):1177-1185.
19. Wang M, Escudero-Ibarz L, Moody S, *et al.* Somatic Mutation Screening Using Archival Formalin-Fixed, Paraffin-Embedded Tissues by Fluidigm Multiplex PCR and Illumina Sequencing. *J Mol Diagn.* 2015;17(5):521-532.
20. Ansell SM, Akasaka T, McPhail E, *et al.* t(X;14)(p11;q32) in MALT lymphoma involving GPR34 reveals a role for GPR34 in tumor cell growth. *Blood.* 2012;120(19):3949-3957.
21. Johansson P, Klein-Hitpass L, Grabellus F, *et al.* Recurrent mutations in NF-kappaB pathway components, KMT2D, and NOTCH1/2 in ocular adnexal MALT-type marginal zone lymphomas. *Oncotarget.* 2016;7(38):62627-62639.
22. Baens M, Finalet FJ, Tousseyn T, *et al.* t(X;14)(p11.4;q32.33) is recurrent in marginal zone lymphoma and up-regulates GPR34. *Haematologica.* 2012;97(2):184-188.
23. Akasaka T, Lee YF, Novak AJ, *et al.* Clinical, histopathological, and molecular features of mucosa-associated lymphoid tissue (MALT) lymphoma carrying the t(X;14) (p11;q32)/GPR34-immunoglobulin heavy chain gene. *Leuk Lymphoma.* 2017;58(9):1-4.
24. Perissi V, Aggarwal A, Glass CK, Rose DW, Rosenfeld MG. A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. *Cell.* 2004;116(4):511-526.
25. Schmidt J, Gong S, Marafioti T, *et al.* Genome-wide analysis of pediatric-type follicular lymphoma reveals low genetic complexity and recurrent alterations of TNFRSF14 gene. *Blood.* 2016;128(8):1101-1111.
26. Shibata D, Weiss LM. Epstein-Barr virus-associated gastric adenocarcinoma. *Am J Pathol.* 1992;140:769-774.
27. Elgizouli M, Lowe DM, Speckmann C, *et al.* Activating PI3Kdelta mutations in a cohort of 669 patients with primary immunodeficiency. *Clin Exp Immunol.* 2016;183(2):221-229.
28. Zhang J, Grubor V, Love CL, *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A.* 2013;110(4):1398-1403.

29. Crank MC, Grossman JK, Moir S, *et al.* Mutations in PIK3CD can cause hyper IgM syndrome (HIGM) associated with increased cancer susceptibility. *J Clin Immunol.* 2014;34(3):272-276.
30. Rajagopal S, Shenoy SK. GPCR desensitization: Acute and prolonged phases. *Cell Signal.* 2018;419-16.
31. Zhou XE, He Y, de Waal PW, *et al.* Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. *Cell.* 2017;170(3):457-469.
32. Hasegawa H, Patel N, Etehadieh E, Li P, Lim AC. Topogenesis and cell surface trafficking of GPR34 are facilitated by positive-inside rule that effects through a tri-basic motif in the first intracellular loop. *Biochim Biophys Acta.* 2016;1863(7 Pt A):1534-1551.
33. Fanelli F, De Benedetti PG, Raimondi F, Seeber M. Computational modeling of intramolecular and intermolecular communication in GPCRs. *Curr Protein Pept Sci.* 2009;10(2):173-185.
34. Lagane B, Chow KY, Balabanian K, *et al.* CXCR4 dimerization and beta-arrestin-mediated signaling account for the enhanced chemotaxis to CXCL12 in WHIM syndrome. *Blood.* 2008;112(1):34-44.
35. Cao Y, Hunter ZR, Liu X, *et al.* The WHIM-like CXCR4(S338X) somatic mutation activates AKT and ERK, and promotes resistance to ibrutinib and other agents used in the treatment of Waldenstrom's Macroglobulinemia. *Leukemia.* 2015;29(1):169-176.
36. Nakagawa M, Schmitz R, Xiao W, *et al.* Gain-of-function CCR4 mutations in adult T cell leukemia/lymphoma. *J Exp Med.* 2014;211(13):2497-2505.
37. Ikubo M, Inoue A, Nakamura S, *et al.* Structure-activity relationships of lysophosphatidylserine analogs as agonists of G-protein-coupled receptors GPR34, P2Y10, and GPR174. *J Med Chem.* 2015;58(10):4204-4219.
38. Aoki J, Nagai Y, Hosono H, Inoue K, Arai H. Structure and function of phosphatidylserine-specific phospholipase A1. *Biochim Biophys Acta.* 2002;1582(1-3):26-32.
39. Lee AY, Phan TK, Hulett MD, Korner H. The relationship between CCR6 and its binding partners: does the CCR6-CCL20 axis have to be extended? *Cytokine.* 2015;72(1):97-101.
40. Zuo B, Li M, Liu Y, *et al.* G-protein coupled receptor 34 activates Erk and phosphatidylinositol 3-kinase/Akt pathways and functions as alternative pathway to mediate p185Bcr-Abl-induced transformation and leukemogenesis. *Leuk Lymphoma.* 2015;56(7):2170-2181.
41. Jung H, Yoo HY, Lee SH, *et al.* The mutational landscape of ocular marginal zone lymphoma identifies frequent alterations in TNFAIP3 followed by mutations in TBL1XR1 and CREBBP. *Oncotarget.* 2017.

