Molecular mechanisms of follicular lymphoma and its transformation

Wrench, David John

Molecular mechanisms of follicular lymphoma and its transformation

David John Wrench

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

Centre for Haemato-Oncology
Barts Cancer Institute, London, UK

2011

Abstract

Follicular lymphoma (FL) is the second most common form of non-Hodgkin lymphoma and at least a third of cases undergo aggressive transformation (t-FL), most frequently to diffuse large B-cell lymphoma. This study examined the role of germline and acquired genetic changes in FL and t-FL to determine prognostically significant events and chart the evolution of transformation.

Assessment of germline polymorphic variation in over 200 FL cases demonstrated two SNPs in the HLA region of chromosome 6p (rs10484561 and rs6457327) associate with FL risk in the UK and identified that rs6457327 predicts both time to and risk of transformation independently of clinical variables, including the FLIPI. Mutation and expression studies of the single known gene in linkage disequilibrium with rs6457327 (*C6orf15*) suggest an alternative mechanism is responsible for this transformation association.

DNA copy number and mutational analysis of FL and t-FL samples then revealed a high prevalence of *TNFSRF14* and *EZH2* mutations at transformation accompanied by frequent loss and gain of their genomic locations on chromosome 1p and 7q, respectively. In a search for co-operating genetic events, genome-wide profiling identified recurrent losses and gains ranging from 4 kb to 60 Mb with gain 2p16.1-p15 (including *REL*) predictive of worse survival in FL that transforms.

In >50% of transformed cases, FL DNA contained either copy number aberrations or mutations that were absent from subsequent t-FL. This suggested FL and t-FL might develop non-sequentially from a common cell of origin. To further explore the evolution of FL and t-FL, *IGH-V* somatic-hypermutation (SHM) analysis was performed in sequential FL / t-FL samples. t-FL clones were detected in FL samples taken many months prior to clinical presentation of transformation and, furthermore, the predicted SHM patterns of putative precursors were detected in both FL and t-FL samples indicating that a (long-lived) common progenitor cell could indeed give rise to both FL and subsequent t-FL by divergent clonal evolution.

Statement of Work Undertaken

The work described in this thesis was performed in my role as a Clinical Research Fellow in the Centre for Haemato-Oncology (formerly Medical Oncology) at Barts Cancer Institute, London.

Sample storage was performed by Ms Sameena Iqbal from the Tissue Storage Unit, as was the majority of DNA extraction, for the ongoing management of the tissue bank at this centre. All allelic discrimination assays and RNA RT-PCR were completed by myself for the studies in Chapter 3. Statistical input for these studies was provided by Ms Rachel Waters and Ms Pamela Leighton from the CRUK Centre for Medical Statistics, University of Oxford and by Professor Christine Skibola from the University of California, Berkley, US.

For the work detailed in Chapter 4, DNA micro-array assays were performed by myself. Hybridisation, overnight staining and scanning of the assay products to arrays was done by Ms Tracy Chaplin as part of a core service provided within the institute. I performed the subsequent micro-array analysis and statistical tests.

The study examining evolution of FL and its transformation (Chapter 5) was undertaken with Dr Emanuela Carlotti. We both performed the various PCR, cloning, homo-hetero-duplex and sequencing experiments and I performed the quantative PCR experiments. The writing of this thesis was my own work.

Acknowledgements

Firstly, I would like to express my gratitude to Dr Jude Fitzgibbon and Professor T. Andrew Lister, my supervisors, who have been a great source of support, guidance, humour and inspiration throughout my PhD project. It has been a privilege to work so closely with them during my time in this centre.

I am indebted to many people for help and advice without which this project would not have been possible, in particular to Professor John Gribben for critical evaluation of my work. For smoothing my transition from clinician to fledgling scientist in the early days of the project I give special thanks to Emanuela Carlotti, Sameena Iqbal and Karin Summers. Throughout the course of the project I have since gained much knowledge and received valued advice through many conversations with Emanuela, and Sameena continues to provide a commendable service regarding tissue samples and the running of the tissue bank in this centre.

All staff in the Centre for Haemato-Oncology have contributed to my time being an enjoyable one and I must specifically thank Manu Gupta for his practical advice regarding the micro-array study, so to Professor Bryan Young who developed software used in analysis and provided guidance on this work.

Robust patient follow-up and an extensive clinical database are a particular advantage of this centre's affiliation with the Department of Haemato-Oncology at St Bartholomew's Hospital. In this regard, particular thanks are due to Ms Janet Matthews for her help in provision of clinical data for correlation with experimental findings. Furthermore, the diligence and care of clinical staff and willingness of patients to provide samples and data for research underpin the excellent research framework in this institution and I express my sincere gratitude to all those involved.

For the work in Chapter 3, the Welcome Trust Case Control Consortium provided access to their data and Professor Christine Skibola from the University of California provided advice and statistical guidance while Dr Jean-Baptiste Cazier helped with statistical tests and I convey my thanks to each of them.

Finally, I thank Cancer Research UK for generously funding this project.

Publications

Wrench D, Leighton P, Skibola CF, Conde L, Cazier JB, Matthews J, Iqbal S, Carlotti E, Bödör C, Montoto S, Calaminici M, Gribben JG, Lister TA, Fitzgibbon J. SNP rs6457327 in the HLA region on chromosome 6p is predictive of the transformation of follicular lymphoma. Blood. 2011 Mar 17;117(11):3147-50. PMID: 21233317

Bödör C, O'Riain C, **Wrench D**, Matthews J, Iyengar S, Tayyib H, Calaminici M, Clear A, Iqbal S, Quentmeier H, Drexler HG, Montoto S, Lister TA, Gribben JG, Matolcsy A, Fitzgibbon J. EZH2 Y641 mutations in follicular lymphoma. Leukemia. 2011, Jan 14. [Epub ahead of print.] PMID: 21233829

Wrench D, Montoto S, Fitzgibbon J. Molecular signatures in the diagnosis and management of follicular lymphoma. Curr Opin Hematol. 2010 Jul;17(4):333-40. PMID: 20571396

Carlotti E, **Wrench D**, Matthews J, Iqbal S, Davies A, Norton A, Hart J, Lai R, Montoto S, Gribben JG, Lister TA, Fitzgibbon J. Transformation of follicular lymphoma to diffuse large B cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone. Blood. 2009 Apr 9;113(15):3553-7. PMID: 19202129

Wrench D, Waters R, Carlotti E, Iqbal S, Matthews J, Calaminici M, Gribben J, Lister TA, Fitzgibbon J. Clinical relevance of MDM2 SNP 309 and TP53 Arg72Pro in follicular lymphoma. Haematologica. 2009 Jan;94(1):148-50. PMID: 19029147

To Bella and Poppy with love

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Term
AD	allelic discrimination
Ag	Antigen
AID	activation-induced cytidine deaminase
ALL	acute lymphoblastic lymphoma
ANOVA	analysis of variance
ASO	allele specific oligonucleotide
aUPD	acquired uniparental disomy
BCL2	B-cell CLL/lymphoma 2
BL	Burkitt lymphoma
BM	bone marrow
BSA	bovine serum albumin
CD20	cluster differentiation 20
cDNA	complementary DNA
CDR	complementarity determining region
CI	confidence interval
CLL	chronic lymphocytic leukaemia
CN	copy number
CNA	copy number aberration
CNV	copy number variation
CPC	common progenitor cell
CREBBP	cAMP response element-binding, binding protein
ddH_2O	double distilled water
DLBCL	diffuse large B cell lymphoma
DNA	deoxyribonucleic acid
dNTPs	deoxy-nucleotide tri-phosphates
EBV	Epstein-Barr virus
EtOH	ethanol
ExoSap	exonuclease - shrimp alkaline phosphatase
EZH2	enhancer of zeste homolog 2
FL	follicular lymphoma
FLIPI	Follicular Lymphoma International Prognostic Index
FR	framework region
GC	germinal centre
GOLF	Genome Orientated Laboratory File
GWAS	genome-wide association study
HBD	hydroxybutyrate dehydrogenase
HH	homo-hetero-duplex
HLA	human leucocyte antigen
HR	hazard ratio
HWE	Hardy-Weinberg equilibrium
Ig	immunoglobulin
IGH	immunoglobulin heavy locus
IGK	immunoglobulin kappa locus
IGL	immunoglobulin lambda locus
ILI	Italian Lymphoma Intergroup
IPI	International Prognostic Index
kb	kilobase
LD	linkage disequilibrium
LDH	lactate dehydrogenase
LN	lymph node
LOH	loss of heterozygosity
MAF	minor allele frequency
MAPD	median absolute pairwise difference
Mb	megabase
MBR	major breakpoint region
mcr	minor cluster region

MCR	minimum common region
MDM2	Mdm2 p53 binding protein homolog (mouse)
MNC	mononuclear cells
MYC	v-myc myelocytomatosis viral oncogene homolog
MYCBP2	MYC binding protein 2
MZL	marginal zone lymphoma
NHL	non-Hodgkin lymphoma
NTC	no template control
PB	peripheral blood
PCR	polymerase chain reaction
PRC2	polycomb repressive complex 2
PS	performance status
OR	odds ratio
OS	overall survival
QC	quality control
qPCR	quantitative PCR
RNA	ribonucleic acid
RNase	ribonuclease
RT-PCR	real time-polymerase chain reaction
SEER	Surveillance, Epidemiology and End Results program
SHM	somatic hypermutation
SLL	small lymphocytic lymphoma
SNP	single nucleotide polymorphism
TBE	tris-borate-EDTA
TE	tris-EDTA
TEMED	tetramethylene diamine
t-FL	transformed follicular lymphoma
TNFRSF14	tumour necrosis family receptor superfamily, member 14
TP53	tumour protein p53
tris	tris(hydroxymethyl)aminomethane
TTT	time to transformation
VH	immunoglobulin heavy chain variable region
WHO	World Health Organisation
WTCCC BC58	Welcome Trust Case Control Consortium 1958 birth cohort

Chapter 1 Introduction

The hallmark of cancer is uncontrolled clonal proliferation of cells which occurs following genetic changes that disrupt normal cell cycle regulation. A single genetic "hit" is typically insufficient to cause cancer and this led to the two-hit hypothesis established by Knudson four decades ago (Knudson 1971). Subsequently, a sequence of cancer development characterised by the accumulation of multiple genetic events, including chromosomal changes and gene mutations, was proposed (Fearon and Vogelstein 1990). These models reflect the observation that "Cancer is, in essence, a genetic disease" (Vogelstein and Kinzler 2004), a statement which highlights the fundamental importance of the genetic events that contribute to tumour development and progression.

Around one in three people develop cancer and one in four people die from it in the UK (Source: Office for National Statistics, 16 April 2010 available at website http://www.statistics.gov.uk/CCI/nugget.asp?ID=915; licensed under the Open Government Licence v.1.0). Cancer is therefore the most common form of genetic disease and insights into the underlying genetic mechanisms are an important aspect of understanding its behaviour. These have been made possible through the steady improvement in resolution of investigative techniques that have developed from the early identification of chromosomes by staining with giemsa giving characteristic signature patterns to each chromosome (so called "G banding"). This enabled the formation of a karyotype which could be examined for abnormalities. Subsequently, refinements to karyotyping have included the use of probes for specific regions (as in fluorescent in-situ hybridisation; FISH) to detect chromosomal rearrangements or, more recently, to compare target and control DNA by hybridisation to metaphase chromosomes (comparative genomic hybridisation; CGH) enabling more focal regions of copy number change to be determined.

The advent of DNA micro-arrays containing many thousands (and now millions) of probes has subsequently enabled a reduction in the resolution of chromosomal study to less than 10 kb in size. Furthermore, the current era of genome-wide sequencing that has developed following the completion of the human genome project (Lander *et al.* 2001) permits pan-genomic inspection at the single base pair level. This is a

significant advance that should reveal much information about the genetics of cancer over the coming years.

Haematological malignancies have been at the forefront of genetic studies in cancer, with leukaemias particularly suitable due to the presence of cells in the peripheral blood which are, therefore, easily accessed and can be applied to several investigative platforms. Indeed, both cytogenetic and molecular features of various leukaemias are now core aspects guiding their treatment.

