Investigation of mutations, copy number variations and translocations by capture-based Next-generation sequencing in patients with chronic B-cell lymphoproliferative disorders

D. WREN

DClinSci 2021

Investigation of mutations, copy number variations and translocations by capture-based Next-generation sequencing in patients with chronic B-cell lymphoproliferative disorders

DÖRTE WREN

A thesis submitted in partial fulfilment of the requirements of Manchester Metropolitan University for the degree of Doctor of Clinical Science

Department of Life Sciences
Manchester Metropolitan University
In collaboration with the Clinical Genomics Department, Royal Marsden NHS Foundation Trust

2021

To Martin & Simeon

(Last one, I promise!)

Für Mama & Papa und Steffen

Abstract

Patients with leukaemic indolent B-NHL who could not be assigned to a specific disease category were identified and 108 cases underwent extensive molecular characterisation using a capture-NGS sequencing approach designed by the EuroClonality NGS consortium.

Without a clear diagnosis such patients have no evidence-based, standardised treatment pathways and have no or limited access to clinical trials and novel therapies as eligibility and approval are based on distinct histological categories or WHO-defined disease entities.

Well-characterised translocations were found in 10 cases supporting a diagnosis of Mantle cell lymphoma in six patients, follicular lymphoma in two patients and atypical CLL in a further 2 cases. In a further case a rare, TRA-CDK6 fusion was detected which was recently confirmed as a recurrent, diagnostic finding in SMZL. Combination of variants in KLF2, deletion of 7q together with IGHV1-2*04 gene usage confirmed the suspected diagnosis of SMZL in one case.

The most frequently mutated genes were TP53 in 16/101 (15%) and the hotspot MYD88 p.(Leu265Pro) variant in 13/101 (13%) followed by variants in genes associated with NF-κB signalling. Integration of the molecular findings together with other (histo)pathological and clinical findings may enable the distinction between LPL and MZL-related disease categories in these patients.

NGS was capable of identifying clonal IG-rearrangements in all cases. Further stereotyping analysis using the VLeader protocol and Sanger sequencing showed most cases carried IGHV3 (52%) and IGHV4 (27%) genes with IGHV4-34 the most common (17%). Evidence of Somatic hypermutations were found in 90% of cases with 77% showing <98% homology to the germline sequence. None of the IGHV rearrangements showed a CLL-stereotype.

Preliminary analysis of copy-number variants showed a high frequency of trisomy 12 (17 cases), 17p deletion events in combination with TP53 variants as well as deletions of 7q, a recurrent feature of MZL.

This prospective study clearly demonstrates the benefit of using NGS-panel approaches for the characterization of patient cohorts which based on routine

diagnostics evaluation cannot be assigned to a defined disease category. An additional 12 patients were given a formal diagnosis based on the molecular findings and further clinical review prompted by the results of this study will likely lead to an increase in this number. The detailed information obtained about the clonal structure, including the specific V(D)J-rearrangements and combinations of molecular aberrations in each case, will enable further research into the aetiology of these diseases. From a diagnostic laboratory perspective in particular, the NGS panel utilised here is capable of reducing the time and cost of the evaluations required by combining the analysis of clonality, single-nucleotide variants, copy-number variants and structural variants in a single assay. Given the breadth of results, this NGS panel is also a useful tool for the characterisation of cohorts in a research setting and the design can be extended in future as novel markers and areas of interest are described in the literature.

Statement

The work presented here forms part of the ENABLE – NGS: Enhancing diagnosis in chronic B-cell lymphoproliferative disorders using Next-Generation Sequencing study funded by the Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and The Institute of Cancer Research, London. Principle Investigator is Dr Sunil Iyengar Consultant Haematologist, RMH.

The data included here has in-part been published as an oral abstract at ASH 2020 in collaboration with Matt Cross, Clinical Research Fellow (see page 160 or https://ash.confex.com/ash/2020/webprogram/Paper141542.html).

The laboratory work was undertaken by staff of the Clinical Genomics laboratory, RMH. The panel used was a collaboration with the EuroClonality NGS consortium who also provided bioinformatic support and assistance.

I am a member of the EuroClonality Consortium and have been involved in the proof-of principle study for the EuroClonality NGS capture panel as well as the validation study.

For this thesis and the abstract published, I have performed the analysis of the NGS data (including the interpretation of SNVs), IG-rearrangement analysis (both from NGS and Sanger sequencing), translocation detection and the preliminary CNV analysis for all samples and the results and interpretation presented in this thesis are my own, original work.

Acknowledgements

I would like to thank Sunil Iyengar for including me as a co-investigator in this study and supporting me through my HSST.

I would like to thank Matt Cross for working with me on this study and sharing his clinical insight.

I would like to thank Nina Dempsey-Hibbert, my academic supervisor, for her guidance.

My thanks go to Grzegorz Pietka for his technical expertise and help utilising the Illumina App for variant identification; to Renata Gomez and Simina Botosneanu for performing the VLeader PCR and Sanger sequencing set-ups.

In addition, I am grateful for access to the bioninformatic tools as part of the EuroClonality team- in particular my thanks go to Nikos Darzentas and Karol Pál for running the samples and providing advice. Similarly, I would like to thank David Gonzalez and Shambhavi Srivastava for giving me access to the CNV component of their analysis pipeline.

I would also specifically like to thank David Taussig, David Gonzalez, Lisa Thompson, Suzanne MacMahon, Mike Hubank and Sunil Iyengar for supporting me in my role as Clinical Scientist within Clinical Genomics at RMH, enabling me to complete my HSST and to achieve Consultant level.

Finally, a big thank you to all my colleagues at RMH for 10 demanding but rewarding years.

Contents

List of Tables

List of Figures

Abbreviations

ALL	Acute Lymphoblastic Leukaemia
AML	Acute Myeloid Leukaemia
amp	Amplification
BAM	Binary Alignment MAP
BCR	B-cell receptor
BL	Burkitt Lymphoma
BM	Bone marrow
bp	Base pairs
c.	cDNA
CBL	Chronic B-lymphocytosis
CD	Cluster Differentiation marker
cDNA	Complementary DNA
CDR	Complementarity-determining region
cen	Centromeric
CHIP	Clonal haemopoesis of indeterminate potential
chr	Chromosome
CLL	Chronic Lymphocytic Lymphoma
CN-LOH	Copy-number neutral loss of heterozygosity
CNV	Copy number variants
COSMIC	Catalogue Of Somatic Mutations In Cancer
CSR	Class-switch recombination
CT	Computertomography
D	Diversity
del	Deletion
der	Derivative
DLBCL	Diffuse Large B-cell Lymphoma
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
EC	EuroClonality
ECNDC	EC-next-generation sequencing DNA capture
EDTA	Ethylenediaminetetraacetic acid
ERIC	European Research Initiative in CLL
FCR	Fluodarabine Cyclophosphamide Rituximab
FISH	Fluorescence-*in situ* hybridisation
FL	Follicular Lymphoma
FR	Framework
fs	Frameshift
FU	Follow-up
GC	Germinal centre
GRCh38	Genome Reference Consortium Human Build 38
HC	High count
HCL	Hairy-cell Leukaemia
IARC	International Agency for Research on Cancer
ICD-O	International Classification of Diseases for. Oncology
ID	Identifier
IGH	Immunoglobulin heavy chain

IGK	Immunoglobulin light chain kappa
IGL	Immunoglobulin light chain lambda
IGV	Integrative Gemonics Viewer
IHC	Immunohistochemistry
IMGT	International ImMunoGeneTics information system®
J	Joining
LC	Low count
LPD	Lymphoproliferative disease
LPL	Lymphoplasmacytic Lymphoma
MALT	Mucosa-associated Lymphoma
MBL	Monoclonal B-cell lymphocytosis
MBR	Major breakpoint region
MCL	Mantle Cell Lymphoma
mcr	Minor breakpoint cluster region
MDS	Myelodysplasia
MGUS	Monoclonal gammopathy of unknown significance
MRD	Measurable-residual disease
MTC	Major translocation cluster
NA	Not available
Neg	Negative
NGS	Next-generation Sequencing
NHL	Non-Hodgkin Lymphoma
NICE	The National Institute for Health and Care Excellence
NMZL	Nodal Marginal Zone Lymphoma
NOS	Not otherwise specific
OS	Overall survival
p	Short arm of chromosome
p.	Protein
PB	Peripheral blood
PCR	Polymerase-chain reaction
PCR	Polymerase-chain reaction
PEST	Proline, glutamic acid, serine and threonine
PFS	Progression-free survival
Pos	Positive
q	Long arm of chromosome
QC	Quality control
R	Rituximab
RAG	Recombination-activation genes
RMH	Royal Marsden Hospital
RNA	Ribonucleic acid
ROI	Region of interest
RT-PCR	Reverse transcription polymerase chain reaction
SDRPL	Splenic Diffuse Red Pulp Lymphoma
SHM	Somatic hypermutation
sIg	Surface immunoglobulin
SIHMDS	Specialist Integrated Haematological Malignancy Diagnostic Service
SLLU	Splenic Leukemia/Lymphoma Unclassifiable

SMZL	Splenic Marginal Zone Lymphoma
SNP	Single nucleotide polymorphisms
SNV	Single nucleotide variants
SOP	Standard operating procedure
SV	Structural variants
tAML	Therapy-related AML
tbc	To be confirmed
tel	Telomeric
Ter	Terminating codon
TLR	Toll-like receptor
tMDS	Therapy-related MDS
TR	T-cell receptor
tris	Trisomy
TTT	Time to treatment
UKAS	United Kingdom Accreditation Service
UTR	Untranslated region
V	Variable
VAF	Variant allele frequency
VUS	Variant if unknown significance
WHO	World Health Organisation
WM	Waldenström's Macroglobulinaemia

1. Introduction

1.1 Classification of Lymphoproliferative disorders (LPDs)

The international World Health Organisation (WHO) Classification of Tumours of Haematopoietic and Lymphoid tissue (4th revised edition (Swerdlow et al., 2016a)) recognises more than 50 distinct lymphoproliferative-disorders entities which are determined by morphological, immunophenotypic, genetic and clinical characteristics. The aim is to categorise diseases into distinct entities to develop and apply standardised treatment pathways to ensure optimal, evidence-based patient care. In the last decades advances in the understanding of the cellular processes giving rise to the different lymphomas have been made, leading to the identification of disease-defining chromosomal aberrations and DNA-variants (table 1.1). Our more detailed understanding of the pathology has also enabled huge shifts in treatment protocols and drug designs with options now including molecularly-targeted drugs like BRAF-, BTK- and BCL2-inhibitors as well as cellular therapy with Chimeric antigen receptor T (CAR-T)-cells for certain diseases.

Patients with suspected or confirmed lymphoproliferative disease (LPD) are assessed and cared for in specialist units within the NHS to ensure highest expertise, access to dedicated testing and integrated reporting as set out in the National Institute of Health and Care Excellence (NICE) guideline 'Haematological cancers: improving outcomes' (NG47, published 25th of May 2016 https://www.nice.org.uk/guidance/ng47, accessed 08/05/2021)).

Despite these advances, cases remain which do not fit into any of the defined categories or where conflicting features mean several differential diagnoses cannot be further delineated.

Table 1.1: Examples of haematological neoplasms where the molecular aberration is a defining feature. Additional morphological and immunophenotype characteristics used for diagnosing these diseases are briefly summarised.

World health organisation (WHO) classification (Swerdlow *et al.* 2008 and 2016)	**Genetic diagnostic markers and their frequency in a given disease category** (from Swerdlow *et al.* 2008 and 2016)	**Diagnostic markers: Morphology** (from Swerdlow et al. 2008 and 2016)	**Diagnostic markers: Immunophenotype** (from Swerdlow et al. 2008 and 2016)
Burkitt lymphoma (BL)	t(8;14) translocation involving *MYC* in >95% of cases; Additional recurrent variants (~70% of cases) in *TCF3* and in *ID3*	Medium-sized cells, cytoplasm is deeply basophilic with lipid vacuoles Very high proliferation rate and mitotic figures "starry sky pattern"	Positive for B-cell antigens (CD19, CD20, CD22, CD79a and PAX5); positive for CD10 negative for CD5 Strong expression of MYC and Ki-67
Mantle cell lymphoma (MCL)	t(11;14) translocation involving *CCND1* seen in 65-95% (also seen in myeloma, diffuse large B cell lymphoma (DLBCL)) Deletions of regions including *CDKN2A, TNFAIP3, ATM* and TP53 as well as gains of *BCL2* and *CCND1*; Mutations in *CCND1, NOTCH1/2* and *TP53*	Monomorphic small- medium sized cells with a number of morphological variants: blastoid, pleomorphic, small-cell and marginal zone-like (refer to Swerdlow *et al.* 2008 and 2016 for further details)	Strongly positive for surface Ig; Positive for CD5, FMC7 and CD43 Negative for CD10, CD23 and BCL6 In particular, positive for SOX11 and CCND1 expression
Follicular lymphoma (FL)	t(14;18) translocation involving *BCL2* seen in 70-95% (also seen in DLBCL) Mutations and /or copy number aberrations in *BCL2, KMT2D, EZH2, CREBBP, BCL5, MEF2B* and *TNFAIP3*	Characterised by follicular pattern which effaces the nodal architecture with poor definition; neoplastic cells are a mix of centrocytes and centroblasts	Positive for surface Ig and B-cell antigens Positive for CD10, BCL2 and BCL6 Negative for CD5 and CD43 Overexpression of BCL2 is the hallmark
Mucosa-associated lymphoid tissue (MALT) lymphoma	t(11;18) translocation involving *API2-MLT* seen in 20-50%	Marginal zone B cells, small-medium in size, irregular nuclei, pale cytoplasm; may show plasmacytic differentiation	Positive for CD20, CD79a and CD43+/- Negative for CD5, CD10 and CD23
Hairy cell leukaemia (HCL)	*BRAF* V600E seen in >98% of cases	Small-medium sized cells often with kidney-shaped nucleus and pale blue cytoplasm (abundant) which shows characteristic projections	Bright positivity for surface Ig, CD20, CD22 and CD11c; Positive for CD103, CD25, CD123, CD200 and CCND1 Negative for CD5 and CD10 Positivity for Annexin A1 is a hallmark
Lymphoplasmacytic lymphoma (LPL)/Waldenstroem's macroglobulinaemia (WM)	*MYD88* L265P seen in >90% Mutations in *CXCR4* in 30% and *ARID1A* in 17%	Small lymphocytes with plasma cell component	Positive for B-cell antigens (CD19, CD20, CD22, CD79a) and surface Ig Negative for CD5, CD10, CD103 and CD23 Plasma cell component CD138+, CD19+ and CD45+

In 2016, a review was performed by a clinical fellow of the Specialist Integrated Haematological Malignancy Diagnostic Service (SIHMDS) at the Royal Marsden Hospital (RMH) which showed that a considerable number of indolent B –LPDs were unclassifiable and were referred to as either CD5+ve B-LPD, B-LPD not-otherwise specified (NOS) and Lymphoplasmacytic Lymphoma/Marginal Zone Lymphoma (LPL/MZL)-type LPD (figure 1.1) when evaluated according to best-practice guidelines. It was acknowledged that additional off-pathway testing including further cytogenetic and immunohistochemistry (IHC) analysis could lead to a defined diagnosis but it was estimated that at least half of the cases would not receive a diagnosis based on the WHO-classification and thus would have no clear, evidence-based treatment pathways and no option to enter a clinical trial.

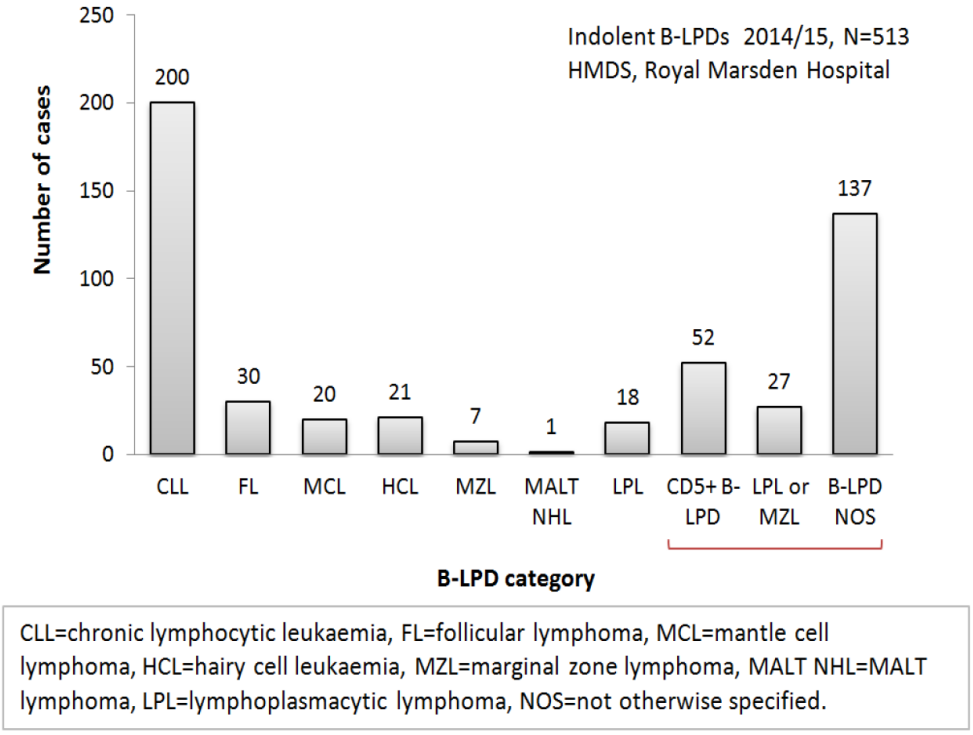


Figure 1.1: Diagnoses assigned to 513 cases of indolent lymphoma assessed according to NICE-guidelines at SIHMDS RMH. The three groups of uncertain classification are highlighted by the red line and represent 42% of cases.

The latest edition of the WHO-classification has, for the first time, recognised and incorporated further precursor scenarios for indolent B-LPD similar to the well-known monoclonal gammopathy of undetermined significance (MGUS) and monoclonal B-lymphocytosis (MBL) which can give rise to plasma cell myeloma and chronic lymphocytic leukaemia (CLL), respectively (table 1.3). Some of the

undiagnosed cases within the RMH cohort may fall into these novel categories as discussed further in chapter 3 and 4.

The “Enhancing diagnosis in chronic B-cell lymphoproliferative disorders using Next-Generation Sequencing” (ENABLE) study, NCT03344809, prospectively collected samples on patients presenting to the SIHMDS who could not be assigned to an established WHO-disease category and instead fitted one of the three groups highlighted in figure 1.1. The study aim is to apply next-generation sequencing to further characterise this group of patients.

The next chapters will briefly explore the differential diagnoses the cases in the three groups identified could fall into, focusing on the relevance of CD5-expression and plasmacytoid features, together with an outline of the cellular pathways that are expected to be affected by molecular aberrations in these disease entities.

1.2 Differential diagnoses for cases with uncertain classifications

B-NHLs arise at different stages during B-cell development (figure 1.2) usually once they have moved to secondary lymphoid tissues. Upon antigen stimulation, germinal centres (GC) are formed within lymph nodes. Here, B-cells undergo somatic hypermutation (SHM) and class-switch recombination (CSR) together with affinity selection and proliferation cycles which generate B-cells producing highly-specific antibodies with autoreactive B-cells being eliminated in the process. Based on phenotype, molecular markers and expression profiling findings it is thought that Follicular Lymphoma (FL) and some Diffuse-large B cell lymphoma (DLBCL) are derived from B-cells undergoing the GC-reaction whereas classic Hodgkin lymphoma (cHL), activated-B-cell-type DLBCL and mediastinal lymphomas develop from B-cells that have completed the GC-reaction with memory B-cells being the cell of origin for plasma cell myeloma B-cells. For Chronic lymphocytic leukaemia (CLL), one group of patients carry SHM in their CLL-resembling GC-B-cell whereas other CLL clones resemble naïve B-cells without evidence of SHM. B-cells that reside in the mantle zone or marginal zone of a lymph node can give rise to mantle cell lymphomas (MCL) or marginal zone lymphomas (MZL), respectively.

As shown in table 1.1 a number of B-NHLs are defined by their specific chromosomal rearrangements like for example *IGH-BCL2* in FL and *CCND1-IGH* in

MCL. In most cases, the translocation moves a strong enhancer of the IG-locus into proximity of a proto-oncogene resulting in upregulation of the expression of the proto-oncogene and driving proliferative and anti-apoptotic signalling processes in the neoplastic cells. In rare cases, mutation of one gene is the hallmark and primary driver of a disease (e.g. *BRAF* in Hairy Cell Leukaemia) whereas most disease subtypes are heterogeneous with regards to their molecular variants and overlap between disease entities is common.

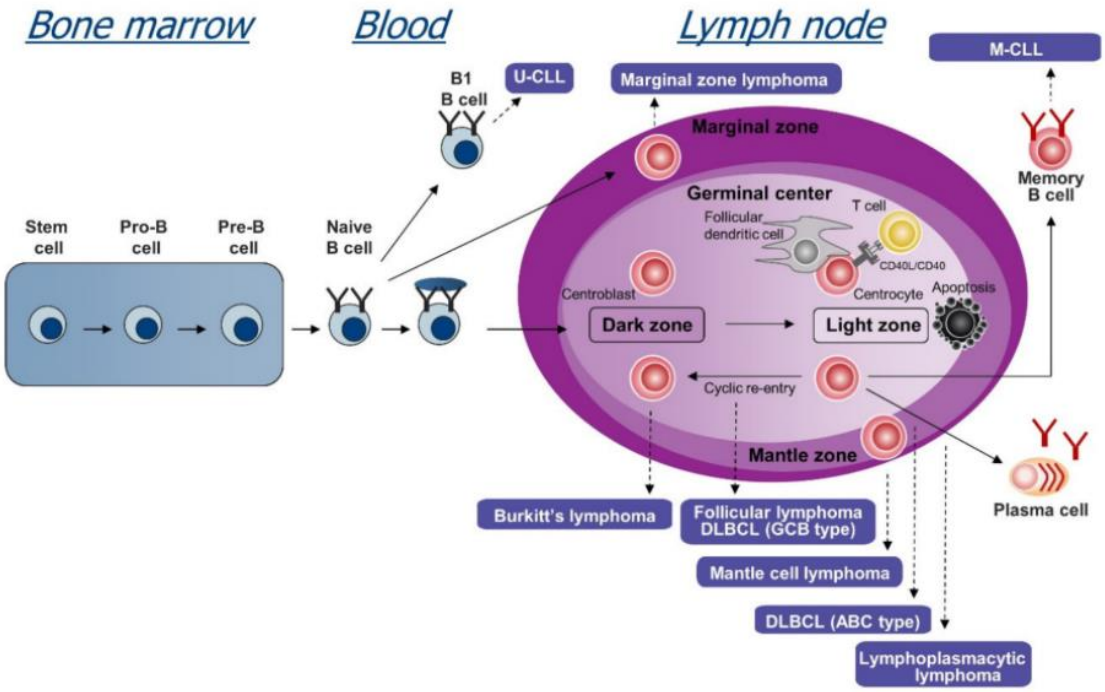


Figure 1.2: Cell of origin for different B-NHLs (from (Efremov et al., 2020) figure 2 page 4). Schematic representation of the steps of B-cell development and maturation in the different environments (bone marrow, peripheral blood and lymph nodes) and the postulated normal counterpart of the neoplastic cells of different B-NHLs.
U-CLL= unmutated CLL, M-CLL= mutated CLL, DLBCL= Diffuse-large B-cell Lymphoma

1.2.1 CD5+ve B-lymphoproliferative disease

Aberrant CD5+ positivity is most characteristic for the B-cell clone of CLL and MCL and these diagnoses are established based on morphology and a panel of immunophenotyping markers including those to establish the CLL score (table 1.2 (Köhnke et al., 2017)). If cells show CD5-positivity but a CLL-score of 3 or less, fluorescence-*in situ* hybridisation (FISH) for t(11;14) is undertaken to rule out MCL.

Table 1.2: CLL score based on Immunophenotyping marker panel. A score of 4 and 5 is expected for CLL; CLL-score of 3 could indicate atypical CLL or MCL and further evaluation is required. In addition, CD200 is positive in CLL but not in MCL. sIg = surface immunoglobulin

Marker	Result	Score
CD5	Positive	1
CD23	Positive	1
FMC7	Negative	1
sIg	Weak	1
CD22/CD79B	Weak	1

Even though it is not a WHO-recognised disease entity, atypical CLL is an accepted concept; it is characterised by a CLL score of 3, a higher degree of abnormal lymphocytes in addition to the CLL-clone (but remaining below 55% prolymphocytes), strong CD20 and FMC7 expression and presence of trisomy 12 in up to 20% of cases (Abruzzo et al., 2018). Clinically, some studies have shown association with thrombocytopenia and a higher incidence of Richter's transformation, especially in cases with *IGHV4-39* gene usage. (Autore et al., 2018, Fazi et al., 2011). In 2019, Fang *et al.* published a review of 1684 patients with CLL seen by the Mayo Clinic (Fang et al., 2019). One group of patients with atypical morphology including plasmacytoid appearance, negativity for CD23, strong CD20 and/or surface immunoglobulin expression (and thus a CLL score of no more than 3) was found to have a high incidence of translocations involving BCL3 transcription coactivator (*BCL3)* (Fang et al., 2019). The group recommended that FISH assessment for *BCL3*-rearrangements should be included in routine testing as well as investigations to rule out MCL and marginal-zone lymphoma (MZL) in these cases due to the overlap in morphological features, however, this is has not yet been incorporated into testing guidelines.

Although uncommon, CD5-positivity is observed in rare cases across other small cell/ low-grade B-Non-Hodgkin Lymphomas (B-NHLs) making diagnosing harder and some may be assigned to B-NHL, not otherwise specified (NOS). In CLL and MCL the majority of cases are CD5+ve whereas <5% of case of Lymphoplasmacytic Lymphoma (LPL) and Marginal zone lymphoma (MZL) usually are (Lim et al., 2019).

1.2.2 Low-grade B-LPDs with plasmacytic features

Plasmacytic differentiation within the neoplastic B-cells can be seen across a number of B-cell lymphomas ranging from Lymphoplasmacytic Lymphoma (LPL), where the majority of the clone are plasma cells, to Mantle Cell Lymphoma (MCL),

where plasmacytic features are infrequent (Swerdlow et al., 2016b). Given the overlapping morphological features, differentiating between Splenic Marginal Zone Lymphoma (SMZL), Lymphoplasmacytic Lymphoma (LPL), nodal Marginal zone lymphoma (NMZL) and other small-cell B-LPDs with plasmacytic features can be difficult (Hamadeh et al., 2015). One example is the rare splenic B-cell lymphoma, unclassifiable (SLLU) which has overlapping features with and requires the exclusion of CLL, Hairy-Cell Leukaemia (HCL), B-cell Prolymphocytic Lymphoma (B-PLL), Splenic Marginal Zone Lymphoma (SMZL) and Lymphoplasmacytic Lymphoma (LPL). Similarly, rare entities like HCL-variant (HCL-v), and Splenic Diffuse Red Pulp small B-cell lymphoma (SDRPL) are difficult to distinguish (Curiel-Olmo et al., 2017). With the exception of CLL they are generally CD5-negative but rare cases with CD5-positivity have been seen across all these categories.

For most of the marginal zone lymphomas – in particular SMZL and SDRPL- a definite diagnosis still requires splenic histology; however, this is rarely performed in routine practice and the diagnosis relies on clinical features and morphological features although there are none that are disease defining (Jaramillo Oquendo et al., 2019). The 2017 workshop of the Society for Hematopathology/European Association for Haematopathology reviewed the role of molecular genetics in the diagnosis of lymphoid neoplasms (Lim et al., 2019). Clinical cases submitted to the panel for discussion highlighted the need for integration of morphology, immunophenotyping and molecular findings to derive the most likely diagnosis but overlapping features and molecular aberrations affecting the same cellular pathways were common. The distinction between SDRPL and Hairy Cell Leukaemia-Variant (HCL-v) was stated to be particularly challenging with *CCND3*-variants favouring the former and presence of *MAP2K1*-variants the latter. The presentation also included cases that were found to be *MYD88*-variant positive but matching a diagnosis of NMZL and Ocular adnexal marginal zone lymphoma rather than LPL. In 2011, Ngo *et al.* described a pathogenic missense variant in the myeloid differentiation primary response 88 (*MYD88*) gene in a cohort of Diffuse Large B-cell Lymphoma – activated B-cell type (DLBCL-ABC) cases (Ngo et al., 2011). The nucleotide change from thymine to cytosine causes an amino acid change of leucine to proline at position 265 (NM_002468.4, also referred to as position 273 in NM_001172567.1) causing increased activity and signalling along the Toll-like receptor (TLR) and Nuclear Factor-κB-(NF-κB) pathways leading to proliferation and

pro-survival signalling (figure 1.3). Several studies published shortly after showed that the highest incidence of this variant was in fact in LPL, where it is the main oncogenic driver (Ruben et al., 2019). Despite the high frequency of > 90% in LPL it is not a diagnostic marker as it does occur at low frequency in CLL, MCL, plasma cell myeloma and across SMLZ and other MZL as well as DLBCL (Ruben et al., 2019). The European Association for Haematopathology /Society for Hematopathology slide workshop held in 2014 and reported by Swerdlow *et al.* in 2016, demonstrated the importance of identifying *MYD88* variants to support a diagnosis of LPL (Swerdlow et al., 2016a). In particular, LPL could be diagnosed even if there was extensive nodal effacement, follicular colonisation and the cells expressed CD5 or CD10, all atypical findings for LPL. Such cases often only showed subtle plasmacytic differentiation but presence of a *MYD88* variant was shown in all cases. It was absent on the other hand in cases that were eventually classified as CLL/SLL, Follicular Lymphoma (FL), SOX11-negative MCL and MZL all of which showed plasmacytic differentiation (Swerdlow et al., 2016b). Presence of variants in *MYD88* is strongly associated with presence of IgM-paraprotein and is almost always seen in IgM-MGUS but never in plasma cell myeloma or IgG/IgA-MGUS (Swerdlow et al., 2016a).

Table 1.3: **Brief overview of the different disease categories that are included in the differential diagnosis of the cases being explored in this study including the MBL-entities**. Examples of usual features- see WHO Classification for details.

WHO disease category	**Usual phenotype**	**Morphology**	**Molecular variants**	**Other**
CLL	CD5+ve	Small, homogeneous	Tris12, del11q, 13	CLL score >3, CD23+ve
MCL	CD5+ve	Small, Clefted nucleus	t(11;14)	CLL score </=3, CD23-ve
FL	CD5-ve	Small, folded/cleaved nucleus	t(14;18)	Plasmacytic differentiation is rare
SMZL	CD5-ve	Villous lymphocytes	KLF2, NOTCH2, IGHV1-2*04	
LPL	CD5-ve	Plasmacytoid features	MYD88	Paraprotein
HCL-v	CD5-ve	Villous lymphocytes	MAP2K1, IGHV4-34	
HCL	CD5-ve	Villous lymphocytes	BRAF V600E	Annexin A 1 positive
SDPRL	CD5-ve	Villous lymphocytes	CCND3	Provisional entity
NMZL	CD5-ve	Plasmacytoid features		
Non-CLL-type MBL	CD5-ve	Heterogeneous	Various	Includes CBL-MZ
Atypical CLL-type MBL	CD5+ve	Heterogeneous	Various	Usually CLL score =3

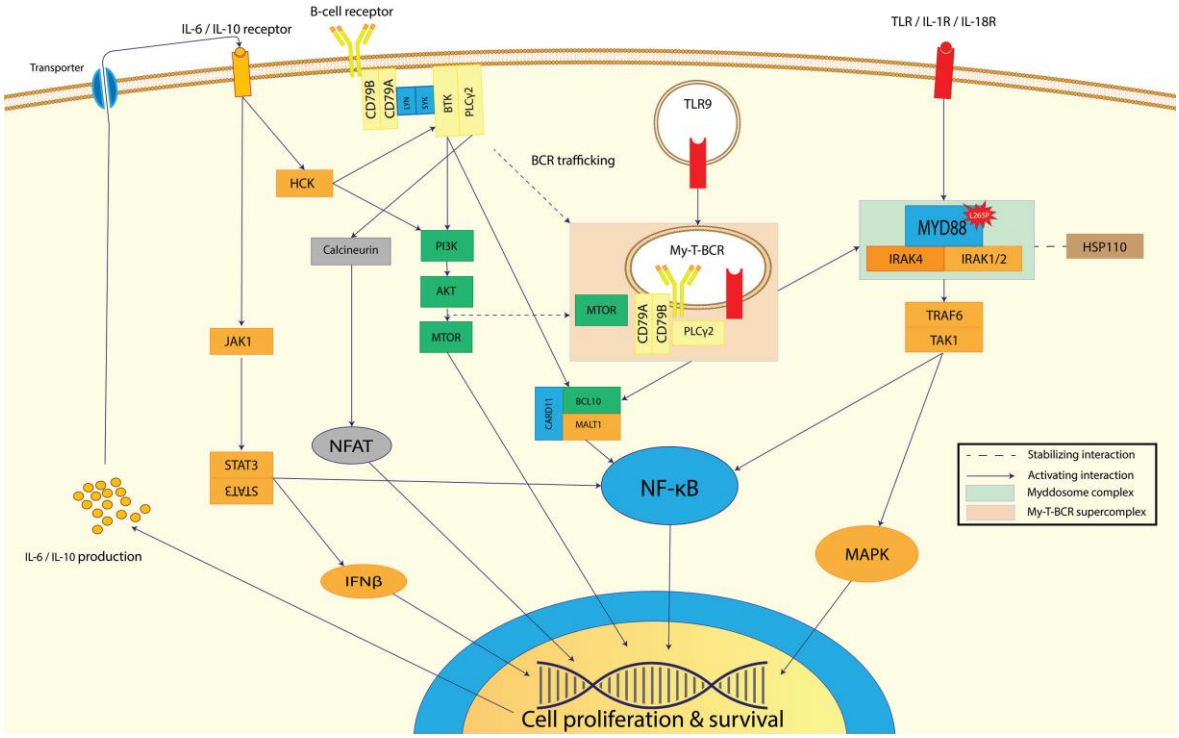


Figure 1.3: **Signalling pathways and their interaction relevant to mature B-cell neoplasms**. B-cell receptor (BCR), interleukin receptors and Toll-like receptors (TLR) are shown on the cell surface together with the components of the intracellular signalling cascades they activate. The NF-kB complex is a central part which can be controlled through most of the pathways with uncontrolled activation leading to increase in cell proliferation and pro-survival signalling. Figure reproduced from (Ruben et al., 2019) page 2338 Figure 1))

1.3 B-cell characteristics: signalling cascades and V(D)J-rearrangements

Cellular pathways like, for example, those controlling the cell-cycle are deregulated across most cancer entities (e.g. loss of or variants in *TP53*) but some are unique to a particular cancer type due to the cell of origin. Most B-NHLs are derived from mature, fully differentiated B-cells requiring pro-survival signalling through their B-cell receptor (BCR) (figure 1.2 and 1.3). The clone is sustained by BCR-triggering due to specific antigens – including autoantigens – and/or by aberrant activation through molecular changes in one of the several signalling cascades (figure 1.3). Figures 1.3 and 1.5 shows examples of genes which are recurrently mutated in the context of B-NHLs and which have been shown to play a pathogenic role in the development and persistence of the clonal B-cell population. These have all been included as targets on the NGS panel used to analyse samples in the ENABLE study (see table A1 in the appendix for the full scope).

Changes and overactivity of these pathways provides pro-survival and pro-proliferative signals which sustain the clone but it has also become evident that the

clones are dependent on these signals (they are 'addicted') and as such blocking of these signalling pathways has been identified as an ideal target for treatment. Figure 1.4 provides a summary of some of the drugs in routine use or in clinical trials for B-NHLs.

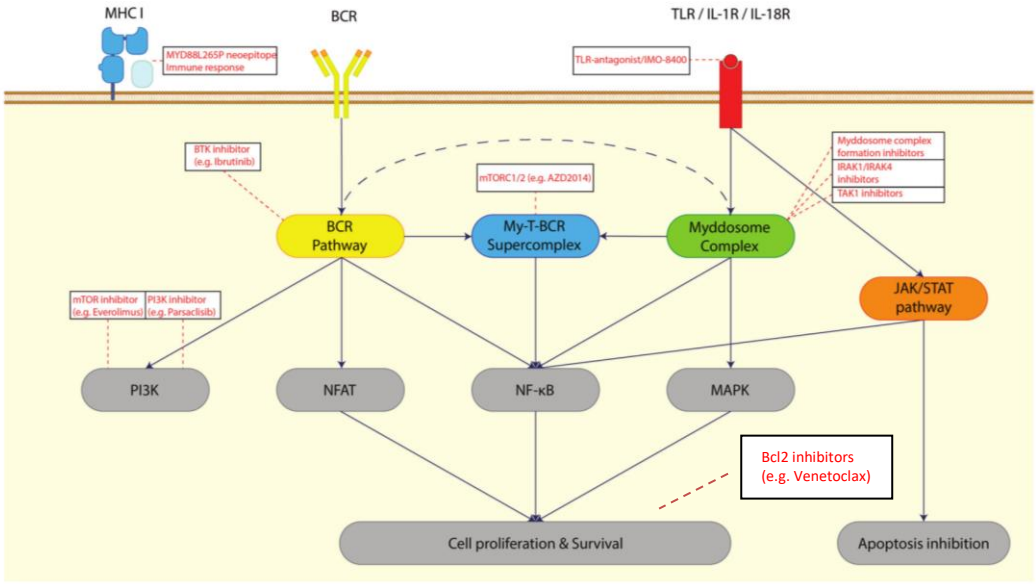


Figure 1.4: **Signalling pathways and their interaction relevant to mature B-cell neoplasms**. Highlighted in red are drugs in use/in clinical trials for the treatment of B-NHLs pointing at the part of the signalling cascades they act upon. Figure reproduced and modified from (Ruben et al., 2019) page 2343, figure 2))

What follows is a brief summary of the key players of the various B-cell specific signalling pathways with a focus on those deregulated by variants found in B-NHLs which could be detected by molecular testing, including next-generation sequencing approaches employed in this study. Good reviews have recently been published on these topics including future treatment options (e.g. (Ruben et al., 2019, Chung, 2020, Casola et al., 2019, Grondona et al., 2018, Sorrentino et al., 2019, Visco et al., 2020)) and the reader is referred to these for more detail with relevant points being picked up in the discussion chapter for cases identified here as part of the ENABLE study.

1.3.1 B-cell receptor (BCR) signalling and NF-κB activation

BCRs are activated through binding of their specific antigen matching their variable regions. The signal is transmitted via CD79A/B with their immunoreceptor tyrosine-

based activation motif (ITAM) domains to spleen tyrosine kinase (SYK) and Lck/Yes-related novel protein tyrosine kinase (LYN) which in turn activate Bruton tyrosine kinase (BTK). BTK triggers phosphoinositide3-kinase (PI3K)/AKT/mammalian target of rapamycin(mTOR) activation as well as signalling to PLCγ2 leading to the recruitment of the caspase recruitment domain family member 11 (CARD11), BCL10, and mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1) and resulting in NF-κB activation (figure 1.3).

When the transcription factors which are part of the NF-κB family are activated, they migrate from the cytoplasm to the nucleus to bind to DNA and initiate transcription of target genes that are involved in B-cell survival and proliferation. Activation of NF-κB signalling can be triggered by several cascades including the canonical pathway through activation of Tumour Necrosis Factor (TNF)-cytokine receptors, via the BCR to Bruton-Tyrosine Kinase (BTK) route as well as the non-canonical pathway via Toll-like receptors (TLR) and the MYD88-containing 'myddosome' complex as shown in figures 1.3 and 1.4 (Diop et al., 2020).

Activity is tightly controlled in healthy B-cells but various aberrations in components of the NF-κB pathway upstream of the transcription factors have been identified as oncogenic drivers leading to constitutive activation of NF-κB signalling without external activation signals through the relevant BCR, TNFR or TLR (figure 1.3). In addition to the pro-survival and pro-proliferative signalling, constitutive activation of NF-κB confers resistance to chemotherapy (Nagel et al., 2014).

Variants in genes coding for several components of the BCR-signalling cascade are found across mature B-cell malignancies including variants in *CD79A/B, CARD11, TCF3* and *ID3* in aggressive lymphomas like DLBCL and Burkitt Lymphoma (BL) and *BRAF* in HCL (figures 1.3 and 1.4). They all lead to constitutive, uncontrolled activation of the respective pathways including NF-κB (Grondona et al., 2018).

In addition to the gain-of function variants, negative regulators of NF-κB activity can be lost through mutations. Examples are TNF alpha-induced protein 3 (*TNFAIP3*) (figure 1.5) and Krüppel-like Factor 2 (*KLF2*, (Clipson et al., 2015)).

Inhibitors that block this aberrant signalling are being developed and some have already entered routine use: The BTK-inhibitor Ibrutinib and the Phosphoinositide 3′ kinase delta (PI3Kδ) inhibitor Idelalisib are licensed for use in CLL, MCL, relapsed MZL and WM for example (Ruben et al., 2019). In addition, proteasome inhibitors like for example Bortezomib and the immune-modulatory drug lenalidomide are able to block overactive NF-κB (Chung, 2020).

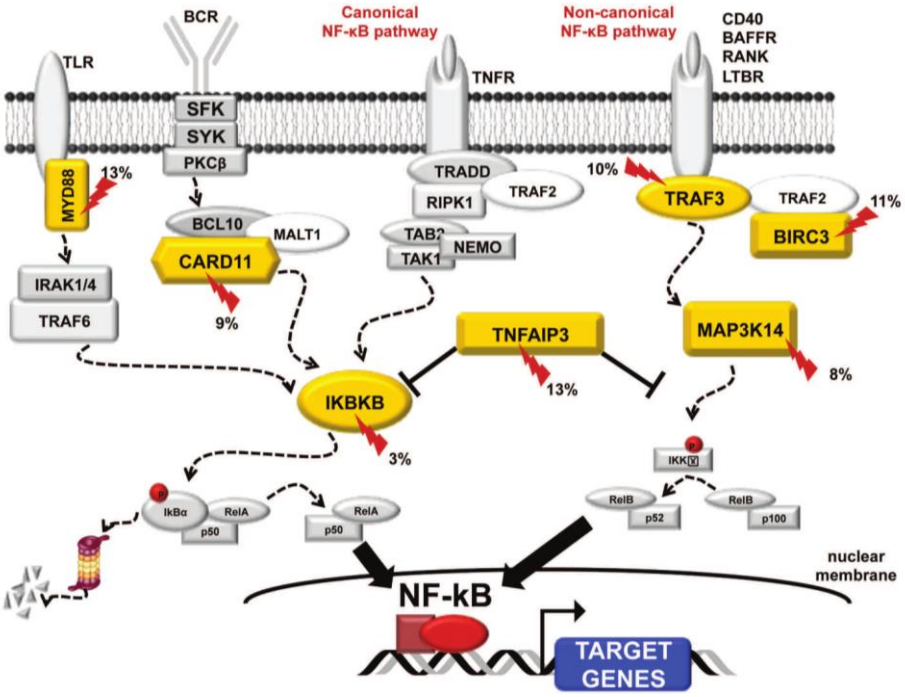


Figure 1.5: **Components and interaction of pathways leading to NF-κB signalling**. Pathogenic variants are found recurrently in various genes (highlighted in yellow). Frequency of variants shown here (red arrows) are for SMZL (Figure reproduced from (Luca and Davide, 2012) page 639 figure 1)

1.3.2 Notch Receptor signalling

Another pathway implicated in B-NHL and best understood in the context of CLL is the NOTCH signalling cascade. Upon activation, the intracellular domain of the NOTCH receptor (which includes the transactivation and proline, glutamate, serine and threonine (PEST) domain) transfers to the nucleus where it initiates gene transcription. In B-cells, both NOTCH1 and NOTCH2 play a role in various developmental and differentiation processes whilst also supporting B-cell survival and proliferation in parallel to the BCR-mechanisms (Arruga et al., 2018).

Variants in NOTCH1 and NOTCH2 in B-NHL lead to truncation of the PEST domain and are most-commonly encountered in CLL, FL and MCL (NOTCH1) and SMZL, (NOTCH2). The loss of the PEST domain removes areas required for phosphorylation by CDK8 and FBXW7 to initiate protein degradation and results in retention of the protein and a longer half-life of activated NOTCH. In contrast to the aberrations discussed for BCR and NF-KB signalling leading to activation independent of ligand-binding, mutant NOTCH still requires activation by its ligand (Arruga et al., 2018). *NOTCH1* PEST-domain variants are associated with poor prognosis in CLL and MCL and negate the benefit of adding rituximab to fludarabine-chlorambucil therapy as demonstrated in the German CLL8 trial (Stilgenbauer et al., 2014).

1.3.3 V(D)J-rearrangements

A unique hallmark of B- and T-lymphocytes is the modification of genomic DNA during the process of V(D)J-recombination (figure 1.6). Variable (V) genes, diversity (D) genes and joining (J) genes are selected at random to form a V(D)J-rearrangement with the interspersing DNA being removed leading to an altered 'germline' DNA sequence. This sequence encodes the variable region of an antibody which defines the specificity of an antibody towards its antigen. As a result of this process each B-cell produces a unique B-cell receptor (BCR)/antibody and no two B-cells carry the identical V(D)J rearrangement, unless they are clonally related. Analysis of these V(D)J-rearrangements at the genomic level can help to distinguish between reactive (polyclonal) and clonal (potentially neoplastic) B-cell or T-cell populations, supporting diagnostic evaluations for possible lymphomas in routine diagnostic practice (Au - Agathangelidis et al., 2018, Langerak et al., 2012).

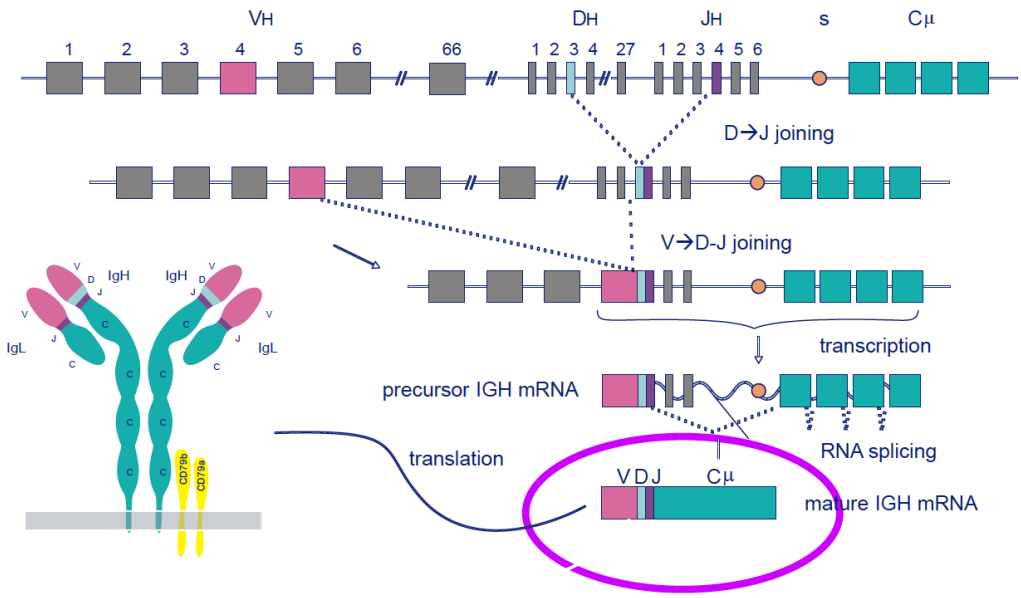


Figure 1.6: Diagram showing the processes of V-D-J-gene joining and generation of a unique V(D)J rearrangement during B-cell maturation. Deletion and insertion events take place at the joints and once a B-cell undergoes the germinal centre reaction in the lymph node following antigen exposure, somatic hypermutations are introduced to further increase the specificity of the antibody. Antibodies are made up of two heavy and two light chains and the variable region (shown in pink bottom left) is encoded by the V(D)Jrearranged genes. The constant region determines the type of immunoglobulin (IgM, IgA etc).

The development of lymphomas is linked to B-cell activation via antigens and shares common features with autoimmune diseases (Singh et al., 2020). In contrast to the huge variety achieved by V(D)J-rearrangements in a healthy immune repertoire, some lymphomas show a skewed usage of certain IGHV-genes: IGHV1-69, IGHV3-30 and IGHV 4-34 are for example highly prevalent in CLL (Sutton et al., 2009), whereas IGHV1-2 is closely linked to SMZL (Bikos et al., 2012). It is therefore likely that a common antigen triggers the selection of these clones, promotes their survival and provides the basis for cancer development. Furthermore, in addition to the same IGHV-gene, these V(D)J-rearrangements show homology (so called stereotypes) in their complementary-determining regions (CDRs), the very area that is finetuned by deletion and insertion events and addition of somatic hypermutations to be as unique and specific towards an antigen as possible. Comparison of the homology between B-cell clones of patients with B-NHL, most extensively studied in CLL, has led to the identification of different stereoytpes of V(D)J-rearrangements, which vary in frequency between different B-NHL subtypes (Zibellini et al., 2010, Rossi and Gaidano, 2010). Importantly, different stereotypes are associated with different prognosis in CLL: for example an IGHV3-21 rearrangement assigned to stereotype

subset #2 carries a particular poor prognosis regardless of somatic hypermutation status compared to a very indolent disease course in patients with IGHV4-34 stereotype subset #4 (see https://station4.arrest.tools/subsets/ for subsets identified in CLL (Sutton et al., 2016, Rossi and Gaidano, 2010)).

Detailed analysis of the V(D)J-rearrangement the B-cell clone of a patient carries can therefore provide evidence of clonality, may support a particular diagnosis, defines prognosis in some cases and sheds light onto the aetiology of the disease, enabling future research into the development of lymphomas.

1.4 Evaluation of lymphoid neoplasms by Next-generation sequencing

Over the last few years, the interest in utilising NGS for the evaluation of lymphoid neoplasms has increased and several studies have been published evaluating the benefit of running NGS-panels in the context of B-NHL with most of them focusing on the detection of single nucleotide variants with known clinical relevance. Sujobert *et al.* on behalf of the French LYSA Lymphoma study association) and GBMHM (Groupe de Biologistes Moléculaires des Hémopathies Malignes) published a proposal for a consensus NGS-panel to evaluate B-and T-cell malignancies by NGS (Sujobert et al., 2018) and the European Research Initiative in CLL (ERIC) consortium compared the performance of different NGS-panels available (Lesley-Ann et al., 2020). Like Guillermin *et al.,* they all saw a clear benefit through the increase in the number of genes that could be analysed for each patient whilst also highlighting the optimisation required and the difficulty that could arise in the interpretation of the findings: firstly, in the interpretation of the individual variants and secondly, in the integration of the molecular variants with other pathology and clinical findings (Sujobert et al., 2018, Guillermin et al., 2018).

Using a 31-gene panel on 147 cases with B- and T-LPD seen in their routine clinic, Jajosky *et al*. identified clinically-relevant variants in more than 60% of cases (Jajosky et al., 2021). The greatest benefit lay in identifying variants helpful in the diagnosis of 16 cases, which were difficult to assign based on current standard of care testing. Nine cases were CD5+positive B-NHL, which based on the presence of the hotspot *MYD88* variant were assigned a diagnosis of LPL or if NF-κB-related variants were found, a diagnosis of MZL was given. Similarly, this mutation pattern

helped to distinguish between MZL and LPL in cases with plasmacytic features. In a similar approach with a 46-gene NGS panel, Bommier *et al.* further characterised 229 cases including 89 small-cell B-LPDs (Bommier et al., 2021). There was high concordance (89.5%) with the histological diagnosis derived by a regional expert panel but in 24 cases the diagnosis changed which had a therapeutic impact in 11 cases. Amongst these were low-grade B-NHL where the diagnosis changed from MZL to LPL and MZL to FL or *vice versa.*

The use of capture-based NGS approaches for the simultaneous detection of SNVs, CNVs, translocations and IG/TR rearrangements in the context of haematological malignancies formed the topic of my Higher Specialist Scientific Training (HSST) C1 innovation project (Wren, 2017). A copy of the write-up is provided in appendix A9. The work included a detailed literature search to establish the current applications and performance characteristics of capture NGS panels (Part A) as well as the review of explorative studies to understand their potential for clinical use in future (Part B). Advantages and disadvantages of the capture-NGS approach for each of the targets of interests (SNVs (HSST C1 chapter 5), CNVs (chapter 4), IG/TR rearrangements and translocations (chapters 6 and 3, respectively)) are presented in Part A Part B discusses the benefits of this approach for the laboratory looking at the advantages in extending the scope of analysis, reduction in costs, streamlining of workflows and promoting automation.

The evidence found through the C1 work clearly indicated that capture-based NGS is an efficient tool to enable the analysis of all aberrations. The main advantages identified were the consolidation of different testing strategies into one assay reducing time and costs and the ability to identify unusual and novel fusion partners and rearrangements not covered by current routine methods. The validation study undertaken by the EuroClonality NGS consortium (Stewart et al., 2021) as well as other studies published recently by Yasuda *et al*. (pan-haematological malignancies), McConnell *et al*. (sarcoma) and Cuenca *et al.* (plasma cell myeloma) for example demonstrate that the technology fulfils the quality and performance criteria required for routine diagnostic use with high concordance rates compared to current standard-of care (SOC) tests and high specificity, sensitivity and reproducibility (table 4.1) (Yasuda et al., 2020, Cuenca et al., 2020, McConnell et al., 2020, Stewart et al., 2019, Stewart et al., 2021).

1.4.1 Principle of capture-based NGS

The capture-based NGS process is shown in figure 1.7 and further detailed in chapter 2.5. The main difference to amplicon or primer-extension approaches for library preparation is that the tiled bait design ensures high and even coverage whilst reducing the number of PCR cycles needed thus reducing artefacts and errors that can be introduced in the process. The EC-NDC panel is designed for the simultaneous detection of IG/TR-rearrangements to assess clonality, SNVs, CNVs and translocations (full scope is detailed in appendix table A1). For the analysis of lymphomas in particular, the ability of translocation detection and rearrangement detection is vital and is explained in figure 1.8 and in more detail within the examples identified in the results section (chapter 3.6). In contrast to standard-of care (SOC) Fluorescent *in situ* hybridisation (FISH) testing, one assay can identify translocations and rearrangements arising from any of the IG-loci whilst also identifying all fusion partners including unusual and novel ones.

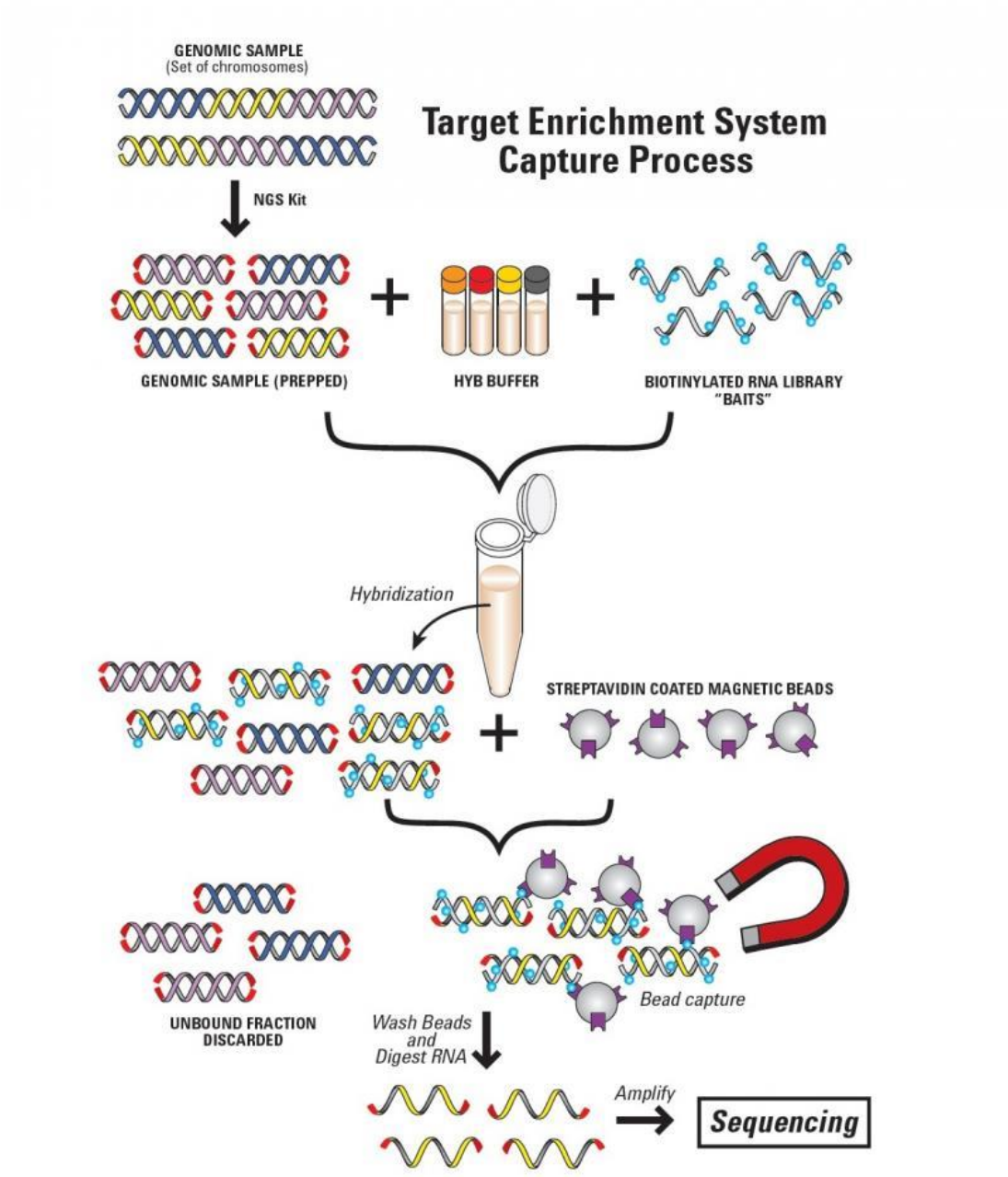


Figure 1.7: Depiction of the capture-NGS workflow using liquid bead capture Genomic DNA is fragmented and NGS platform adaptors, barcodes for patient identification and PCR/Sequencing primer binding sites added (Library preparation). Biotinylated baits (also referred to as probes) designed to be complementary to the regions of interest are added and allowed to hybridise to the libraries. Following hybridisation, streptavidin-coated, magnetic beads allow the selection of DNA-Bait dimers which should represent the regions of interest and undergo NGS sequencing (figure and text reproduced from C1 HSST project page 3).

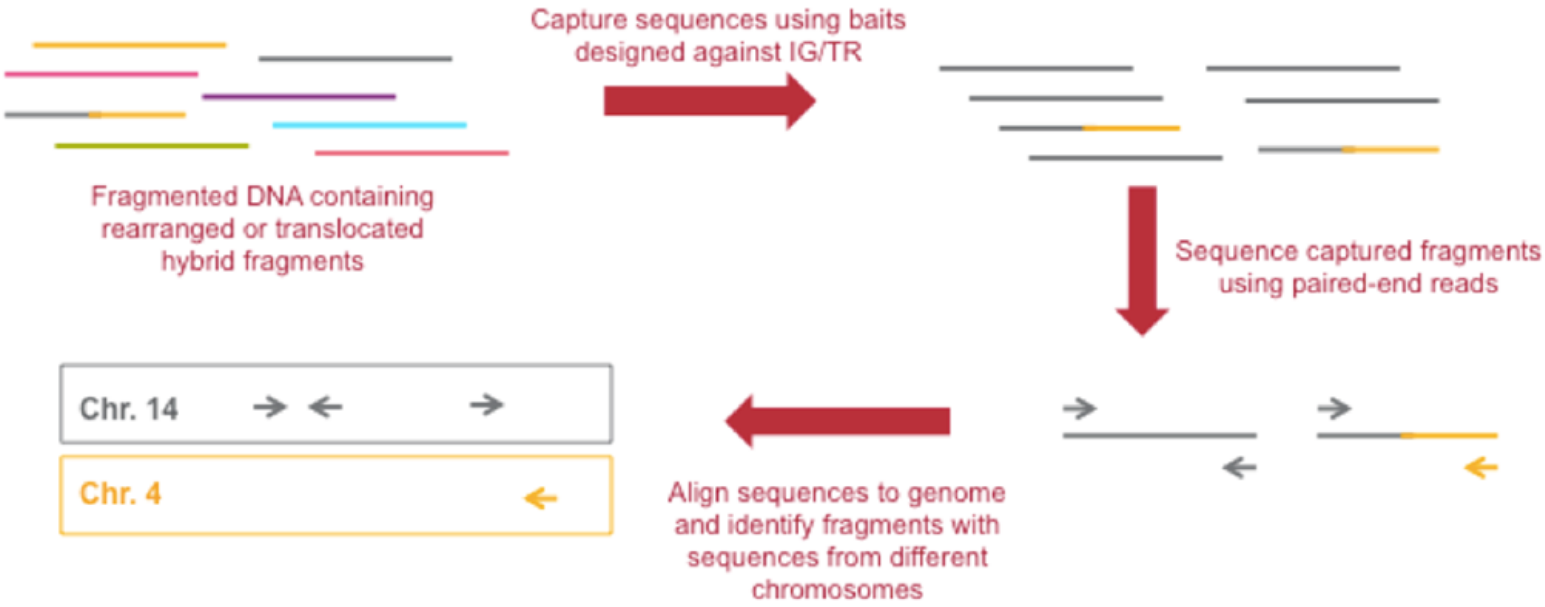


Figure 1.8: The Euroclonality Capture-based NGS panel is capable of detecting all translocations arising from IG- and TR-loci regardless of partner. In this example, baits would have been designed against the IGH-locus on chromosome 14 (in grey) and only DNA fragments containing 'grey' DNA sequence will hybridise to the baits and be sequenced. If a translocation is present, some of the DNA fragments generated in the first step will contain DNA sequences from the IGH-locus as well as the partner chromosomes, in this case chromosome 4 (yellow). The 'yellow' part of the fragment will be subjected to sequencing even though no baits were designed for its detection (figure reproduced and text modified from C1 HSST project page 3).

1.5 Aim of the ENABLE study

The overall aim of the ENABLE study is to enhance the diagnosis of unclassifiable, non-CLL B-LPDs using next-generation sequencing technology. By using a comprehensive NGS-panel designed for the detection of clonality, single-nucleotide variants, copy-number variants and chromosomal rearrangements events this subgroup of patients without a definite diagnosis can be characterised in greater detail. In addition, the study will evaluate the potential benefit of using this novel technology in routine practice to enable comprehensive evaluation of the molecular changes in a cost-effective and timely way. Given the known complexity of molecular aberrations in B-cell lymphoproliferative disease there is a need to be able to analyse patterns and combinations rather than single variants for which the EuroClonality capture-NGS panel is an ideal tool.

The study aims to identify molecular changes that may assign a disease category to patients, may open-up access to clinical trials for patients or identify possible treatment targets to initiate further research and clinical studies.

The long-term follow-up data will help us understand the rate of progression and need for treatment in this cohort of patients expected to show heterogeneous clinical behaviour. Results from this large, prospective study will aid advance our insight into defining new subgroups and precursor lesions and enable the incorporation of molecular findings into diagnosis and prognostic evaluation of patients with unclassifiable, non-CLL B-LPDs.

This thesis presents the findings of the NGS-panel analysis from 108 out of the 120 patients recruited into the ENABLE study between 2017 and 2021. It includes a discussion on the potential clinical utility of the findings but due to the lack of clinical information, the discussion of specific patient cases is limited and the power of the study to assign a definite diagnosis or disease subgroup may thus be underestimated.

2. Materials and Methods

2.1 Study design and cohort

The study was based at The Royal Marsden Hospital (RMH) NHS Foundation Trust and led by Chief Investigator Dr Sunil Iyengar (Consultant Haematologist, RMH). Ethics approval was granted by the London-Camberwell St Giles Research Ethics Committee (REC), reference number 16/LO/0375. Informed consent was obtained from all patients participating as stipulated in the Human Tissue Act 2004 and the Code of Practice for consent (Human Tissue Authority). Samples were assigned and labelled with trial ID based on their trial centre e.g. RMH, followed by the CCR study number 4383 and a case specific number (example RMH-4383-0001). Patients could be enrolled and consented into the study from any of the authorized sites which were already part of the routine RMH Specialist Integrated Haematological Malignancy Diagnostic Service (SIHMDS).

2.1.1 Inclusion criteria

Patients with a morphologically low-grade, CD19+ve/CD20+ve clonal B-cell neoplasm with a CLL score of three or below and a Hairy-Cell leukaemia (HCL) score below 4 were eligible for study entry. This includes patients with low-grade CD5+ve B-NHL, B-NHL not otherwise specified (NOS) and low-grade B-NHL with plasmacytic differentiation, all of which are categories that do not fit into any of the defined WHO categories based on current standard-of care testing results. In the most recent WHO classification update, three different types of monoclonal B-cell lymphocytosis (MBL) have been included: CLL-type, atypical CLL-type and non-CLL-type. They are differentiated based on their different phenotypes (see chapter 1.2) and atypical CLL-type and non-CLL-type MBL would be eligible for this study. Up-front investigations were performed as routine patient assessments as part of their evaluation of lymphocytosis according to the SIHMDS pathways in place at the time.

In order to be included in the study, samples should have had a minimum of 20% infiltration with clonal B-cells, however, this was not reviewed upon receipt at the RMH SIHMDS and results show that a small number of samples with lower infiltration may have been received (refer to 3.1 for further details).

2.1.2. Patient cohort and samples

A total of 108 patients were enrolled in the trial. Peripheral blood (4-10 mL) or bone marrow samples (2 mL), anti-coagulated with EDTA (with at least 20% neoplastic infiltration) were collected. A germline control sample was obtained by CD15+ve cell selection from the blood sample. A repeat sample had to be requested in case the sample was >48 hours old upon receipt to ensure optimal cell selection or if the extraction failed to obtain the required amount of DNA.

Out of the 108 samples included here, only three samples were bone marrow, the rest were peripheral blood samples.

As the study had not completed recruitment at the time of writing, no access to the clinical data and follow-up data for the patients was yet available. At study entry, samples were anonymised and given a trial number before being received in the laboratory. Only the study's principal investigator and clinical fellow working on the study were allowed to review the clinical data and provided me with the information used in this thesis, including the case studies (e.g. chapter 4.4.3, 4.4.6, table 3.6 and the limitations in this write-up arising due to this are discussed in chapter 4.6).

2.2 Cell separation

Magnetic cell separation for CD15+ve cells was performed using the StemCell EasySep™ Whole blood positive selection kit (Catalogue #17751, StemCell™ Technologies, Vancouver, Canada).

To 4 mL of peripheral blood or 2 mL bone marrow, the equal amount of 1X EasySep™ RBC Lysis Buffer was added, followed by 200 µL of EasySep™ Release Human CD15 Positive Selection Cocktail. The tube was inverted 6 times to mix and incubated at room temperature for 3 minutes. Next, 200 µL of EasySep™ Releasable RapidSpheres™ 50201 were added and tube contents mixed by inversion. Following a 3-minute incubation at room temperature, the tube was transferred into the EasySep™ Magnet and left to stand for 5 minutes. The supernatant was transferred into an appropriately labelled 15 mL Falcon tube and 500 µL aliquotted into a 1.5 mL screw cap tube for downstream DNA extraction (CD15+ve depleted blood was used as the tumour sample). To the tube in the

magnet, 5 mL EasySep™ Separation Buffer was added, the tube was removed from the magnet and the contents mixed by pipetting up and down five times and the tube returned to the magnet. Following a 5-minutes incubation the supernatant was discarded and 5 mL of Separation Buffer added, and the process repeated twice more. Following the third supernatant disposal the tube was removed from the magnet and the cells were resuspended in 400 µL of Separation buffer and transferred into an appropriately labelled 1.5 mL screw cap tube (CD15+ve positive cells to be used as germline control).

The purity of the CD15+ve positive cell fraction was checked by Biomedical scientists of the immunophenotyping laboratory RMH and also assessed for the presence of CD19+ve and CD20+ve cells by flow cytometry. Purity had to be >90% with less than 5% of CD19+ve or CD20+ve cells to minimise the risk of contamination of tumour in the germline sample.

2.3 Extraction of DNA

Cell counts were performed using the HemoCue® WBC system (ID 123001, HemoCue AB) to avoid overloading the extraction system. The HemoCure cell counter stains the white cells, which are counted by image analysis. A volume of 10µL of neat sample is required.

DNA was extracted from the cell fractions using the QIAsymphony DNA Midi kit (#937255, Qiagen) on the automated QIAsymphony SP instrument (ID 9001297, Qiagen) running the QIAsymphony DNA-Blood_400_V6_DSP protocol (version December 2017). The system uses magnetic-particle technology for nucleic acid extraction and purification. In brief, cells in the sample are lysed to release nucleic acids that bind to the surface of magnetic particles. These are washed to remove contaminants and eluted. Input volume was 400 µL containing a maximum of 13x 10^9cells/ L. Elution volumes was 100 µL.

Samples were quantified using the Qubit fluorometer (ID Q33216, Invitrogen™ Qubit™ 3 Fluorometer, Thermo Fisher Scientific) using the dsDNA broad range kit (#Q33265, Invitrogen™ Qubit™, Thermo Fisher Scientific)

Samples were stored at -20°C for short term and -80°C for long-term storage.

2.4 IGHV mutational status analysis

Analysis of the IGHV rearrangement including somatic hypermutation status was performed according to the diagnostic service standard-operating procedure (SOP) used routinely at the clinical Genomics laboratory at RMH, following the guidance published by European Research Initiative in CLL (ERIC, (Ghia et al., 2007, Au - Agathangelidis et al., 2018)). For each sample a PCR product is generated using the multiplex PCR primer combination and parameters as shown in table 2.1.Each VLeader primer enables the amplification for the respective V-gene family and input is 100ng of DNA. In cases where the first set-up showed two overlapping sequences indicating the presence of two V(D)J rearrangements on the two alleles both utilizing the same V-gene, RNA was extracted from stored mononuclear cells suspensions using the RNAeasy kit (#74104, Qiagen) and cDNA made using the Super Script™ II protocol (#18064022, ThermoFisher). 4 µL of neat cDNA was then subjected to the same PCR as outlined for genomic DNA. Agencourt AMPure XP bead clean-up of the PCR product was performed following the manufacturer's instructions (#A63881, Beckham Coulter).

Table 2.1: Primer sequences for VLeader multiplex PCR mix (top), PCR master mix (bottom left) and PCR cycling conditions (bottom right)

Primer Mix	Primer Stock	Primer sequence (5′ to 3′)
IGH Leader	**JH CONS**	CTTACCTGAGGAGACGGTGACC
	IGHL1/7	CTCACCATGGACTGSAYYTGGAG
	IGHL2	ATGGACAYACTTTGYTMCACRCTCC
	IGHL3a	ATGGARTTKGGGCTKWGCTGGGTTT
	IGHL3b	GGCTGAGCTGGGTTTTCCTTGTTGC
	IGHL4	CTGTGGTTCTTYCTBCTSCTGGTGG
	IGHL5	CCTCCTCCTRGCTRTTCTCCAAG
	IGHL6	CTGTCTCCTTCCTCATCTTCCTGCC

Reagent	*Final concentration*	1 x (µl)
Sterile $_dH_2O$	-	10.88
10x Buffer	1 x	2
$MgCl_2$ 25mM	1.5 mM	1.2
dNTP 10mM	0.2 mM	0.4
IGHL Primer mix (12.5 µM)	0.2 µM	0.32
AmpliTaq Gold® (5U/µl)	1.25 U	0.2
DNA (20ng/µl)	100 ng	5
Total		**20**

Cycles	Temp	Time
Hold	95°C	5min
35	95°C	45sec
	60°C	45sec
	72°C	60sec
Hold	75°C	10min
Hold	4°C	∞

2.4.1 Sanger sequencing

Two rounds of sequencing reactions were performed. First, the sequencing reaction was run using the JHcons primer which allows the identification of the V-gene used in each clonal rearrangement. The sequencing reaction is then repeated utilizing the specific family V-gene primer to ensure bi-directional sequencing of the whole stretch of the IGHV construct is completed and the V-primer sequencing read should be used for assessing somatic hypermutation status and stereotyping as outlined below.

The sequencing reaction reagents and cycling conditions are provided in table 2.2. Following ethanol clean-up the product was denatured 95°C for 10 min, mixed with 10 µl of formamide and run on the Applied Biosystems® Sanger Sequencing 3500 genetic analyser (Thermo Fisher).

Table 2.2: Reagent mix for Sanger sequencing reaction and cycling conditions

Sequencing reaction mix	Volume (µL)	Sequencing cycling conditions
Buffer Big Dye SeqPrimer (1 µM stock) Sterile dH_2O	3.75 0.25 2.0 3.0	96°C for 2mins 96°C for 10sec 55°C for 20sec 60°C for 4min } 30 cycles Hold at 15°C

2.4.2 Analysis of IGHV Sanger sequencing data

Analysis was performed through IMGT V-Quest (http://www.imgt.org/IMGT_vquest/input, (Brochet et al., 2008)) and ARResT/AssignSubset (http://tools.bat.infspire.org/arrest/assignsubsets/, (Bystry et al., 2016)) according to the recommended guidelines (Au - Agathangelidis et al., 2018). SeqScanner software (ThermoFisher) was used to visualize the sanger sequencing traces and to assess quality of each run. The IGHV-, D and J- gene was recorded for each clonal rearrangement detected, together with identity to germline and the result of the stereotype comparison. The amino acid sequence for the CDR3 was compared to the NGS panel data to ensure the identical clonal rearrangement was detected for each case. Details of the data obtained from each software are

shown in more detail in case study 1 and the results for the cohort are provided in table A7 in the appendix.