FIGURE LEGENDS:

Figure 1: Distinct mutation profiles in MALT lymphoma of various sites. A total of 27 genes including the 10 MALT lymphoma associated genes indentified in the present study were screened for mutation by targeted sequencing, but only those showing recurrent changes are presented here. API2-MALT1 denotes t(11;18)(q21;q21) from previous studies,^{11,12} *GPR34* results in MALT lymphoma of the salivary gland include both mutation and t(X;14)(p11;q32)/*GPR34-IGH*. MALT lymphomas of salivary gland are featured by frequent *TBL1XR1* and *GPR34* mutations, while those of thyroid by frequent *TET2*, *TNFRSF14* and *PIK3CD* mutations, and ocular adnexa by frequent *TNFAIP3* mutation.

Figure 2: Nature and distribution of mutations in newly identified MALT lymphoma associated genes. The site of MALT lymphoma in which mutation was identified is indicated by colour: salivary gland in blue, thyroid in red, ocular adnexa in green, stomach in orange, lung in purple and other sites in grey. TM: transmembrane domain, ABD: adaptor binding domain, RBD: ras binding domain, DSBH: double-stranded β -helix.

Figure 3. Nature and distribution of *TBL1XR1* mutations in MALT lymphoma, and their potential functional impact.

A) *TBL1XR1* mutations identified in this study. The site of MALT lymphoma in which mutation was identified is indicated by colour: salivary gland in blue, thyroid in red, ocular adnexa in green, stomach in orange, lung in purple and other sites in grey.

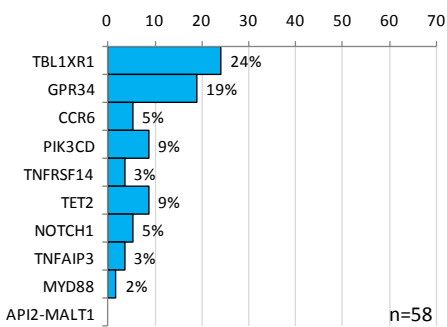
B) Detailed analyses of *TBL1XR1* mutations within the WD40 domains. *TBL1XR1* mutations identified in the present and a previous study by Jung et al are included in the analyses.⁴¹ The majority of *TBL1XR1* mutations are localised in the regions of structural importance, such as HSDW tetrad indicated by arrows and the residues predicted to be top facing (WDSP predictor, indicated in

red text). The approximate position in the HMM logo for WD40 repeats (Prosite PS00678) is shown below the sequence.