However, in the most common form of haematological malignancy, non-Hodgkin lymphoma (NHL), although chromosomal translocations are well characterised and aid diagnosis they do not typically influence subsequent management and considerably less is known as to the relevance of other genetic changes in disease. As many forms of NHL remain incurable and it has shown an increased incidence in recent decades, there is a pressing need for accurate prognostic and predictive models which can guide the development and application of therapy. Indeed, several clinical tools can group cases in to prognostic groups but these still do not influence routine patient management.

Consequently, as has occurred in the leukaemias, characterising the critical genetic events in NHL can be expected to better risk stratify cases both for upfront management as well as treatment later in the course of disease. This is particularly important for a prevalent form of NHL, follicular lymphoma (FL), which shows a spectrum in both its natural history and clinical course, despite similar upfront clinical and histological parameters in the majority of cases.

Studies in FL to date show that it is accompanied by frequent and diverse genetic changes but the specific events driving its progression and outcome are largely uncharacterised. The degree of genetic instability in FL may explain the range of clinical outcomes observed in this disease, however, the challenge is identifying the critical changes which could provide suitable therapeutic targets and help improve its outcome. In this regard, the examination of both heritable and acquired genetic variants in this thesis provides novel insights in to FL and its progression that help explain previous reports and provide an important model of FL evolution which should guide future studies.

1.1 Follicular lymphoma

FL was first described over eighty years ago (Brill *et al.* 1925; Symmers 1927) and has most recently been defined as a "neoplasm composed of follicle centre ... B-cells ... which usually has at least a partially follicular pattern" in the World Health Organisation (WHO) Classification of Tumours (Swerdlow 2008).

1.1.1 Histopathology

This current classification follows early efforts using histology alone (Gall and Mallory 1942) that distinguished FL from reactive hyperplasia and termed it a nodular lymphoma (Hicks *et al.* 1956). This was then refined by the incorporation of immunological cell type (Lukes and Collins 1974) and recognition that FL arose from the germinal centre (Lennert *et al.* 1975), a structure found in the follicles of secondary lymphoid organs as discussed by MacLennan (MacLennan 1994). Subsequent attempts to improve the clinical utility of NHL classification commenced with the working formulation ("The Non-Hodgkin's Lymphoma Pathologic Classification Project" 1982) before a large international collaboration proposed the Revised European-American classification of lymphoid neoplasms (REAL) (Harris *et al.* 1994). This and the earlier WHO classification (Jaffe *et al.* 2001) incorporated histological, immunological, genetic and clinical features of disease, an approach which forms the current diagnostic standard (Swerdlow 2008).

As FL is a malignant expansion of the germinal centre B cells from lymphoid follicles, it includes both cell types normally found in these structures, namely centrocytes and centroblasts. Centrocytes are small cells often with cleaved nuclei and scant cytoplasm and form the bulk of the tumour. Centroblasts are larger and show prominent nucleoli. The varying proportions of centroblasts in FL tumour specimens, as initially proposed by Berard and colleagues (Glick *et al.* 1981; Mann and Berard 1983), has led to three different grades of tumour being described (Jaffe *et al.* 2001), summarised as follows:

Grade*	Criteria
1	≤ 5 centroblasts per hpf
2	6 - 15 centroblasts per hpf
3	> 15 centroblasts per hpf
3A	Centrocytes present
3B	Solid sheets of centroblasts

*Distinction between 1 and 2 no longer required as there are no important clinical differences between them. hpf indicates high power field of 0.159mm^2 (which is the view obtained using a ×40 objective lens)

Tumour grading is important as grade 3 may progress more readily than grade 1 or 2 (Bartlett *et al.* 1994). Furthermore, within grade 3 solid sheets of centroblasts distinguish a 3B subtype (Jaffe *et al.* 2001), which has molecular features analogous to the most common aggressive form of B-NHL, diffuse large B cell lymphoma (DLBCL) (Ott *et al.* 2002; Bosga-Bouwer *et al.* 2003; Karube *et al.* 2007) and frequently contains areas suggestive of DLBCL (Katzenberger *et al.* 2004). Consequently, FL grade 3B has been managed in a similar manner to *de novo* DLBCL. Indeed, the presence of any regions of DLBCL in association with FL now requires a formal diagnosis of DLBCL (with accompanying FL). Nevertheless, the majority of FL cases (at least 80%) present as grade 1 or 2 disease. Furthermore, grade 3B constitutes only around one quarter of the remaining cases (Hans *et al.* 2003) and it seems probable that this proportion will reduce as cases are formally diagnosed as DLBCL in light of the recent WHO classification update.

However, as identified by Relander and colleagues in a recent review (Relander *et al.* 2010), there have been persistent challenges in accurately reproducing FL tumour grading between centres. Indeed, despite observed histopathological and molecular differences between grades 3A and 3B, one recent report suggests that outcomes of these two subtypes are comparable (Shustik *et al.* 2010). Consequently, the role of grade in determining FL prognosis and management remains unclear.

Immuno-histochemical staining is an important aspect of diagnosis, particularly in the minority of cases which do not have a clear follicular architectural pattern on morphological examination, with FL cells expressing the B cell markers CD20, CD22, CD79a and immunoglobulin. The majority of cases express CD10 while CD5 and cyclinD1 are not expressed which can, at least in part, help distinguish them from other B-NHL while the expression BCL2 (Lai *et al.* 1998) and BCL6 (Cattoretti *et al.* 1995) are characteristic, but not specific, findings in FL.

1.1.2 Presentation demographics

FL is the most common indolent NHL and the second most common form of all NHL after DLBCL, as illustrated in **Figure 1.1** ("The Non-Hodgkin's Lymphoma Classification Project" 1997; Swerdlow 2008). The median survival of around one decade reflects the relapsing-remitting nature that is typically observed over a disease course of a number of years (Johnson *et al.* 1995). As NHL constitutes approximately 4% of all cancers diagnosed annually in the UK (Source: Office for National Statistics "Cancer and mortality in the United Kingdom 2005-2007" Statistical Bulletin 24 August 2010; licensed under the Open Government Licence v.1.0.), FL forms only around 1% of new cancer but due to its prolonged clinical course has a comparatively high prevalence (approximately 20 to 25 cases per 100000 population) which represents an important overall clinical burden.

With a median age at presentation in the 6th decade (Friedberg *et al.* 2009) FL is predominantly a disease of older adults, although paediatric cases are occasionally reported (Swerdlow 2004). It has a higher incidence in western countries (Anderson *et al.* 1998), has a slight male preponderance and occurs more commonly in whites (Groves *et al.* 2000). Importantly, Groves and colleagues also demonstrated, using data from the Surveillance, Epidemiology and End Results (SEER) program, that the incidence of FL increased significantly over the period 1978 to 1995 (Groves *et al.* 2000) and a more recent study indicates a similar trend for FL may have continued in to this century, with a rate of increase in incidence of 1.8% per year from 1992 to 2001 (Morton *et al.* 2006). Indeed, in a subsequent SEER analysis extended to 2007 (available at website http://seer.cancer.gov/csr/1975_2007/), the incidence of all NHL continues to increase (Altekruse *et al.* 2009).

1.1.3 Outcome

The majority of patients will experience disease relapses and remissions over the course of a number of years, with remission duration after therapy tending to shorten after each disease recurrence (Johnson *et al.* 1995). In contrast, some patients may never require treatment and some can even experience spontaneous disease remission (Horning and Rosenberg 1984) whereas others succumb early to rapidly progressive

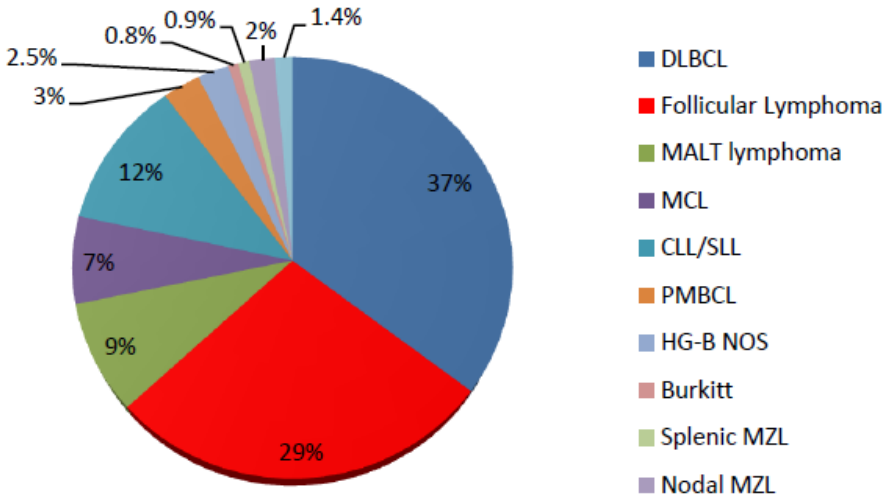


Figure 1.1 Subtypes of B-cell non-Hodgkin lymphoma

Illustrated are differing lymphoma subtypes and the respective percentages they form of B-NHL. Adapted from the WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (Swerdlow 2008). As T-NHL make up in the region of 10% of NHL, FL constitutes around one quarter of *all* NHL.
DLBCL indicates diffuse large B cell lymphoma; MALT, mucosal associated lymphoid tissue; MCL, mantle cell lymphoma; CLL/SLL, chronic lymphocytic leukaemia/small lymphocytic lymphoma; PMBCL, primary mediastinal B cell lymphoma; HG-B NOS, high grade B cell lymphoma not otherwise specified; and MZL, marginal zone lymphoma.

disease that is resistant to therapy or develop aggressive transformation which has a dismal prognosis (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008).

1.1.4 Clinical aspects

Painless lymphadenopathy is the typical presentation of FL and accompanying symptoms are uncommon. Other involved sites may include the bone marrow in around half of cases (Federico *et al.* 2000; Federico *et al.* 2009) and, less frequently, the spleen and tonsillar tissue. Disease may sometimes be detectable in peripheral blood but can present at any extra-nodal site, with isolated primary skin or intestinal involvement identifying specific variants with particularly good prognosis (Shia *et al.* 2002; Willemze *et al.* 2005). The majority of cases present with advanced disease, as Stage 3 or 4 when applying the Ann Arbor Classification (Gallagher *et al.* 1986; The Non-Hodgkin's Lymphoma Classification Project" 1997), which is associated with a worse prognosis compared to disease of limited extent (Stages 1 or 2) (Bastion *et al.* 1997; Solal-Celigny *et al.* 2004). Other indicators of high disease burden including a large (>5cm) peripheral lymph node, extensive bone marrow involvement and two or more extra-nodal sites of disease also associate with worse outcome (Romaguera *et al.* 1991). Age and male sex are further negative predictors of outcome (Leonard *et al.* 1991; Decaudin *et al.* 1999) as are elevated serum LDH, ESR and β_2-microglobulin as well as anaemia and hypo-albuminaemia (Leonard *et al.* 1991; Bastion *et al.* 1997) while a higher absolute peripheral blood lymphocyte count may predict better overall survival (Siddiqui *et al.* 2006).