2.5 EuroClonality-next-generation sequencing DNA capture (ECNDC) assay

The design of the NGS capture panel was part of the NGS capture workpackage of the EuroClonality (EC) consortium (see https://euroclonalityngs.org/usr/pub/pub.php) and the proof-of-principle study was published on behalf of the EuroClonality consortium in 2014 (Wren et al., 2014, Wren et al., 2017). The ENABLE study discussed here was run in parallel to a large multi-centre validation study by the EuroClonality consortium, of which the scientists who performed the ENABLE study were part of. The results of the EC validation study have since been published (Stewart et al., 2019).

The work was performed following the validated standard-operating procedure (SOP) in use for diagnostic analysis in the Clinical Genomics Department, RMH. This protocol consists of the following parts as outlined in figure 2.1:

- Preparation of library
- Enrichment using the EuroClonality custom probe panel
- Massive parallel sequencing using Illumina NextSeq 500

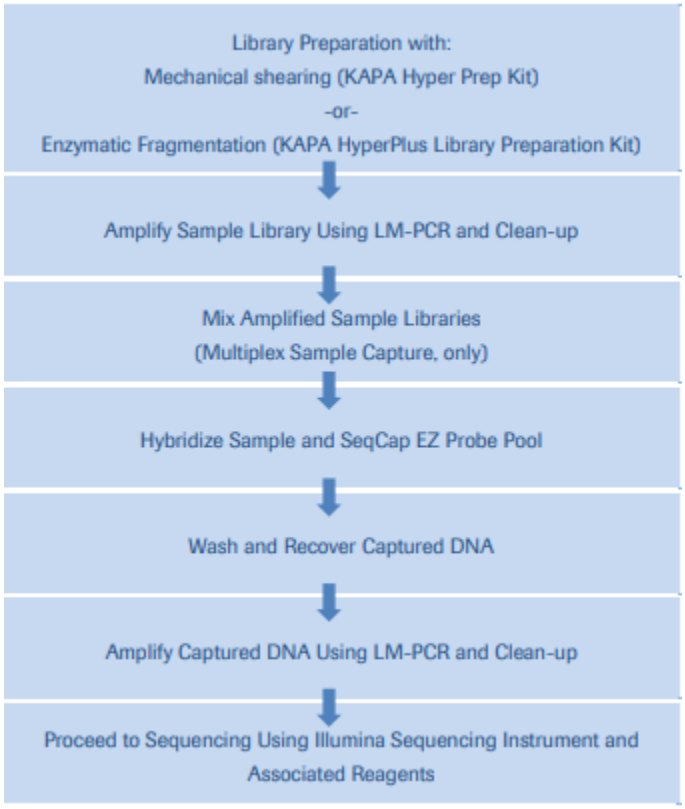


Fig. 2.1: Overview of the NGS capture laboratory workflow

2.5.1 Part 1: Preparation of targeted capture DNA library using KAPA™ HyperPlus Preparation Kit (#07962428001, Roche Sequencing Solutions) incorporating KAPA Unique Dual-Indexed Adapters (08861919702, Roche Sequencing Solutions)

In brief, 100 ng of high-molecular weight DNA obtained through the extraction process underwent enzymatic fragmentation to produce double-stranded (ds) DNA fragments of around 250-300 bp to enable short-read sequencing on the Illumina NextSeq platform. The fragmentation reaction was held at 37°C for exactly 20 min to ensure optimal fragmentation. These fragments were end-repaired followed by an A-tailing reaction to create an A-nucleotide overhang to allow adaptors designed with thymine overhangs to attach during the ligation of adaptors step. Adaptors containing the required sequences to allow the library fragments to attach to the sequencing flow cell as well as indexes for unique sample identification to permit multiplexing were attached to both ends of the DNA fragments during the adaptor ligation step. To each sample a unique adaptor (UDI) was assigned and documented on the worksheet and the sample incubated at 20°C for 15 min to allow

for adaptor ligation. Double-sided size selection followed using AmpureXP beads to only retain DNA fragments >250 bp and <450 bp.

The sample library was then amplified using Ligation-mediated (LM) PCR (see table 2.3) followed by another purification with AmpureXP beads.

Table 2.3: **Ligation-mediated PCR cycling parameters. Reaction volume was 50 µL and the heated lid was set to 99°C**

Step	Temperature	Duration	Cycles
Initial denaturation	98°C	45sec	1
Denaturation	98°C	15sec	6
Annealing	60°C	30sec	
Extension	72°C	30sec	
Final extension	72°C	60sec	1
Hold	4°C	∞	1

The purified, amplified libraries were quantified using the Qubit dsDNA High Sensitivity (HS) reagents (#Q3285, Invitrogen™ Qubit™, Thermo Fisher Scientific). In order to continue, the libraries had to have a fragment size distribution centring around 320 bp and a yield of at least 1µg.

2.5.2 Part 2: Capture-based enrichment of libraries using SeqCap EZ Choice Library capture probes (NimbleGen, Roche Sequencing solutions) and HyperCap enrichment kit

The target regions were identified based on expert review by the EuroClonality consortium (table A1 in the appendix). The design was submitted to NimbleGen/Roche for probe design against the GRCh38/hg38 genome assembly (SeqCap EZ choice Library, Roche Sequencing solutions).

Equal amounts of each library were mixed to obtain a sample pool of 1 µg. Number of samples pooled ranged from 12 to 22 paired (tumour + germline) samples per run.

To prevent annealing of target-specific probes to the adaptors, samples were first treated with Universal blocking oligos and COT Human DNA and the samples dehydrated using the SpeedVac Concentrator Plus (ID 5305000363, Eppendorf) set to 60°C. Following resuspension with 10.5 µL of hybridization mix, this was added

to 4.5 µL of the SeqCap EZ capture probe pool and left for hybridization at 47°C for 16-20 hours with the lid of the thermocycler set at 57°C to reduce evaporation.

The capture beads contained in the Kapa HyperPlus kit were prepared according to the kit instructions and 50 µL added to the hybridization sample pool and the mixture left at 47°C for 15 min. SeqCap EZ capture probes are biotinylated and can thus be retrieved using streptavidin-coated magnetic particles as outlined in the steps below.

Following the incubation of the pools with the magnetic beads, several wash steps were performed following the HyperPlus protocol, to remove unbound libraries thus selecting for the libraries that contain sequences of the defined target regions (enrichment for target regions). The magnetic-bead bound libraries were eluted in 40 µ L of PCR-grade water and 20 µL of this was taken forward into the amplification by LM-PCR using the parameters provided in tables 2.4 and 2.5. A final bead clean-up was performed post PCR using 90 µL of SeqCap EZ purification beads and eluted in 25 µL.

Yield and size distribution of the capture library pool was assessed on the 4200 TapeStation system (ID G2991AA, Agilent) using HighSensitivity D1000 reagents (#5067-5584 and #5067-5585, Agilent) and the pool quantified using HS Qubit reagents (as above).

For downstream processes the library pools were diluted to 4 nM.

Table 2.4: **Post-capture LM-PCR master mix**

Reagent	**Volume per library pool (µL)**
KAPA HiFi HotStart Ready Mix (2x)	25
Post-LM-PCR Oligos (5 µM)	5

Table 2.5: **Post-capture Ligation-mediated PCR cycling parameters**. Reaction volume was 50 µL and the heated lid was set to 99°C

Step	**Temperature**	**Duration**	**Cycles**
Initial denaturation	98°C	45sec	1
Denaturation	98°C	15sec	14
Annealing	60°C	30sec	
Extension	72°C	30sec	
Final extension	72°C	60sec	1
Hold	4°C	∞	1

Yield and size distribution of the capture library pool was assessed on the 4200 TapeStation system (ID G2991AA, Agilent) using HighSensitivity D1000 reagents

(#5067-5584 and #5067-5585, Agilent) and the pool quantified using HS Qubit reagents (as above).

For downstream processes the library pools were diluted to 4 nM.

2.5.3 Part 3: Sequencing and bioinformatic analysis

Sequencing was performed using 75 bp paired-end sequencing on a Mid Output NextSeq kit, 300 cycles and pools run on the NextSeq 500 (Illlumina). FASTQ files generated with FASTQ Generation Version: 1.0.0 (Illumina) were used for downstream analysis pipelines.

For the detection of IG-rearrangements and translocations from the NGS data, the ARResT/Interrogate, ARREsT/AssignSubsets and complementary pipelines developed by the Bioinformatics analysis team (BAT) led by Nikos Darzentas in collaboration with EuroClonality-NGS Working Group were accessed and analysis results generation was done by the BAT team (special thanks to Nikos Darzentas and Karol Pál). For further details about the pipelines please refer to the EuroClonality publication and the relevant bioinformatics papers (Bystry et al., 2016, Bystry et al., 2015, Stewart et al., 2019). Calls were compared to the Sanger sequencing data for the *IGHV*-rearrangements; light-chain restriction based on immunophenotyping results or immunohistochemistry (IHC) results and translocation calls reviewed manually using the Binary Alignment Map (BAM) files visualised through the Integrative Genomics Viewer (IGV, https://software.broadinstitute.org/software/igv/) from the Broad Institute (Robinson et al., 2017).

For each V(D)J-rearrangement, the pipeline provides detailed information on what is termed the ‘clonotype’: junction class (rearrangement type), 5’ gene, junctional segmentation and N-(D)-N region statistics (5’ gene deletions, N-(D)-N length, 3’ gene deletions), 3’ gene, junction amino acid sequence, and rearrangement productivity (see chapter 3, case study one and table 3.1 for an example).

For the single nucleotide variant analysis, the Illumina App Enrichment Version: 3.1.0 was run via Basespace (Illumina). It uses Isaac to align the reads against the reference genomes (a custom target manifest had to be included as the design of

the panel was against GRCh38). Somatic variants are called using Pisces and germline variants using Starling. The Illumina Annotation Engines provides the annotation of the variants with the Pluggable Universal Metrics Analyzer (PUMA) calculating alignment and variant metrics. Further details and citations for the components of the pipeline can be found here: https://emea.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/enrichment.html

Filter settings were as follows: Variant Read Frequency > 0.049 (for tumour samples) and Variant Read Frequency > 0.19 (for germline samples) including all consequences as follows:

a. Initiator codon (ATG) loss
b. Missense
c. Deleterious
d. Deleterious - Low Confidence
e. Probably Damaging
f. Possibly Damaging
g. Incomplete terminal codon
h. Stop gained
i. Stop loss
j. Splice donor
k. Splice acceptor
l. Splice region
m. Frameshift Indels
n. Inframe deletion
o. Inframe insertion

Variants called by the Illumina App were exported into Excel for tumour and germline samples and the calls of the germline sample compared to the tumour samples and subtracted. The remaining variant calls were subjected to variant analysis and interpretation according to the SOP in place for the diagnostic service at Clinical Genomics (RMH) for somatic variant analysis in accordance with the relevant published guidelines (Li et al., 2017, Froyen et al., 2019). Variant analysis was performed by a qualified, state-registered Clinical Scientist. Variants classed as pathogenic/likely pathogenic or with potential clinical actionability were included in this thesis.

For the copy number variant analysis FASTQ files were submitted to the EuroClonality pipeline (hosted by Precision Medicine Centre of Excellence, Centre for Cancer Research & Cell Biology, Queens University Belfast, special thanks to Shambhavi Srivastava for running these samples). In addition, coverage data obtained for the tumour and germline sample was compared and the ratio plotted

via excel to visualise any changes. This analysis is ongoing and only preliminary examples are included here (see results chapter for further details).

3. Results

3.1 Sample cohort

In total 108 samples underwent NGS testing. Results were included in this review for 100 samples: Two samples (REF-4383-0001 and RA2-4383-0003) were excluded from the study due to lack of neoplastic cell content confirmed by flow cytometry. NGS set up failed due to a technical error for RAX-4383-0001 (VLeader data available) and REF-4383-0004. In three additional samples, tumour infiltration is assumed to be low: For R1K-4383-0003 no flow cytometry is available, and the white cell count was $7x10^9$/L with a lymphocyte count of $1.4x10^9$/L. No clonal markers were found for this sample by either the NGS panel or Sanger sequencing for Immunoglobulin (IG) rearrangements. Similarly, no evidence of clonality was seen for RVR-4383-0003. For all three samples the mean coverage depth was >500x indicating that the NGS analysis performance was within acceptable limits.

For RMH-4383-0003, NGS analysis showed SNVs but all with VAFs below 10% (one classed as likely pathogenic, two variants of unknown significance (VUS)), and VLeader failed to amplify a clonal rearrangement. NGS analysis for IG-rearrangements is pending. The results were only included on the SNV cohort (n=101).

Uncertain results were obtained for RVR-4383-0024 due to the low number of supporting reads in the NGS and no amplification of a clonal product by VLeader PCR. For this case, a pathogenic *MYD88* variant was detected on a different sample and no evidence of this variant was found in this sample. See tables 3.3 and A2 for an overview of results per sample.

3.2 Panel performance

The full validation of the panel has been published (Stewart et al., 2021) and is thus not discussed here in detail. Figure 4.1 shows a summary of the analytical performance of the EC-NDC assay as published by Stewart et al. (Stewart et al., 2019, Stewart et al., 2021). The panel achieved excellent sensitivity and specificity for the detection of single-nucleotide changes (SNVs) down to a variant-allele

frequency (VAF) of 5% (98% and 100%, respectively). With the same cut-off of 5% VAF, sensitivity and specificity were 97% for the detection of clonal Ig/TR rearrangements and 95% and 99% for the detection of translocations. The panel detected 145 out of 152 known translocation and was able to identify two cases with translocations that were not detectable by standard-of care testing. The translocations detected by other methods and not found by NGS were likely below the level of detection in the sample analysed or may have breakpoints in a location not included in the panel design. For IGH-rearrangements results were compared against the current gold-standard PCR methodology (EuroClonality/Biomed2) and a low false negative rate of 8 out of 162 cases was seen. For 41 cases without clonal IGH rearrangement by PCR, NGS showed clonal rearrangements in 11 and ddPCR confirmed the NGS results in 63% of these, pointing to a higher sensitivity of NGS compared to PCR-based clonality assessment.

A high specificity was achieved for copy-number changes when a cut-off of 20% was used (100% specificity) but the sensitivity was only 88% due to an error in the design of the baits resulting in a high number of false negatives for 13q deletions (24 out of 43 cases with 13q deletions, see table 4 in (Stewart et al., 2021)for details)

A mean target coverage of >500x was used as a cut off for acceptable performance (validation study 'Pass') in the EuroClonality (EC) validation study and a level of detection of 4% for SNVs was confirmed with tumour infiltration ≥20%. The mean coverage across all runs was 1193x for the tumour samples. A higher coverage mean was deemed appropriate as the tumour infiltration was not assessed for each sample in the ENABLE study.

Suboptimally-performing target regions are summarised in the supplementary information of the EC validation paper and were identical in the performance of the runs in the ENABLE study (table A5 in the appendix). The only coding regions were exon 2 and 9 for *TCF3* and variants in these regions could have been missed. For the validation study regions were classed as underperforming if the coverage for the region was greater than 2 standard deviations below the mean in half of the samples. These targets were excluded in the evaluation of CNVs.

3.3 Detection of clonal IG-rearrangements

The design of the EuroClonality capture panel includes the IGH-locus as well as the IGK- and IGL-light-chain loci to allow for the detection of clonality and to be able to characterize the specific IG-rearrangement (table A7 in the appendix provides the regions of interest (ROI)). The NGS data was run through the ARResT/Interrogate pipeline to identify clonal rearrangements and to delineate the complementarity-determining region 3 (CDR3) (Bystry et al., 2016). Given the short read lengths of 75 bp paired-end reads, assessment of somatic hypermutation status in accordance with the published guidelines (Ghia et al., 2007) was not possible and additional analysis using the recommended Sanger sequencing approach using VLeader primers was performed in parallel. This also served as a control for the identification of the clonal *IGHV*-rearrangement and gene usage using this novel capture-NGS approach and the ARResT/Interrogate pipeline as the validation study by the EuroClonality NGS consortium was still in progress (now in press).

The ARResT/Interrogate pipeline performs two alignment steps to improve the detection of and differentiation between V(D)J-rearrangements. First reads are aligned to the germline sequence using the IMGT database reference genomes and then reads are aligned against each other to identify clonotypes. Clonotypes are defined as V(D)J-rearrangement which use the same V-gene and have identical CDR3-regions based on their amino-acid sequence. An example of the result output is provided in table 3.1. A minimum of 5 supporting reads had to be called for a rearrangement to be included in the output and >10% in locus and in class were expected for the dominant rearrangement. Similar rearrangement listed separately due to small changes in the CDR3 sequence as exemplified in table 3.1 were considered to represent the same rearrangement if the V- and J-gene was identical to the main call and the latter was the productive rearrangement. The dominant rearrangements were concordant with the VLeader analysis in all cases as discussed in more detail below.

Table 3.1: ARResT/Interrogate pipeline output: For each sample the number of reads in which a junction has been identified is provided (usable reads, column A). From the FASTQ data, reads have already been deduplicated and filtered based on standard quality criteria at this point. The class of rearrangement (e.g. V-J, D-J) and its locus is listed together with the specific rearrangement details: Gene usage, CDR3 amino acid sequence (column K 'junction') as well as the insertions/deletions that have occurred at the junctions (column J 'event1'); for example -1/26/-5 means that 1 nucleotide was deleted from the V-gene, the junction N-D-gene-N contained 26 nucleotides and 5 nucleotides were deleted from the J-gene. The number of reads containing the specific rearrangement are listed (column G 'fragments') together with the % this rearrangement represent of the total number of reads with a rearrangement from the same locus (e.g. IGH, column H) as well as the % in class, which is the proportion of e.g. D-J rearrangement this one represents (column I '%class'). The main clonal rearrangement for each locus is highlighted in green. Rearrangements with minor changes compared to the dominant rearrangement are listed separately and some examples are included in the table above. Highly similar calls were merged to increase the number of supporting reads for the rearrangement for the analysis. Column M ('event comments') states whether the rearrangement is deemed productive or unproductive with additional information if anchors (Cys104 and Trp118) are missing or a stop codon occurs.

	A	B	C	D	E	F	G	H	I	J
1	usable	class1	class2	locus	gene	gene partner	fragments	%	%class	event1
2	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	128	33.9	67.7249	D3-9 -10/19/-14 J6
3	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	6	1.59	3.1746	D3-9 -10/19/-14 J6
4	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
5	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
6	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
7	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
8	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
9	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
10	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
11	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
12	1464	R	DJ:Dh-Jh	IGH	IGHD3-9	IGHJ6	1	0.26	0.5291	D3-9 -10/19/-14 J6
13	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	150	39.7	79.3651	V3-23 -1/26/-5 J4
14	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/25/-4 J4
15	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
16	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
17	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
18	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
19	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
20	1464	R	VJ:Vh-(Dh)-Jh	IGH	IGHV3-23	IGHJ4	1	0.26	0.5291	V3-23 -1/26/-5 J4
21	1464	R	VJ:Vk-Jk	IGK	IGKV1-17	IGKJ2	220	41.4	80	V1-17 -5/2/-7 J2
22	1464	R	intron-Kde	IGK	intron	Kde	112	21.1	43.75	intron -3/0/-1 Kde
23	1464	R	intron-Kde	IGK	intron	Kde	8	1.5	3.125	intron -3/0/-1 Kde

Table 3.1 (continued): ARResT/Interrogate pipeline output (continued)

	A	B	K	L	M
1	usable	class1	junction	junction nt seq	event comments
2	1464	R	VVLRYFDWLVGCSS#YGMDVW	gtggtattacgatattttgactgg..... CTCGTTGGGTGTTCCTCTATC .....##cggtatggacgtctgg	potentially productive;cap=unk;miss5
3	1464	R	VVLRYFDWLVGCSSIRYGR	gtggtattacgatattttgactgg..... CTCGTTGGGTGTTCCTCTA .....tacggtatggacgn	potentially productive;cap=unk;missing 2nd anchor;missing GXG;unsafeJ;trimmed;miss5;miss3
4	1464	R	VELR*FDWLVACSS#YGMAVW	gtggaattacgataatttgactgg..... CTCGTTGCGTGTTCCTCTA .....##tacggtatggccgtctgg	potentially productive;cap=unk
5	1464	R	VELRYFDWIVGFSS#YGMDVW	gtggaattacgatattttgactgg..... ATCGTTGGGTTTTCCTCTA .....##tacggtatggacgtctgg	potentially productive;cap=unk;missing GXG;miss5;miss3
6	1464	R	VELRYFEWLVLCSS#YGMDVW	gtggaattacgatattttgaatgg..... CTCGTTTTGTGTTCCTCTA .....##tacggtatggacgtctgg	potentially productive;cap=unk;miss5
7	1464	R	VLLLYFDWLVGCSS#YGMDVW	gtgttattactatattttgactgg..... CTCGTTGGGTGTTCCTCTA .....##tacggtatggacgtctgg	potentially productive;cap=unk;miss5
8	1464	R	VVIRYLDWVVGCAS#YGMDVW	gtggtaatacgatatttagactgg..... GTCGTTGGGTGTGCGTCTA .....##tacggtatggacgtctgg	potentially productive;cap=unk;miss5
9	1464	R	VVLRFFDWLVGCSS#*GLDVW	gtggtattacgatttttttgactgg..... CTCGTTGGGTGTTCCTCTA .....##tagggtttggacgtctgg	unproductive;cap=unk;stop codon NJ
10	1464	R	VVLRYFDWLLGCTS#YGMDFW	gtggtattacgatattttgactgg..... CTCCTTGGCTGTACCTCTA .....##tacggtatggacttctgg	potentially productive;cap=unk
11	1464	R	VVLRYFDWLVGCSS#*GLEVW	gtggtattacgatattttgactgg..... CTCGTTGGGTGTTCCTCTA .....##tagggtttggaggtctgg	unproductive;cap=unk;stop codon NJ
12	1464	R	VVLRYFDWLVGCSS#YGMAVW	gtggtattacgatattttgactgg..... CTCGTTGGGTGTTCCTCTA .....##tacggaatggccgtctgg	potentially productive;cap=unk
13	1464	R	CAKAFSSGWFWGFDYW	tgtgcgaaag. CCTTTAGCAGTGGCTGGTTTTGGGGG .....tttgactactgg	productive;cap=unk
14	1464	R	SAKAFTSGRFSGFDYM	tctgctaaag. CCTTTACCAGTGGCAGGTTTTCTGG ....ctttgactacatg	productive;cap=unk;missing anchor(s);missing 2nd anchor;miss5;miss3
15	1464	R	*AKAVSSGGFWGFDDR	tgagcgaaag. CCGTTAGCAGTGGCGGGTTTTGGGGG .....tttgacgacagg	unproductive;cap=unk;stop codon ju;missing anchor(s);missing 2nd anchor;miss5
16	1464	R	C*KAVSTGGIWGFDYW	tgttagaaag. CCGTTAGCACTGGCGGGATTTGGGGG .....tttgactactgg	unproductive;cap=unk;stop codon ju;miss5
17	1464	R	CAKAFSSGWFWVFDYW	tgtgcgaaag. CTTTTAGCAGTGGCTGGTTTTGGGTG .....tttgactactgg	productive;cap=unk;miss5
18	1464	R	CAKAFSTGSFWAFDYS	tgtgcgaaag. CCTTTAGCACTGGCTCGTTTTGGGCG .....tttgactactcg	productive;cap=unk;missing anchor(s);missing 2nd anchor;missing GXG;miss5
19	1464	R	CAKAISSGWLWGFDYW	tgtgcgaaag. CCATTAGCAGTGGCTGGTTATGGGGG .....tttgactactgg	productive;cap=unk;miss5;miss3
20	1464	R	CPKAFSSGWIWRFDYW	tgtccgaaag. CCTTTAGCAGTGGCTGGATATGGCGG .....tttgactactgg	productive;cap=unk;miss5
21	1464	R	CLQHNSYVF	tgtctacagcataatagttac..... GT .....tttt	productive;cap=unk
22	1464	R	PCVCPIDAAVASFP#SPSGSPGR	gcctgtgtctgcccgattgatgctgccgtagccagctttcctg... .#gagccctagtggcagcccagggcga	unknown;cap=unsafe
23	1464	R	PCVCPIDAAVASFPGALVAAQG	ccctgtgtctgcccgattgatgctgccgtagccagctttcctg... .gagccctagtggcagcccagggcgn	unknown;cap=unsafe;trimmed;miss3

NGS analysis was available in 97 samples and in all cases at least one clonal rearrangement at an IG-locus was identified. Clonal light-chain rearrangements involving the IGK-locus and the IGL-locus, if lambda-expressing, were identified in all 97 samples. In three samples, IGK and/or IGL were the only clonal IG -rearrangements detected with NGS and VLeader analysis unable to identify a clonal *IGHV*-rearrangement suggesting that the cases had a high SHM status preventing annealing of the PCR primers as well as making reliable alignment of the NGS reads impossible.

For 70 cases the NGS and VLeader results were concordant in identifying the same productive *IGHV*-rearrangement including an identical CDR3 region. In an additional 16 cases, the CDR3 amino acid and J-gene usage was concordant, but NGS analysis was unable to unequivocally assign a V-gene. As one of the possible options always included the V-gene identified by VLeader testing, these cases were also considered concordant. Nine of the 16 cases had a high SHM load above 5% which may have contributed to the different V-gene assignment by NGS; in addition ongoing clonal diversification could be present which is well-documented in various B-cell NHL entities (Sutton et al., 2009, Granai et al., 2020). This was not further investigated in this study. It is also important to note that shorter read lengths of 75 bp were employed instead of the recommended 150 bp reads and that the bait design did not include V-genes. It is therefore possible that fewer reads cover the full V-gene sequence were present in this group of cases making assignment to a V-gene less reliable. The EuroClonality NGS group has worked through a comparison of 150 bp and 75 bp reads as part of their validation and has shown that all rearrangements and translocations were detected successfully but the design was aimed at the detection of clonality and not at the characterisation of SHM/IGHV gene-usage and for this the 150 bp reads maybe more appropriate (personal communication Prof. D. Gonzalez).

For 5 samples the analysis by the ARResT/Interrogate pipeline had not yet been performed. VLeader analysis was successful in 4/5 samples with one sample showing no amplification. This could be a technical error although the only flow cytometry data available was 12 months old but showed 35% clonal B-cells at the time and the patient had not received any treatment in the meantime to the best of our knowledge. Analysis of SNVs showed several variants but all were <10% VAF

supporting the notion of low tumour burden in this sample although subclonality cannot be excluded.

In the remaining 8 cases where NGS was able to detect a clonal *IGHV*-rearrangement, no amplification was seen by Sanger in 3 cases: In two, the most likely explanation is a mutation located under the primer binding site as the analysis for each samples was repeated and failed to amplify a clonal product. In one case, the repeat analysis has not yet been performed and technical error cannot be excluded. In 5 cases the VLeader set up amplified more than one rearrangement and the analysis will need to be repeated on RNA to confirm the productive rearrangement.

3.4 Light chain restriction

Presence of a productive kappa- or a productive lambda-rearrangement together with unproductive kappa rearrangements was used to predict the light chain restriction of the clone from the NGS data. When compared to the light chain restriction as demonstrated by flow cytometry or IHC, which was available in 66 cases, concordance was seen in 65 of the 66 cases (table A7 in the appendix). One case, BEL-4383-0003 was predicted to have kappa light-chain restriction based on NGS finding of a single, productive kappa-rearrangement which was in agreement with the flow cytometry results; however, IHC reported lambda light-chain restriction. Review of the case showed that it had a CLL score of 4 (and thus should not have been enrolled in the study) but that there was also a possible small clonal plasma cell population present alongside which could explain the different findings. NGS however did not show evidence of a second clone and no further clonal markers were detected.

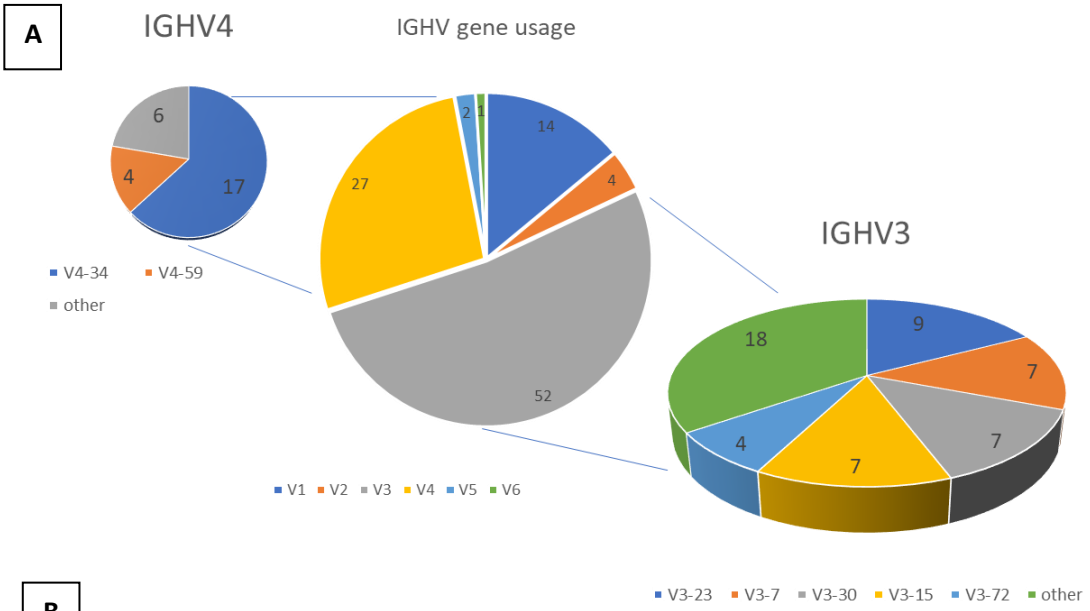
Overall, 61 cases showed kappa light-chain restriction: 45 cases with immunophenotyping/IHC analysis and an additional 15 cases based on NGS alone. Lambda light-chain restriction was seen in the remaining 36 cases, including 29 with concordant immunophenotyping/IHC analysis and 5 by NGS only. In healthy B-cells the ratio of kappa: lambda is roughly 2:1 and although the ratio is slightly lower, the increase in lambda-expressing clones is unlikely be significant.

3.5 Somatic hypermutation (SHM) status and stereotyping analysis

Combining the NGS and Sanger data, for the 100 samples where a clonal *IGHV*-rearrangement was identified, the V-gene family usage is shown in figure 3.1. The majority of V-genes were from the V3 family (n=52, 52%), followed by V4 (n=27, 27%) and V1 (n=15, 15%), which is in-line with the expected distribution in healthy subjects as well as in cases with neoplastic disease, e.g. CLL. For the V3-gene family, *V3-23* and *V3-7/3-30* were the most common and for the V4-gene family it was mainly *V4-34* and *V1-2* and *V1-69* for the V1-gene family. In particular *IGHV4-34* is known to have a high prevalence in neoplastic disease, particularly CLL and MCL and it is associated with recognition of auto-antigens (Sutton et al., 2009, Defrancesco et al., 2020).

SHM status assessment was possible in 90/100 cases and was assessed using IMGT-VQuest (Brochet et al., 2008) together with ARResT/Interrogate for the assignment of stereotypes (Bystry et al., 2015). Analysis was performed according to the SOP in place for assessment of *IGHV* mutational status in CLL in routine clinical practice at RMH which follows the guidelines published by the ERIC consortium (Ghia et al., 2007, Rosenquist et al., 2017) and uses the cut-offs described in figure 3.1. The majority of cases had mutated *IGHV*, with 12 of the mutated cases showing very high-hypermutation >10%. Of the 21 unmutated case, nine showed complete homology to the germline (SHM 0%). Importantly, none of the cases were assigned to the 18 main stereotypes associated with CLL (Rossi and Gaidano, 2010, ten Hacken et al., 2019) which points towards a different etiology – perhaps a different antigen- in these cases compared to CLL, even in those cases potentially falling into the atypical CLL category (see below). In the 6 MCL cases, none carried a stereotyped *IGHV3-21* or *IGHV4-34* which have been reported in around 10% of MCL cases (Hadzidimitriou et al., 2011).

A



B

Gene family	IGHV gene	Cases
V1		14
	V1-2	6
	V1-69	5
	other	3
V2		4
	V2-5	2
	V2-70	2
V3		52
	V3-23	9
	V3-7	7
	V3-30	7
	V3-15	7
	V3-72	4
	other	18
V4		27
	V4-34	17
	V4-59	4
	other	6
V5	V5-51	2
V6	V6-1	1

C

	SHM	% (90 cases)
Unmutated	21	23
Borderline	9	10
Mutated	60	67

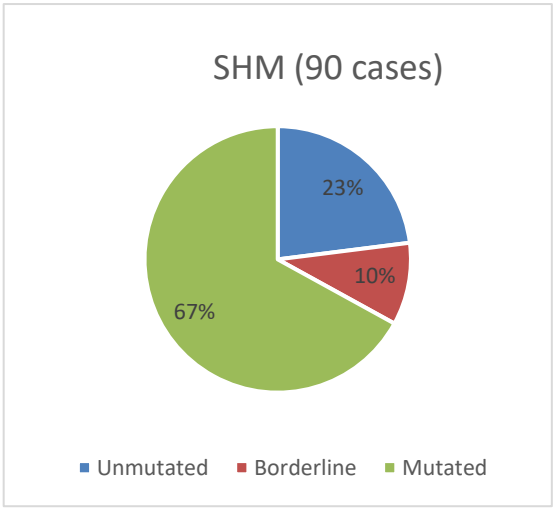


Fig. 3.1: IGHV gene usage and SHM status. A shows the distribution of *IGHV*-gene usage across the cohort of 100 cases for which a *IGHV* could be identified by NGS and/or VLeader analysis. The largest group are V3- and V4-family genes and the breakdown into individual *IGHV*-genes is provided in B. Table and figure under C shows the number of unmutated, mutated and borderline SHM cases. Unmutated >98% homology to germline, borderline <98>97.1% and mutated ≤97.

As these clonal neoplasms exhibit features of mature germinal-centre-derived B-cells, they are in general expected to have undergone clonal selection within the germinal centre including somatic hypermutation additions (see figure 1.2, chapter 1.2). Accordingly, some level of SHM was detected in 81 of the 90 cases as only nine cases showed no evidence of SHM. Nevertheless, it is well known that in CLL, around half of cases show unmutated *IGHV*. If the cut-offs in use for prognostication in CLL are applied here, then 77% of cases in this cohort were mutated (including borderline mutated) but other studies outside CLL have considered any level of SHM as mutated which increases the percentage of mutated cases to 90%. Unmutated cases could either be derived from early-GC where no or only minimal SHM has occurred or have undergone differentiation to memory B-cells outside the germinal centre, a process that has been demonstrated for CLL (Gemenetzi et al., 2020).

3.6 Detection of translocations

Chromosomal translocations arising from the IG-loci are a hall mark of B-cell lymphomas and are often associated with a specific subtype (table 1.1). Their presence is routinely investigated in a variety of B- and T-cell neoplasms and in addition to diagnostic information these can carry prognostic information. For the ENABLE cohort given the lack of B-symptoms and with a diagnosis of an indolent small-cell B-NHL, no or limited FISH analysis for the detection of particular translocations was included in their diagnostic workup as they did not fulfill the relevant criteria for further testing according to the SIHMDS pathways.

The design of the capture panel included the D- and J-genes as well as the switch regions for all the IG-constant regions. In addition, the well-known breakpoints of common fusion partners e.g. *BCL2* and *CCND1* were covered by target-enrichment baits to increase the sensitivity of the assay (table A4 in the appendix). The EuroClonality NGS capture approach thus allows for the detection of IG-fusions arising from the IGH/IGK and IGL loci regardless of location of breakpoint and fusion partner (Wren et al., 2014, Wren et al., 2017, Stewart et al., 2019).

Output from the ARResT/Interrogate pipeline was used for the detection of translocations and in addition, a manual review of the traces on IGV focusing on the J-genes, all switch regions as well as *BCL2, CCND1* and *BCL6* breakpoint locations

was performed for all cases by visualising the BAM files in IGV. This mimicked the approach taking in the EC-validation study which was progressing in parallel. All translocations were identified by the pipeline and no additional findings were made by manual analysis.

In the 100 cases included, a translocation between *IGH* and chromosome 11 (*CCND1*) was detected in 6 cases, two showed t(14;18) *IGH-BCL2* and two cases carried a t(14;19) *IGH-BCL3* translocation. In an additional four cases, chromosomal rearrangements were detected between unusual partners, one arising from the *TRA/D* locus and involving *CDK6* which has now been published as a recurring rearrangement in B-cell-lymphomas (Gaillard et al., 2021), two involving chromosome 7 and one arising from a non-Ig locus on chromosome 14 and breaking into chromosome 11. The latter three have not yet been further investigated and will not be discussed in this thesis. A summary of fusion breakpoints and partners are provided in table 3.2. The case with *TRA/D-CDK6* will be discussed as a case study 2 below.

Table 3.2: **Summary of samples in which a well-established translocation was detected by the capture NGS approach**. Usable reads= number of unique sequencing reads containing a rearrangement (either V(D)J or translocations) as defined by the ARResT/Interrogate pipeline (requires presence of a RSS recombination signal sequence and presence of a full CDR3 junction or cross-chromosomal junctions). Breakpoints and translocation partners are shown together with the combination on each derivative chromosome. Supporting reads are reads going across the junction, not split reads which explains the lower number of reads called by the pipeline compared to the additional split reads visible on IGV. MBR= major breakpoint region in *BCL2*, MTC= major translocation cluster in *CCND1*; NA= not available as translocation appears unbalanced. FISH results are provided with 4 fusions confirmed, 4 results pending and two undetected by FISH (see text for details).

Trial ID	usable reads	Fusion	Break Gene 1	Location	Break gene 2	Location	supporting reads	Break Gene 1	Location	Break gene 2	Location	supporting reads	FISH
RMH-4383-0011	1182	IGH-BCL2	IGHJ5	chr14:105,864,086	BCL2 centromeric	chr18:63,083,439	59	BCL2 centromeric	chr18:63,083,432	IGHD3-10	chr14:105,888,555	40	Positive
REF-4383-0007	206	IGH-BCL2	IGHJ5	chr14:105,863,259	3'MBR BCL2	chr18:63,121,516	16	3'MBR BCL2	chr18:63,121,491	IGHD2-2	chr14:105,916,833	10	Negative
REF-4383-0003	520	IGH-BCL3	centromeric BCL3	chr19:44,729,189	IGHA1switch	chr14:105,710,539	4	IGHM switch	chr14:105,859,342	centromeric BCL3	chr19:44,729,148	34	Pending
RVR-4383-0016	1104	IGH-BCL3	BCL3 5'UTR	chr19:44,747,969	IGHA1switch	chr14:105,711,127	0	IGHM switch	chr14:105,859,968	BCL3 5'UTR	chr19:44,747,949	37	Pending
RMH-4383-0005	*1763*	*CCND1-IGH*	*unknown*	*unknown*	*unknown*	*unknown*	*NA*	*Centromeric*	*chr11:69,639,244*	*IGHD2-2*	*chr14:105,916,836*	*75*	*Negative*
RVR-4383-0002	216	CCND1-IGH	IGHJ4	chr14:105,864,264	CCND1 MTC	chr11:69,532,139	47	CCND1 MTC	chr11:69,532,134	IGHD2-21	chr14:105,888,555	3	Positive
RAX-4383-0003	584	CCND1-IGH	IGHJ4	chr14:105,864,249	CCND1 centromeric	chr11:69,624,290	45	CCND1 centromeric	chr11:69,624,281	IGHD3-22	chr14:105,886,049	42	Pending
RDU-4383-0005	868	CCND1-IGH	IGHJ6	chr14:105,863,255	CCND1 centromeric	chr11:69,586,533	76	NA	NA	NA	NA	NA	Positive
REF-4383-0013	1230	CCND1-IGH	IGJH5	chr14:105,863,858	CCND1 centromeric	chr11:69,626,719	719	CCND1 centromeric	chr11:69,626,721	IGHD3-22	chr14:105,886,053	69	Positive
RHU-4383-0006	1816	CCND1-IGH	IGHJ5	chr14:105,863,899	CCND1 centromeric	chr11:69,619,627	174	CCND1 telomeric	chr11:69,686,152	IGHJ4	chr14:105,864,057	5	Pending

3.6.1 Chromosomal rearrangements involving *IGH* and *BCL2*

In two cases, a t(14;18) translocation arising from the *IGH*-locus was detected. The translocation is most-commonly detected in Follicular Lymphoma but can also be seen in DLBCL and in rare cases of CLL (Swerdlow et al., 2016a). They usually arise during D-J-recombination involving recombination-activation genes (RAG)-mediated processes and several breakpoint clusters have been described for *BCL2* (Espinet et al., 2008).

Figures 3.2 and 3.3 show the construction of the rearrangement in the two samples in this cohort.

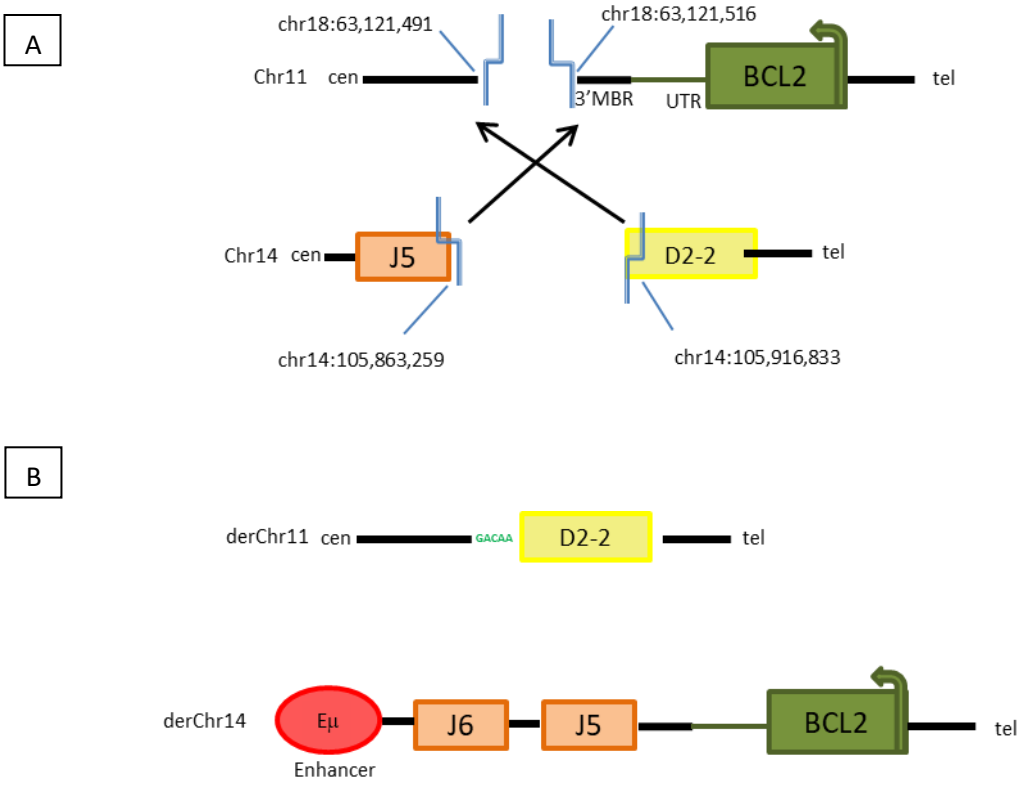
For case REF-4383-0007, the translocation occurred during the D-J recombination and involved the *D2-2* gene and *IGHJ5*. The double-strand break on chromosome 18 occurred in the 3' major breakpoint region (3'MBR) around 4kb downstream of the MBR located in the middle of the untranslated part of the last exon of *BCL2* which is the most commonly reported site for the translocation. Insertion of 5 nucleotides occurred at the fusion site creating derivate chromosome 11. A larger number of nucleotides appears to have been inserted on derivative chromosome 14, but the exact sequence could not be determined due to the lack of overlapping reads. As a result of this translocation the super enhancer Eμ located in the intronic region between the J-genes and the constant regions on the IG-locus drives the uncontrolled expression of proto-oncogene *BCL2* as part of the neoplastic process (Magnoli et al., 2019).

The low number of supporting reads (26 in total, 10 for D2-2- break and 16 for the *BCL2* site called by the pipeline) could indicate low tumour infiltration in the sample which could also explain the failure of Sanger sequencing to identify an *IGHV* rearrangement for this case (table A7). Usable reads containing a rearrangement as identified by the ARResT/Interrogate pipeline for V(D)J-rearrangements or translocations also had the lowest number of any sample with translocation with a total of 206. NGS mean coverage overall was 511x which was at the lower end but still acceptable. FISH analysis using the Abbott Molecular IGH-BCL2 dual-colour, dual fusion probe failed to demonstrated the presence of a t(14;18). The examination included 150 nuclei and was performed at the genomics laboratory in Bristol. The probe is designed and has been validated to detect breaks occurring in

the 3'MBR and D-J and the fusion seen here is thus in scope for the assay. In the rare cases of atypical CLL with an *IGH-BCL2* fusion, these have been shown to be secondary events and it may therefore be possible that the translocation is present in a subclone below the level of detection for FISH analysis (Swerdlow et al., 2016a, Fang et al., 2019). No further information was available to review the morphology or B-clone phenotype for this case at this point.

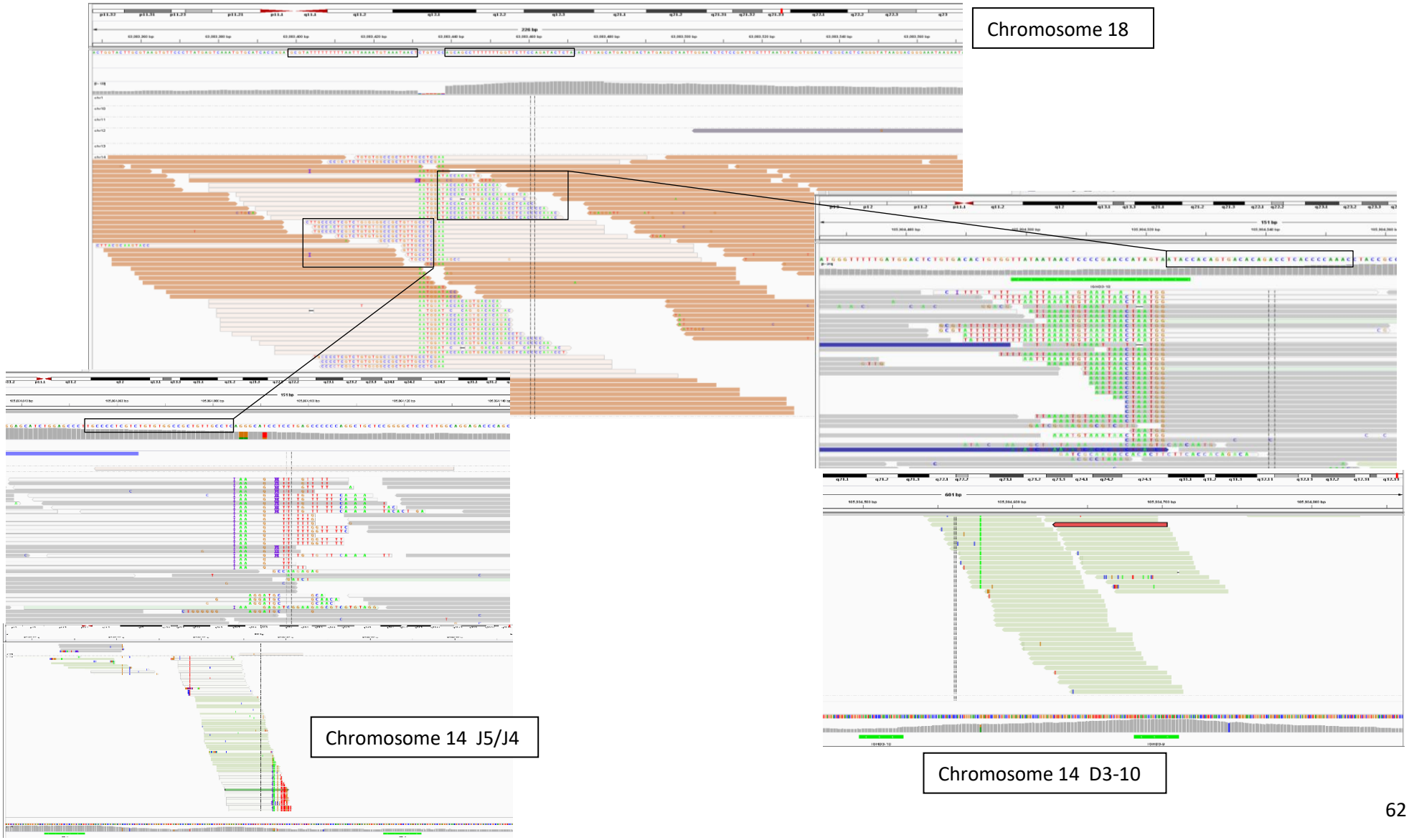
The second sample carrying a t(14;18) *IGH-BCL2* fusion was RMH-4383-0011. The break on chromosome 14 occurred during D-J rearrangement involving *D3-10* and *IGHJ5*: the break on chromosome 18 was outside and further centromeric to the MBR and mcr regions of *BCL2* around 45 kb away from the end of exon 3 of *BCL2*. Detail of the breakpoints and balanced translocation products are shown in figure 3.3. There was good read support for this event with 40 read for derivative chr18 (derchr18) and 59 for derchr14. The result was confirmed by FISH using the Vysis dual fusion probe (cytogenetics department RMH). Of 169 cell scored, 54% had an *IGH-BCL2* fusion signal pattern consistent with the presence of a t(14;18)(q32;q21) translocation. The diagnosis for this patient was amended from B-NHL, NOS to Follicular lymphoma WHO ICD-O 9690/3 upon review of all findings and the patient is being treated accordingly.

Fig. 3.2: Example of a t(14;18) IGH-BCL2 translocation detected in the study. Schematic representation (not to scale) of the translocation breakpoints and partners (A) and the resulting derivative chromosomes (B). The double-strand DNA break on chromosome 14 occurred during D-J recombination involving *J5* and *D2-2*. The break on chromosome 18 is located in the 3'MBR region centromeric of *BCL2*. As a result of the fusion the IGH super enhancer Eμ influences the expression of *BCL2* now located downstream of the enhancer on the derivative chromosome 14, leading to BCL2 overexpression. The super enhancer Eμ is depicted in red, the J genes are shown as orange and the D-genes as yellow boxes. *BCL2* is shown in dark green. Black lines depict the DNA sequence and blue jagged lines the double strand breaks. Cen- = centromeric; tel = telomeric; der = derivative



Next page:

Fig. 3.3: Example of precise mapping of the breakpoints for a t(14;18) *IGH-BCL2* translocation detected in the study. The top section shows the read alignments for chromosome 18 visualised on IGV. All reads are coloured by chromosome of mate and orange reads highlight that the paired read is located in chromosome 14 (visualisation and colour-coding from IGV). To the right screenshots of the breakpoint region around *IGHD3-11* on chromosome 14 showing split-reads and fusion-spanning reads; Pale green colouring means that the partner read is located in chromosome 18 and unmatching bases are shown within reads. On the left the same is shown for the second break on chromosome 14 located between *IGHJ5* and *IGHJ4*. Boxes show where the sequences reads match the germline sequence on the partner chromosome; sequences can be read across the breakpoints specifically showing that on chromosome 18, 7 bases were deleted and on derivative chr18, 5 bases were inserted (ATGG) and for derivative chr14 three based (GAA) were inserted. The banding pattern of the chromosome is shown at the top with the DNA germline sequence according to GRCh38 below. The track underneath shows the coverage data for the sample analysed.


Chromosome 18
Chromosome 14 J5/J4
Chromosome 14 D3-10

3.6.2 Chromosomal rearrangements involving *IGH* and *BCL3*

In two cases, a t(14;19) translocation arising from the IGH locus was detected. The translocation is described as a recurrent albeit rarer finding in B-LPDs and is most-commonly assigned a diagnosis CLL or atypical CLL (M et al., 2016). Like *BCL2*, *BCL3* is a proto-oncogene which under the influence of the Eμ enhancer becomes overexpressed in B-cells carrying the translocation.

In contrast to the *IGH-BCL2* translocations discussed above, the two *IGH-BCL3* fusions likely arose during the process of class-switch recombination (CSR) rather than D-J rearrangements. The breakpoints lie within the switch region of the constant genes *IgM/IgA/IgG* although the exact location is difficult to identify due to the homology of the areas. In both cases the break on chromosome 19 has occurred in well-defined regions: one directly adjacent to the start of the open reading frame of *BCL3* (case RVR-4383-0016) and the other 15 kb centromeric to *BCL3* (REF-4383-0003, not shown).

For case RVR-4383-0016, the pipeline identified 37 supporting reads but on IGV the number is higher (figure 3.4). The two breaks for the balanced translocation occurred in *IGHM*switch and *IGHA1*switch regions. This finding suggests that the translocation arose during class switch recombination and that the intervening DNA sequence has been deleted during the process (for further details on translocation mechanisms see for example (Küppers and Dalla-Favera, 2001, Walker et al., 2013)). The 3' enhancer centromeric to the *IGH*-locus is responsible for the overexpression of *BCL3* with the other IGH-enhancer now being located on derivate chromosome 19. A scenario whereby the oncogenic activation of genes on both derivative chromosomes is required has been described for t(4;14) in myeloma ((Walker et al., 2013)), however no target gene has yet been identified for chromosome 19.

The breakpoint for case REF-4383-0003 lies 15kb centromeric to *BCL3* an area that was not included in the design of the panel and no baits capture this region during library preparation and enrichment. The reads that can be seen aligning to chromosome 19 (figure 3.5) are therefore only present due to the partner of the paired-end read being aligned to chromosome 14 and the fragment spanning the translocation breakpoint. In contrast to the RVR-4383-0016 case (figure 3.4) no grey reads indicating complete matching of both paired reads to the same chromosome

are therefore present and no reads are seen for this region in any other sample. Nevertheless, the read depth for this sample is high (>240 reads, with 38 reads across the fusion) indicating that there was a high infiltration with tumour cells in this sample.

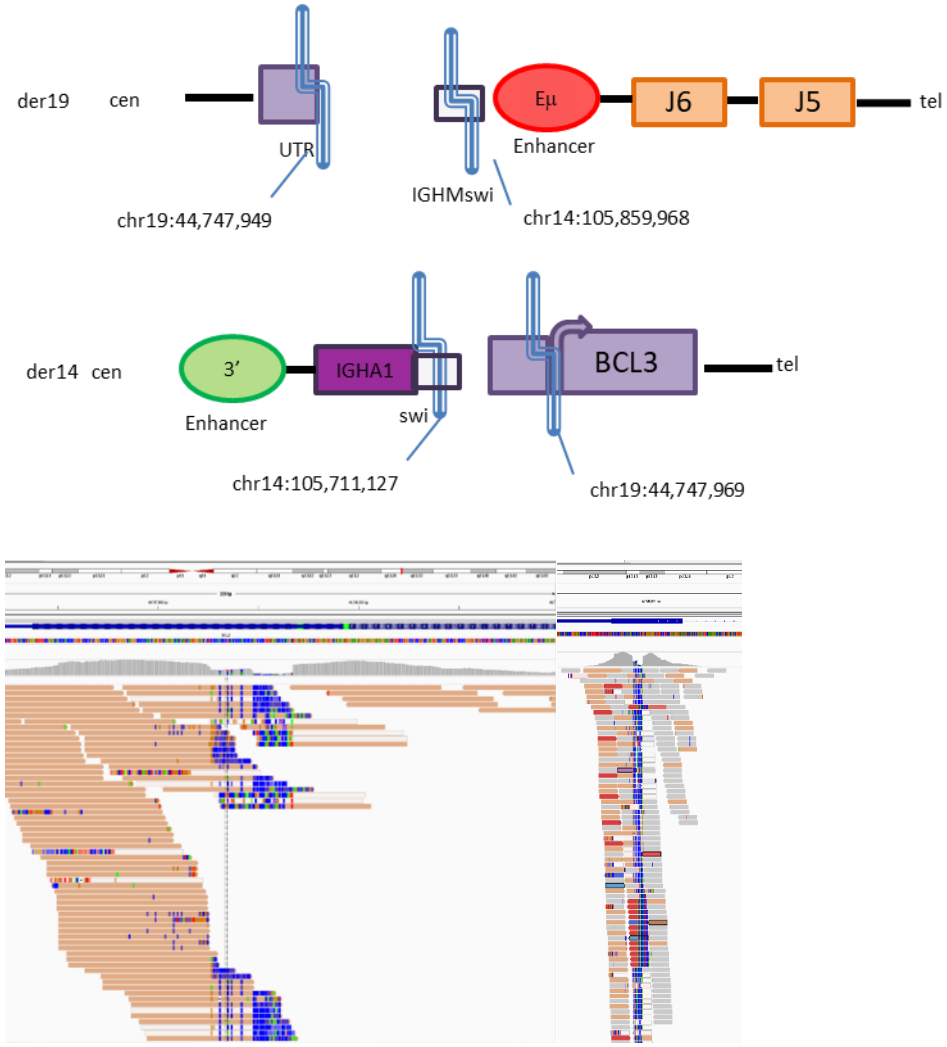


Fig. 3.4: Example of a t(14;19) translocation detected in the study. A Schematic depiction (not to scale) of the location of breakpoints and the derivative chromosomes 19 and 14 resulting from the balanced translocation occurring during CSR events. Case RVR-4383-0016 is shown. B Screenshot from IGV genome viewer showing reads aligning to chromosome 19 in the 5'UTR region of *BCL3*. On the left, reads are grouped by chromosome of mate, on the right the reads are not grouped. The region was included in the design of the studies and baits were included for this region hence there are grey reads present which are not involved in the fusion.

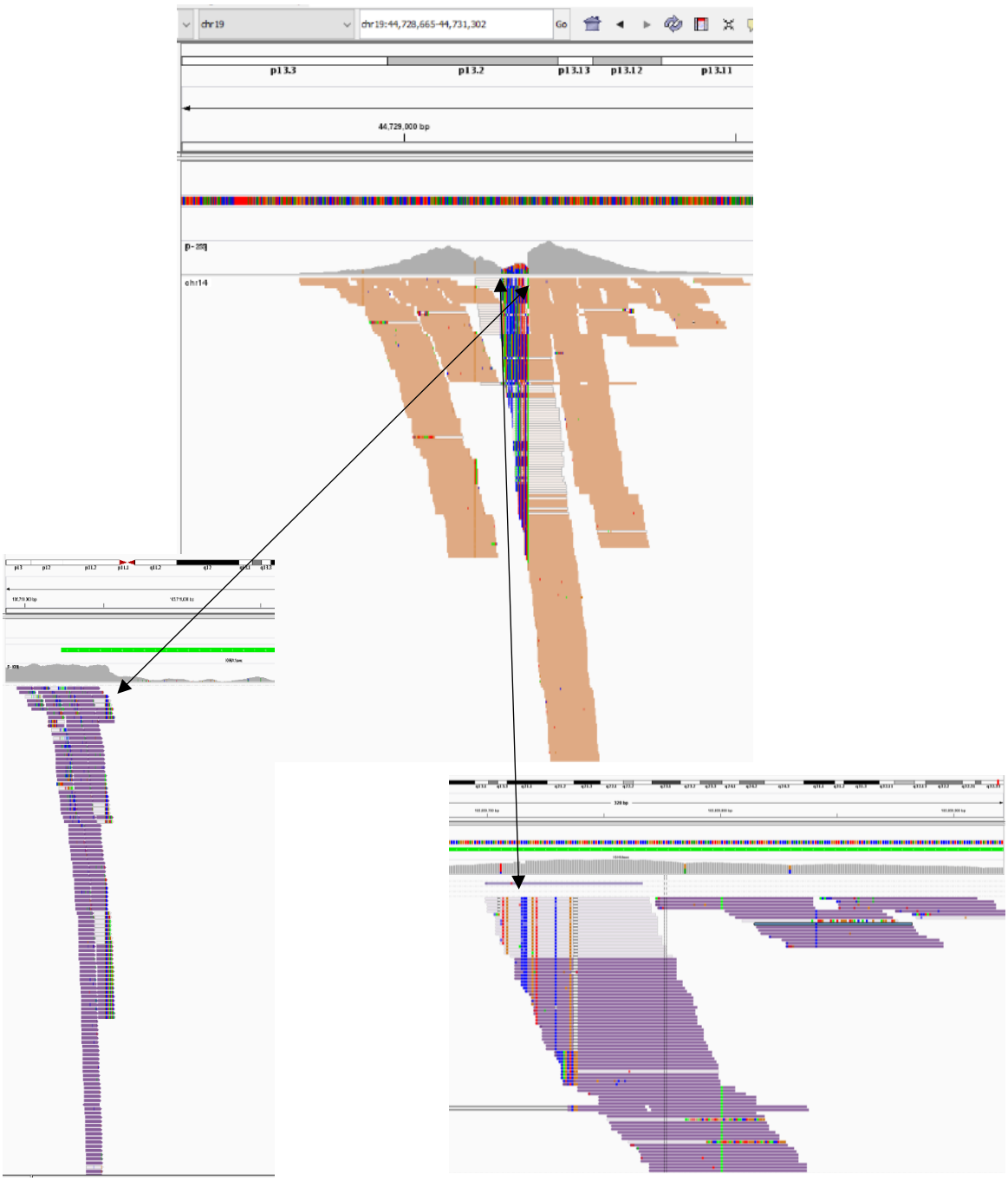


Fig. 3.5: Example of a t(14;19) translocation detected in the study. Case REF-4383-0003 is shown. Screenshots from IGV genome viewer showing reads aligning to chromosome 19 (top) and the two switch regions on chromosome 14 (IGHA to the bottom left and IGHM to the bottom right). The breakpoint region for *BCL3* in this case was not included in the bait design in contrast to the previous example. All reads aligning to this region on chromosome 19 are therefore paired-end reads from sequencing fragments spanning the fusion points. Black arrows show the direction of fusion leading to the IGHA reads going across into chromosome 19 creating the derchr14 where the 3' IGH enhancer controls the expression of the *BCL3* gene. As was shown in more detail for the *BCL2* fusions, the exact breakpoints can be identified and any deletions and insertions of nucleotides at the junctions recorded (details not shown).

Back in 2007, Martin-Subero *et al.* published their characterisation of 58 patients with a *IGH-BCL3*-translocation positive B-NHL (Martín-Subero et al., 2007). Although most of these cases tend to be diagnosed with CLL, their cohort showed a great deal of heterogeneity with only 3 out of 25 cases showing typical CLL morphology and 6 cases having features of MZL. CLL with *BCL3*-translocations can be CD5-positive or negative but tend to be CD23-negative with a high frequency of trisomy 12 as seen in RVR-4383-0016, which is absent in all other B-NHL with *BCL3*-translocations (Huh et al., 2007). In addition, cases more closely resembling CLL carried unmutated *IGHV* rearrangements (0% SHM) and a simply karyotype with all MZL-related cases showing some degree of SHM and a higher degree of additional abnormalities (Martín-Subero et al., 2007). Both cases found in this cohort have *IGHV*-rearrangements with complete homology to the germline (unmutated) and only one additional SNV change was found in RVR-4383-0016. Based on these details the diagnosis is likely that of (atypical) CLL with t(14;19), however further review of the morphology and immunophenotyping data is required.

A retrospective review published by the Mayo Clinic identified 19 cases of CLL with t(14;19) over 14 years. These patients had a higher 5-year probability of requiring treatment and a reduced 5-year survival compared to patients without the translocation (Fang et al., 2019). The authors therefore suggest that testing for the presence of *BCL3*-translocations should be incorporated into routine testing for patients with CLL.

3.6.3 Chromosomal rearrangements involving *IGH* and *CCND1*

Translocation involving the *IGH*-locus and *CCND1* on chromosome 11 are characteristic of Mantle Cell lymphoma (MCL) and are also observed with high frequency in myeloma but occur very rarely in other mature B-cell malignancies (Swerdlow et al., 2016a). In our cohort, 6 samples showed evidence of a rearrangement between chromosome 14 and chromosome 11 (table 3.2). Three cases had breakpoints at D- and J-gene segments suggesting that the translocation happened during D-J recombination processes. Two of these cases show balanced translocations where the fusion on both derivative chromosomes could be found. In one case the fusion appears to be unbalanced with only derivative chromosome 14 being identified. The fusion constructs for the above three cases are shown as

graphics in figure 3.6. Only in one case out of the 6 cases was the breakpoint on chromosome 11 located in the major translocation cluster (MTC). FISH confirmed the presence of a t(11;14) translocation in the case with unbalanced translocation and two of the balanced ones with two yet to be analysed. For these patients the diagnosis will likely change to Mantle Cell Lymphoma ICD-O 9673/3 upon review. Only one case, RAX-4383-0003 had unmutated *IGHV*, which is associated with nodal MCL and more aggressive disease (Silkenstedt et al., 2021). All other cases had either borderline (2 cases) or mutated *IGHV* (3 cases). Presence of missense variants in *TP53* is associated with poor prognosis in MCL (cases RHU-4383-0006 and REF-4383-0013) (Silkenstedt et al., 2021) even in cases with the more indolent non-nodal, leukaemia variant (Sakhdari et al., 2019).

Case RHU-4383-0006 showed an unusual translocation construct. The breaks on chromosome 11 occurred centromeric and telomeric of *CCND1* and the fragment containing *CCND1* was inserted into the chromosome 14 between the J5- and J4-gene (figure 3.7). The second functional V(D)J-rearrangement on the second allele is also shown in the figure for this sample. The mechanism underlying this event has not been further investigated but similar constructs were recently reported in the context of MCL with t(11;14) (Fuster et al., 2020, Miao et al., 2016, Martín-Garcia et al., 2019, Peterson et al., 2019). FISH analysis was requested but no result was received so far.

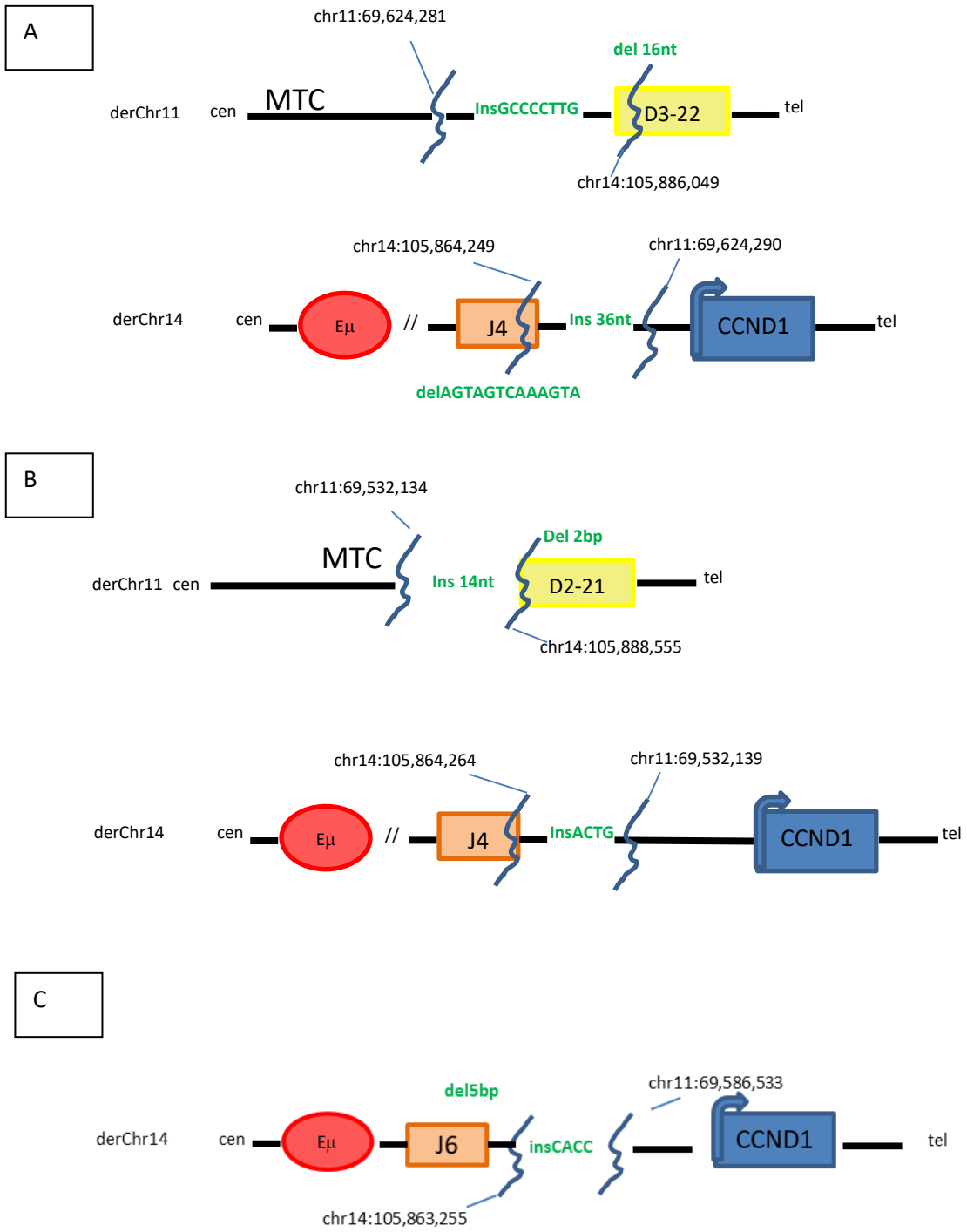


Fig. 3.6: Schematic representation of three of the t(11:14) IGH-CCND1 translocations detected in the study. Schematic representation (not to scale) of the translocation breakpoints and partners (and the resulting derivative chromosomes). Cases shown in A + B show a balanced translocation whereas in case shown in panel C only the derivative chromosome 14 was detected. The double-strand DNA break on chromosome 14 occurred during D-J recombination involving different J- and D-genes. The break on chromosome 11 was located in the well-characterised MTC locus in case B. As a result of the fusion the IGH super enhancer Eμ influences the expression of *CCND1* now located downstream of the enhancer on the derivative chromosome 14, leading to CCND1 overexpression. Deletions and insertion of nucleotides compared to the germline sequence occurs at the fusion points (details in green). The super enhancer Eμ is depicted in red, the J-genes are shown as orange and the D-genes as yellow boxes. *CCND1* is shown in blue with an arrow indicating the direction of transcription. Black lines depict the DNA sequence and blue squiggles the double strand breaks. Cen- = centromeric; tel = telomeric; der = derivative

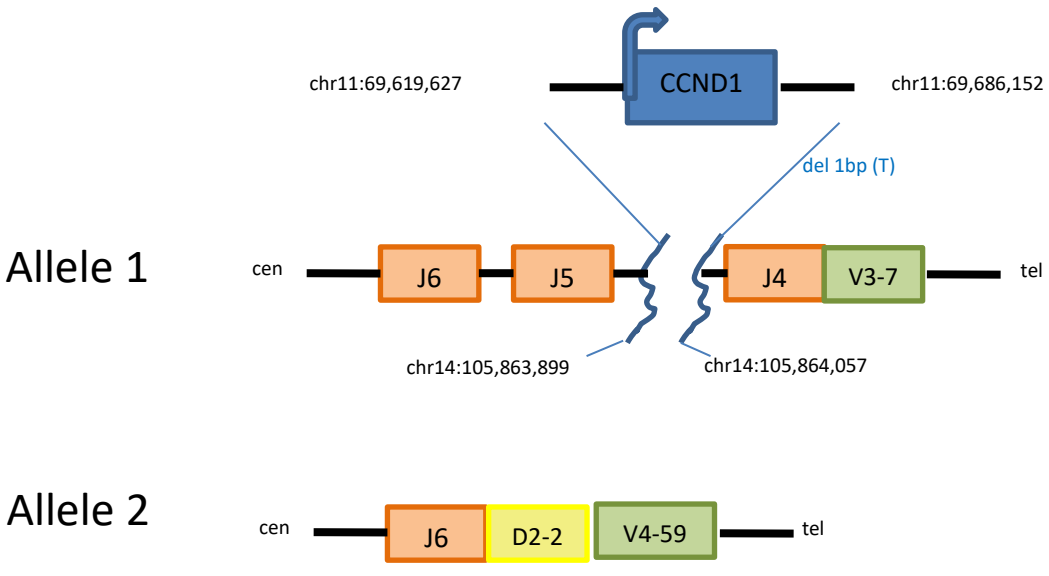
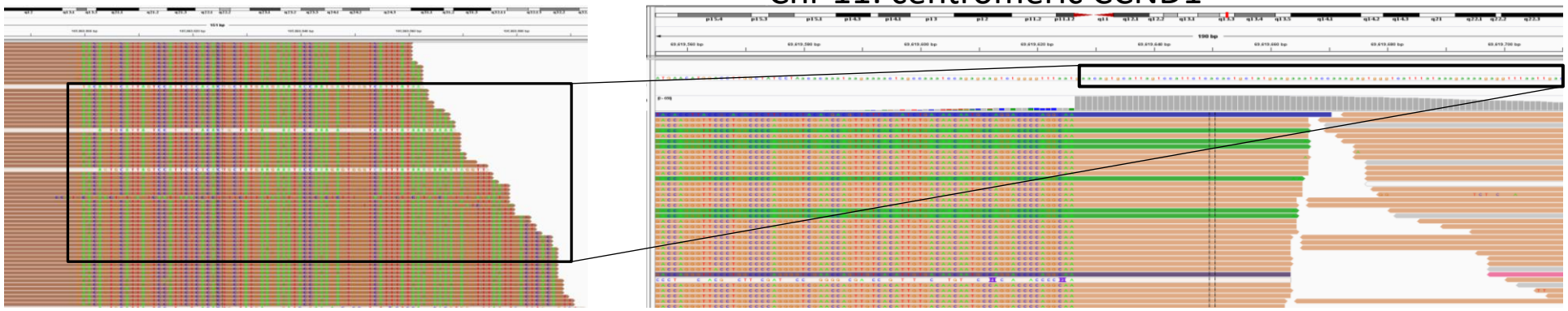


Fig. 3.7: Description of the unusual t(14;19) translocation detected in the study. Case RHU-4383-0006 is shown. Schematic representation of the two alleles present in the B cells of this case (not to scale). *CCND1* from chromosome 11 has been inserted between J5 and J4 on chromosome 14. The second allele shows a complete, productive V(D)J-rearrangement. Centromere and telomere location are indicated for orientation and the double strand DNA breaks are indicated by blue lines.