Figure 4. Significant association among genetic changes in MALT lymphoma of the salivary gland and ocular adnexa. In both cohorts, samples lacking the described changes are not included. *P* value in red colour indicates a positive correlation, while *P* value in black colour denotes a negative correlation, with bold font indicating significant values. X indicates inactivating frameshift and nonsense mutations. d indicates deletion, s indicates splice site mutations and T indicates translocation.

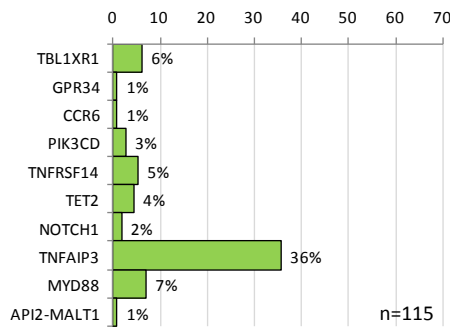
Salivary gland

Mutation Frequency (%)



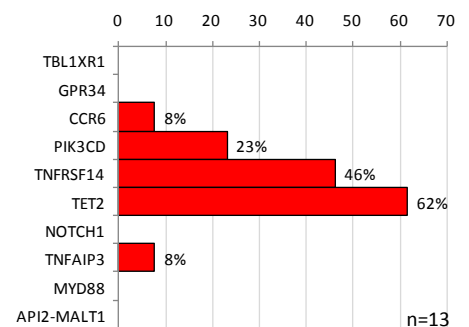
Ocular adnexa

Mutation Frequency (%)



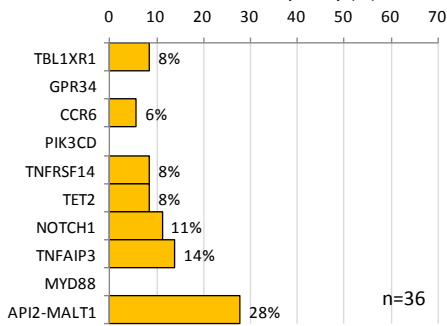
Thyroid

Mutation Frequency (%)



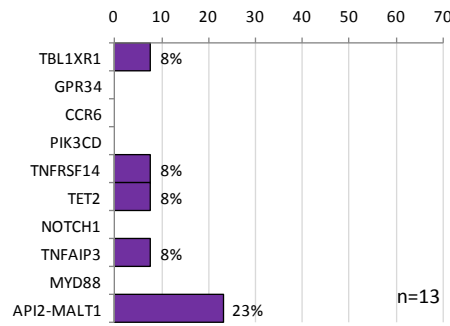
Stomach

Mutation Frequency (%)



Lung

Mutation Frequency (%)



Other

Mutation Frequency (%)

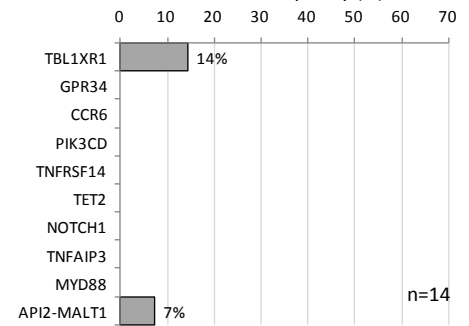


Figure 1. Distinct mutation profiles in MALT lymphoma of various sites. A total of 27 genes including the 10 MALT lymphoma associated genes identified in the present study were screened for mutation by targeted sequencing, but only those showing recurrent changes are presented here. API2-MALT1 denotes $t(11;18)(q21;q21)$ from previous studies,^{11,12} *GPR34* results in MALT lymphoma of the salivary gland include both mutation and $t(X;14)(p11;q32)/GPR34-IGH$. MALT lymphomas of salivary gland are featured by frequent *TBL1XR1* and *GPR34* mutations, while those of thyroid by frequent *TET2*, *TNFRSF14* and *PIK3CD* mutations, and ocular adnexa by frequent *TNFAIP3* mutation.