As cases with apparently similar morphology at initial presentation can have markedly different disease courses, the identification of these outcome associations has led to the development of clinical prognostic indices that can risk stratify patient groups. These indices typically include several factors reflecting patient fitness, disease bulk and disease aggressiveness. Each factor contributes to an overall score which is used to risk stratify patients with FL into prognostic groups. Four large studies that have generated prognostic tools are summarised in **Table 1.1**. The first of these was the International Prognostic Index (IPI) which was initially developed for prognostication of aggressive NHL ("The International Non-Hodgkin's Lymphoma Prognostic Factors Project" 1993) and was subsequently shown to define risk groups

Table 1.1 Comparison of clinical prognostic indices in follicular lymphoma

Study group	IPI			ILI			FLIPI			FLIPI-2			
Prognostic factors[§]	*Age*	> 60		*Age*	> 60		*Age*	≥ 60		*Age*	> 60		
	Stage	3 or 4		*Sex*	Male		*Stage*	3 or 4		*BMI*	present		
	LDH (iu)	> ULN		*LDH (iu)*	> ULN		*LDH (iu)*	> ULN		*β_2M*	> ULN		
	PS	2 to 4		*B sx*	Present		*Hb (g/dl)*	< 12		*Hb (g / dl)*	< 12		
	Extranodal	≥ 2		*Extra-nodal*	≥ 2		*LN sites*	≥ 5		*LN size (cm)*	> 6		
				ESR (mm/hr)	≥ 30								
		Cases	10 yr OS		Cases	10 yr OS		Cases	10 yr OS		Cases	5 yr OS	5 yr PFS
Risk groups*	*Low (0 or 1)*	36	75	*Low (0 or 1)*	64	65	*Low (0 or 1)*	36	71	*Low (0)*	18	98	79
	Low-int (2)	32	47	*Int (2)*	23	54	*Int (2)*	37	51	*Int (1 or 2)*	62	88	51
	High-int (3)	19	55	*High (3 to 5)*	13	11	*High (3 to 5)*	27	36	*High (3 to 5)*	20	77	20
	High (4 or 5)	13	0										
Study type	Retrospective single-centre			Retrospective multi-centre			Retrospective multi-centre			Prospective multi-centre			

[§]The values for each factor that give a score of 1 in risk group stratification are indicated for the respective prognostic indices. The maximum score for each study was five (note: although six factors were included in the ILI study, no case scored more than five).

*The risk group categories are indicated for each of the four clinical prognostic indices. The respective cumulative scores established from the prognostic factors are given in parentheses. There are four groups for the IPI and three each for the remaining three indices. Figures indicate the percentages of cases included in each risk group and the percentage survival (OS or PFS at five or ten years, as indicated).

LDH indicates lactate dehydrogenase; iu, international units; PS performance status as per the Eastern Cooperative Oncology Group (ECOG) criteria (Oken *et al.* 1982); int, intermediate; ULN, upper limit of normal; sx, symptoms; extra-nodal, number of extra-nodal sites involved by disease; ESR, erythrocyte sedimentation rate; Hb, haemoglobin, LN sites, number of lymph node sites involved by disease; BMI, bone marrow involvement by disease; β_2M, beta-2 microglobulin; LN size, longest diameter of the largest involved lymph node; and int, intermediate.

in FL (Lopez-Guillermo *et al.* 1994). However, only a relatively small proportion of FL cases (13%) are classified as high risk using this model, which has limited its clinical application. An FL specific prognostic index which identified three risk groups was developed over a decade ago by the Italian Lymphoma Intergroup (ILI) (Federico *et al.* 2000) with the poor risk group constituting 13% of cases. Following this, a large international multi-centre study (Solal-Celigny *et al.* 2004) developed the Follicular Lymphoma International Prognostic Index (FLIPI). This includes five clinical parameters and classifies cases in to low (36% of cases), intermediate (37%) and high (27%) risk groups with overall survival at ten years of 71%, 51%, and 36%, respectively. Recently, the FLIPI-2 has been reported which used prospectively acquired data to develop a model which predicted both overall survival and progression free survival (Federico *et al.* 2009). However, although all these prognostic tools robustly stratify patients in to survival risk groups and can be easily applied in the clinical setting, they have not yet led to risk adapted treatment. Moreover, there is heterogeneity of outcome within the risk groups themselves which limits their predictive power at the level of individual cases. So, while they are valuable adjuncts to prospective clinical studies, these clinical prognostic indices are yet to impact on routine patient management.

1.1.5 Management

The management of FL is diverse reflecting the heterogeneity of clinical behaviour that can be observed over the course of the illness. However, a particularly important determinant of initial management is the presenting disease stage.

1.1.6 Limited stage disease

Although the minority of cases present with limited (stage 1 or 2) disease, these cases may enter prolonged remission following local radiotherapy (Richards *et al.* 1989; MacManus and Hoppe 1996). This has led to the recommendation that local radiotherapy should be offered to patients with limited stage FL at diagnosis (Dreyling 2010). Intriguingly, some cases with stage 1 disease may experience long remissions following excision of the single disease site alone (Soubeyran *et al.*

1996), giving a "surgical remission", but this approach is not considered standard of care.

1.1.7 Advanced stage disease

As the majority of FL cases have stage 3 or 4 disease at diagnosis they are not suitable for treatment with radiotherapy. At present there remains no recommended curative option for these cases but, following many years of uncertainty regarding how best to manage advanced stage FL, recent advances have brought at least some consensus as to initial treatment. The need for treatment is largely determined by the presence of B symptoms, organ compromise, haematopoietic impairment, bulky disease, ascites / pleural effusion or rapid lymphoma progression and, in the absence of these features, management by observation alone may be appropriate. Such a "watch and wait" approach has long been employed (Portlock and Rosenberg 1979; O'Brien *et al.* 1991), may not adversely influence subsequent disease events (Horning and Rosenberg 1984) and prospective randomised studies have demonstrated similar outcomes compared to upfront therapy (Young *et al.* 1988; Brice *et al.* 1997; Ardeshna *et al.* 2003). However, this approach is being challenged following the recent demonstration that single agent rituximab (an anti-CD20 monoclonal antibody) given upfront at initial presentation may be beneficial in those patients who would otherwise be candidates for watch and wait (Ardeshna *et al.* 2010).

Following reports beginning more than three decades ago into the activity of both single agent (Lister *et al.* 1978) and combination chemotherapy (Portlock and Rosenberg 1976) regimes in FL, the advent of rituximab immunotherapy has significantly improved the treatment of patients with advanced stage disease. Rituximab itself has activity as a single agent in FL (McLaughlin *et al.* 1998) but it is in combination with chemotherapy that it has shown most promise. Following the first report of its use with cyclophosphamide, doxorubicin, vincristine and prednisolone (CHOP) in FL (Czuczman *et al.* 1999), four randomised studies have demonstrated significant survival benefits of rituximab when added to combination chemotherapy as initial treatment (Hiddemann *et al.* 2005; Marcus *et al.* 2005; Herold *et al.* 2007; Marcus *et al.* 2008; Salles *et al.* 2008) and a benefit on overall

survival is supported by a systematic meta-analysis (Schulz *et al.* 2007). This data has led to Rituximab given in combination with chemotherapy being recommended as first line treatment in advanced stage previously untreated FL that requires therapy ("NICE Technology Appraisal Guidance 137" 2008).

The optimal management of relapsed disease is less clear. However, rituximab has a role either in combination with chemotherapy (Forstpointner *et al.* 2006; van Oers *et al.* 2006; van Oers *et al.* 2010) or as a single agent (Ghielmini *et al.* 2004; Martinelli *et al.* 2010) and non-cross resistant chemotherapy is considered, particularly if the disease recurrence occurs early (< 6 months) following prior treatment.

Augmentation of remission has been attempted by maintenance therapy with either chemotherapy (Steward *et al.* 1988) or interferon but these have been limited by significant risks of myelosuppression or poor tolerability, although interferon may improve both progression free and overall survival particularly in the pre-rituximab era (Rohatiner *et al.* 2007; Baldo *et al.* 2010). However, maintenance therapy with rituximab following induction of remission appears to be beneficial and well tolerated after both treatment with combination chemotherapy (Forstpointner *et al.* 2006; van Oers *et al.* 2006; van Oers *et al.* 2010) or rituximab monotherapy (Hainsworth *et al.* 2002) following FL relapse as well as after rituximab monotherapy as first line therapy after diagnosis (Ghielmini *et al.* 2004; Martinelli *et al.* 2010). Furthermore, rituximab maintenance is beneficial in previously untreated patients with FL following combination chemotherapy induction (Hochster *et al.* 2009). Importantly, a recent large international multi-centre study demonstrates that following remission induction with rituximab and combination chemotherapy for previously untreated advanced stage FL, subsequent rituximab maintenance is beneficial (Salles *et al.* 2011) suggesting this approach may become a standard of care.

Options for patients with FL who may not be suitable candidates for combination chemotherapy include single agent chemotherapy (such as alkylating agents) or rituximab monotherapy which may achieve satisfactory disease control. Monoclonal anti-CD20 antibody conjugated with radio-isotopes are further therapies which have been approved for use either in relapsed disease (in the form of either ^{90}Yttrium-ibritumomab tiuxetan or ^{131}Iodine-tositumomab) or as consolidation of remission

after first-line chemotherapy (^{90}Yttrium-ibritumomab tiuxetan) and have the potential to provide long-term disease control following a single application (Kaminski *et al.* 2001; Davies *et al.* 2004; Morschhauser *et al.* 2008). Other therapies also show promise in FL and the preliminary report on a randomised study of the use of bendamustine in combination with rituximab is encouraging. It indicates this regime has similar efficacy but a better side-effect profile than the widely used R-CHOP regime when employed as initial therapy in advanced stage disease (Rummel *et al.* 2009) and may be particularly relevant to those patients less suited to more intensive approaches.

1.1.8 Role of haematopoietic stem cell transplantation

By contrast, a subset of patients may be candidates for intensive high-dose chemotherapy (HDT) augmented by autologous haematopoietic stem cell transplantation (auto-HSCT). This attempts to provide long-term disease remission but the increased toxicity of this approach is an important limiting factor. Consequently, it is typically reserved for a cohort of younger, fitter patients who have poor risk disease. However, as first line therapy, it cannot yet be recommended outside the context of clinical trials as no overall survival benefit has been demonstrated to date in randomised studies although an improvement in PFS with a plateau indicates that at least for some patients it may be beneficial (Deconinck *et al.* 2005). However, prolonged follow-up confirms an excess of second malignancies following this approach when total body irradiation (TBI) is included in the conditioning regime (Gyan *et al.* 2009). At relapse, HDT with auto-SCT has been investigated in randomised studies and a benefit in overall survival compared to chemotherapy has been reported (Schouten *et al.* 2003). This study again utilised TBI conditioning and other studies with prolonged follow-up after auto-HSCT administered at FL relapse have reported an increase in second malignancies (at a rate of around 15%) associated with a similar approach (Montoto *et al.* 2007a; Rohatiner *et al.* 2007). Conditioning regimes that excluded TBI had a lower rate (3.5%) of secondary malignancies (Montoto *et al.* 2007a) but the role of this approach requires clarification in the era of rituximab immunotherapy.

Allogeneic HSCT (allo-HSCT) represents a potentially curative option in patients with advanced stage FL and its use has recently been reviewed by van Besien (van Besien 2009). It provides a graft versus lymphoma effect, eliminates the risk of graft contamination by lymphoma cells and is associated with lower rates of relapse. However, these effects have frequently not translated in to a survival benefit due to non-relapse related mortality (NRM) rates caused by the intensive, myelo-ablative conditioning and risk of graft-versus-host disease (GvHD). This has prompted investigation of reduced intensity conditioning (RIC) allo-HSCT, an approach which shows promise as indicated by recent studies in relapsed FL. One report, with a median of five years follow-up, provides estimated OS and PFS of 85% and 83% respectively (Khouri *et al.* 2008) and another study, using T cell depletion in an attempt to reduce the risk and severity of GvHD, demonstrates low NRM (15%) and an encouraging PFS (76%) at 4 years (Thomson *et al.* 2010).

Few studies have compared the two transplant approaches, with one randomised study closing early due to sub-optimal recruitment (n = 30) (Tomblyn *et al.* 2010). A two-centre retrospective study compared HDT and auto-HSCT to RIC-allo-HSCT, over a median follow-up of 7 and 3 years respectively, using regimens that did not contain TBI (Ingram *et al.* 2008). The respective disease free (67% v 69%) and overall survival (56% v 58%) are similar for both approaches, reflecting a lower relapse rate but higher NRM of the allogeneic approach. In the auto-HSCT group after censoring at 5.5 years, there were five secondary malignancies recorded (representing 6% cases) and only one in the Allo-SCT group, perhaps reflecting the shorter follow-up.

Encouragingly, the overall clinical outcome of FL is improving as demonstrated for cases managed at St Bartholomew's Hospital over the last four decades **Figure 1.2** and this effect is probably a consequence of recent improvements in both treatment and supportive management of patients with this disease (Fisher *et al.* 2005; Liu *et al.* 2006). However, FL survival still lags behind that of the general population (Swenson *et al.* 2005) and an important issue in this regard is the current inability to reliably predict adverse events over the prolonged course of this disease.