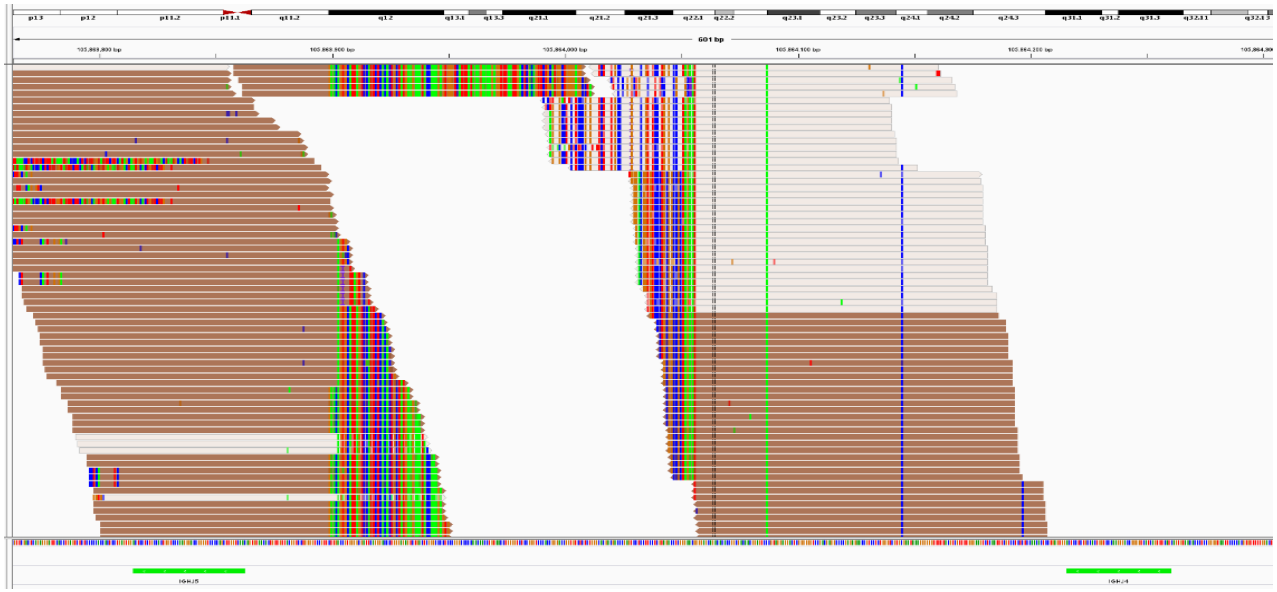
Fig 3.8 (next page): Screenshots from IGV genome viewer showing the split reads going across the junction for the case shown in figure 3.7. The panel at the bottom provides a zoomed-out view to show the two breaks that have occurred on chromosome 14 between the J5- and the J4-gene. The top panel shows reads aligned to chromosome 14 on the left and chromosome 11 on the right. Reads were coloured by chromosome of mate: Brown indicates that the second read in the pair is located on chromosome 11 instead of chromosome 14 and vice versa for the orange reads aligning to chromosome 11 in the top right panel. The nucleotides that are highlighted as mismatches in the reads show that these parts of the sequencing reads align to the germline sequence of chromosome 11 (black boxes) and *IGHJ5* (arrow with germline sequence underneath), respectively. Second break on chromosome 14 and break telomeric of *CCND1* on chromosome 11 not shown.

Chr 14: IGHJ5

Chr 11: centromeric CCND1



IGHJ5



For two cases, REF-4383-0013 and RMH-4383-0005 the complete translocation construct could not be identified. REF-4383-0013 was confirmed to be positive by FISH for the t(11;14). However, the NGS data provides the results depicted in figure 3.9. Here, two breaks occur centromeric to *CCND1* with one reading into *D3-22* on chromosome 14 but the other being fused to the centromeric portion of chromosome 14, close to *IGHJ5*. There are 719 reads supporting this call with an additional 69 confirming the second break and fusion to *D3-22*. From the NGS data it is unclear what happened to *CCND1* which is telomeric in relation to both breaks and does not appear to be involved in the fusion constructs (figure 3.9). Likely, there are further reads that have bene discarded by the pipeline or they were not incorporated due to odd locations of breakpoints not covered in the design of the panel. Long-read sequencing could help elucidate this further.

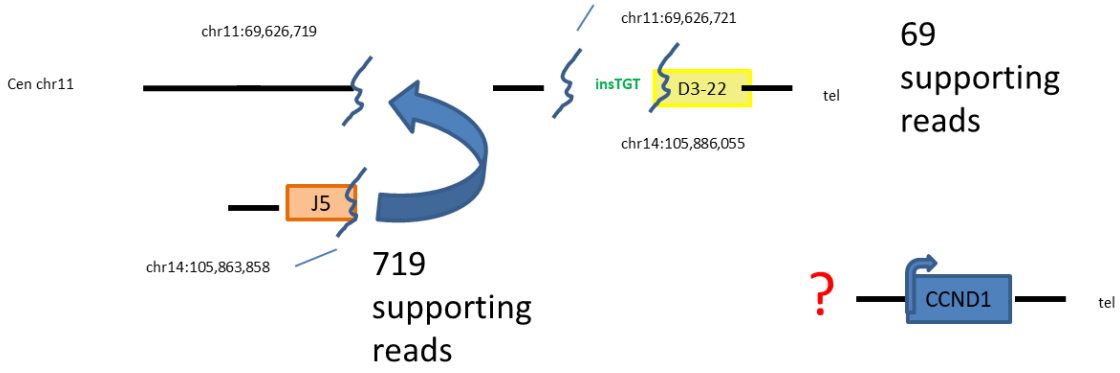


Fig. 3.9: Schematic representation of the rearrangement based on the NGS data for case REF-4383-0013. Two breaks can be delineated on chromosome 14 at *IGHJ5* and *D3-22* both with sequencing reads across the break intro chromosome 11. However, no reads supporting the reciprocal fusion incorporating *CCND1* were present. cen= centromeric, tel= telomeric. Black lines depict DNA and blue squiggles the dsDNA break points. The *D3-22*-gene is shown as a yellow box, the *J5*-gene in orange (not to scale).

The second case, RMH-4383-0005 also has a construct where the fusion of *CCND1* to chromosome 14 cannot be demonstrated based on the reads obtained from NGS testing. Here, it appears that the fusion was initiated during the D-J rearrangement process and reads align to *D2-2* that either break into *J6* or – at the same nucleotide position- break and fuse to chromosome 11 (figure 3.10). An additional complete V(D)J-rearrangement is also detected which raises the possibility of a subclonal translocation. FISH was negative for this case using the IGH/CCND1 XT, Vysis dual fusion (312 cells scored, analysis performed by the Cytogenetics laboratory, Clinical

Genomics, The Royal Marsden NHS Foundation Trust). Repeat analysis and further studies utilising long-read sequencing or whole-genome sequencing will be required to clarify the rearrangement present in this case.

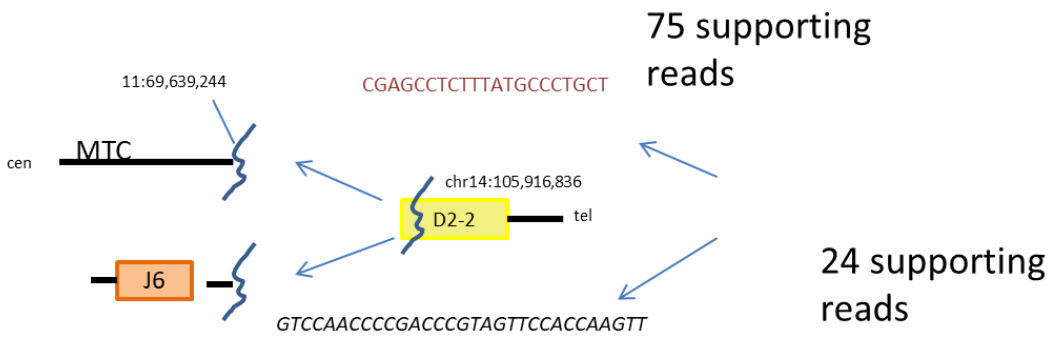


Fig. 3.10: Representation of the reads aligned to *D2-2* for case RMH-4383-0005. Supporting reads show the fusion to J6-gene as well as chromosome 11 arising from the same *D2-2* locus. MTC= major translocation locus on chromosome 11; cen= centromeric, tel= telomeric. Black lines depict DNA and blue squiggles the dsDNA break points. The D2-2 genes is shown as a yellow box, the J6-gene in orange (not to scale). The two different DNA sequences D2-2 joins to as found by NGS are provided,

Table 3.3 (next page): Overview of molecular aberrations found in the cohort of samples including translocations, SNVs and results from preliminary CNV analysis. Only variants deemed pathogenic/likely pathogenic are listed, VUS are not included. Presence of trisomy12 and deletion 17p (del17p) is listed but the CNV analysis is not yet complete and further CNVs may be present in all samples. For reference sequences please refer to table A8 in the appendix. As detailed in the text, 7 samples were removed from the analysis: this table shows results for 101 out of the 108 cases enrolled and run by NGS to date (the case with SNV calls but potential low tumour infiltration is included here). Of note, RVR-4383-0009 had suboptimal mean coverage <400x and may need to be classed as fail on NGS although the clonal *IGHV*-rearrangement was detected (with low read support) and confirmed by Sanger sequencing. Repeat analysis will be attempted.

ID	Translocation	Singel nucleotide variants (SNVs) and indels	Variant allele frequency (VAF) %	CNV
REF-4383-0007	t(14;18) IGH-BCL2	ARID1Ac.553C>Tp.(Gln185Ter)	20	Not yet review ed
		CDKN2Ac.387C>Ap.(Tyr129Ter)	6	
		KMT2Dc.15289C>Tp.(Arg5097Ter)	19	
		KMT2Dc.10437_10440+2delp.?	15	
		EP300c.4399T>Ap.(Tyr1467Asn)	23	
RMH-4383-0011	t(14;18) IGH-BCL2	none	0	Trisomie 12
RVR-4383-0002	t(11;14) CCND1-IGH	none	0	Not yet review ed
RMH-4383-0005	t(11;14) CCND1-IGH	none	NA	CCND1 amplification
RHU-4383-0006	t(11;14) CCND1-IGH	TP53c.818G>Ap.(Arg273His)	32	17pdel, CCND1 amplification
		TP53c.476C>Tp.(Ala159Val)	30	
REF-4383-0013	t(11;14) CCND1-IGH	TP53c.707A>Gp.(Tyr236Cys)	84	17pdel
RDU-4383-0005	t(11;14) CCND1-IGH	none	NA	Not yet review ed
RAX-4383-0003	t(11;14) CCND1-IGH	none	NA	Not yet review ed
RVR-4383-0016	t(14;19) IGH-BCL3	none	NA	Trisomie 12
REF-4383-0003	t(14;19) IGH-BCL3	none	NA	Not yet review ed
BEL-4383-0008	t(7;14) CDK6-TRAD	TP53c.420delp.(Cys141AlafsTer29)	46	17pdel
RVR-4383-0013	t(11;14) RAD51-BIRC3	KMT2Dc.15256C>Tp.(Arg5086Ter)	57	Trisomie 12
		FBXW7c.1393C>Tp.(Arg465Cys)	32	
REF-4383-0009	t(7;14) ??	none	NA	Trisomie12
RMH-4383-0004	t(7;14) ??	ARID1Ac.60_62delp.(Pro21del)	7	
		MYD88c.896A>Gp.Asn299Ser	20	
		MYD88c.904A>Cp.Thr302Pro	22	
REF-4383-0014	No	MYD88c.818T>Cp.(Leu273Pro)	34	Not yet review ed
RHU-4383-0003	No	MYD88c.818T>Cp.(Leu273Pro)	27	None
RA2-4383-0002	No	MYD88c.818T>Cp.(Leu273Pro)	39	Not yet review ed
RJ6-4383-0001	No	MYD88c.818T>Cp.(Leu273Pro)	46	Not yet review ed
RVR-4383-0011	No	MYD88c.818T>Cp.(Leu273Pro)	33	Not yet review ed
RVR-4383-0027	No	MYD88c.818T>Cp.(Leu273Pro)	28	Not yet review ed
RTP-4383-0002	No	MYD88c.818T>Cp.(Leu273Pro)	20	Not yet review ed
RDD-4383-0003	No	MYD88c.818T>Cp.(Leu273Pro)	27	Not yet review ed
REF-4383-0002	No	MYD88c.818T>Cp.(Leu273Pro)	40	Not yet review ed
RMH-4383-0013	No	MYD88c.818T>Cp.(Leu273Pro)	32	Not yet review ed
		CD79Bc.590A>Gp.(Tyr197Cys)	64	
RJ6-4383-0002	No	MYD88c.818T>Cp.(Leu273Pro)	23	Not yet review ed
		CBLc.1259G>Ap.(Arg420Gln)	25	
		PAX5c.540dupp.(Ile181HisfsTer62)	21	
RMH-4383-0001	No	MYD88c.818T>Cp.Leu273Pro	38	Not yet review ed
		FAT1c.9075G>Tp.Lys3025Asn	39	
		CD79Bc.599T>Cp.Leu200Pro	43	
RVR-4383-0005	No	MYD88c.818T>Cp.(Leu273Pro)	28	Not yet review ed
		NFKBIEc.681delp.(Glu228ArgfsTer17)	24	
RA2-4383-0004	No	TP53c.114dupp.(Ala39SerfsTer4)	22	17pdel
BEL-4383-0001	No	KLF2 c.76-1G>A p.?	29	del7q (POT1, BRAF, EZH2)
		TP53c.707A>Gp.(Tyr236Cys)	8	
BEL-4383-0002	No	KRASc.38G>Ap.(Gly13Asp)	22	Trisomy 12
BEL-4383-0005	No	CCND3c.850C>Tp.(Pro284Ser)	23	Not yet review ed
BEL-4383-0006	No	NOTCH2c.7165C>Tp.(Gln2389Ter)	38	Not yet review ed
		NFKBIEc.762C>Gp.(Tyr254Ter)	25	
		TRAF2c.189-2A>G p.?	7	
BEL-4383-0007	No	TP53c.797G>Ap.(Gly266Glu)	32	17pdel
R1K-4383-0004	No	CXCR4c.1018G>Tp.(Gly340Ter)	29	Not yet review ed
R1K-4383-0007	No	BRAFc.1781A>Gp.(Asp594Gly)	9	Trisomie12, partial del7q (POT1)
		ARID1Ac.154_155insTCTTTCGp.(Ala52ValfsTer61)	28	
		BRAFc.1395_1400delp.(Gly466_Ser467del)	11	

ID	Translocation	Singel nucleotide variants (SNVs) and indels	Variant allele frequency (VAF) %	CNV
RA2-4383-0001	No	CARD11c.1070A>Tp.(Asp357Val)	32	Not yet reviewed
RA2-4383-0005	No	TP53c.715_720delp.(Asn239_Ser240del)	20	17pdel
		CREBBPc.445C>Tp.(Gln149Ter)	23	
RAX-4383-0006	No	CD79Bc.574_577delp.(Glu192LysfsTer19)	21	Not yet reviewed
RAX-4383-0007	No	KMT2Dc.13450C>Tp.(Arg4484Ter)	22	Trisomie12
RDD-4383-0001	No	CCND3c.838C>Tp.(Gln280Ter)	38	del7q (POT1, BRAF, EZH2)
		BRAFc.1405_1406delGGinsTCp.(Gly469Ser)	76	
RDD-4383-0002	No	CREBBPc.6895C>Tp.(Gln2299Ter)	26	Not yet reviewed
RDU-4383-0004	No	BIRC3c.1284_1288delp.(Glu429GlyfsTer7)	10	Not yet reviewed
RDU-4383-0001	No	TP53c.818G>Tp.(Arg273Leu)	72	17pdel, 17q amp
		CCND3c.860T>Gp.(Val287Gly)	37	
		CREBBPc.4574A>Cp.(Gln1525Pro)	38	
REF-4383-0005	No	NRASc.181C>Ap.(Gln61Lys)	21	Trisomie12
		NRASc.182A>Tp.(Gln61Leu)	8	
		NOTCH1c.7541_7542delp.(Pro2514ArgfsTer4)	29	
REF-4383-0006	No	CARD11c.734T>Cp.(Leu245Pro)	30	Not yet reviewed
REF-4383-0008	No	CCND3c.838C>Tp.(Gln280Ter)	38	Not yet reviewed
REF-4383-0010	No	TRAF3c.1032_1039delp.(Trp344CysfsTer43)	9	Not yet reviewed
		TRAF3c.1463G>Ap.(Trp488Ter)	16	
REF-4383-0012	No	ARID1Ac.2732+1G>Cp.?	32	Trisomie12
		CXCR4c.1000C>Tp.(Arg334Ter)	7	
		BIRC3c.1639delp.(Gln547AsnfsTer21)	16	
RHU-4383-0001	No	BIRC3c.1139delp.(Pro380LeufsTer2)	6	Trisomie12
		NOTCH2c.7198C>Tp.(Arg2400Ter)	28	
		NOTCH2c.6973C>Tp.(Gln2325Ter)	8	
REF-4383-0016	No	TP53c.919+2T>Ap.?	34	17pdel
		TP53c.159G>Ap.(Trp53Ter)	5	
RHU-4383-0005	No	ARID1Ac.5965C>Tp.(Arg1989Ter)	43	None
		CCND3c.852_853insGTACTp.(Thr285ValfsTer21)	11	
		CCND3c.811dupp.(Arg271ProfsTer53)	27	
		CREBBPc.3098_3099delp.(Lys1033ArgfsTer17)	35	
RMH-4383-0002	No	TNFAIP3c.986+1G>C	6	Not yet reviewed
		NOTCH1c.7105C>Ap.Pro2369Thr	23	
		NOTCH1c.6963_6964delCCinsTTp.Gln2322Ter	24	
		MAP3K14c.199G>Ap.Ala67Thr	10	
		KLF2c.883_892+7delp.?	15	
RMH-4383-0003	No	BIRC3c.1285dupp.Glu429GlyfsTer9	8	potentially low tumour inifltraon
RMH-4383-0007	No	CCND3c.872_873delp.His291ProfsTer32	12	17pdel
RMH-4383-0008	No	TP53c.584T>Cp.Ile195Thr	33	Not yet reviewed
RMH-4383-0009	No	CCND3c.852_854delp.Thr285del	16	Not yet reviewed
		BIRC3c.1681delp.Met561TrpfsTer7	20	
RMH-4383-0012	No	TP53c.838A>Tp.(Arg280Ter)	82	17pdel
		CCND3c.857A>Gp.(Asp286Gly)	38	
RTP-4383-0001	No	KLF2c.862C>Tp.(His288Tyr)	9	Not yet reviewed
RMH-4383-0015	No	TP53c.722C>Tp.(Ser241Phe)	88	17pdel, 17q amp
RTP-4383-0005	No	CREBBPc.2415delp.(Met806Ter)	37	Not yet reviewed
RTP-4383-0006	No	BRAFc.1406G>Cp.(Gly469Ala)	34	Not yet reviewed
RTP-4383-0007	No	TP53c.725G>Tp.(Cys242Phe)	10	17pdel, 17q amp, del7q (POT1)
		TP53c.621_622dupp.(Asp208ValfsTer40)	45	
		CCND3c.843_871delp.(Ser282ProfsTer32)	12	
RTP-4383-0009	No	TP53c.733G>Ap.(Gly245Ser)	84	17pdel
		SF3B1c.1866G>Cp.(Glu622Asp)	45	

ID	Translocation	Singel nucleotide variants (SNVs) and indels	Variant allele frequency (VAF) %	CNV
RVR-4383-0004	No	BIRC3c.1285dupp.(Glu429GlyfsTer9)	NA	Not yet review ed
RVR-4383-0014	No	ARID1Ac.5161C>Tp.(Arg1721Ter)	21	Not yet review ed
		KMT2Dc.15874G>Tp.(Glu5292Ter)	22	
		KMT2Dc.11566C>Tp.(Gln3856Ter)	25	
		CREBBPc.4337G>Ap.(Arg1446His)	46	
RVR-4383-0015	No	NOTCH2c.6851_6867delp.(Pro2284ArgfsTer23)	33	Not yet review ed
		KMT2Dc.11814_11837delp.(Gln3947_Gln3954del)	6	
		TRAF3c.328delp.(Ile110PhefsTer19)	27	
		TRAF3c.854dupp.(Asn285LysfsTer13)	7	
RVR-4383-0018	No	ARID1Ac.6147G>Ap.(Trp2049Ter)	11	Not yet review ed
		TNFAIP3c.1226_1229delp.(Glu409GlyfsTer12)	13	
RVR-4383-0020	No	TP53c.722C>Gp.(Ser241Cys)	43	Not yet review ed
RVR-4383-0022	No	NOTCH2c.6909dupp.(Ile2304HisfsTer9)	29	Not yet review ed
		TNFAIP3c.1247_1251delp.(Asn416ThrfsTer11)	40	
RVR-4383-0023	No	CARD11c.352A>Gp.(Ile118Val)	11	Trisomie 12
RVR-4383-0025	No	TRAF3c.423C>Ap.(Cys141Ter)	18	Not yet review ed
RVR-4383-0006	No	NFKBIEc.759_762delp.(Tyr254SerfsTer13)	22	Not yet review ed
RJ6-4383-0003	No	KLF2c.872C>Tp.(Ala291Val)	45	Trisomie12, partial del7q (POT1)
RVR-4383-0012	No	TP53c.626_627delp.(Arg209LysfsTer6)	72	17pdel
RVR-4383-0026	No	none	NA	Trisomie 12
RHU-4383-0004	No	none	NA	Trisomie12
RDU-4383-0003	No	none	NA	Trisomie12
RMH-4383-0014	No	none	NA	Trisomie 12
RAX-4383-0004	No	none	NA	Tris12, Amplification BCL2 (partially anlaysed)
RMH-4383-0010	No	none	NA	17pdel
RHU-4383-0002	No	none	NA	BCL2 amplification
RA2-4383-0006	No	none	NA	BCL2 amplification, partial deletion chr13 (RB1), possible tris3 (amplification MYD88, PIK3CA)
BEL-4383-0003	No	none	NA	Not yet review ed
RVR-4383-0021	No	none	NA	Not yet review ed
RVR-4383-0019	No	none	NA	Not yet review ed
RVR-4383-0017	No	none	NA	Not yet review ed
RVR-4383-0007	No	none	NA	Not yet review ed
RVR-4383-0008	No	none	NA	Not yet review ed
RVR-4383-0009	No	none	NA	Not yet review ed
RVR-4383-0010	No	none	NA	Not yet review ed
RVR-4383-0001	No	none	NA	Not yet review ed
RTP-4383-0003	No	none	NA	Not yet review ed
RTP-4383-0004	No	none	NA	Not yet review ed
RMH-4383-0006	No	none	NA	Not yet review ed
RDU-4383-0002	No	none	NA	Not yet review ed
RAX-4383-0005	No	none	NA	Not yet review ed
RDD-4383-0004	No	none	NA	Not yet review ed
REF-4383-0011	No	none	NA	Not yet review ed
R1K-4383-0001	No	none	NA	Not yet review ed
R1K-4383-0002	No	none	NA	Not yet review ed
RAX-4383-0002	No	none	NA	Not yet review ed

3.7 Detection of Single-nucleotide variants (SNVs)

Seventy-two genes deemed to be of clinico-pathological interest were collated by the experts of EuroClonality consortium for inclusion in the panel in 2014 (list and region of interest provided in the appendix table A1)(Wren et al., 2014)). The design aimed to include genes with known relevance for diagnosis, prognosis or therapeutic options in the context of mature B- and T- NHLs as well as B-ALL and T-ALL. Table A1 in the appendix provides an overview of the genes chosen and the relevance assigned to them at the time. As only clonal B-lymphocytosis cases were eligible for the ENABLE study, not all genes are relevant or expected to be mutated in these cases and given that the design is >7 years old, additional candidates like for example *MEF2B* and *FOXO1*, which have now been shown to be recurrently mutated in B-NHL and carry prognostic information in FL, were not yet sufficiently characterized to be included in this panel at the time (Chung, 2020). Similarly, receptor-type protein tyrosine phosphatase delta (*PTPRD*) had not yet been described as a recurrent finding in nodal marginal zone lymphoma (NMZL) (Bertoni et al., 2018).

Coverage analysis performed as part of the validation studies showed that all regions of interest (ROI) performed well with the target mean coverage depth set to 500x (Wren et al., 2017, Stewart et al., 2019). Using Horizon blends the limit of detection was determined to be 4% (Stewart P. et al Blood Advances *in press*).

Out of 108 cases enrolled, 101 samples were included in SNV analysis. Seven samples were excluded as explained in more detail in chapter 3.1 above. Case RMH-4383-0003 with low-VAF SNVs has been included in this part of the analysis.

Driver variants were detected in 60 out of the 101 samples (59%, table 3.3). Thirty-seven samples had a single variant, 14 carried two, 10 cases three and in 5 cases four or more SNVs were detected. The majority of variants were missense (50 cases), followed by frameshifts seen in 28 cases, 26 nonsense variants, 7 potential splice variants in 6 in-frame insertion/deletion events.

Variants with VAF over 5% were included. Twenty-four variants had VAF between 5-10% and nine variants were found to have a VAF above 50% which indicates a loss of the second allele or copy-neutral loss of heterozygosity (CN-LOH). Most of

the variants with VAF above 50% were found in *TP53* and CNV analysis showed evidence of del(17p) in all cases (see below).

In 8 of the 9 cases with variants <10% VAF, additional driver variants with higher VAFs were present suggesting that clonal evolution with acquisition of additional variants had occurred. In RMH-4383-0003, the only pathogenic variant seen was in *BIRC3* with a VAF of 8%. As discuss previously, no clonal IG-rearrangements were detected for this case and it is likely that this case has a low tumour percentage which would explain the low VAF observed.

Additional SNVs were detected in 4 cases with a chromosomal translocation: Two of the *IGH-CCND1*-positive samples also had *TP53* variants and REF-4383-0007 with an *IGH-BCL2* rearrangement carried multiple additional changes (see table 3.3). BEL-4383-0008 with *CDK6-TRAD* also showed mutated *TP53* and this case will be presented in more detail as a case study 2 below.

In 35 cases no SNVs classed as pathogenic or likely pathogenic/clinical actionable were detected. In 5 cases SNVs were detected but classed as variants of unknown significance and these are not further discussed here.

In eight of these samples CNV analysis (see below) indicated the presence of trisomy 12 (5 cases), deletion of the short arm of chromosome 17 (del17p, 1 case) and amplification of *BCL2* on chromosome 18 in two cases.

Full CNV analysis has yet to be performed for all cases and in the 19 cases without pathogenic SNVs or translocations further CNVs may be discovered.

A total of 117 SNVs were detected across the 72 genes analysed. Variants in three genes, *TP53, MYD88* and *CCND3* accounted for 37% of all SNVs in the study cohort (figure 3.11). The 18 variants in *TP53* were found across 16 patients (2 patients had two variants each, table 3.4); the 10 variants in *CCND3* were seen in 9 patients (1 patient with two variants in *CCND3*, table 3.5) and the 15 variants in *MYD88* were seen in 14 patients (chapter 3.7.3.). For the one case with two concomitant *MYD88* variants, both were not the common hotspot driver variant p.(Leu265Pro).

Figure 3.12 provides a visual representation of the frequency of aberrations in each of the pathways found in this study cohort. Combining the number of variants found (figure 3.11) according to the pathways they impact shows that, as expected for B-

NHL, variants were predominantly found in genes associated with BCR-signalling processes and NF-κB activation (figures 3.11 and 3.12, chapter 1.3). Each of these pathway and variant groups will be discussed in more detail below: Given the high frequency of variants in the respective genes, cases with *TP53, CCND3* and *MYD88* variants are discussed in individual chapters (3.71, 3.72 and 3.73, respectively) The discussion also includes additional variants detectedalongside these driver variants and for *MYD88*, association with particular *IGHV*-gene usage is shown in detail. Variants in the other genes have been combined into one chapter (3.74) with cases being reviewed individually with regards to additional findings.

No variants were detected in *MYC, TCF3* and *ID3*; genes which are associated with Burkitt Lymphoma (BL); *ATM* and *EZH2*, genes commonly mutated in MCL, FL and DLBCL, respectively were also not found in this study. As would be expected for treatment-naïve patients, no variants associated with treatment-resistance, in genes like *BTK, PLCG2* and *BCL2* were seen. None of the *MYD88*-mutated cases had a variant in *CXCR4*, which occurs in around 30% of LPL/WM cases (Hunter et al., 2014, Kaiser et al., 2021).

Of note, as neutrophils (CD15+ve) were used as the source of germline-DNA, variants in the three clonal haematopoesis of indeterminate potential (CHIP)-associated genes *DNMT3A, TET2* and *ASXL1* may be under-represented here as they commonly occur in the myeloid series of elderly individuals and may therefore be filtered from the output due to their presence in the germline control sample (Karner et al., 2019).

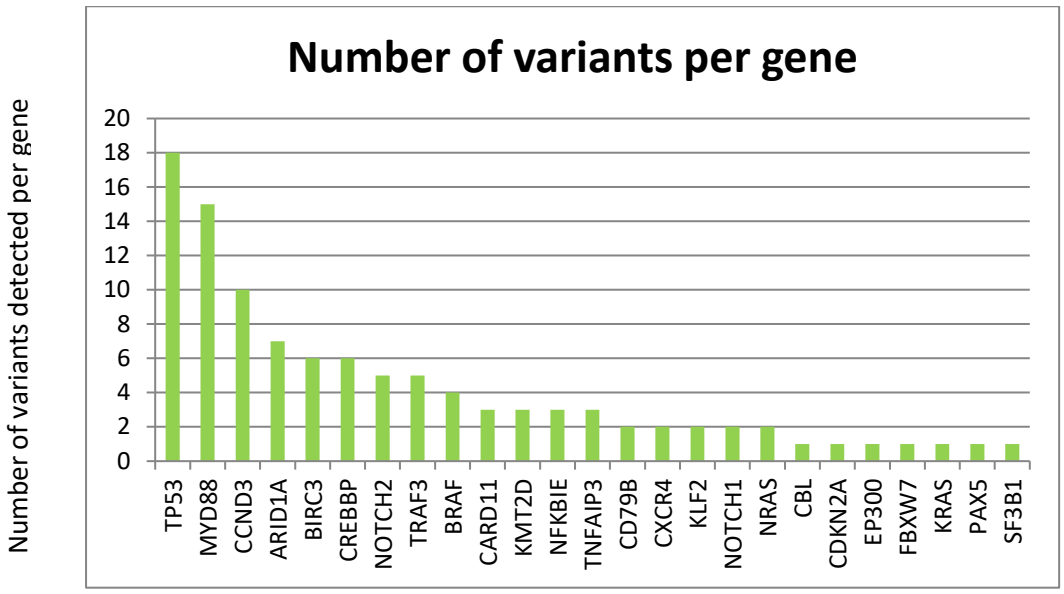


Fig. 3.11: Number of variants found per gene across all samples analysed so far. . Variants were found in 25 different genes and 37% of variants clustered in just three genes: *TP53* (18 variants), *MYD88* (15 variants) and *CCND3* (10 variants).For seven genes, only one case each carried a variant (*CBL, CDKN2A, EP300, FBXW7, KRAS, PAX5* and *SF3B1*).

Fig. 3.12: Proportional representation of the number of variants in each gene belonging to the different cellular pathways. The majority affect BCR- and/or TLR -induced NF-κB activation followed by *TP53* and *CCND3* aberrations affecting cell cycle controls.

3.7.1 Variants in *TP53*

A total of 19 SNVs were detected in*TP53*. The VAF ranged from 8-88% (table 3.4). The high VAF can be explained by the presence of CNV on the second allele (see below). Three of the variants seen fell within the well-known cancer hotspot codons 273 (2 cases) and 245 (one case) (PETITJEAN et al., 2007).

The remaining variants were 8 missense (classed as likely pathogenic based on the ACGS *TP53*-specific guidelines and evidence on the IARC TP53 database (Fortuno et al., 2021)), 4 frameshift variants, 2 nonsense variants and one potential splice variant. The latter three types are all considered pathogenic for *TP53*. For the indel variant seen in case RA2-4383-0005, the in-frame deletion does affect the DNA-binding site which is a crucial protein domain and the area recurrently affected by mutations in cancer and the variant was therefore classed as likely pathogenic (of note the case also has CNV affecting *TP53*)(Malcikova et al., 2018).

The additional missense variants at codons 266, 241 and 195 are all well characterized changes seen across various cancers and in the context of Li-Fraumeni in the germline setting (see International Agency for Research on Cancer (IARC) https://p53.iarc.fr/ and ClinVar database https://www.ncbi.nlm.nih.gov/clinvar/). All lie in the DNA-binding domain, show absence of transactivation activity and cannot suppress cell growth in cell culture experiments (Kato et al., 2003). They are thus all classed as pathogenic.

Table 3.4: Summary of cases with *TP53* SNVs and/or CNVs detected by the capture NGS panel. Additional variants, *IGHV*-gene usage, somatic hypermutation status (SHM) and light chain restriction information are provided. 16 cases had at least one SNV and 15 showed CNV on 17p affecting the *TP53* gene locus. Thirteen of the 16 cases with a SNV also had evidence of a deletion on 17p leading to the loss of the *TP53* gene on the second allele and are assumed to be biallelic *TP53* aberrations.

ID	TP53 variant present?	TP53 variant	VAF (%)	17pdel present?	Additional variants	Gene (VAF %)	IGHV	SHM	Light chain
RA2-4383-0005	Yes	p.(Asn239_Ser240del)	20	Yes	Yes	CREBBP (23)	IGHV3-23	Mutated	Lambda
RDU-4383-0001	Yes	p.(Arg273Leu)	72	Yes	Yes	CCND3 (37), CREBBP (38)	IGHV4-39	NA	Kappa
REF-4383-0013	Yes	p.(Tyr236Cys)	84	Yes	Yes	MCL t(11;14)	IGHV4-59	Mutated	Kappa
RHU-4383-0006	Yes	p.(Ala159Val), p.(Arg273His)	32, 30	Yes	Yes	MCL t(11;14)	IGHV4-59=IGHV4-61	Mutated	Kappa
RMH-4383-0012	Yes	p.(Arg280Ter)	82	Yes	Yes	CCND3 (38)	IGHV3-30=IGHV3-64	Mutated	Kappa
RTP-4383-0007	Yes	p.(Cys242Phe), p.(Asp208ValfsTer40)	10, 45	Yes	Yes	CCND3 (12)	IGHV3-23	Mutated	Kappa
RTP-4383-0009	Yes	p.(Gly245Ser)	84	Yes	Yes	SF3B1 (45)	IGHV3-21	Unmutated	Lambda
BEL-4383-0007	Yes	p.(Gly266Glu)	32	Yes	No	NA	IGHV3-53	Mutated	Kappa
BEL-4383-0008	Yes	p.(Cys141AlafsTer29)	46	Yes	Yes	CDK6-TRA translocaion	IGHV3-23	Mutated	Kappa
RA2-4383-0004	Yes	p.(Ala39SerfsTer4)	22	Yes	No	NA	IGHV3-7	Mutated	Kappa
REF-4383-0016	Yes	p.(Trp53Ter), p.? (c.919+2T>A)	33, 34	Yes	No	NA	IGHV1-8	Unmutated	Kappa
RMH-4383-0015	Yes	p.(Ser241Phe)	88	Yes	No	NA	IGHV4-39	Mutated	Kappa
RVR-4383-0012	Yes	p.(Arg209LysfsTer6)	72	Yes	No	NA	IGHV1-18	Mutated	Kappa
BEL-4383-0001	Yes	p.(Tyr236Cys)	8	No	Yes	KLF2 (29)	IGHV1-2	Unmutated	Kappa
RVR-4383-0020	Yes	p.(Ser241Cys)	42	No	Yes	other CNV	IGHV3-48	Unmutated	Lambda
RMH-4383-0008	Yes	p.(Ile195Thr)	33	No	No	NA	IGHV4-59	Mutated	Kappa
RMH-4383-0007	No	NA	NA	Yes	Yes	CCND3 (12)	IGHV4-34	Mutated	Lambda
RMH-4383-0010	No	NA	NA	Yes	No	NA	IGHV3-23	Borderline	Kappa

Two approaches for the analysis of copy-number variants (CNVs) were employed. NGS data was submitted to the EuroClonality pipeline (hosted by Precision Medicine Centre of Excellence, Centre for Cancer Research & Cell Biology, Queens University Belfast). This pipeline compares the coverage results for a given sample to a control dataset generated from 150 samples without CNVs. It can currently only utilise the tumour data and compares this against the artificial, normal control. In addition, the coverage data obtained for the tumour and germline samples for each case was compared by retrieving the normalised ratio for each target from the Illumina basespace App pipeline analysis and calculating the ratio between the tumour and germline sample (not shown). For the few samples analysis so far, both approaches show very comparable results, with the direct comparison of tumour versus germline sample the optimal approach allowing for sample specific comparison whilst also adjusting for run variation. Examples of the traces obtained are shown in figure 3.13. For case RHU-4383-0006, the data shows that the deletion does not affect the whole arm of 17p as telomeric and centromeric data points show no drop in ratio (panels A and b figure 3.13). All data points for the *TP53* do show a consistent drop indicating the loss of the complete *TP53* gene. The pipeline combines the different target regions assigned to 17p telomeric (several polymorphic SNPs), *TP53* gene locus and 17p centromeric (several polymorphic SNPs) to one datapoint each (panel B in figure 3.13). In the second example (panel C), the drop in coverage occurs across the three regions indicating the loss of the complete short arm of chromosome 17.

Only 3 cases had a *TP53* mutation without accompanying del17p and for BEL-4383-0001 further pathogenic changes were detected (discussed further below) whereas IN RMH-4383-0008 and RVR-4383-0020 a single *TP53* variant was the sole molecular driver variant found.

Two cases have a CNV change on 17p only. In one case, an additional variant in *CCND3* was detected whereas there were no further findings in the second case.

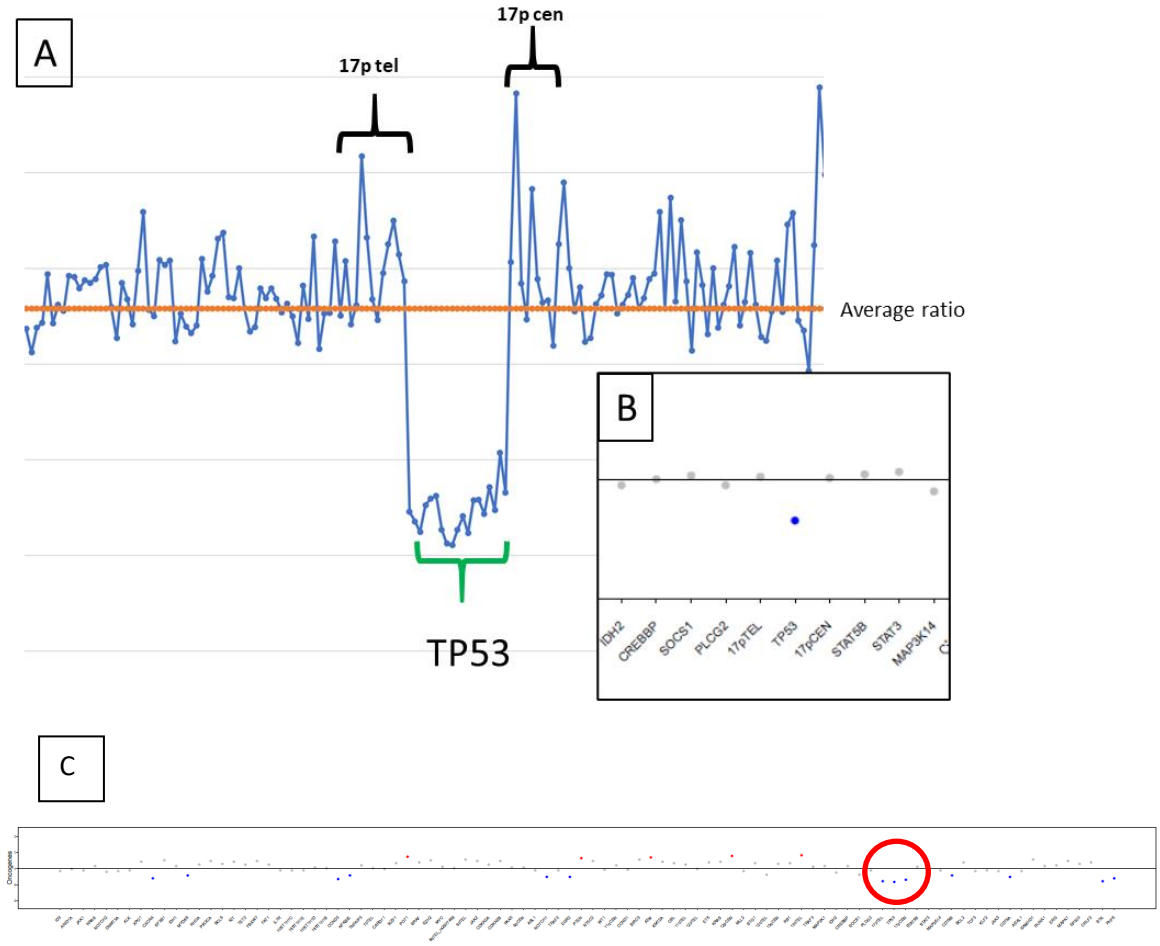


Fig. 3.13: CNV analysis shows the deletion of *TP53* (Panels A+B) and del17p in panel C. The plot in panel A is based on the ration of coverage between the tumour and germline sample at each data point using the mean coverage ratio values. The average ratio for the sample is plotted in orange. The 17p telomeric and centromeric datapoints are highlighted by the black brackets above the graph and show no drop in coverage in this case. In contrast, the complete *TP53* gene (green bracket) shows a drop in coverage. Panel B shows the pipeline report for the same sample. It is concordant to the plot in panel A showing a drop in coverage for *TP53* only. For this analysis the different data points for each target were merged which was not the case in the analysis performed in A. For comparison panel C shows a case with evidence of del17p where the telomeric and centromeric data points show the same drop in coverage as for *TP53*.

The cohort of cases presented here shows a strong correlation between the presence of a SNV and deletion events on 17p affecting the *TP53* gene. Although *TP53* is tumour suppressor, missense variants are the most-commonly found change in the context of haematological cancers and appear to have the most profound effect (Kato et al., 2003, Malcikova et al., 2018, PETITJEAN et al., 2007, Rodrigues et al., 2020). This characteristic has yet to be fully understood but it appears that TP53-proteins carrying missense variants are more stable, accumulate in the cells and can exert a dominant negative on wildtype TP53 in the formation of functional protein complexes. In particular in the context of MCL presence of *TP53*

missense variants is associated with shorter progression-free survival (PFS), a higher risk of relapse and shorter overall survival (OS) compared to patients without *TP53* mutations but this correlation is not seen for del17p in multi-variate analysis (Rodrigues et al., 2020, Eskelund et al., 2017). Akin to CLL, MCL shows poor responses to immunochemotherapy regimen in the presence of *TP53* aberrations (Hill et al., 2020, Roué and Sola, 2020). For cases RHU-4383-0006 and REF-4383-0013 detection of the *IGH-CCND1* fusion as well as the *TP53* mutations is therefore of upmost importance for proper clinical management.

However, for the remainder of the cases in the ENABLE study, long-term follow-up is needed to help us understand the impact of the *TP53* aberrations on clinical course. Notwithstanding the high frequency of *TP53* variants and gene deletions, based on their clinical evaluation and other pathological findings these patients are considered to have low-risk, indolent disease. In CD5+ve MBL and early CLL, *TP53* variants have been shown to be associated with indolent disease if present in *IGHV*-mutated cases (Dagklis et al., 2009, Best et al., 2009, Fazi et al., 2011) which does match the molecular characteristics of this cohort.

3.7.2 Variants in cyclin D3 (*CCND3*)

CCND3 is required for germinal centre (GC) formation and is needed for transition from G1- to S-phase in the cell cycle through its interaction with CDK4 and CDK6 (Singh et al., 2020). Pae *et al.* have recently shown that CCND3 activity is required for continued proliferation of B-cells undergoing affinity maturation in the germinal centres once they have transferred into the dark zone where they continue to proliferative and undergo SHM before returning to the light zone for further selection (Pae et al., 2021). When an activating mutation in *CCND3* was introduced into a mouse model, proliferation of B-cells in the dark zone was increased together with uncontrolled clonal proliferation post GC-reaction (Pae et al., 2021).

Variants in *CCND3* are more commonly seen in aggressive lymphomas e.g. BL and DLBCL but are recurrently detected in small cell B-NHL, especially in those that do not fit into one of the WHO-categories and those previously proposed to be called CBL-MZ (Parker et al., 2018). Curiel-Olmos *et al.* found a distinct pattern of variants in SDRPL of which mutations in the proline, glutamic acid, serine and threonine (PEST) domain of *CCND3* were the most common, seen in 24% of their SDRPL

cases (Curiel-Olmo et al., 2017). They demonstrated increased CCND3 expression in the mutated but also in some of the *CCND3*-wildtype cases of their SDRPL cohort. *IGHV3-43* and *IGHV3-23* were over-represented.

Parker et al. (Parker et al., 2018) published findings of CBL-MZ cases with long-term follow-up which included 3 cases with *CCND3* PEST domain mutations; two also had *TP53* mutations one together with deletion of the short arm of chromosome 7 (del7q) and trisomy 12 (tris12) and the other with an isochromosome 17q; the third case showed trisomy 12 as an additional finding. Importantly, none of the cases had typical features of SDRLP and all had stable disease after 50 months of follow-up.

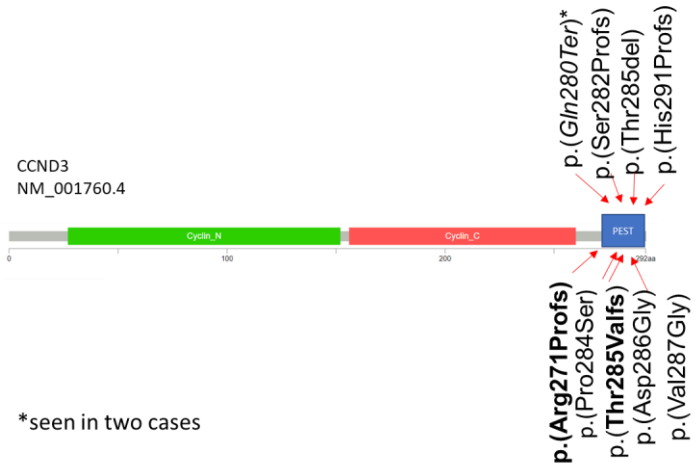


Fig. 3.14: Location of *CCND3* variants detected in this cohort. Schematic of CCND3 protein with location of the PEST domain highlighted in blue. All variants detected fall within the PEST domain and include the phosphorylation motifs. Variants in bold were found in the same case and shown to be biallelic. Nonsense variant at Gln280 was detected in two cases (see table 3.5 for details).

In this cohort, *CCND3* variants were detected in 9 cases (table 3.5 and figure 3.14): Three frameshift variants (two being detected in RHU-4383-0005), two nonsense variants at the same codon 280 and 4 SNVs one of which was an in-frame deletion of Thr285. The SNVs included Pro284 which has been shown to be a crucial part of the phosphorylation motif required for CCND3-protein stability and degradation (Schmitz et al., 2014). Sequencing studies in BL and DLBCL have reported frameshift and nonsense variants which remove the carboxyl terminus including the conserved phosphorylation codons 283, 284 and 290 (Schmitz et al., 2014). All variants characterised so far lead to stabilisation of the mutant isoform of the protein and accumulation in BL cells. In four cases aberrations of *TP53* were detected alongside *CCND3*. Two cases also showed del7q but for two cases the *CCND3* variant was the only change detected so far. The variant seen in RMH-4383-0007

is unusual in that it affects the penultimate codon and is predicted to lead to an elongated protein without loss of any of the phosphorylation sites. Such variants have been reported in BL (Catalogue Of Somatic Mutations In Cancer (COSMIC) and cbioportal databases accessed at https://cancer.sanger.ac.uk/cosmic and https://www.cbioportal.org/, respectively) but no functional evaluation has been reported and their exact pathogenic mechanisms is therefore uncertain.

Limited clinical information was available for four cases (BEL-4383-005, RDD-4383-001, RMH-4383-0007 and RMH-4383-0009, table 3.6). They were all male, aged between 71-83 years and none of them has shown disease progression to date (Case RDD-4383-0001 died of metastatic thyroid cancer). Apart from RDD-4383-0001, all patients were asymptomatic at diagnosis. Bone marrow aspirate review performed at RMH for RDD-4383-001 favoured a diagnosis of CD5+-SMZL. Presence of a *CCND3* variant together with del7q would fit this diagnosis but potentially SDRPL could be in the differential given the high prevalence in the latter and the overlap in features making the distinction between these difficult (Curiel-Olmo et al., 2017). Case RMH-4383-0007 carries a mutated *IGHV4-34* rearrangement, has a *CCND3* PEST domain variant and villous lymphocytes (CD5-ve); all characteristics that together could point to a diagnosis of SDRPL although there is lack of splenomegaly (Swerdlow et al., 2016a).

Table 3.5: Details of the nine cases of the cohort in which *CCND3* variants were present. Details of the variant including VAF is provided together with additional variants and CNVs (preliminary analysis focusing on del17p, del7q and trisomy 12 only)

ID	Singel nucleotide variants (SNVs) and indels	VAF	CNV
BEL-4383-0005	CCND3c.850C>Tp.(Pro284Ser)	23	Not yet reviewed
RDD-4383-0001	CCND3c.838C>Tp.(Gln280Ter)	38	del7q (POT1, BRAF, EZH2)
	BRAFc.1405_1406delGGinsTCp.(Gly469Ser)	76	
RDU-4383-0001	TP53c.818G>Tp.(Arg273Leu)	72	17pdel, 17q amp
	CCND3c.860T>Gp.(Val287Gly)	37	
	CREBBPc.4574A>Cp.(Gln1525Pro)	38	
REF-4383-0008	CCND3c.838C>Tp.(Gln280Ter)	38	Not yet reviewed
RHU-4383-0005	ARID1Ac.5965C>Tp.(Arg1989Ter)	43	None
	CCND3c.852_853insGTACTp.(Thr285ValfsTer21)	11	
	CCND3c.811dupp.(Arg271ProfsTer53)	27	
	CREBBPc.3098_3099delp.(Lys1033ArgfsTer17)	35	
RMH-4383-0007	CCND3c.872_873delp.His291ProfsTer32	12	17pdel
RMH-4383-0009	CCND3c.852_854delp.Thr285del	16	Not yet reviewed
	BIRC3c.1681delp.Met561TrpfsTer7	20	
RMH-4383-0012	TP53c.838A>Tp.(Arg280Ter)	82	17pdel
	CCND3c.857A>Gp.(Asp286Gly)	38	
RTP-4383-0007	TP53c.725G>Tp.(Cys242Phe)	10	17pdel, 17q amp, del7q (POT1)
	TP53c.621_622dupp.(Asp208ValfsTer40)	45	
	CCND3c.843_871delp.(Ser282ProfsTer32)	12	

Table 3.6: Summary of the clinical data available for four cases with *CCND3* variants. Data is from time of enrolment to ENABLE study. Details on immunophenotype and morphological features are provided.

Case ID	Additional variants	Splenomegaly /LN	Immunophenotype	Morphological features
BEL-4383-0005	No	No	CD5-ve, CD10-ve, CD20+ve, CLL score 0/5, HCL score 0/4	Small to medium sized lymphocytes
RDD-4383-0001	*BRAF,* del7q	Splenomegaly no B-symptoms	CD5+ve, CD10-ve, CD200+ve	Medium-sized lymphocytes some with callous cytoplasm; BM aspirate review favours diagnosis of SMZL
RMH-4383-0007	No	No	CD5-ve, CD10-ve, CD23-ve, CD200+ve. HCL score 0/4	Small/medium sized lymphocytes with some villous projection
RMH-4383-0009	*BIRC3*	No	CD5+ve, CD10 –ve, CD23-ve	Small to medium sized lymphocytes BM biopsy: B-NHL, NOS

BRAF variants are usually absent in SMZL and SDRPL (Bühler et al., 2020). The *BRAF* p.(Gly469Ser) variant is a mutational hotspot seen commonly in lung cancer and melanomas. It falls within the tyrosine kinase domain of BRAF and *in vitro* experiments have confirmed an increase in ERK-phosphorylation driving proliferation (Berger et al., 2016). Given the difficulty of distinguishing between SMZL and NMZL but also CBL-MZ in some cases due to the lack of specific, disease- defining markers, the presence of a variant in *BRAF* could point towards NMZL as reported by Pillonel *et al.* (Pillonel et al., 2018). Despite not being helpful in identifying a definite diagnosis in this case, it could be postulated that the knowledge of variants like an activating *BRAF* change may in future be able to identify possible treatment options should the need arise in these patients. Similarly, the presence of a *BIRC3* variant in RMH-4383-0009 does not aide in classifying this case further given that *BIRC3* variants are seen across all types of indolent, small-cell B-NHLs.

3.7.3 Variants in *MYD88*

Although most common in LPL/WM (incidence >90% (Gertz, 2019)), *MYD88* variants are found in up to 10% of marginal-zone lymphomas, including SMZL where they are exclusive to variants in *KLF2* and *NOTCH2* but associated with *TP53* (Clipson et al., 2015). In addition, *MYD88* variants can be seen in a small number of cases in CLL where they are associated with unusual phenotype (Shuai et al., 2020). They are a helpful to distinguish WM (the IgM-serum protein producing form of LPL) from IgM paraprotein-positive myeloma where *MYD88* is absent (Kaiser et al., 2021). *MYD88* plays a role in several cellular pathways including the innate immune response via triggering of Toll-like receptor (TLR) pathways but for B-NHL and especially LPL, activation of NF-κB signalling as well as enhancing BTK-activity appear to be the main pathogenic changes (Rossi, 2014). This is supported by the high efficacy of BTK-inhibitors like Ibrutinib in the treatment of LPL and WM (Gertz, 2019).

The most common missense variant in *MYD88*, p.(Leu265Pro), was seen in 13 cases with a mean VAF of 32% (range 20-46%). In nine cases, the *MYD88* variant was the only driver change detected and the VAF distribution fits with the monoallelic driver variant scenario described for *MYD88*-driven B-cell neoplasms

(Rossi, 2014). In three cases pathogenic/likely pathogenic variants were detected in other genes involved in BCR- and NF-κB signalling with two cases showing variants in *CD79B* (table 3.7). One case had a *NFKBIE* frameshift variant alongside the *MYD88* hotspot variant, and another showed variants in *CBL* and *PAX5*. The combination of variants would be unusual for LPL/WM cases and perhaps more in-keeping with CLL/CLL-MBL but further review of morphology and flow cytometry is required to integrate all findings. None had a variant in *CXCR4* seen in ~30% of LPL (Hunter et al., 2014).

V(D)J-rearrangements and CD79B *variants associated with the presence of* MYD88 *variants*

The majority of V(D)J-rearrangements carried a V3-gene (9/13) with 5/7 having a *V3-74* and *V3-23* (table 3.7). It has been shown that LPL/WM has a high V3-gene usage with *V3-23* and *V3-74* being the most commonly found V-genes (Gachard et al., 2013, Petrikkos et al., 2014). In addition, LPLs are characterized by hypermutated *IGHV*, often high-hypermutated >5% and short CDR3 amino acid sequence lengths of less than 17 (Rollett et al., 2006). All of the V3-cases are classed as mutated with 4 being high-hypermutated (range 8-13%). CDR3 amino acid sequences for the 7 cases were between 8 and 16 amino acids in length. All but three case include *IGHJ4*02* and no additional SNV or CNVs were detected amongst the seven cases. For two cases, RVR-4383-0005 and RJ6-4383-0001, no Sanger sequencing data is available and the assignment of the *IGHD* and *IGHJ* gene was not possible unequivocally by NGS whilst the SHM status could also not be determined. Both cases showed additional SNV: in *NFKBIE* (VAF 24%) for case RVR-4383-0005 and *CBL* (VAF 25%) and *PAX5* (VAF 21%) for RJ6-4383-0001. In the final case, REF-4383-0002, neither Sanger sequencing nor NGS were able to identify the *IGHV*-rearrangement. Clonality was demonstrated by NGS by the presence of a clonal *VK1-3/KDel* and a productive lambda rearrangement (*VL2-14/J2* or *J3*) by NGS which matches the lambda expression identified by flow cytometry.

In all three cases where Sanger sequencing failed to amplify the *IGHV*-rearrangement, it is likely that the PCR reaction failed due to the presence of a nucleotide change under the primer binding site due to the extensive level of somatic

hypermutations described for LPL/WM (Gachard et al., 2013) and displayed by the majority of other cases in the current *MYD88*-mutated cohort presented here. This would also explain the inability of the NGS pipeline to exactly define which V3-gene is present. Further detailed analysis of the individual NGS reads may be able to demonstrate the location of the variants to confirm this but this was not yet performed. Another possible explanation is clonal evolution with modification of the *IGHV*-rearrangement including ongoing somatic hypermutation acquisition which has been shown to occur and more-detailed analysis of individual NGS sequencing reads could enable further insight.

The *IGHV*-rearrangement could be determined in an additional three cases all with *IGHV4-34* gene usage. Two cases had SHM of 6% with one being borderline at 2%, the latter also had a short CDR3 of 13 amino acids whereas the other two carried the two longest CDR3 sequences with 17 and 22 amino acids. Both SHM-high cases had an additional variant in *CD79B* at VAF similar to the *MYD88* variant in one case and double in the other (table 3.7). No further variants were found in the third case with *IGHV4-34* although CNV analysis is yet to be completed.

The co-occurrence of *MYD88* and *CD79B* variants is well documented for DLBCL-ABC (Takeuchi et al., 2017) but also LPL/WM but in less than 10% of samples. Bommier et al. (Bommier et al., 2021) reviewed the utility of adding a targeted NGS panel analysis to cases of lymphoma with unknown subtypes after expert review. From their findings they state that the presence of a *MYD88* variant together with a *CD79B* variant is strongly suggestive of a diagnosis of LPL over MZL.

The variant detected in case RMH-4383-0013 (p.(Tyr197Cys), VAF 64%) is the main hotspot variant within the ITAM domain and is seen in 75% of *CD79B*-mutated cases in DLBCL-ABC (Wang et al., 2017). Patients with this variant combination have been shown to have a particular good response to the BTK-inhibitor Ibrutinib (Visco et al., 2020). The *IGHV4-34* gene is overrepresented in B-NHL likely due to its autoreactivity recognizing erythrocyte I-i polysaccharide antigen and human B lymphocyte CD45 O-linked polysaccharide antigen (Grondona et al., 2018). It appears that the combination of *MYD88* and *CD79B* variants may play a part in allowing the survival of a self-reactive B-cell (Wang et al., 2017).

The second case has a less-commonly encountered *CD79B* variant p.(Leu200Pro) (VAF 43%); it has been detected in four DLBCL cases and falls within the ITAM domain close to the Tyr197 hotspot.

None of the cases with a *MYD88* variants showed evidence of trisomy 12 or a deletion on the short arm of chromosome 17 involving *TP53*, both of which have been described in the context of LPL/WM albeit at low frequency (Florence et al., 2013, Gustine et al., 2019).

MYD88 is reported in a small percentage of SMZL and *V3-23*-usage is prevalent in SMZL (Zibellini et al., 2010). These findings thus support the suggested diagnosis of SZML based on histopathological results in case RVR-4383-0027 (table 3.7). Similarly, for case RMH-4383-0001 the usage of *IGHV4-34* is rarer in LPL and would fit with a conclusion of CD5+ve SMZL as provided by the lymph node biopsy. Following on from that for case RMH-4383-0013 the usage of *V4-34* would make SMZL more likely but given the differential between LPL and SZML histopathologically, and the presence of a *MYD88* hotspot variant, the differential remains the same.

Based on the molecular features of a single driver variant in *MYD88* together with *IGHV3-74* usage a diagnosis of LPL is a likely diagnosis for cases RDD-4383-0003 and RA2-4383-0002 but as all the above cases exemplify the differential diagnosis is complex and molecular variants and *IGHV*-usage is not distinct enough to make a final call, it can merely support. RA2-4383-0002 was also one of the few cases where bone marrow was available for analysis, making a diagnosis of LPL very plausible. Nevertheless, detection of *MYD88* variants in the peripheral blood of patients with LPL/WM or IgM-MGUS and other B-cell LPDs has been demonstrated by several studies despite the incidence of a true leukaemic phase being low in some of these diseases (Xu et al., 2014, Shekhar et al., 2021). Given the limited number of cases for which clinical information is available, no further conclusion can be drawn, and the differential remains between LPL, CLL and SZML or their precursor states (e.g. non-CLL MBL/CBL-MZ).

Non-hotspot MYD88 *variants*

Case RMH-4383-0004 was found to have two missense changes in *MYD88* not affecting Leu265: one changing Asn299 into Serine (VAF 20%) and one changing Thr302 to Proline (VAF 22%). Both have been described in the context of haematological malignancies (entries on COSMIC and cbioportal databases accessed at https://cancer.sanger.ac.uk/cosmic and https://www.cbioportal.org/, respectively) and Thr302Pro has been shown to be activating in *in-vitro* experiments although at a lower level than the Leu265Pro variant (Shuai et al., 2020)).

This case also carries a rearrangement between chromosome 7 and 14 arising from the IGMswitch locus and breaking into MICALL2/INTS1. It is at this point not known which effect, if any, this has but the location of the breakpoint on chromosome 14 would point towards initiation of the class-switch reaction (CSR) creating double strand breaks in the switch regions of the constant genes.

The NGS analysis for IG-rearrangements was not completed in time to be included in this thesis. Sanger sequencing results show a functional rearrangement using *IGHV3-74, IGHD3-10*01* and *IGHJ4* which has a high prevalence in LPL. The rearrangement was mutated with a somatic hypermutation level of 6.9 and a CDR3 length of 11 amino acids (CAGFDWGAYFR) which again fits with the expected characteristics of short, highly mutated CDR3s in LPL. Nevertheless, reports of non-hotspot *MYD88* variants are usually from non-LPL cases like CLL, SMZL and DLBCL (Rossi, 2014).

The patient is a 55-year old man who was diagnosed back in 2007 and enrolled onto the ENABLE study in December 2016. He has an IgG kappa paraprotein that has been steadily increasing from 7g/L at diagnosis to 16g/L at enrolment. Lymphadenopathy was present and previous cytogenetic analysis showed 13q14 loss (15%) and *TP53* loss (7%). His bone marrow trephine examination IHC suggested that it was most likely CLL with an unusual distribution and atypical phenotype. Morphologically, medium sized lymphocytes with moderate amounts of cytoplasm and a few prolymphocytes were reported. CD5+, CD19+, CD20+ (mod-strong), CD23+ (partial), CD200+, CD22+, CD79b+, FMC7+, CD10-, CD43-. CLL score 2/5. Analysis of a large cohort of 1179 CLL cases by Shuai *et al*. (Shuai et al., 2020) found *MYD88* variants in 3.1%. The comparison between cases with Leu265Pro hotspot variants compared to other *MYD88* variants showed that those

with the Leu265Pro variant tended to be younger, had mutated *IGHV* and carried an isolated del(13q14.3). Interestingly, this case carries the deletion on 13q based on cytogenetic analysis, has mutated *IGHV* and is only 55 years old but does not carry the hotspot variant. It could be postulated that the acquisition of the del17p involving *TP53* may be one factor indicated in the progression of the disease as testified by the increase in paraprotein. Although some features would be compatible with LPL, most findings including the morphological and histopathological evaluation favour CLL, which would be supported by the combination of an unusual *MYD88* variant together with del13q14.

Table 3.7 (next page): ***MYD88* variants.** Top panel: Summary of the 13 samples with the pathogenic *MYD88* hotspot variant p.(Leu265Pro). The VAF of the variant found by NGS is provided together with the V-D-J gene, SHM status and CDR3 length with amino acid sequence. Light chain restriction was concordant between NGS and immunophenotyping data. Additional SNVs were found as listed. * NGS data available only for IG-rearrangement; # MYD88 variants other than p.(Leu265Pro). Bottom panel: Clinical information and results of other tests available for 6 of the samples with *MYD88* variants.

Sample ID	*MYD88* VAF (%)	V-gene	D-gene	J-gene	% SHM	CDR3 length	Amino acid sequence	Light chain restriction	Other SNVs (VAF)/CNVs
RA2-4383-0002	39	IGHV3-74*01	IGHD6-6*01	IGHJ4*02	3	15	CVRGGSYSTSSGDYW	Lambda	none
RVR-4383-0027	28	IGHV3-23*01	IGHD7-27*01	IGHJ6*02	4	16	CAKESREYYYYGMDVW	Kappa	none
RDD-4383-0003	27	IGHV3-74*01	IGHD3-22*01	IGHJ4*02	5	13	CARSPGKGSSNFW	Kappa	none
RTP-4383-0002	20	IGHV3-48*02	IGHD3-16*02	IGHJ4*02	8	16	CARGVTWGTERYIDSW	Lambda	none
REF-4383-0014	34	IGHV3-7*01	IGHD2-2*01	IGHJ4*02	12	14	CVSSEDLLRNFHHW	Kappa	none
RHU-4383-0003	27	IGHV3-15*01	IGHD5-12*01	IGHJ4*02	12	15	CMTDPIVGTTPGDFW	Lambda	none
RJ6-4383-0002	23	IGHV3-53*02	IGHD4-11*01	IGHJ4*02	13	8	CTRHETIW	Lambda	CBL (25%), PAX5 (21%)
RVR-4383-0005*	28	IGHV3-23	NA	IGHJ4/5	NA	11	CAKSLKEDMYW	Kappa	NFKBIE (24%)
RJ6-4383-0001*	46	IGHV3-23	NA	IGHJ3	NA	15	CARDAETGWGLFDMW	Kappa	none
RVR-4383-0011	33	IGHV4-34*03	IGHD2-8*02	IGHJ5*02	2	13	CATGGGPISWFDPW	Lambda	none
RMH-4383-0001	38	IGHV4-34*02	IGHD2-21*02	IGHJ2*01	6	17	CAKGPSPQGPGRDFDLW	Lambda	CD79B (43%)
RMH-4383-0013	32	IGHV4-34*01	IGHD4-17*01	IGHJ2*01	6	22	CARGLSLGNTVTTPPYWYFDLW	Kappa	CD79B (64%)
REF-4383-0002	40	NA	NA	NA	NA	NA	NA	Lambda	none
RMH-4383-0004#	20,22	IGHV3-74*01	IGHD3-10*01	IGHJ4*02	6.9	11	CAGFDWGAYFR	NA	? ARID1A

Sample ID	*MYD88* VAF (%)	V gene	SHM (%)	CD5	CLL score	Paraprotein	Morphology	Comments
RA2-4383-0002	39	IGHV3-74*01	3	Pos	2	No	Low-grade B-NHL no distinguishing features	Sister has CLL*, splenomegaly
RVR-4383-0027	28	IGHV3-23*01	4	Neg	0	NA	Villous cytoplasm, Bone marrow trephine favours SZML	Splenomegaly, HCL score 2/4; likely SMZL
RDD-4383-0003	27	IGHV3-74*01	5	Pos	2	IgG Lambda	NA	Progressive lymphocytosis
RMH-4383-0001	38	IGHV4-34*02	6	Pos	NA	IgM Kappa	CD5+ SMZL based on LN biopsy	Multiple previous episodes of lymphocytosis with varying CD5 status
RMH-4383-0013	32	IGHV4-34*01	6	Neg	NA	No	Differential lies between SMZL or LPL based on LN biopsy	Presented with lymphadenopathy
RMH-4383-0004#	20,22	IGHV3-74*01	6.9	Pos	2	IgG Kappa	IHC in-keeping with CLL, otherwise atypical, CD5+, CD23 partial, CD200+	13q14 loss (15%) and *TP53* loss (7%)

* Familial predisposition has been shown to exist in CLL as well as LPL/WM (Hanzis et al., 2011, Law and Houlston, 2019)

3.7.4 Variants in genes associated with NF-κB and BCR-signalling pathways

As shown in figure 3.11, several variants were found in genes associated with NF-κB- and BCR-signalling pathways, including *TRAF2, TRAF3, NFKBIE, BIRC3* and *TNFAIP3*. The most frequently mutated gene, *MYD88*, has already been discussed separately above (chapter 3.7.3). Although variants in these pathways have been described across different B-NHL entities and none of these variants are disease-defining, they differ in frequency between subtypes and, in the context of this study of small-cell B-NHL, the differential includes the various types of marginal zone lymphomas alongside atypical CLL, HCL-v and MCL. Presence of variants can lend support to a particular diagnosis: for example, presence of variants in *NOTCH2* or *KLF2* has been shown to have a high positive-predictive value for a diagnosis of SMZL (Spina et al., 2021).

Table 3.8 provides an overview of the 20 cases included in this chapter. Variant details including VAF for each case can be found in table 3.3 whilst the IG-rearrangement information is available in appendix table A7.

Table 3.8: Distribution of variants in 20 cases with driver variants in the BCR-signalling cascade or NF-κB pathways. Y= variant detected; tbc= to be confirmed; not all cases have been assessed for del7q and trisomy 12. Deletion of the long arm of chromosome 7 is a recurrent feature in SMZL whereas trisomy 12 can be found across various B-NHL types including the MZL and atypical CLL. Additional findings not associated with NF-κB- or BCR-signalling are listed in the comments column and are discussed in more detail in the text.

Case ID	KLF2	NOTCH2	del7q	tris12	BIRC3	TRAF2/3	NFKBIE	TNFAIP3	CARD11	Comments
BEL-4383-0001	Y		Y	N						IGHV1-2*04
RJ6-4383-0003	Y		Y	Y						
RMH-4383-0002	Y		Y	N				Y		NOTCH1
RTP-4383-0001	Y		tbc	tbc						
BEL-4383-0006		Y	tbc	tbc		Y	Y			
RVR-4383-0022		Y	tbc	tbc				Y		IGHV1-2*2
RVR-4383-0015		Y	tbc	tbc		Y				
RHU-4383-0001		Y	N	Y	Y					
REF-4383-0010			tbc	tbc		Y				
RVR-4383-0025			tbc	tbc		Y				
RVR-4383-0018			tbc	tbc				Y		
RVR-4383-0006			tbc	tbc			Y			
RA2-4383-0001			tbc	tbc					Y	
RVR-4383-0023			N	Y					Y	
REF-4383-0006			tbc	tbc					Y	
RDU-4383-0004			tbc	tbc	Y					
REF-4383-0012			N	Y	Y					
RMH-4383-0003			tbc	tbc	Y					low tumour infiltration
RMH-4383-0009			tbc	tbc	Y					17pdel, CCND3
RVR-4383-0004			tbc	tbc	Y					

*Variants in Krüppel-like factor 2 (*KLF2*)*

KLF2 acts as a tumour suppressor and truncating variants are most commonly seen in lymphoid malignancies (Höllein et al., 2017). As shown by Clipson *et al.* in an *in-vitro* reporter assay, wild-type KLF2 is a potent inhibitor of NF-κB signalling even in the presence of *CARD11* and *MYD88* variants but the suppression is lost if KLF2 is mutated allowing for uncontrolled signalling (Clipson et al., 2015). Within the cases identified in this cohort, variants in *KLF2* included a 17bp-deletion seen in case RMH-4383-0002 which lies across the exon-intron boundary with loss of the 5' splice acceptor site for exon 2 (table 3.3, figure 3.15). It is assumed that deletion of this variants will cause aberrant splicing and potentially loss of translation altogether (Anna and Monika, 2018). Another potential splice variant was seen in BEL-4383-0001 c.76-1 (VAF 29%). In the other two cases, missense variants were identified at codons 288 and 291, two of the most-commonly reported locations of variants in SMZL (Piva et al., 2015, Stamatopoulos et al., 2004). Piva *et al*. first reported the high prevalence of mutations in *KLF2* in SMZL. In their cohort of 96 SMZL cases, 19 were positive for a *KLF2* variant and *KLF2* variants were also found in other B-NHLs including HCL, DLBCL and NMZL in their study extension. Codons His288 and Ala291 are evolutionary conserved and lie within the first zinc-finger domain of KLF2 which is involved in DNA binding. Cell-culture experiments demonstrated that mutant proteins caused by the presence of a variant at codon 288 were unable to access the nucleus and could thus no longer act as a transcription factor suppressing NF-κB activity (Piva et al., 2015). Piva *et al.* also already established the correlation of *KLF2* variants with *NOTCH2* variants, del7q and *IGHV1-2*04* usage as hallmarks of SZML, which has received support from additional studies since and is discussed further in case study BEL-4383-0001 below (Clipson et al., 2015).

The variant p.(Ala291Val) is recurrently reported in the context of SMZL (Jaramillo Oquendo et al., 2019), however, the *in-vitro* reporter assays utilized by Clipson *et al.* failed to show a reduction in suppression of NF-κB activity compared to wildtype KLF2. In their cohort additional variants in *KLF2* were present in addition to this variant which was not the case in case RJ6-4383-0003 (Clipson et al., 2015).