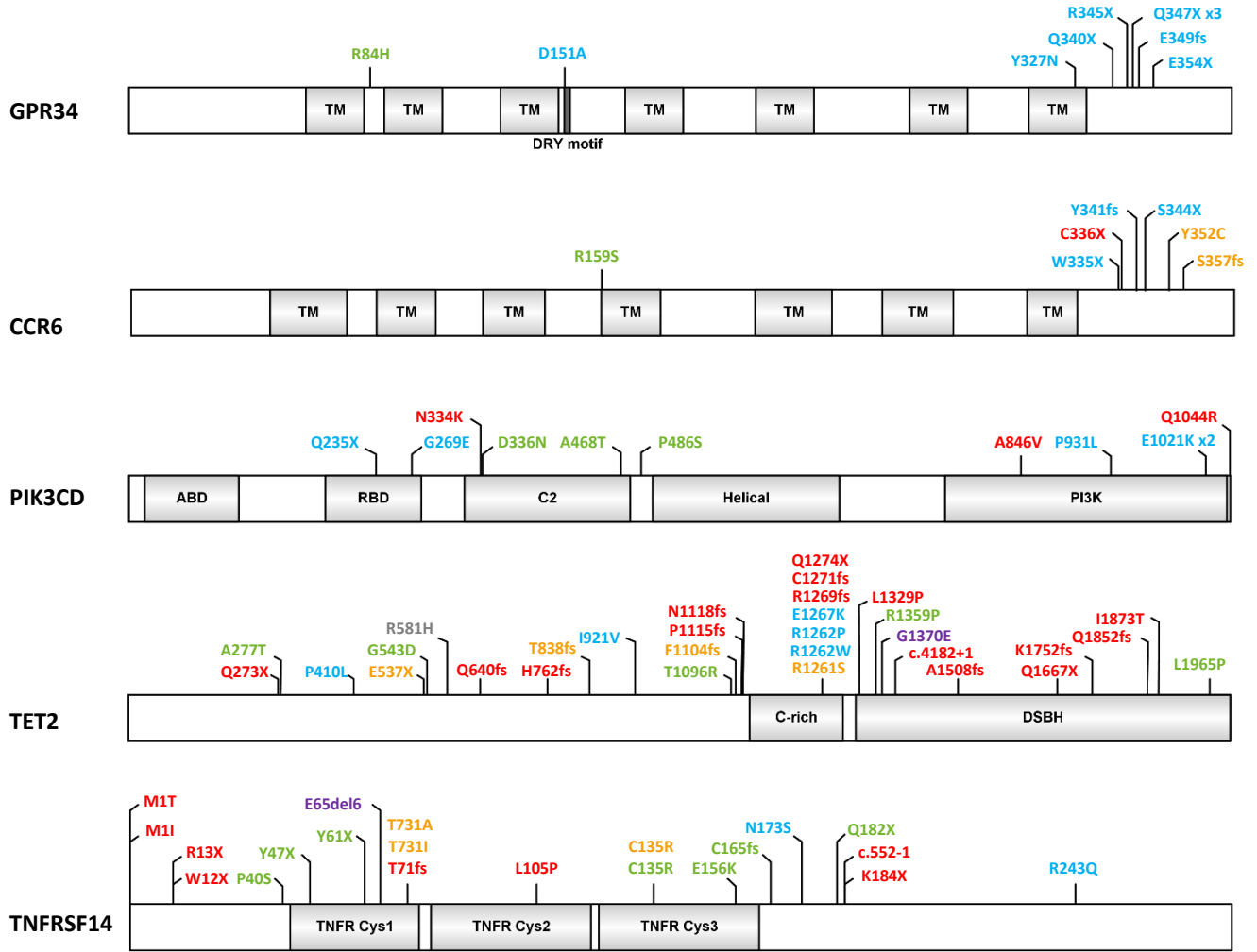


Figure 2. Nature and distribution of mutations in newly identified MALT lymphoma associated genes. The site of MALT lymphoma in which mutation was identified is indicated by colour: salivary gland in blue, thyroid in red, ocular adnexa in green, stomach in orange, lung in purple and other sites in grey. TM: transmembrane domain, ABD: adaptor binding domain, RBD: ras binding domain, DSBH: double-stranded β -helix.