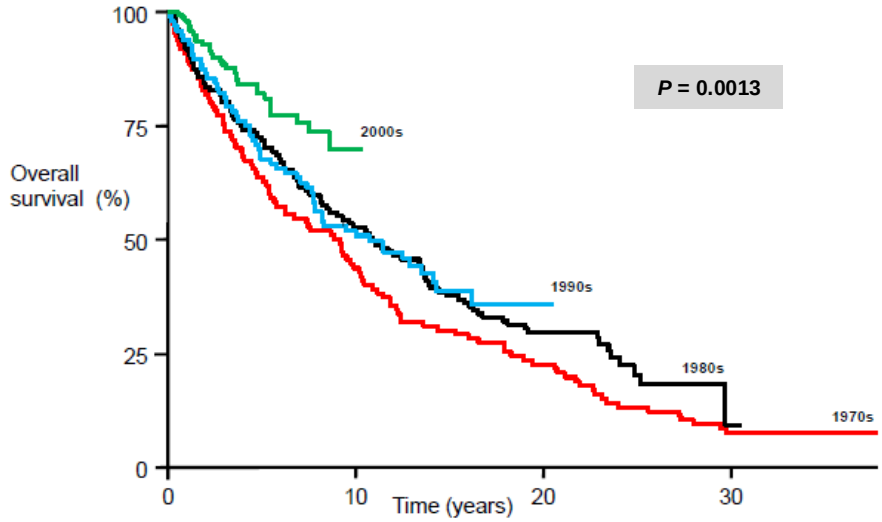


Figure 1.2 Improving overall survival in FL

OS for all patients with FL (n = 834) presenting to Barts and the London NHS trust over four decades (as indicated). Data (unpublished) were extracted in July 2010 from the clinical database previously maintained by the ICRF / CRUK Medical Oncology Unit (now Department of Haemato-Oncology) based at St Bartholomew's Hospital.

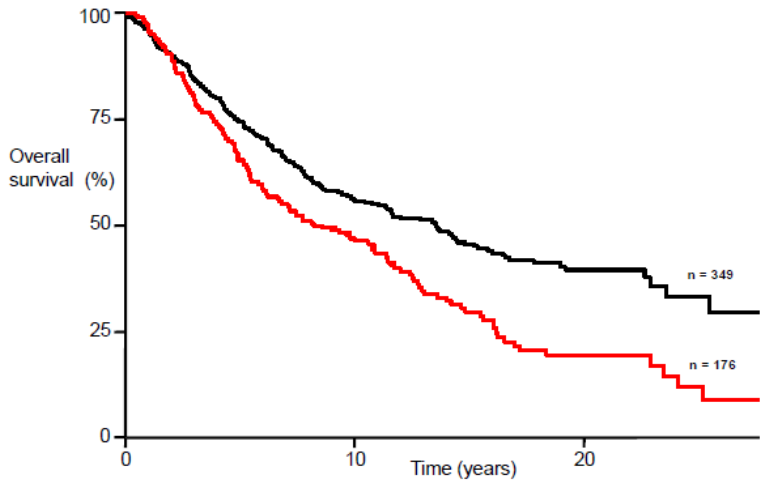


Figure 1.3 Survival of FL which transforms and does not transform

Illustrated is the OS for patients diagnosed with FL and seen at St Bartholomew's Hospital between 1981 and 2001 (reflecting the period from which FL / t-FL samples are investigated in this thesis project). The survival of those patients who never developed transformation (black line) is significantly better than those whose disease transformed at any period (red line). The data represent an update of a report from this institution (Montoto *et al.* 2007b).

1.2 Transformation

A particularly significant adverse disease event that can present as clinical relapse or progression at any stage of FL is its transformation to aggressive lymphoma (t-FL). The morphological appearances of such tumour samples resemble *de novo* DLBCL in the vast majority of cases although transformation to Burkitt lymphoma and acute lymphoblastic lymphoma or even Hodgkin lymphoma can occur, albeit rarely (LeBrun *et al.* 1994; Kroft *et al.* 2000). Two recent large studies with prolonged follow-up indicate transformation develops at a median of 3 to 4 years following FL diagnosis (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008). Both reports indicate that around one third of FL will develop t-FL at some point in the course of their disease, with one suggesting a persistent risk of around 3% per year (Al-Tourah *et al.* 2008). Importantly, FL cases that develop t-FL have worse survival compared to those cases that do not **Figure 1.3** and the outcome following transformation is particularly poor, with a median survival from this event of around 1.5 years.

Clinical indicators of t-FL include an isolated enlarging lymph node, the onset of B symptoms (night sweats, fevers, unexplained loss of weight), unusual sites of new disease or organ involvement as well as biochemical abnormalities including hypercalcaemia and rising serum lactate dehydrogenase (LDH) level which should prompt further evaluation of the patient. Although the presence of at least one of these features distinguishes a group of FL patients who have similar outcome to those with biopsy proven transformation (Al-Tourah *et al.* 2008), until this approach is prospectively validated, obtaining histological confirmation of transformation will remain an important aspect of patient management.

Treatment of t-FL is a clinical challenge and it remains to be seen whether the advent of rituximab immunotherapy will reduce the incidence of this event or improve its outcome. As it is usually histologically and clinically similar to *de novo* DLBCL, therapies developed for DLBCL are frequently applied to t-FL. However, treatment is often complicated by prior therapies and patients' older age with attendant co-morbidities. When prolonged remission is an appropriate goal then combination immuno-chemotherapy, which is consolidated by HDT and auto-HSCT or even RIC-allo-HSCT in some centres, is appropriate (Montoto *et al.* 2007a; Thomson *et al.* 2009). If symptom control is the primary objective then local radiotherapy alone or

attenuated immuno-chemotherapeutic regimes may be appropriate. However, few reports deal exclusively with the management of t-FL and there is a pressing need for co-operative studies addressing this issue. In this regard, current survival data for FL cases on the clinical database at St Bartholomew's Hospital suggests there has been no improvement in survival of FL that transforms when cases from each of the last four decades are compared, log rank $P = 0.44$ (unpublished). Indeed, t-FL is a significant contributor to death in patients with FL as a previous study identified that 42 / 95 (44%) deaths in FL over a period of 9 years from diagnosis were attributable to t-FL (Bastion *et al.* 1997) and, in an earlier study, 68% of cases who died after disease progression had evidence of histological transformation on post-mortem examination (Garvin *et al.* 1983).

Prediction of transformation might facilitate better disease management and clinical parameters from FL diagnosis that are routinely available have been examined for their prognostic potential. Indeed, higher FLIPI scores associate with increased risk of subsequent t-FL (Gine *et al.* 2006; Montoto *et al.* 2007b) as does advanced stage (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008) and, as previously reviewed (Muller-Hermelink and Marx 2000), other studies suggest that grade 3b FL or FL with diffuse areas (before its re-classification (Swerdlow 2008)), may associate with transformation. However, in an earlier study, histological sub-types of FL determined according to the previously used Working Formulation ("The Non-Hodgkin's Lymphoma Pathologic Classification Project" 1982) did not associate with t-FL, and serum β_2M was the only diagnostic FL parameter to correlate with risk of this event (Bastion *et al.* 1997).

The same study also revealed that a high serum LDH level at transformation, as well as the type of and response to therapy administered as treatment of t-FL, correlate with outcome after transformation. A more recent study also suggests response to salvage therapy at t-FL is predictive of outcome (Montoto *et al.* 2007b) while another demonstrates that limited disease stage at t-FL associates with better survival (Al-Tourah *et al.* 2008). Nevertheless, despite these prognostic factors, it is not yet possible to reliably predict t-FL occurrence or outcome for individual patients and this suggests these surrogate markers of disease have limited utility in its management.

Some recent reports provide insights in to the biological mechanisms underlying t-FL, as reviewed by Bernstein and Burack (Bernstein and Burack 2009), but a definitive picture of the critical molecular aspects of transformation remains elusive. Nonetheless, these studies perhaps represent the first steps towards molecular profiling, an approach already used in other haematological malignancies to guide patient treatment. In FL, such a tool could be particularly relevant in developing risk adapted management through identifying, firstly, those FL cases destined to transform or not and, secondly, the most appropriate management options in those cases that develop t-FL. These are important goals as transformation remains a significant cause of morbidity and mortality in FL.

1.3 Normal B cell development

Before addressing the molecular aspects of FL / t-FL, it is important to consider the normal B lymphocyte as FL arises as a malignant clone of these cells and the genetic aspects of normal B cell ontogeny have a role in the development and progression of FL.

The fundamental role of B cells is to produce immunoglobulin molecules as antigen-specific antibodies that contribute to the immune response. B cells initially develop in the bone marrow from haematopoietic stem cells and undergo two distinct phases of development as have been extensively reported (Tonegawa 1983; Rajewsky 1996; van Zelm *et al.* 2007) and which are summarised here.

Within the bone marrow an antigen independent differentiation phase includes rearrangement of the immunoglobulin (Ig) gene loci. There are three Ig gene loci *IGH@* (on chromosome 14q32.3), *IGK@* (chromosome 2p) and *IGL@* (chromosome 22q) the protein products of which form Ig molecules initially expressed on the cell surface as functional Ig receptors. An antigen dependent phase of differentiation then follows after passage of immature B cells from the bone marrow, via the blood, to secondary lymphoid organs. When these B cells bind compatible antigen in the peripheral compartment they undergo clonal proliferation and further differentiation in the germinal centre (GC) region of the lymph node (MacLennan 1994). It is in the GC that naive B cells develop in to centroblasts and subsequently centrocytes which

then form either plasma cells or long-lived memory B cells that each produce and secrete the antigen-specific Ig molecules as antibodies (Allen *et al.* 2007).

The process of immunoglobulin formation is critical to B cell maturation and function and involves a complex series of steps involving both genetic events and post-translational glycosylation. The mature Ig molecule includes two heavy chains (IgH) and two light (either Igκ or Igλ) chains which are held together by disulphide bonds between cysteine residues in a "Y" shape. Indeed, the incorporation of only one or other light chain in to the Ig molecule is used routinely as an adjunct to histological diagnosis as, in B cell tumours, clonal expansion of the malignant population restricts them to surface expression of either Igκ or Igλ. By using probes against the κ and λ segments of the Ig molecule, clonality can then be inferred by detection of only one or other in the (presumed) malignant cells, helping to distinguish them from polyclonal B cell expansions (van Dongen *et al.* 2003). There is also diversity within the IgH molecules which creates five antibody subtypes (IgA, IgD, IgE, IgG, IgM), two of which require further structural modification to form dimers (IgA) or pentamers (IgM). These subtypes impart additional specifications on the anatomical locations of action and downstream functional consequences of antigen binding for the respective mature Ig molecules.

The base of each Ig molecule is termed the F_C region and is a conserved structure formed by the two heavy chains. The F_C region interacts with immune effector mechanisms via F_C-receptors on other cells including macrophages or interacting with proteins of the complement system to effect a range of physiological immune processes, including lysis of cells, recognition of opsonised particles and activation of other immune mediators (eosinophils, basophils, mast cells). The two arms of the Ig molecule are termed the F_{AB} regions and each is each formed by heavy and light chains. They contain conserved (constant) domains but the heavy and light chain variable domains (V_H and V_L respectively) are the determinants of antigen specificity. Within these V_H and V_L regions, three areas bind antigen and are termed the complementarity determining regions (CDRs) and these are flanked by conserved sequences termed framework regions (FRs).