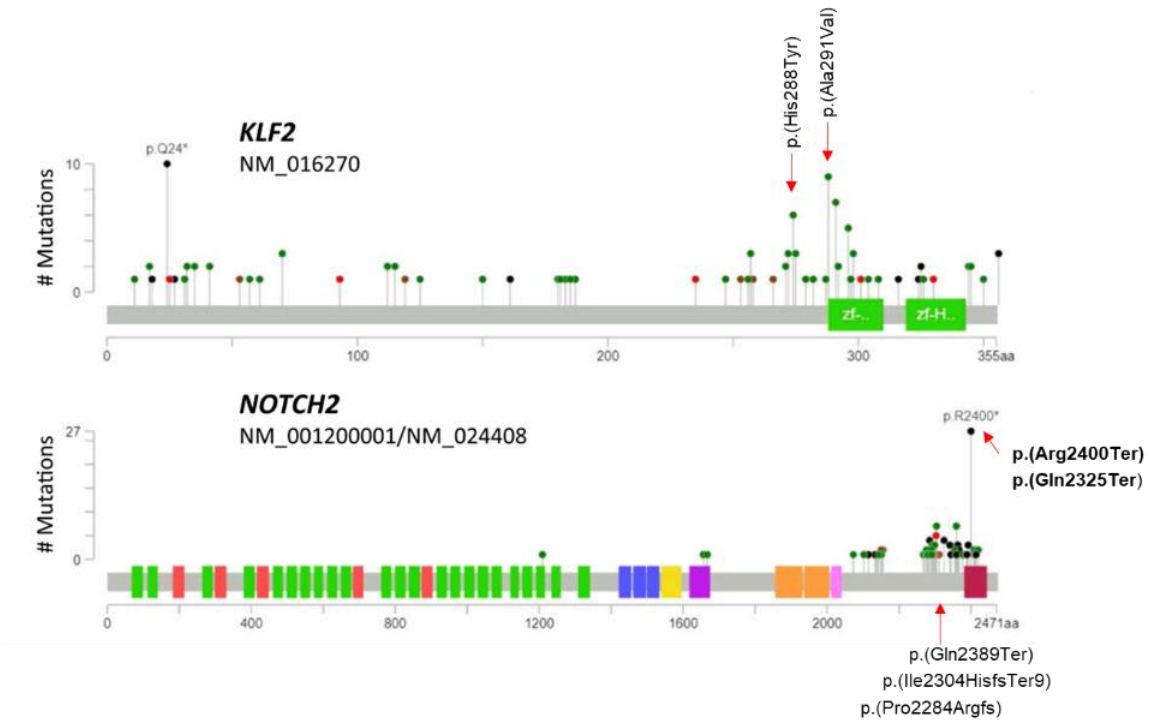


Fig.3.15: **Variants in *KLF2* and *NOTCH2***.This diagram was taken from the review of mutations in SMZL published by Jaramillo et al in 2019 (Figure 3 page 6, (Jaramillo Oquendo et al., 2019)). Superimposed are the two hotspot variants found in this cohort in *KLF2* as well as the PEST domain variants in *NOTCH2* which all fall within the areas frequently changed by variants in SMZL and other B-NHLs.

Variants in NOTCH2

Similar to *KLF2*, variants in *NOTCH2* are most prevalent in SMZL but can be found across other MZL -in particular NMZL- and various B-NHLs at lower frequencies (Hurwitz et al., 2021, Bühler et al., 2020). Variants in *NOTCH2* and the related *NOTCH1* have been best characterised in SMLZ and CLL, respectively (Arruga et al., 2018). They usually lie within exon 34 which encodes the PEST-domain of all Notch proteins (Kiel et al., 2012) and result in impaired degradation of the Notch intracellular domain (NICD) which is cleaved upon receptor activation and relocates to the nucleus to act as a transcription factor (Arruga et al., 2018, Shanmugam et al., 2021). In the cohort presented here, four cases were found to have *NOTCH2* variants (figure 3.15). All variants were located in exon 34. Three were nonsense and two were frameshift variants with a range in VAF of 8-38%. The low VAF of 8% was present in case RHU-4383-0001 which carried two nonsense variants; the one with a VAF of 28% fits more with a *bone fide* driver variant and these could be biallelic although this cannot be confirmed by the assay employed.

In RMH-4383-0002, two variants were detected in *NOTCH1* alongside the *KLF2* and *TNFAIP3* splice variants. One is a missense p.(Pro2369Thr) and one is a nonsense variant, both with similar VAF and located in the PEST domain. In CLL, more than 80% of *NOTCH1* PEST domain mutations consist of a 2bp deletion which is different to the nonsense variant seen here (Sorrentino et al., 2019). *NOTCH1* variants are most frequent in CLL and DLBCL but are seen in about 5% of SMZL (Onaindia et al., 2017) but not the other MZL subtypes (Spina et al., 2021).

*Variants in tumour necrosis factor receptor associated factor (*TRAF*) genes*

TRAF2 and *TRAF3* are cytoplasmic signal-transduction proteins involved in several cell signalling cascades including TLR signal transfer and negative control of NF-κB activity (figures 1.3 and 1.5 (Zhu et al., 2018)). Inactivation through mutations results in elevated NF-κB activity which is well documented in MCL and DLBCL but also MALT and SMZL (Zhu et al., 2018, Jaramillo Oquendo et al., 2019). In SMZL, variants in these genes are commonly found alongside variants in *KLF2* and *NOTCH2* (Clipson et al., 2015).

A total of 5 *TRAF* variants was detected (4 in *TRAF3* and one in *TRAF2*), all of which are frameshifts or nonsense variants which fits with loss-of-function (LOF) as the expected mechanism. The location of the variants detected in *TRAF3* is shown in figure 3.16.

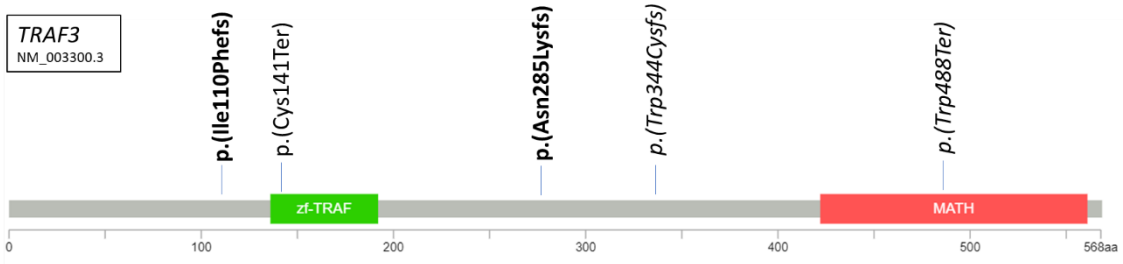


Fig. 3.16: **Schematic representation of the TRAF3 protein including functional domains.** Location of the 5 variants detected in the cohort are shown. In two cases two variants were present, which may be biallelic; these are highlighted as pairs in bold and italics.

In REF-4384-0010 and RVR-4383-0015 two *TRAF3* variants were detected with an additional *NOTCH2* variant seen in the latter. In both cases one variant had a

significantly higher VAF than the second, which is thus assumed to be a secondary variant acquired during clonal evolution (table 3.3). For the third case RVR-4383-0025 the nonsense variant in *TRAF3* was the only driver change found so far (VAF 18%). Variants have been described throughout the gene without particular hotspots, although it has been noted that variants from SMZL and other rare B-lymphomas are currently underrepresented in the main databases like COSMIC for example (Jaramillo Oquendo et al., 2019). All known pathogenic variants result in the loss or disruption of the C-terminal MATH domain which is required to interact with MAP3K14 to initiate its degradation via BIRC3 within the non-canonical NF-κB pathway leading to uncontrolled NF-κB signalling (Bertoni et al., 2018).

The only *TRAF2* variant found in the cohort was present in BEL-4383-0006. It is a potential splice variant with a VAF of 6% that is present alongside nonsense variants in *NFKIEB* and *NOTCH2* with 25% and 38% VAF, respectively. It is likely that the *NOTCH2* and *NFKBIE* changes are the founder variants in this case with the *TRAF2* variant having been acquired at a later stage (subclonal).

*Variants in TNF alpha-induced protein 3 (*TNFAIP3*)*

TNFAIP3 is a tumour suppressor and acts as a negative regular of NF-κB-activity (figure 1.5 (Grondona et al., 2018)). It encodes A20 and variants in this gene are mainly found in lymphomas and leukaemias but not in solid cancers. Variants reported are frameshift and nonsense variants across the length of the gene. This study identified three cases with *TNFAIP3* variants (table 3.3 and 3.8): One splice and two frameshift variants. In two cases this finding was associated with the presence of variants in one of the *NOTCH* genes with a *KLF2* variant seen in one case. In the third case without *NOTCH* or *KLF2* variants, a nonsense variant in *ARID1A* was also seen but this has not yet been evaluated or reviewed further.

TNFAIP3 variants have also been reported with highest frequencies in MZL-type B-NHL with one study showing a particular high prevalence in MALT of the ocular adnexa where it was associated with restricted *IGHV*-gene usage in particular *IGHV4-34* (Moody et al., 2017). The same study found *IGHV1-69* to be prevalent in MALT of the salivary gland however, the number of cases was low. *IGHV3-30* has been shown to be associated with gastric MALT (Zucca and Bertoni, 2016). IGHV-

genes for the three cases detected in our cohort were *IGHV1-69, IGHV1-2*2* and *IGHV3-30* and no further comparison with regards to CDR3 stereotypes has yet been undertaken.

In addition to *TNFAIP3* inactivation through SNVs, deletion of the gene locus at chromosome 6q23 is a recurrent finding in MZL/MALT (Zucca and Bertoni, 2016, Rinaldi et al., 2011) and further analysis to look for the presence of such deletions will be undertaking in future.

*Variants in baculoviral IAP repeat containing 3 (*BIRC3*)*

BIRC3 encodes an inhibitor of apoptosis (IAP) cellular protein Iap2. It is relevant across various cancer entities due to its role in the regulation of apoptosis but for B-NHL its effect on non-canonical NF-κB signalling is the most important (Honma et al., 2009). *BIRC3* variants in B-NHL cases always lack the RING-domain which, similar to mutations in *TRAF3*, result in the inability to ubiquitinate and thus inactivate MAP3K14 within the NF-κB pathway (Bertoni et al., 2018).

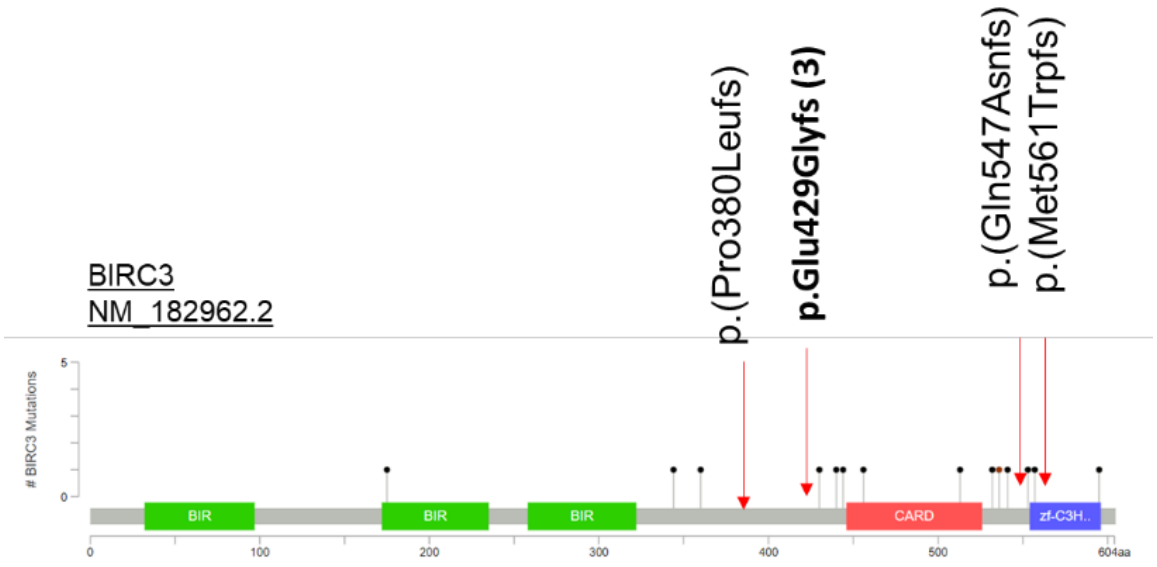


Fig. 3.17: **Schematic representation of the BIRC3 protein including functional domains**. Location of the 6 variants detected in the cohort is shown. Codon 429 was affected three times. Figure copied from OncoKB and data is from MSK-IMPACT Clinical Sequencing Cohort (Zehir et al., Nature Medicine, 2017)(Zehir et al., 2017). The area referred to as RING domain is highlighted in blue.

As shown in figure 3.17 the frameshift and nonsense variants seen in our cohort are predicted to lead to truncation upstream of the RING domain. They all fall within the two areas identified as mutations hotspots in CLL between aa 367-438 and 537-564 (Diop et al., 2020). Mutations are seen in 10% of SMZL and 5% of NMZL but are most commonly seen in CLL and MCL where they have been shown to carry prognostic significance and are associated with chemoimmuno-refractoriness (Diop et al., 2020, Tausch and Stilgenbauer, 2020). *BIRC3* variants occurred alongside *NOTCH2* variants and trisomy 12 in once case. In another case a variant in *CCND3* and *TP53* was present together with del17p and in another a nonsense variant in *CXCR4* and a potential splice variant in *ARID1A* together with trisomy12 were seen but in three cases *BIRC3* was the sole driver variant identified.

The molecular hallmark of MALT lymphomas are translocations between *BIRC3* and *MALT1* with the fusion leading to the loss of RING domain of *BIRC3*. The result is increased NF-κB-signalling through sustained MAP3K14 activity as would be predicted for the *BIRC3* single-nucleotide variants seen here. The NGS panel used in this study is capable of detecting *BIRC3-MALT1* fusions and other translocations reported for MALT but none were found amongst the 101 samples analysed. In RVR-4383-0013 a chromosomal rearrangement involving *BIRC3* was detected. The break on chromosome 14 occurs in the large intron 8 of *RAD51B* (genomic location chr14:68,253,931) with the break in *BIRC3* located in the Major breakpoint region described for the *BIRC3-MALT1* fusions. However, no reads showing the derivative chr11 are found and this could represent an unbalanced fusion. Whether or not the remaining *BIRC3* is transcribed and functional in the same way as a truncated *BIRC3* gene with a frameshift variant needs to be investigated further.

*Variants in caspase recruitment domain family member 11 (*CARD11*)*

The *CARD11* gene product is crucial for the signal transduction form activated BCRs and TRs resulting in the activation of various cellular signalling cascades including NF-κB (Bedsaul et al., 2018). Driver variants in *CARD11* are gain-of-function changes that trigger downstream pathway activation in the absence of ligand binding to BCRs and TRs (Chung, 2020). In addition to being part of the pathology of the clonal neoplasm, variants in *CARD11* are also a mechanism for the clone to escape suppression when patients are treated with BCR-inhibitors like Ibrutinib as

well as the protein kinase C inhibitor sotrastaurin (Naylor et al., 2011). In particular in CD5-negative cases of ABC-DLBCL, a high incidence of *CARD11* has been reported and may contribute to the clinical phenotype of the disease (Lenz et al., 2008, Takeuchi et al., 2017). Mutations in *CARD11, CCND3* as well as *TNFAIP3* have recently been shown to be acquired in B-cells producing autoantibodies against rheumatoid factor which provides evidence for a shared route of development between autoimmune disease and neoplastic lymphoma (Singh et al., 2020).

Activating variants are pre-dominantly missense changes in the coiled-coil domain in both DLBCL and SZML (Rossi et al., 2012, Lenz et al., 2008). Differences in the degree of autonomous signalling caused by different variants have been documented (Bedsaul et al., 2018). In DLBCL and SZML, *CARD11* variants are usually secondary events and are encountered together with other variants. However, the three cases analysed here (figure 3.18) show no further SNVs with the full CNV analysis still to be performed (one case, RVR-4383-0023, had evidence of trisomy12 and no CNV on 7q, table 3.3).

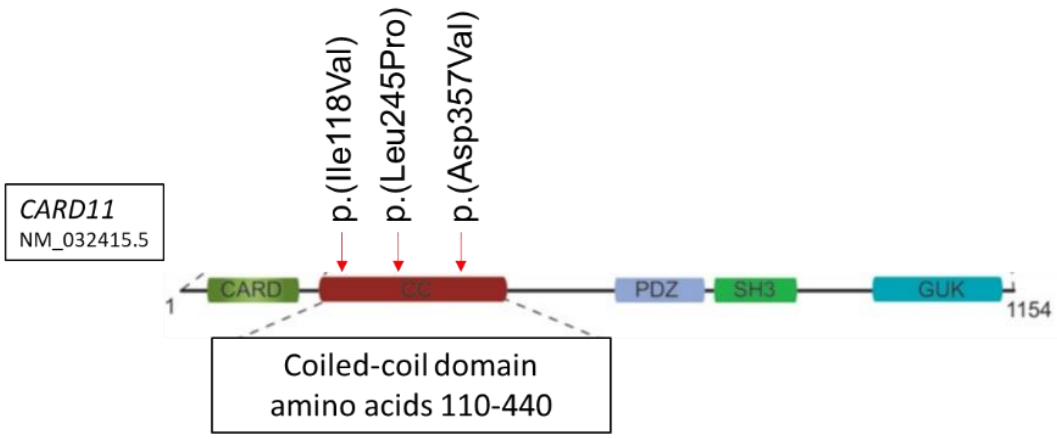


Fig. 3.18: Schematic representation of the CARD11 protein including functional domains. All three variants detected are missense variants affecting the coiled-coil domain. Upon activation through BCR/TR triggers, the coiled-coil domain of the Card11 becomes accessible and can interact with the downstream effector molecules (Bedsaul et al., 2018). Point mutations in the CC domain abolish the function of the autoinhibitory loop mechanism and keep the protein in an active state making the CC domain available for interactions without appropriate stimulus leading to overactivity of downstream pathways including NF-κB (Bedsaul et al., 2018)

3.8 Copy-number variants (examples of preliminary analysis)

3.8.1 Copy-number involving the long arm of chromosome 7

As described for the CNV plots examples for the del17p in chapter 3.7.1, figure 3.19 shows the CNV plot for case RMH-4383-0002 (preliminary analysis only). A clear drop in ratio is visible across all targets on 7q, whereas the coverage for 7p is unchanged. Based on the design of the panel, the results confirm that the deletion includes the genes *POT1, BRAF* and *EZH2*. Further potential CNVs are seen, which need to be scrutinized further and the case has now been submitted for analysis using the validated Bioinformatics pipeline.

CNVs in 7q were seen in two additional cases with *KLF2* variants (table 3.3 and 3.8). In RJ6-4383-0003 the deletion was focal and included *POT1* but not *BRAF* and *EZH2*, which was the same area deleted in case R1K-4383-0007. In the latter two *BRAF* variants and a change in *ARID1A* were present alongside del7q. RDD-4383-0001 showed loss of all three genes with concomitant variants in *CCND3* and *BRAF*. Relevance of these findings has not yet been reviewed further.

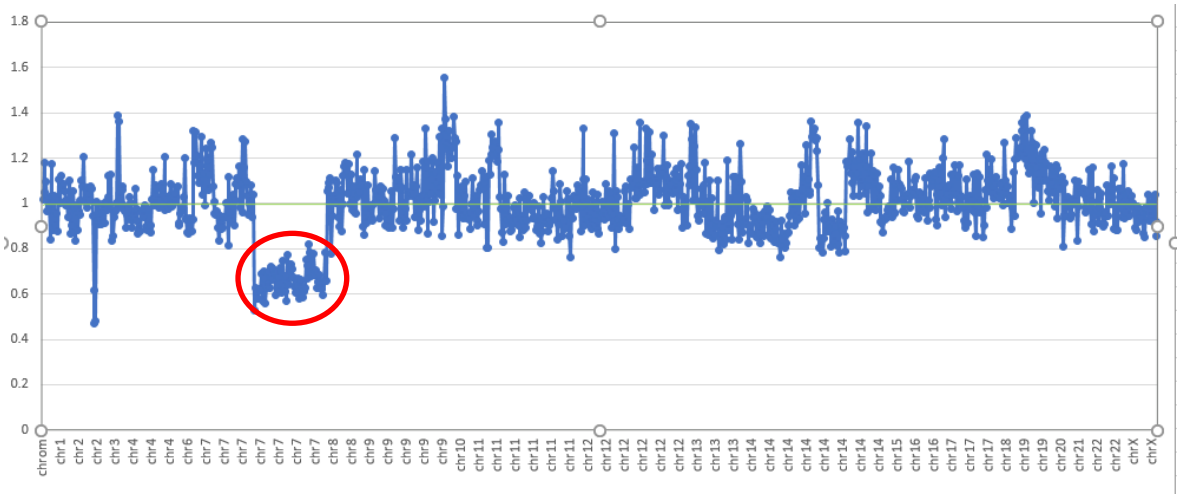


Fig. 3.19: Example of the CNV analysis plot comparing tumour and germline coverage at each target. Normalized coverage for each sample was used and the ratio plotted. The green line represents the average ratio. The data suggests deletion of chr7q circled in red (including *POT1, BRAF* and *EZH2*), possible amplification on chr9 including *TRAF2* and chr11 (*CCND1*) as well as some amplification on chr19 including *KLF2* and *TCF3* loci. Given this is not a full molecular karyotype, whether this is focal amplification, whole arm or even chromosome cannot be ascertained at present. Further analysis is underway.

Clinical details were only available for three of the cases discussed in this chapter: RMH-4383-002 (82-year old lady), diagnosed back in 2012. She was asymptomatic at diagnosis and splenomegaly was an incidental finding on a CT scan performed for breast cancer follow-up. Her clonal B cell population was CD5+ve, CD10+ve and

CD23-ve without any significant morphological features. A bone marrow biopsy performed in 2015 showed 40% involvement with MZL or LPL. She received Rituximab monotherapy in 2016 due to progressive splenomegaly and cytopaenias. Her lymphocytosis resolved and her counts recovered. The patient received no further treatment for her B-NHL and died of metastatic breast cancer in 2018. The clonal *IGHV* rearrangement utilized *V1-69* which is common in both CLL and SMZL with the former usually unmutated and mutated in the latter (Zibellini et al., 2010). The SHM for this case was 6.6%. Presence of variants in *KLF2, TNFAIP3* and del7q supports the MZL-origin for this case; the presence of two *NOTCH1* variants, however, is unusual in this context and would be expected in CLL. From our data there is no evidence of two separate clonal B-cell populations and together with the histopathological data the likely diagnosis remains MZL with *KLF2* variant and del7q supporting SMZL (Clipson et al., 2015, Höllein et al., 2017).

Cases RVR-4383-0015 and RVR-4383-0022 show *NOTCH2* variants together with *TRAF* or *TNFAIP3* variants, which suggests MZL-related disease. RVR-4383-0015 carried a mutated *IGHV4-39* rearrangement; RVR-4383-0022 an unmutated *IGHV1-2*02*; both of these but in particular *V1-2*02* have been found with high frequencies in SMZL (Zibellini et al., 2010). RVR-4383-0015 presented with splenomegaly, B-symptoms (lymphocytosis and anaemia) and weight loss at the age of 60. The film showed small mature lymphocytes, some with clefted nucleus and some larger forms with prolymphocytic features. CD5 was positive, CD10, CD200 and CD23 negative by flow cytometry. The bone marrow results suggested a diagnosis of SMZ. The patient received treatment with R-Bendamustine in July 2018 with no further follow-up data yet available.

RVR-4383-0002 was diagnosed in 2010 with lymphocytosis and mild cytopaenias although displayed no clinical symptoms. A CT-scan showed splenomegaly and abdominal lymphadenopathy at enrolment in 2018 (patient was 66 years old at the time). Bone marrow investigations favoured a diagnosis of SMZL with LPL in the differential, although *MYD88* was negative. The clonal B cells were negative for CD5, CD10 and positive for CD23 and CD200. The patient did not required treatment for her B-NHL but died of metastatic thyroid carcinoma in April 2020. Based on the molecular investigations and the presence of *IGHV1-2*02*, a diagnosis of SMZL is likely.

3.8.2 Copy-number variants involving chromosome 12

Evidence of amplification events on chromosome 12 was detected in 17 cases out of the 40 cases investigated for CNVs so far (43%, table 3.3). For 16 cases the data suggest complete trisomy 12 whereas the amplification appears restricted to the long arm in one case (RMH-4383-00014). Table 3.9 provides the details for these cases. *IGHV4* and *IGHV3* represent the majority of *IGHV*-genes used with 6 cases carrying a *IGHV4-34* (35%). *IGHV4-34* is usually mutated in CLL but only 2 of the case show mutated V(D)J, one is borderline and two show complete homology with the germline, which is unusual.

Overall, four cases were truly unmutated with complete homology to germline (SHM level of 0%), one case had a borderline SHM status (1.4%) and 7 were classed as hypermutated. For 5 cases no SHM was available (see chapter 3.3 for details). In 8 cases the trisomy 12 was the only pathogenic change identified (no evidence of SNVs or chromosomal translocations with the full CNV analysis yet to be completed). Two cases had a translocation arising out of the IGH-locus (partnered with *BCL3* or *BCL2*). In the remaining cases additional SNVs were present in a wide variety of genes and there was no evidence of del17p in any of the cases.

In rare cases (<2%), translocations between *IGH-BCL2* and *IGH-BCL3* have been described in patients diagnosed with CLL and they appear to be secondary changes in this context (Swerdlow et al., 2016a). As is the case for the two cases with these translocations in this group, cases with *BCL2* have mutated *IGHV* whereas *BCL3* cases have unmutated *IGHV* and presence of trisomy 12 alongside these translocations has been documented before. According to the entry criteria for this study all cases in this cohort have a CLL score of 3 or below. The diagnosis of CLL/SLL according to WHO classification uses the Matutes' (also called RMH) scoring system based on flow cytometric markers shown in table 1.2 and 3.9 (Köhnke et al., 2017). CLL is expected to co-express CD5 and CD23 with absence of CD22, CD79b and FMC7 and low (dim) staining for surface immunoglobulins. The only other mature B-NHL that consistently expresses CD5 is MCL and in contrast to CLL strongly expresses sCD22, CD79b and FMC7 with negativity for CD23.

Table 3.9: Overview of cases with trisomy 12 in the ENABLE cohort. *IGHV* gene usage, SHM and CDR3 lengths are provided together with any additional SNVs detected in this study. The CLL score together with the relevant immunophenotyping markers required for calculation (see text for details) is listed with the addition of CD200. Brief comments on the morphological features are provided but the full diagnostic work-up was not available. Information is incomplete for non-molecular findings.

Case ID	IGHV usage	Mutational status (%)	CDR3 length (aa)	Additional Variants (SNVs only)	CLL score	Morphologically	CD5 Pos =1, Neg = 0)	CD23 Pos =1, Neg = 0	FMC7 (Neg =1, Pos =0)	CD79b or CD22 (neg/weak =1, Strong =0	slg (weak =1, Strong = 0)	CD200	Other
RVR-4383-0026	V4-4*07	9.5	8	None	3	Not atypical	Pos	Pos	Pos	Pos	Mod	Pos	CD43+
RVR-4383-0023	V4-34*01	7.4	13	CARD11	1	Some nucleoli	Pos	Neg	Pos	Pos	Strong	Pos	CD43 -
RVR-4383-0016	V4-34*01	0	24	BCL3 translocation	0	Not typical of CLL	Neg	Neg	Pos	Pos	Strong	Pos	CD43 - Possible follicular lymphoma on skin biopsy
RVR-4383-0013	V4-34*02	11	21	KMT2D, FBXW7	3	Clefted/irregular nuclei and frequent nucleoli. Morphology not typical of CLL.	Pos	Pos	Pos	Weak	Mod	Pos	CD43-
RMH-4383-0014	V4-34*01	NA	NA	None	1	Not atypical	Pos	Neg	ND	Pos	Strong	Pos	CD43+
REF-4383-0012	V4-34*01	0	18	ARID1A, CXCR4, BIRC3	1 (or 0)	Hairy morphology on PB	Neg	Dim Pos (Neg	Pos	Pos	Bright	? ND	CD38 dim
RAX-4383-0004	V4-34*01	1.4	18	None	1	Not atypical	Pos	Neg	ND	Pos	Mod	Pos	CD43-
RMH-4383-0011	V3-9*01	8.7	18	BCL2 translocation	1	Not atypical	Neg	Neg	ND	Partial	Mode	Pos	Partial CD43+
REF-4383-0009	V3-9*01	4.8	18	None	? 2 or 3	Not clear	Pos	Pos	Pos	Pos	Unclear	?	CD38 dim
RAX-4383-0007	V3-15*01	8.8	17	KMT2D	3	Occasional nucleoli	Pos	Neg	ND	Weak	Mod	Pos	CD43- CD38 dim
RDU-4383-0003	V3-48*02	4.9	11	None	1	Not atypical	Pos	Neg	Pos	Pos	Mod	Pos	CD43-
RJ6-4383-0003	V3-30	NA	NA	KLF2	1	Not atypical. Some irregular nuclear edge.	Neg	Neg	Pos	Pos	Weak	Neg	CD43-
RHU-4383-0004	V3	NA	NA	None	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details
RHU-4383-0001	V3	NA	NA	BIRC3, NOTCH2 (2)	?2 or 3	Awaited	Pos	Pos	Pos	Pos	UNK	Pos	CD38 partial + CD43-
BEL-4383-0002	NA	NA	NA	KRAS	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details
R1K-4383-0007	V6-1*01	0	14	ARID1A, BRAF (2)	0	Not atypical	Neg	Neg	ND	Pos	Mod	Pos	CD43 -
REF-4383-0005	V1-69*-1	0	24	NRAS (2), NOTCH1	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details	Not got details

Cases exist which display features in between CLL and MCL. Typically, prolymphocyte levels are increased (5-55% above that would be B-PLL) and cases can be CD23-ve, FMC7+ve, CD79b+ve, strong surface Ig and some are CD5-ve. CD200 is an additional useful marker that allows the distinction between CLL (including atypical CLL, see below) and MCL cases as CD200 is positive in CLL but almost always absent in MCL. CD200 expression was assessed for 12 of the 17 cases with trisomy 12 and 11 were positive with RJ6-4383-0003 the only case negative for CD200 (table 3.9). This case carries the *KLF2* p.(Ala291Val) variant discussed above (hotspot variant in SMZL but inconsistent functional data) together with a potential partial deletion on chromosome 7q involving *POT1* based on the manual CNV analysis. The pipeline CNV analysis did not call this deletion so cytogenetic studies or arrays are required for confirmation. Presence of *KLF2* and CNV on 7q could indicate a MZL/SMZL-type neoplasm or precursor version of CBL-MZ in which trisomy 12 has been described as a finding (Bertoni et al., 2018)

Although not a WHO-defined disease entity, atypical CLL is a known concept and refers to a group of cases that present with lower CLL scores (usually score of 3) due to abnormal phenotypes and with unusual morphological features such as cleaved nuclei and lymphoplasmacytoid features (Abruzzo et al., 2018, Xochelli et al., 2014). As shown by Autore *et al.,* atypical CLL is associated with trisomy 12 in 20% of cases of which 40% also carry *NOTCH1* mutations with variants in *FBXW7* seen in a smaller subset (Autore et al., 2018). Patients in this group have unmutated *IGHV*, a high proliferative rate and a higher risk of Richter's transformation. Interestingly these cases were shown to predominantly utilise *IGHV4-39* with specific stereotyped receptors (Autore et al., 2018). None of the cases with trisomy 12 in this cohort show this particular V-gene and no stereotype association was found suggesting that CLL with trisomy 12 and the case with potentially atypical CLL/non-CLL MBL with trisomy 12 have different characteristics.

Compared to these descriptions of atypical CLL the cases summarised here do not fulfil the major features apart from the presence of trisomy 12 and CD5+ in nine of the cases. None of the clones utilised the *IGHV4-39* gene as its rearrangement and only a single case, REF-4383-0005, carried a *NOTCH1* mutation (PEST domain mutation in exon 34, VAF 29%) together with two hotspot *NRAS* variants at codon 61. The latter are either present in the same clone as biallelic variants or raise the possibility of the

presence of two separate clonal B cell populations. Vendramini *et al.* showed a high incidence of activating RAS/MAP kinase pathway variants in CLL with trisomy 12 (Vendramini et al., 2019). The case presented here fit their result in that it carries an unmutated *IGHV* rearrangement, however in Vendramini's study presence of Trisomy 12 and *RAS* mutations was mutually exclusive of *NOTCH1* mutations in contrast to the case found here. When looking at disease course, CLL with trisomy 12 and *KRAS* or *NRAS* activating hotspot variants showed a shorter TFS and *RAS* variants have been associated with poorer responses to immunochemotherapy and development of resistance to treatment (Vendramini et al., 2019).

A recent publication by Degaud *et al.* characterised a group of patients sharing several characteristics with these cases identified in the ENABLE study: The group published a detailed clinical and molecular analysis of a cohort of 18 patients who had been diagnosed with unclassifiable isolated monoclonal lymphocytosis (Degaud et al., 2019). Their patients carried features of both of CLL and MZL, had a CLL score below 3, strong CD20 staining and CD43 negativity. Morphologically all samples had CLL-like lymphocytes, but all also included a component of atypical lymphocytes. Trisomy 12 was present in 17 out of their 18 patients. CD5 was positive in 15 cases. In a few cases del13q or del17p were present but importantly no case showed trisomy 3 or 18 and there was no evidence of del7q making a diagnosis of MZL less likely. As was the case in this study, one case also had an *IGH-BCL3* translocation but *NOTCH1* variants and *MYD88* hotspot variants were not detected amongst the cases positive for trisomy 12. In contrast to our findings, all cases either used *IGHV3*- or *IGHV1*-genes with no *IGHV4-34*. The authors concluded that these were examples of either atypical CLL or CBL-MZ with a diagnosis of CD5+ve-MZL (early phase without organomegaly) another, less likely, possibility.

Based on the additional findings in our cohort with trisomy 12, the two cases with the translocations can be assigned a diagnosis of Follicular lymphoma and, likely atypical CLL (discussed above). With regards to the SNVs, the presence of *NOTCH2* mutations together with *BIRC3* would be compatible with a diagnosis of MZL, likely SMZL or more appropriately CBL-MZ as RHU-4383-0001 presented without symptoms, splenomegaly or lymphadenopathy. Case RVR-4383-0013 with *KMT2D* and *FBXW7* variants and atypical morphology more likely fits into the atypical CLL

category rather than MZL given the prevalence of variants in these genes in CLL compared to MZL.

Two cases stand-out as they are both CD5-negative: one has a rare *IGHV6*-1 and the other an *IGHV4-34* rearrangement both with 0% SHM. The latter carries a *BIRC3, CXCR4* and *ARID1A* variants but morphologically shows villous cytology ('hairy' cells). Given the presence of an *ARID1A* splice variant and a *CXCR4* nonsense variant together with the morphological features would support a diagnosis of either LPL or SMZL. The other case shows no specific morphological features, has a CLL score of 0 together with a *ARID1A* and an activating *BRAF* variant. CNV analysis suggested the presence of a deletion in 7q involving *POT1* but not *EZH2* or *BRAF*. Differential remains between CBL-MZ and a lymphoma of MZ-type.

3.9 Case studies

Case study 1: BEL-4383-0001

The patient is a 75-year-old female who was found to have a lymphocytosis but was asymptomatic at the time of diagnosis. On examination she had no B-symptoms, no organomegaly and no enlarged lymph nodes. CT showed borderline splenomegaly (13.5 cm) but no lymphadenopathy. The full blood count provided in table 3.10 is from the time of enrolment into the ENABLE study in September 2017. Based on bone marrow investigations the patient was diagnosed with a CD5-negative, CD10-negative B-NHL. Morphology noted the presence of small to medium-sized lymphocytes with mature nuclei and clefted nuclei in a proportion of cells. Serum electrophoresis did not demonstrate a paraprotein.

Table 3.10: Full blood count values at study enrolment in September 2017 for case BEL-4383-0001. The exact reference values were not available and therefore average published reference ranges are provided instead (bottom row). Hb- Haemoglobin, WCC= white cell count

Hb g/dL	WCC $x10^9$/L	Platelets $x10^9$/L	Lymphocytes $x10^9$/L	Neutrophils $x10^9$/L
13.9	17.2	170	13.7	2.5
11.5-16.5	4-11	100-450	1.5-4.5	2-7.5

NGS analysis found two potential driver variants, a missense change in *TP53* (c.707A>G p.(Tyr236Cys)) and a potential splice variant in *KLF2* (c.76-1G>A p.?). The *TP53* variant lies within exon 7 which encodes the DNA-binding site of *TP53* and is a known recurrent cancer hotspot (IARC database https://p53.iarc.fr/ accessed 3.3.2021). The change has been reported in the context of Li-Fraumeni syndrome as a germline variant and as a somatic change in solid cancers in addition to haematological neoplasms (cbioportal https://www.cbioportal.org/ accessed 3.3.2021). By comparing the germline to the somatic samples for this patient the somatic origin of this variant is confirmed. The variant is present with a low VAF of 8% compared to the *KLF2* variant with 29% VAF which suggests the *TP53* variant has been acquired later and is present in a subclone. The *KLF2* variant is a potential splice site changing the recognition site AG into AA as shown in figure 3.20. The different algorithms predict the loss of the 3'-splice acceptor site at the start of exon 2. It is likely that this will alter the mRNA transcribed by either omission of exon 2 or activation of an alternative splice site for example. This could result in the loss of a relevant functional protein domain or lead to nonsense-mediated decay. As *KLF2* is a tumour suppressor, loss-of-function mutations are the expected mechanism.

Fig. 3.20: **Splice change predictions by different algorithms accessed through Alamut Visual 2.11**. The variant found in this sample was entered into the Alamut Visual 2.11 software. The top half shows the germline or reference sequence with exon 2 highlighted in blue. Hits from the different algorithms show the high confidence in the detection of a 3' acceptor site (green bars underneath the reference sequence). The bottom half shows the mutated sequence with the nucleotide change from G to A at the acceptor site. All prediction models agree that the acceptor site has been lost (no green bars against any of the algorithms) leading to an alteration of the mRNA.

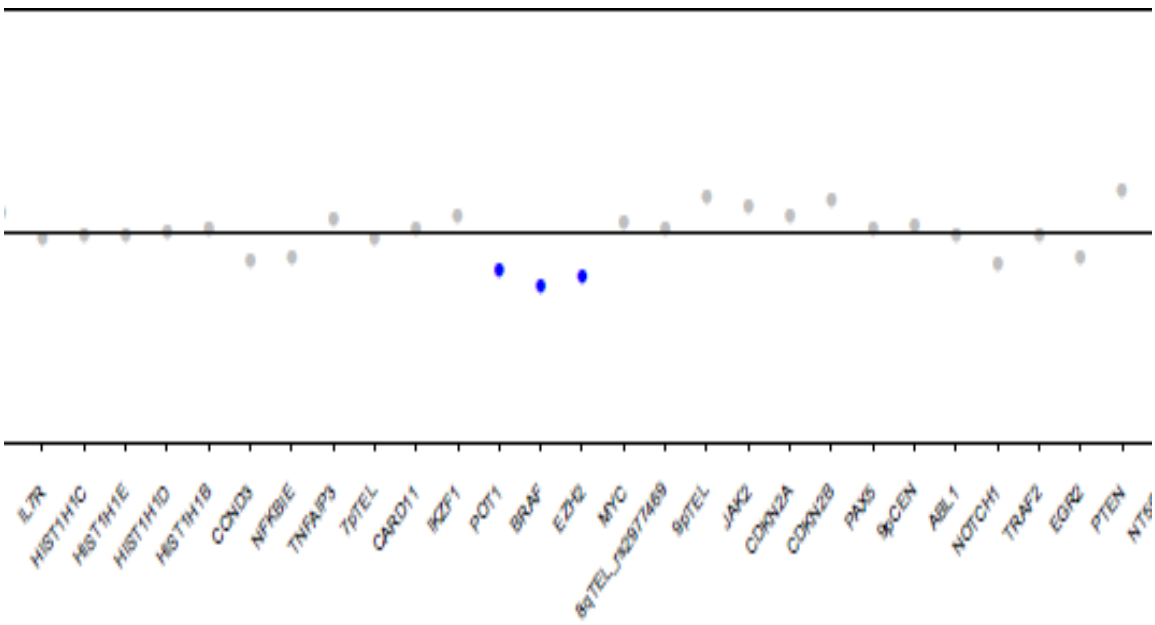


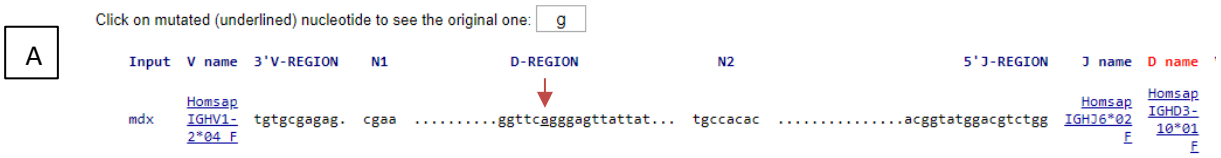
Fig. 3.21: Part of the CNV plot for case BEL-4383-0001 from the bioinformatics pipeline. Coverage for the targets *POT1, BRAF* and *EZH2* on chromosome 7q are below the threshold and are highlighted in blue to signal a potential deletion. The pipeline compares the coverage of a samples against a pooled normal sample.

In addition to the two potential driver variants, there is evidence of deletion of the long arm of chromosome 7, including the genes *BRAF, POT1* and *EZH2* (figure 3.21). No chromosomal rearrangements were detected within the scope of the NGS assay. For the clonal IG-rearrangements, the following assignments were made based on Sanger sequencing as well as the NGS data. The IG-heavy chain rearrangement is made up of the *IGHV1-2*04* gene, *IGHD3-10*01* gene and *IGHJ6*02* gene. The CDR3 amino acid sequence is 19 amino acids long (including the two anchors). Details are provided in figure 3.22.

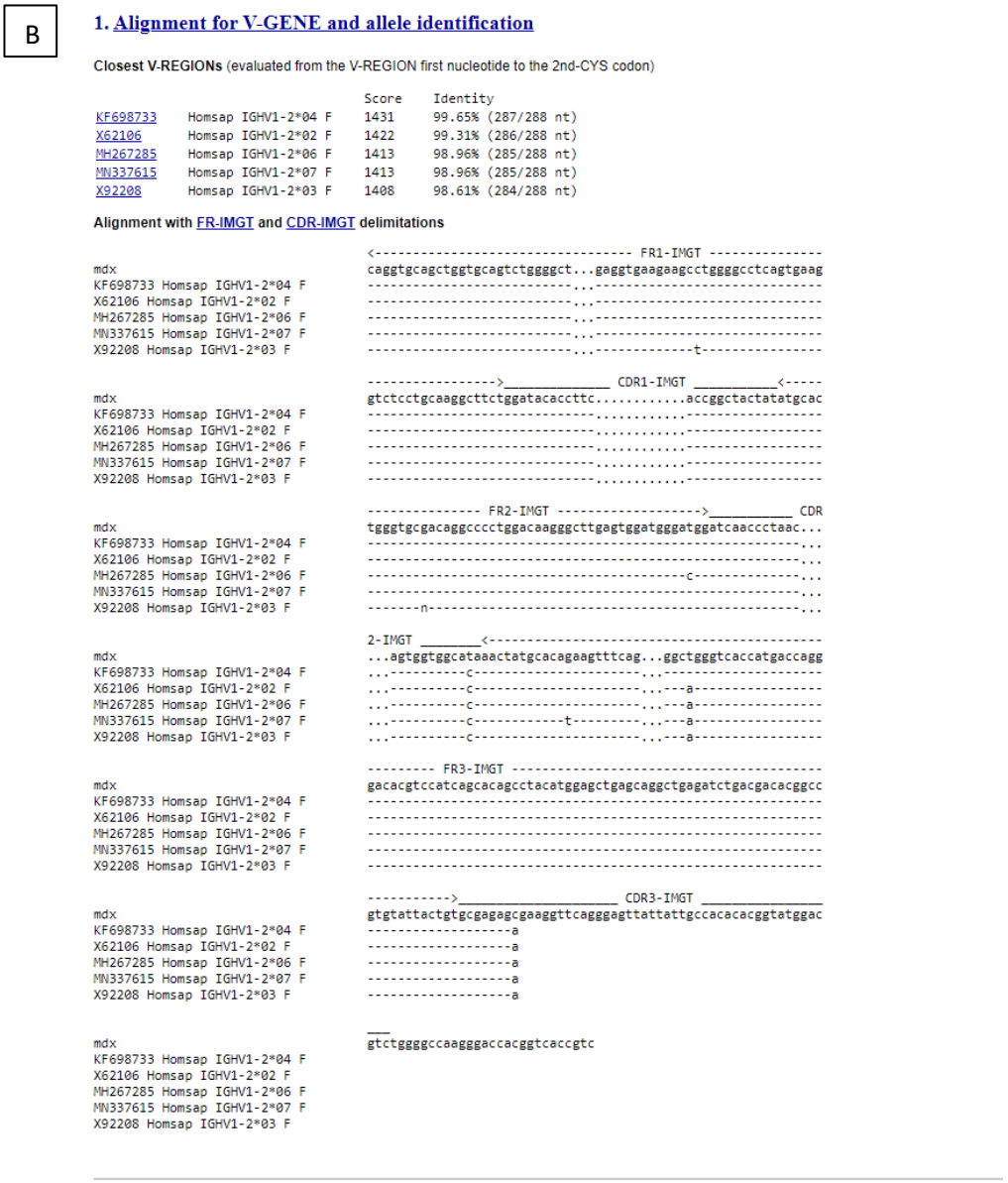
As shown in figure 3.22 there is a single nucleotide change in the V-sequence compared to the *IGHV1-2*04* germline sequence with one further G>A change within the CDR3 region (results output from IMGT database V-Quest). This means this case is minimally mutated (or borderline) according to Bikos *et al*. who found this to be the case in 71% of their *IGHV1-2*04* SMZL cases (figure 3.22 and 3.23 (Bikos et al., 2012)).

Fig. 3.22: Detailed IGHV rearrangement analysis. Panel A shows the nucleotide sequence across the CDR3 region as defined by IMGT V-Quest. The rearrangement is made up of *IGHV1-2*04* with a removal of 1 nucleotide at the junction and the addition of 4 nucleotides; the *IGHD3-10*01* gene is present with -10 nucleotides and -3 nucleotides at either end and the addition of 8 nucleotides in the junction with the *IGHJ6*02* gene. There is a deletion of 15 nucleotides from the J-gene (full description is (10)-1(4)-10(18)-3(8)-15(17)). The nucleotide that differs compared to the germline is underlined and highlighted with a red arrow (a) and the germline nucleotide (g) is displayed in the box above. Panel B shows the alignment of the V-gene sequence against the best matches obtained compared to the different *IGHV* germline references sequences within IMGT. Analysis performed with IMGT V-Quest. Panel C shows the amino acid sequence of the CDR3 region including the two anchors at 104 and 118. The change in the amino acid 110 from Arg (R) to Gln (Q) is highlighted.

A



B



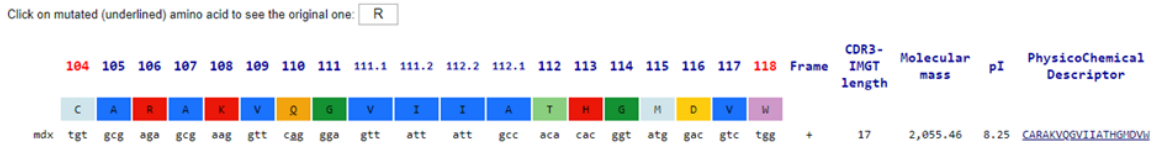


Table 2. Biased IGHV1-2*04/IGHD3-10 gene associations leading to restricted antigen-binding site motifs in SMZL

Germline	*V CAR*	*N1*	*D ITMVRGVII*	*N2*	*J*	
SMZLLS22	---	AV	----GV--	SRNGY	DAFDIW	IGHJ3
MZL114	---	GGL	---GV-	NLLSGGGDRDA	AFHIW	IGHJ3
SLVL27	---	GGV	--GV-	RSAFYSEGEGAK	FDYW	IGHJ4
SMZLLS24	---	GGV	--GV--	TSVLGVGQLYNT	DVW	IGHJ4
30041	---	GDV	--GV--	TVVRGSLT	HFDYW	IGHJ4
SLVL07	---	G	----GV--	SRLVGAEY	YFDYW	IGHJ4
30202	---	A	--QGVK	H	YYYGMDVW	IGHJ6
BEL01	---	AK	QGVII	ATM	GMDVW	IGHJ6

Abbreviations: IGHV, IG heavy variable; SMZL, splenic marginal-zone lymphoma. Shared amino-acid motifs were identified in the antigen-binding sites of IGHV1-2*04 rearrangements using the IGHD3-10 in RF3. Similar to IGHV1-2*04/IGHD3-3 rearrangements (Table 1), an 'IGHD-derived' GV dipeptide appears on the tip of the VH CDR3 loop in all cases (highlighted in light gray). Biochemically equivalent amino acids are highlighted in dark gray. Dashes indicate identity in keeping with IMGT/V-QUEST output format.

Fig. 3.23: **Comparison of the CDR3 amino acid sequence seen in BEL-4383-0001 to the biased gene usage and restricted antigen-binding site motifs as published by Bikos *et al.*(Bikos et al., 2012).** The table (table 2) is reproduced from the study by Bikos *et al*. which provides the most detailed insight into the homology of *IGHV*-rearrangements across SMZL patients to date. The clonal CDR3 in our patient shows a high degree of similarity to case 30202 in particular and importantly contains the "GV" motif that appears to be a defining feature of the antibody encoded by this subgroup of SMZL cases.

Case discussion

Studies published by Stamatopoulos *et al*. (Stamatopoulos et al., 2004), Bikos *et al* (Bikos et al., 2012) and Thieblemot *et al*. (Thieblemont et al., 2014) have demonstrated a skewed repertoire of *IGHV* and *IGK/LV*-gene usage in SMZL. *IGHV1-2* and in particular *IGHV1-2*04*, as seen in the case presented here, are found in 30-40% of SMZL but are uncommon in other B-NHLs even in the closely related NMZL where IGHV4-34 and V1-69 are the most dominant V-genes (Bertoni et al., 2018). In addition, in SMZL long CDR3 regions are seen (as in this case) with complementary features as shown in figure 3.23 with common motifs. In addition there is clear bias in use of light chains alongside this particular *IGHV* gene: *IGKV3-20> IGKV1-8 > IGL2-14* (Thieblemont, 2017). In case BEL-4383-0001, kappa light chain restriction was found with the productive rearrangement containing *IGKV1-8*. The clonal lymphocytosis seen in this patient thus has arisen from a B-cell which carries a particular IG-rearrangements that likely plays a part in the development of SMZL and it points towards a common antigen that is triggering the selection and proliferation of such clones in the pathogenesis of SMZL.

A recent systematic review of studies on SNVs in SMZL, confirmed the high prevalence of mutations in *KLF2, NOTCH2* and *TP53* in this subset of patients (Jaramillo Oquendo et al., 2019). The authors noted that the study published by Clipson *et al*. (Clipson et al., 2015) was the only one that reported on *KLF2* variants. Not all studies included this gene in their design, but Jaramillo Oquendo *et al.* postulated that studies may also have missed variants due to the high GC-content of the gene leading to suboptimal sequencing performance (Jaramillo Oquendo et al., 2019). They also highlighted that *KLF2* and *NOTCH2* variants are mostly mutually exclusive in SMZL but either can be see together with *TP53* variants, the latter occurs in around 15% of SMZL.

30-40% of cases with SZML have del(7q) and it is the most common chromosomal change found across all studies although the target gene has still not been identified. The majority of SMZL carry further chromosomal changes including gains of 3/3q, 9q,12q, and 18q, and losses of 6q, 8p, 14q, and 17p with >10% of cases presenting with a complex karyotype (Thieblemont, 2017). Although the full CNV analysis is pending the CNV traces for this case do not show additional prominent changes.

Given the lack of further evidence for a diagnosis of SMZL in this patient, this case may be better classed as a Chronic B-lymphocytosis with marginal zone features (CBL-MZ) as proposed by Xochelli *et al.* (Xochelli et al., 2014). Although it is now accepted that not all non-CLL MBLs fit this category and that there is a huge heterogeneity between patients, the case presented here does fit the molecular hallmarks of SMZL and it is likely correct to consider this as a CBL-MZ precursor to SMZL.

In a recent study by Parker *et al.,* 37 patients with CBL-MZ were investigated and followed-up for a median of 9.6 years. Only 9 cases progressed, amongst them one case with *IGHV1-2*04* (98.1% homology to germline) and, unusually, a *NOTCH2* and a *KLF2* mutation. Del14q was also present. This patient did progress to SMZL with a time to progression from diagnosis of 47 months which was the shortest timeframe of any of the cases that progressed in the study. Overall survival for this case was 51 months. An additional 5 out of the 9 cases progressed to SMZL but these did not carry the combination of IG-gene rearrangements and variants closely associated with SMZL.

The patient present in this case study has so far shown no signs of disease progression and has not received any treatment. She was last seen in clinic in August 2020, two years after enrolment. However, as exemplified in the study by Parker *et al.* (Parker et al., 2018) and supported by other studies looking at non-CLL MBL/CBL-MZ cases, the latency and time to progression is long with indolent disease courses in the majority of patients. Nevertheless, follow-up of this patient is indicated. Differences in the types and combinations of variants present in the B-clone plus external factors like the clonal microenvironment will likely play a yet uncharacterised role in defining the disease course in each of these patients. Nevertheless, this case does demonstrate the common features present in B-cell clones that give rise to clonal lymphocytosis and which have the potential to progress into overt neoplasia in a subset of patients. From a scientific and research point of view this case also highlights the wealth of information that can be obtained from a single test using specially-designed NGS-panels. Utilising such panels in routine diagnostics will enable clinicians and scientist to gather real-life outcomes in a prospective way improving on studies limited by small case numbers and divergent testing approaches. Ultimately, this will allow us to gain a better understanding of the likelihood of progression of individual patients and to identify the different evolutionary trajectories B-cell clones can take. This information will be useful to inform patient management and, eventually, treatment.

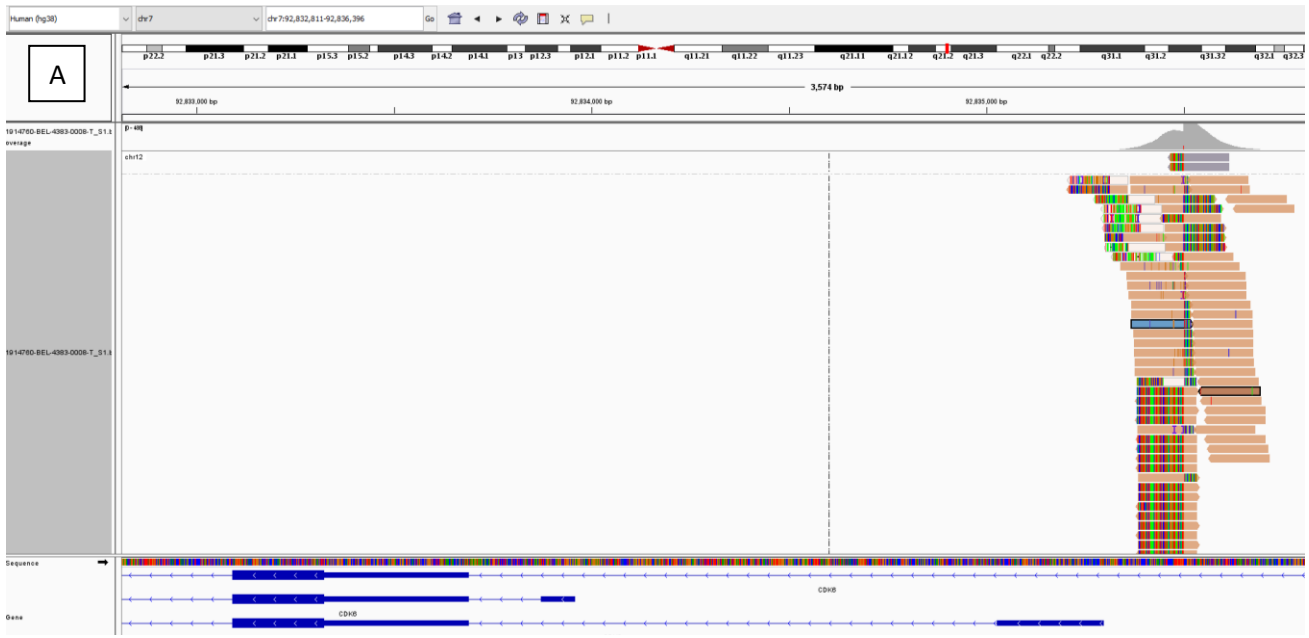
Case study 2: BEL-4383-0008

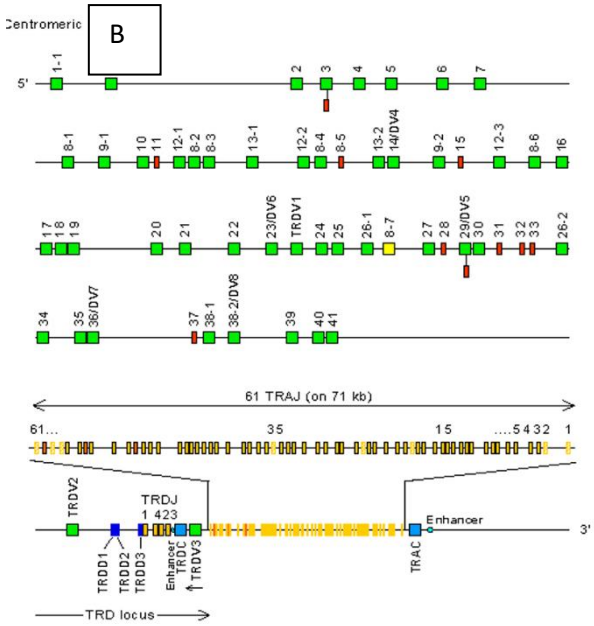
In December 2020 Gaillard *et al*. published their findings on a cohort of 57 cases with B-NHL carrying cyclin-dependant-kinase 6 (*CDK6)*-translocations (Gailllard et al., 2021). Amongst them were four cases with *CDK6*-rearrangements involving the *TRA/TRD* locus as the partner as seen in case BEL-4383-0008 in the ENABLE study (figure 3.24). The case also had a frameshift *TP53* variant, *TP53* c.420del p.(Cys141AlafsTer29) with a VAF of 46% together with del17p (figure 3.25). The productive *IGHV*-rearrangement identified by NGS *was IGHV3-23/IGHJ5* together with *IGKV1-16/IGKJ4* (kappa light-chain restricted). No further clinical information was available at this point.

Translocations involving *CDK6* are rare and most commonly arise from the *IGK* locus to form a t(2;7) translocation. Unusually in this case, it is not one of the IG-loci but the TRA/D locus that is involved suggesting that despite the B-cell origin double-strand breaks occurred here; potentially through the initiation of TR-rearrangements mechanisms and activity of recombination-activation genes (RAG) and /or terminal deoxynucleotidyl transferase (TdT) as has been described for cross-lineage rearrangements in ALL and CLL for example.

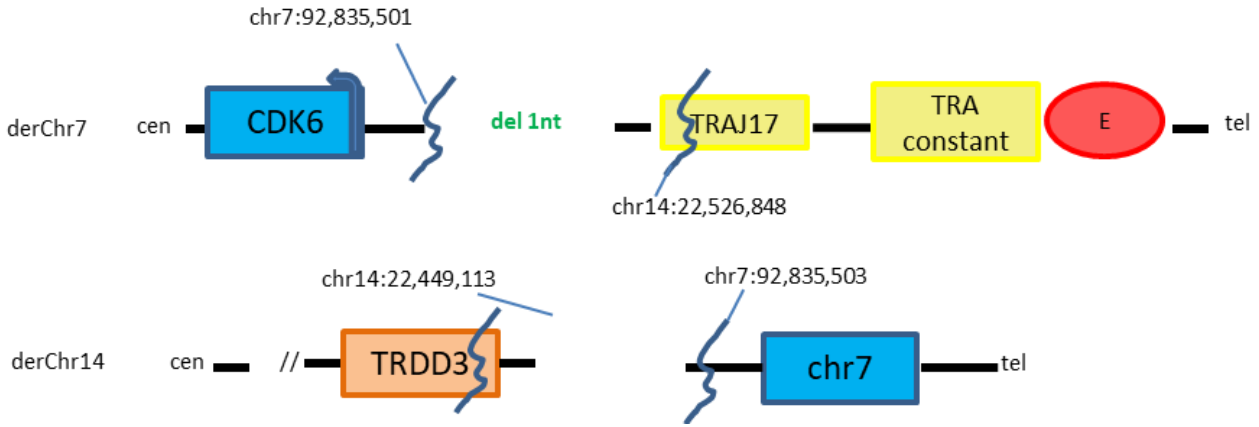
In Gaillard's study, two of the four cases with *TRAD/CDK6* translocation fitted the CLL/SLL/MBL category and two the MZL/small B-Cell-lymphoma category based on morphology, *IGHV* usage and additional variants and chromosomal changes found. All were CD5+ve. One of the two SZML cases had a *TP53* mutation and deletion of 13q14, the latter was also present in both MBL cases with *CDK6*-translocations. Both MBL cases had *IGHV4-34* usage and one had a complex karyotype.

Fig. 3.24 (next page): Translocation between chr 14 *TRA/D* locus and chr 7 *CDK6* in case BEL-4383-0008. Panel A shows the reads aligned to chromosomes 7 all resulting from paired read partners in chromosome 14 (IGV screenshot). The break occurs close to the *CDK6* transcription start. The genomic layout of the *TRA* and *TRD* locus is provided (panel B) and a diagram of the fusion is displayed (panel C). Although not a legitimate D-J fusion, double strand breaks must have occurred at T*RAJ17* and *TRDD3* locus which then fused with chromosome 7 bringing the *CDK6* expression under the influence of the enhancer in the *TRA* constant region. This is a recently described fusion and no functional data is yet available. It cannot be excluded that there is an effect on a gene in chr7 now on derivative chr14 although the second enhancer witin the *TRA/D* locus is likely lost in this rearrangement as the break occurs centromeric in *TRDD3.* See chapter 3.6 for further details on the diagram shown.


A
3,574 bp
92,833,000 bp
92,834,000 bp
92,835,000 bp
chr12
CDK6


B
Centromeric
5'
61 TRAJ (on 71 kb)
61...
35
15
....5 4 3 2
1
TRDV2
TRDJ
1 423
Enhancer
3'
TRDD1
TRDD2
TRDD3
Enhancer
TRDC
TRDV3
TRAC
TRD locus

C


chr7:92,835,501
derChr7
cen
CDK6
del 1nt
TRAJ17
TRA
constant
E
tel
chr14:22,526,848
chr14:22,449,113
derChr14
cen
//
TRDD3
chr7:92,835,503
chr7
tel

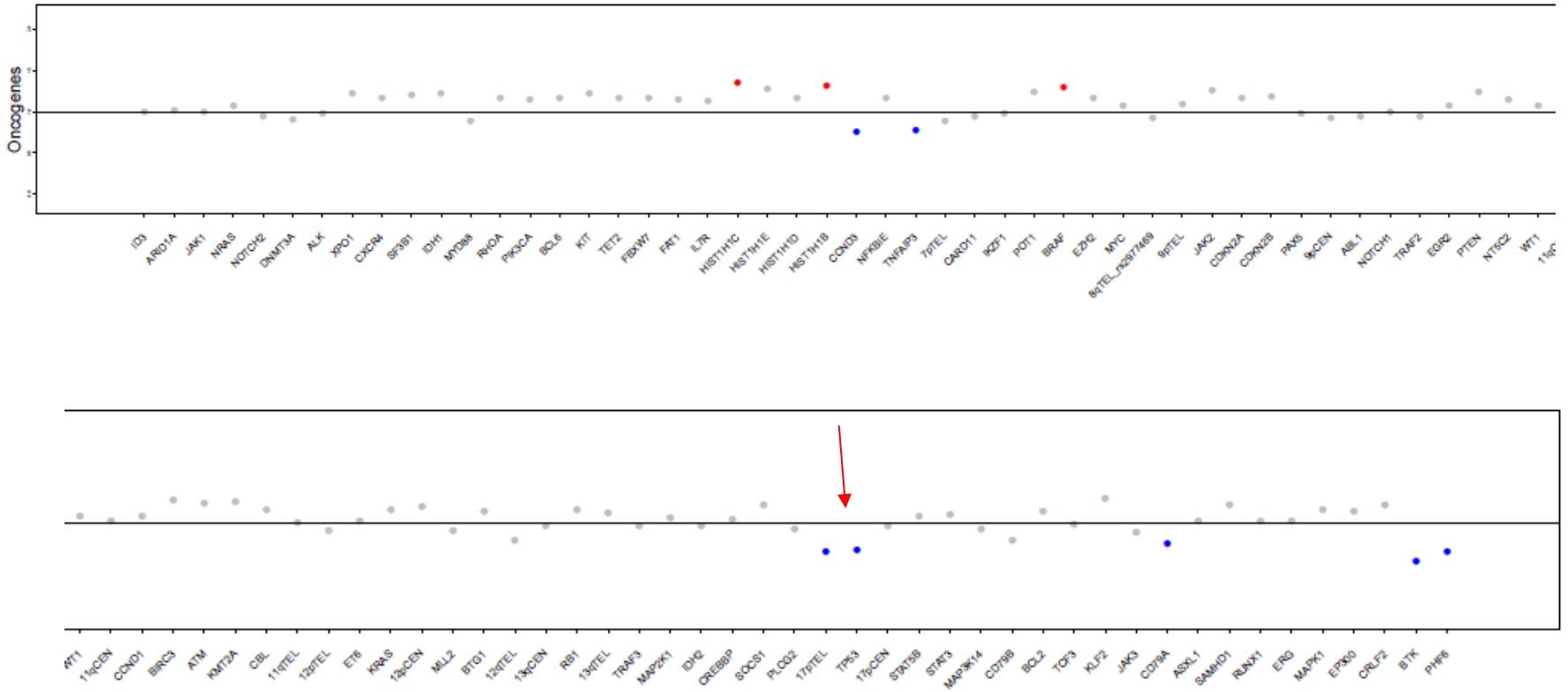


Fig. 3.25: CNV plot for case study 2. CNV plot shows *TP53* deletion (including telomeric but not centromeric areas, red arrow) and potentially *CCND3* and *TNFAIP3* deletions amongst others (further analysis pending). The del(17p) was also confirmed in the manual analysis (data not shown). CNVPanelizer of the bioinformatics pipeline uses a threshold of 0.8 to indicate loss (highlighted in blue) and 1.2 to indicate gain (red dots).

In their full cohort, 55% of cases had *TP53* changes and the majority had mutations and deletions. The most frequently used *IGHV* was *V3-23*. Review of their cases assigned the majority to MZL with a strong suggestion of SMZL. However, the *CDK6*-translocation positive cases with MZL phenotype had specific characteristics: They were CD5+ve, showed presence of prolymphocytic cells and most cases had *TP53* mutations together with deletion of 17p. Based on the few cases identified so far, prognosis appears to be good with an indolent disease course and an OS at 5 years of >80% although given the very low samples numbers further studies are needed to confirm this.

It is postulated that the rearrangement results in overexpression of *CDK6* driven by the control of the TR-locus enhancer that has been brought into close proximity to the start of the open reading frame of *CDK6*. However, expression studies have not yet been performed on the samples included in the original publication nor on the ENABLE case. FISH studies are pending to confirm the *CDK6* break in BEL-4383-0008 and the possibility of IHC studies is being explored. Overexpression of CDK4 and CDK6 is found across haematological malignancies leading to defects in cell cycle and transcription regulation and CDK4/6 inhibitors have been developed and are being used in clinical trials (Nebenfuehr et al., 2020). If CDK6-overexpression is confirmed, targeted therapy may be an option for these cases in the future if required.

4. Discussion

4.1 The Next-generation sequencing approach

Our understanding of the complexity of molecular changes implicated in the development of LPDs has increased dramatically in recent years and some of these findings have been incorporated into the latest updates of the WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue (table 1.1 (Swerdlow et al., 2016a)). Translating this into the clinic requires an expansion in the number and a broadening of the scope of tests performed which increases the cost to the health service but can also extend the time it takes until these results are available to the clinician. In the context of lymphomas biopsy material can be limited restricting the number of the test that can be performed. Recent advances in technology, like for example next-generation sequencing panels, can help diagnostic laboratories adapt to meet the growing demand in molecular and cytogenetic testing.

The cohort of patients in this study will have undergone investigations to exclude the presence of a *bona fide* B-NHL as classified by the WHO. Given the expected indolent nature of their disease, the likely frailty of the patient group making bone marrow sampling difficult and the lack of clinical symptoms requiring treatment in most cases, no further investigations will routinely be performed. There is often no direct clinical need as results would not change clinical management and comprehensive evaluation using molecular and cytogenetic approaches is currently not cost- nor time-efficient. This does mean, however, that our understanding of the large group of B-NHL, NOS and their precursor lesions is currently limited and patient management and treatment, if required, is not standardized. Furthermore, patients in this cohort who do require treatment have no access to clinical trials as these often have specific diagnoses as eligibility criteria.

The aim of the EuroClonality-NGS DNA Capture (EC-NDC) panel approach was to design a tool that would be able to analyse the broad spectrum of molecular markers required for the investigation of possible or confirmed B- and T-cell lymphomas (Wren et al., 2014 on behalf of the EuroClonality NGS workpackage). For this it needed to be able to identify clonal IG- and TR-rearrangements as well as detect the various classes of molecular changes implicated in diagnosis, prognosis and

therapeutic evaluation of lymphomas in clinical practice: SNVs, CNVs and chromosomal translocations (Wren et al., 2017, Stewart et al., 2019). Several other groups and commercial companies have developed NGS-based capture approaches to aid the expansion of molecular testing in routine laboratories for haematological cancers. A detailed literature search and discussion of this topic was presented as part of my C1 Innovation Project and the reader is referred to this for further details.

Due to its design the EC NGS capture panel is an excellent tool to analyse in a broad, non-targeted way, different molecular changes associated with clonal B-cell lymphoproliferations to characterise a patient cohort that would normally undergo limited molecular testing. As the investigation is a single laboratory set-up, time and expenditure is limited and this could easily be transferred into clinical practice if deemed clinically relevant.

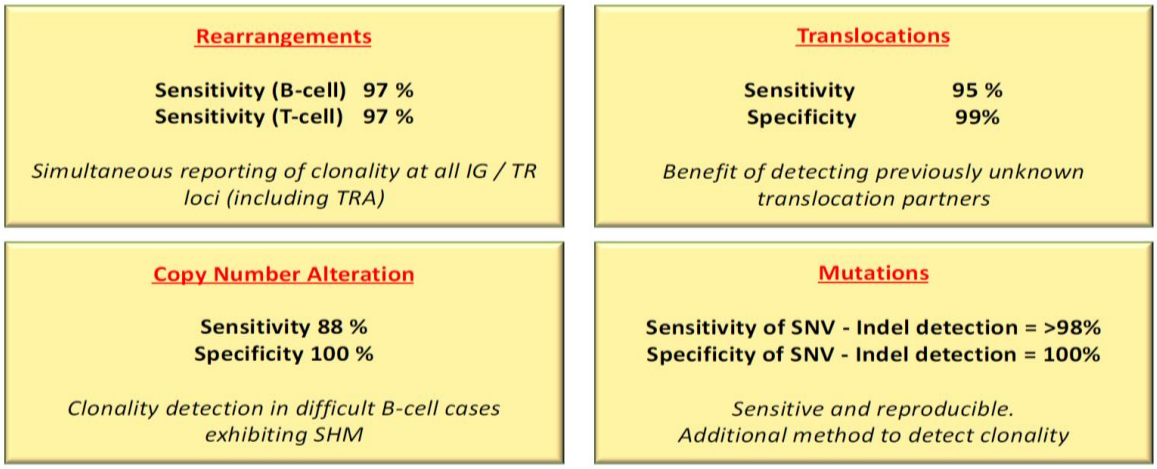


Fig. 4.1: Performance of the EuroClonality-NGS DNA Capture Panel compared to SOC testing results as presented at ASH 2019 (taken from (Stewart et al., 2019)). The EuroClonality NGS consortium has since completed a Pan-European validation study across 7 diagnostic laboratories, analysing 280 samples from B- and T-cell lymphomas (Stewart et al., 2019 and Stewart et al., 2021).

The full validation has now been published and the panel performance is therefore not be discussed further here (*(Stewart et al., 2021)*). Concordance with standard-of-care (SOC) testing for all aberrations was very high (figure 4.1) with the exception of del(13q) which had a lower pick up rate by NGS compared to FISH due to the NGS design only including the *RB1* locus but not *DLEU1* resulting in false-negative calls by NGS. A re-design of the NGS panel to include more of the relevant genomic regions on 13q will likely be able to improve the pick-up rate.

The laboratory work for this study was performed at the Clinical Genomics laboratory at the Royal Marsden NHS Foundation Trust, one of the seven laboratories involved in the proficiency and validation study performed by EuroClonality-NGS and was run in accordance with the QC-performance criteria stipulated in the proficiency runs (see supplementary information in (Stewart et al., 2019)). The only difference was that the design of the ENABLE study panel did not include the V-genes of the *IGH*-, *IGK*- and *IGL*-loci and paired-end read lengths were 75 bp instead of 150 bp for the NextSeq runs performed. The samples of the EuroClonality-NGS validation study have been re-run at 75 bp in the lead laboratory with identical results confirming that the approach here was appropriate for the detection of clonality, SNVs, CNVs and translocations (unpublished result, personal communication Prof. D. Gonzalez).