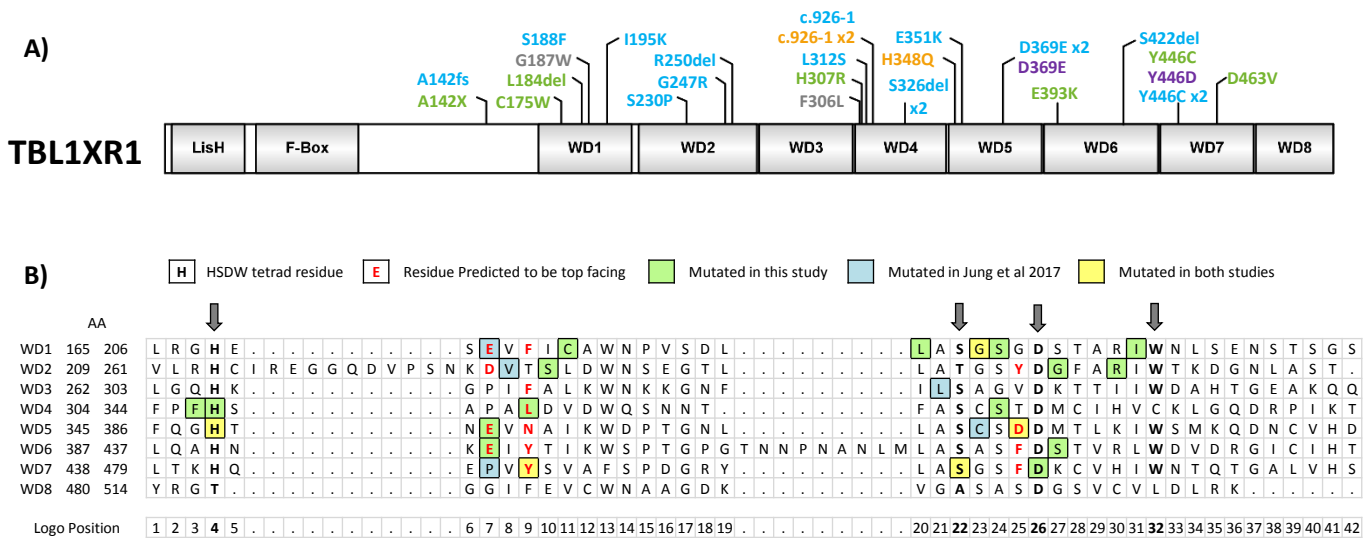


Figure 3. Nature and distribution of *TBL1XR1* mutations in MALT lymphoma, and their potential functional impact.

A) *TBL1XR1* mutations identified in this study. The site of MALT lymphoma in which mutation was identified is indicated by colour: salivary gland in blue, thyroid in red, ocular adnexa in green, stomach in orange, lung in purple and other sites in grey.

B) Detailed analyses of *TBL1XR1* mutations within the WD40 domains. *TBL1XR1* mutations identified in the present and a previous study by Jung et al are included in the analyses.⁴¹ The majority of *TBL1XR1* mutations are localised in the regions of structural importance, such as HSDW tetrad indicated by arrows and the residues predicted to be top facing (WDSP predictor, indicated in red text). The approximate position in the HMM logo for WD40 repeats (Prosite PS00678) is shown below the sequence.