There is a huge diversity in the antigen specificity of Ig molecules. This reflects the almost limitless number of antigens that might be encountered by a B cell and it is

physiological genetic changes that drive this variability. There are three stages in this process:

a) The variable region of the Ig heavy chain includes three segments termed variable (V), diversity (D) and joining (J) while the light chains consist of only V and J segments. The first stage in selecting for antigen specificity in Ig molecules involves rearrangements of variable regions of the respective gene loci, first identified over thirty years ago (Hozumi and Tonegawa 1976; Brack *et al.* 1978). There are multiple V, D and J segments with, for example, a minimum of 66, 27 and 6 respective segments for *IGH*. Each developing B cell selects only one of the V / D / J or V / J gene segments for *IGH* and *IGK*/*IGL*, respectively. This is a random process termed V(D)J recombination **Figure 1.4**, which occurs within the developing B cells in the bone marrow, and the many different potential VDJ rearrangements translates in to a "combinatorial repertoire" reflecting around 2×10^6 Ig molecules (van Dongen *et al.* 2003). The V(D)J recombination process is mediated by the protein products of the recombination activating genes, *RAG1* and *RAG2*, which cause transient breaks in DNA to enable the selection of the different segments. Further diversity is conferred on the variable region by the occurrence of small deletions and nucleotide insertions at the junction regions of V / D / J segments during the recombination process (van Dongen *et al.* 2003). VDJ recombination occurs first in *IGH* and then, after production of a functional IgH protein, VJ recombination occurs in *IGK*. At this stage *IGK* deletion may occur and, if so, VJ recombination occurs in *IGL*. Following V(D)J recombination, the complementary allele (which has not undergone rearrangement) is effectively silenced so the B cell can only produce one form of variable heavy and light chains (Bergman and Cedar 2004). This immature B cell containing the rearranged *IGH* and either *IGK* or *IGL* loci then passes from the bone marrow.

The V(D)J process is relevant in B cell malignancy as clonal expansion means the tumour cells have a common *IGH* rearrangement with its detection already used as a management tool to monitor levels of minimal residual disease (MRD) in acute lymphoblastic leukaemia (Patel *et al.* 2010). Furthermore, the multiple *IGH-VH* segments can be grouped in to sub-family groups (1 to 7) based on the sequence similarity of three FRs in the VH segments. Consequently, the FR sites have been used for design of primers that are complementary to all VH segments within the

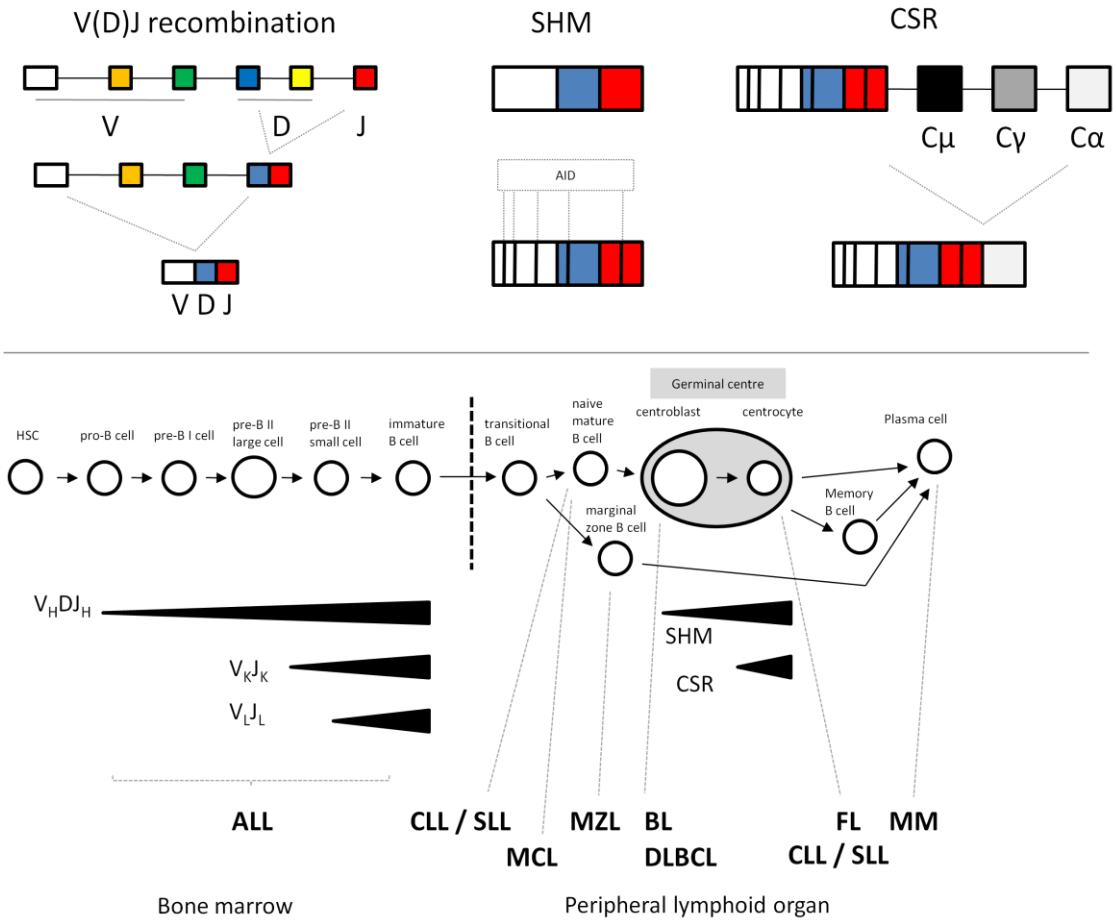


Figure 1.4 B cell development and B-NHL

Top, schematic representation of the processes underlying B cell antibody affinity maturation. For illustrative purposes limited V (n = 3), D (n = 2) and J (n = 1) segments (highlighted by different colours) are shown to illustrate V(D)J recombination, here for the *IGH* variable region and vertical black lines represent the sites of SHMs created by AID. Similarly, limited constant segments representing IgM (Cμ), IgG (Cγ) and IgA (Cα) are used to illustrate the process of CSR, here selecting for an IgA heavy chain molecule. Bottom, B cell ontogeny in the bone marrow and periphery is indicated with circles representing the various cell stages. The sites and stages of the three aspects of V(D)J recombination (for *IGH*, *IGK* and *IGL*), SHM and CSR are indicated. The germinal centre is highlighted in grey and examples of different B-NHL corresponding to their cell of origin are given at the foot of the figure. The figure is an amalgam adapted from figures by Kuppers and colleagues, Kuppers and Dalla-Favera and van Zelm and colleagues (Kuppers *et al.* 1999; Kuppers and Dalla-Favera 2001; van Zelm *et al.* 2007). ALL indicates acute lymphoblastic leukemia; CLL / SLL, chronic lymphocytic leukemia / small lymphocytic lymphoma; MCL, mantle cell lymphoma; MZL, marginal zone lymphoma; BL, Burkitt lymphoma; DLBCL, diffuse large B cell lymphoma; FL, follicular lymphoma; and MM, multiple myeloma.

respective *IGH-VH* subgroup (van Dongen *et al.* 2003). Application of these primers can then distinguish clonal from polyclonal B-cell proliferations as clonal expansions will reveal distinct, single PCR products that represent only one VH sub-family compared to heterogeneous PCR products obtained from polyclonal proliferations.

b) The second phase of antigen diversity is mediated by the process of somatic hypermutation (SHM). This occurs in the GC of lymph nodes, requires T cell help through CD40-CD40L mediated interactions with B cells and involves the introduction of single nucleotide mutations in the variable regions of the Ig genes (Berek *et al.* 1991; Jacob *et al.* 1991; Kuppers *et al.* 1993), preferentially in the CDRs, by the activation-induced cytidine deaminase (AID) enzyme (Muramatsu *et al.* 1999; Muramatsu *et al.* 2000; Arakawa *et al.* 2002; Nagaoka *et al.* 2002) as illustrated in **Figure 1.4**. This can lead to changes in the amino-acid sequence of the antigen binding sites in the surface Ig receptor molecule and attendant changes in its antigen binding affinity. If the affinity is reduced then the B cell is selected for destruction by apoptosis. However, B cells with increased antigen affinity of the Ig molecule undergo preferential selection (Liu *et al.* 1989) and clonal expansion then subsequently secrete antibody (Kuppers and Dalla-Favera 2001).

c) A final phase contributing to antibody diversity is termed class switch recombination (CSR). This also occurs in the GC and involves selection of single constant regions (**Figure 1.4**). The antigen affinity is not changed but the Ig molecule subtype can change from the IgM or IgD expressed on the surface of naive mature B cells to one of IgA, IgE and IgG. CSR involves deletion of sections of the *IGH* constant domain by the creation of DNA breaks and their subsequent repair thought to be mediated by AID and uracil DNA glycosylase (Casali and Zan 2004). This creates B cell progeny which have the same antigen affinity but differing Ig subtypes (termed isotypes). The differing immune effector mechanisms conferred by the respective isotypes provide alternative means by which the same antigen may be cleared, giving an additional degree of diversity in the antibody response to antigen. The antibody isotype has clinical utility, particularly in the most mature B cell malignancy, multiple myeloma (which is a clonal proliferation of plasma cells). The presence of excessive amounts of one antibody isotype in serum, or as secreted light chain fragments in the urine, suggests clonality and is used to monitor disease and its response to therapy.

As both V(D)J recombination and SHM also occur in malignant B cells arising from the GC, these features are used later in this thesis to establish the clonal relationship of FL and t-FL samples obtained from patients and to model the evolution of transformation.

1.4 Genetic changes in FL

The evidence that these physiological processes occur in various forms of B-NHL has contributed to the determination of normal B cell counterparts for a number of these tumours, as reviewed by Kuppers and colleagues (Kuppers *et al.* 1999) and illustrated in **Figure 1.4**. In FL, a report from 1986 was the first to demonstrate the presence of SHM sites in tumour cells by sequencing the *IGH* variable regions (Cleary *et al.* 1986b). Subsequently, malignant clonal expansion was shown to occur after the initiation of SHM (Zelenetz *et al.* 1992) a process which is ongoing in FL cells (Kon *et al.* 1987; Cleary *et al.* 1988; Zelenetz *et al.* 1992; Ottensmeier *et al.* 1998). This indicates that FL has a GC B cell origin and that intra-clonal genetic diversity occurs within the tumour and can be detected by sequencing SHM sites in the variable region. Moreover, a recent report demonstrates that AID activity is required for the development of GC derived lymphoma in mouse models, providing experimental evidence that AID induced events (namely CSR and SHM) are required to generate the secondary genetic lesions (such as translocations and the ongoing SHM) occurring in the malignant transformation of B cells at the GC stage of differentiation (Pasqualucci *et al.* 2008).

1.4.1 Translocation (14;18)

Translocations of chromosomal material occur frequently in cancer ("Mitelman Database" 2011) and may drive the disease through disruption of normal gene function particularly through creation of fusion genes or dysregulation of gene expression (often the up-regulation of oncogenes). Many examples have followed the first description of a cancer associated translocation, the Philadelphia chromosome in chronic myeloid leukaemia (Tough *et al.* 1961) which juxtaposes the *BCR* and *ABL* genes leading to a fusion protein product that has a critical role in driving this malignancy. In B-NHL translocations involving *IGH* (and less commonly *IGK* / *IGL*)

are frequent events (Willis and Dyer 2000) reflecting the disruption and strand breaks caused to DNA during V(D)J recombination, SHM and CSR (Kuppers and Dalla-Favera 2001).

In FL, around 85% of cases (Weiss *et al.* 1987; Yunis *et al.* 1987) harbour the translocation (14;18)(q32;q21), referred to herein as t(14;18), which occurs as a result of mistakes in VDJ recombination (Tsujimoto *et al.* 1985b) and juxtaposes the *BCL2* gene to the *IGH* locus (Tsujimoto *et al.* 1984) as illustrated in **Figure 1.5**. This typically creates two derivative chromosomes, with the *IGH* breakpoints on chromosome 14 occurring in the J_H region. The *BCL2* breakpoints on chromosome 18 cluster together principally in the 3' UTR region of the gene (Tsujimoto *et al.* 1985a; Cleary *et al.* 1986c), within a 150 bp site termed the major breakpoint region (MBR), although a second cluster occurs around 20 kb downstream of *BCL2* (Cleary *et al.* 1986a; Buchonnet *et al.* 2000), termed the minor breakpoint region (mbr) and, more rarely still, breakpoints may be "atypical" (such as those within the gene's large 3rd exon). Significantly, a report has now demonstrated that the *BCL2* MBR site can acquire altered DNA structure that can then be cut by RAG nucleases and so provides a potential mechanism underlying the consistent involvement of this site in translocations (Raghavan *et al.* 2004). Furthermore, t(14;18) can be detected in the blood of patients with FL and so has been used as a molecular marker of disease response, particularly following attempts to achieve long-term remission through HDT with Auto-SCT (Gribben *et al.* 1993; Mahe *et al.* 2003).