4.1.1 Utility of the germline sample for NGS analysis

In contrast to the validation study, a germline sample was run alongside each patient's tumour sample in the ENABLE study to allow for the subtraction of constitutional variants present in the germline of a patient. Germline samples were obtained by CD15+ve cell-selection (neutrophils). Samples were run at lower depth compared to the tumour sample as only heterozygous/homozygous changes with minimum VAF of 50% are expected and a lower coverage reduces the cost and time required for bioinformatic analysis. This approach will, however, remove any variants that have arisen in a non-lineage specific precursor. Variants associated with clonal haemopoiesis of indeterminate potential (CHIP), like for example *SF3B1, TET2* and *DNMT3A*, could have been removed as 'false' germline variants if they were present in the neutrophil compartment. This is expected to be a rare event and the benefit of reducing the complexity of the variant analysis by reducing the number of calls and removing likely-benign, private SNVs is helpful and can avoid overinterpretation of changes not implicated in the pathogenesis. Following variant analysis according to published guidelines (Li et al., 2017, Froyen et al., 2019), only variants assigned as pathogenic or likely pathogenic ('driver variants') as well as clinically-actionable variants were reported to try and identify those variants likely implicated in the aetiology of the clonal B-cell lymphocytosis.

The pipeline utilised for CNV analysis cannot at present compare germline and tumour samples and it will be helpful to continue comparing the CNV manually to ascertain that none of the changes observed are present in the germline of a patient.

As no positive selection for B-cells was performed for the tumour sample we cannot be completely certain that the variants detected reside in the clonal population of interest, especially in cases where the lymphocytosis or percentage of clonal B-cells in the sample analysed may have been low. No evaluation of the infiltration level was performed on the ENABLE sample at enrollment which means that the exact percentage of clonal B-cells in a given sample is not known.

The purity of the germline sample was assessed during the validation but not routinely for each of the samples processed. It is therefore potentially possible that samples lacking driver variants may be false negatives if the germline sample was contaminated with enough tumour material to remove variants detected in both samples. Given that the tumour and germline variant analysis was processed through the pipeline separately with a semi-manual excel-sheet based process for sorting and comparing variants, an unusual VAF of variants in the germline sample (not 50% or 100%) would likely have been spotted. In addition, all germline samples underwent analysis for IG V(D)J-rearrangements and no case was found that showed a clonal IG-rearrangement in the germline sample or indeed showed the same rearrangement as was present in the clonal tumour samples. Given the sequencing coverage obtained for the germline sample, a medium/high level of contamination above 20% or more can therefore be ruled out with confidence whereas low-level contamination would remain a possibility. Nevertheless, if it were deemed necessary to run paired samples for routine diagnostics, a more thorough method to assess for germline contamination should be employed for each sample together with bioinformatic evaluation of variant VAF and type of variants encountered.

Apart from *TP53*, the scope of the panel does not include genes implicated in cancer predisposition syndromes and incidental findings unrelated to the condition being investigated are therefore not expected. Of note, the consent did not include the analysis of germline variants and therefore any *TP53* variants present in the germline would not be investigated nor reported.

4.1.2 Evaluation of performance: Detection of clonality by NGS

The panel had a high success rate in identifying clonal markers. NGS detected evidence in one or more IG-targets in all 97 samples for which data was available. In 95 out of the 97 samples, clonal rearrangements at the *IGH*-locus as well as the light-chain loci were detected and in the two cases without *IGH*-clonal rearrangement the presence of a clonal B-cell population was confirmed by *IGK/IGL* rearrangements. Importantly, the identification of *IGHV*-rearrangements was highly concordant with the current gold standard of VLeader Sanger sequencing. However, given the differences in read length (75 bp instead of 150 bp) and due to using D-J-baits only, the assignment of the V-genes was less robust in the NGS analysis compared to VLeader and the results from the EuroClonality NGS DNA capture panel validation. The precise V-gene usage is not required for clonality assessment and, as shown by the validation study, can be improved by longer read lengths and the inclusion of V-gene baits in the design of the panel.

As shown in the EuroClonality capture NGS panel validation and seen here for the assessment of B-clonality, the EuroClonality capture NGS panel is capable of replacing the Biomed2 PCR for clonality assessment which is currently the gold standard for clonality assessments in mature B- and T-cell proliferations (Langerak et al., 2012, van Dongen et al., 2003). The data includes all the details required for the identification of potential measurable-residual disease (MRD) monitoring markers which could allow the design of patient-specific assays if IG-based MRD-evaluation becomes routine for CLL and beyond (Brüggemann et al., 2019, Chase and Armand, 2018).

In the context of the ENABLE study, the presence of a clonal B-cell population was already known through previous work-up including immunophenotypic analysis when the patient first presented. Pragmatically, the identification of the clonal IG-rearrangements by NGS ensures that the sample contains a high-enough proportion of clonal B-cells thus making the NGS findings more reliable.

Additional clonal markers in the form of SNVs, CNVs and translocations were detected in most samples (81% currently). Apart from 19 samples in which the CNV analysis is still pending all samples had either a translocation, likely pathogenic driver variant or a CNV in regions associated with lymphoid disease (Swerdlow et al., 2016).

4.2 *IGHV*-gene usage and Stereotyping analysis

Although the NGS was capable of identifying clonal IG-rearrangements in all samples, the design of the panel used did not allow for the analysis of SHM-status and stereotype of the *IGHV*-rearrangement. The VLeader Sanger sequencing was therefore run in parallel to provide this data and it showed a high rate of concordance between the *IGHV*-rearrangements identified by NGS and the VLeader assay, including the CDR3 amino acid sequences.

Stereotyping analysis proved that the etiology of the clonal B-cell lymphoproliferations seen in this cohort of patients differ from CLL as none of the *IGHV*-rearrangements matched the 18 stereotypes described in CLL (Rossi and Gaidano, 2010, Ghia et al., 2007). Further analysis could be done to compare the *IGHV*-sequences with other stereotypes reported for MCL (Hadzidimitriou et al., 2011) and SMZL (Zibellini et al., 2010) for example and to investigate whether there are any commonalities between cases e.g. between cases of the large group with *IGHV4-34* usage (Bystry et al., 2015).

In one case, BEL-4383-0001 the highly specific *IGHV* used (*V1-2*04*) directly supports a diagnosis of a SMZL (see case study 1, chapter 3.9). For other cases, as discussed in more details in the results chapter and below, the specific *IGHV*-gene used together with the mutational status can lend support to a particular diagnosis within a differential if one *IGHV*-gene is more prevalent in one than the other (Zibellini et al., 2010, Gemenetzi et al., 2020).

4.3 Evaluation of performance: Detection of variants by NGS

4.3.1 Detection of translocations

FISH confirmed the presence of the translocations detected by the EuroClonality capture NGS in all cases with t(14;18) and so far for three cases with t(11;14); for two the analysis is pending and for two FISH did not show the presence of a t(11;14) (see table 3.2 and chapter 3.6.3). It is known that cryptic rearrangements or rearrangements with unusual breakpoints may not be detectable by FISH analysis (Peterson et al., 2019, King et al., 2019). *CCND1* expression can be reviewed although overexpression of *CCND1* in B-NHL has been described in the absence of t(11;14) thus IHC and RNAseq evaluations would not be ideal for confirmation of the

rearrangement seen (Hill et al., 2020). Alternative technologies like Bionano Genome Mapping (Bionano® Genomics, San Diego) or long-read sequencing (e.g. Oxford NanoPore technologies) may be helpful to fully identify the chromosomal change present in this case.

For the rare case with the *CDK6-TRA/D* fusion it would be useful to demonstrate upregulation of CDK6 by IHC for example as this had not yet been proven by the study that first published this translocation in 4 cases of SMZL in December 2020 (Gaillard et al., 2021). FISH probes are available but are not routinely used in diagnostic laboratories and the fusion has yet to be confirmed for the case found in the ENABLE study.

Additional cases with unusual rearrangements were identified and are awaiting further evaluation but many studies have now confirmed the suitability of capture-based NGS approaches for the detection of translocations including those arising from IG- and TR- loci (for example see (Szankasi et al., 2019, McConnell et al., 2020, Chow et al., 2017, Cuenca et al., 2020, Yasuda et al., 2020)).

4.3.2 Detection of SNVs

Clinically relevant variants like *MYD88* and *TP53* in the MCL cases were confirmed on United Kingdom Accreditation Service (UKAS)-approved, routine NGS-panels run by the Clinical Genomics laboratory at RMH. Larger panels allow for the detection of variants that lie outside the common regions or variant hotspots. For example, two variants were detected in *MYD88* in case RMH-4383-0004, both likely pathogenic with good supporting evidence for pathogenicity. Parker *et al.* and Shuai *et al*. reported on the distribution of *MYD88* variants whereby Leu265Pro hotspot variants were the predominant one and in some studies sole change found in LPL/WM patients whereas other B-NHL entities with a lower frequency of *MYD88* variants overall had a much higher prevalence of *MYD88* variants other than Leu265Pro (Shuai et al., 2020, Parker et al., 2018). Although not discussed in this thesis, *BRAF* variants outside the Val600 hotspot were detected in two samples and the variants detected were likely activating (gain-of-function) variants like Val600.

NGS-panels can easily be adapted to incorporate further genes of interest which is essential especially if they are to be used in routine diagnostics. For example,

PTPRD has now been shown to be prevalent in NMZL and should thus be included in the next design of this panel (Bertoni et al., 2018).

Many of the patients included in this cohort will remain on a watch-and-wait approach with regular follow-up and review but without the need for treatment. Serial samples could be obtained in these cases to assess clonal evolution, which may help elucidate further which patients are going to progress and whose condition remains indolent for a long time. For such studies the level of detection of the NGS panel may need to be re-assessed if smaller subclones need to be detected.

Only either peripheral blood or bone marrow samples were analysed in this study for each of the patients. No lymph node or spleen-tissue specimen were tested in cases with lymphadenopathy or splenomegaly. It is therefore possible that different clones with other or additional variants are present in these patients and that such clones may not be circulating but may be selected for should treatment be initiated, or they develop a growth and/or survival advantage over other clones by acquisition of additional aberrations.

It is also important to note that this technique cannot confirm that the variants seen in the same patient are present in the same B-cell clone. It is theoretically possible that several clones exist, each carrying separate genetic changes. Some changes may arise in early stem cells and variants could be present across lineages (discussed for *SF3B1* in chapter 4.1.1) whilst several clones may arise in elderly patients with MBL or CHIP.

4.3.3 Detection of CNVs

So far only the presence of trisomy 12 has been confirmed for two cases (RHU-4383-0001 and BEL-4383-0002) as karyotype studies were performed for these prior to study enrolment. The analysis for CNVs is ongoing and other changes have yet to be confirmed by karyotype, FISH or array analysis. For B-NHLs presence of *TP53* mutations alongside del17p is a common finding and the high rate of case with this combination and the absence of *TP53* and del17p in other cases indirectly suggests the panel is performing well with regards to the detection of CNVs.

4.4 Comparison with other studies and relevance of findings

For the ENABLE study all cases with clonal B-cell lymphocytosis and a CLL score <4 that could not be classified according to WHO could be enrolled. Thus, the patient cohort will include non-CLL MBL-type cases with isolated lymphocytosis or presence of a paraprotein without other features right up to cases with lymphadenopathy, organomegaly and B-symptoms that fall into B-NHL, NOS.

Most published studies have focused on particular patient cohorts and have more stringent entry criteria often including CLL score of 2 or below and absence of splenomegaly, lymphadenopathy and cytopaenias. These studies therefore mainly include MBL cases and have helped increase our knowledge about the different subtypes and heterogeneity within the three MBL-categories now included in the WHO-classification (Swerdlow et al., 2016a). Examples are shown in table 4.1 (next page, references are: (Xochelli et al., 2014, Bruscaggin et al., 2014, Parker et al., 2018, Kalpadakis et al., 2017, Defrancesco et al., 2020, Degaud et al., 2019, Kalpadakis et al., 2019, Kostopoulos et al., 2017, Merli et al., 2019)).

Table 4.1: Overview of studies performed in indolent B-NHL and MBL cases including in-depth characterisation and published in recent years. Cohort size, assays and cases incorporated are listed as are the main findings. FU= Follow-up

Study	Cohort	Investigations	Characteristics	Findings
Xochelli 2014	102	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM	MZ features on morphology/flow CLL <3 CD5-ve Asymptomatic No lymphadenopathy No splenomegaly	Proposal of CBL-MZ as disease entity Stable disease course for majority Progression: Splenomegaly (del(7q)/complex karyotype) V4-34, mutated Tris12, tris3, Query whether SMZL or SLLU und SDRPL
Bruscaggin 2014	16	NGS panel *NOTCH2, NOTCH1, BIRC3, TNFAIP3, TRAF3, IKBKB, MYD88, CD79A, CD79B, CARD11, BRAF* and *MAP2K1*	As above	*MYD88* variants were associated with IgM paraprotein NOTCH2 V4-34, mutated FU mean 44 months 3 progression with splenomegaly
Parker 2018	37	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM Large NGS panel	CD5-ve MZ features on morphology/flow	FU mean 9.6 years: 76% stable disease Progression to MZL, 5 with diagnosis of SMZL *MYD88* most common variant
Kalpadakis 2017	53	*MYD88* L265P	CBL-MZ as above	10 cases with *MYD88*, all with IgM paraproteinaemia Differential between CBL-MZ and LPL/WM
DeFrancesco 2020	28	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM NGS panel	Non-CLL MBL (CD5+ve n=8; CD5-ve N-=20) CD5-ve cases with MZL features CLL<3	HIGH prevalence of *MYD88*, mostly non-hotspot FU mean 35 months, 1 progression to SLLU V4-34, mutated No NOTCH2 variants
Degaud 2019	18	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM NGS panel	CLL <3 CD5+ Asymptomatic CLL-like and atypical lymphocytes in all cases	FU mean 48 months: 5 patients required treatment Trisomy 12 most common finding , 13q14 (9), *TP53* most common variant, no NOTCH2/*MYD88* hotspot variants
Kalpadakis 2019	100	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM NGS panel	CD5-ve LPD	FU mean 42 months 30 patients with signs of progression, 10 actively treated Splenomegaly-SMZL Lymphadenopathy-NMZL
Kostopoulos 2017	227	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM	All three MBL categories excluding low-count MBL	Median follow-up 76 months Cytogenetic changes seen least in non-CLL MBL Trisomy 12 seen across all three groups
Merli 2019 (ASH abstract only)	77	Morphology Immunophenotyping Cytogenetics *IGHV* usage/SHM NGS panel	Comparing CBL-MZ (CD5-ve, n=33) versus SMZL (n=44)	FU median 1.6 years; CBL-MZ 6 cases progressed (SMZL/NMZL) Full study not yet published

4.4.1 Relevance of testing: Detection of translocations

The detection of several well-defined translocations highlights the benefits of this panel being able to detect all the different translocations partners - including novel and unusual fusions (e.g. *CDK6-TRA/D*) - in the same assay making the work-up of a heterogeneous group of patients more straight-forward. No information was available as to the presentation of these cases, but it is likely that the presenting features did not strongly suggest the possibility of a MCL- or FL- associated disorder and thus these investigations may not have been performed at diagnosis.

Although both t(14;18) and t(11;14) have been detected in the blood of healthy individuals at low levels by RT-PCR (Szankasi et al., 2019), the NGS approach is not sensitive enough for the detection at such low levels. Together with the presence of lymphocytosis and a confirmed clonal B-cell population, the presence of a translocation is unlikely to be an unrelated, incidental finding.

IGH-CCND1 translocations are recurrent findings in studies investigating patients with non-CLL MBL or atypical MBL where they are associated with CD5+ve cases (Maitre and Troussard, 2019). Given the heterogeneity in presentation and the atypical features seen in non-CLL/atypical CLL MBL, Xochelli *et al*. strongly advocate routine testing for MCL- and FL-associated translocations in patients presenting with persistent lymphocytosis beyond 3 months especially in cases with CD5-positivity and atypical (non-CLL) features (Xochelli et al., 2017). Similarly, the proposal by Fang *et al.* to test all CLL for the presence of *IGH-BCL3* translocations due shorter time-to-first treatment and OS should be extrapolated to include the MBL presentations given the detection of two cases in our cohort of 100 patients (Fang et al., 2019).

Having an NGS-panel available to test for this in a non-target specific way will make this feasible and cost-effective. It will also reduce the tailoring of testing based on specific parameters at presentation, allowing a blanket-test for patients presenting with clonal B-lymphocytosis.

Furthermore, case REF-4383-0003 with a t(11;14) exemplifies the ability of this technology to detect rearrangements with unusual breakpoints which may be missed by FISH and which would not be amenable to detection by RT-PCR, as the PCR-design is breakpoint specific (Espinet et al., 2008, Pulsoni et al., 2020).

By using the sequencing approach, the exact breakpoint of a translocation including the insertion/deletion events can be mapped which enables the identification of patient-specific MRD targets akin to IG/TR-rearrangements (Pulsoni et al., 2020). Scientifically, we gain further details about the mechanisms underlying the rearrangement event including the recognition sequences, increasing our understanding of disease aetiology.

4.4.2 Relevance of testing: Detection of *IGHV*-rearrangements

Activation of and signalling through the B-cell receptor (BCR) has been shown to be essential for survival in mature B-cells but also plays a role in the initiation and driving of neoplastic development (Casola et al., 2019, Gemenetzi et al., 2020). Based on the somatic hypermutation status two types of CLL are recognised: Unmutated-CLL and Mutated-CLL (Delgado et al., 2020). They display different disease outcomes and together with the identification of specific *IGHV*-stereotypes, remain one of the most important prognostic factors in CLL with unmutated *IGHV* being associated with poorer prognosis regardless of treatment regimens (Delgado et al., 2020). Antigens involved in the initiation of lymphoma can be external as is the case for many of the MZL (e.g. *Hepatitis C* infection in SMZL/NMZL and *Helicobacter pylori* infection in MALT (Bertoni et al., 2018)) but often involve autoantigens including those present on the B-cell surface (Casola et al., 2019, Singh et al., 2020). Self-reactive BCR are related to *IGHV4-34* gene usage as its antigens include human erythrocyte I-i polysaccharide antigen and human B lymphocyte CD45 O-linked polysaccharide antigen (Wang et al., 2017). Self-reactive cells need to escape control mechanisms in place to remove them and there is now evidence that presence of *CD79B* variants in combination with *CARD11* variants as well as cases with combination of *MYD88* and *CD79A/B* variants provide a way to prevent apoptosis and allow for survival and proliferation of self-reactive B cells (Wang et al., 2017, Singh et al., 2020). Research into the antibodies expressed by lymphoma cells in parallel to investigations into autoimmune disease will help us understand the commonalities (Singh et al., 2020). Given that the type of antigen plays a key role in survival, eradication of *Helicobacter pylori* is known to be able to treat and in some cases lead to regression of MALT/MZL lymphomas and a greater understanding of the antigens triggering the clonal B-cell receptors in B-NHL may allow for better risk stratification and, potentially, novel treatment approaches.

As shown in several examples in the study, *IGHV*-gene usage differs across disease entities and can, just like SNV detection, support a diagnosis or make another less likely (Gemenetzi et al., 2020). As is the case for the SNVs, however, there is overlap between various diseases subtypes and careful integrations with all available data is required to arrive at the most likely diagnosis.

4.4.3 Relevance of testing: Monoclonal B-cell lymphocytosis

Three types of monoclonal B-cell lymphocytosis (MBL) are now recognized by the WHO classification (Swerdlow et al., 2016a). The most common, and best characterised is CLL-MBL where the clonal B cells show a CLL-like phenotype (CD5+, CD23+. CD10+, CD200+) and molecular aberrations are consistent with those seen in CLL. For this entity the monoclonal B-cell population has to be $<5x10^9$/L with low count (LC) and high count (HC= 0.5–5.0x10^9/L) being distinguished with HC having a higher risk of transformation to CLL (Maitre and Troussard, 2019). As one of the entry criteria for ENABLE was a CLL score of 3 or lower, such cases will not have been included in this study.

The second type of MBL, termed atypical CLL-MBL, expresses CD5+ve but is usually CD23-negative. Cases show features of CLL and often have a CLL score of 3 with atypical morphology being highlighted. Strong association with trisomy 12 has been identified (Xochelli et al., 2017).

The third type is non-CLL MBL which does incorporate cases that are CD5-ve and have lymphoplasmacytic features and are referred to as CBL-MZ (Xochelli et al., 2014). This group is very heterogeneous clinically and remains the least characterized of the MBLs (Angelillo et al., 2018, Xochelli et al., 2017)

Non-CLL MBL includes cases which are often CD5-ve and have plasmacytic features including villi and an immunophenotype that suggested the B-cell clone is of MZ-origin. The term CBL-MZ for this group of patients was first proposed by Xochelli *et al*. (Xochelli et al., 2014). In a cohort of 102 patients presenting with CD5-negative clonal lymphocytosis with MZ features but no lymphadenopathy or splenomegaly they found recurrent molecular changes concordant with MZ: del7q, tris18 and tris 3. At the time no mutational analysis was undertaken. At 5-year follow-up, progression was only seen in 17% of patients and always involved splenomegaly

pointing to a potential diagnosis of SMZL or splenic lymphoma/leukemia unclassifiable (SLLU). They also investigated the possibility of this being a presentation of occult MALT lymphoma and although the majority of cases investigated showed sign of gastritis with 1/3rd of cases positive for *Helicobacter pylori*, a causal relationship could not be proven and only one case received a diagnosis of MALT .

In contrast to SMZL which shows a high prevalence of *IGHV1-2*04* gene usage (Bikos et al., 2012), this is rarely detected in cases that are part of MBL studies and this is concordant with our findings in that only 1/100 cases carried this *IGHV*-gene (see case study 1 on BEL-4383-0001). We detected no *MAP2K1* variants within our cohort which are seen in up to 50% of SLLU (Bruscaggin et al., 2014) making this diagnosis less likely. The study by Bruscaggin *et al.* reported on 16 cases with CBL-MZ showing a high incidence of *NOTCH2* variants but, crucially, their NGS-panel did not yet include *KLF2* in its scope.

Parker *et al.* published the study with the longest follow-up to date (mean FU 9.6 years (Parker et al., 2018)). Their cohort of 37 patients was well-characterised, including larger NGS-panel investigations and demonstrated a long latency and indolent course with progression rates of <20%. Case that did progress to overt disease were diagnosed with SMZL or MZL and *MYD88* variants were the most common finding across the whole cohort.

The clinical data, including evaluation performed at enrolments, was not yet available for most of the ENABLE patients. It is therefore not possible to compare how many patients were categorized as one of the MBLs at diagnosis, how many presented with splenomegaly or lymphadenopathy and how many developed these symptoms over time. Long-term follow-up will be able to address this aspect in detail and will add to the published data on this cohort of patients (table 4.1). Extrapolating from the few cases where clinical data was available, most cases had limited splenomegaly or lymphadenopathy and very few required treatment or indeed showed clinical progression of their disease so far but follow-up is short and the final analysis will be performed after a mean follow-up of 5 years. The extremely indolent course of many of the MBL-type cases has been highlighted across various studies (table 4.1 for relevant studies in this context).

4.4.4 Relevance of testing: *MYD88* variants

MYD88 is recurrently mutated in cases assigned to CBL-MZ, however the incidence rate differs between studies with a high rate seen in Kalpadakis *et al.* (Kalpadakis et al., 2017) and Bruscaggin *et al.* (Bruscaggin et al., 2014) but not by Xochelli *et al*. (Xochelli et al., 2014). DeFrancesco *et al.* reported a high incidence of non-Leu265Pro variants which have been associated with atypical CLL in other studies (Defrancesco et al., 2020, Rossi, 2014).

Our findings agree with a high prevalence of *MYD88* variants, predominantly Leu265Pro hotspot variants. Kalpadakis *et al.* reported a close association of *MYD88* with IgM paraprotein (Kalpadakis et al., 2017). Currently this information was not available for the *MYD88*-mutated cases but will be included in the long-terms study.

Initial studies investigating the prevalence of *MYD88* Leu265Pro variants across B-NHL showed an almost complete restriction to LPL/WM (Ondrejka et al., 2013), which was supported by data from Hamadeh *et al.* (Hamadeh et al., 2015). Hamadeh *et al*. investigated a cohort of 44 splenectomy-defined SMZL, 19 of which had plasmacytic differentiation and none carried the *MYD88* variant. However, further, larger studies reported the presence of *MYD88* variants in up to 15% of SMZL together with IgM and plasmacytic differentiation (Martinez-Lopez et al., 2015). The ongoing difficulty in differentiating between small B-cell lymphomas with plasmacytic differentiation was discussed at the 2014 European Association for Haematopathology/Society for Hematopathology slide workshop and although the value of *MYD88* testing was acknowledged and a high number of cases were assigned a diagnosis of LPL based on the presence of a *MYD88* variant, cases remained that do not fit the relevant criteria for classification (Swerdlow et al., 2016b).

As a result of the demonstration of the *MYD88* Leu265Pro hotspot variants, cases with a suggested diagnosis of SMZL based in histopathology assessment should be re-reviewed (table 3.8). Swerdlow *et al*. reported that a proportion of LPL can be CD5+ve, CD10+ve and /or CD23+ve (Swerdlow et al., 2016b) and, similarly, around 5% of SMZL can show CD5+ positivity (Bertoni et al., 2018). Of note, out of the five ENABLE cases with Leu265Pro *MYD88* variant for which clinical information was available, three are CD5+ve, all were CD10-negative. Despite the additional

information obtained, the differential diagnosis between LPL and SMZL remains. Further analysis to look for CNV changes associated with MZL (including del7q, trisomy 3 and 18) is required.

4.4.5 Relevance of testing: Marginal-zone lymphomas (MZLs)

Several cases for which clinical information was available suggested a possible diagnosis of SMZL and progression of cases in all CBL-MZ/non-CLL MBL studies most commonly involves splenomegaly with some cases eventually being diagnosed with SMZL (table 4.1). As was the case in our cohort, a number of these are CD5+ve SMZL (overall ~5% of SMZL are CD5+). Such cases have been shown to have a higher lymphocyte count at diagnosis and more diffuse bone marrow infiltration compared to CD5-ve SMZL, however, their outcome appears to be comparable (Baseggio et al., 2010).

Molecular testing has provided a lot of detail about the different MZL disease entities and can to a degree support the decision between the differential diagnoses as the overlap in morphology and phenotype is high (Swerdlow et al., 2016b). Figure 4.2 taken from Bertoni *et al*. clearly demonstrates the characteristics of the three MZL and the commonalities in pathway aberrations (Bertoni et al., 2018). Some of the SNVs may carry prognostic information as some studies have shown a higher risk of progression to aggressive lymphoma in SMZL with *KLF2* and *NOTCH2* variants (Jaramillo Oquendo et al., 2019). As these are rare diseases, they are underrepresented in most databases with Jaramillo Oquendo *et al.* finding that only 23% of reported SMZL variants are included in the COSMIC database for example making the analysis and interpretation of variants more challenging (Jaramillo Oquendo et al., 2019). Incorporating NGS panels like the one employed in the ENABLE study in routine diagnostic use would make sure this data is generated with the help of the clinical-scientific community it can be documented and entered into databases to further our understanding of such rarer disease entities and those not readily assigned to a given disease category.

As more studies are being published the evidence is increasing to support the suggestion that CBL-MZ should be considered as an indolent precursor to SMZL and other MZL (Merli et al., 2019).

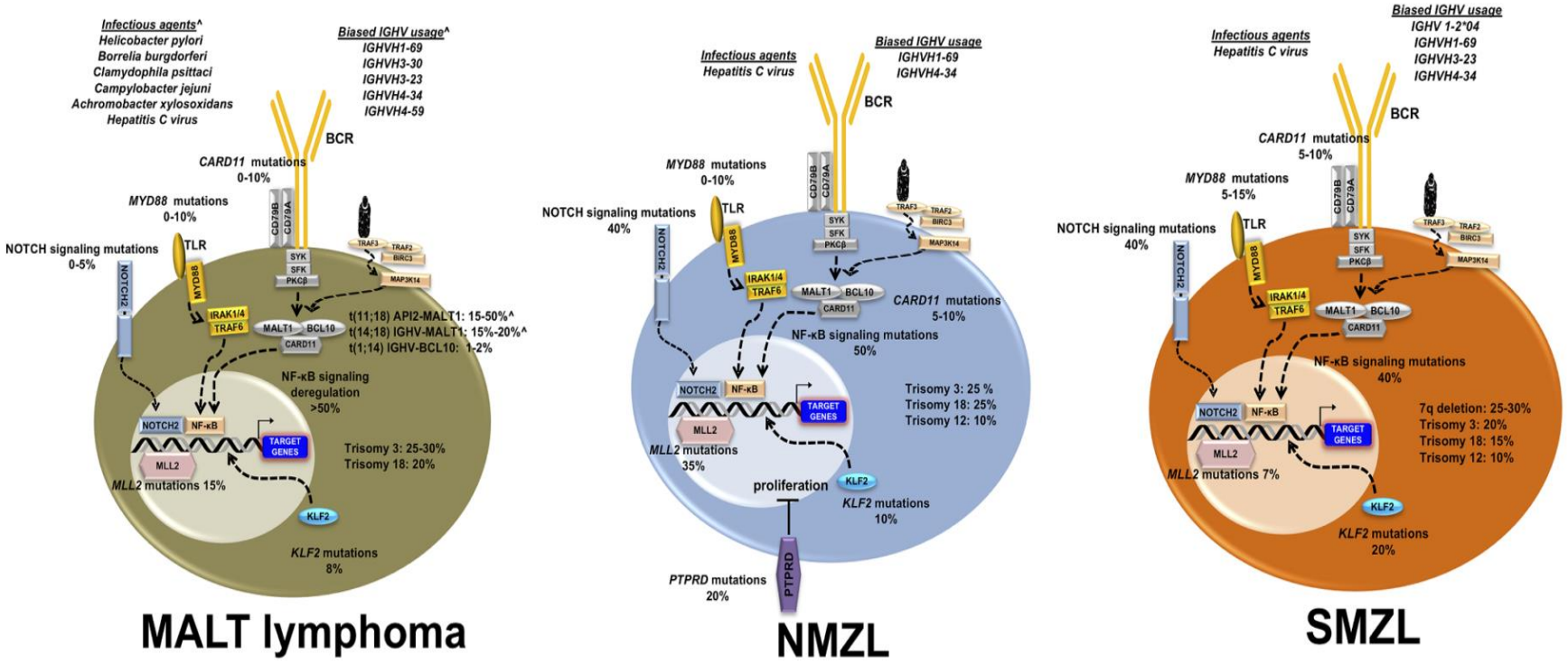


Figure 1. Summary of the main genetic and biologic features characterizing marginal zone lymphomas. ^Depending on the anatomical site. BCR, B-cell receptor; IGHV, immunoglobulin heavy variable; MALT, mucosa-associated lymphoid tissue; NF-κB, nuclear factor kappa B; NMZL, nodal marginal zone lymphoma; SMZL, splenic marginal zone lymphoma; TLR, Toll-like receptor.

Fig 4.2: **Genetic features in MZL**. Reproduced from Bertoni *et al*. ((Bertoni et al., 2018) figure 1 page 3). Note the difference in IGHV-usage between the three entities as well as the prevalence of *MYD88* and *NOTCH* variants and frequency of trisomy 12.

4.4.6 Relevance of testing: Summary

Apart from the translocations that may assign a particular WHO-defined diagnosis, none of the molecular changes on its own can classify the cases analysed as part of the ENABLE study. Further detailed analysis of the CNVs is pending and full integration with the clinical details and other laboratory and histopathological findings is paramount to try and separate these cases into diseases entities. Long-term follow-up will provide further evidence of the likely behaviour of different cases within this group of patients. The data presented so far shows clear alignment with findings published on atypical CLL/MBL and non-CLL/MBL with overlap to the MZL categories. It highlights the heterogeneity seen across patients with small B-cell lymphocytosis molecularly but also points towards the common pathways that are implicated in the development of these diseases. Table 4.2 is an attempt of an overview to highlight the findings and benefits of the ENABLE study analysis so far.

As is the case for the diagnosis and the management of haematological malignancies currently, broader testing of molecular changes in each patient will only provide real value if interpreted and assessed in conjunction with other findings. These include analysis of the cell phenotype (immunophenotyping by flow cytometry and immunohistochemistry), enumeration of cell types in a sample and review of their morphology (full blood count including differential, analysis of blood film or bone marrow aspirate smear). For tissue samples, including bone marrow trephines, review of the tissue architecture and the type, pattern and location of a B-cell infiltration can help point towards a likely diagnosis (table 1.1 and details in (Swerdlow et al., 2016a)) The technique and analysis presented here is therefore an addition to the current testing strategy but does not replace the need for the standard evaluation with different testing modalities. Further large-scale studies- including whole-genome sequencing, which is now available for patients with haematological malignancies for whom standard-of-care-testing failed to provide a diagnosis (https://www.england.nhs.uk/publication/national-genomic-test-directories/; accessed 20.11.2021)- may provide the required data to further define the disease subtypes (or their precursor scenarios) the patients studied here represent.

Table 4.2: Likely diagnosis based on the findings of the ENABLE study and the potential benefit derived from the analysis in the clinical context

Findings ENABLE study	**Likely /differential diagnosis**	**Features**	**Benefit**
IGH-BCL2	**FL**, atypical CLL		Diagnosis
IGH-BCL3	**Atypical CLL**/atypical CLL-MBL	Trisomy12, *TP53*; Disease course characterised	Diagnosis and prognosis
IGH-CCND1	**MCL**/Atypical CLL-MBL	*TP53*	Diagnosis and impact in treatment/prognosis if MCL
CDK6-TRA/D	SMZL	*TP53*, indolent course	Diagnosis and prognosis
MYD88 L265P	**LPL/WM**>> SMZL, CLL		Review required
MYD88 L265P + *CD79B*	**LPL/WM** >> MZL		Diagnosis
KLF2	SMZL/ non-CLL MBL or CBL-MZ		Diagnosis
NOTCH2	SMZL, MZL/ non-CLL MBL or CBL-MZ		Diagnosis
Trisomy 12	Atypical CLL/others	Variable	Review required
*IGHV*1-2/*IGHV*1-2*04	SMZL (V1-2*04)	KLF2, NOTCH2	Diagnosis
CCND3 + *IGHV*4-34 or *IGHV3-23,* mutated	B-SSLU/SDRPL	CD5-ve, villous lymphocytes	Diagnosis
Combination NF-κB/BCR variants	CBL-MZ	CD5-ve, CD10-ve, CD23-ver *IGHV*4-34, mutated	Likely indolent disease course

4.5 Potential relevance in future

For the cohort of cases with NF-κB- and BCR-signalling-related variants presented here, and although *KLF2, NOTCH2* and *MYD88* have a positive predictive value for SMZL and LPL for example, the diagnosis remains one based on histopathology and requires the integration of all findings and clinical symptoms. Treatment regimen are currently not standardized but are similar in that Rituximab is used either as monotherapy or in combination with Dexamethasone, Bendamustine or Lenalidomide. As these are indolent diseases a watch-and-wait approach is adopted in most patient with treatment required due to developing cytopenias in a proportion of cases. For a proportion of patients discharge to the community for follow-up by their General Practitioner (GP) could be appropriate. One crucial aspect of gathering data on these cases is to identify patterns that would predict the few cases with less indolent disease. Currently these models predominantly rely on the degree of lymphocytosis and lactase-dehydrogenase (LDH)-levels (Kalpadakis et al., 2019)

but data on non-CLL MBL and atypical CLL-MBL progression is limited (Merli et al., 2019, Kostopoulos et al., 2017).

In their prospective study of 227 cases with patients fitting into one of the three MBL classes but excluding LC-MB, Kostopoulos *et al*. found that after a median follow-up of 76 months, 78 out of the 227 cases progressed to symptomatic disease and 27 required treatment for their symptoms (Kostopoulos et al., 2017). CD5-ve MBL had the least number of cytogenetic aberrations and also had the lowest rate of disease progression. Across all three groups there was a correlation between a higher clonal B-cell count and reduced progression-free and treatment-free survival and evidence of clonal evolution was associated with poorer outcome. One drawback of the study was the lack of sequencing data meaning no evaluation of SNVs was included.

In addition to supporting a particular diagnosis, the presence or absence of variants can have an impact on response to different treatment regime. For example, in CLL and MCL, detection of missense variants in *TP53* variants makes a shift from standard-first line therapies like FCR to BTK-inhibitors mandatory (Delgado et al., 2020, Malcikova et al., 2018). Acquisition of *BIRC3* variants causes non-canonical NF-κB pathway activation and causes resistance to FCR and bypasses the block of BTK by Ibrutinib leading to treatment failure in these diseases (Diop et al., 2020). Such cases may benefit from anti-BCL2 (venetoclax)-based therapy approaches (Asslaber et al., 2019).

Co-occurring variants can also influence the response to treatment (Defrancesco et al., 2020, Guillermin et al., 2018). A recent example is the combination of *MYD88* Leu265Pro variants and *CXCR4* variants in LPL/WM: Presence of *MYD88* without a concurrent *CXCR4* variant provides the highest sensitivity to Ibrutinib compared to patients wildtype for both or with variants in both genes (Kaiser et al., 2021). Presence of *CXCR4* variants in asymptomatic patients with WM has been shown to relate to a distinct phenotype with higher bone marrow involvement and crucially a shorter time-to-treatment (Kaiser et al., 2021). One ENABLE case, not further discussed here, carried a *CXCR4* variant without *MYD88* changes (R1K-4383-0004). Similarly, in DLBCL, responses to Ibrutinib were only observed in *MYD88*-mutated cases if variants were also present in *CD79A* or *CD79B* (Guillermin et al., 2018).

Given the heterogeneity of the findings in this study (as in other studies) elucidating this level of detail will require consortium efforts and good, well annotated database that would be achievable if diagnostic testing could take a broader approach as shown in this study. Greater understanding of the pathways involved in driving neoplastic proliferation and survival will also pave the way for novel treatment avenues as well as providing a rationale for the combination of different targeted drugs to avoid escape and resistance (Chung, 2020). It may also pave the way for designing basket trials combining cases and assigning drug regimen based on molecular aberrations as was pioneered in the Cancer Research UK Matrix Trial for lung cancer (Middleton et al., 2020).

One interesting aspect is the high frequency of *TP53* aberrations given *TP53*'s crucial role in cell-cycle control and apoptosis. Long-term-follow up will help to elucidate their impact on disease behaviour and the likelihood of progression especially in those case with likely biallelic *TP53* loss due to variants and del(17p). Paradoxically, CLL-MBL with del(17p) and *TP53* variants has been shown to follow a particular indolent course; yet when these changes are present in fully-developed CLL, their presence is associated with shorter time to treatment, lack of response to FCR and poorer prognosis even with novel treatment regimens (e.g. BTK-inhibitors and BCL2-inhibitors). Co-factors modifying and contributing to this effect are currently unknown and may well lie beyond simple molecular changes for example in the epigenetic modifications of a clone or even in its microenvironment. Further research to investigate these components is already progressing.

Even in patients not needing treatment for the neoplastic disorder there is an increased risk of recurrent infections, autoimmune reactions and possibly a predisposition to other cancers: a higher risk of development of tMDS or tAML has been demonstrated in patients receiving treatment for solid tumours if they have an underlying CHIP (Coombs et al., 2017, Hammond and Loghavi, 2021).

Although the risk of overt disease appears low for most of these patients, some will show signs of disease progression and a fraction will require treatment with a small proportion transforming to overt high-risk aggressive disease like DLBCL. Thus, monitoring is needed in these patients and a better understanding of the individual components of the clone in each patient may make better risk predictions possible in future.

It may also eventually be possible to treat at an earlier stage with molecularly-targeted drugs rather than systemic therapy once symptoms have develop to reduce side effects and overall impact in an already elderly population. This concept has been proposed for MGUS and multiple myeloma recently (Maura et al., 2021).

4.6 Potential developments in technology

Within the field of molecular assay technology and in particular DNA sequencing, technologies continue to evolve at pace. Together with the reduction in cost, library preparation and sequencing technology is becoming quicker and more automated, improving its applicability into routine testing ensuring results can be delivered in a timeframe that is clinically relevant.

The next improvement for the panel presented here is the addition of targets across all chromosomes to enable 'molecular karyotyping' using shallow whole-genome sequencing in addition to the panel already established (Stewart et al., abstract 272 ASH 2021). Currently, the panel can detect a limited number of copy-number variants (CNVs) selected based on clinical utility. Review of ploidy and CNVs across all chromosomes could add further details to the cases presented here and aid their classification including prognosis. It would also reduce the need for FISH and other cytogenetic testing, reducing cost and time required for the evaluation of these cases. Similarly, developments in long-read sequencing will enable the detection of complex insertion-deletion events, inversions and large rearrangements currently not within scope of short-read sequencing panels like the one used in this study.

Ultimately, whole-genome sequencing (WGS) may remove the need for smaller panels. In October 2021, eligibility for WGS was extended in the context of haematological malignancies to those patients who could not be assigned to a WHO diagnosis based on standard of care testing (https://www.england.nhs.uk/publication/national-genomic-test-directories/; accessed 20.11.2021). From a clinical laboratory perspective, it would be ideal to re-analyse the cohort of this study to demonstrate the additional information obtained by WGS and to ascertain the added value of WGS over targeted panels. From a research perspective, WGS will provide the broadest review possible of genomic aberrations and together with expression studies via RNA-sequencing and

epigenetic testing (e.g methylation analysis) will enable characterisation of disease subtypes in more detail than ever before. However, one limitation of WGS will be the detection of subclonal variants which may be present at levels below the limit of detection of WGS, potentially reducing the predictive power and clinical utility of WGS in some clinical settings like for example longitudinal monitoring of patients on watch-and-wait.

4.7 Review of study and considerations for routine use

The ENABLE study includes a large cohort of routine patient-cases that represent the common presentations seen at a SIHMDS clinic. The detailed characterization of patients by morphology, immunophenotyping and NGS for molecular and cytogenetic variants together with the planned, long follow up will provide useful information to increase our knowledge of the different MBL types. In contrast to most other studies published this study is collecting patients' data prospectively, with long follow-up. One drawback is the recruitment of patients from different centres leading to inconsistencies of the testing performed for each patient and a central review of the morphology is therefore planned although not all patients underwent a bone marrow examination. There is currently lack of follow-up data reducing the power of the analysis. Given the high incidence of clonal variants, repeat analysis by NGS at a later time-point or at the point of progression could have been valuable but was not included in the study design. Given that a large proportion of the cohort appear to have MZL-related disease, further markers specific for this group should ideally be included in future designs of the NGS panel. This includes novel genes but also the review of the design for specific copy-number changes associated with MZL-like disease. A molecular karyotype by NGS could be the ideal addition to the testing performed so far (Mareschal et al., 2021, Liquori et al., 2017).

As already shown in the C1 Innovation Project, the NGS-capture approach is cost-efficient given the reduction in orthogonal testing (especially FISH). At present, patient expectations would need to be carefully managed given the limited information available to accurately define their specific disease subtype and likely disease course. Although the molecular alterations are common across B-NHLs and

predominantly affect BCR- and NF-κB signalling pathways, not all drugs will be applicable given the different combinations of molecular changes but also patients without defined disease may not be eligible for clinical trials. Novel or unusual findings may also be difficult to confirm and evaluate for impact.

4.8 Future work

The crucial next step for this work is the addition of the clinical data especially the long-term follow-up to further understand the behaviour of the disease in each patient. In addition to the DNA-based analysis presented here, aliquots of cells have been frozen to enable assessment of changes beyond the DNA-sequence level, including methylation arrays to assess for potential epigenetic changes.

The expansion of NGS-based testing will identify additional and sometimes novel variants that, although recurrent, are seen at such low frequencies that these were up to now not routinely tested for. Novel fusions and CNVs will provide further details on the development of the disease and the combination of variants will provide insight into the likely behaviour of the disease on a very individual level enabling personalised medicine to become a reality.

In future it may be possible to group patients based on aberrations in the same cellular pathway rather than based on phenotypic subtypes as is the main criterion currently. This could enable basket-trial designs which would allow treating patients with aberrations in the same pathways with the same therapy, regardless of the specific diagnosis.

Expanding studies such as this one to include RNA-material would allow for analysis of gene expression which would not only help to provide evidence on pathogenic mechanisms of splice variants for example but would also show the actual effect of an altered genomic profile in the neoplastic cells. Elucidating the characteristics of the B-clones from “all angles” will provide a much better picture of the drivers for disease but also help us understand the interplay of various changes better which ultimately dictate the behaviour of a clone in each individual patient.

The information about the V(D)J-rearrangements provided through this NGS-analysis could be explored further to see whether specific stereotypes associated

with a particular mutation pattern can be identified and to compare the rearrangements found to those seen in other diseases like for example autoimmune conditions. Similarities between different B-cell clones at the level of V(D)J-rearrangements would point towards a common etiology (potentially the same antigen) and would further our understanding of the development of these diseases.

For some lymphomas, like CLL for example, a genomic predisposition exists, as family members of patients with CLL have a 8.5 fold higher risk to develop the disease compared to the general population (Slager etal., 2021, Hamadou et al., 2021). Although clonal haematopiesis, MBL and MGUS are common among the general population beyond 60 years of age, higher incidences of certain malignancies are observed in families and different ethnicities and genetic predisposition alleles are just beginning to be understood. Evaluation of family members of the patients included in this study could therefore be of interest.

With the current data and the type of consent in place, germline analysis is not feasible and the depth to which the germline control samples were sequenced would not allow for a full assessment. Running matched tumour-normal samples routinely would on the one hand facilitate standard-of-care analysis but more importantly would generate the data required to screen a wider group for predisposition variants, a path we are now embarking on with the help of whole-genome sequencing in certain tumour types.

5. Conclusion

Based on the detection of the translocations alone, 11% of patients could be assigned to a known disease category and gain benefit from inclusion within associated, established treatment pathways (see https://www.royalmarsden.nhs.uk/ngs-technology for feedback from one of the patients included in this study) or access to clinical trials. This figure is likely to increase once the cohort of *MYD88*-variant- positive cases has been reviewed.

Given the atypical presentation of these cases, the less selective approach of this NGS panel is helpful to rule out or identify common lesions. In the case of BCL3, routine testing by FISH is not currently recommended given the rarity of a fusion and the associated costs for testing. The universal NGS-design employed here facilitates the cost-effective and timely characterization of patients in routine diagnostics.

The NGS-capture approach performed well with few technical problems and had a high pickup rate of clonal markers: 100% of cases for confirmation of clonality by IG rearrangements and additional clonal marker (SNV/CNV or SV) in 83% of cases (table 3.2) with further analysis for CNVs still pending.

As important as identifying patients with specific disease entities is the reassurance for patients who, based on all findings, may be predicted to have a very indolent course of disease with no or very little impact on life expectancy and performance status. Identifying such patients has the potential to reduce burden on hospital clinics by moving follow up to the community.

Studies like ENABLE are also increasing our understanding of precursor scenarios in the context of haematological malignancies and how they impact on the development of overt lymphoma, frequency of transformation and their relevance beyond a specific neoplasm in terms of co-morbidities and potential repercussions in other disease and treatment scenarios.

6. References

ABRUZZO, L. V., HERLING, C. D., CALIN, G. A., OAKES, C., BARRON, L. L., BANKS, H. E., KATJU, V., KEATING, M. J. & COOMBES, K. R. 2018. Trisomy 12 chronic lymphocytic leukemia expresses a unique set of activated and targetable pathways. *Haematologica,* 103**,** 2069-2078.

ANGELILLO, P., CAPASSO, A., GHIA, P. & SCARFÒ, L. 2018. Monoclonal B-cell lymphocytosis: Does the elderly patient need a specialistic approach? *European Journal of Internal Medicine,* 58**,** 2-6.

ANNA, A. & MONIKA, G. 2018. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *Journal of applied genetics,* 59**,** 253-268.

ARRUGA, F., VAISITTI, T. & DEAGLIO, S. 2018. The NOTCH Pathway and Its Mutations in Mature B Cell Malignancies. *Frontiers in Oncology,* 8.

ASSLABER, D., WACHT, N., LEISCH, M., QI, Y., MAEDING, N., HUFNAGL, C., JANSKO, B., ZABORSKY, N., VILLUNGER, A., HARTMANN, T. N., GREIL, R. & EGLE, A. 2019. BIRC3 Expression Predicts CLL Progression and Defines Treatment Sensitivity via Enhanced NF-κB Nuclear Translocation. *Clin Cancer Res,* 25**,** 1901-1912.

AU - AGATHANGELIDIS, A., AU - SUTTON, L. A., AU - HADZIDIMITRIOU, A., AU - TRESOLDI, C., AU - LANGERAK, A. W., AU - BELESSI, C., AU - DAVI, F., AU - ROSENQUIST, R., AU - STAMATOPOULOS, K. & AU - GHIA, P. 2018. Immunoglobulin Gene Sequence Analysis In Chronic Lymphocytic Leukemia: From Patient Material To Sequence Interpretation. *JoVE***,** e57787.

AUTORE, F., STRATI, P., LAURENTI, L. & FERRAJOLI, A. 2018. Morphological, immunophenotypic, and genetic features of chronic lymphocytic leukemia with trisomy 12: a comprehensive review. *Haematologica,* 103**,** 931-938.

BASEGGIO, L., TRAVERSE-GLEHEN, A., PETINATAUD, F., CALLET-BAUCHU, E., BERGER, F., FFRENCH, M., COURIS, C. M., THIEBLEMONT, C., MOREL, D., COIFFIER, B., SALLES, G. & FELMAN, P. 2010. CD5 expression identifies a subset of splenic marginal zone lymphomas with higher lymphocytosis: a clinico-pathological, cytogenetic and molecular study of 24 cases. *Haematologica,* 95**,** 604-12.

BEDSAUL, J. R., CARTER, N. M., DEIBEL, K. E., HUTCHERSON, S. M., JONES, T. A., WANG, Z., YANG, C., YANG, Y.-K. & POMERANTZ, J. L. 2018. Mechanisms of Regulated and Dysregulated CARD11 Signaling in Adaptive Immunity and Disease. *Frontiers in immunology,* 9**,** 2105-2105.

BERGER, A. H., BROOKS, A. N., WU, X., SHRESTHA, Y., CHOUINARD, C., PICCIONI, F., BAGUL, M., KAMBUROV, A., IMIELINSKI, M., HOGSTROM, L., ZHU, C., YANG, X., PANTEL, S., SAKAI, R., WATSON, J., KAPLAN, N., CAMPBELL, J. D., SINGH, S., ROOT, D. E., NARAYAN, R., NATOLI, T., LAHR, D. L., TIROSH, I., TAMAYO, P., GETZ, G., WONG, B., DOENCH, J., SUBRAMANIAN, A., GOLUB, T. R., MEYERSON, M. & BOEHM, J. S. 2016. High-throughput Phenotyping of Lung Cancer Somatic Mutations. *Cancer Cell,* 30**,** 214-228.

BERTONI, F., ROSSI, D. & ZUCCA, E. 2018. Recent advances in understanding the biology of marginal zone lymphoma. *F1000Res,* 7**,** 406.

BEST, O. G., GARDINER, A. C., DAVIS, Z. A., TRACY, I., IBBOTSON, R. E., MAJID, A., DYER, M. J. & OSCIER, D. G. 2009. A subset of Binet stage A CLL patients with TP53 abnormalities and mutated IGHV genes have stable disease. *Leukemia,* 23**,** 212-4.

BIKOS, V., DARZENTAS, N., HADZIDIMITRIOU, A., DAVIS, Z., HOCKLEY, S., TRAVERSE-GLEHEN, A., ALGARA, P., SANTORO, A., GONZALEZ, D., MOLLEJO, M., DAGKLIS, A., GANGEMI, F., BOSLER, D. S., BOURIKAS, G., ANAGNOSTOPOULOS, A., TSAFTARIS, A., IANNITTO, E., PONZONI, M., FELMAN, P., BERGER, F., BELESSI, C., GHIA, P., PAPADAKI, T., DOGAN, A.,

DEGANO, M., MATUTES, E., PIRIS, M. A., OSCIER, D. & STAMATOPOULOS, K. 2012. Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: ontogenetic implications. *Leukemia,* 26**,** 1638-46.

BOMMIER, C., MAUDUIT, C., FONTAINE, J., BOURBON, E., SUJOBERT, P., HUET, S., BASEGGIO, L., HAYETTE, S., LAURENT, C., BACHY, E., GHESQUIÈRES, H., THIEBLEMONT, C., SALLES, G. & TRAVERSE-GLEHEN, A. 2021. Real-life targeted next-generation sequencing for lymphoma diagnosis over 1 year from the French Lymphoma Network. *Br J Haematol*.

BROCHET, X., LEFRANC, M. P. & GIUDICELLI, V. 2008. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res,* 36**,** W503-8.

BRÜGGEMANN, M., KOTROVÁ, M., KNECHT, H., BARTRAM, J., BOUDJOGRHA, M., BYSTRY, V., FAZIO, G., FROŇKOVÁ, E., GIRAUD, M., GRIONI, A., HANCOCK, J., HERRMANN, D., JIMÉNEZ, C., KREJCI, A., MOPPETT, J., REIGL, T., SALSON, M., SCHEIJEN, B., SCHWARZ, M., SONGIA, S., SVATON, M., VAN DONGEN, J. J. M., VILLARESE, P., WAKEMAN, S., WRIGHT, G., CAZZANIGA, G., DAVI, F., GARCÍA-SANZ, R., GONZALEZ, D., GROENEN, P. J. T. A., HUMMEL, M., MACINTYRE, E. A., STAMATOPOULOS, K., POTT, C., TRKA, J., DARZENTAS, N., LANGERAK, A. W. & ON BEHALF OF THE EUROCLONALITY, N. G. S. W. G. 2019. Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. *Leukemia,* 33**,** 2241-2253.

BRUSCAGGIN, A., MONTI, S., ARCAINI, L., RAMPONI, A., RATTOTTI, S., LUCIONI, M., PAULLI, M., GAIDANO, G. & ROSSI, D. 2014. Molecular lesions of signalling pathway genes in clonal B-cell lymphocytosis with marginal zone features. *Br J Haematol,* 167**,** 718-20.

BÜHLER, M., MATUTES, E., ROZMAN, M. & CAMPO, E. 2020. Pathology of primary splenic B-cell lymphomas: a review. *Diagnostic Histopathology,* 26**,** 398-406.

BYSTRY, V., AGATHANGELIDIS, A., BIKOS, V., SUTTON, L. A., BALIAKAS, P., HADZIDIMITRIOU, A., STAMATOPOULOS, K. & DARZENTAS, N. 2015. ARResT/AssignSubsets: a novel application for robust subclassification of chronic lymphocytic leukemia based on B cell receptor IG stereotypy. *Bioinformatics,* 31**,** 3844-6.

BYSTRY, V., REIGL, T., KREJCI, A., DEMKO, M., HANAKOVA, B., GRIONI, A., KNECHT, H., SCHLITT, M., DREGER, P., SELLNER, L., HERRMANN, D., PINGEON, M., BOUDJOGHRA, M., RIJNTJES, J., POTT, C., LANGERAK, A. W., GROENEN, P. J. T. A., DAVI, F., BRÜGGEMANN, M., DARZENTAS, N. & EUROCLONALITY-NGS 2016. ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data. *Bioinformatics,* 33**,** 435-437.

CASOLA, S., PERUCHO, L., TRIPODO, C., SINDACO, P., PONZONI, M. & FACCHETTI, F. 2019. The B-cell receptor in control of tumor B-cell fitness: Biology and clinical relevance. *Immunological Reviews,* 288**,** 198-213.

CHASE, M. L. & ARMAND, P. 2018. Minimal residual disease in non-Hodgkin lymphoma – current applications and future directions. *British Journal of Haematology,* 180**,** 177-188.

CHOW, S., REECE, D., KEATS, J. J., BERGSAGEL, P. L., BAHLIS, N. J., NERI, P., PIZA RODRIGUEZ, G., SEBAG, M., WHITE, D. J., SONG, K., COMEAU, T., VENNER, C. P., CHEN, C. I., ROY, J., LOUZADA, M. L., STAKIW, J., KEW, A., KUKRETI, V., PRICA, A., TIEDEMANN, R., PANTOJA, M. A., PAUL, H. K., PUGH, T. J. & TRUDEL, S. 2017. Targeted-Capture Sequencing Approach to V(D)J Rearrangement, Immunoglobulin (Ig) Translocation and Mutation Detection in Multiple Myeloma (MM): Potential Applications in Minimal Residual Disease (MRD) Analysis and Comparison to Multiparameter Flow Cytometry (MFC) in the Myeloma Canada Research Network (MCRN) 001 Study with Intravenous Busulfan and Melphalan (BuMel) Conditioning Followed By Lenalidomide Maintenance in Newly Diagnosed Multiple Myeloma (NDMM). *Blood,* 130**,** 3057-3057.

CHUNG, C. 2020. Driving toward precision medicine for B cell lymphomas: Targeting the molecular pathogenesis at the gene level. *Journal of Oncology Pharmacy Practice,* 26**,** 943-966.

CLIPSON, A., WANG, M., DE LEVAL, L., ASHTON-KEY, M., WOTHERSPOON, A., VASSILIOU, G., BOLLI, N., GROVE, C., MOODY, S., ESCUDERO-IBARZ, L., GUNDEM, G., BRUGGER, K., XUE, X., MI, E., BENCH, A., SCOTT, M., LIU, H., FOLLOWS, G., ROBLES, E. F., MARTINEZ-CLIMENT, J. A., OSCIER, D., WATKINS, A. J. & DU, M. Q. 2015. KLF2 mutation is the most frequent somatic change in splenic marginal zone lymphoma and identifies a subset with distinct genotype. *Leukemia,* 29**,** 1177-1185.

COOMBS, C. C., ZEHIR, A., DEVLIN, S. M., KISHTAGARI, A., SYED, A., JONSSON, P., HYMAN, D. M., SOLIT, D. B., ROBSON, M. E., BASELGA, J., ARCILA, M. E., LADANYI, M., TALLMAN, M. S., LEVINE, R. L. & BERGER, M. F. 2017. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell,* 21**,** 374-382.e4.

CUENCA, I., MEDINA, A., ROSINOL, L., VAZQUEZ, I., SANCHEZ, R., GUTIERREZ, N. C., RUIZ-HEREDIA, Y., BARRIO, S., ORIOL, A., MARTIN-RAMOS, M. L., FERNÁNDEZ-GUIJARRO, M., BLANCHARD, M. J., RIOS, R., SUREDA BALARI, A., HERNÁNDEZ, M. T., DE LA RUBIA, J., ALKORTA, G., AGIRRE, X., ORDÓÑEZ, G. R., VALDÉS-MAS, R., MATEOS, M.-V., BLADE CREIXENTI, J., LAHUERTA, J. J., SAN-MIGUEL, J. F., CALASANZ, M. J., GARCIA-SANZ, R. & MARTINEZ-LOPEZ, J. 2020. Clinical Validation of a NGS Capture Panel to Identify Mutations, Copy Number Variations and Translocations in Patients with Multiple Myeloma. *Blood,* 136**,** 13-14.

CURIEL-OLMO, S., MONDÉJAR, R., ALMARAZ, C., MOLLEJO, M., CERECEDA, L., MARÈS, R., DERDAK, S., CAMPOS-MARTÍN, Y., BATLLE, A., GONZÁLEZ DE VILLAMBROSÍA, S., GUT, M., BLANC, J., TRAVERSE-GLEHEN, A., VERNEY, A., BASEGGIO, L., CAMACHO, F. I., WOTHERSPOON, A., STAMATOPOULOS, K., XOCHELLI, A., PAPADAKI, T., KANELLIS, G., PONZONI, M., GARCÍA-COSÍO, M., VAQUÉ, J. P., BELTRÁN, S., GUT, I., PIRIS, M. A. & MARTÍNEZ, N. 2017. Splenic diffuse red pulp small B-cell lymphoma displays increased expression of cyclin D3 and recurrent CCND3 mutations. *Blood,* 129**,** 1042-1045.

DAGKLIS, A., FAZI, C., SALA, C., CANTARELLI, V., SCIELZO, C., MASSACANE, R., TONIOLO, D., CALIGARIS-CAPPIO, F., STAMATOPOULOS, K. & GHIA, P. 2009. The immunoglobulin gene repertoire of low-count chronic lymphocytic leukemia (CLL)–like monoclonal B lymphocytosis is different from CLL: diagnostic implications for clinical monitoring. *Blood,* 114**,** 26-32.

DEFRANCESCO, I., ZIBELLINI, S., BOVERI, E., FRIGENI, M., FERRETTI, V. V., RIZZO, E., BONOMETTI, A., CAPUANO, F., CANDIDO, C., RATTOTTI, S., TENORE, A., PICONE, C., FLOSPERGHER, E., ZERBI, C., BERGAMINI, F., FABBRI, N., CRISTINELLI, C., VARETTONI, M., PAULLI, M. & ARCAINI, L. 2020. Targeted next-generation sequencing reveals molecular heterogeneity in non-chronic lymphocytic leukemia clonal B-cell lymphocytosis. *Hematol Oncol,* 38**,** 689-697.

DEGAUD, M., BASEGGIO, L., GRANGE, B., MANZONI, D., HUET, S., CALLET-BAUCHU, E., TRAVERSE-GLEHEN, A., DAVI, F., GHESQUIÈRES, H., SALLES, G. & SUJOBERT, P. 2019. Unclassifiable Isolated Monoclonal Lymphocytosis: Comprehensive Description of a Retrospective Cohort. *Cancers (Basel),* 11.

DELGADO, J., NADEU, F., COLOMER, D. & CAMPO, E. 2020. Chronic lymphocytic leukemia: from molecular pathogenesis to novel therapeutic strategies. *Haematologica,* 105**,** 2205-2217.

DIOP, F., MOIA, R., FAVINI, C., SPACCAROTELLA, E., DE PAOLI, L., BRUSCAGGIN, A., SPINA, V., TERZI-DI-BERGAMO, L., ARRUGA, F., TARANTELLI, C., DEAMBROGI, C., RASI, S., ADHINAVENI, R., PATRIARCA, A., FAVINI, S., SAGIRAJU, S., JABANGWE, C., KODIPAD, A. A., PERONI, D., MAURO, F. R., GIUDICE, I. D., FORCONI, F., CORTELEZZI, A., ZAJA, F., BOMBEN, R., ROSSI, F. M., VISCO, C., CHIARENZA, A., RIGOLIN, G. M., MARASCA, R., COSCIA, M.,

PERBELLINI, O., TEDESCHI, A., LAURENTI, L., MOTTA, M., DONALDSON, D., WEIR, P., MILLS, K., THORNTON, P., LAWLESS, S., BERTONI, F., POETA, G. D., CUNEO, A., FOLLENZI, A., GATTEI, V., BOLDORINI, R. L., CATHERWOOD, M., DEAGLIO, S., FOÀ, R., GAIDANO, G. & ROSSI, D. 2020. Biological and clinical implications of BIRC3 mutations in chronic lymphocytic leukemia. *Haematologica,* 105**,** 448-456.

EFREMOV, D. G., TURKALJ, S. & LAURENTI, L. 2020. Mechanisms of B Cell Receptor Activation and Responses to B Cell Receptor Inhibitors in B Cell Malignancies. *Cancers,* 12**,** 1396.

ESKELUND, C. W., DAHL, C., HANSEN, J. W., WESTMAN, M., KOLSTAD, A., PEDERSEN, L. B., MONTANO-ALMENDRAS, C. P., HUSBY, S., FREIBURGHAUS, C., EK, S., PEDERSEN, A., NIEMANN, C., RÄTY, R., BROWN, P., GEISLER, C. H., ANDERSEN, M. K., GULDBERG, P., JERKEMAN, M. & GRØNBÆK, K. 2017. TP53 mutations identify younger mantle cell lymphoma patients who do not benefit from intensive chemoimmunotherapy. *Blood,* 130**,** 1903-1910.

ESPINET, B., BELLOSILLO, B., MELERO, C., VELA, M. C., PEDRO, C., SALIDO, M., PIJUAN, L., FLORENSA, L., BESSES, C., SERRANO, S. & SOLÉ, F. 2008. FISH is better than BIOMED-2 PCR to detect IgH/BCL2 translocation in follicular lymphoma at diagnosis using paraffin-embedded tissue sections. *Leukemia Research,* 32**,** 737-742.

FANG, H., REICHARD, K. K., RABE, K. G., HANSON, C. A., CALL, T. G., DING, W., KENDERIAN, S. S., MUCHTAR, E., SCHWAGER, S. M., LEIS, J. F., CHANAN-KHAN, A. A., SLAGER, S. L., BRAGGIO, E., SMOLEY, S. A., KAY, N. E., SHANAFELT, T. D., VAN DYKE, D. L. & PARIKH, S. A. 2019. IGH translocations in chronic lymphocytic leukemia: Clinicopathologic features and clinical outcomes. *Am J Hematol,* 94**,** 338-345.

FAZI, C., SCARFÒ, L., PECCIARINI, L., COTTINI, F., DAGKLIS, A., JANUS, A., TALARICO, A., SCIELZO, C., SALA, C., TONIOLO, D., CALIGARIS-CAPPIO, F. & GHIA, P. 2011. General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. *Blood,* 118**,** 6618-25.

FLORENCE, N.-K., JEROME, L., ELISE, C., AURORE, G., SARAH, M., CAROLE, B., AGNES, D., NATHALIE, G., STÉPHANIE, S., CATHERINE, H., DOMINIQUE, P., HOSSEIN, M., JORIS, A., VIRGINIE, E., CHRYSTÈLE, B.-N., ISABELLE, L., CHRISTINE, T., LAURENCE, B., FRANCINE, M., JEAN, C., MARIE-JOELLE, M., EVELYNE, C.-B., LAUREN, V., HÉLÈNE, B., ROGER, O., JULIE, L., SYLVIE, C., HÉLÈNE, M.-B. & VÉRONIQUE, L. 2013. Chromosomal aberrations and their prognostic value in a series of 174 untreated patients with Waldenström's macroglobulinemia. *Haematologica,* 98**,** 649-654.

FORTUNO, C., LEE, K., OLIVIER, M., PESARAN, T., MAI, P. L., DE ANDRADE, K. C., ATTARDI, L. D., CROWLEY, S., EVANS, D. G., FENG, B. J., FOREMAN, A. K. M., FRONE, M. N., HUETHER, R., JAMES, P. A., MCGOLDRICK, K., MESTER, J., SEIFERT, B. A., SLAVIN, T. P., WITKOWSKI, L., ZHANG, L., PLON, S. E., SPURDLE, A. B. & SAVAGE, S. A. 2021. Specifications of the ACMG/AMP variant interpretation guidelines for germline TP53 variants. *Hum Mutat,* 42**,** 223-236.

FROYEN, G., LE MERCIER, M., LIERMAN, E., VANDEPOELE, K., NOLLET, F., BOONE, E., VAN DER MEULEN, J., JACOBS, K., LAMBIN, S., VANDER BORGHT, S., VAN VALCKENBORGH, E., ANTONIOU, A. & HÉBRANT, A. 2019. Standardization of Somatic Variant Classifications in Solid and Haematological Tumours by a Two-Level Approach of Biological and Clinical Classes: An Initiative of the Belgian ComPerMed Expert Panel. *Cancers,* 11**,** 2030.

FUSTER, C., MARTÍN-GARCIA, D., BALAGUÉ, O., NAVARRO, A., NADEU, F., COSTA, D., PRIETO, M., SALAVERRIA, I., ESPINET, B., RIVAS-DELGADO, A., TEROL, M. J., GINÉ, E., FORCADA, P., ASHTON-KEY, M., PUENTE, X. S., SWERDLOW, S. H., BEÀ, S. & CAMPO, E. 2020. Cryptic insertions of the immunoglobulin light chain

enhancer region near CCND1 in t(11;14)-negative mantle cell lymphoma. *Haematologica,* 105**,** e408-e411.

GACHARD, N., PARRENS, M., SOUBEYRAN, I., PETIT, B., MARFAK, A., RIZZO, D., DEVESA, M., DELAGE-CORRE, M., COSTE, V., LAFORÊT, M. P., DE MASCAREL, A., MERLIO, J. P., BOUABDHALLA, K., MILPIED, N., SOUBEYRAN, P., SCHMITT, A., BORDESSOULE, D., COGNÉ, M. & FEUILLARD, J. 2013. IGHV gene features and MYD88 L265P mutation separate the three marginal zone lymphoma entities and Waldenström macroglobulinemia/lymphoplasmacytic lymphomas. *Leukemia,* 27**,** 183-189.

GAILLLARD, B., CORNILLET-LEFEBVRE, P., LE, Q.-H., MALOUM, K., PANNETIER, M., LECOQ-LAFON, C., GRANGE, B., JONDREVILLE, L., MICHAUX, L., NADAL, N., ITTEL, A., LUQUET, I., STRUSKI, S., LEFEBVRE, C., GAILLARD, J.-B., LAFAGE-POCHITALOFF, M., BALDUCCI, E., PENTHER, D., BARIN, C., COLLONGE-RAME, M. A., JIMENEZ-POQUET, M., RICHEBOURG, S., LEMAIRE, P., DEFASQUE, S., RADFORD-WEISS, I., BIDET, A., SUSIN, S. A., NGUYEN-KHAC, F., CHAPIRO, E. & HÉMATOLOGIQUE, T. G. F. D. C. 2021. Clinical and biological features of B-cell neoplasms with CDK6 translocations: an association with a subgroup of splenic marginal zone lymphomas displaying frequent CD5 expression, prolymphocytic cells, and TP53 abnormalities. *British Journal of Haematology,* 193**,** 72-82.

GEMENETZI, K., AGATHANGELIDIS, A., ZARAGOZA-INFANTE, L., SOFOU, E., PAPAIOANNOU, M., CHATZIDIMITRIOU, A. & STAMATOPOULOS, K. 2020. B Cell Receptor Immunogenetics in B Cell Lymphomas: Immunoglobulin Genes as Key to Ontogeny and Clinical Decision Making. *Frontiers in Oncology,* 10.

GERTZ, M. A. 2019. Waldenström macroglobulinemia: 2019 update on diagnosis, risk stratification, and management. *Am J Hematol,* 94**,** 266-276.

GHIA, P., STAMATOPOULOS, K., BELESSI, C., MORENO, C., STILGENBAUER, S., STEVENSON, F., DAVI, F., ROSENQUIST, R. & ON BEHALF OF THE EUROPEAN RESEARCH INITIATIVE ON, C. L. L. 2007. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia,* 21**,** 1-3.

GRANAI, M., AMATO, T., DI NAPOLI, A., SANTI, R., VERGONI, F., DI STEFANO, G., MANCINI, V., KOVALCHUK, S., CENCINI, E., CARTA, A. G., AVERSA, S., ZIEPERT, M., CEVENINI, G., LAZZI, S., LEONCINI, L. & BELLAN, C. 2020. IGHV mutational status of nodal marginal zone lymphoma by NGS reveals distinct pathogenic pathways with different prognostic implications. *Virchows Archiv,* 477**,** 143-150.

GRONDONA, P., BUCHER, P., SCHULZE-OSTHOFF, K., HAILFINGER, S. & SCHMITT, A. 2018. NF-κB Activation in Lymphoid Malignancies: Genetics, Signaling, and Targeted Therapy. *Biomedicines,* 6**,** 38.

GUILLERMIN, Y., LOPEZ, J., CHABANE, K., HAYETTE, S., BARDEL, C., SALLES, G., SUJOBERT, P. & HUET, S. 2018. What Does This Mutation Mean? The Tools and Pitfalls of Variant Interpretation in Lymphoid Malignancies. *Int J Mol Sci,* 19.

GUSTINE, J. N., TSAKMAKLIS, N., DEMOS, M. G., KOFIDES, A., CHEN, J. G., LIU, X., MUNSHI, M., GUERRERA, M. L., CHAN, G. G., PATTERSON, C. J., MEID, K., DUBEAU, T., YANG, G., HUNTER, Z. R., TREON, S. P., CASTILLO, J. J. & XU, L. 2019. TP53 mutations are associated with mutated MYD88 and CXCR4, and confer an adverse outcome in Waldenström macroglobulinaemia. *British Journal of Haematology,* 184**,** 242-245.

HADZIDIMITRIOU, A., AGATHANGELIDIS, A., DARZENTAS, N., MURRAY, F., DELFAU-LARUE, M.-H., PEDERSEN, L. B., LOPEZ, A. N., DAGKLIS, A., ROMBOUT, P., BELDJORD, K., KOLSTAD, A., DREYLING, M. H., ANAGNOSTOPOULOS, A., TSAFTARIS, A., MAVRAGANI-TSIPIDOU, P., ROSENWALD, A., PONZONI, M., GROENEN, P., GHIA, P., SANDER, B., PAPADAKI, T., CAMPO, E., GEISLER, C., ROSENQUIST, R., DAVI, F., POTT, C. & STAMATOPOULOS, K. 2011. Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. *Blood,* 118**,** 3088-3095.

HAMADEH, F., MACNAMARA, S. P., AGUILERA, N. S., SWERDLOW, S. H. & COOK, J. R. 2015. MYD88 L265P mutation analysis helps define nodal lymphoplasmacytic lymphoma. *Mod Pathol,* 28**,** 564-74.

HAMMOND, D. & LOGHAVI, S. 2021. Clonal haematopoiesis of emerging significance. *Pathology,* 53**,** 300-311.

HANZIS, C., OJHA, R. P., HUNTER, Z., MANNING, R., LEWICKI, M., BRODSKY, P., IOAKIMIDIS, L., TRIPSAS, C., PATTERSON, C. J., SHEEHY, P. & TREON, S. P. 2011. Associated malignancies in patients with Waldenström's macroglobulinemia and their kin. *Clin Lymphoma Myeloma Leuk,* 11**,** 88-92.

HILL, H. A., QI, X., JAIN, P., NOMIE, K., WANG, Y., ZHOU, S. & WANG, M. L. 2020. Genetic mutations and features of mantle cell lymphoma: a systematic review and meta-analysis. *Blood Advances,* 4**,** 2927-2938.

HÖLLEIN, A., STENGEL, A., MEGGENDORFER, M., FASAN, A., KERN, W., HAFERLACH, T. & HAFERLACH, C. 2017. Deletion 7q Is Associated with KLF2 and NOTCH2 Mutations and Is Strongly Correlated with Splenic Marginal Zone Lymphoma but Also Found in Lymphoplasmacytic Lymphoma and Hairy Cell Leukemia Variant. *Blood,* 130**,** 1465-1465.