The translocation leads to the over-expression of the anti-apoptotic BCL2 protein conferring a survival advantage on affected cells. The significance of BCL2 was illustrated in a series of publications through the 1980s and early 1990s (Tsujimoto and Croce 1986; Alnemri *et al.* 1992) with evidence that it was growth promoting in lymphoid cells (Nunez *et al.* 1989) and tumorigenic in mouse models (Reed *et al.* 1988). However, despite its frequency in FL, an important study demonstrated that although t(14;18) promoted polyclonal lymphoid hyperplasia in mice, progression to lymphoma was associated with secondary changes. This suggests that t(14;18) in isolation is insufficient to cause lymphoma (McDonnell and Korsmeyer 1991). Support for this model in humans is provided by the lack of t(14;18), or sub-karyotypic *BCL2-IGH* rearrangements, in around 10 to 15% of FL cases and a recent report proposing that a specific t(14;18) negative disease subtype exists that is

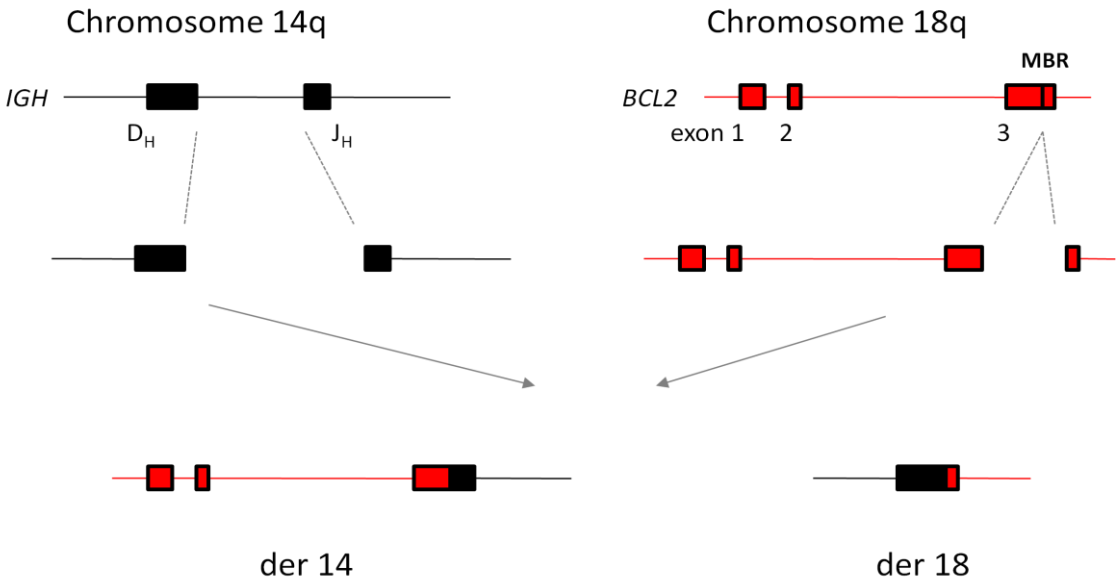


Figure 1.5 t(14;18)

Illustrated is the process of translocation between the *IGH* and *BCL2* loci. Left, a DNA double strand break in *IGH* occurs as part of V(D)J recombination. Right, the MBR situated in the 3' region of *BCL2* is prone to similar breaks (Raghavan *et al.* 2004). Bottom, aberrant repair of the breaks leads to the translocation, possibly facilitated by the spatial proximity of the *IGH* and *BCL2* loci in B cells (Roix *et al.* 2003), and formation of derivative (der) chromosomes 14 and 18. For illustrative purposes both *IGH* and *BCL2* are orientated 5' to 3' but they each occur on reverse DNA strands hence der 14 contains *IGH-BCL2*. The figure is adapted from Lieber and colleagues (Lieber *et al.* 2006).

characterised by inguinal lymphadenopathy, diffuse histological features and presence of chromosome 1p deletion (Katzenberger *et al.* 2009). Furthermore, t(14;18) can be detected at low levels in the peripheral blood mononuclear cells of normal, healthy individuals (Limpens *et al.* 1995) at a prevalence of 23 to 50% (Dolken *et al.* 1996; Summers *et al.* 2001) as compared to the much lower rate of FL in the population (Morton *et al.* 2006; Friedberg *et al.* 2009), indicating other factors are important in disease development. Nonetheless, an age associated increase in the prevalence of t(14;18) in healthy individuals (Schuler *et al.* 2009), which parallels the distribution of FL cases across the population, provides epidemiological support for this translocation as a predisposing factor for this lymphoma. This is further supported by the demonstration that t(14;18) occurs in an atypical expanding population of B cells with FL-like characteristics in healthy individuals (Roulland *et al.* 2006) suggesting these constitute a pre-malignant cell pool from which FL can develop.

1.4.2 Other genetic aspects of FL

Additional cytogenetic events are common in FL. Instead of being translocated to the *IGH* region, *BCL2* may occasionally be juxtaposed to *IGK* or *IGL* by t(2;18)(p11;q21) and t(18;22)(q21;q11), respectively (Hillion *et al.* 1991; Knutsen 1997; Lin *et al.* 2008). Other translocations include those between 3q27 and the immunoglobulin gene loci (Otsuki *et al.* 1995), disrupting the *BCL6* oncogene, and rare events such as t(16;18)(p13;q21) (Mahmoodi *et al.* 2004). Rearrangement of the *MYC* oncogene is observed in aggressive B-NHL (Burkitt lymphoma and intermediate BL/DLBCL) and is associated with a poor prognosis in DLBCL (Savage *et al.* 2009; Barrans *et al.* 2010). So called "double-hit" lymphomas containing both *MYC* and *BCL2* rearrangements have a particularly poor outcome (Johnson *et al.* 2009b). Consistent with this, studies in FL suggest that *MYC* rearrangements might confer a more aggressive clinical course if present at diagnosis (Christie *et al.* 2008), the rearrangements can occur at disease progression (De Jong *et al.* 1988; Yano *et al.* 1992) and high level amplification of the *MYC* locus has been reported in FL (Bentz *et al.* 1996).

However, the most common events are chromosomal gains or losses with an early study suggesting that loss 6q and gains 7, 12q and 2p often accompanied a more clinically aggressive disease course (Yunis *et al.* 1987). A report then addressed the prognostic value of karyotypic changes, revealing that loss or disruption of 6q23-26 or 17p associated with worse survival and predicted shorter time to transformation (Tilly *et al.* 1994) and around the same period a study of multiple NHL subtypes showed that loss 6q, gain 7 and gain 12 were enriched in FL (Johansson *et al.* 1995). In a large subsequent study (n = 165) assessing only t(14;18) containing FL diagnostic samples, almost all cases (97%) had additional cytogenetic changes (Horsman *et al.* 2001) at a median of 8 per case. A follow-up investigation (n = 336) revealed that secondary cytogenetic events delineate differing initial routes of disease development. In this study, loss 1p (20%) and loss 6q (30%) were the most common additional changes but only loss 17p or gain 12 associated with inferior survival (Hoglund *et al.* 2004). However, this study included a cohort of non-diagnostic cases who had received treatment prior to the included biopsy sample and prior treatment itself associates with an increase in karyotypic changes (Offit *et al.* 1991).

These cytogenetic studies identified a large number of heterogeneous gains and losses. However, only a subset are recurrent and even less have prognostic relevance. Other studies have employed newer, higher resolution techniques in attempts to better distinguish regions of genetic change which may be important in FL. Using CGH (Kallioniemi *et al.* 1992), Viardot and colleagues identified 7p / 7q or 6q as most frequently affected by gain or loss, respectively. Furthermore, they could determine minimum common regions affected within these areas with loss 6q25 the single cytogenetic event to independently associate with worse survival (Viardot *et al.* 2002).

During the course of this thesis, micro-array-CGH (array-CGH) (Pinkel *et al.* 1998) has been used to investigate diagnostic FL samples in a study recently reported by the British Columbia Cancer Agency (BCCA). This identified a number of regions of recurrent genomic copy number change (some as small as 184 kb) with two, loss 1p36.22-p36.33 and loss 6q21-q24.3, predicting both overall survival and risk of future disease transformation independently of the IPI (Cheung *et al.* 2009b). The prognostic significance of these two regions builds on the previous reports identifying loss affecting 6q as a recurrent event with another recent array-CGH

study identifying that loss of 6q25 and 6q26 (as well as 9p21) associate with worse survival (Schwaenen *et al.* 2009). Furthermore, another micro-array study has identified that in 50% of FL loss of heterozygosity (LOH), as either copy number loss or copy neutral LOH, affects chromosome 1p36 (Ross *et al.* 2007) and therefore represents the second most common chromosomal change described in FL. Copy neutral LOH (termed acquired uniparental disomy; aUPD) affecting 1p36 has since been shown by our group to associate with worse survival in FL and aUPD 16p predicts shorter time to progression and increased risk of transformation (O'Shea *et al.* 2009).

The frequency and prognostic value of the various cytogenetic associations is an important aspect of FL, however, a recent large study (n = 210) demonstrates that the complexity of cytogenetic aberrations does not associate with outcome (Johnson *et al.* 2008b). This emphasises that unravelling the underlying genetic events is a key aspect towards understanding the mechanisms that drive this disease. In this regard, although disruption of gene function is important in cancer, until recently little has been known as to its relevance in FL. However, mutations of the important tumour suppressor gene *TP53*, although only occurring in 6% of FL at diagnosis, are now known to associate with worse outcome (O'Shea *et al.* 2008) which is in keeping with the prognostic significance attributed to loss of chromosome 17p (Tilly *et al.* 1994) which contains the *TP53* locus.

Other mechanisms may be important in FL but, although changes in RNA expression may reflect underlying gene disruption, a recent profiling study reveals a highly uniform pattern of gene expression at FL diagnosis (Glas *et al.* 2007) suggesting this approach may not be able to discriminate the critical events driving the malignant cells in this disease. Moreover, although DNA methylation appears important (Killian *et al.* 2009; O'Riain *et al.* 2009), the precise role of this phenomenon in FL remains unclear. Consequently, the genetic aspects of FL remain a fundamental area for exploration. Indeed, since commencement of this thesis project, other groups have identified mutations in two genes, *FAS* (Johnson *et al.* 2009a) and *TNFRSF14* (Cheung *et al.* 2010b), that predict FL outcome. Even more recently, genome-wide sequencing has revealed novel mutations in *EZH2* (Morin *et al.* 2010a), *CREBBP* (Pasqualucci *et al.* 2011), *MLL2* and *MEF2B* (Mendez-Lago *et al.* 2010; Morin *et al.*

2010b) at varying frequencies but all with potential functional significance giving intriguing new insights into FL which are discussed later in this thesis.

1.5 Genetics of transformation

The demonstration that loss 1p36 or 6q21-24 (Cheung *et al.* 2009b) as well as aUPD 16p (O'Shea *et al.* 2009), when present at FL diagnosis, are independently predictive of future t-FL strongly suggests these regions contain factors that predispose to transformation. Sometime ago, loss 6q23-26 or 17p in FL were also associated with risk of transformation although insufficient events occurred to enable multivariate analysis (Tilly *et al.* 1994) and somatic mutations occurring in the first intron of *BCL6* have since been shown to independently predict risk of t-FL (Jardin *et al.* 2005). These molecular prognosticators may, in future, help guide a pre-emptive strategy aimed at preventing transformation. However, at present these genetic predictors have not been validated in prospective studies and do not form part of the standard diagnostic process, so they have yet to influence routine patient management.

Other studies have attempted to determine transformation specific events which might therefore identify targets for therapy in the event that t-FL occurs. These have compared paired FL and subsequent t-FL biopsy samples from cases whose disease transformed.