HONMA, K., TSUZUKI, S., NAKAGAWA, M., TAGAWA, H., NAKAMURA, S., MORISHIMA, Y. & SETO, M. 2009. TNFAIP3/A20 functions as a novel tumor suppressor gene in several subtypes of non-Hodgkin lymphomas. *Blood,* 114**,** 2467-75.

HUH, Y. O., ABRUZZO, L. V., RASSIDAKIS, G. Z., PARRY-JONES, N., SCHLETTE, E., BRITO-BAPABULLE, V., MATUTES, E., WOTHERSPOON, A., KEATING, M. J., MEDEIROS, L. J. & CATOVSKY, D. 2007. The t(14;19)(q32;q13)-positive small B-cell leukaemia: a clinicopathologic and cytogenetic study of seven cases. *Br J Haematol,* 136**,** 220-8.

HUNTER, Z. R., XU, L., YANG, G., ZHOU, Y., LIU, X., CAO, Y., MANNING, R. J., TRIPSAS, C., PATTERSON, C. J., SHEEHY, P. & TREON, S. P. 2014. The genomic landscape of Waldenstrom macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood,* 123**,** 1637-46.

HURWITZ, S. N., CAPONETTI, G. C., SMITH, L., QUALTIERI, J., MORRISSETTE, J. J. D., LEE, W. S., FRANK, D. M. & BAGG, A. 2021. Mutational Analysis Reinforces the Diagnosis of Nodal Marginal Zone Lymphoma With Robust PD1-positive T-Cell Hyperplasia. *The American Journal of Surgical Pathology,* 45**,** 143-145.

JAJOSKY, A. N., HAVENS, N. P., SADRI, N., ODURO, K. A., JR, MOORE, E. M., BECK, R. C. & MEYERSON, H. J. 2021. Clinical Utility of Targeted Next-Generation Sequencing in the Evaluation of Low-Grade Lymphoproliferative Disorders. *American Journal of Clinical Pathology*.

JARAMILLO OQUENDO, C., PARKER, H., OSCIER, D., ENNIS, S., GIBSON, J. & STREFFORD, J. C. 2019. Systematic Review of Somatic Mutations in Splenic Marginal Zone Lymphoma. *Scientific reports,* 9**,** 10444-10444.

KAISER, L. M., HUNTER, Z. R., TREON, S. P. & BUSKE, C. 2021. CXCR4 in Waldenström's Macroglobulinema: chances and challenges. *Leukemia,* 35**,** 333-345.

KALPADAKIS, C., PANGALIS, G. A., KONSTANTINOU, E., VASSILAKOPOULOS, T. P., TELONIS, V., SACHANAS, S., TSIRKINIDIS, P., MOSCHOGIANNIS, M., YIAKOUMIS, X., SIAKANTARIS, M., TSAFTARIDIS, P., ILIAKIS, T., BEFANI, M., KYRTSONIS, M., DIMITRAKOPOULOU, A., BOUTSIKAS, G., KORKOLOPOULOU, P., KONTOPIDOU, F., KOULIERIS, E., PONTIKOGLOU, C., XIMERI, M., PSYLAKI, M., ROUMELIOTI, M., RONTOGIANNIS, D., PAPADAKI, H., PANAGIOTIDIS, P. & ANGELOPOULOU, M. K. 2019. CLONAL B-CELL LYMPHOCYTOSIS OF MARGINAL ZONE ORIGIN (CBL-MZ): DESCRIPTION OF MAIN CLINICAL FEATURES, DISEASE EVOLUTION AND OUTCOME IN A SERIES OF 100 PATIENTS. *Hematological Oncology,* 37**,** 223-224.

KALPADAKIS, C., PANGALIS, G. A., VASSILAKOPOULOS, T. P., ROUMELIOTI, M., SACHANAS, S., KORKOLOPOULOU, P., KOULIERIS, E., MOSCHOGIANNIS, M., YIAKOUMIS, X., TSIRKINIDIS, P., PONTIKOGLOU, C., RONDOYIANNI, D.,

PAPADAKI, H. A., PANAYIOTIDIS, P. & ANGELOPOULOU, M. K. 2017. Detection of L265P MYD-88 mutation in a series of clonal B-cell lymphocytosis of marginal zone origin (CBL-MZ). *Hematol Oncol,* 35**,** 542-547.

KARNER, K., GEORGE, T. I. & PATEL, J. L. 2019. Current Aspects of Clonal Hematopoiesis: Implications for Clinical Diagnosis. *Annals of Laboratory Medicine,* 39**,** 509-514.

KATO, S., HAN, S.-Y., LIU, W., OTSUKA, K., SHIBATA, H., KANAMARU, R. & ISHIOKA, C. 2003. Understanding the function–structure and function–mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proceedings of the National Academy of Sciences,* 100**,** 8424-8429.

KING, R. L., MCPHAIL, E. D., MEYER, R. G., VASMATZIS, G., PEARCE, K., SMADBECK, J. B., KETTERLING, R. P., SMOLEY, S. A., GREIPP, P. T., HOPPMAN, N. L., PETERSON, J. F. & BAUGHN, L. B. 2019. False-negative rates for MYC fluorescence in situ hybridization probes in B-cell neoplasms. *Haematologica,* 104**,** e248-e251.

KÖHNKE, T., WITTMANN, V. K., BÜCKLEIN, V. L., LICHTENEGGER, F., PASALIC, Z., HIDDEMANN, W., SPIEKERMANN, K. & SUBKLEWE, M. 2017. Diagnosis of CLL revisited: increased specificity by a modified five-marker scoring system including CD200. *Br J Haematol,* 179**,** 480-487.

KOSTOPOULOS, I. V., PATERAKIS, G., PAVLIDIS, D., KASTRITIS, E., TERPOS, E., TSITSILONIS, O. E. & PAPADHIMITRIOU, S. I. 2017. Clonal evolution is a prognostic factor for the clinical progression of monoclonal B-cell lymphocytosis. *Blood Cancer Journal,* 7**,** e597-e597.

KÜPPERS, R. & DALLA-FAVERA, R. 2001. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene,* 20**,** 5580-5594.

LANGERAK, A. W., GROENEN, P. J. T. A., BRÜGGEMANN, M., BELDJORD, K., BELLAN, C., BONELLO, L., BOONE, E., CARTER, G. I., CATHERWOOD, M., DAVI, F., DELFAU-LARUE, M. H., DISS, T., EVANS, P. A. S., GAMEIRO, P., GARCIA SANZ, R., GONZALEZ, D., GRAND, D., HÅKANSSON, Å., HUMMEL, M., LIU, H., LOMBARDIA, L., MACINTYRE, E. A., MILNER, B. J., MONTES-MORENO, S., SCHUURING, E., SPAARGAREN, M., HODGES, E. & VAN DONGEN, J. J. M. 2012. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. *Leukemia,* 26**,** 2159-2171.

LAW, P. J. & HOULSTON, R. S. 2019. Genetic predisposition to chronic lymphocytic leukemia. *HemaSphere,* 3**,** 37-39.

LENZ, G., DAVIS, R. E., NGO, V. N., LAM, L., GEORGE, T. C., WRIGHT, G. W., DAVE, S. S., ZHAO, H., XU, W., ROSENWALD, A., OTT, G., MULLER-HERMELINK, H. K., GASCOYNE, R. D., CONNORS, J. M., RIMSZA, L. M., CAMPO, E., JAFFE, E. S., DELABIE, J., SMELAND, E. B., FISHER, R. I., CHAN, W. C. & STAUDT, L. M. 2008. Oncogenic CARD11 Mutations in Human Diffuse Large B Cell Lymphoma. *Science,* 319**,** 1676-1679.

LESLEY-ANN, S., VIKTOR, L., ANNA, E., DIEGO, C., ARON, S., EUGEN, T., KATERINA STANO, K., FERRAN, N., MARINE, A., JIKTA, M., TATJANA, P., JADE, F., ZADIE, D., DAVID, O., DAVIDE, R., PAOLO, G., JONATHAN, C. S., SARKA, P., STEPHAN, S., FREDERIC, D., ELIAS, C., KOSTAS, S., RICHARD, R. & EUROPEAN RESEARCH INITIATIVE ON, C. L. L. 2020. Comparative analysis of targeted next-generation sequencing panels for the detection of gene mutations in chronic lymphocytic leukemia: an ERIC multi-center study. *Haematologica,* 106**,** 682-691.

LI, M. M., DATTO, M., DUNCAVAGE, E. J., KULKARNI, S., LINDEMAN, N. I., ROY, S., TSIMBERIDOU, A. M., VNENCAK-JONES, C. L., WOLFF, D. J., YOUNES, A. & NIKIFOROVA, M. N. 2017. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *The Journal of molecular diagnostics : JMD,* 19**,** 4-23.

LIM, M. S., BAILEY, N. G., KING, R. L. & PIRIS, M. 2019. Molecular Genetics in the Diagnosis and Biology of Lymphoid Neoplasms: 2017 Society for

Hematopathology/European Association for Haematopathology Workshop Report. *American Journal of Clinical Pathology,* 152**,** 277-301.

LIQUORI, A., SUCH, E., PALOMO, L., LESENDE, I., NEEF, A., IBAÑEZ, M., AVETISYAN, G., COMPANY, D., MOREAU, S., PEDROLA, L., SELLES, J., SAUS, A., AYALA, J., ACHA, P., SANJUAN, A., BOLUDA, M., DE MATTEO, B., GONZALEZ, E., SANZ, M. A., SOLE, F., SANZ, G. & CERVERA, J. 2017. A Novel Single Next-Generation Sequencing (NGS) Approach for the Molecular Karyotyping and Genotyping of Patients with Myelodysplastic Syndromes. *Blood,* 130**,** 1700-1700.

LUCA, A. & DAVIDE, R. 2012. Nuclear factor-κB dysregulation in splenic marginal zone lymphoma: new therapeutic opportunities. *Haematologica,* 97**,** 638-640.

M, D. E. B., TOUS, C., GUÉGANIC, N., MJ, L. E. B., BASINKO, A., MOREL, F. & DOUET-GUILBERT, N. 2016. Immunoglobulin gene translocations in chronic lymphocytic leukemia: A report of 35 patients and review of the literature. *Mol Clin Oncol,* 4**,** 682-694.

MAGNOLI, F., TIBILETTI, M. G. & UCCELLA, S. 2019. Unraveling Tumor Heterogeneity in an Apparently Monolithic Disease: BCL2 and Other Players in the Genetic Landscape of Nodal Follicular Lymphoma. *Frontiers in Medicine,* 6.

MAITRE, E. & TROUSSARD, X. 2019. Monoclonal B-cell lymphocytosis. *Best Pract Res Clin Haematol,* 32**,** 229-238.

MALCIKOVA, J., TAUSCH, E., ROSSI, D., SUTTON, L. A., SOUSSI, T., ZENZ, T., KATER, A. P., NIEMANN, C. U., GONZALEZ, D., DAVI, F., GONZALEZ DIAZ, M., MORENO, C., GAIDANO, G., STAMATOPOULOS, K., ROSENQUIST, R., STILGENBAUER, S., GHIA, P. & POSPISILOVA, S. 2018. ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia-update on methodological approaches and results interpretation. *Leukemia,* 32**,** 1070-1080.

MARESCHAL, S., PALAU, A., LINDBERG, J., RUMINY, P., NILSSON, C., BENGTZÉN, S., ENGVALL, M., ERIKSSON, A., NEDDERMEYER, A., MARCHAND, V., JANSSON, M., BJÖRKLUND, M., JARDIN, F., RANTALAINEN, M., LENNARTSSON, A., CAVELIER, L., GRÖNBERG, H. & LEHMANN, S. 2021. Challenging conventional karyotyping by next-generation karyotyping in 281 intensively treated patients with AML. *Blood Advances,* 5**,** 1003-1016.

MARTÍN-GARCIA, D., NAVARRO, A., VALDÉS-MAS, R., CLOT, G., GUTIÉRREZ-ABRIL, J., PRIETO, M., RIBERA-CORTADA, I., WORONIECKA, R., RYMKIEWICZ, G., BENS, S., DE LEVAL, L., ROSENWALD, A., FERRY, J. A., HSI, E. D., FU, K., DELABIE, J., WEISENBURGER, D., DE JONG, D., CLIMENT, F., O'CONNOR, S. J., SWERDLOW, S. H., TORRENTS, D., BELTRAN, S., ESPINET, B., GONZÁLEZ-FARRÉ, B., VELOZA, L., COSTA, D., MATUTES, E., SIEBERT, R., OTT, G., QUINTANILLA-MARTINEZ, L., JAFFE, E. S., LÓPEZ-OTÍN, C., SALAVERRIA, I., PUENTE, X. S., CAMPO, E. & BEÀ, S. 2019. CCND2 and CCND3 hijack immunoglobulin light-chain enhancers in cyclin D1(-) mantle cell lymphoma. *Blood,* 133**,** 940-951.

MARTÍN-SUBERO, J. I., IBBOTSON, R., KLAPPER, W., MICHAUX, L., CALLET-BAUCHU, E., BERGER, F., CALASANZ, M. J., DE WOLF-PEETERS, C., DYER, M. J., FELMAN, P., GARDINER, A., GASCOYNE, R. D., GESK, S., HARDER, L., HORSMAN, D. E., KNEBA, M., KÜPPERS, R., MAJID, A., PARRY-JONES, N., RITGEN, M., SALIDO, M., SOLÉ, F., THIEL, G., WACKER, H. H., OSCIER, D., WLODARSKA, I. & SIEBERT, R. 2007. A comprehensive genetic and histopathologic analysis identifies two subgroups of B-cell malignancies carrying a t(14;19)(q32;q13) or variant BCL3-translocation. *Leukemia,* 21**,** 1532-1544.

MARTINEZ-LOPEZ, A., CURIEL-OLMO, S., MOLLEJO, M., CERECEDA, L., MARTINEZ, N., MONTES-MORENO, S., ALMARAZ, C., REVERT, J. B. & PIRIS, M. A. 2015. MYD88 (L265P) somatic mutation in marginal zone B-cell lymphoma. *Am J Surg Pathol,* 39**,** 644-51.

MAURA, F., LANDGREN, O. & MORGAN, G. J. 2021. Designing Evolutionary-based Interception Strategies to Block the Transition from Precursor Phases to Multiple Myeloma. *Clinical Cancer Research,* 27**,** 15-23.

MCCONNELL, L., HOUGHTON, O., STEWART, P., GAZDOVA, J., SRIVASTAVA, S., KIM, C., CATHERWOOD, M., STROBL, A., FLANAGAN, A. M., ONISCU, A., KROEZE, L. I., GROENEN, P., TANIERE, P., SALTO-TELLEZ, M. & GONZALEZ, D. 2020. A novel next generation sequencing approach to improve sarcoma diagnosis. *Modern Pathology,* 33**,** 1350-1359.

MERLI, M., BIANCHI, B., BERTÙ, L., FERRARIO, A., MORA, B., UCCELLA, S., FURLAN, D., TIBILETTI, M. G., CATTANEO, S., SESSA, F. & PASSAMONTI, F. 2019. Clonal B-Cell Lymphocytosis with Mariginal-Zone Features: Comparison with Overt Splenic Marginal-Zone Lymphomas in 77 Patients from a Monocentric Series. *Blood,* 134**,** 4017-4017.

MIAO, Y., LIN, P., WANG, W., JEFFREY MEDEIROS, L. & LU, X. 2016. CCND1-IGH Fusion-Amplification and MYC Copy Number Gain in a Case of Pleomorphic Variant Mantle Cell Lymphoma. *Am J Clin Pathol,* 146**,** 747-752.

MIDDLETON, G., FLETCHER, P., POPAT, S., SAVAGE, J., SUMMERS, Y., GREYSTOKE, A., GILLIGAN, D., CAVE, J., O'ROURKE, N., BREWSTER, A., TOY, E., SPICER, J., JAIN, P., DANGOOR, A., MACKEAN, M., FORSTER, M., FARLEY, A., WHERTON, D., MEHMI, M., SHARPE, R., MILLS, T. C., CERONE, M. A., YAP, T. A., WATKINS, T. B. K., LIM, E., SWANTON, C. & BILLINGHAM, L. 2020. The National Lung Matrix Trial of personalized therapy in lung cancer. *Nature,* 583**,** 807-812.

MOODY, S., ESCUDERO-IBARZ, L., WANG, M., CLIPSON, A., OCHOA RUIZ, E., DUNN-WALTERS, D., XUE, X., ZENG, N., ROBSON, A., CHUANG, S.-S., COGLIATTI, S., LIU, H., GOODLAD, J., ASHTON-KEY, M., RADERER, M., BI, Y. & DU, M.-Q. 2017. Significant association between TNFAIP3 inactivation and biased immunoglobulin heavy chain variable region 4-34 usage in mucosa-associated lymphoid tissue lymphoma. *The Journal of Pathology,* 243**,** 3-8.

NAGEL, D., VINCENDEAU, M., EITELHUBER, A. C. & KRAPPMANN, D. 2014. Mechanisms and consequences of constitutive NF-κB activation in B-cell lymphoid malignancies. *Oncogene,* 33**,** 5655-5665.

NAYLOR, T. L., TANG, H., RATSCH, B. A., ENNS, A., LOO, A., CHEN, L., LENZ, P., WATERS, N. J., SCHULER, W., DÖRKEN, B., YAO, Y. M., WARMUTH, M., LENZ, G. & STEGMEIER, F. 2011. Protein kinase C inhibitor sotrastaurin selectively inhibits the growth of CD79 mutant diffuse large B-cell lymphomas. *Cancer Res,* 71**,** 2643-53.

NEBENFUEHR, S., KOLLMANN, K. & SEXL, V. 2020. The role of CDK6 in cancer. *International Journal of Cancer,* 147**,** 2988-2995.

NGO, V. N., YOUNG, R. M., SCHMITZ, R., JHAVAR, S., XIAO, W., LIM, K.-H., KOHLHAMMER, H., XU, W., YANG, Y., ZHAO, H., SHAFFER, A. L., ROMESSER, P., WRIGHT, G., POWELL, J., ROSENWALD, A., MULLER-HERMELINK, H. K., OTT, G., GASCOYNE, R. D., CONNORS, J. M., RIMSZA, L. M., CAMPO, E., JAFFE, E. S., DELABIE, J., SMELAND, E. B., FISHER, R. I., BRAZIEL, R. M., TUBBS, R. R., COOK, J. R., WEISENBURGER, D. D., CHAN, W. C. & STAUDT, L. M. 2011. Oncogenically active MYD88 mutations in human lymphoma. *Nature,* 470**,** 115-119.

ONAINDIA, A., MEDEIROS, L. J. & PATEL, K. P. 2017. Clinical utility of recently identified diagnostic, prognostic, and predictive molecular biomarkers in mature B-cell neoplasms. *Mod Pathol,* 30**,** 1338-1366.

ONDREJKA, S. L., LIN, J. J., WARDEN, D. W., DURKIN, L., COOK, J. R. & HSI, E. D. 2013. MYD88 L265P somatic mutation: its usefulness in the differential diagnosis of bone marrow involvement by B-cell lymphoproliferative disorders. *Am J Clin Pathol,* 140**,** 387-94.

PAE, J., ERSCHING, J., CASTRO, T. B. R., SCHIPS, M., MESIN, L., ALLON, S. J., ORDOVAS-MONTANES, J., MLYNARCZYK, C., MELNICK, A., EFEYAN, A., SHALEK, A. K., MEYER-HERMANN, M. & VICTORA, G. D. 2021. Cyclin D3 drives inertial cell cycling in dark zone germinal center B cells. *J Exp Med,* 218.

PARKER, H., MCIVER-BROWN, N. R., DAVIS, Z. A., PARRY, M., ROSE-ZERILLI, M. J. J., XOCHELLI, A., GIBSON, J., WALEWSKA, R., STREFFORD, J. C. & OSCIER,

D. G. 2018. CBL-MZ is not a single biological entity: evidence from genomic analysis and prolonged clinical follow-up. *Blood advances,* 2**,** 1116-1119.

PETERSON, J. F., BAUGHN, L. B., KETTERLING, R. P., PITEL, B. A., SMOLEY, S. A., VASMATZIS, G., SMADBECK, J. B., GREIPP, P. T., MANGAONKAR, A. A., THOMPSON, C. A., PARIKH, S. A., CHEN, D. & VISWANATHA, D. S. 2019. Characterization of a cryptic IGH/CCND1 rearrangement in a case of mantle cell lymphoma with negative CCND1 FISH studies. *Blood Adv,* 3**,** 1298-1302.

PETITJEAN, A., HAINAUT, P. & OLIVIER, M. 2007. The IARC TP53 mutation database: a resource for studying the significance of TP53 mutations in human cancers. *Iatreia,* 20**,** s41-s42.

PETRIKKOS, L., KYRTSONIS, M. C., ROUMELIOTI, M., GEORGIOU, G., EFTHYMIOU, A., TZENOU, T. & PANAYIOTIDIS, P. 2014. Clonotypic analysis of immunoglobulin heavy chain sequences in patients with Waldenström's macroglobulinemia: correlation with MYD88 L265P somatic mutation status, clinical features, and outcome. *Biomed Res Int,* 2014**,** 809103.

PILLONEL, V., JUSKEVICIUS, D., NG, C. K. Y., BODMER, A., ZETTL, A., JUCKER, D., DIRNHOFER, S. & TZANKOV, A. 2018. High-throughput sequencing of nodal marginal zone lymphomas identifies recurrent BRAF mutations. *Leukemia,* 32**,** 2412-2426.

PIVA, R., DEAGLIO, S., FAMÀ, R., BUONINCONTRI, R., SCARFÒ, I., BRUSCAGGIN, A., MEREU, E., SERRA, S., SPINA, V., BRUSA, D., GARAFFO, G., MONTI, S., DAL BO, M., MARASCA, R., ARCAINI, L., NERI, A., GATTEI, V., PAULLI, M., TIACCI, E., BERTONI, F., PILERI, S. A., FOÀ, R., INGHIRAMI, G., GAIDANO, G. & ROSSI, D. 2015. The Krüppel-like factor 2 transcription factor gene is recurrently mutated in splenic marginal zone lymphoma. *Leukemia,* 29**,** 503-7.

PULSONI, A., DELLA STARZA, I., CAPPELLI, L. V., TOSTI, M. E., ANNECHINI, G., CAVALLI, M., DE NOVI, L. A., D'ELIA, G. M., GRAPULIN, L., GUARINI, A., DEL GIUDICE, I. & FOÀ, R. 2020. Minimal residual disease monitoring in early stage follicular lymphoma can predict prognosis and drive treatment with rituximab after radiotherapy. *British Journal of Haematology,* 188**,** 249-258.

RINALDI, A., MIAN, M., CHIGRINOVA, E., ARCAINI, L., BHAGAT, G., NOVAK, U., RANCOITA, P. M., DE CAMPOS, C. P., FORCONI, F., GASCOYNE, R. D., FACCHETTI, F., PONZONI, M., GOVI, S., FERRERI, A. J., MOLLEJO, M., PIRIS, M. A., BALDINI, L., SOULIER, J., THIEBLEMONT, C., CANZONIERI, V., GATTEI, V., MARASCA, R., FRANCESCHETTI, S., GAIDANO, G., TUCCI, A., UCCELLA, S., TIBILETTI, M. G., DIRNHOFER, S., TRIPODO, C., DOGLIONI, C., DALLA FAVERA, R., CAVALLI, F., ZUCCA, E., KWEE, I. & BERTONI, F. 2011. Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact on the clinical outcome. *Blood,* 117**,** 1595-604.

ROBINSON, J. T., THORVALDSDÓTTIR, H., WENGER, A. M., ZEHIR, A. & MESIROV, J. P. 2017. Variant Review with the Integrative Genomics Viewer. *Cancer Research,* 77**,** e31-e34.

RODRIGUES, J. M., HASSAN, M., FREIBURGHAUS, C., ESKELUND, C. W., GEISLER, C., RÄTY, R., KOLSTAD, A., SUNDSTRÖM, C., GLIMELIUS, I., GRØNBÆK, K., KWIECINSKA, A., PORWIT, A., JERKEMAN, M. & EK, S. 2020. p53 is associated with high-risk and pinpoints TP53 missense mutations in mantle cell lymphoma. *British Journal of Haematology,* 191**,** 796-805.

ROLLETT, R. A., WILKINSON, E. J., GONZALEZ, D., FENTON, J. A. L., SHORT, M. A., EVANS, P. A. S., RAWSTRON, A. C. & OWEN, R. G. 2006. Immunoglobulin Heavy Chain Sequence Analysis in Waldenström's Macroglobulinemia and Immunoglobulin M Monoclonal Gammopathy of Undetermined Significance. *Clinical Lymphoma and Myeloma,* 7**,** 70-72.

ROSENQUIST, R., GHIA, P., HADZIDIMITRIOU, A., SUTTON, L. A., AGATHANGELIDIS, A., BALIAKAS, P., DARZENTAS, N., GIUDICELLI, V., LEFRANC, M. P., LANGERAK, A. W., BELESSI, C., DAVI, F. & STAMATOPOULOS, K. 2017. Immunoglobulin gene sequence analysis in chronic lymphocytic leukemia: updated ERIC recommendations. *Leukemia,* 31**,** 1477-1481.

ROSSI, D. 2014. Role of MYD88 in lymphoplasmacytic lymphoma diagnosis and pathogenesis. *Hematology Am Soc Hematol Educ Program,* 2014**,** 113-8.
ROSSI, D. & GAIDANO, G. 2010. Biological and clinical significance of stereotyped B-cell receptors in chronic lymphocytic leukemia. *Haematologica,* 95**,** 1992-1995.
ROSSI, D., TRIFONOV, V., FANGAZIO, M., BRUSCAGGIN, A., RASI, S., SPINA, V., MONTI, S., VAISITTI, T., ARRUGA, F., FAMÀ, R., CIARDULLO, C., GRECO, M., CRESTA, S., PIRANDA, D., HOLMES, A., FABBRI, G., MESSINA, M., RINALDI, A., WANG, J., AGOSTINELLI, C., PICCALUGA, P. P., LUCIONI, M., TABBÒ, F., SERRA, R., FRANCESCHETTI, S., DEAMBROGI, C., DANIELE, G., GATTEI, V., MARASCA, R., FACCHETTI, F., ARCAINI, L., INGHIRAMI, G., BERTONI, F., PILERI, S. A., DEAGLIO, S., FOÀ, R., DALLA-FAVERA, R., PASQUALUCCI, L., RABADAN, R. & GAIDANO, G. 2012. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *The Journal of experimental medicine,* 209**,** 1537-1551.
ROUÉ, G. & SOLA, B. 2020. Management of Drug Resistance in Mantle Cell Lymphoma. *Cancers,* 12**,** 1565.
RUBEN, A. L. D. G., ANNE, M. R. S., MARIE JOSÉ, K., STEVEN, T. P. & JOOST, S. P. V. 2019. MYD88 in the driver's seat of B-cell lymphomagenesis: from molecular mechanisms to clinical implications. *Haematologica,* 104**,** 2337-2348.
SAKHDARI, A., OK, C. Y., PATEL, K. P., KANAGAL-SHAMANNA, R., YIN, C. C., ZUO, Z., HU, S., ROUTBORT, M. J., LUTHRA, R., MEDEIROS, L. J., KHOURY, J. D. & LOGHAVI, S. 2019. TP53 mutations are common in mantle cell lymphoma, including the indolent leukemic non-nodal variant. *Annals of Diagnostic Pathology,* 41**,** 38-42.
SCHMITZ, R., CERIBELLI, M., PITTALUGA, S., WRIGHT, G. & STAUDT, L. M. 2014. Oncogenic Mechanisms in Burkitt Lymphoma. *Cold Spring Harbor Perspectives in Medicine,* 4.
SHANMUGAM, V., CRAIG, J. W., HILTON, L. K., NGUYEN, M. H., RUSHTON, C. K., FAHIMDANESH, K., LOVITCH, S., FERLAND, B., SCOTT, D. W. & ASTER, J. C. 2021. Notch activation is pervasive in SMZL and uncommon in DLBCL: implications for Notch signaling in B-cell tumors. *Blood Advances,* 5**,** 71-83.
SHUAI, W., LIN, P., STRATI, P., PATEL, K. P., ROUTBORT, M. J., HU, S., WEI, P., KHOURY, J. D., YOU, M. J., LOGHAVI, S., TANG, Z., FANG, H., THAKRAL, B., MEDEIROS, L. J. & WANG, W. 2020. Clinicopathological characterization of chronic lymphocytic leukemia with MYD88 mutations: L265P and non-L265P mutations are associated with different features. *Blood Cancer Journal,* 10**,** 86.
SILKENSTEDT, E., LINTON, K. & DREYLING, M. 2021. Mantle cell lymphoma – advances in molecular biology, prognostication and treatment approaches. *British Journal of Haematology,* n/a.
SINGH, M., JACKSON, K. J. L., WANG, J. J., SCHOFIELD, P., FIELD, M. A., KOPPSTEIN, D., PETERS, T. J., BURNETT, D. L., RIZZETTO, S., NEVOLTRIS, D., MASLE-FARQUHAR, E., FAULKS, M. L., RUSSELL, A., GOKAL, D., HANIOKA, A., HORIKAWA, K., COLELLA, A. D., CHATAWAY, T. K., BLACKBURN, J., MERCER, T. R., LANGLEY, D. B., GOODALL, D. M., JEFFERIS, R., GANGADHARAN KOMALA, M., KELLEHER, A. D., SUAN, D., RISCHMUELLER, M., CHRIST, D., BRINK, R., LUCIANI, F., GORDON, T. P., GOODNOW, C. C. & REED, J. H. 2020. Lymphoma Driver Mutations in the Pathogenic Evolution of an Iconic Human Autoantibody. *Cell,* 180**,** 878-894.e19.
SORRENTINO, C., CUNEO, A. & ROTI, G. 2019. Therapeutic Targeting of Notch Signaling Pathway in Hematological Malignancies. *Mediterranean journal of hematology and infectious diseases,* 11**,** e2019037-e2019037.
SPINA, V., MENSAH, A. A. & ARRIBAS, A. J. 2021. Biology of splenic and nodal marginal zone lymphomas. *Annals of Lymphoma,* 5.
STAMATOPOULOS, K., BELESSI, C., PAPADAKI, T., KALAGIAKOU, E., STAVROYIANNI, N., DOUKA, V., AFENDAKI, S., SALOUM, R., PARASI, A., ANAGNOSTOU, D., LAOUTARIS, N., FASSAS, A. & ANAGNOSTOPOULOS, A. 2004. Immunoglobulin heavy- and light-chain repertoire in splenic marginal zone lymphoma. *Molecular medicine (Cambridge, Mass.),* 10**,** 89-95.

STEWART, J. P., GAZDOVA, J., DARZENTAS, N., WREN, D., PROSZEK, P., FAZIO, G., SONGIA, S., ALCOCEBA, M., SARASQUETE, M. E., VILLARESE, P., VAN DER KLIFT, M. Y., HEEZEN, K. C., MCCAFFERTY, N., PAL, K., CATHERWOOD, M., KIM, C. S., SRIVASTAVA, S., KROEZE, L. I., HODGES, E., STAMATOPOULOS, K., KLAPPER, W., GENUARDI, E., FERRERO, S., VAN DEN BRAND, M., CAZZANIGA, G., DAVI, F., SUTTON, L.-A., GARCIA-SANZ, R., GROENEN, P. J. T. A., MACINTYRE, E. A., BRÜGGEMANN, M., POTT, C., LANGERAK, A. W., GONZALEZ, D. & GROUP, O. B. O. T. E.-N. W. 2021. Validation of the EuroClonality-NGS DNA capture panel as an integrated genomic tool for lymphoproliferative disorders. *Blood Advances,* 5**,** 3188-3198.

STEWART, P., GAZDOVA, J., DARZENTAS, N., WREN, D., PROSZEK, P., FAZIO, G., SONGIA, S., ALCOCEBA, M., SARASQUETE, M. E., VILLARESE, P., VAN DER KLIFT, M. Y., HEEZEN, K., MCCAFFERTY, N., PAL, K., CATHERWOOD, M., KIM, C. S., SRIVASTAVA, S., HODGES, E., STAMATOPOULOS, K., KLAPPER, W., FERRERO, S., VAN DEN BRAND, M., CAZZANIGA, G., DAVI, F., GARCIA-SANZ, R., GROENEN, P., MACINTYRE, E., BRÜGGEMANN, M., POTT, C., GENUARDI, E., LANGERAK, A. W. & GONZALEZ, D. 2019. Euroclonality-NGS DNA Capture Panel for Integrated Analysis of IG/TR Rearrangements, Translocations, Copy Number and Sequence Variation in Lymphoproliferative Disorders. *Blood,* 134**,** 888-888.

STILGENBAUER, S., SCHNAITER, A., PASCHKA, P., ZENZ, T., ROSSI, M., DÖHNER, K., BÜHLER, A., BÖTTCHER, S., RITGEN, M., KNEBA, M., WINKLER, D., TAUSCH, E., HOTH, P., EDELMANN, J., MERTENS, D., BULLINGER, L., BERGMANN, M., KLESS, S., MACK, S., JÄGER, U., PATTEN, N., WU, L., WENGER, M. K., FINGERLE-ROWSON, G., LICHTER, P., CAZZOLA, M., WENDTNER, C. M., FINK, A. M., FISCHER, K., BUSCH, R., HALLEK, M. & DÖHNER, H. 2014. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood,* 123**,** 3247-54.

SUJOBERT, P., LE BRIS, Y., DE LEVAL, L., GROS, A., MERLIO, J. P., PASTORET, C., HUET, S., SARKOZY, C., DAVI, F., CALLANAN, M., THIEBLEMONT, C., SIBON, D., ASNAFI, V., PREUDHOMME, C., GAULARD, P., JARDIN, F., SALLES, G. & MACINTYRE, E. 2018. The Need for a Consensus Next-generation Sequencing Panel for Mature Lymphoid Malignancies. *HemaSphere,* 3**,** e169-e169.

SUTTON, L.-A., KOSTARELI, E., HADZIDIMITRIOU, A., DARZENTAS, N., TSAFTARIS, A., ANAGNOSTOPOULOS, A., ROSENQUIST, R. & STAMATOPOULOS, K. 2009. Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. *Blood,* 114**,** 4460-4468.

SUTTON, L.-A., YOUNG, E., BALIAKAS, P., HADZIDIMITRIOU, A., MOYSIADIS, T., PLEVOVA, K., ROSSI, D., KMINKOVA, J., STALIKA, E., PEDERSEN, L. B., MALCIKOVA, J., AGATHANGELIDIS, A., DAVIS, Z., MANSOURI, L., SCARFÒ, L., BOUDJOGHRA, M., NAVARRO, A., MUGGEN, A. F., YAN, X.-J., NGUYEN-KHAC, F., LARRAYOZ, M., PANAGIOTIDIS, P., CHIORAZZI, N., NIEMANN, C. U., BELESSI, C., CAMPO, E., STREFFORD, J. C., LANGERAK, A. W., OSCIER, D., GAIDANO, G., POSPISILOVA, S., DAVI, F., GHIA, P., STAMATOPOULOS, K., ROSENQUIST, R. & ERIC, T. E. R. I. O. C. L. L. 2016. Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors. *Haematologica,* 101**,** 959-967.

SWERDLOW, S. H., CAMPO, E., PILERI, S. A., HARRIS, N. L., STEIN, H., SIEBERT, R., ADVANI, R., GHIELMINI, M., SALLES, G. A., ZELENETZ, A. D. & JAFFE, E. S. 2016a. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood,* 127**,** 2375-90.

SWERDLOW, S. H., KUZU, I., DOGAN, A., DIRNHOFER, S., CHAN, J. K. C., SANDER, B., OTT, G., XERRI, L., QUINTANILLA-MARTINEZ, L. & CAMPO, E. 2016b. The many faces of small B cell lymphomas with plasmacytic differentiation and the contribution of MYD88 testing. *Virchows Archiv,* 468**,** 259-275.

SZANKASI, P., BOLIA, A., LIEW, M., SCHUMACHER, J. A., GEE, E. P. S., MATYNIA, A. P., LI, K. D., PATEL, J. L., XU, X., SALAMA, M. E. & KELLEY, T. W. 2019. Comprehensive detection of chromosomal translocations in lymphoproliferative disorders by massively parallel sequencing. *Journal of Hematopathology,* 12**,** 121-133.

TAKEUCHI, T., YAMAGUCHI, M., KOBAYASHI, K., MIYAZAKI, K., TAWARA, I., IMAI, H., ONO, R., NOSAKA, T., TANAKA, K. & KATAYAMA, N. 2017. MYD88, CD79B, and CARD11 gene mutations in CD5-positive diffuse large B-cell lymphoma. *Cancer,* 123**,** 1166-1173.

TAUSCH, E. & STILGENBAUER, S. 2020. BIRC3 mutations in chronic lymphocytic leukemia - uncommon and unfavorable. *Haematologica,* 105**,** 255-256.

TEN HACKEN, E., GOUNARI, M., GHIA, P. & BURGER, J. A. 2019. The importance of B cell receptor isotypes and stereotypes in chronic lymphocytic leukemia. *Leukemia,* 33**,** 287-298.

THIEBLEMONT, C. 2017. Improved biological insight and influence on management in indolent lymphoma. Talk 3: update on nodal and splenic marginal zone lymphoma. *Hematology Am Soc Hematol Educ Program,* 2017**,** 371-378.

THIEBLEMONT, C., BERTONI, F., COPIE-BERGMAN, C., FERRERI, A. J. & PONZONI, M. 2014. Chronic inflammation and extra-nodal marginal-zone lymphomas of MALT-type. *Semin Cancer Biol,* 24**,** 33-42.

VAN DONGEN, J. J. M., LANGERAK, A. W., BRÜGGEMANN, M., EVANS, P. A. S., HUMMEL, M., LAVENDER, F. L., DELABESSE, E., DAVI, F., SCHUURING, E., GARCÍA-SANZ, R., VAN KRIEKEN, J. H. J. M., DROESE, J., GONZÁLEZ, D., BASTARD, C., WHITE, H. E., SPAARGAREN, M., GONZÁLEZ, M., PARREIRA, A., SMITH, J. L., MORGAN, G. J., KNEBA, M. & MACINTYRE, E. A. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia,* 17**,** 2257-2317.

VENDRAMINI, E., BOMBEN, R., POZZO, F., BENEDETTI, D., BITTOLO, T., ROSSI, F. M., DAL BO, M., RABE, K. G., POZZATO, G., ZAJA, F., CHIARENZA, A., DI RAIMONDO, F., BRAGGIO, E., PARIKH, S. A., KAY, N. E., SHANAFELT, T. D., DEL POETA, G., GATTEI, V. & ZUCCHETTO, A. 2019. KRAS, NRAS, and BRAF mutations are highly enriched in trisomy 12 chronic lymphocytic leukemia and are associated with shorter treatment-free survival. *Leukemia,* 33**,** 2111-2115.

VISCO, C., TANASI, I., QUAGLIA, F. M., FERRARINI, I., FRAENZA, C. & KRAMPERA, M. 2020. Oncogenic Mutations of MYD88 and CD79B in Diffuse Large B-Cell Lymphoma and Implications for Clinical Practice. *Cancers (Basel),* 12.

WALKER, B. A., WARDELL, C. P., JOHNSON, D. C., KAISER, M. F., BEGUM, D. B., DAHIR, N. B., ROSS, F. M., DAVIES, F. E., GONZALEZ, D. & MORGAN, G. J. 2013. Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. *Blood,* 121**,** 3413-3419.

WANG, J. Q., JEELALL, Y. S., HUMBURG, P., BATCHELOR, E. L., KAYA, S. M., YOO, H. M., GOODNOW, C. C. & HORIKAWA, K. 2017. Synergistic cooperation and crosstalk between MYD88(L265P) and mutations that dysregulate CD79B and surface IgM. *J Exp Med,* 214**,** 2759-2776.

WREN, D. 2017. *Capture Next-Generation Sequencing as a tool for the simultaneous detection of SNV, CNV, translocations and Immunoglobulin/T-cell receptor rearrangements in B- and T-cell lymphoproliferative diseases.* HSST C1 Innovcation project, Manchester Metropolitan University (MMU).

WREN, D., WALKER, B. A., BRÜGGEMANN, M., CATHERWOOD, M., POTT, C., STAMATOPOULOS, K., LANGERAK, A. W. & GONZALEZ, D. 2014. Translocations and Clonality Detection in Lymphoproliferative Disorders By Capture-Based Next-Generation Sequencing. a Pilot Study By the Euroclonality-NGS Consortium. *Blood,* 124**,** 5169-5169.

WREN, D., WALKER, B. A., BRÜGGEMANN, M., CATHERWOOD, M. A., POTT, C., STAMATOPOULOS, K., LANGERAK, A. W., GONZALEZ, D. & EUROCLONALITY,

N. G. S. C. 2017. Comprehensive translocation and clonality detection in lymphoproliferative disorders by next-generation sequencing. *Haematologica,* 102**,** e57-e60.

XOCHELLI, A., KALPADAKIS, C., GARDINER, A., BALIAKAS, P., VASSILAKOPOULOS, T. P., MOULD, S., DAVIS, Z., STALIKA, E., KANELLIS, G., ANGELOPOULOU, M. K., MCIVER-BROWN, N., IBBOTSON, R., SACHANAS, S., KORKOLOPOULOU, P., ATHANASIADOU, A., ANAGNOSTOPOULOS, A., PAPADAKI, H. A., PAPADAKI, T., STAMATOPOULOS, K., PANGALIS, G. A. & OSCIER, D. 2014. Clonal B-cell lymphocytosis exhibiting immunophenotypic features consistent with a marginal-zone origin: is this a distinct entity? *Blood,* 123**,** 1199-206.

XOCHELLI, A., OSCIER, D. & STAMATOPOULOS, K. 2017. Clonal B-cell lymphocytosis of marginal zone origin. *Best Pract Res Clin Haematol,* 30**,** 77-83.

YASUDA, T., SANADA, M., NISHIJIMA, D., KANAMORI, T., IIJIMA, Y., HATTORI, H., SAITO, A., MIYOSHI, H., ISHIKAWA, Y., ASOU, N., USUKI, K., HIRABAYASHI, S., KATO, M., RI, M., HANDA, H., ISHIDA, T., SHIBAYAMA, H., ABE, M., IRIYAMA, C., KARUBE, K., NISHIKORI, M., OHSHIMA, K., KATAOKA, K., YOSHIDA, K., SHIRAISHI, Y., GOTO, H., ADACHI, S., KOBAYASHI, R., KIYOI, H., MIYAZAKI, Y., OGAWA, S., KURAHASHI, H., YOKOYAMA, H., MANABE, A., IIDA, S., TOMITA, A. & HORIBE, K. 2020. Clinical utility of target capture-based panel sequencing in hematological malignancies: A multicenter feasibility study. *Cancer science,* 111**,** 3367-3378.

ZEHIR, A., BENAYED, R., SHAH, R. H., SYED, A., MIDDHA, S., KIM, H. R., SRINIVASAN, P., GAO, J., CHAKRAVARTY, D., DEVLIN, S. M., HELLMANN, M. D., BARRON, D. A., SCHRAM, A. M., HAMEED, M., DOGAN, S., ROSS, D. S., HECHTMAN, J. F., DELAIR, D. F., YAO, J., MANDELKER, D. L., CHENG, D. T., CHANDRAMOHAN, R., MOHANTY, A. S., PTASHKIN, R. N., JAYAKUMARAN, G., PRASAD, M., SYED, M. H., REMA, A. B., LIU, Z. Y., NAFA, K., BORSU, L., SADOWSKA, J., CASANOVA, J., BACARES, R., KIECKA, I. J., RAZUMOVA, A., SON, J. B., STEWART, L., BALDI, T., MULLANEY, K. A., AL-AHMADIE, H., VAKIANI, E., ABESHOUSE, A. A., PENSON, A. V., JONSSON, P., CAMACHO, N., CHANG, M. T., WON, H. H., GROSS, B. E., KUNDRA, R., HEINS, Z. J., CHEN, H. W., PHILLIPS, S., ZHANG, H., WANG, J., OCHOA, A., WILLS, J., EUBANK, M., THOMAS, S. B., GARDOS, S. M., REALES, D. N., GALLE, J., DURANY, R., CAMBRIA, R., ABIDA, W., CERCEK, A., FELDMAN, D. R., GOUNDER, M. M., HAKIMI, A. A., HARDING, J. J., IYER, G., JANJIGIAN, Y. Y., JORDAN, E. J., KELLY, C. M., LOWERY, M. A., MORRIS, L. G. T., OMURO, A. M., RAJ, N., RAZAVI, P., SHOUSHTARI, A. N., SHUKLA, N., SOUMERAI, T. E., VARGHESE, A. M., YAEGER, R., COLEMAN, J., BOCHNER, B., RIELY, G. J., SALTZ, L. B., SCHER, H. I., SABBATINI, P. J., ROBSON, M. E., KLIMSTRA, D. S., TAYLOR, B. S., BASELGA, J., SCHULTZ, N., HYMAN, D. M., ARCILA, M. E., SOLIT, D. B., LADANYI, M. & BERGER, M. F. 2017. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med,* 23**,** 703-713.

ZHU, S., JIN, J., GOKHALE, S., LU, A. M., SHAN, H., FENG, J. & XIE, P. 2018. Genetic Alterations of TRAF Proteins in Human Cancers. *Frontiers in Immunology,* 9.

ZIBELLINI, S., CAPELLO, D., FORCONI, F., MARCATILI, P., ROSSI, D., RATTOTTI, S., FRANCESCHETTI, S., SOZZI, E., CENCINI, E., MARASCA, R., BALDINI, L., TUCCI, A., BERTONI, F., PASSAMONTI, F., ORLANDI, E., VARETTONI, M., MERLI, M., RIZZI, S., GATTEI, V., TRAMONTANO, A., PAULLI, M., GAIDANO, G. & ARCAINI, L. 2010. Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma. *Haematologica,* 95**,** 1792-1796.

ZUCCA, E. & BERTONI, F. 2016. The spectrum of MALT lymphoma at different sites: biological and therapeutic relevance. *Blood,* 127**,** 2082-2092.

7. Appendix

Table A1: Scope of the ECNDC (EuroClonality Capture NGS) panel. Genes and the region included for the bait design is listed together with the relevance in the context of lymphoid malignancies. The genes were selected by the experts of the EuroClonality Consortium and were deemed to be of clinical interest in acute as well as chronic lymphoid malignancies or of interest across various pan-haem pathways. Highlighted in green are genes in which variants were found in the cohort of patients included in this study.

Gene	Target region	Relevance	Disease	Context
AIRD1A	Full coding region	Prognosis (FL)	FL, LPL/WM, SMZL	Chromatin remodellling
ASXL1	Mutatoin hotpost in last exon	NA	NA	Epigenetic
ATM	Full coding region		CLL, MCL	DNA repair
BCL2	Full coding region	treatment Venetoclax	FL, DLBCL	Anti-apoptotic signalling
BIRC3	Exons 8-9	Prognosis (CLL), Treatment (CLL?MCL)	BL,MCL, CLL, SMZL	NF-kB signalling
BRAF	Exons 11,15	Diagnosis (HCL)	HCL, CLL	MAPK signaling /signal transduction
BTG1	Full coding region	NA	ALL	Transcription factor
BTK	Full coding region	Ibrutinib resistance	FL, CLL, MCL, DLBCL, LPL/WM	BCR signallling
CARD11	Exons 2-16	Prognosis (FL), Treatment	MCL, FL, DLBCL, LPL, T-NHL	BCR/NF-kB signalling
CBL	Exons 7-9		rare in lymhpoid	Signal transduction
CCND3	Exon 5		BL, DLBLCL, SLLU	Cell cycle contol
CD79A	Exons 4,5		Fl, DLBCL, SMZL	BCR signalling
CD79B	exons 4-6		Fl, DLBCL, SMZL, LPL/WM	BCR signalling
CDKN2A	Full coding region	Deletions, prognosis in MCL/DLBCL	MCl, DLBCL, PCNSL	Cell cycle contol
CDKN2B	Full coding region		MCl, DLBCL, PCNSL	Cell cycle contol
CREBBP	Full coding region	Prognosis (FL, DLBCL)	FL, DLBCL, CLL, BL	Chromatin remodellling
CXCR4	Full coding region	Treatment	LPL/WM	Cell signalling receptor
DNMT3A	Exons 8-23		AITL, PTLC-NOS,T -NHL, CHIP	Epigenetic
EGR2	Mutation hotspot in exon 2	prognosis	CLL	Transcription factor
EP300	Full coding region	Prognosis (FL)	FL, DLBCL, SMZL	Chromatin remodellling
ERG	Full coding region		ALL	Transcription factor
EZH2	Full coding region	Prognosis (FL)	FL, DLBCL	Chromatin remodellling
FAT1	Full coding region		across cancer	TM protein of cadherin family
FBXW7	Full coding region	Prognosis (CLL), Treatment (CLL?MCL)	CLL, ETP-ALL	Signal transduction
HIST1H1B	Full coding region		across cancer	Chromatin remodellling
HIST1H1C	Full coding region		across cancer	Chromatin remodellling
HIST1H1D	Full coding region		across cancer	Chromatin remodellling
HIST1H1E	Full coding region		across cancer	Chromatin remodellling
ID3	Full coding region	Diagnosis (BL)	BL, NMZL	Transcription factor
IDH1	Exon 4	potential treatment		Epigenetic
IDH2	Exons 4,7	potential treatment	PTCL, AITL	Epigenetic
IKZF1	Full coding region	Prognosis	ALL	Transcription factor
IL7R	Exon 6		ALL	Signal transduction
JAK1	Exons 12-15			Signal transduction
JAK2	Exons 12, 14			Signal transduction
JAK3	Exons 12-14		T-NHL	Signal transduction
KIT	Exons 9,11,13,17,18	Diagnosis	SM	Transcription factor
KLF2	Full coding region	Diagnosis (MZL)	SMZL, NMZL, HCL, MCL	Transcription factor
KMT2A	Full coding region			Epigenetic
KMT2D	Full coding region	Prognoisis and treatment	FL, DLBCL	Epigenetic
KRAS	Exons 2-4		MM, DLBCL	Signal transduction
MAP2K1	Exons 1-8	Diagnosis	HCLv, paed FL, Histiocytosis	Signal transduction
MAP3K14	Full coding region	Diagnosis	Histiocytosis	Signal transduction
MYC	Full coding region	Prognosis/Treatment	BL, MCL, DLBCL	Cell cycle contol
MYD88	Exon 5	Diagnosis/ Treatment	LPL/WM, DLBCL, MZl	NfkB signalling
NFKBIE	Exon 1	Diagnosis	HL, PMBL	NfkB signalling
NOTCH1	Exons 26, 27, 34	Prognosis (CLL, MCL, T-ALL)	CLL, MCL, T-ALL	Signal transduction
NOTCH2	Exon 34	diagnosis/prognosis (MZL/MCL)	SMZL, NMZL, LPL/WM	Signal transduction
NRAS	Exons 2-4	Diagnosis	Histiocytosis	Signal transduction
NT5C2	Exons 9-15		ALL, myeloid	Nucleotide processing
PAX5	Full coding region		ALL	Transcription factor
PHF6	Full coding region		(ETP) T-ALL	Cell cycle contol
PIK3CA	Exons 2,5,10, 21	Risk progression	MCL	Signal transduction
PLCG2	Exons19, 20, 24	Ibrutinib resistance	FL, CLL, MCL, DLBCL, LPL/WM	BCR signallling
POT1	Full coding region		CLL	Telomere stabilitys
PTEN	Full coding region		T-ALL, DLBCL	Signal transduction
RHOA	Full coding region	Diagnosis (AITL)	T-NHL	Signal transduction
RUNX1	Full coding region		ALL, myeloid	Transcription factor
SAMHD1	Full coding region		CLL	Nucleotide processing
SF3B1	Exons 12-17	Diagnosis	CLL	Spliceosome
SOCS1	Full coding region		DLBCL	JAK/STAT signalingpathway
STAT3	Exons 10-16, 20,21	Diagnosis	T-NHL	Signal transduction
STAT5B	Exons 13-19	Diagnosis	T-NHL	Signal transduction
TCF3	Full coding region	Diagnosis (BL)	BL	Transcription factor
TET2	Full coding region	Diagnosis (AITL)	AITL and other T-NHL, CHIP	Epigenetic
TNFAIP3	Full coding region	Diagnosis/Treatment	HL, PMBL, DLBCL, NMZL, SZML	NF-kB signalling
TP53	Full coding region	Prognosis, Treatment (CLL)	All cancers	Cell cycle control
TRAF2	Full coding region	Diagnosis	LPL/WM, SMZL, DLBCL, CLL	NF-kB signalling
TRAF3	Full coding region	Diagnosis	LPL/WM, SMZL, DLBCL, CLL	NF-kB signalling
WT1	Full coding region	?	ALL, myeloid	Transcription factor
XPO1	Exons 15-16	Diagnosis/Treatment	CLL, mediastinal DLBCL, HL	RNA metabolism

Table A2: IG- and TR-genes including in the bait design

IGH chr14	IGHD1-1	IGHD4-17
	IGHD2-2	IGHD5-18
	IGHD3-3	IGHD6-19
	IGHD4-4	IGHD2-21
	IGHD5-5	IGHD3-22
	IGHD6-6	IGHD4-23
	IGHD1-7	IGHD5-24
	IGHD2-8	IGHD6-25
	IGHD3-9	IGHD1-26
	IGHD3-10	
	IGHD4-11	IGHJ1
	IGHD5-12	IGHJ2
	IGHD6-13	IGHJ3
	IGHD1-14	IGHJ4
	IGHD2-15	IGHJ5
	IGHD3-16	IGHJ6

IGK chr 2	IGKJ1	Kde
	IGKJ2	IntronRSS
	IGKJ3	
	IGKJ4	
	IGKJ5	

IGL chr22	IGLJ1
	IGLJ2
	IGLJ3
	IGLJ4
	IGLJ5
	IGLJ6
	IGLJ7

IGHswitch	IGHA2
	IGHE
	IGHG4
	IGHG2
	IGHA1
	IGHG1
	IGHG3

TRA chr14	TRAJ1	TRAJ26	TRAJ52
	TRAJ2	TRAJ27	TRAJ53
	TRAJ3	TRAJ28	TRAJ54
	TRAJ4	TRAJ29	TRAJ56
	TRAJ5	TRAJ30	TRAJ57
	TRAJ6	TRAJ31	TRAJ58
	TRAJ7	TRAJ32	TRAJ59
	TRAJ8	TRAJ33	TRAJ61
	TRAJ9	TRAJ34	
	TRAJ10	TRAJ35	
	TRAJ11	TRAJ36	
	TRAJ12	TRAJ37	
	TRAJ13	TRAJ38	
	TRAJ14	TRAJ39	
	TRAJ15	TRAJ40	
	TRAJ16	TRAJ41	
	TRAJ17	TRAJ42	
	TRAJ18	TRAJ43	
	TRAJ19	TRAJ44	
	TRAJ20	TRAJ45	
	TRAJ21	TRAJ46	
	TRAJ22	TRAJ47	
	TRAJ23	TRAJ48	
	TRAJ24	TRAJ49	
	TRAJ25	TRAJ50	

TRD chr14	TRDD1	TRDJ1
	TRDD2	TRDJ2
	TRDD3	TRDJ3
		TRDJ4

TRB chr7	TRBD1	TRBJ1-1	TRBJ2-1
	TRBD2	TRBJ1-2	TRBJ2-2
		TRBJ1-3	TRBJ2-3
		TRBJ1-4	TRBJ2-4
		TRBJ1-5	TRBJ2-5
		TRBJ1-6	TRBJ2-6
			TRBJ2-7

TRG chr7	TRGJ1
	TRGJ1P
	TRGJP
	TRGJ2
	TRGJ2P

Table A3: Regions included in bait design for CNV assessment

Target	Chromosome
PIK3CA	chr3
IKZF1	chr7
7p TEL	chr7
MYC	chr8
8q TEL	chr8
9p TEL	chr9
CDKN2A	chr9
CDKN2B	chr9
PAX5	chr9
9p CEN	chr9
11q CEN	chr11
CCND1	chr11
ATM	chr11
11q TEL	chr11
12p TEL	chr12
ETV6	chr12
12p CEN	chr12
12q CEN	chr12
BTG1	chr12
12q TEL	chr12
13q CEN	chr13
RB1	chr13
13q TEL	chr13
17p TEL	chr17
TP53	chr17
17p CEN	chr17

Table A4: Regions included in bait design for translocation detection

Target	Chromosome
STIL	chr1
ALK	chr2
BCL6	chr3
BCL1	chr11
BIRC3	chr11
KMT2A	chr11
BCL2	chr18
CRLF2	chrX

Table A6: Regions with suboptimal coverage. Regions achieved less than 500x across samples and these areas were already identified as poor performers in the EuroClonality Validation study (Stewart P. et al. Blood Advances *in press*). In the validation study targets were highlighted as suboptimal coverage if coverage was >2 standard deviations (SD) below the mean in ≥ 50% samples. Apart from *TCF3* exon 2 and 9 all other targets are CNV SNPs with additional targets included in the design.

Chromosome	Start	End	Target
chr8	140542050	140542150	8qTEL_rs2977475
chr8	140549209	140549309	8qTEL_rs2977490
chr8	140556338	140556438	8qTEL_rs2271736
chr11	132697125	132697225	11qTEL_rs1940150
chr12	43448951	43449051	12qCEN_rs2220865
chr12	131033171	131033271	12qTEL_rs7960677
chr12	131030078	131030178	12qTEL_rs1195889
chr13	23741443	23741543	13qCEN_rs36116586
chr13	48457646	48457746	RB1CN_rs9568042
chr13	48461409	48461509	RB1CN_rs9568043
chr17	714749	714849	17pTEL_rs2474694
chr17	20211440	20211540	17pCEN_rs2703812
chr19	1622303	1622425	TCF3 exon 9 NM_003200.5
chr19	1650158	1650258	TCF3 exon 2 NM_003200.5

Table A7 (Page 155-158): Details of the V(D)J rearrangements at the IGH and IGK/L loci for each sample including SHM status and light chain restriction. Results from NGS, Sanger sequencing and IHC/Immunophenotyping analysis.

ID	Allele	IGHV	IGHD	IGHJ	SHM%	IGKV	IGKJ	IGLV	IGLJ	Predicted light chain (NGS)	Light chain (Flow, IHC, serum paraprotein)	Comments
BEL-4383-0001	1 (functional)	IGHV1-2		IGHJ6	0.3	IGKV1-8	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)					IGKV2-29	IGKJ2					
BEL-4383-0002	1 (functional)	IGHV4-34=IGHV4-61		IGHJ3	>10	IGKV3-20	IGKJ4	IGLV1-44	IGLJ3	lambda	Lambda (Flow)	IGHV assignment uncertain likely due to high SHM load
	2 (non-functional)		IGHD3-22	IGHJ2		IGKV4-1	Kde					
						Kde	IntronRSS					
BEL-4383-0003	1 (functional)	IGHV3-7=IGHV3-11		IGHJ4	6.2	IGKV3-15	IGKJ4			kappa	Kappa weak (Flow) and	CLL score of 4 shouldn't have been enrolled
	2 (non-functional)										Lambda (IHC)	Possible additional MGUS
BEL-4383-0005	1 (functional)	IGHV5-51		IGHJ4	5.9	IGKV1-5	Kde	IGLV2-14	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD5-24	IGHJ6		IntronRSS	Kde					
BEL-4383-0006	1 (functional)	IGHV3-7		IGHJ4	3.8			IGLV1-40	IGLJ1	lambda	Unknown	
	2 (non-functional)		IGHD3-10	IGHJ4		IGKV2-18	Kde	IGLV7-43	IGLJ3			
BEL-4383-0007	1 (functional)	IGHV3-53		IGHJ5	8.1	IGKV2-28	IGKJ4			kappa	Kappa (Flow)	Productive IGK could not be identified
	2 (non-functional)		IGHD2-21	IGHJ3		IGKV1-12	IGKJ4					
						IntronRSS	Kde					
BEL-4383-0008	1 (functional)	IGHV3-23		IGHJ5	5.2	IGKV1-16	IGKJ4			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD1-26	IGHJ6								
R1K-4383-0001	1 (functional)	IGHV4-4=IGHV4-39		IGHJ2	2.1	IGKV1-39	IGKJ4			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD5-24	IGHJ4								
R1K-4383-0002	1 (functional)	IGHV5-51		IGHJ4	3.1	IGKV3-20	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ3								
R1K-4383-0004	1 (functional)	IGHV3-21		IGHJ4	6.9			IGLV1-44	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD6-13	IGHJ5		IGKV2-29	IGKJ2					
						Kde	IntronRSS					
R1K-4383-0007	1 (functional)	IGHV6-1		IGHJ3	0			IGLV1-51	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-22	IGHJ3		IGKV2-30	IGKJ4					
RA2-4383-0001	1 (functional)	IGHV4-34		IGHJ4	NA	IGKV1-27	IGKJ1			kappa	Kappa (Flow)	PCR/Sanger approach amplified two V4 fragments, RNA not tested
	2 (non-functional)		IGHD6-6	IGHJ4								
RA2-4383-0002	1 (functional)	IGHV3-13=V3-74=V6-1		IGHJ4	2.8	IGKV1-8	Kde	IGLV2-11	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)	IGHV1-8		IGHJ6		IGKV1-17	Kde	IGLV9-49	IGLJ2=IGLJ3			
RA2-4383-0004	1 (functional)	IGHV3-7		IGHJ4	4.9	IGKV4-1	IGKJ2			kappa	Kappa (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD6-19	IGHJ4		IGKV3-11	Kde					
RA2-4383-0005	1 (functional)	IGHV3-23		IGHJ5	10.7	IGKV1-13	IGKJ4	IGLV1-47	IGLJ7	lambda	Lambda (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD2-2	IGHJ6		IGKV2-29	IGKJ4					
	3					Kde	IntronRSS					
RA2-4383-0006	1 (functional)	IGHV2-5		IGHJ4	1.7	IGKV2-30	IGKJ5	IGLV3-10	IGLJ1	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-3	IGHJ4		IGKV2-26	Kde					
						Kde	IntronRSS					
RAX-4383-0001	1 (functional)	IGHV3-30		IGHJ6	8.3						Lambda (Flow)	NGS failed
RAX-4383-0002	1 (functional)	IGHV3-15		IGHJ5	10.1	IGKV1-33	IGKJ3	IGLV3-21	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ1		IGKV7-3	Kde					
RAX-4383-0003	1 (functional)	IGHV4-34		IGHJ4	0	IGKV3-20	IGKJ5			kappa	Kappa (Flow)	
	2 (non-functional)					IGKV2-29	IGKJ1					
RAX-4383-0004	1 (functional)	IGHV4-34		IGHJ5	1.4	IGKV3-11	IGKJ5			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD3-22	IGHJ1								
RAX-4383-0005	1 (functional)	IGHV4-34		IGHJ5	6.3	IGKV4-1	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)											
RAX-4383-0006	1 (functional)	IGHV3-30		IGHJ6	3.8	IGKV4-1	IGKJ1			kappa	Unknown	NGS could not identify IGHV and whether productive/not productive; CDR3 amino acid sequence not available through NGS but Sanger productive for IGHV3-30 rearrangement
	2 (non-functional)	IGHV1-69		IGHJ4								
RAX-4383-0007	1 (functional)	IGHV3-15		IGHJ4/5	8.8	IGKV2-29	IGKJ2	IGLV7-46	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)					IGKV3(D)-11	IGKJ5	IGLV3-10	IGLJ2=IGLJ3			
	3					Kde	IntronRSS					
RDD-4383-0001	1 (functional)	IGHV1-8 or IGHV1-46		IGHJ4	0	IGKV3-11	IGKJ1			kappa	Kappa (Flow)	NGS shows two productive rearrangements with J4 and different V genes; Sanger has IGHV1-8 but unprodctive according to IMGT as Trp118 not identfied
	2 (non-functional)	IGHV3-23		IGHJ4								

ID	Allele	IGHV	IGHD	IGHJ	SHM%	IGKV	IGKJ	IGLV	IGLJ	Predicted light chain (NGS)	Light chain (Flow, IHC, serum paraprotein)	Comments
RDD-4383-0002	1 (functional)	IGHV1-69		IGHJ4	11.5	IGKV1-5	IGKJ1			kappa	Kappa (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD1-7	IGHJ4								
RDD-4383-0003	1 (functional)	IGHV3-74		IGHJ4	4.5	IGKV3-20	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)											
RDD-4383-0004	1 (functional)	IGHV3-7		IGHJ6	5.6	IGKV1-9=IGK	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ5								
RDU-4383-0001	1 (functional)	IGHV4-39		IGHJ1	NA	IGKV4-1	IGKJ4			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD6-6	IGHJ6		IGKV1-5	IGKJ4					
						intron	Kde					
RDU-4383-0002	1 (functional)	IGHV3-15		IGHJ1	10.5	(IGKV2-26)	(Kde)			(lambda)	Lambda (Flow)	PCR/Sanger with the same rearrangement found by NGS but low coverage overall
	2 (non-functional)											
RDU-4383-0003	1 (functional)	IGHV3-48		IGHJ4	4.9	IGKV3-15	IGKJ5			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-15	IGHJ4								
RDU-4383-0004	1 (functional)	IGHV3-72		IGHJ4	NA	IGKV1-16	IGKJ4			kappa	Unknown	PCR/Sanger failed
	2 (non-functional)		IGHD5-12	IGHJ6		IGKV2-24	Kde					
RDU-4383-0005	1 (functional)	IGHV3-7		IGHJ4	2.4	IGKV4-1	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)											
REF-4383-0002	1 (functional)	?iGHV3-30		IGHJ4	NA			IGLV2-14	IGLJ2/3	lambda	Unknown	IGHV gene not identified; neither NGS nor PCR/Sanger detected IGHV likley due to high SHM
	2 (non-functional)	IGHV-1-8/1-18?		IGHJ4		IGKV1-33	Kde					
REF-4383-0003	1 (functional)	IGHV1-69-2		IGHJ6	0	IGKV2-28	IGKJ1			kappa	Unknown	
	2 (non-functional)		IGHD6-13	IGHJ3								
REF-4383-0005	1 (functional)	IGHV1-69		IGHJ5	0	IGKV1-16	IGKJ5			kappa	Unknown	
	2 (non-functional)		IGHD3-9	IGHJ4		IGKV1-37	IGKJ1					
REF-4383-0006	1 (functional)	IGHV1-2		IGHJ4	NA	IGKV3-20	IGKJ3			kappa	Unknown	Two rearrangements even on RNA, NGS data not clear whether productive; likely high SHM or biclonal
	2 (non-functional)	IGHV4-39		IGHJ4								
REF-4383-0007	1 (functional)	?	?	?	NA	IGKV4-1	IGKJ3			kappa	Unknown	IGHV gene not identified; neither NGS nor PCR/Sanger detected IGHV likley due to high SHM
	2 (non-functional)	?	?	?		IGKV2-30	IGKJ2/3					
REF-4383-0008	1 (functional)	IGHV3-48		IGHJ4	1	IGKV1-33	IGKJ1			kappa	Unknown	
	2 (non-functional)	IGHV1-3		IGHJ1		IGKV4-1	Kde					
REF-4383-0009	1 (functional)	IGHV3-9		IGHJ4	4.8	IGKV1-33	IGKJ5	IGLV1-51	IGLJ3	lambda	Unknown	
	2 (non-functional)		D2-15	IGHJ5		IGKV3-20	Kde					
	3					Kde	IntronRSS					
REF-4383-0010	1 (functional)	IGHV3-72		IGHJ3	2.4	IGKV4-1	IGKJ3			kappa	Unknown	
	2 (non-functional)		IGHD6-13	IGHJ3								
REF-4383-0011	1 (functional)	IGHV3-7 or 3-11		IGHJ6	2	IGKV3-15	IGKJ5			kappa	Unknown	
	2 (non-functional)		D3-9	IGHJ5								
REF-4383-0012	1 (functional)	IGHV4-34		IGHJ4	0	IGKV1-17	IGKJ1			kappa	Unknown	
	2 (non-functional)	IGHV3-33		IGHJ4								
REF-4383-0013	1 (functional)	IGHV4-59		IGHJ5	3.9	IGKV4-1	IGKJ4			kappa	Unknown	
	2 (non-functional)											
REF-4383-0014	1 (functional)	IGHV3-7		IGHJ4	11.5	IGKV4-1	IGKJ4			kappa	Unknown	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)					IGKV1-8	Kde					
REF-4383-0016	1 (functional)	IGHV1-8		IGHJ5	0.3	IGKV4-1	IGKJ2			kappa	Unknown	
	2 (non-functional)		IGHD1-7	IGHJ6		IGKV2-29	IGKJ1					
RHU-4383-0001	1 (functional)	IGHV3-72		IGHJ6	NA	IGKV1-16	IGKJ4			kappa	Kappa (Flow)	No PCR/Sanger results- repeat on RNA required
	2 (non-functional)		IGHD2-2	IGHJ6		IGKV4-1	IGKJ2					
RHU-4383-0002	1 (functional)	IGHV3-15		IGHJ4	4.4			IGLV2-18	IGLJ1	lambda	IgG Lambda (Paraprotein)	
	2 (non-functional)		IGHD2-15	IGHJ4		IGKV3-11	Kde					
RHU-4383-0003	1 (functional)	IGHV3-15		IGHJ4	11.9	IGKV7-3	IGKJ4	IGLV3-10	IGLJ3/J2	lambda	Unknown	
	2 (non-functional)		IGHD2-15	IGHJ4		IGKV4-1	Kde					
						Kde	IntronRSS					
RHU-4383-0004	1 (functional)	IGHV3-9		IGHJ3	NA	IGKV3-20	IGKJ3	IGLV1-51	IGLJ2=IGLJ3	lambda	IgG Lambda (Paraprotein)	PCR/Sanger only amplifies the unproductive rearragement; RNA not performed
	2 (non-functional)	IGHV3-52		IGHJ3		IGKV2-24	IGKJ4	IGLV4-3	IGLJ2=IGLJ3			
	3					Kde	IntronRSS					
RHU-4383-0005	1 (functional)	IGHV3-23		IGHJ4	0	IGKV1-17	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ6		intron	Kde					

ID	Allele	IGHV	IGHD	IGHJ	SHM%	IGKV	IGKJ	IGLV	IGLJ	Predicted light chain (NGS)	Light chain (Flow, IHC, serum paraprotein)	Comments
RHU-4383-0006	1 (functional)	IGHV4-59=IGHV4-61		IGHJ6	5.3	IGKV5-2	IGKJ2			kappa	Unknown	
	2 (non-functional)					IntronRSS	Kde					
RJ6-4383-0001	1 (functional)	IGHV3-23		IGHJ3	NA	IGKV3-11	IGKJ5			kappa	Kappa (Flow)	No PCR/Sanger results- likley due to mutation under primer binding site
	2 (non-functional)		IGHD3-22	IGHJ6								
RJ6-4383-0002	1 (functional)	IGHV3-30		IGHJ4	13	IGKV4-1	IGKJ4	IGLV5-45	IGLJ3	lambda	Lambda (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)					IGKV3-11	Kde					
	3					intron	Kde					
RJ6-4383-0003	1 (functional)	IGHV3-30		IGHJ6	0.7	IGKV1-33	IGKJ2	IGLV7-46	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)	IGHV6-1		IGHJ5		IGKV4-1	IGKJ4					
						IntronRSS	Kde					
RMH-4383-0001	1 (functional)	IGHV4-34		IGHJ2	5.6						Kappa (Flow)	no NGS IGHV analysis performed
RMH-4383-0002	1 (functional)	IGHV1-69		IGHJ3	6.6						Lambda (Flow)	no NGS IGHV analysis performed
RMH-4383-0004	1 (functional)	IGHV3-74		IGHJ4	6.9						Unknown	no NGS IGHV analysis performed
RMH-4383-0005	1 (functional)	IGHV1-2/IGHV1-58		IGHJ4	2	IGKV1-33	IGKJ3	IGLV3-27	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ6		IntronRSS	Kde					
RMH-4383-0006	1 (functional)	IGHV3-15		IGHJ4	5.8						Lambda (Flow)	no NGS IGHV analysis performed
RMH-4383-0007	1 (functional)	IGHV4-34		IGHJ4	7	IGKV1-33	Kde	IGLV5-45	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD6-25	IGHJ4								
RMH-4383-0008	1 (functional)	IGHV4-59		IGHJ4	3.5	IGKV4-1	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD3-10	IGHJ5		IGKV2-28	Kde					
RMH-4383-0009	1 (functional)	IGHV1-69		IGHJ3	6.5	IGKV1-27	IGKJ3			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ4		IGKV1-5	IGKJ2					
RMH-4383-0010	1 (functional)	IGHV3-23		IGHJ5	2.8	IGKV4-1	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-15	IGHJ4								
RMH-4383-0011	1 (functional)	IGHV3-9	IGHJ6		8.7	IGKV4-1	IGKJ4	IGLV1-51	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)					IGKV2-30	IGKJ2					
						IntronRSS	Kde					
RMH-4383-0012	1 (functional)	IGHV3-30=IGHV3-64	IGHJ4		7.3	IGKV4-1	IGKJ3			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-15	IGHJ5								
RMH-4383-0013	1 (functional)	IGHV4-34		IGHJ2	5.6	IGKV1-16	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-8	IGHJ4		IGKV1-5	IGKJ1					
RMH-4383-0014	1 (functional)	IGHV4-34		IGHJ4	NA	IGKV3-15	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)					IGKV4-1	IGKJ2					
RMH-4383-0015	1 (functional)	IGHV4-39		IGHJ6	10	IGKV4-1	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD6-19	IGHJ4			x					
RTP-4383-0001	1 (functional)	IGHV4-59		IGHJ4	5.5	IGKV4-1	IGKJ4	IGLV3-21	IGLJ3	lambda	Lambda (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)					IGKV1-17	Kde					
	3					IntronRSS	Kde					
RTP-4383-0002	1 (functional)	IGHV3-48		IGHJ5	8.7	IGKV2-30	Kde	IGLV4-69	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ5		IGKV5-2	Kde					
RTP-4383-0003	1 (functional)	IGHV3-48		IGHJ6	0	IGKV1-33	Kde	IGLV1-40	IGLJ1	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD2-21	IGHJ5		IGKV5-2	Kde	IGLV5-45	IGLJ3			
RTP-4383-0004	1 (functional)	IGHV4-34		J4/J5	10.5	IGKV1-12	IGKJ4			kappa	Kappa (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD2-2	IGHJ5		IGKV1-39	IGKJ3					
RTP-4383-0005	1 (functional)	IGHV2-70		IGHJ5	0.3	IGKV4-1	IGKJ1	IGLV2-23	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		D5-18/5-5/5-2	IGHJ4		IGKV2-30	Kde					
	3					IGKV4-1	IntronRSS					
RTP-4383-0006	1 (functional)	IGHV4-34		IGHJ6	4.2	IGKV3-20	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ4		IGKV2-30	Kde					
RTP-4383-0007	1 (functional)	IGHV3-23		IGHJ4	5.6	IGKV4-1	IGKJ4			kappa	Kappa (Flow)	
	2 (non-functional)	IGHV3-30/3-66		IGHJ4		IGKV3-15	Kde					
RTP-4383-0009	1 (functional)	IGHV3-21		IGHJ5	1.5	IGKV5-2	IGKJ4	IGLV3-19	IGLJ2=IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ4		IGKV1-33	IGKJ4					
						IntronRSS	Kde					
RVR-4383-0001	1 (functional)	IGHV3-66		IGHJ4	8.8	IGKV1-17	IGKJ3			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD4-23	IGHJ2								

ID	Allele	IGHV	IGHD	IGHJ	SHM%	IGKV	IGKJ	IGLV	IGLJ	Predicted light chain (NGS)	Light chain (Flow, IHC, serum paraprotein)	Comments
RVR-4383-0002	1 (functional)	IGHV3-23		IGHJ4	4.5	IGKV4-1	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)											
RVR-4383-0003	1 (functional)	IGHV1-3		IGHJ4	NA	IGKV4-1	IGKJ4	IGLV2-8	IGLJ2=IGLJ3	lambda	Lambda (Flow)	IGHV uncertain due to very low read support by NGS; likley low tumour infiltration: PCR/Sanger looks polyclonal
	2 (non-functional)		IGHD2-21	IGHJ4		IGKV2-28	IGKJ3					
RVR-4383-0004	1 (functional)	IGHV3-11		IGHJ6	9	IGKV2-28	IGKJ5			kappa	Kappa (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)					IGKV2-29	Kde					
RVR-4383-0005	1 (functional)	IGHV3-23		IGHJ4	>10	IGKV2-24	IGKJ4			kappa	Kappa (Flow)	PCR/Sanger only amplifies the unproductive rearragement; RNA not performed
	2 (non-functional)	IGHV3-7/other		IGHJ4								
RVR-4383-0006	1 (functional)	IGHV3-73		IGHJ4	11.6	IGKV4-1	IGKJ3			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD1-26	IGHJ4		IGKV3-20	IGKJ3					
						IntronRSS	Kde					
RVR-4383-0007	1 (functional)	IGHV2-70		IGHJ4	4.1	IGKV3-20	IGKJ4			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD1-26	IGHJ1								
RVR-4383-0008	1 (functional)	IGHV3-15		IGHJ6	3.4	IGKV1-39	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ6								
RVR-4383-0009	1 (functional)	IGHV2-5		IGHJ5	7.9	IGKV1-5	IGKJ5			kappa	Kappa (Flow)	suboptimal NGS mean coverage
	2 (non-functional)											
RVR-4383-0010	1 (functional)	IGHV3-30		IGHJ4	7	IGKV3-20	IGKJ1			kappa	Kappa (Flow)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD3-9	IGHJ6		IntronRSS	Kde					
RVR-4383-0011	1 (functional)	IGHV4-61		IGHJ5	2.5	IGKV1-37	IGKJ5	IGLV1-40	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ4								
RVR-4383-0012	1 (functional)	IGHV1-18		IGHJ4	5.2						Kappa (Flow)	
RVR-4383-0013	1 (functional)	IGHV4-34		IGHJ6	11	IGKV3-15	Kde	IGLV2-8	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-9	IGHJ4		IGKV4-1	Kde					
RVR-4383-0014	1 (functional)	IGHV3-33=IGHV3-48		IGHJ4	2.1	IGKV4-1	IGKJ3			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD6-25	IGHJ3		IGKV1-8	Kde					
RVR-4383-0015	1 (functional)	IGHV1-2		IGHJ4	2.1	IGKV3-20	IGKJ1			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD3-16	IGHJ4		IGKV3-20	IGKJ4					
RVR-4383-0016	1 (functional)	IGHV4-34		IGHJ6	0	IGKV1-8	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD1-26	IGHJ5								
RVR-4383-0017	1 (functional)	IGHV3-72		IGHJ5	6.1	IGKV4-1	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD2-15	IGHJ4		IGKV1-33	Kde					
RVR-4383-0018	1 (functional)	IGHV3-30		IGHJ4	5.9	IGKV4-1	IGKJ2			kappa	Kappa (IHC)	IGHV gene assignment based on PCR/Sanger; identical CDR3 between NGS and PCR/Sanger
	2 (non-functional)		IGHD2-21	IGHJ3								
RVR-4383-0019	1 (functional)	IGHV1-2		IGHJ6	1.4	none		IGLV2-8	IGLJ1	lambda	Lambda (Flow)	NGS could not identify productive versus non productive IGHV, no evidence of kappa rearrangement; Sanger supports IGHV1-2 as productive
	2 (non-functional)	IGHV3-53		IGHJ4								
RVR-4383-0020	1 (functional)	IGHV3-48		IGHJ5	0.7	IGKV1-37	IGKJ3	IGLV3-10	IGLJ3/J2	lambda	Lambda (Flow)	
	2 (non-functional)					IGKV1-39	IGKJ5	IGLV1-47	IGLJ3/J3			
						Kde	IntronRSS					
RVR-4383-0021	1 (functional)	IGHV4-4		IGHJ2	2.4	IGKV2-29	IGKJ3	IGLV7-46	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD2-2	IGHJ5		IGKV2-26	IGKJ4	IGLV4-60	IGLJ3			
	3					Kde	IntronRSS					
RVR-4383-0022	1 (functional)	IGHV1-2*2		IGHJ4	NA	IGKV3-20	IGKJ3			kappa	Kappa (Flow)	No PCR/Sanger results- likley due to mutation under primer binding site
	2 (non-functional)		IGHD3-9	IGHJ6								
RVR-4383-0023	1 (functional)	IGHV4-34		IGHJ6	7.7	IGKV3-11	IGKJ4	IGLV1-44	IGLJ1	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD5-18/5-5	IGHJ4		IGKV1-33	IGKJ5					
	3					Kde	IntronRSS					
RVR-4383-0024	1 (functional)	IGHV3-64		IGHJ4	NA	(IGKV1-16)	(IGKJ4)			(lambda)	Unknown	Uncertain due to very low read support by NGS; likley low tumour infiltration
	2 (non-functional)					(IGKV2-29)	(Kde)					Pathogenic variant detected by another lab in another samples which was not present here
RVR-4383-0025	1 (functional)	IGHV3-73		IGHJ4	6.5	IGKV1-16	IGKJ4	IGLV3-25	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD3-16	IGHJ4		IGKV1-39	Kde					
						IntronRSS	Kde					
RVR-4383-0026	1 (functional)	IGHV4-4		IGHJ5	9.5			IGLV7-43	IGLJ3	lambda	Lambda (Flow)	
	2 (non-functional)		IGHD5-12	IGHJ6		IGKV3-20	Kde					
RVR-4383-0027	1 (functional)	IGHV3-23		IGHJ6	4.2	IGKV1-12	IGKJ2			kappa	Kappa (Flow)	
	2 (non-functional)		IGHD5-24	IGHJ4		IGKV1-33	Kde					

Table A8: Reference sequence for genes

Gene	Transcript
ARID1A	NM_006015.4
ATM	NM_000051.3
BIRC3	NM_001165.4
BIRC3	NM_001165.4
BRAF	NM_004333.4
BTK	NM_001287344.1
CARD11	NM_032415.5
CBL	NM_005188.3
CCND3	NM_001760.4
CD79B	NM_001039933.1
CDKN2A	NM_000077.4
CREBBP	NM_004380.2
CXCR4	NM_001008540.1
DNMT3A	NM_175629.2
EP300	NM_001429.3
FAT1	NM_005245.3
FBXW7	NM_033632.3
HIST1H1C	NM_005319.3
HIST1H1D	NM_005320.2
HIST1H1E	NM_005321.2
HIST1H1E	NM_005321.2
KLF2	NM_016270.2
KMT2D	NM_003482.3
KRAS	NM_033360.3
MAP3K14	NM_003954.4
MYD88	NM_001172567.1
NFKBIE	NM_004556.2
NOTCH1	NM_017617.4
NOTCH2	NM_024408.3
NRAS	NM_002524.4
PAX5	NM_016734.2
POT1	NM_015450.2
SAMHD1	NM_015474.3
SF3B1	NM_012433.3
TNFAIP3	NM_006290.3
TP53	NM_000546.5
TRAF2	NM_021138.3
TRAF3	NM_003300.3

Appendix A9: Write-up of HSST C1 Innovation project evaluating capture-based NGS like EC-NDC as a tool for routine diagnostics

HSST C1 Project

Dorte Wren

30.6.2017

Capture Next-Generation Sequencing as a tool for the simultaneous detection of SNV, CNV, translocations and Immunoglobulin/T-cell receptor rearrangements in B- and T-cell lymphoproliferative diseases

1. Background

Next-generation sequencing (NGS) technology has primarily been used in a research context to broaden our understanding of genomic aberrations underlying different cancer entities and examples of pivotal studies performed in haematological cancers are listed in table 1. As our understanding of the relevance of mutations with regards to specific disease pathologies, prognosis and therapy response continues to increase, so does the repertoire of mutations and genomic variants to be assessed in order to optimise patient management (Vaque *et al.* 2013, Rosenquist *et al*., 2016). Over the last 5 years NGS has therefore started to move into clinic (Gray *et al.,* 2015, Cottrell *et al*., 2014) as a tool capable of interrogating large numbers of regions of interest in parallel, with high sensitivity and accuracy.