CGH studies have identified copy number changes detectable only at transformation. Gain 12q13-14 (including *CDK4* and *MDM2*) occurred at transformation in half of the cases from a set (n = 24) with paired FL / t-FL samples (Hough *et al.* 2001). A small study (n = 6) then used micro-dissected FL tumour cells for comparison against paired t-FL samples and CGH analysis identified heterogeneous copy number changes acquired at transformation, including gain 7 in five cases (all initially grade 1 or 2) (Boonstra *et al.* 2003). Another CGH study (n = 30) detected loss 6q16-21 and gain 7pter-q22 only in t-FL (Berglund *et al.* 2007). Following the association of gain 7 with FL (Johansson *et al.* 1995), this was assessed further in FL and t-FL using a FISH probe which detected gain 7 occurring at FL progression and transformation (Bernell *et al.* 1998). As previously, new techniques have improved resolution for the detection of copy number changes affecting chromosomes and an

array-CGH study (n = 12) identified a range of copy number changes in both FL and t-FL (Martinez-Climent *et al.* 2003). In particular, 14 gains or losses that were not found in the diagnostic FL samples (or FL relapse) were present in t-FL samples with gains 12pter-12q12 and Xq25-q26 each detected in 2/10 t-FL but not in FL samples. Losses of 1p36 and 6q22-24 were present in 5/12 and 4/12 FL cases respectively suggesting enrichment in cases destined to transform. In a more recent study, LOH was detected more often at t-FL than FL and renders mutations of *CDKN2A* (9p) and *TP53* (17p) homozygous in a subset of cases (Fitzgibbon *et al.* 2007).

Mutations affecting *TP53* were the first to be associated with transformation as reported by Lo Coco and colleagues. They detected *TP53* mutation only at transformation in 3/5 t-FL cases and in both FL / t-FL samples from one of the two remaining cases whereas no mutations were detected in paired samples from a cohort of FL (n = 18) that had progressed but not transformed (Lo Coco *et al.* 1993). Other reports indicate that *TP53* mutations occur only rarely at FL diagnosis or progression, being more common at transformation (Sander *et al.* 1993; Davies *et al.* 2005). Similarly, mutations and deletions of *CDKN2A* have been associated with t-FL (Elenitoba-Johnson *et al.* 1998; Pinyol *et al.* 1998) and, indeed, loss of *CDKN2A* expression occurs with disease progression (Villuendas *et al.* 1998). Other mutational targets at transformation include *BCL6* (Lossos and Levy 2000), *BCL2* (Matolcsy *et al.* 1996) and *MYC* (Lossos *et al.* 2002; Rossi *et al.* 2006) which is also frequently rearranged at FL progression (Lee *et al.* 1989; Yano *et al.* 1992).

1.6 Evolution of t-FL

The acquisition of additional genetic events at t-FL represents the culmination of the conventional model of FL progression by direct clonal evolution. Other studies have attempted to define functional models of transformation through expression profiling. These suggest roles for dysregulation of genes important in cell-cycle control including *MYC* and its target genes (Lossos *et al.* 2002; Davies *et al.* 2007) as well as *MAPK* (Elenitoba-Johnson *et al.* 2003) but the overall importance of these pathways in t-FL remains unclear. Moreover, the less common cases of composite, discordant and down-grade lymphoma (defined with regard to FL as: FL and DLBCL in same biopsy specimen; FL and DLBCL present simultaneously at different anatomical sites; and *de novo* DLBCL presenting subsequently with FL,

respectively) hint at complex evolutionary behaviour, while certain studies indicate genetic changes in FL / t-FL may occur non-sequentially.

This was first suggested just over a decade ago by Matolcsy and colleagues who compared the SHM patterns from the *IGH* variable region in FL and DLBCL clones from a case of FL that transformed. This revealed certain SHMs were common to FL and DLBCL, some occurred exclusively in DLBCL but, importantly, others detected in FL were absent from DLBCL and a "common malignant ancestor" was proposed, in which divergent evolution of an FL subclone gave rise to DLBCL (Matolcsy *et al.* 1999). Support for this model comes from the observation that a large proportion of t-FL samples lack cytogenetic or molecular alterations found in the preceding FL samples (Martinez-Climent *et al.* 2003; Fitzgibbon *et al.* 2007). Genetic changes in FL / t-FL can therefore be acquired in a pattern that does not mirror the clinical course of disease. Furthermore, a long-lived FL initiating cell can give rise to clonally divergent disease in separate patients, following transmission from donor to recipient of a haematopoietic stem cell transplant (Hart *et al.* 2007). Consequently, divergent (indirect) evolution from a common origin represents an alternative mechanism underlying transformation, a concept that is addressed further in Chapter 5.

1.7 Microenvironment

In addition to the specific attributes of malignant cells, the cells which constitute the surrounding tumour microenvironment also have a significant role in FL. Significantly, Levy and colleagues demonstrated almost two decades ago that an immune response could be induced against autologous tumoral antigens in patients who had been treated for FL and this could lead to tumour regression in certain cases (Kwak *et al.* 1992). Subsequently, a large study by Dave and colleagues identified that RNA expression profiles derived from tumour infiltrating (non-malignant) immune cells independently associate with survival in FL (Dave *et al.* 2004) with favourable outcome for a profile enriched with T cell markers compared to a worse outcome with a profile containing predominantly macrophage derived markers. A more recent study that provides confirmation as to the importance of this association also addressed expression at the protein level, demonstrating that infiltrating macrophages associate with worse outcome (Byers *et al.* 2008). Several other studies

using immuno-histochemical staining have identified associations between tumour infiltrating immune cells and FL survival (Farinha *et al.* 2005; Alvaro *et al.* 2006; Lee *et al.* 2006a; Carreras *et al.* 2009). Although their results at times appear contradictory, a potential explanation is offered by one recent study which suggests the impact of tumour microenvironment on prognosis in FL depends on specific treatment protocols (de Jong *et al.* 2009). Furthermore, two relatively large studies have shown patterns of infiltrating T cells in FL correlate specifically with risk of future t-FL (Glas *et al.* 2007; Farinha *et al.* 2010) indicating it is not only tumour cell specific events that influence the process of transformation.

Although the precise role of the cells that constitute the surrounding microenvironment in FL is not clear, an elegant review by Burger and colleagues (Burger *et al.* 2009) proposes three differing patterns of relationship between haematological cancers and the microenvironment. Two scenarios involve either tumour proliferation independent of stromal signalling, as in the case of Burkitt Lymphoma, or a dysfunctional environment in which tumour cells retain a dependence on stromal signalling for growth but "parasitize" in the microenvironment through altered stromal signalling, as in the case of the acute leukaemias. The third scenario (attributable to FL) involves a regulated co-existence of tumour cells with the microenvironment in which tumour proliferation is critically dependent on signals from the surrounding stroma. Furthermore, the association of a stromal cell ligand (APRIL) and FL cell receptor (TACI), causing proliferation of FL cells and cyclin D1 upregulation within them (Gupta *et al.* 2009), demonstrates one mechanism by which this might occur.

Given the importance of the immune cell repertoire in normal B cell development, determining whether the acquired genetic changes that occur in FL modulate the interactions of these malignant B cells with the microenvironment may provide valuable insights into the progression of this disease. Alternatively, the associations between immune response and disease outcome might (at least to a degree) reflect innate differences in host ability to clear the malignant cells, suggesting an intrinsic immune predisposition to FL development. However there remains much to learn about the microenvironment and its role in FL, and the sometimes contradictory findings require clarification so that their clinical relevance can be optimally established.

1.8 Predisposition to FL

The potential relevance of immune predisposition to FL reflects a wider spectrum of both external (environmental) and internal (inherited) predisposing factors that contribute to cancer and which, in addition to the acquired genetic events already described, are relevant in FL.

1.8.1 Environmental factors

Significantly, the incidence of FL shows geographical variation, being higher in Western countries (Anderson *et al.* 1998), and studies hinting at an increased risk of NHL in spouses of cases have included occasional husband / wife FL pairs (Friedman and Quesenberry 1999; Villeneuve *et al.* 2009). Consequently, environmental factors may be important in FL and these have been examined in two recent comprehensive reviews of NHL epidemiology (Fisher and Fisher 2004; Alexander *et al.* 2007).

Smoking is a major environmental carcinogenic determinant, as identified in a seminal report on lung cancer almost half a century ago (Doll and Hill 1964) and an association with FL risk is demonstrated in a recent meta-analysis (Morton *et al.* 2005). Dietary factors may be also be relevant (Chang *et al.* 2005) as may use of hormone replacement therapy (Cerhan *et al.* 2002). In particular, the potential immune-modulatory properties of (allogeneic) blood transfusion have led to several studies investigating risk of FL following this procedure but a recent meta-analysis shows no association (Castillo *et al.* 2010). Furthermore, the increased use of pesticides over recent decades, their potential toxic effects and their early use in the Western world have, in combination with the concurrent increase in NHL incidence, prompted several association studies. Although pesticide exposure has been inconsistently associated with *overall* NHL risk, reports suggest a higher prevalence of t(14;18) containing cells in farmers exposed to pesticides (Roulland *et al.* 2004) and that development of t(14;18)-positive NHL is specifically associated with pesticide exposure (Schroeder *et al.* 2001; Chiu *et al.* 2006). Importantly, Agopian and colleagues have since demonstrated that pesticide exposure leads to clonal expansion of t(14;18)-containing cells and that these show similar properties to malignant FL cells (Agopian *et al.* 2009) providing a direct link between

environmental carcinogen exposure, t(14;18) frequency in peripheral blood and malignant progression which has significant public health implications.

1.8.2 Familial predisposition

Inherited predisposition may contribute to cancer and it is now four decades since Knudson first showed this for familial genetic determinants in retinoblastoma (Knudson 1971). Although it has been recognised for some time that germline factors contribute to risk of NHL (Linet and Pottern 1992), until recently less was known regarding the different subtypes.

A report from the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute suggested that relatives of cases with FL had higher incidence of NHL (Chatterjee *et al.* 2004), and the multi-centre EpiLymph study showed an association between family history of haematological malignancy and risk of all lymphoid malignancies (although not specifically with FL) (Casey *et al.* 2006).

However, an important report from the Swedish Cancer Registry (SCR) demonstrates that parental history of FL is associated with an increased risk of FL in offspring (Altieri *et al.* 2005). This was a cohort study surveying NHL cases diagnosed between 1993 and 2002 according to uniform histopathological criteria (Jaffe *et al.* 2001) from a pool of almost 7 million offspring and over 6 million parents included in the registry. Moreover, a subsequent study examining a larger cohort of FL cases (n = 2668) diagnosed between 1958 and 2004, and included in the (SCR), has demonstrated a fourfold increase of FL in parents and offspring (although not siblings) of FL cases as compared to control populations (Goldin *et al.* 2009). Consequently, these data suggest that a familial component does indeed contribute to risk of developing FL.

1.9 Single nucleotide polymorphisms (SNPs)

Studies have attempted to establish what heritable genetic determinants might confer the familial risk seen in FL. Two previous case reports of familial FL, one in a set of male monozygotic twins (Marco *et al.* 1999) and the other in a mother and daughter pair (Last *et al.* 2000), could not reveal any common genetic events. However, the

completion of the human genome project (Lander *et al.* 2001; Venter *et al.* 2001) has subsequently enabled germline genetic polymorphic variation between individuals to be examined in detail, particularly through the study of single nucleotide polymorphisms (SNPs).

SNPs are termed “single-base differences in the DNA sequence that can be observed between individuals in the population” (Roses 2000) and represent heritable variation in the human genome. Within SNPs, the base with the higher frequency across a population can be termed the “A” (major) allele and that with the lower frequency the “a” (minor) allele giving the three possible genotypes of AA, Aa and aa (**Figure 1.6**). Although SNPs may rarely involve more than two bases (by having alternative variant alleles) the vast majority involve only one major and one minor base, with this variation indicated as "A>a". SNPs therefore represent bi-allelic genomic markers and blocks of several SNPs that are consistently transmitted together can identify regions of the genome in high linkage disequilibrium (LD) termed haplotypes.

Over 3 million SNPs have so far been described (Frazer *et al.* 2007) by the international haplotype mapping project ("The International HapMap Project" 2003), a consortium dedicated to determining human genomic variation with a view to identifying genetic determinants associated with human disease. At least 10 million SNPs are predicted to actually occur, with an estimated frequency of around one SNP every 100 to 300 bp.