At present morphology, immunophenotyping and histopathological assessments are the fundamental modalities in the investigation and diagnosis of B- and T cell neoplasms with only a few pivotal genetic markers for diagnosis (table 2, NICE (2016), Swerdlow *et al.,* 2008 and 2016). Gaining in importance are changes of the DNA structure at chromosomal and nucleotide level which can help with the diagnosis but are especially important for accurate prognosis and optimising treatment choice (predictive testing) enabling us to move towards personalised medicine (examples given in table 3, Vaque *et al.* 2013, Rosenquist *et al*., 2016).

This review will focus on the applicability and potential advantages of using targeted capture NGS (from now on referred to as CapNGS, figure 1) for the detection of copy number variants (CNVs), translocations and Immunoglobulin-gene (IG) and T-cell receptor (TCR) gene rearrangements in addition to single nucleotide variants (SNVs), with the view to implement this method for evaluation of possible and known lymphomas in the clinic. As the review is not aimed to discuss the relevance of molecular testing in lymphomas table 3 provides some examples of how a broader molecular approach could benefit patients.

Table 1: Studies performing NGS analysis to identify new genetic lesions and prevalence of molecular aberrations in lymphoid malignancies and relevant reviews.

Lead author (et al.)	Disease entity	Year published
Du M.Q.	MALT	2017
Wang M.	AITL	2017
Bogusz A.M.	MCL	2016
Karmali R.	DLBCL	2017
Kiel M.J.	SMZL	2012
Rossi D.	CLL	2017

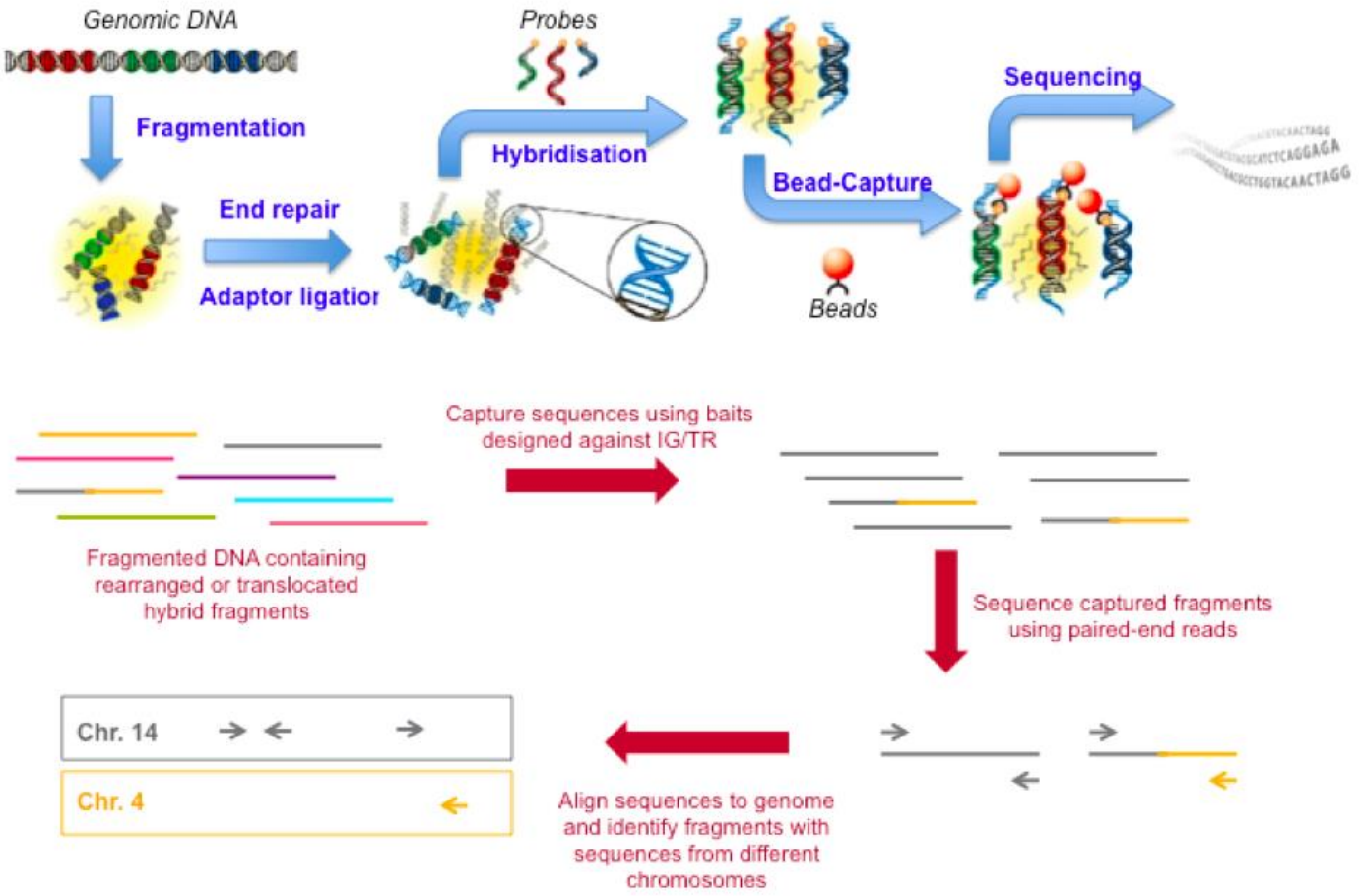


Figure 1: Depiction of the CapNGS workflow using liquid bead capture (top). Genomic DNA is fragmented and NGS platform adaptors, barcodes for patient identification and PCR/Sequencing primer binding sites added (Library preparation). Biotinylated baits (also referred to as probes) designed to be complementary to the regions of interest are added and allowed to hybridise to the libraries. Following hybridisation, streptavidin-coated, magnetic beads allow the selection of DNA-Bait dimers which should represent the regions of interest and undergo NGS sequencing.

Explanation of the principle of translocation detection using CapNGS (bottom). In this example, baits would have been designed against the IGH loci on chromosome 14 (in grey) and only DNA fragments containing 'grey' DNA sequence will hybridise to the baits and be sequenced. If a translocation is present, some of the DNA fragments generated in the first step will contain DNA sequences from the IGH locus as well as the partner chromosomes, in this case chromosome 4 (yellow). The 'yellow' part of the fragment will be subjected to sequencing even though no baits were designed for its detection, enabling CapNGS approaches to detect any (previously unknown) translocation partner.

Table 2: Genetic markers employed for diagnosing haematological malignancies

World health organisation (WHO) classification (Swerdlow *et al.* 2008 and 2016)	**Diagnostic marker**	**Frequency** (from Swerdlow *et al.* 2008 and 2016)
Burkitt lymphoma (BL)	t(8;14) translocation involving *MYC*	>95%
Mantle cell lymphoma (MCL)	t(11;14) translocation involving *CCND1*	65-95% (also seen in myeloma, diffuse large B cell lymphoma (DLBCL))
Follicular lymphoma (FL)	t(14;18) translocation involving *BCL2*	70-95% (also seen in DLBCL)
Mucosa-associated lymphoid tissue (MALT) lymphoma	t(11;18) translocation involving *API2-MLT*	20-50%
Hairy cell leukaemia (HCL)	*BRAF* V600E	>98%
Waldenstroem's macroglobulinaemia (WM)	*MYD88* L265P	>90%
T-Large granular lymphocyte (LGL) leukaemia	*STAT3*	~40%

Table 3: Examples of known and novel molecular markers with clinical relevance in B-LPDs

Disease	**Molecular Marker**	**Clinical relevance**		**Reference**
CLL	*IGHV* mutational status	Prognosis	Somatic hypermutation (SHM) >2% superior prognosis	Rossi D *et al.* (2017)
	IG Stereotype	Prognosis	Subset#2 carries poor prognosis regardless of SHM status	
	TP53 mutations/17p deletions	Predictive + Prognostic	No response to standard FCR immunochemotherapy; require Ibrutinib/Idelalisib regimen (NICE approved)	
WM	*MYD88* and *CXCR4* mutations	Diagnostic + Predictive	MYD88 is a diagnostic marker; patients with an additional mutation in CXCR4 have delayed and reduced responses to Ibrutinib	Yun *et al.* (2017)
DLBCL	*BTK* and *PLCG2* mutations	Predictive	Mutations confer resistance to BTK inhibitors	Castillo *et al.* (2016)
MGUS	*MYD88* L265P mutation	Prognostic	Higher risk of progression to LPD	
FL	7M-FLIPI scoring system	Prognostic	Addition of 7 mutations to the current prognostic scoring system approves patient delineation	Pastore *et al.* 2015

2. Literature review

Search terms, criteria and databases used are listed in table 4. Studies were included when at least SNV plus CNVs were evaluated, preferably in the clinical context rather than explorative studies. Very few papers reported on the applicability and validation of a targeted CapNGS panel in the clinical context and only a few proof-of-principles studies evaluated translocation detection (see table A1 for an overview of most relevant studies).

This review did not include SNV detection as this is an established application for CapNGS and the study was also not intended to provide a technical evaluation of different CapNGS methods.

Table 4: Details of literature search. Searches were limited to the last 10years to reflect the recent developments in NGS technology and its application.

A			
Search terms	**Plus**	**Focus**	**Findings**
Capture NGS	Diagnostic (tool) Clinical applications	Studies reporting application in clinical /routine setting	Few publications (4)
Capture NGS	CNV copy number detection Translocation detection Clonality /clonal VDJ rearrangements/IG/TCR rearrangements/minimal residual disease (MRD)	Proof of principle studies to provide evidence that this approach does work In the context of cancer (not constitutional disease)	SNV>>>translocations (n=12)>>CNV (n= 7)>IG/TR rearrangements (n=2)
Capture NGS	Fluorescence in-situ hybridisation (FISH) karyotype/aCGH /single nucleotide polymorphism (SNP) arrays	To search for studies comparing the novel methods with established assays (used in diagnostic context)	All studies identified above compared their results to a method in routine use
B Databases			
Pubmed https://www.ncbi.nlm.nih.gov/pubmed/, Specific Expert consortia (European research initiative in CLL (ERIC), Euroclonality (EC)), Cancer Research-UK (CR-UK) Stratified medicine programme, 100000 Genomics England (GEL) project Guidelines from Association of clinical genetic science (ACGS), British Society for clinical Haematology (BSCH), National institute for health and care excellence (NICE), National institute for health research (NIHR), NCCN, WHO haematological neoplasms, Conferences: European Haematology Association (EHA) 2015-2016 and American Society of haematology (ASH) 2015-2016 abstract searches			

3. Focus on translocation detection

An early, proof-of-principle study for the detection of translocations using CapNGS was published by Duncavage *et al.* (2012), which demonstrated the possibility of concurrent translocations (*PML-RARA, RUNX1-RUNX1T1*) and mutations detection in AML cell lines. The authors already postulated that targeted CapNGS protocols could provide diagnostic and prognostic information in leukaemias on a single platform with improved efficiency and greater scalability than current methods. Grossmann *et al.* (2011), also working in AML, successfully identified all known (based on prior FISH, karyotyping and RT-PCR analysis) recurrent translocations in 6 patient samples but also identified novel *RUNX1* fusion partners in four cases highlighting this important advantage of CapNGS (see figure 1). A further benefit of NGS is achieving nucleotide resolution: A deletion found by CapNGS in a patient with a known *KMT2A* translocation based on FISH affected the RT-PCR primer site explaining the false-negative result in the RT-PCR MRD assay (Grossman *et al*., 2011). Jang *et al.* (2016) re-investigated 16 NSCLC cases with discordant *ALK* translocations status based on conventional FISH and IHC analysis. Thirteen cases were confirmed to be false-positive IHC results, a mistake that could be detrimental for the patient. CapNGS revealed three cases positive by FISH but IHC negative, did not to carry the standard *EML4-ALK* fusion but other unusual partners. Such details will help us to understand why patients' responses to ALK inhibitors can be varied despite identical FISH results and highlight the advantage CapNGS could bring to disease like *KMT2A*-positive ALLs, where the specific fusion partner affects the prognosis and thus the treatment regime of a patient (Tamai and Inokuchi, 2010).

Pfarr's *et al.* study (2016) is important in that it demonstrates the feasibility of using CapNGS on FFPE specimen, achieving 100% concordance with FISH for the detection of *ALK, ROS* and *RET* rearrangements when validated in clinical cohorts of >150 NSCLC cases. An earlier, smaller study by Abel *et al.* (2014) showed lower sensitivity (93%) for *ALK* translocation detection in cases where the tumour infiltration was below 50% and DNA concentrations were less than 100ng; two important observations now well recognised and incorporated into best practice guidelines (Jennings *et al.,* 2017, Cottrell *et al.,* 2014). Clonality detection is routinely performed on FFPE sections from lymph node biopsies or excision samples and skin biopsies play a major role in the diagnosis of Mycosis fungoides. Especially in the latter, the DNA yield can be minimal and may hamper CapNGS analysis. Abel's et al., (2014) is one of the few studies discussing the technical performance of their CapNGS panel in detail and comparing different bioinformatics approaches, information that is valuable for any diagnostic laboratory wanting to implement a CapNGS method. By bioinformatically downsampling their reads they concluded that a minimum coverage of 600x across breakpoint regions should result in >90% sensitivity in specimen with >20% tumour infiltration and that the overall coverage often cited by other studies can be misleading and attention needs to be given to coverage achieved across breakpoint regions as

their panel required a re-design and increased density of baits to achieve suitable read depths and avoid false-positive calls. In comparison to the advocated >1000x coverage for the detection of somatic SNV (Hageman *et al.,* 2014) this read depth appears low, however the reported coverage across the studies reporting the successful detection of translocation is very variable and often only mean coverage values are cited which may not reflect the read depth across an intronic breakpoint region.

Two studies of particular interest in the context of B- and T lymphoproliferative disorders, focussed on myeloma (Walker *et al.,* 2013) and DLBCL (Bouamar *et al.,* 2013). Using a Sureselect (Agilent) RNA-bait capture-NGS design and 150bp paired-end sequencing (Hiseq, Illumina) Walker *et al*. (2013) were able to detect all IG-translocations in their 65 myeloma samples previously analysed by FISH or RT-PCR. Based on the detailed sequencing information they demonstrated that, although the majority of translocations arose during class-switch recombination (CSR) events, unexpectedly in 12 cases the location of the breakpoints suggested non–CSR mechanisms comprising DH-JH rearrangements, receptor revision and SHM. This demonstrates that despite being a post-germinal centre neoplasm, initiating rearrangements occur at the Pro-B cell stage in myeloma. When comparing the t(11;14) found in myeloma and MCL, MCL events all took place during V(D)J recombination despite MCL being another post GC neoplasm; furthermore the MCL and myeloma V(D)J breakpoints were distinct (Walker *et al.,* 2015). Such detailed observations will help us to understand the pathogenesis of diseases better and could prove relevant in the clinical context: A particular t(14;20) translocation arising from a V(D)J recombination event occurs more frequently in MGUS than myeloma and these patients exhibit stable disease without transformation (Walker *et al.,* 2015).

Bouamar *et al.* (2013) identified 31 fusions in 28 DLBCL biopsies with detailed breakpoint including clinically important double-hit constructs involving *BCL2* and *MYC*. Concordance was achieved in 21 out of the 23 cases cytogenetic data was available for. In one case NGS failed to detect a fusion with an unusual breakpoint due to low read depth (on the edge of the probed region). The other case was one of three where CapNGS identified a novel IG translocations, two of which involved *IRF8*. The rare but recurrent nature of this translocation was subsequently confirmed by Tinguely *et al.* (2013) in a subset of GC-DLBCL. They employed long-range PCR and Sanger sequencing to characterise the fusion, a far more labour-intensive approach.

In our own small pilot study we were able to identify known translocations in 18 out of 21 cases (Wren *et al.,* 2017, see appendix). The three discrepancies could be explained by technical limitations, predominantly low DNA concentrations. A novel fusion was seen in a DLBCL case with a break into 6p25, lying in the vicinity of *IRF4*. Activating *IRF4* translocation are known to occur in GC-derived LPD and this rearrangement could have the same effect but no further studies were performed. Importantly, three T-LPD cases were included with translocations arising from either the TRB, TRG or TRD locus. All were successfully detected by CapNGS without the need for separate assays for each of the loci and all samples were prepared and sequenced on the same run, making translocation detection from any of the IG/TR loci feasible.

Table 5: Summary of advantages and potential problems using CapNGS for translocation detection based on published data

Advantages	**Difficulties**
Able to detect translocations arising from all IG/TR loci if baits cover switch regions and V(D)J elements	Large area to be probed ~200kb
Able to identify novel fusion partners	Breakpoints may lay outside the probed region
Able to analyse all different loci at the same time (*IG, TR, BCL2, MYC* depending on bait design)	
Works on PB/BM-derived DNA, but also on FFPE material	DNA concentrations appear to be a limiting factor; limited evidence for its use on FFPE-derived DNA
Analysis on nucleotide level providing additional information on potential target genes, disease pathogenesis and MRD markers	Detailed breakpoint information currently has limited/no utility in the clinical context

4. Focus on CNV detection

In addition to the routinely used karyotyping and FISH analysis for common CNVs, aCGH and SNP arrays are considered the gold standard for CNV detection but are not routinely available in all diagnostic laboratories despite the clinical relevance of CNVs in haematopoietic cancers (for a review see Song and Shao, 2016). Many proof-of-principle studies have been published showing excellent concordance between conventional methods (aCGH, SNP arrays, FISH/karyotyping) and WGS (Hayes, 2013; Uzilov *et al.,* 2016) but also WES (Vosberg *et al.,* 2016, Wenric *et al.,* 2017) with some promising results also published with targeted CapNGS (Shen, *et al*., 2016 Bolli *et al.,* 2015, Chen *et al.,* 2017). An essential study was performed by Samorodnitsky *et al.* (2015) who compared four different, commercially available custom, capture-targeted DNA sequencing NGS approaches based on their performance in SNV and CNV detection. Library preparation by sonication consistently provided greater uniformity in capture and the two methods utilising enzymatic steps for DNA fragmentation missed SNV due to lack of coverage. CNV detection was however not affected and all methods performed equally well.

Nevertheless, analytical difficulties need to be recognised especially in the tumour setting (Liu *et al.* 2013): WGS provides the easiest interpretation of CNVs as read depths can be compared across the whole genome whereas targeted approaches and even WES suffer from bias due to reduced coverage in GC rich regions and uneven coverage close to bait boundaries (Abel and Duncavage, 2014). Bioinformatic input and adjustment is possible for the analysis of CNV from WES and targeted panels (Zhao *et al.,* 2013) however, a recent review by Zare *et al.* (2017) concluded that the performance of current tools is limited with sensitivities between 50-80% and a high number of false positive calls. Analysis becomes very difficult if no matched germline

sample is run for comparison due to the genomic complexity of tumours caused by abnormal ploidy, varying tumour content and tumour heterogeneity (Wenric *et al.*, 2017, Liu *et al.* 2013). A discussion of different bioinformatics tools is beyond this review but see Liu *et al.* (2013), Zhao *et al.,* (2013) and Zare *et al.* (2017) with regards to challenges in CNV detection.

Two studies with rational approaches for CNV detection in targeted CapNGS panels were published by Bolli *et al.,* (2015) and McKerrell *et al.* (2016): both groups selected highly polymorphic SNPs (minor allele frequencies 0.40-0.45) covering the entire genome (2538 and 8673 SNPs, respectively) from the 1000-Genome data set and compared read depths across these regions. Although CapNGS read-outs were concordant with aCGH (Bolli *et al.,* 2015) and FISH/karyotyping data (McKerrell *et al.*, 2016) the gains and losses were called against assumed normal ploidy (CapNGS cannot establish the ploidy of cells), and thus absolute copy numbers could not be determined. To improve on this, McKerrell *et al.* (2016) selected a group of SNPs for each individual patient that were heterozygous in the germline and derived zygosity values for each sample from this allowing for more accurate CNV calls with sensitivities at least equal to conventional karyotyping: in one sample the study found copy-neutral loss of heterozygosity of 2p and 11q which had not been detected by karyotyping.

As both designs cover autosomal and X chromosomes and are not disease-specific this design could be adopted and standardised for any CapNGS panel with the possibility of adding additional SNPs around the area of particular interest for each study to increase sensitivity (McKerrell *et al.*, 2016). However, further studies comparing this approach to the gold standard are warranted also with regards to bioinformatics tools as both studies used pipelines developed in-house. In NSCLC samples, Chen's *et al.* (2017) study was able to detected *ERBB2* amplification by CapNGS having included 1000 SNPs to allow CNV detection, a much lower number than in the two studies above. Although looking feasible, the optimal design and extent of SNP analysis on CapNGS needs to be further evaluated as an increase in baits will increase the cost of running the panel. On the plus side, including SNP coverage will help in identifying sample mix-ups and low-level contamination as the SNP pattern will be patient specific (Chen *et al.,* 2017).

Table 6: Summary of advantages and potential problems using CapNGS for CNV detection based on published data

Advantages	**Difficulties**
Less sequencing capacity needed in comparison to WGS (cheaper)	WGS results probably superior
Universal standardises panels could work	Bioinformatic approaches specific for targeted NGS needed and further optimisation required
Added information making results patient specific (reduce sample mix and erroneous results due to contamination)	
Can detect CN-LOH/Uniparental disomy (UPD)	
	Limited data published

5. Focus on combining SNV, CNV and translocation detection for clinical practice

Few studies were found that discuss validation of a targeted CapNGS tool for the detection of multiple molecular changes in the diagnostic setting (table A1). The lack of publications could be due to slow adoption of CapNGS as opposed to amplicon-based NGS into routine laboratories, the use of off-the-shelf panels purchased from companies and the move towards WGS in larger centres.

The four studies presented here all provided more detailed data on performance characteristics of their CapNGS assay (average read depth, coverage statistics, areas targeted including overlap of baits and flanking regions) as well as their choice of bioinformatic algorithms than the proof-of-principle studies discussed earlier, which is important due to the current lack of standardisation of laboratory procedures and analysis pipelines (Abel and Duncavage, 2013). Cheng *et al.* (2015) working on solid tumours and He *et al.* (2016) studying haematological malignancies, performed extensive validations to show applicability of the methods in PB/BM as well as FFPE tissue, to define the level of detection for selected examples of SNV and CNVs and to ascertain assay reproducibility in accordance with newly developed guidelines (Jennings *et al.,* 2017, Cottrell *et al.,* 2014). Cheng *et al.* (2015) showed that by increasing the pre-hybridisation PCR cycle number even small samples can be analysed and suitable coverage obtained albeit extra care has to be take to remove PCR duplicates and importantly they only pooled two samples per Miseq run which is unlikely to be cost effective for diagnostic laboratories and a panel of this size (>1Mb) likely requires a more powerful NGS instrument to be available. Based on their evaluation the studies were able to derive sensitivity values resulting in defined thresholds for variant calling to be used in the clinical context, even differentiating between known and novel changes (table 7). In all studies the established thresholds for variant calling are low enough to provide clinically relevant information, despite all groups making fairly conservative calls.

He's *et al.* study (2016) provides the largest set of data in the clinical context analysing 3696 samples from a range of myeloid and lymphoid malignancies achieving a CapNGS performance success rate of 93%. Failures occurred due to suboptimal RNA preparation and coverage below the cut-off of 250x and this failure rate is in-line with current RNA-based tests and FISH technology. Their study is unique in that they perform two separate hybridisations, one on DNA and one on RNA on each sample, which increases the cost. The benefit of RNA is obvious for fusion genes with large introns (e.g. *ABL1*) which would require a huge number of baits in an DNA approach, however DNA is required for the identification of fusion constructs that do not result in a fusion transcript e.g. most translocations in lymphoma arising from IG/TCR loci. Sensitivity established by dilution experiments with cell lines is excellent and interestingly DNA-based sequencing was more sensitive than RNA-based approaches for *KMT2A*- and *PDGFRA* translocation detection but no possible explanation for this is provided. Looking at the clinical

utility, they identified at least one driver mutation in 95% of cases of which 77% identified a suitable targeted therapy (including clinical trials), most of them in their myeloid cohort. 66% of cases carried a variant with known prognostic relevance of which 19% were *TP53* mutations in NHLs. The study identified novel IGH rearrangements in 17% of cases- for most though the effect is unknown making it difficult to utilise this information for more than evidence of a clonal cell population at present.

Table 7: Performance characteristics of the CapNGS approaches in the 5 different studies and their reporting thresholds.

Study	**Coverage statistics**	**Reporting thresholds**	**Reproducibility assessed?**	**Control**
Cheng *et al.* (2015)		Minimum depth of 20x Tier 1: 2% VAF with at least 8 mutant reads Tier 2: 5% VAF with at least 10 mutant reads	Yes SD 0.01-0.025	Matched normal
He *et al.* (2016)	Median coverage (DNA) 500x with minimum cut-off 250x for novel and 150x for hotspots variants	SNV: MAF 1% (known), 5% (novel) Indels: MAF 3% (known), 10% novel Translocations: 100% if tumour 20-100%, 98% if tumour 10%	Yes 97% concordance, including 100% in long-term assessment	None
Pritchard *et al.* (2014)	Mean coverage 150x with minimum 50x for positive calls	10%	NA	No: variants filtered against internal database, exome variant server and 1000 Genome data
McKerrell *et al.* (2016)	>30x for 94% of target exons and 98% of SNPs, >70x for 75% overall	5%, down to 1% for known hotspots	NA	Universal, normal cord blood sample and in-house MIDAS software

Pritchard *et al.* (2014) and McKerrell *et al.* (2016) describe how they employed specific bioinformatics tools for the detection of the different types of genetic variants consecutively to improve detection rates, ease and reliability of analysis. In contrast to most other studies, neither ran a matched germline sample alongside the tumour sample for subtracting constitutional variants. Pritchard *et al.* (2014) relied on extensive reviews for each change against several recognised databases whereas McKerrell *et al.* (2016) included a cord blood samples as a universal control and devised a dedicated software tool trained on data obtained from normal

controls to select relevant variants from the tumour (AML) sample. Without expert bioinformatics understanding it is difficult to evaluate these approaches and the current recommendation, certainly for WGS, is to run a matched germline control (Hamblin, 2017). However, myeloid disease cause a particular problem in this regard: Using remission samples for comparison in a cohort of 24 AMLs, McKerrell *et al.* (2016) found that clinically relevant mutations had been filtered from the analysis as they were present in the remission sample was well as the tumour sample in 5 cases. This persistence of founder mutations has not bee described in lymphoid neoplasms although it seems prudent to utilise DNA derived from spit samples or CD15+ selected cells as comparators (Hamblin, 2017).

Unfortunately, none of the studies provide a cost analysis of running their panels in the routine diagnostic setting. 'Karyogene' (McKerrell *et al.*, 2016) designed as a 'one-stop' diagnostic tool for AML and MDS covers 49 genes, 9111 SNPS and intronic regions for the four most relevant translocations resulting in a capture area of 2.3Mbp, which will be reflected in the price of the baits and the required sequencing power (number of samples that can be pooled per run/size of run cartridge/type of NGS instrument), making this expensive to run. Their results for 62 AML and 50 MDS cases convincingly show though that the panel is capable of detecting all clinically relevant changes it is designed for and that the performance matches current technology with additional translocations and CN-LOH being detected. In-line with the findings on CNV detection earlier, CapNGS only achieves 93% sensitivity compared to karyotyping and the latter remains valuable in providing a broad overview of aberrations as for example inv3 were not included in the 'Karyogene' design and therefore not detected. For targeted panels the trade off is between size and therefore price versus complete coverage of all clinically relevant loci, with the latter continuously expanding as new data becomes available requiring re-design of panels.

Table 8: Summary of advantages and potential problems using CapNGS in the clinical context

Advantages	**Difficulties**
Able to analyse all different variants in one set-up on one DNA sample	Limitations due to panel design (not all areas covered) and DNA quantity/quality
Sensitivity, Specificity and failure rates are acceptable	CNV analysis has limitations
Pan-cancer design is possible	WGS could be easier/cheaper/more established
New software tools being designed for specific CapNGS applications improving analysis	Complicated bioinformatic pipelines required-expert input mandatory
Can combine DNA and RNA hybridisation in parallel workflows	Evaluating the need for a control sample and defining the best option for different disease categories
Pan-cancer CapNGS designs are being commercialised by companies providing technical and bioinformatics support	Although more data is being published and recommendations are being drawn up by experts, optimal coverage is panel dependant and performance needs to be carefully evaluated by each user

6. Focus on IG/TCR rearrangement detection

Analysis of IG/TCR loci has numerous applications (see table 9) and is commonly performed according to Euroclonality (EC) protocols and guidelines (van Dongen *et al.,* 2003; Langerak *et al.,* 2012). Currently multiplex PCR reactions are used and the same approach has been adapted in the amplicon-based NGS context (Invivoscribe https://www.invivoscribe.com; Adaptivebiotech http://www.adaptivebiotech.com), although no peer-reviewed direct comparisons between the EC method and commercial NGS methods could be found. The main focus of NGS applications has been on marker detection and MRD measurements in ALL due to the high sensitivity achievable by NGS and this has been compared against current RQ-PCR protocols by a number of studies and validated for routine service (Ladetto *et al.,* 2014, Shin *et al.,* 2017, Faham *et al.,* 2012).

The lack of publications around clonality testing by NGS is unsurprising as the field is dominated by the EC expert consortium and the commercial companies (one of which holding the patent on PCR-based clonality testing) and the fact that the IG/TCR loci represent unique stretches of DNA undergoing modification at the germline level together with a high degree of homology between genes in each loci making their analysis particularly challenging and not amenable to analysis with standard NGS primer design and common bioinformatics pipelines (Langerak *et al.,* 2017).

A proof-of principle analysis by the EuroClonality-NGS consortium (Wren *et al.,* 2017) confirmed that capture NGS approaches are capable of identifying IGH V-J and incomplete D-J, IGK, IGL, TRB V-J and incomplete TRB D-J and TRG clonal rearrangements showing complete concordance with current EC protocol and Southern blot results in all 14 samples. At the time, due to the lack of specialised software, analysis was performed manually on Integrative genomics viewer (Broad Institute), however, since then multiple bioinformatics tools have been developed (Bystry *et al*., 2017, Duez *et al.,* 2016, Alamyar *et al.,* 2012) with the EC group currently validating two of these (Gonzalez D., personal communication). As discussed by Langerak *et al.* (2017), CapNGS including baits for the V, D and J genes of each IG/TCR locus will enable the identification of clonal rearrangements for clonality assessment and MRD marker assignment (including unproductive rearrangements) for any lymphoid disease whilst also providing details on the number of clonal B/T cells within the whole B/T cell pool. The authors scrutinise how the greater resolution delivered by NGS-based clonality assessment will require the reappraisal of the meaning of clonality in the context of healthy individuals and disease entities whilst also making standardisation of bioinformatics pipelines, result interpretation and visualisation mandatory. Challenges associated with clonality analysis by CapNGS are summarised in table 10.

Montalvo *et al.* (2016) used a CapNGS approach to identify IG/TCR rearrangements for marker identification in ALL patients to use the information for the design of MRD monitoring assays based on unique patient specific IG/TCR rearrangements. Although only a small study with 10

patients, CapNGS identified all 50 clonal rearrangements previously reported but was also capable of finding additional rearrangements providing hope that this method could be used to help MRD monitoring in the ~15% of cases for whom no marker can be found by traditional methods. Identification of minor clones may also carry prognostic information and if monitored longitudinally could provide alerts for disease progression. They used Illumina V3 chemistry performing 300bp paired-end reads, much longer than the 75bp-120bp paired-end reads run by Wren *et al.* (2017). Neither the papers presenting the translocation detection studies (see above) nor studies working on IG/TCR rearrangement detection provide any reason for using particular read-lengths; however comparisons would be useful to establish whether all read lengths are equally suited to rearrangement detection.

Table 9: Role of IG/TCR rearrangement analysis and potential advantages of CapNGS

Application	Current methods	Advantages of CapNGS
Clonality testing	• Identification of a clonal cell population required in ~15% of NHL diagnoses • Detection of a clonal relationship between different lesions	• Analysis of all IG/TR, run in a single set-up (capture) • More detailed analysis of clonal relationship possible based on sequence information • Investigation of intraclonal diversity and dynamics • Higher sensitivity (amplicon >capture) • Less affected by somatic hypermutations (capture NGS)
MRD monitoring	• Clonal V(D)J rearrangements in IG/TR loci serve as markers for sensitive MRD detection on a patient specific basis	• Greater sensitivity (amplicon) • Identification and monitoring of minor clones possible • All IG/TR loci are interrogated increasing the availability of MRD markers
Repertoire/Stereo-typing analysis*	• Characterisation of specific IG/TCR rearrangements associated with disease pathogenesis and prognosis	• More detailed analysis of clonotypes/repertoire skewing • Analysis of clonal diversification and clonal relationship

Table 10: Challenges identified

Challenges for CapNGS applications for clonality detection (see Langerak *et al.,* 2017)
Bioinformatic pipelines- validation and standardization especially for CapNGS
Limit of detection needs to be defined
Validation in the context of varying tumor load, different DNA quality (including FFPE)
Error correction to enable correct assignment of clonotypes, stereotypes and IGHV SHM load; differentiation between technical (PCR/Sequencing) artifact and clonal evolution
Definition and interpretation of clonality needs to be re-evaluated in the context of minor clones and clonotype information
Standardisation of workflow including sample preparation, bioinformatics pipeline, visualization and interpretation

7. Conclusion

The aim was to review studies using CapNGS methods to find evidence for the feasibility of applying CapNGS as a tool for the concurrent analysis of SNVs, CNVs, translocations and clonal IG/TR rearrangements. Proof-of-principle data is published for all applications with larger bodies of evidence available for SNV and translocation detection both in solid and haematological neoplasms. For CNV detection the data is not consistent and WGS approaches seem to be favoured. However, the result from McKerell's *et al.* (2017) design incorporating genome-wide SNPs appear sound and if reproducible would provide a transferable approach. Data on clonality detection is very limited for the reasons discussed above and any advances will rely heavily on the work of the EC-NGS group.

The search was hampered by ambiguity in the terminology with CapNGS being referred to as capture-enrichment, pull-down and hybridisation NGS, which may have resulted in studies being missed. The search also tried in particular to identify studies interrogating multiple different genomic aberrations in haematological cancers, which limited the number of studies identified and may have caused useful studies employing CapNGS in smaller applications to be overlooked.

Studies comparing the different commercially available CapNGS protocols are scarce, however, the studies presented here mainly used the two most prominent CapNGS methods, namely Agilent SureSelect (Agilent Technologies, Santa Clara, CA) and Roche NImbleGen SeqCap EZ technology (Roche, Madison, WI) which have been proven to yield very similar read depths (Hagemann *et al.,* 2014). Still, more thorough investigations, like the one performed by

Samorodnitsky *et al.* (2015) comparing the performance of different CapNGS on CNV detection are warranted to understand the advantages and disadvantages of each method for each type of variant detection (table 11). NGS has been a disruptive technology and this is evident by the lack of standardisation in sample processing protocols, laboratory and sequencing methods and bioinformatics pipelines. Nevertheless, the results of the studies summarised here support the applicability of CapNGS methods for the detection of the different variants as well as IG/TCR rearrangements and numerous advantages over current methodologies were identified. A design encompassing all clinically relevant genes, translocation breakpoints whilst also incorporating IG/TCR loci and SNPs for CNV assessment could provide a very efficient and clinically valuable tool in the context of B- and T lymphoproliferative diseases.

Table 11: Information lacking from studies/limitations

Information not included /not published
Limited data on sensitivity and specificity and how values were derived
Limited information on coverage other than mean coverage and the performance of panels for different variants
Lack of information on cost
Lack of information on performance in a routine laboratory (workflow, batching, turn-around time)
As most studies were proof-of-principle, no validation in retrospective and prospective cohorts
No data on direct comparison between targeted CapNGS and WGS for translocation detection
Limited data on RNA versus DNA baits for translocation detection
Comparison of data difficult as bioinformatics analyses performed on in-house tools/high level of bioinformatics expertise required to understand their limitations
Several different CapNGS protocols are in use and only one studies reports on a direct comparison between them in the context of SNV and CNV detection (Samorodnitsky *et al.,* 2015)
Affect of percentage tumour infiltration/blast percentage is not assessed
Read lengths for paired-end reads on Illumina chemistry differed between 75bp and 300bp for the paired-end reads but no comparison has been performed to establish the best read length for translocation and IG/TR rearrangement detection.
Focus on B-cell malignancies with no relevant studies identified for T-cell LPDs

Table A1: Overview of the most relevant proof-of principle studies and studies validating CapNGS in clinic are highlighted in bold

Authors (et al.)	Journal	Year	Context	Summary	Title	Comment
Cheng DT	**J Mol Diagn**	**2015**	**Solid tumours**	**Panel included 341 cancer-related genes plus 33 intronic regions known to habour recurrent rearrangements. Good info on clinical validation and sensitivity down to 2%**	**Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology**	**Used in clinic**
Frampton GM	Nat Biotechnol	2013	Solid tumours	Panel included 287 cancer-related genes plus intronic regions known to habour recurrent rearrangements. At least 1 CRGA was detected in 76% of cases. Foundation One panel DNA only	Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing.	Proof-of principle study
Luthra R	J Mol Diagn	2017	Solid tumours	OncoMine Comprehensive Assay 143 cancer-related genes	A Targeted High-Throughput Next-Generation Sequencing Panel for Clinical Screening of Mutations, Gene Amplifications, and Fusions in Solid Tumors.	
Drilon A	Clin Cancer Res	2015	Lung cancer	"Foundation One" Panel included287 cancer-related genes plus 47 intronic regions known to habour recurrent rearrangements. All patients had tested negative for standatd 11 gene tests by NGS. In 26% of patients a directly actionable mutation was detected by this panel and in a further 39% results indicated applicability for clinical trial.	Broad, Hybrid Capture-Based Next-Generation Sequencing Identifies Actionable Genomic Alterations in Lung Adenocarcinomas Otherwise Negative for Such Alterations by Other Genomic Testing Approaches.	Proof-of principle study
Authors (et al.)	**Journal**	**Year**	**Context**	**Summary**	**Title**	**Comment**
Hirshfield KM	Oncologist	2016	Solid tumours	"FoundationOne" Clinical Laboratory Improvement Amendments (CLIA)-certified panel; Good discussion of clinical relevance and the need for tumour MDT meetings; Implementable clinical action in 35% of patients (diagnostic modification, targeted therapy, need for germline testing)	Clinical Actionability of Comprehensive Genomic Profiling for Management of Rare or Refractory Cancers.	**Used in clinic**
Pritchard CC	**J Mol Diagn**	**2014**	**Pan-Cancer**	**UW-OncoPlex 194 cancer-related genes plus intronic regions known to habour recurrent rearrangements (solid and Haematological cancers). Good guidance on how to establish tier system for variants and how to interpret findings in the clinical context.Sensitivity 99.2% compared to current standard methods for all alterations**	**Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens.**	**Used in clinic**
Pfarr N	**Genes Chromosomes Cancer**	**2016**	**NSCLC**	**Targeted for ALK, ROS, RET fusions as well as standard somatic mutations locations; 100% sensitivity compared to current assays including FISH; they conclude that this can replace conventional methods**	**High-throughput diagnostic profiling of clinically actionable gene fusions in lung cancer.**	**Used in clinic**
He J	**Blood**	**2016**	**Haematological malignancies**	**DNA/RNA profiliing for all with large >3000 clinical cohort of patient samples**	**Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting**	**Clinical Validation**
Walker B	**Blood**	**2013**	**Myeloma**	**Capture design for the detection of translocations arising from IG/TR loci**	**Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells**	
Abel HJ	**J Mol Diagn**	**2014**	**NSCLC, ALL**	**Capture design for the detection of ALK and KMT2A rearrangements**	**Using NGS capture for detection of rearrangements in ALK (lung cases FFPE) and KMT2A (ALL, liquid) from DNA**	Proof-of principle study
Authors (et al.)	**Journal**	**Year**	**Context**	**Summary**	**Title**	**Comment**
Bouamar H	**Blood**	**2013**	**DLBCL**	**Capture design for the detection of translocations arising from IGH locus; identification of new partners**	**A capture-sequencing strategy identifies *IRF8*, *EBF1*, and *APRIL* as novel *IGH* fusion partners in B-cell lymphoma**	Proof-of principle study
Bolli N	**Blood Cancer J**	**2016**	**Myeloma**	**246 genes, 2538 SNPs for CNV and entire IGH locus captured; performed on cell lines with the view to identify novel biomarkers**	**A DNA target-enrichment approach to detect mutations, copy number changes and immunoglobulin translocations in multiple myeloma**	Proof-of principle study
Bolli N	**Haematologica**	**2015**	**AML**	**Mutations and CNVs; Haloplex approach**	**Characterization of gene mutations and copy number changes in acute myeloid leukemia using a rapid target enrichment protocol**	Proof-of principle study
Conte N	**Leukemia**	**2013**	**AML**	**CN-AMLanalysed for SNVs and CNVs**	**A DNA target-enrichment approach to detect mutations, copy number changes and immunoglobulin translocations in multiple myeloma**	Proof-of principle study
McKerrell T	**Blood**	**2016**	**AML/MDS**	**"Karyogene" ,SNV, CNV and translocation and all indels etc; 93% concordance with karyotyping, inv3 not included in design**	**Development and validation of a comprehensive genomic diagnostic tool for myeloid malignancies**	**Aim to be used in clinic**
Wren D	**Haematologica**	**2017**	**LPDs**	**Capture design for the detection of IG/TCR translocation and IG/TCR rearrangements**	**Comprehensive translocation and clonality detection in lymphoproliferative disorders by next-generation sequencing.**	Proof-of principle study
Duncavage EJ	**Modern Pathology**	**2012**	**Cell lines**	**Proof of principle study demonstrating capture NGS's ability to detect translocations and mutations; focus on t(9;11), t(9;22) and t(15;17) only**	**Targeted next generation sequencing of clinically significant gene mutations and translocations in leukemia**	Proof-of principle study
Grossman V	**Leukemia**	**2011**	**AML**	**Proof of principle that capture NGS can detect SNV, indels and translocations in AML. Description of the detection of a cryptic MLL t(11;19) rearrangement indetified by karyotype/FISH but not evaluable by RT-PCR. They also described 4 novel RUNX1 fusions in their series of AML cases.**	**Targeted next-generation sequencing detects point mutations, insertions, deletions and balanced chromosomal rearrangements as well as identifies novel leukemia-specific fusion genes in a single procedure**	Proof-of principle study
Jang J	J Thoracic Oncology	2016	NSCLC	FISH+/IHC- cases were shown to carry an ALK translocation other than the EML4-ALK by NGS (3 cases). FISH-/IHC+ cases were confirmed to be false-positive by IHC (13 cases)	Custom Gene Capture and Next-Generation Sequencing to Resolve Discordant ALK Status by FISH and IHC in Lung Adenocarcinoma	

List of targets for the current version of the CapNGS panel

SNV	Translocation	SNPs for CNV detection
ABL1_1	CCND1	9p_TEL_rs7019901
ARID1A_1	BCL2	CDKN2A_rs2811710
ATM_2	KMT2Atx	CDKN2B_rs2518723
BCL2_1	TCF3tx	9p_CEN_rs10973772
BIRC3_2	CRLF2tx	17p_TEL_rs7222601
BRAF_1	STILtx	TP53_rs1794289
BTG1_1	BIRC3tx	17p_CEN_rs1037037
BTG1_1	ALKtx	11q_CEN_rs10792269
BTG1_2	PDGFRBtx	11q_TEL_rs2512702
BTK_2	BCL6tx	CCND1_Copy_rs1352075
CARD11_2	TP63	ATM_Copy_rs228598
CCND1_1	DUSP22	8q_TEL_rs2292779ß
CCND3_1		MYC_rs113337520
CD79A_1		IKZF1_rs7782210
CD79B_1	**SNV (continued)**	7p_TEL_rs10227104
CDKN2A_1	MLL2_1	PAX5_rs6476588
CDKN2B_1	MYC_1	13q_CEN_rs36116586
CREBBP_1	MYD88_4	RB1_rs7994141
CXCR4_1	NFKBIE_1	13q_TEL_rs9514776
DNMT3A_2	NFKBIE_6	BTG1_rs28399539
EGR2_2	NOTCH1_1	12q_TEL_rs1725791
EP300_1	NOTCH2_34	ETV6_rs2855750
ERG_3	NRAS_5	12pTEL_rs10848531
EZH2_2	NT5C2_2	12pCEN_rs7138369
FAT1_2	PAX5_1	PIK3CA_rs2677760
FBXW7_2	PHF6	

FLT3_1	PIK3CA_2
HIST1H1B_1	POT1_5
HIST1H1C_1	PTEN_1
HIST1H1D_1	RHOA_2
HIST1H1E_1	RUNX1_1
ID3_2	SAMHD1_1
IDH1_3	SF3B1_1
IDH2_1	SOCS1_2
IKZF1_2	STAT3_2
IL7R_1	STAT5B_2
JAK1_25	TCF3_2
JAK2_3	TET2_3
JAK3_2	MAP3K14_1
KIT_1	MAPK1_1
KLF2_1	TNFAIP3_2
KRAS_2	TP53_2
MAP2K1_1	TRAF2_1
XPO1_2	TRAF3_2
WT1_1a	

Abbreviations

ABC	Activated B cell
ACGS	Association of clinical genetic science
AITL	Angioimmunoblastic lymphoma
ALCL	Anaplastic large cell lymphoma
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
ASH	American Society of haematology
BCSH	British Society for clinical Haematology
BL	Burkitt lymphoma
CLL	Chronic lymphocytic leukaemia
CN-LOH	copy neutral- loss of heterozygosity
CNV	Copy number variant
CSR	class-switch recombination
DLBCL	Diffuse large B cell lymphoma
EC	Euroclonality consortium
EHA	European Haematology Association
ERIC	European research initiative in CLL
FFPE	Formalin fixed paraffin embedded
FISH	Fluorescence in-situ hybridisation
FL	Follicular lymphoma
GC	Germinal Centre
GEL	Genomics England
HCL	Hairy cell leukaemia
IG	Immunoglobulin
IHC	Immunohistochemistry
LGL	Large granular lymphocyte
LPD	Lymphoproliferative disease
MALT	Mucosa-associated lymphoid tissue lymphoma
MBL	Monoclonal B lymphocytosis

MCL	Mantle cell lymphoma
MDS	Myelodysplasia
MGUS	Monoclonal gammopathy of unknown significance
MRD	Minimal residual disease
NGS	Next-generation sequencing
NHL	Non-Hodgkin Lymphoma
NICE	National Institute for Health and Care Excellence
NIHR	National institute for health research
NSCLC	Non small cell lung cancer
PCR	Polymerase chain reaction
SMZL	Splenic marginal zone lymphoma
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SOP	Standard operating procedure
TCR	T-cell receptor
UKAS	United Kingdom accreditation service
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World health organisation
WM	Waldenstroem's macroglobulinaemia

References

Abel H.J. and Duncavage E.J. (2013) 'Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches.' *Cancer Genetics*, 206(12) pp. 432-40.

Abel H.J., Al-Kateb H., Cottrell C.E., Bredemeyer A.J., Pritchard C.C., Grossmann A.H., Wallander M.L., Pfeifer J.D., Lockwood C.M. and Duncavage E.J. (2014) 'Detection of gene rearrangements in targeted clinical next-generation sequencing.' *The Journal of Molecular Diagnostics* 16(4) pp. 405-417

Alamyar E., Guidicelli V., Li S., Duroux P and Lefranc M-P. (2012) 'IMGT/HighV-QUEST: The IMGT® web portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high throughput and deep sequencing. '*Immunome Research*. 8:1:2 pp. 1-16

Bogusz A.M. and Bagg A. (2016) 'Genetic aberrations in small B-cell lymphomas and leukemias: molecular pathology, clinical relevance and therapeutic targets.' *Leukemia Lymphoma* 57(9) pp. 1991-2013

Bolli N., Li Y., Sathiaseelan V., Raine K., Jones D., Ganly P., Cocito F., Bignell G., Chapman M.A., Sperling A.S., Anderson K.C., Avet-Loiseau H., Minvielle S., Campbell P.J. and Munshi N.C. (2016) A DNA target-enrichment approach to detect mutations, copy number changes and immunoglobulin translocations in multiple myeloma. *Journal of Blood Cancer* J. 2016 6(9) pp. e467-475.

Bolli N., Manes N., McKerrell T., Chi, J., Park N., Gundem G., Quail M.A., Sathiaseelan V., Herman B., Crawley C., Craig J.I.O., Conte N., Grove C., Papaemmanuil E., Campbell P.J., Varela I., P., and Vassiliou G.S. (2015) 'Characterization of gene mutations and copy number changes in acute myeloid leukemia using a rapid target enrichment protocol'

Haematologica, 100(2) pp. 214–222.

Bouamar H., Abbas S., Lin A.P., Wang L., Jiang D., Holder K.N., Kinney M.C., Hunicke-Smith S. and Aguiar R.C. (2013) 'A capture-sequencing strategy identifies IRF8, EBF1, and APRIL as novel IGH fusion partners in B-cell lymphoma.' *Blood*, 122(5) pp. 726-733.

Bystry V., Reigl T., Krejci A., Demko M., Hanakova B., Grioni A., Knecht H., Schlitt M., Dreger P., Sellner L., Herrmann D., Pingeon M., Boudjoghra M., Rijntjes J., Pott C., Langerak A.W., Groenen J., Davi F., Brüggemann M. and Darzentas N. (2017) 'ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data.' *Bioinformatics,* 33(3) pp. 435-437

Castillo J.J., Treon S.P. and Davids M.S. (2016) Inhibition of the Bruton Tyrosine Kinase Pathway in B-Cell Lymphoproliferative Disorders. *Cancer Journal*, 22(1) pp. 34-39.

Chen Y., Zhao L., Wang Y., Cao M., Gelowani V., Xu M., Agrwal S.A., Li Y., Diager S.P., Gibbs R., Wang F. and Chen R. (2017) '*SeqCNV:* a novel method for identification of copy number variations in targeted next-generation sequencing data. *BMC Bioinformatics*. 18 pp. 147-156

Cheng D.T., Prasad M., Chekaluk Y., Benayed R., Sadowska J., Zehir A., Syed A., Wang Y.E., Somar J., Li Y., Yelskaya Z., Wong D., Robson M.E., Offit K., Berger M.F., Nafa K., Ladanyi M. and Zhang L. (2017) 'Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing' *BMC Medical Genomics*, 10 pp. 33-39.

Cheng D.T., Mitchell T.N., Zehir A., Shah R.H., Benayed R., Syed A., Chandramohan R., Liu Z.Y., Won H.H., Scott S.N., Brannon A.R., O'Reilly C., Sadowska J., Casanova J., Yannes A., Hechtman J.F., Yao J., Song W., Ross D.S., Oultache A., Dogan S., Borsu L., Hameed M., Nafa K., Arcila M.E., Ladanyi M. and Berger M.F. (2015) 'Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology.' *The Journal of Molecular Diagnostics*,17(3) pp. 251-264

Cottrell C.E., Al-Kateb .H, Bredemeyer A.J., Duncavage E.J., Spencer D.H., Abel H.J., Lockwood C.M., Hagemann I.S., O'Guin S.M., Burcea L.C., Sawyer C.S., Oschwald D.M., Stratman J.L., Sher D.A., Johnson M.R., Brown J.T., Cliften P.F., George B., McIntosh L.D., Shrivastava S., Nguyen T.T., Payton J.E., Watson M.A., Crosby S.D., Head R.D., Mitra R.D., Nagarajan R., Kulkarni S., Seibert K., Virgin H.W. 4th, Milbrandt J. and Pfeifer J.D. (2014) 'Validation of a next-generation sequencing assay for clinical molecular oncology.' *The Journal of Molecular Diagnostics,* 16(1) pp. 89-105

Drilon A., Wang L., Arcila M.E., Balasubramanian S., Greenbowe J.R., Ross J.S., Stephens P., Lipson D., Miller V.A., Kris M.G., Ladanyi M. and Rizvi N.A. (2015) 'Broad, Hybrid Capture-Based Next-Generation Sequencing Identifies Actionable Genomic Alterations in Lung Adenocarcinomas Otherwise Negative for Such Alterations by Other Genomic Testing Approaches.' *Clinical Cancer Research*, 21(16) pp. 3631-9

Du M.Q. (2017) 'MALT lymphoma: Genetic abnormalities, immunological stimulation and molecular mechanism.' *Best Practice Research in Clinical Haematology* 30(1-2) pp. 13-23

Duncavage E.J., Abel H.J., Szankasi P., Kelley T.W. and Pfeifer J.D. (2012) 'Targeted next generation sequencing of clinically significant gene mutations and translocations in leukemia.

Modern Pathology, 25(6) pp. 795-804

Duez M., Giraud M., Herbert R., Rocher T., Salson M. and Thonier F. (2016) 'Vidjil: A Web Platform for Analysis of High-Throughput Repertoire Sequencing.' *PLoS One,* 11(11) pp. 1-12

Faham, M., J. Zheng, M. Moorhead, V. E. Carlton, P. Stow, E. Coustan- Smith, C. H. Pui, and D. Campana. (2012). 'Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia.' *Blood* 120 pp. 5173–5180

Frampton G.M., Fichtenholtz A., Otto G.A., Wang K., Downing S.R., He J., Schnall-Levin M., White J., Sanford E.M., An P., Sun J., Juhn F., Brennan K., Iwanik K., Maillet A., Buell J., White E., Zhao M., Balasubramanian S., Terzic S., Richards T., Banning V., Garcia L., Mahoney K., Zwirko Z., Donahue A., Beltran H., Mosquera J.M., Rubin M.A., Dogan S., Hedvat C.V., Berger M.F., Pusztai L., Lechner M., Boshoff C., Jarosz M., Vietz C., Parker A., Miller V.A., Ross J.S., Curran J., Cronin M.T., Stephens P.J., Lipson D. and Yelensky R. (2013) 'Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing.' *Nature Biotechnology,* 31(11) pp.1023-1031

Gray P.N., Dunlop C.L.M. and Elliott A.M. (2015) 'Not All Next Generation Sequencing Diagnostics are Created Equal: Understanding the Nuances of Solid Tumor Assay Design for Somatic Mutation Detection '*Cancers, 7* pp. 1313-1332;

Grossmann V., Kohlmann A., Klein H.U., Schindela S., Schnittger S., Dicker F., Dugas M., Kern W., Haferlach T. and Haferlach C. (2011) 'Targeted next-generation sequencing detects point mutations, insertions, deletions and balanced chromosomal rearrangements as well as identifies novel leukemia-specific fusion genes in a single procedure.' *Leukemia,* 25(4) pp. 671-680

Hagemann I.A., Cottrell C.E., and Lockwood C.M. (2014) 'Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy' *Cancer Genetics,* 206 pp. 420-431

Hagemann I.A., Cottrell C.E., and Lockwood C.M. (2014) 'Design of targeted, capture-based, next generation sequencing tests for precision cancer therapy' *Cancer Genetics,* 206 pp. 420-431

Hamblin A. (2017) *Haem-onc Programme: Eligibility criteria, sample requirements & sample handling by type.* Presentation at the GEL Haematological Cancer Workshop 21st June

Hamblin A., Wordsworth S., Fermont J.M., Page S., Kaur K., Camps C., Kaisaki P., Gupta A., Talbot D., Middleton M., Henderson S., Cutts A., Vavoulis D.V., Housby N., Tomlinson I., Taylor J.C. and Schuh A. (2017) 'Clinical applicability and cost of a 46-gene panel for genomic analysis of solid tumours: Retrospective validation and prospective audit in the UK National Health Service.' *PLoS Medicine,* 14(2) pp1-26

Hayes J.L. (2013) 'Diagnosis of copy number variation by Illumina next generation sequencing is comparable in performance to oligonucleotide array comparative genomic hybridisation.' *Genomics,* 102(3) pp. 174-181

He J., Abdel-Wahab O., Nahas M.K., Wang K., Rampal R.K., Intlekofer A.M., Patel J., Krivstov A., Frampton G.M., Young L.E., Zhong S., Bailey M., White J.R., Roels S., Deffenbaugh J., Fichtenholtz A., Brennan T., Rosenzweig M., Pelak K., Knapp K.M., Brennan K.W., Donahue A.L., Young G., Garcia L., Beckstrom S.T., Zhao M., White E., Banning V., Buell J., Iwanik K., Ross J.S., Morosini D., Younes A., Hanash A.M., Paietta E., Roberts K., Mullighan C., Dogan A., Armstrong S.A., Mughal T., Vergilio J.A., Labrecque E., Erlich R., Vietz C., Yelensky R., Stephens P.J., Miller V.A., van den Brink M.R., Otto G.A., Lipson D. and Levine R.L. (2016) 'Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting.' *Blood*, 127(24) pp. 3004-3014.

Hoelzer D., Bassan R., Dombret H, Fielding A., Ribera J.M. and Buske C. (2016) 'Acute Lymphoblastic Leukaemia: ESMO Clinical Practice Guidelines' *Annals of Oncology,* 27(suppl 5) pp. v69-v82

Hovelson D.H., McDaniel A.S., Cani A.K., Johnson B., Rhodes K., Williams P.D., Bandla S., Bien G., Choppa P., Hyland F., Gottimukkala R., Liu G., Manivannan M., Schageman J., Ballesteros-Villagrana E., Grasso C.S., Quist M.J., Yadati V., Amin A., Siddiqui J., Betz B.L., Knudsen K.E., Cooney K.A., Feng F.Y., Roh M.H., Nelson P.S., Liu C.J., Beer D.G., Wyngaard P., Chinnaiyan A.M., Sadis S., Rhodes D.R. and Tomlins S.A. (2015) Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. *Neoplasia.* 17(4) pp. 385-99.

Jang J.S., Wang X., Vedell P.T., Wen J., Zhang J., Ellison D.W., Evans J.M., Johnson S.H., Yang P., Sukov W.R., Oliveira A.M., Vasmatzis G., Sun Z., Jen J. and Yi E.S. (2016) 'Custom Gene Capture and Next-Generation Sequencing to Resolve Discordant ALK Status by FISH and IHC in Lung Adenocarcinoma.' *Journal of Thoracic Oncology*, 11(11) pp.1891-1900

Jennings L.J., Arcila M.E., Corless C., Kamel-Reid S., Lubin I.M., Pfeifer J., Temple-Smolkin R.L., Voelkerding K.V. and Nikiforova M.N. (2017) 'Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists.' *The Journal of Molecular Diagnostics* 19(3) pp. 341-365

Karmali R. and Gordon L.I. (2017) 'Molecular Subtyping in Diffuse Large B Cell Lymphoma: Closer to an Approach of Precision Therapy.' *Current Treatment Options Oncology* 18(2) pp.11-16

Kiel M.J., Velusamy T., Betz B.L., Zhao L., Weigelin H.G., Chiang M.Y., Huebner-Chan D.R., Bailey N.G., Yang D.T., Bhagat G., Miranda R.N., Bahler D.W., Medeiros L.J., Lim M.S. and Elenitoba-Johnson K.S. (2012) 'Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma.' *Journal of Experimental Medicine* 209(9) pp. 1553-1565

Kotelnikova E.A., Pyatnitskiy M., Paleeva, A., Kremenetskaya O. and Vinogradov D. (2016) 'Practical aspects of NGS-based pathways analysis for personalized cancer science and medicine' *Oncotarget*, 7(32) pp. 52493–52516

Ladetto, M., Bruggemann M., Monitillo L., Ferrero S., Pepin F., Drandi D., Barbero D., Palumbo A., Passera R., Boccadoro M., Ritgen M., Gokbuget N., Zheng J., Carlton V., Traumtann H., Faham M. and Pott C. (2014). Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia* 28 pp. 1299–1307. ,

Langerak A.W., Groenen P.J., Brüggemann M., Beldjord K., Bellan C., Bonello L., Boone E., Carter G.I., Catherwood M., Davi F., Delfau-Larue M.H., Diss T., Evans P.A., Gameiro P., Garcia Sanz R., Gonzalez D., Grand D., Håkansson A., Hummel M., Liu H., Lombardia L., Macintyre E.A., Milner B.J., Montes-Moreno S., Schuuring E., Spaargaren M., Hodges E. and van Dongen J.J. (2012) 'EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations.' *Leukemia,* 26(10) pp. 2159-2171

Langerak A.W., Brüggemann M., Davi F., Darzentas N., van Dongen J.J.M., Gonzalez D., Cazzaniga G., Giudicelli V., Lefranc M.P., Giraud M., Macintyre E.A., Hummel M., Pott C., Groenen P.J.T.A. and Stamatopoulos K. (EuroClonality-NGS Consortium). (2017) 'High-Throughput Immunogenetics for Clinical and Research Applications in Immunohematology: Potential and Challenges.' *Journal of Immunology,* 198(10) pp. 3765-3774

Liu, B., Morrison, C.D., Johnson, C.S., Trump, D.L., Qin, M., Conroy, J.C., Wang, J. and Liu, S. (2013) 'Computational methods for detecting copy number variations in cancer genome using next generation sequencing: Principles and challenges.' *Oncotarget* 4 pp. 1868–1881

Luthra R., Patel K.P., Routbort M.J., Broaddus R.R., Yau J., Simien C., Chen W., Hatfield D.Z., Medeiros L.J. and Singh R.R.(2017) 'A Targeted High-Throughput Next-Generation Sequencing Panel for Clinical Screening of Mutations, Gene Amplifications, and Fusions in Solid Tumors.' *The Journal of Molecular Diagnostics*. 19(2) pp. 255-264.

McCabe M.T., Ott H.M., Ganji G., Korenchuk S., Thompson C., Van Aller G.S., Liu Y., Graves A.P., Pietra A.D., Diaz E., LaFrance L.V., Mellinger M., Duquenne C., Tian X., Kruger R.G., McHugh C.F., Brandt M., Miller W.H., Dhanak D., Verma S.K., Tummino P.J. and Creasy C.L. (2012) 'EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations.' *Nature* 492 pp. 108-112

McKerrell T., Moreno T., Ponstingl H., Bolli N., Dias J.M., Tischler G., Colonna V., Manasse B., Bench A., Bloxham D., Herman B., Fletcher D., Park N., Quail M.A., Manes N., Hodkinson C., Baxter J., Sierra J., Foukaneli T., Warren A.J., Chi J., Costeas P., Rad R., Huntly B., Grove C., Ning Z., Tyler-Smith C., Varela I., Scott M., Nomdedeu J., Mustonen V. and Vassiliou G.S. (2016) 'Development and validation of a comprehensive genomic diagnostic tool for myeloid malignancies.' *Blood*, 128(1) pp. e1-9.

Montalvo M.L.G., Paris M., Tosi M., Lilliu S., Peruta B., Cavagna R., Belotti C., Zanghi P., Elidi L., Salmoiraghi S., INtermesoli T., Bassan R., Rambaldi A. and Spinelli O. (2016) 'Capture-based next

generation sequencing (NGS) identifies also rare IG/TCR rearrangements in adult acute lymphoblastic leukemia patients' Poster presented at: EHA, Vienna, 19th May

NICE. (2016) *Non-Hodgkin's lymphoma: diagnosis and management.* NICE guideline [NG52] [Online] [Accessed 3.6.2017] https://www.nice.org.uk/guidance/ng52

Pastore A., Jurinovic V., Kridel R., Hoster E., Staiger AM., Szczepanowski M., Pott C., Kopp N., Murakami M., Horn H., Leich E., Moccia AA., Mottok A., Hiddemann W., Gascoyne RD., Weinstock DM., and Weigert O. (2015) Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. *Lancet Oncology,* 16(9) pp. 1111-1112

Pedersen M.B., Hamilton-Dutoit S.J., Bendix K., Ketterling R.P., Bedroske P.P., Luoma I.M., Sattler C.A., Boddicker R.L., Bennani N.N., Nørgaard P., Møller M.B., Steiniche T., d'Amore F. and Feldman A.L. (2017) DUSP22 and TP63 rearrangements predict outcome of ALK-negative anaplastic large cell lymphoma: a Danish cohort study. *Blood,* [Online] 'First online' published 18th May 2017] DOI: 10.1182/blood-2016-12-755496

Pfarr N., Stenzinger A., Penzel R., Warth A., Dienemann H., Schirmacher P., Weichert W. and Endris V. (2016), 'High-throughput diagnostic profiling of clinically actionable gene fusions in lung cancer' *Genes Chromosomes Cancer,* 55(1) pp. 30-44

Pritchard, C. C., Salipante, S. J., Koehler, K., Smith, C., Scroggins, S., Wood, B., Wu D., Lee K.L., Dintzis S., Adey A., Liu Y., Eaton K.D., Martins R., Stricker K., Margolin K.A., Hoffman N., Churpek J.E., Tait J.F., King M-C. and Walsh, T. (2014). 'Validation and Implementation of Targeted Capture and Sequencing for the Detection of Actionable Mutation, Copy Number Variation, and Gene Rearrangement in Clinical Cancer Specimens. '*The Journal of Molecular Diagnostics*, 16(1) pp. 56–67.

Richards S., Aziz N., Bale S., Bick D., Das S., Gastier-Foster J., Grody W.W., Hegde M., Lyon E., Spector E., Voelkerding K., and Rehm H.L. (2015) 'Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology' *Genetics in Medicine,* 17(5) pp. 405-424

Robinson J.T, Thorvaldsdóttir H., Winckler W., Guttman M., Lander E.S., Getz G. and Mesirov J.P. (2011) Integrative Genomics Viewer. *Nature Biotechnology* 29, pp. 24–26

Rosenquist R., Rosenwald A., Du M.Q., Gaidano G., Groenen P., Wotherspoon A., Ghia P., Gaulard P., Campo E. and Stamatopoulos K. (2016) 'Clinical impact of recurrently mutated genes on lymphoma diagnostics: state-of-the-art and beyond.' *Haematologica,* 101(9) pp. 1002-1009

Rossi D., Gerber B. and Stüssi G. (2017) 'Predictive and prognostic biomarkers in the era of new targeted therapies for chronic lymphocytic leukemia.' *Leukemia Lymphoma* 58(7) pp.1548-1560

Rossi D, Ciardullo C and Gaidano G. (2013) 'Genetic aberrations of signaling pathways in lymphomagenesis: revelations from next generation sequencing studies.' *Seminars in Cancer Biology,* 23(6) pp. 422-430.

Samorodnitsky E., Datta J., Jewell B.M., Hagopian R., Miya J., Wing M.R., Damodaran S., Lippus J.M., Reeser J.W., Bhatt D., Timmers C.D. and Roychowdhury S. (2015) 'Comparison of custom capture for targeted next-generation DNA sequencing.' *The Journal of Molecular Diagnostics,* 17(1) pp. 64-75

Shen W., Szankasi P., Sederberg M., Schumacher J., Frizzell K.A., Gee E.P., Patel J.L., South S.T., Xu X. and Kelley T.W. (2016) 'Concurrent detection of targeted copy number variants and mutations using a myeloid malignancy next generation sequencing panel allows comprehensive genetic analysis using a single testing strategy.' *British Journal of Haematology*, 173(1) pp. 49-58.

Shin S., Hwang I.S., Kim J., Lee K.A., Lee S.T. and Choi J.R. (2017) 'Detection of Immunoglobulin Heavy Chain Gene Clonality by Next-Generation Sequencing for Minimal Residual Disease Monitoring in B-Lymphoblastic Leukemia.' *Annals of Laboratory Medicine,* 37(4) pp. 331-335

Song J. and Shao H. (2016) 'SNP Array in Hematopoietic Neoplasms: A Review.' *Microarrays*. 5(1) pp. 1-23

Strom S.P. (2016) 'Current practices and guidelines for clinical next-generation sequencing oncology testing' *Cancer, Biology & Medicine,* 13(1) pp. 3-11

Swerdlow S.H., Campo E., Harris N. L., Jaffe E.S., Pileri S.A., Stein H., Thiele J. and Vardiman J.W. (2008) *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed., Lyon: IARCPress.

Swerdlow S.H., Campo E., Pileri S.A., Harris N.L., Stein H., Siebert R., Advani R., Ghielmini M., Salles G.A., Zelenetz A.D. and Jaffe E.S. (2016) 'The 2016 revision of the World Health Organization classification of lymphoid neoplasms' *Blood*, 127(20) pp. 2375-2390

Tamai H. and Inokuchi K. (2010) '11q23/MLL acute leukemia : update of clinical aspects.'

Journal of Clinical Experimental Hematopathology, 50(2) pp. 91-8.

Thompson P.A. and Wierda W.G. (2016) 'Eliminating minimal residual disease as a therapeutic end point: working toward cure for patients with CLL' *Blood,* 127 pp. 279-286

Tinguely M., Thies S., Frigerio S., Reineke T., Korol D. and Zimmermann D.R. (2013) 'IRF8 is associated with germinal center B-cell-like type of diffuse large B-cell lymphoma and exceptionally involved in translocation t(14;16)(q32.33;q24.1).' *Leukemia Lymphoma*, 55(1) pp. 136-142

Uzilov A.V., Ding W., Fink M.Y., Antipin Y., Brohl A.S., Davis C., Lau C.Y., Pandya C., Shah H., Kasai Y., Powell J., Micchelli M., Castellanos R., Zhang Z., Linderman M., Kinoshita Y., Zweig M., Raustad K., Cheung K., Castillo D., Wooten M., Bourzgui I., Newman L.C., Deikus G., Mathew B., Zhu J., Glicksberg B.S., Moe A.S., Liao J., Edelmann L., Dudley J.T., Maki R.G., Kasarskis A., Holcombe R.F., Mahajan M., Hao K., Reva B., Longtine J., Starcevic D., Sebra R., Donovan M.J., Li S., Schadt E.E. and Chen R. (2016) 'Development and clinical application of an integrative genomic approach to personalized cancer therapy.' *Genome Medicine,* 8(1) pp. 62-72

van Dongen J.J., Langerak A.W., Brüggemann M., Evans P.A., Hummel M., Lavender F.L., Delabesse E., Davi F., Schuuring E., García-Sanz R., van Krieken J.H., Droese J., González D., Bastard C., White H.E., Spaargaren M., González M., Parreira A., Smith J.L., Morgan G.J., Kneba M. and Macintyre E.A. (2003) 'Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936.' *Leukemia,* 17(12) pp. 2257-2317.

Vaqué J.P., Martínez N., Batlle-López A., Pérez C., Montes-Moreno S., Sánchez-Beato M. and Piris M.A. (2014) 'B-cell lymphoma mutations: improving diagnostics and enabling targeted therapies.' *Haematologica*, 99(2) pp. 222-31

Vosberg S., Herold T., Hartmann L., Neumann M., Opatz S., Metzeler K.H., Schneider S., Graf A., Krebs S., Blum H., Baldus C.D., Hiddemann W., Spiekermann K., Bohlander S.K., Mansmann U. and Greif P.A. (2016) 'Close correlation of copy number aberrations detected by next-generation sequencing with results from routine cytogenetics in acute myeloid leukemia.' *Genes Chromosomes Cancer*, 55(7) pp. 553-567.

Walker B.A., Wardell C.P., Johnson D.C., Kaiser M.F., Begum D.B., Dahir N.B., Ross F.M., Davies F.E., Gonzalez D. and Morgan G.J. (2013) 'Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells.' *Blood,* 121(17) pp. 3413-3419

Walker B.A., Boyle E.M., Wardell C.P., Murison A., Begum D.B., Dahir N.M., Proszek P.Z., Johnson D.C., Kaiser M.F., Melchor L., Aronson L.I., Scales M., Pawlyn C., Mirabella F., Jones J.R., Brioli A., Mikulasova A., Cairns D.A., Gregory W.M., Quartilho A., Drayson M.T., Russell N., Cook G., Jackson G.H., Leleu X., Davies F.E. and Morgan G.J. (2015) 'Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma.' *Journal of Clinical Oncology,* 33(33) pp. 3911-20

Wallis Y., Payne S., McAnulty C., Bodmer D., Sistermans E., Robertson K., Moore D., Abbs S., Dean

Z. and Devereau A. (2013). *Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics.* [Online] [Accessed 2.3. 2017]

http://www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf

Wang M., Zhang S., Chuang S.S., Ashton-Key M., Ochoa E., Bolli N., Vassiliou G., Gao Z. and Du M.Q. (2017) 'Angioimmunoblastic T cell lymphoma: novel molecular insights by mutation profiling.' *Oncotarget*. 14;8(11) pp. 17763-17770

Wenric S., Sticca T., Caberg J.H., Josse C., Fasquelle C., Herens C., Jamar M., Max S., Gothot A., Caers J. and Bours V. (2017) 'Exome copy number variation detection: Use of a pool of unrelated healthy tissue as reference sample.' *Epidemiology*, 41(1) pp. 35-40

Woyach J.A. and Johnson A.J. (2015) 'Targeted therapies in CLL: mechanisms of resistance and strategies for management' *Blood,* 126(4) pp. 471-477

Wren D., Walker B.A., Brüggemann M., Catherwood M.A., Pott C., Stamatopoulos K., Langerak A.W. and Gonzalez D.; EuroClonality-NGS consortium. (2017) 'Comprehensive translocation and clonality detection in lymphoproliferative disorders by next-generation sequencing.' *Haematologica*, 102(2) pp.e57-e60.

Yun S., Johnson A.C., Okolo O.N., Arnold S.J., McBride A., Zhang L., Baz R.C. and Anwer F. (2017) 'Waldenström Macroglobulinemia: Review of Pathogenesis and Management.' *Clinical Lymphoma Myeloma Leukemia*, 17(5) pp. 252-262.

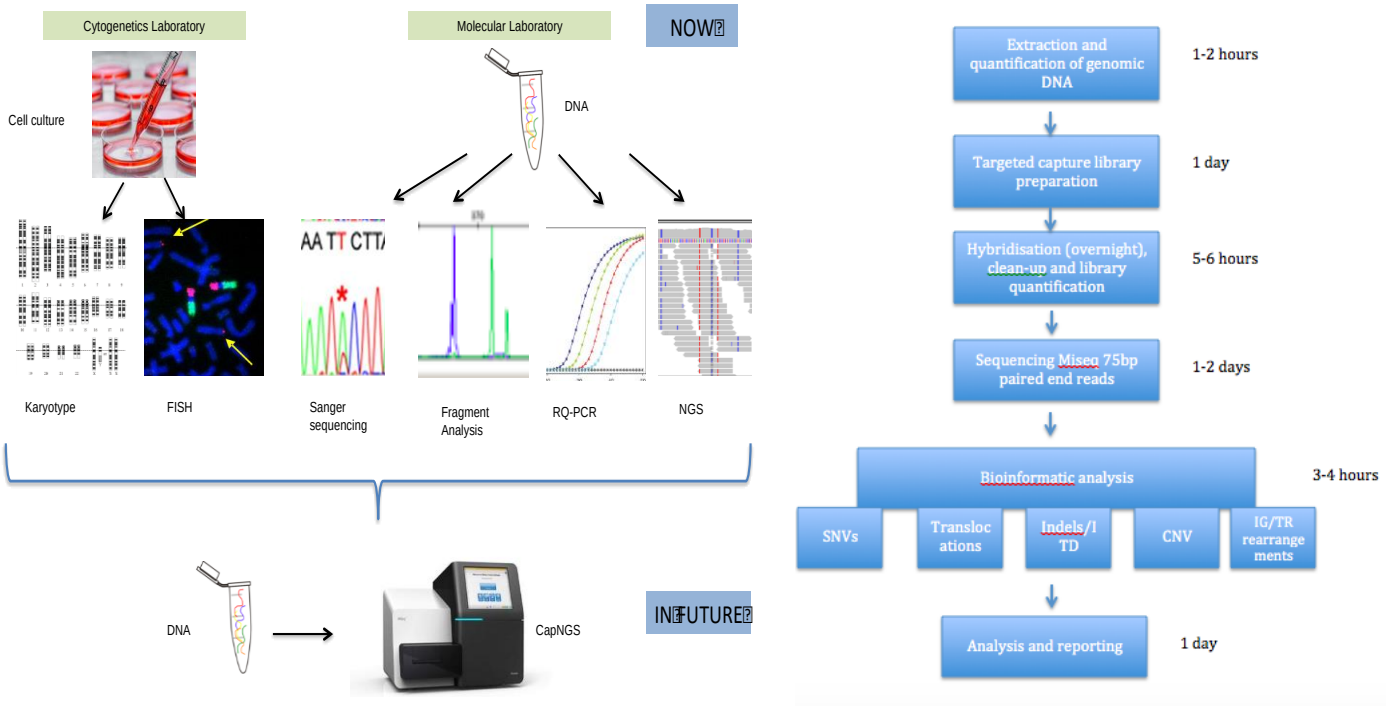
Zare F., Dow M., Monteleone N., Hosny A. and Nabavi S. (2017) 'An evaluation of copy number variation detection tools for cancer using whole exome sequencing data.' *BMC Bioinformatics*, 18(1) pp. 286-270

Zhao, M., Wang, Q., Wang, Q., Jia, P., and Zhao, Z. (2013) 'Computational tools for copy number variation (cnv) detection using next-generation sequencing data: Features and perspectives.' *BMC Bioinformatics* 14, S1.

Context of the innovation

Next generation sequencing technologies are rapidly being incorporated into diagnostic laboratories for the detection of clinically relevant SNV and CNVs in solid cancers (Sahm *et al.,* Passinen-Sohns *et al.,* 2017) and myeloid disorders (Hamblin *et al.,* 2017). In the context of lymphomas, two additional targets provide essential clinical information: translocations and Immunoglobulin-gene (IG) and T-cell receptor (TCR) gene rearrangements (Swerdlow *et al.,* 2016, Langerak *et al.,* 2012). Evidence collated in the literature review shows that a capture-based NGS approach is capable of detecting translocations, including those arising from the IG/TR loci, as well as IG/TCR gene rearrangements in addition to SNV and CNVs. Although further validating studies need to be performed to demonstrate equal performance of CapNGS compared to established methodologies for these targets there are many potential benefits associated with such an approach: (a) A large number of genes and loci associated with SNVs and CNVs in lymphomas can be analysed in parallel; (b) Known translocations arising from any of the IG/TCR loci as well as other lymphoma-associated translocations can be detected, including those where the translocation partners is unknown (Walker *et al.,* 2013); (c) It offers analysis of IG/TCR rearrangements analogous to current clonality testing with the advantage of providing the nucleotide sequence enabling a more detailed investigation of the clonal architecture and the relationship between clonal cell populations (Langerak et al., 2017). The proposed innovation is to design and validate a novel CapNGS panel capable of the *concurrent* detection of clinically relevant SNV, CNVs and translocations whilst also providing detailed information on the clonal IG/TR rearrangement present in the sample with the view, if successful, to replace the current multi-modality approach with one universal CapNGS set-up for all lymphomas samples received for molecular testing by SIHMDS (figure B1). This approach and specifically clonality and translocation detection by CapNGS has not been used in the diagnostic setting before.

Figure B1: SIHMDS testing pathways showing the different processes and methods currently employed compared to the CapNGS workflow.



Benefits of the innovation for the laboratory

Improving the scope of laboratory testing

The list of targets to be included in the CapNGS design covers all clinically relevant targets (Swerdlow *et al.,* 2016, Rosenquist *et al.,* 2016) together with targets likely to be relevant in future based on current research data (for example *DUSP22* and *TP63* rearrangements in *ALK*-negative ALCL (Pederson *et al.,* 2017)) or novel targeted treatment approaches (e.g. EZH2 inhibitors (McCabe *et al.,* 2012)) and covers xxbp in total. In addition to mutational hotspots, whole genes are included to enable the detection of rare mutations that are likely to have the same effect (e.g. Ibrutinib-resistance causing mutations in *BTK* (Woyach and Johnson, 2015)) with overlap of 15bp at each side of the exons to detect splice site mutations (Kotelnikova *et al.,* 2016). The design for CNV analysis is under review to establish the best selection of genome-wide SNPs akin to the model proposed by McKerrell *et al.* (2016). In contrast to FISH analysis, detection of translocations by CapNGS does not require the knowledge of the translocation partner, allowing for the detection of any translocation arising from the IG/TCR loci (Walker *et al.,* 2013). The bait design covers the entire V(D)J genes (thus enabling clonality analysis) plus the switch regions of the constant genes to cover translocations arising during V(D)J rearrangement, SHM or class-switch recombination (Walker *et al.,* 2015). The scalability of NGS methods (Hovelson *et al.,* 2015) will provide a flexible tool to add additional targets as they are described. The inclusion of markers not yet in clinical practice provides an excellent basis to be incorporated in research studies, collaborations or clinical trials. Long-term this flexibility will provide a competitive edge for the SIHMDS against competition from other NHS and private laboratories.

Streamlining laboratory workflows

The workflow for CapNGS is outlined in figure B1 with estimated timings for each of the steps involved. Based on our experience with other NGS capture panels for SNV detection, performing one set up per week reliably allows for reports to be released within 10-12 working days, which is in-line with clinical requirements (NICE, 2016) and faster than the average FISH and Sanger sequencing TAT of 15 working days thus enabling integrated reports to be released well within the expected 4 week timeframe (NICE, 2016). Under current workflows, a lymphoma sample is received by the general specimen reception and booked onto the hospital database. If required, a sample is send to the cytogenetics department for FISH analysis and another to the molecular department for mutation or clonality assessment. Following the completion of the respective laboratory investigations, the results are reported as stand-alone reports, transferred onto the hospital database under each separate laboratory tab and the information needs to be retrieved when the integrated report is written. CapNGS produces all results in a single workflow, which can be reported together facilitating integrated report writing and reducing delays and potential errors due to missing information. The new workflow would bring the haemato-oncology work of the department in-line with the solid cancer workstreams, most of which are already NGS-based. This would reduce the number of instruments required in the laboratory releasing much –needed space, reduce the amount of paperwork associated with each separate procedure (SOPs, maintenance and training documentation), reduce the requirement for cross-training of staff on different methods and minimise the number of different kits and reagents bought from different suppliers with the potential of bulk discounts for NGS reagents. Streamlining of processes would also reduce the work and documentation required for UKAS accreditation and could free up technical staff to take on other duties.

Semi-automated analysis

Once designed and validated, the bioinformatics pipeline combines the different molecular variants found in a sample in one output file (Mckerrell *et al.,* 2016), making analysis and reporting easier, less time consuming and by reducing the number of data transfers less error prone. In particular for clonality analysis, the availability of the exact sequence information for each clonotype together with the design of specific bioinformatics algorithms, is predicted to improve read-outs by making interpretation less subjective (Langerak *et al.,* 2017). A pre-defined tier-system for variants (Tier 1 = known pathogenic, Tier 2= likely pathogenic, Tier 3= Variant of unknown significance, Tier 4= known SNP) should be established and ratified by the medical Consultants and Clinical scientists in accordance with best-practice guidelines (Richards *et al.,* 2015) and based on regular review of the published literature. This could be fed into the pipeline providing automated annotation for variants detected and reducing the number of variants that need to be reviewed with regards to their likely relevance (Wallis *et al.,* 2013).

Costs

At this stage the cost for the panel is difficult to establish as the final design needs to be agreed and the validation in the clinical context performed to establish required read depths which will have an impact on the number of patient samples that can be pooled on each run affecting costs. The aim is to achieve a run cost of around £500 which is higher than the stipulated £300 target set by Cancer Research UK (Hamblin *et al.,*

2017), however this panel provides a larger body of information and the specific requirement of IG/TR loci investigation makes this a special case. Only one study has been published looking at the cost of running a small targeted NGS panel for mutation detection in a NHS laboratory (Hamblin *et al.,* 2017) and their data suggests that as soon as more than one target needs to be investigated, NGS-based approaches are a viable alternative to multimodality testing. Our current assay costs are provided in the appendix and even without further analysis it is clear that if a patient with myeloma requires an IG translocation work-up by FISH together with additional mutations like *KRAS* and *TP53* then the CapNGS panel could be cheaper (current cost around £600 personal communication K.B.).

Benefits of the innovation for patients

The key benefit for patients is the incorporation of all molecular targets currently required for diagnosis and to assess prognosis and predict treatment response, together with additional markers that could suggest experimental treatment options or allow entry into a clinical trial (Pritchard *et al.,* 2014), especially in those patients with disorders in the 'not otherwise specific' group. The panel will specifically help in the differential diagnosis of HCL (*BRAF+/MAP2K-)* versus variant HCL (*MAP2K+/BRAF-),* separating WM from myeloma and SMZL if a *MYD88* mutation can be demonstrated and in picking out MCL with a t(11;14) translocation in CD5+ B-NHL and FL/MALT IG-translocation positive cases in CD5- B-NHLs. Patients will also welcome a faster turn-around time and a more reliable availability of all their results for the integrated report. As only DNA is required, stored DNA can be reassessed using this panel when available in patients that could benefit from clinical trials due to disease progression or treatment resistance. The detailed analysis of translocation breakpoints and clonal IG/TCR rearrangements sequences will be available for all patients even if not required for the diagnosis and could provide the basis for patient specific MRD assays akin to those performed in ALL (Hoelzer *et al.,* 2016) and CLL (Thompson and Wierda, 2016). From this information the relationship between two separate lesions can be established with greater certainty than before, which will be crucial in selecting high-dose regimen for patients with clonally-related neoplasms as opposed to those presenting with two individual clones, who will receive standard first-line therapy as needed. Patients with MBL and MGUS may also benefit from additional information gleaned form this panel in terms of mutations, translocations and IG stereotype to understand their risk of disease progression.

Evidence of stakeholder engagement

<u>Haemato-Oncology Consultants</u>

An interview was conducted with the lead consultant of the lymphoma service and questionnaires were send to all other haemato-oncology and histopathology consultants (see appendix for template). All differentiated between current use and predicted greater potential use in future. Four out of 5 replies stated that histopathology and immunophenotyping currently remain sufficient for the diagnosis in the majority of patients but with *BRAF* and *MYD88* already included in the new WHO classification (Swerdlow *et al.,* 2016) and further data published everyone was in agreement that there is an increasing need for more comprehensive testing with one example being the identification of *MYC, BCL2*, and *BCL6* translocations in double-hit

DLBCL, which has been incorporated into the latest NICE guidance for DLBCL management (NICE, 2016). Similarly, it was deemed a very useful tool to expand the evaluation of those cases currently undergoing clonality testing to distinguish reactive from clonal and it is expected that the NGS results will provide a clearer answer than current methodologies, particularly in T-LPD. The main advantage identified by all was the reduced need for multiple FISH tests (reducing costs) and the potential to extend the CNV testing, as we do not offer arrays in house. The main disadvantage identified was the surplus information obtained, especially in patients not currently needing treatment as it was felt that this could cause anxiety in patients, although the lead clinician did not feel that this was money wasted as the information could become relevant at a later date.

Euroclonality Expert consortium

My design of the current panel was based on an extensive literature review and input from all the leading experts within the EC consortium: telephone conferences were held and the final design was agreed at one of the EC meetings. The panel is being validated for translocation, SNV and clonality detection in a pan-European study conducted by EC together with the development of a tailored bioinformatics pipeline.

Implementation plan

There are two gateways for the implementation of this innovation project: One to design and validate the CapNGS panel in the clinical context and as part of the EC study to ensure equal or better sensitivity, specificity and accuracy compared to current established methods and two, if this is successful, to evaluate whether the replacement of current technologies and workflows by CapNGS is cost effective in the SIHMDS setting. Potential barriers for each stage are summarized in the table below:

Potential barriers – Gateway 1
Validation through the collaborative effort of the EC consortium may take longer than if it were done in-house
EC-developed dedicated bioinformatics pipeline is an essential tool and has yet to be fully released
Cost
Potential barriers – Gateway 2
CNV detection is not included in the EC validation thus needs to be performed in house in collaboration with a laboratory with expertise in arrays
Cost of additional validation
Competition by commercial companies/private laboratories although clonality testing is a niche market
WGS may be introduced sooner than expected potentially rendering this panel obsolete
Workload for the validation and the introduction requires dedicated staff and time
Failure to obtain UKAS accreditation
Reduction in FISH testing may lead to resentment amongst staff in the cytogenetic department, loss of expertise, need for re-training
Price may be too high to make this approach viable in enough cases to cover costs; external referrers may stop sending samples
Lack of NHS commissioning for NGS panel testing

Executive summary/lay person summary

New insight into the molecular changes and the cellular pathways implicated in the different types of lymphomas has increased our understanding of the underlying pathogenesis but has also highlighted the complexity and diversity of molecular aberrations present in each patient. Importantly, we are beginning to

understand the significance of these molecular changes in the context of different disease entities. This helps us define the prognosis for the patient more accurately and to predict effectiveness of specific drugs and treatment regimes allowing us to move towards personalised medicine increasing the efficacy of treatment whilst reducing toxicity. This not only improves outcome and quality of life for patients but also reduces the amount of money spent on treatment unlikely to provide any benefit. An ever expanding list of recurrent mutations at the chromosomal and DNA sequence level is emerging and whilst only a minority of these have a direct impact on patient management currently, it is clear that their evaluation will become mandatory in future to provide the best care to each patient. At present, the standard laboratory testing repertoire for the diagnosis and management of patients with lymphoma as part of the Specialist Integrated Haematological Malignancies Diagnostic Service (SIHMDS) consists of more than twenty different targets, analysed by ten different laboratory techniques, each requiring dedicated equipment, workflows and expertise situated within different laboratory departments (figure B1). Next-generation sequencing (NGS) methods have been shown to be capable of analysing all the different types of molecular targets clinically relevant in the context of lymphomas and it is proposed to design and evaluate a capture-based NGS panel incorporating all molecular markers currently needed for the diagnostic, prognostic and predictive investigations performed in patients with lymphoma. The design will include genes and hotspots, known to be affected by recurrent mutations (e.g. *TP53, MYD88*), provide information on copy-number variations (e.g. 17pdel and gain(1q) as poor prognostic markers in CLL and myeloma, respectively) and allow the identification of lymphoma-associated translocations (e.g. *IGH-BCL2, IGH-CCND1*) and clonal Immunoglobulin and TCR- gene rearrangements to ascertain the clonal, neoplastic origin of a lymphocytosis and to provide information on potential minimal-residual disease markers to monitor treatment responses.

The implementation of such an assay would dramatically reduce the complexity of workflows within SIHMDS, shorten turn-around times and would provide results simultaneously, facilitating integrated reporting as required by NICE. Not only is such an assay likely to result in a reduction in cost but it also requires less patient material making it more applicable to different sample types (including small biopsies). Finally, the scalability of NGS panels allows for the addition of further molecular markers, ensuring patients will benefit from new knowledge and evidence, whilst also ensuring the SIHMDS remains competitive providing the best possible, state-of-the-art laboratory service.

References are provided at the end of Part A

Appendix 1: Overview of prices for the current cytogenetic and molecular tests performed for lymphoma investigations as part of the SIHMDS. Prices include costs for equipment, maintenance, staff (incl. training, turn-over and annual leave), and overheads.

Test	**Method**	**Price**	**Price charged (£)**
NGS panel	Capture NGS	(Target £500)	
Clonality	Genescan	£106 (B) + £111 (T)	£116 (B) + £116 (T) (>70% of cases order both)
IGHV SHM	Sanger sequencing	£78	88
TP53 mutations	NGS panel (CRC amplicon)	£105 (NGS)	158
17pdel	FISH	£222	222
IGH break	FISH	£440	440
CNV/large chromosomal aberrations	Karyotype/FISH	£495-770 (karyotype + multiple FISH	495-770 (karyotype + multiple FISH
KIT	ddPCR (external)	£100	120
MYD88	Sanger sequencing (external)	£120	100
BRAF	COBAS	£120	120
MRD maker identification (IG/TR)	(external)	Unknown	Unknown
EZH2	Not offered	NA	NA
BTK/PLCG2	Not offered	NA	NA

Appendix 2: Questionnaire distributed to haemato-oncology and histopathology consultants who are part of the SIHMDS.

Capture NGS panel providing analysis of:

- *Mutations in genes relevant for diagnosis/prognosis/predictive (including TP53, MYD88, CXCR4, BRAF, BTK,PLCg2, CARD11, PIK3, NOTCH1 + 2 and a list of other mutations described in the literature; <u>assume</u> all "clinically relevant" mutations are included)*
- *Translocations arising from all IG (IGH, IGK, IGL) and TR (TRG, TRB, TRA, TRD) loci*
- *Translocation of MALT (11;18), CCND1, ALK and BCL2/6 even if not partnered with IG or TR*
- *V(D)J rearrangement info for all IG/TR loci: clonality + sequence details to look in more detail at oligoclonality/subclones/relationship between clones; able to identify MRD markers*
- *Detect CNVs (9p (CDKN2A/B), 17p (TP53), 11q (CCND1, ATM), 8q (MYC), 7p (PAX5), 13q (BTG1), 12q(ETV6, PIK3cA))- others could be added*

TAT 10-15 working days, DNA from PB/BM/FFPE

Rather difficult to have an idea on price for this at the moment but based on current pricing let's hypothesise this would be somewhere in the region of the combined price of B+T clonality + VDJ+ TP53 + Single FISH study or FISH panel + TP53; (min £500)

1. Could you see this as a tool used for:
    a. Patients with lymphocytosis (what I mean is a more generic first-line work up where lymphoma has not been diagnosed but is a possibility; e.g. when clonality studies are done now)
    b. Providing a diagnosis according to WHO (for which cases would this be particularly useful?)
    c. Deciding between different differential diagnoses (again most useful examples)
    d. Comprehensive investigation done for all lymphomas at diagnosis
        i. If not all which ones do you think this would be most useful for?

   ii. Would you think this would be rather more relevant for (instead of diagnostic scenario)
      1. Treatment decision
      2. Difficult cases/relapse cases/several diagnostic options only (i.e. selected cases rather than a broad-tool)

2. Would you ask for karyotype in all cases/when would you ask for a karyotype/additional FISH studies if you had run this panel?

3. If this demonstrates same or better sensitivity and specificity for IG/TR translocations compared to FISH, do you see any reason not to replace FISH with this approach? (Some rare translocations will be missed if unusual breakpoints fall outside the capture area)
4. If this demonstrates the same or better sensitivity and specificity for CNVs in the targeted regions compared to karyotype/aCGH assays could you see a role for this approach in the routine work up instead of aCGH and karyotyping?

5. Should a 'tier'- system be predefined for the different aberrations? e.g. tier one = variants with known clinical impact (prognosis/treatment), tier two = part of same pathway as known mutations, not tested prospectively, tier 3= likely pathological/described in other cancers, tier 4= others/no knowledge of relevance
   a. Should the laboratory only report pre-defined variants and exclude targets with no known clinical relevance from the analysis (those that would not alter patient management at present)?

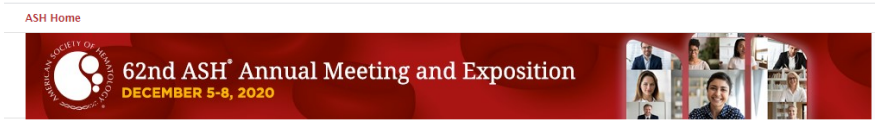
6. Do you think patients would be happy to consent to use the data derived on other targets for research and study purposes (would that be included in current consent?)
7. In your opinion what are the main advantages/disadvantages of this approach?

8. Do you think there is a need/will be a need for
   a. routine MRD monitoring on B/T cell lymphomas and if so in which type of disease is this most likely to become routine practice (if any)?
   b. Stereotype analysis of VDJ rearrangements.
   c. Defining subclones/oligoclonality.
   d. Molecular characterisation in scenarios like MGUS/MBL to try and understand which patients will progress/catch progression early.

Appendix A10
ASH abstract 2020

1109 Targeted Next Generation Sequencing Improves Diagnosis in Unclassifiable Leukemic Indolent B-Cell Non-Hodgkin Lymphoma and Identifies a Subset with Recurrent MYD88 Mutations in a Prospective Multicentre Study

Available at https://ash.confex.com/ash/2020/webprogram/Paper141542.html



-Author name in bold denotes the presenting author
-Asterisk * with author name denotes a Non-ASH member
denotes an abstract that is clinically relevant.

denotes that this is a recommended PHD Trainee Session.

denotes that this is a ticketed session.

1109 Targeted Next Generation Sequencing Improves Diagnosis in Unclassifiable Leukemic Indolent B-Cell Non-Hodgkin Lymphoma and Identifies a Subset with Recurrent MYD88 Mutations in a Prospective Multicentre Study

Program: Oral and Poster Abstracts
Session: 621. Lymphoma—Genetic/Epigenetic Biology: Poster I
Hematology Disease Topics & Pathways:
Diseases, Non-Hodgkin Lymphoma, Technology and Procedures, Lymphoid Malignancies, Clinically relevant, NGS

Saturday, December 5, 2020, 7:00 AM-3:30 PM

Matthew J Cross, MBBS, BSc[1*], *Grzegorz Pietka, PhD*[2*], *Sarah Dunne*[3*], *Leonora Conneely*[3*], *Michael Hubank, PhD*[4*], *Corinne De Lord, MD, FRCPath*[5*], *Michelle Furtado, MD, FRCPath, PhD*[6*], *Pawel Kaczmarek, MD, PhD*[7*], *Dima El-Sharkawi*[3*], *Claire E. Dearden, MD, FRCP, FRCPath*[7*], *Kabir Mohammed*[8*], *Alan Stewart Dunlop, MSc, BS*[9*], *Ricardo Morilla*[10*], *Karol Pal, PhD*[11,12*], *Nikos Darzentas, PhD*[12*], *David Gonzalez, PhD*[13*], *Dorte Wren, PhD FRCPath*[2*] *and Sunil Iyengar, MD, FRCPath, PhD*[3*]

[1]*Department of Haemato-Oncology, Royal Marsden Hospital, Sutton, ENG, United Kingdom*
[2]*Clinical Genomics, Royal Marsden Hospital, Sutton, United Kingdom*
[3]*Department of Haemato-Oncology, Royal Marsden Hospital, Sutton, United Kingdom*
[4]*Clinical Genomics, The Royal Marsden Hospital, Sutton, United Kingdom*
[5]*Epsom and St Helier University Hospitals NHS Trust, Carshalton, United Kingdom*
[6]*Royal Cornwall Hospital, Truro, United Kingdom*
[7]*Department of Haematology, Surrey and Sussex Healthcare NHS Trust, Redhill, United Kingdom*
[8]*Royal Marsden Hospital, London, United Kingdom*
[9]*Immunophenotyping, Royal Marsden Hospital, Sutton, ENG, United Kingdom*
[10]*Immunophenotyping, Royal Marsden Hospital, Sutton, United Kingdom*
[11]*Department of Molecular Medicine, CEITEC MU – Central European Institute of Technology, Masaryk University, Brno, Czech Republic*
[12]*Department of Hematology, University Hospital Schleswig-Holstein, Kiel, Germany*
[13]*Precision Medicine Centre of Excellence, Centre for Cancer Research & Cell Biology, Queens University Belfast, Belfast, United Kingdom*

Introduction: Patients with leukemic indolent B-cell Non-Hodgkin Lymphomas (LI B-NHL) that do not easily fall into a distinct pathological category with standard diagnostics can pose a challenge as their optimal management is uncertain. Lack of a clear diagnostic label can deny patients access to novel therapies as regulatory approval of drugs is largely granted within histological categories. Moreover, these patients may be excluded from participation in clinical trials. We report findings from a prospective study evaluating targeted next generation sequencing (NGS) of circulating malignant B-cells in this setting.

Methods: Over a period of 4 years, 108 patients with LI B-NHL, deemed unclassifiable on standard peripheral blood (PB) diagnostics, were prospectively enrolled from 14 centres around the UK including ours. Patients with non-clonal lymphocytosis, confirmed CLL-type MBL, CLL (Matutes score 4 or 5), MCL, HCL, high grade lymphoma on standard PB diagnostics were excluded. Clinical and morphological characteristics were analysed. PB immunophenotyping was used for characterisation including ROR1 expression. CD15+ cells were positively selected with magnetic beads for germline DNA. Tumor-germline pairs were analysed with the Euroclonality-NGS DNA capture panel, and ARResT/Interrogate and complementary pipelines, to characterise clonal immunoglobulin rearrangements, translocations and somatic mutations (Stewart *et al.* ASH annual meeting 2019). Germline variants were subtracted to improve somatic mutation detection.

Results: Median age was 72yrs with M:F ratio of 2:1. Incidental lymphocytosis on a routine blood count was the commonest mode of presentation with only a third of patients symptomatic at presentation. Median lymphocyte count was 12.4 x 109/L at the time of sampling. Just under 50% of cases had partial or complete CD5 expression, CD19 and 20 expression was moderate to strong in most cases while ROR1 was negative in most cases. NGS was able to assign a diagnostic category in 11/90 (12%) cases based on detection of *IGH;CCND1* (6), *IGH;BCL2* (2), *IGH;BCL3* (2) translocations and in one case a *BRAF* mutation with confirmation of hairy cell leukaemia on immuno-morphology. These tests were not triggered by standard diagnostic pathways at initial presentation. Two novel translocations of uncertain significance were detected – *RAD51B:BIRC3* and *IGHM*switch-*MICALL2/INTS1*.

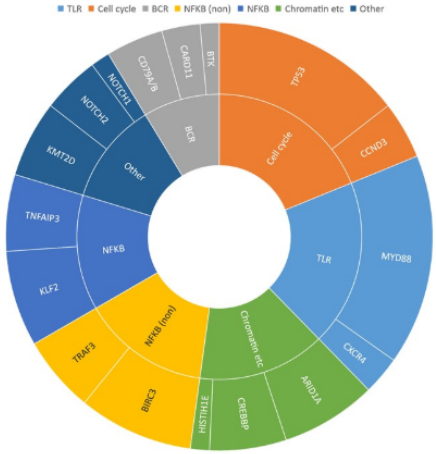
Clonal *IGHV-IGHD-IGHJ* rearrangements were detected by NGS and/or Sanger sequencing with predominantly *IGHV3* and *IGHV4* gene usage, *IGHV4-34* being the most common with 15 cases. The majority of rearrangements showed somatic hypermutations above 2% and none could be assigned to the 19 major CLL-associated stereotyped subsets (using ARResT/AssignSubsets). Somatic mutations were detected in 74/108 cases. The most frequently mutated genes were *TP53* in 16/108 (15%) and the hotspot *MYD88* L265P mutation in 14/108 (13%) with VAF of 20-46%. Other variants detected were distributed among genes associated with NFkB signalling and chromatin remodelling (Figure 1).

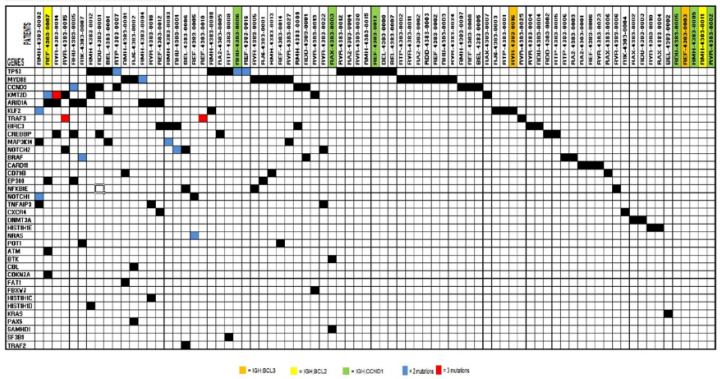
The *MYD88* L265P mutation was present in a fifth of cases with mutations detected (14/74) and was the sole change seen in 7 cases. Accompanying mutations in the remaining cases included *CD79B* (2), *POT1* (1), *NFKBIE* (1). No concomitant *CXCR4* mutations were detected. Median age of the *MYD88* L265P cohort was 68 years with male predominance (M:F=10:4). 9/14 presented with an incidental lymphocytosis and median count of 11.18 x 109/L. Lymphadenopathy was present in 3/14 (21%) while 7/14 (50%) had splenomegaly. A paraprotein was detected in 6/14 cases (4 IgM, 1 IgG, 1 IgA). CD5 was expressed in 6/14 cases by flow cytometry. CD19 and 20 were uniformly positive, CD79b was variable and 12/14 cases tested did not express ROR1.

Conclusion: This prospective study outlines the value of a targeted NGS panel in enhancing the diagnosis of unclassifiable LI B-NHL thereby improving patient access to novel therapies and clinical trials. It highlights recurrent *MYD88* mutations in a subset of patients that have splenomegaly as a frequent feature. 5 year clinical follow up data, currently being gathered in both the *MYD88* mutated and overall cohort, will be valuable in further characterisation and risk stratification to inform management of these patients.

Figure 1. Overview of mutations, translocations detected and key pathways involved.

Pathway	Gene	Number of variants detected
	MYD88	14
TLR	CXCR4	2
	TP53	16
Cell cycle	CCND3	9
	CARD11	3
	BTK	1
BCR	CD79B	3
	BIRC3	6
NFKB (non)	TRAF3	3
	TNFAIP3	3
NFKB	KLF2	6
	CREBBP	6
	ARID1A	7
Chromatin etc	HISTIH1E	2
	KMT2D	6
	NOTCH2	4
Other	NOTCH1	2





Disclosures: **Furtado**: *Abbvie:* Other: Conference Support. **El-Sharkawi**: *Abbvie:* Consultancy, Honoraria, Speakers Bureau; *AstraZeneca:* Consultancy, Honoraria, Speakers Bureau; *Janssen:* Consultancy, Speakers Bureau; *Takeda:* Honoraria, Speakers Bureau; *Innate:* Consultancy. **Iyengar**: *Abbvie:* Honoraria; *Gilead:* Consultancy, Honoraria, Speakers Bureau; *Takeda:* Consultancy, Honoraria, Speakers Bureau; *Janssen:* Honoraria; *Beigene:* Consultancy.

Appendix 11: (Wren et al., 2017)

© Ferrata Storti Foundation

Comprehensive translocation and clonality detection in lymphoproliferative disorders by next-generation sequencing

Detection and characterization of clonal immunoglobulin (IG)/T-cell receptor (TR) rearrangements and translocations in lymphoproliferative neoplasms provide critical information in the diagnostic pathway and are valuable tools for addressing research questions involving B- and T-cell lymphoproliferative disorders (LPD).[1,2] These include ascertaining the clonal nature of lymphoid proliferations,[3,4] characterization of translocations in lymphomas and leukemias,[4] characterization of CDR3 regions for minimal residual disease target identification[5] and stereotyping analysis.[6] Until recently, collecting this information required a combination of different methodologies, such as gene-scanning/heteroduplex analysis, fluorescence *in situ* hybridization (FISH) and Sanger sequencing. The incorporation of next-generation sequencing (NGS) in clinical laboratories opens up new possibilities since an integrated NGS approach can provide data on sequence and structural variation in a single assay, including translocations and IG/TR rearrangements, and has been shown to be successful for the characterisation of IG translocations in myeloma and lymphomas.[7,8]

Within the EuroClonality-NGS consortium, we have designed a capture-based protocol covering the coding V, D and J genes of the IG/TR loci, as well as switch regions in the IGH locus. This design allows the identification of D-J and V-(D)-J rearrangements as well as chromosomal translocations involving IG/TR genes by sequencing through the breakpoint regions in genomic DNA. We piloted this approach using a sample cohort (n=24) consisting of three B-cell precursor acute lymphoblastic leukemias (BCP-ALL), four Burkitt lymphomas, eight chronic lymphocytic leukemias (CLL), two splenic marginal zone lymphomas (SMZL), two diffuse large B-cell lymphomas (DLBCL), two follicular lymphomas, two T-cell acute lymphoblastic leukemias (T-ALL) and one T-cell non-Hodgkin lymphoma (T-NHL). Twenty-one cases were known to carry a translocation arising within the IG/TR loci with the remaining three cases being included for their well characterized D-J or V-(D)-J gene rearrangements. Libraries were constructed from 1 µg of genomic DNA which was fragmented to an average of 200 bp using an E220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). Fragmented DNA was processed using the TruSeq DNA LT sample preparation kit (Illumina, Cambridge, UK). Libraries were hybridized to a custom-designed EZ SeqCap gene panel (Roche-Nimblegen, Madison, MI, USA) which encompassed 180 kb containing the V, D, J and constant regions of the IG and TR loci as well as the switch regions of the IGH locus. Enriched samples were sequenced on a MiSeq (Illumina) using 75 bp or 120 bp paired-end reads. Reads were aligned to the reference genome (hg19) with translocations and variants called using a previously described bioinformatics pipeline.[9] The average depth of sequencing in the 21 samples with successful NGS results was 322x. IG/TR gene rearrangements were also determined by polymerase chain reaction (PCR) analysis and Sanger sequencing using the BIOMED-2 protocol in 14 cases.[1] For the detailed characterization of D-J/V(D)J gene rearrangements the IMGT V-Quest software was used.[5] IG/TR

Table 1. Translocations detected by karyotyping or FISH and the EuroClonality NGS panel.

			NGS Capture results							
				der(IG/TR)				der(partner chromosome)		
Sample	Diagnosis	Karyotyping or FISH results	Break 1	Location	Break 2	Location	Break 1	Location	Break 2	Location
EC20	BCP-ALL	t(X;14)(p22;q32)	IGHJ6	chr14:106329465	*CRLF2*	chrXY:1358468	IGHD3-9	chr14:106370544	*CRLF2*	chrXY:1358472
EC19	BCP-ALL	t(Y;14)	IGHJ5	chr14:106330070	*CRLF2*	chrXY:1333757	IGHD6-19	chr14:106357574	*CRLF2*	chrXY:1333695
EC21	Burkitt lymphoma	t(8;14)	ND	ND	ND	ND	IGHA1	chr14:106177928	*MYC*	chr8:128749373
EC23	Burkitt lymphoma	t(8;14)	IGHG1 and IGHD2-2	chr14: complex	*MYC*	chr8:128746892	IGHG2/3*	chr14	*MYC*	chr8:128746903
EC18	Burkitt lymphoma	t(8;14)	IGHJ4	chr14:106330462	*MYC*	chr8:128746504	ND	ND	ND	ND
EC17	Burkitt lymphoma	t(8;14)	IGHA1	chr14:106177015	*MYC*	chr8:128749477	IGHA1	chr14:106177496	*MYC*	chr8:128749448
EC30	CLL	t(2;14)(p13;32)	IGHM	chr14:106327487	*BCL11A*	chr2:60781261	IGHM	chr14:106327500	*BCL11A*	chr2:60781273
EC31	CLL	t(14;18)(q32;q21)	IGHJ6	chr14:106329459	*BCL2*	chr18:60793474	ND	ND	ND	ND
EC29	CLL	t(14;18)(q32;q21)	IGHJ5	chr14:106330071	*BCL2*	chr18:60769465	IGHD2-15	chr14: 106363819	*BCL2*	chr18:60769465
EC5	DLBCL	t(14;18)	IGHJ5	chr14:106330072	*BCL2*	chr18:60793494	IGHV3-21	chr14:106691673	*BCL2*	chr18:60793498
EC34	SMZL	t(6;14)(p21;q32)	IGHA2*	chr14	*CCND3*	chr6:41942244	ND	ND	ND	ND
EC3	DLBCL	IGH break‡	IGHM	chr14:106326706	*IRF4*	chr6:472682	IGHG4*	chr14	*EXOC2*	chr6
EC25	CLL	t(14;16)(q32;q22)	ND	ND	ND	ND	IGHD2-2	chr14: 106382689	unknown†	chr16:69479932
EC24	CLL	t(1;14)(q32;q32)	IGHM or IGHA1*	chr14	unknown°	chr1:206286226	IGHM or IGHA1*	chr14	unknown†	chr1:206286210
EC33	SMZL	t(5;14)(q13;q32)	IGHM	chr14:106326138	unknown°	chr5:88608990	IGHM	chr14:106326162	unknown†	chr5:88608986
EC11	T-ALL	inv14/t(14;14)	TRDD3	chr14:22918106	*BCL11B*	chr14:99689173	ND	ND	ND	ND
EC12	T-ALL	t(7;10)	TRBJb2.5	chr7:142494805	*TLX1*	chr10:102902431	ND	ND	ND	ND
EC14	T-NHL	inv7	TRGV8	ND	TRBJb2.7	ND	ND	ND	ND	ND

*Alignment equivocal due to sequence homology. °No known lymphoma/leukemia-associated gene in proximity. †demonstrated by FISH IGH break-apart probe. ND: exact breakpoint could not be determined due to insufficient or lack of aligned reads. BCP-ALL: B-cell precursor acute lymphoblastic leukemia; CLL: chronic lymphocytic leukemia; DLBCL: diffuse large B-cell lymphoma; SMZL: splenic marginal zone lymphoma; T-ALL: T-cell acute lymphoblastic leukemia; T-NHL: T-cell non-Hodgkin lymphoma.

translocations had been previously determined by routine FISH, karyotyping and/or Sanger sequencing in the referring laboratories.

In 18 out of 21 samples with a known translocation, identified by either FISH or karyotyping, the breakpoints were identified by the NGS capture panel (Table 1). Of the three samples that failed to produce results, two were fresh frozen samples from lymphomas with degraded and low-quantity DNA (<500 ng total DNA and <1 kb median fragment size by TapeStation analysis) and one was a case of BCP-ALL in which a technical error occurred due to evaporation during the hybridization step.

Of the 18 samples that yielded results, 15 (83%) were from patients with B-cell LPD, of which 11 had well-

Table 2. IGH and IGK rearrangements detected by Sanger sequencing and the EuroClonality NGS panel.

			IGH						IGK			
			Allele 1			Allele 2			Allele 1		Allele 2	
Sample	Diagnosis	Results	IGHV	IGHD	IGHJ	IGHV	IGHD	IGHJ	IGKV	IGKJ	IGKV	IGKJ
EC21	Burkitt lymphoma	Sanger Seq.	3-15	3-22	4	-	3-16	4				
		NGS	3-15	3-22	4	-	3-16	4	1-9	2	D1-13	2
EC23	Burkitt lymphoma	Sanger Seq.	3-23	-	4	-	-	-	4-1	3	-	-
		NGS	3-23	4-23	4	-	-	-	4-1	3	-	-
EC18	Burkitt lymphoma	Sanger Seq.	3-22	6-13	4	-	-	-	4-1	2	intron RSS-KDE*	
		NGS	3-22	6-13	4	-	-	-	4-1	2	intron RSS/VK1-8	
EC30	CLL	Sanger Seq.	4-39	6-25	4	-	-	-				
		NGS	4-39	6-25	4	-	-	-	4-1	1	-	-
EC31	CLL	Sanger Seq.	3-30	2-2	4	-	-	-				
		NGS	3-30	-	4	-	-	-	4-1	3	KDE-VK1-8	
EC29	CLL	Sanger Seq.	5-51	4-17	4	-	-	-				
		NGS	5-51	4-17	4	-	-	-	1-16	4	-	-
EC25	CLL	Sanger Seq.	4-34	5-18	6	-	-	-				
		NGS	4-34	5-18	6	-	2-8	4	4-1	2	-	-
EC24	CLL	Sanger Seq.	4-61	6-19	5	5-51	5-12	4				
		NGS	4-61	6-19	5	5-51	5-12	4	2-30	2	4-1	3
EC33	SMZL	Sanger Seq.	3-23	3-22	3	-	-	-				
		NGS	3-23	3-22	3	-	-	-	1-5	1	-	-
EC32	SMZL	Sanger Seq.	2-5	6-19	2	-	-	-				
		NGS	2-5	6-19	2	-	-	-	2-28	2	-	-
EC34	SMZL	Sanger Seq.	5-51	4-4	6	-	-	-				
		NGS	5-51	4-4	6	-	-	-	-	-	-	-
EC27	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	5-18	6	1-39 4-1†	1 2	1-17	1
EC26	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	2-2	5	1-39 5-2†	1 1	1D-8	2
EC28	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	3-16	4	1-39	4	1-17	1
EC20	BCP-ALL	Sanger Seq.										
		NGS	3-30	2-2	5	-	-	-	KDE-VK2-30 and IGKV7-3-IGKV2-28			
EC22	BCP-ALL	Sanger Seq.										
		NGS	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL
EC19	BCP-ALL	Sanger Seq.										
		NGS	3-30	3-9	5	-	-	-	3-20	1	KDE-IGKJ2 and KDE-IGKV3-20	
EC17	Burkitt lymphoma	Sanger Seq.										
		NGS	3-72	6-13	4	-	3-22	4	1-5	4	KDE-VK3-11	
EC3	DLBCL	Sanger Seq.										
		NGS	1-46	6-6	4	-	-	-	3-7	4	-	-
EC5	DLBCL	Sanger Seq.										
		NGS	3-48	-	4	-	-	-	1-6	3	-	-

Blank cells: No Sanger sequencing data available. *intron RSS-Kde not detectable by NGS as no probes against this region were included in the original panel design. †More than two rearrangements no detected. CLL: chronic lymphocytic leukemia; SMZL: splenic marginal zone lymphoma; BCP-ALL: B-cell precursor acute lymphoblastic leukemia; DLBCL: diffuse large B-cell lymphoma; Sanger Seq.: Sanger sequencing; NGS: next-generation sequencing.

© Ferrata Storti Foundation

known translocations partners: *CRLF2* (2 BCP-ALL), *MYC* (4 Burkitt), *BCL11A* (1 CLL), *BCL2* (2 CLL and 1 DLBCL) and *CCND3* (1 SMZL). In these samples, the exact location of different breakpoints could be delineated from the sequencing reads. In addition, analysis of the reads mapping to the IGH locus showed that in seven cases (47%) the break involved an IGHJ and/or IGHD gene whereas in eight cases (53%) the break lay within the switch regions. The NGS capture approach thus provided additional information about the timing and aberrant cellular processes giving rise to the translocation, either occurring at the time of the D-J or V-DJ recombination or during class-switch recombination.[7] Identification of the specific breakpoint was possible in the majority of cases, although due to the high homology between the different switch regions it was not always possible to specifically map reads unequivocally.

In three of the 15 B-cell LPD samples (20%) only karyo-typing results were available and the location of the breakpoint on the partner chromosome as defined by NGS was mapped to chromosomes 5q14.2 (SMZL), 1q32.1 (CLL) or 16q22.1 (CLL). However, no known lymphoma-associated genes lay in proximity to these breakpoints, highlighting the potential of this approach to identify novel translocation partners and/or mechanisms of disease. Alternatively, these chromosomal alterations may represent "passenger" recombination events, where genomic rearrangements are present in the IGH loci due to a by-product of the processing of double-stranded DNA breaks by the enzymatic machinery, not resulting in an oncogenic translocation.

In the remaining B-cell LPD sample - a case with DLBCL - NGS analysis was able to detect a translocation between IGH and chromosomal location 6p25. The break occurred downstream of *EXOC2* and thus in the vicinity of *IRF4*, a constellation similar to the activating *IRF4* translocations described in germinal-center-derived LPD.[10] FISH analysis of the material using an IGH break-apart probe (Abbott Molecular, Maidenhead, UK) showed the presence of a translocation with no evidence of a *MYC*, *BCL2* or *BCL6* rearrangement, supporting the NGS findings and highlighting the strength of NGS capture approaches to identify novel or uncommon translocation partners and breakpoints in one single analysis.

The three T-cell LPD included in the cohort consisted of two T-ALL and one T-NHL case. The translocations in each of the cases arose from either the TRD, TRG or the TRB locus emphasizing the benefit of the NGS capture panel of being able to interrogate all IG/TR loci in samples from different diseases at the same time. In all three samples the NGS approach identified the same rearrangement as demonstrated previously by karyotyping: t(7;10)(q34;q24) involving TRBJ2-5 and *TLX1*, inv(14)(q11;q32) involving TRDD3 and *BCL11B*[11] and an inversion on chromosome 7 involving TRGV8 (7p14.2) and TRBJ2-7 (7q34).[12]

In all 14 samples with well-characterized V-(D)-J and/or D-J IG gene rearrangements by PCR and Sanger sequencing, NGS was able to detect the same IGH and IGK/IGL gene rearrangements (Table 2), including four CLL cases with hypermutated VDJ rearrangements (90% to 96% homology). All cases included in the study were diagnostic material with tumor infiltration >60% and in the remaining six samples clonal V-(D)-J rearrangement(s) were also identified by NGS, consistent with the clonal nature of the disorder; however, PCR and Sanger sequencing data were not available for these six cases for comparison due to insufficient DNA. This version of the EuroClonality-NGS panel did not include probes for the intron RSS or KDE sequence, explaining why the intron RSS-KDE rearrangement found by Sanger sequencing in a case of Burkitt lymphoma was not detected by the NGS approach. In two CLL cases, a total of three IGK locus gene rearrangements were detected by NGS in each case, raising the possibility of more than one clonal population being present, as previously described in chronic B-cell LPD.[13] Additionally, aberrant clonal rearrangements were seen that were not detected with conventional PCR-based approaches (e.g. IGKV to IGK intron), which warrant further analysis. In the three T-cell LPD cases, Sanger sequencing identified TRDV1-TRDJ1 [T-ALL with inv(14)], a TRBV5-TRBJ1-6 [T-ALL with t(7;10)] and a TRBV5-1-TRBJ2-5 (T-NHL) rearrangement, all of which were also identified in the NGS analysis. In addition, NGS reads demonstrated functional rearrangements in TRG in both T-ALL cases (TRGV11-TRGJ1 and TRGV2-TRGJ1/J2), and a non-functional TRDV1-TRDJ1 rearrangement in one of the T-ALL. No confirmatory Sanger sequencing analysis was feasible for these latter rearrangements due to insufficient DNA.

In summary, this pilot study demonstrates the ability of the EuroClonality-NGS capture approach to simultaneously detect IG/TR translocations and V-(D)-J rearrangements in diagnostic clinical specimens from a range of malignant LPD, including cases with hypermutated VDJ rearrangements. By using capture probes against the V, D and J gene regions of the TR and IG loci (with additional switch regions for IGH), clonal rearrangements and chromosomal translocations arising from these loci can be detected and at the same time the genomic breakpoint sequence involved in the rearrangements and translocations can be identified without the need for additional tests. Other technologies such as target locus amplification have also recently demostrated the ability to detect structural variants and translocations in cancer.[14] An important advantage of these approaches lies in the fact that no prior knowledge of the translocation partner is needed and, therefore, novel or rare chromosomal rearrangements can also be identified by this method, improving their diagnostic value. Sequencing of the V-(D)-J gene rearrangements in any of the IG/TR loci can be used not only to assess clonality and enable a more in-depth analysis of clonal relationships and clonal evolution, but also to identify targets for minimal residual disease monitoring and analysis of the IG/TR repertoire of diverse lymphoid populations. Additional information, for example the somatic hypermutation status of the IGHV-IGHD-IGHJ gene rearrangements, relevant for prognosis in CLL[6] can also be obtained. A new version of this EuroClonality-NGS panel is being designed to include common non-IG/TR translocations as well as genes relevant for diagnosis and prognosis in LPD and a clinical multicenter validation study is now underway within the EuroClonality-NGS consortium.

Dörte Wren,[1] Brian A. Walker,[1,2*] Monika Brüggemann,[3] Mark A. Catherwood,[4] Christiane Pott,[3] Kostas Stamatopoulos,[5] Anton W. Langerak[6†] and David Gonzalez[1,7†] (on behalf of the EuroClonality-NGS consortium)*

[1]The Centre for Molecular Pathology, The Royal Marsden NHS Foundation Trust, London, UK; [2]Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, AR, USA; [3]Second Medical Department, University Hospital Schleswig-Holstein,

© Ferrata Storti Foundation

Campus Kiel, Germany; [4]Department of Haematology, Belfast City Hospital, UK; [5]Institute of Applied Biosciences, CERTH, Thessaloniki, Greece; [6]Department of Immunology, Laboratory for Medical Immunology, Erasmus MC, University Medical Center Rotterdam, the Netherlands and [7]Centre for Cancer Research and Cell Biology, Queen's University Belfast, UK

Correspondence: d.gonzalezdecastro@qub.ac.uk
doi:10.3324/haematol.2016.155424

**These authors contributed equally to this work. [#]These authors contributed equally to this work and share senior authorship of this article.*

Funding: This work was supported by the EuroClonality-NGS consortium. This work was also supported by the NIHR Biomedical Research Centre at the Royal Marsden NHS Foundation Trust and the Institute of Cancer Research.

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12):2257-2317.
2. Langerak AW, Groenen PJ, Bruggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia. 2012;26(10):2159-2171.
3. Bruggemann M, White H, Gaulard P, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies. Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia. 2007;21(2):215-221.
4. Evans PA, Pott C, Groenen PJ, et al. Significantly improved PCR-based clonality testing in B-cell malignancies by use of multiple immunoglobulin gene targets. Report of the BIOMED-2 Concerted Action BHM4-CT98-3936. Leukemia. 2007;21(2):207-214.
5. Bruggemann M, Schrauder A, Raff T, et al. Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. Leukemia. 2010;24(3):521-535.
6. Vardi A, Agathangelidis A, Sutton LA, Ghia P, Rosenquist R, Stamatopoulos K. Immunogenetic studies of chronic lymphocytic leukemia: revelations and speculations about ontogeny and clinical evolution. Cancer Res. 2014;74(16):4211-4216.
7. Walker BA, Wardell CP, Johnson DC, et al. Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. Blood. 2013;121(17):3413-3419.
8. Bouamar H, Abbas S, Lin AP, et al. A capture-sequencing strategy identifies IRF8, EBF1, and APRIL as novel IGH fusion partners in B-cell lymphoma. Blood. 2013;122(5):726-733.
9. Walker BA, Boyle EM, Wardell CP, et al. Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. 2015;33(33):3911-3920.
10. Salaverria I, Philipp C, Oschlies I, et al. Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. Blood. 2011;118(1):139-147.
11. Przybylski GK, Dik WA, Wanzeck J, et al. Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL. Leukemia. 2005;19(2):201-208.
12. Stern MH, Lipkowitz S, Aurias A, Griscelli C, Thomas G, Kirsch IR. Inversion of chromosome 7 in ataxia telangiectasia is generated by a rearrangement between T-cell receptor beta and T-cell receptor gamma genes. Blood. 1989;74(6):2076-2080.
13. Sanchez ML, Almeida J, Gonzalez D, et al. Incidence and clinicobiologic characteristics of leukemic B-cell chronic lymphoproliferative disorders with more than one B-cell clone. Blood. 2003;102(8):2994-3002.
14. de Vree PJ, de Wit E, Yilmaz M, et al. Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. Nat Biotechnol. 2014;32(10):1019-1025.

© Ferrata Storti Foundation

Appendix 12:(Wren et al., 2014)

Translocations and Clonality Detection in Lymphoproliferative Disorders By Capture-Based Next-Generation Sequencing. a Pilot Study By the Euroclonality-NGS Consortium

Dörte Wren, BSc, Brian A Walker, PhD, Monika Brüggemann, Mark Catherwood, Christiane Pott, M..D, PhD, Kostas Stamatopoulos, Anton W Langerak, David Gonzalez, PhD

Check for updates

Blood (2014) 124 (21): 5169.

https://doi.org/10.1182/blood.V124.21.5169.5169

Split-Screen Share Tools

Abstract

Background: Detection and characterization of clonal IG/TR rearrangements and translocations in lymphoproliferative neoplasms provides critical information in the diagnostic pathway in several clinical scenarios and is a valuable tool to address research questions around B and T cells. This includes ascertaining the clonal nature of lymphoid proliferations, characterization of translocations in lymphomas and leukemias, characterization of CDR3 regions for MRD target identification and stereotyping analysis, amongst others. Until now, collecting this information required a combination of different methodologies, such as Gene-scanning/heteroduplex analysis, FISH and Sanger sequencing.

Material and methods: As part of the EuroClonality-NGS consortium, we have designed a capture-based protocol covering the coding V, D and J genes of the IG/TR regions, as well as switch regions in the IGH locus. The assay uses Nimblegen (Roche Molecular Systems) capture baits spanning a total of 180kb and the products are analyzed on a MiSeq Illumina sequencer and MiSeq v3 sequencing chemistry with 2 x 120bp pair-end reads. This design allows the identification of D-J and V(D)J rearrangements as well as chromosomal translocations involving IG/TR genes by Next-generation sequencing (NGS). We piloted this approach for clonality and translocation detection in a cohort of 21 peripheral blood, bone marrow and fresh-frozen samples (3 precursor B-cell acute lymphoblastic leukemias [B-ALL], 4 Burkitt lymphoma [BL], 5 chronic lymphocytic leukemias [CLL], 2 splenic marginal zone lymphomas [SMZL], 2 diffuse large B cell lymphomas [DLBCL], 2 follicular lymphomas [FL], 2 precursor T-cell lymphoblastic leukemias [T-ALL] and 1 T-cell non-Hodgkin lymphoma [T-HNL]) with well-characterized translocations by FISH/Karyotype and/or clonal rearrangements by PCR and Sanger sequencing.

Results: We were able to detect the described IG/TR translocation in 18/21 samples, including translocations into the J, D and switch regions of the IG/TR genes. Three samples failed to produce results, two concerned fresh-frozen lymphomas with low quality DNA, and one concerned a technical error in a B-ALL case. The translocation partner and the breakpoint was identified in 15/18 evaluable cases, including *CRLF2, MYC, BCL11A, BCL2, CCND3, IRF4, BCL11B,* and *TLX1*. In three cases with only karyotyping data available, a translocation involving the IGH locus was identified with no clear leukemia/lymphoma-related genes in the neighboring regions of the reciprocal chromosome, suggesting the potential for new translocation partners and/or mechanism of disease. In all 14 samples with well-characterized D-J/V(D)J rearrangements by PCR and sequencing, NGS was able to detect the same rearrangements. These included IGH and IGK/IGL rearrangements in B-cell proliferations and TRB, TRG and TRD in T-cell proliferations. Additionally, aberrant clonal rearrangements were seen that were not detected with conventional PCR-based approaches (e.g. IGKV to IGK intron).

Conclusion: The EuroClonality NGS IG/TR capture-based approach is a promising tool for the simultaneous detection and characterization of IG/TR translocation and rearrangements in a clinical setting. A formal pan-European validation study is underway within the EuroClonality-NGS consortium.

Disclosures

No relevant conflicts of interest to declare.

Appendix 13:

Euroclonality-NGS DNA Capture Panel for Integrated Analysis of IG/TR Rearrangements, Translocations, Copy Number and Sequence Variation in Lymphoproliferative Disorders

Peter Stewart, PhD, Jana Gazdova, MSc, Nikos Darzentas, PhD, Dorte Wren, PhDFRCPath, Paula Proszek, Grazia Fazio, Simona Songia, Miguel Alcoceba, Maria Eugenia Sarasquete, PhD, Patrick Villarese, PhD, Michèle Y van der Klift, Kim Heezen, Neil McCafferty, MSc, Karol Pal, MSc, Mark Catherwood, PhD, Chang Sik Kim, PhD, Shambhavi Srivastava, PhD, Elizabeth Hodges, PhD, Kostas Stamatopoulos, Wolfram Klapper, Simone Ferrero, MD, Michiel van den Brand, MD, Giovanni Cazzaniga, PhD, Frederic Davi, MD PhD, Ramon Garcia-Sanz, MD PhD, Patricia Groenen, Elizabeth Macintyre, PhDMD, Monika Brüggemann, Christiane Pott, MD PhD, Elisa Genuardi, Anton W Langerak, PhD, David Gonzalez, PhD

Check for updates

Blood (2019) 134 (Supplement_1): 888.

https://doi.org/10.1182/blood-2019-127822

Split-Screen Share Tools

Introduction:

Current diagnostic standards for lymphoproliferative disorders include detection of clonal immunoglobulin (IG) and/or T cell receptor (TR) rearrangements, translocations, copy number alterations (CNA) and somatic mutations. These analyses frequently require a series of separate tests such as clonality PCR, fluorescence in situ hybridisation and/or immunohistochemistry, MLPA or SNParrays and sequencing. The EuroClonality-NGS DNA capture (EuroClonality-NDC) panel, developed by the EuroClonality-NGS Working Group, was designed to characterise all these alterations by capturing variable, diversity and joining IG and TR genes along with additional clinically relevant genes for CNA and mutation analysis.

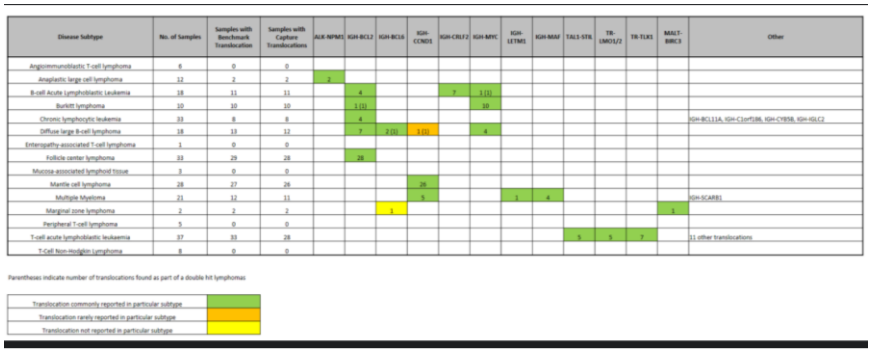
Methods:

Well characterised B and T cell lines (n=14) representing a diverse repertoire of IG/TR rearrangements were used as a proficiency assessment to ensure 7 testing EuroClonality centres achieved optimal sequencing performance using the EuroClonality-NDC optimised and standardised protocol. A set of 56 IG/TR rearrangements across the 14 cell lines were compiled based on detection by Sanger, amplicon-NGS and capture-NGS sequencing technologies. For clinical validation of the NGS panel, clinical samples representing both B and T cell malignancies (n=280), with ≥ 5% tumour infiltration were collected from 10 European laboratories, with 88 (31%) being formalin fixed paraffin-embedded samples. Samples were distributed to the 7 centres for library preparation, hybridisation with the EuroClonality-NDC panel and sequencing on a NextSeq 500, using the EuroClonality-NDC standard protocol. Sequencing data were analysed using a customised version of ARResT/Interrogate, with independent review of the results by 2 centres. All cases exhibiting discordance between the benchmark and capture NGS results were submitted to an internal review committee comprising members of all participating centres.

Results:

All 7 testing centres detected all 56 rearrangements of the proficiency assessment and continued through to the validation phase. A total of 10/280 (3.5%) samples were removed from the validation analysis due to NGS failures (n=1), tumour infiltration < 5% (n=7), and sample misidentification (n=2). The EuroClonality-NDC panel detected B cell clonality (i.e. detection of at least one clonal rearrangement at IGH, IGK or IGL loci) in 189/197 (96%) B cell malignancies. Seven of the 8 discordant cases were post-germinal centre malignancies exhibiting Ig somatic hypermutation. The EuroClonality-NDC panel detected T cell clonality (i.e. detection of at least one clonal rearrangement at TRA, TRB, TRD or TRG loci) in 70/73 (96%) T cell malignancies. In all 3 discordant cases analysis of benchmark PCR data was not able to detect clonality at any TR loci. Next, we examined whether the EuroClonality-NDC panel could detect clonality at each of the individual loci, resulting in sensitivity values of 95% or higher for all IG/TR loci, with the exception of those where limited benchmark data were available, i.e. IGL (n=3) and TRA (n=7). The specificity of the panel was assessed on benign reactive lesions (n=21) that did not contain clonal IG/TR rearrangements based on BIOMED-2/EuroClonality PCR results; no clonality was observed by EuroClonality-NDC in any of the 21 cases. Limit of detection (LOD) assessment to detect IG/TR rearrangements was performed using cell line blends with each of the 7 centres receiving blended cell lines diluted to 10%, 5.0%, 2.5% and 1.25%. Across all 7 centres the overall detection rate was 100%, 94.1%, 76.5% and 32.4% respectively, giving an overall LOD of 5%. Sufficient data were available in 239 samples for the analysis of translocations. The correct translocation was detected in 137 out of 145 cases, resulting in a sensitivity of 95%. Table 1 shows how translocations identified by the EuroClonality-NDC protocol were restricted to disease subtypes known to harbour those types of translocations. Analysis of CNA and somatic mutations in all samples is underway and will be presented at the meeting.

Conclusions:



Disease Subtype	No. of Samples	Samples with Benchmark Translocation	Samples with Capture Translocations	ALK-NPM1	IGH-BCL2	IGH-BCL6	IGH-CCND1	IGH-CRLF2	IGH-MYC	IGH-LETM1	IGH-MAF	TAL1-STIL	TR-LMO1/2	TR-TLX1	MALT-BIRC3	Other
Angioimmunoblastic T-cell lymphoma	6	0	0													
Anaplastic large cell lymphoma	12	2	2	2												
B-cell Acute Lymphoblastic Leukemia	18	11	11		4			7	1 (1)							
Burkitt lymphoma	10	10	10		1 (1)				10							
Chronic lymphocytic leukemia	33	8	8		4											IGH-BCL11A, IGH-C1orf186, IGH-CYB5B, IGH-IGLC2
Diffuse large B-cell lymphoma	18	13	12		7	2 (1)	1 (1)		4							
Enteropathy-associated T-cell lymphoma	1	0	0													
Follicle center lymphoma	33	29	28		28											
Mucosa-associated lymphoid tissue	3	0	0													
Mantle cell lymphoma	28	27	26				26									
Multiple Myeloma	21	12	11				5			1	4					IGH-SCARB1
Marginal zone lymphoma	2	2	2			1									1	
Peripheral T-cell lymphoma	5	0	0													
T-cell acute lymphoblastic leukaemia	37	33	28									5	5	7		11 other translocations
T-Cell Non-Hodgkin Lymphoma	8	0	0													

Parentheses indicate number of translocations found as part of a double hit lymphomas

Translocation commonly reported in particular subtype	
Translocation rarely reported in particular subtype	
Translocation not reported in particular subtype	

The EuroClonality-NDC panel, with an optimised laboratory protocol and bioinformatics pipeline, detects IG and TR rearrangements and translocations with high sensitivity and specificity with a LOD ≤ 5% and provides a single end-to-end workflow for the simultaneous detection of IG/TR rearrangements, translocations, CNA and sequence variants.