There is an increasing recognition that SNPs contribute to phenotypic differences in individuals. An example of this is the demonstration that the alleles of a single SNP (termed rs12913832) can differentiate blue and brown eye colour in a Northern European derived population (Sturm *et al.* 2008). This phenomenon appears to be mediated by the SNP minor allele (C) modulating DNA binding of a transcription factor at an intronic site 21 kb upstream of *OCA2* that contributes to the expression of this gene (which is known to be important in human pigmentation).

Moreover, a large study has examined the contribution of genome-wide germline genetic variation (using SNPs from the HapMap project) to variation in gene expression. It suggests that allele frequency differences may represent widespread functional variation in the human genome that can contribute to phenotypic

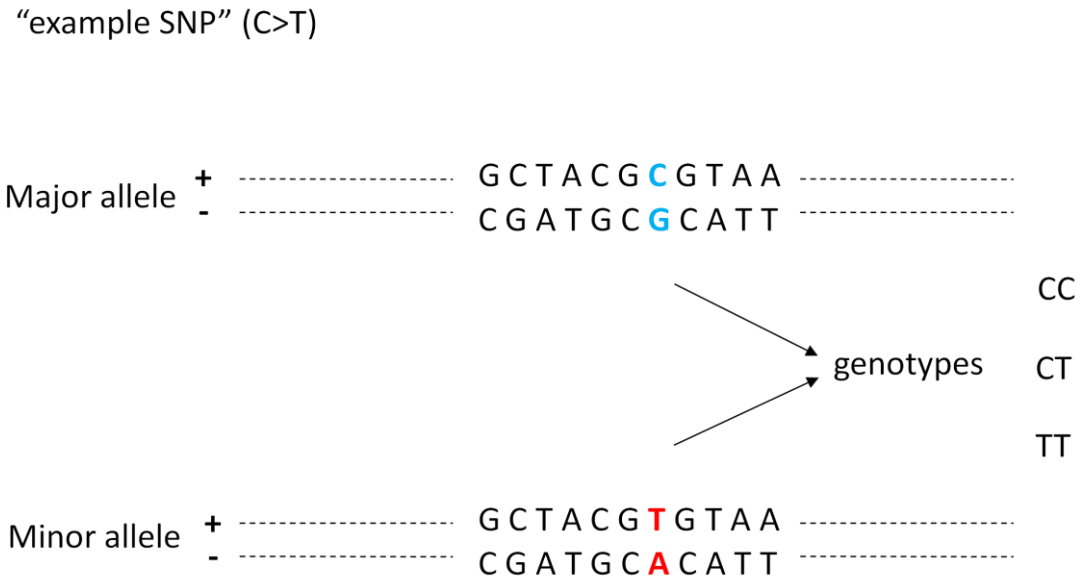


Figure 1.6 Single nucleotide polymorphism

Illustration of a DNA sequence including one SNP with base C forming the major allele (in blue) and base T the minor, or variant, allele (in red) as indicated. Examples of surrounding DNA base sequences are given for each allele with the dashed lines representing the continuing dsDNA sequence. The SNP is termed with reference to the sense (+) strands. The three possible genotypes arising from the SNP (equivalent to AA, Aa, aa as detailed in the text) are illustrated.

differences both within and between populations (Stranger *et al.* 2007b). Notably population dependent variation in SNPs (as observed across the four different populations that were included in the HapMap project) is an important factor to consider in their investigation, giving a requirement for appropriately matched control populations in association studies.

1.9.1 SNPs in FL

The functional associations of SNPs may also be relevant in disease by influencing predisposition or clinical outcome. Indeed, the observed familial associations with risk of FL may, at least in part, be determined by germline SNP variation. To date at least thirty studies have reported risk associations with SNPs in FL (Lossos *et al.* 2001; Moller *et al.* 2002; Lincz *et al.* 2003; Matsuo *et al.* 2004; Skibola *et al.* 2005a; Skibola *et al.* 2005b; Willett *et al.* 2005; Nieters *et al.* 2006; Rollinson *et al.* 2006; Shen *et al.* 2006; Spink *et al.* 2006; Wang *et al.* 2006a; Wang *et al.* 2006b; Cerhan *et al.* 2007a; Lan *et al.* 2007; Lee *et al.* 2007; Novik *et al.* 2007; Purdue *et al.* 2007; Cerhan *et al.* 2008; Skibola *et al.* 2008a; Skibola *et al.* 2008b; Hoffman *et al.* 2009; Ismail *et al.* 2009; Morton *et al.* 2009; Novak *et al.* 2009; Skibola *et al.* 2009; Bracci *et al.* 2010; Conde *et al.* 2010; Kane *et al.* 2010; Nieters *et al.* ; Wang *et al.* 2010) and many of these have targeted SNPs in immune related genes, reflecting the relevance of the immune response in FL. However, considerably less is known regarding the impact of SNPs on clinical outcome in FL with few associations reported so far (Weng and Levy 2003; Jardin *et al.* 2005; Cerhan *et al.* 2007b; Hohaus *et al.* 2007; Racila *et al.* 2008; Han *et al.* 2010) and these, together with the risk associations, require independent validation.

These studies have investigated small numbers of specific SNPs based on their association with genes which were either known or hypothesised to be relevant in FL. However, establishing the relevance of SNPs on a global genomic scale has benefited from a more systematic approach. This has been developed through the application of SNPs as genomic markers which, with the incorporation of specific probes for multiple different SNPs in micro-arrays, has led to the development of genome-wide association studies (GWAS). These studies have compared the allelic frequencies of hundreds of thousands of SNPs in case and control populations to

identify haplotypes and specific SNPs which associate with disease. Haplotypes can then be examined for specific genetic determinants of disease risk.

GWAS have identified associations in both non-malignant conditions (for example bipolar disorder, coronary artery disease, Crohn's disease, rheumatoid arthritis, type 1 and type 2 diabetes ("Wellcome Trust Case Control Consortium" 2007)) and various cancers, as recently reviewed by Stadler and colleagues (Stadler *et al.* 2010). Indeed, the first GWAS in FL has recently been reported and identifies polymorphic variation within the immune gene rich human leucocyte antigen (HLA) region on chromosome 6p21 as particularly important in FL (Skibola *et al.* 2009).

1.10 Human Leucocyte Antigen (HLA) System

The HLA region in humans, also termed the major histocompatibility complex (MHC) region, is associated with human disease. This region has been mapped ("The MHC sequencing consortium" 1999) and contains > 200 loci including the human leucocyte antigen (*HLA*) genes. These encode antigen presenting molecules which play a significant role in immune responses. Similar functional properties of these proteins enable their grouping as Class 1 (*HLA-A*, *B* and *C*) which present non-self antigen to CD8+ T lymphocytes (cytotoxic T lymphocytes; CTLs) and Class 2 (*HLA-DR*, *DQ* and *DP*) which present short antigen fragments to CD4+ T lymphocytes (helper T cells; T_H) (Klein and Sato 2000a). The *HLA* loci show a high degree of polymorphic variability with multiple allelic variants described for each locus and this led to the introduction of standardised nomenclature which utilises a four digit code to distinguish *HLA* alleles that differ in the proteins they encode ("Nomenclature for factors of the HLA system" 1988; Marsh *et al.* 2010). This diversity may affect antigen presentation and appears to be an important evolutionary phenomenon with differing alleles under selective pressure in different geographic regions (Parham and Ohta 1996) suggesting individuals possessing specific allele combinations gain a beneficial effect against deleterious events (such as infection).

1.10.1 *HLA* and disease

Specific *HLA* alleles predispose to a number of human diseases, particularly those with an autoimmune component or an infective basis. The former involve an

aberrant response to self antigens whereas clearance of infection typically requires a targeted response to non-self antigens. Differing *HLA* alleles may influence the presentation of both antigen groups and attendant modulations in immune responses may contribute to their association with disease risk. Therefore, perhaps not surprisingly, *HLA* associations with susceptibility to a wide range of auto-immune conditions as well as infections have been reported (Klein and Sato 2000b; Shiina *et al.* 2004; de Bakker *et al.* 2006). In certain instances an association with the clinical course of disease has also been attributed to HLA subtypes, such as progression of primary biliary cirrhosis (Nakamura *et al.* 2010). Furthermore, associations are also described with various tumours, indeed, a specific role for the HLA system in cancer predisposition has been recognised for some time with early reports indicating HLA antigen subsets associated with increased risk of nasopharyngeal carcinoma (NPC) in specific patient populations (Henderson *et al.* 1976) and over-representation of HLA-DR5 in renal cell carcinoma (RCC) (DeWolf *et al.* 1981). Since these reports, the advent of PCR based DNA typing in the 1990s has greatly facilitated the high resolution discrimination of *HLA* alleles, precipitating the demonstration of both Class 1 and Class 2 associations with several other cancers (Shiina *et al.* 2004; de Bakker *et al.* 2006).

1.10.2 *HLA* associations with lymphoma

Within lymphoid malignancies, HLA antigen subtypes were recognised as important in the risk of endemic Burkitt Lymphoma three decades ago (Jones *et al.* 1980; Jones *et al.* 1985) with a link to childhood ALL also proposed at that time (Von Fliedner *et al.* 1983). More recently, specific *HLA* allele associations have been demonstrated for risk of Hodgkin Lymphoma (*HLA-A*02*) (Niens *et al.* 2007) and DLBCL (*HLA-DRB1*) (Riemersma *et al.* 2006). Further study of MHC loci in NHL indicates both *HLA* and non-*HLA* loci are important in NHL. The *TNF* gene is a Class 3 MHC locus located on chromosome 6p21.3 between *HLA-B* and *HLA-DR*. The role of TNF in lymphoid development, inflammation and cancer led to a study demonstrating polymorphic variation (*TNF* G-308A) predicts disease outcome in NHL and DLBCL (but not FL) (Warzocha *et al.* 1998). The variant allele (A) confers an increased risk of all NHL as well as DLBCL (Rothman *et al.* 2006; Skibola *et al.* 2010) and marginal zone lymphoma (Skibola *et al.* 2010) subtypes. In addition to the relevance

of the *HLA* loci on NHL risk, their influence on disease outcome has been assessed with *HLA-DRB1*02* (Juszczynski *et al.* 2002) and a component of the AH 8.1 ancestral haplotype (which includes *HLA-A*01* and *HLA-B*08*) (Nowak *et al.* 2007) predicting both worse PFS and OS in NHL / DLBCL and NHL respectively (an effect independent of the IPI in a multivariate analyses). However, the role of the *HLA* loci in the risk and outcome of FL (including its transformation to DLBCL) has not been fully established.

1.10.3 rs6457327 and rs10484561

The GWAS performed by Skibola and colleagues (Skibola *et al.* 2009) identified a novel FL susceptibility locus within the MHC region on chromosome 6p21.33 confined to a 26kb region of high linkage disequilibrium containing a single coding region, chromosome 6 open reading frame 15 (*C6orf15*). The validation of findings in three additional, independent cohorts conferred robust statistical significance, with a combined allelic *P*-value of 4.7×10^{-11} across the entire study of 645 cases and 3377 controls for the association of this region with risk of FL. Two SNPs (rs6457327 and rs2517448) which are in complete linkage disequilibrium (LD) showed significant association with FL in the first phase of the study and subsequent genotyping of additional SNPs in their proximity delineated the 26kb risk locus. The sole gene within this locus, *C6orf15*, has also been termed *STG* following the initial identification of a homologue (*rmSTG*) expressed in simian taste buds (Neira *et al.* 2001). It is located just telomeric to the psoriasis susceptibility region 1 (*PSORS1*) but does not itself associate with this disease (Sanchez *et al.* 2004).

A further study has since expanded the initial GWAS through incorporation of a larger number of cases and performing analysis on non-pooled DNA in the first stage of the investigation (Conde *et al.* 2010). This confirmed the associations with FL from the initial study and identified that two further SNPs (rs10484561 and rs7755224) associate with risk of this disease. These are in complete LD and locate to chromosome 6p21.32, again in the MHC region, upstream and within 30kb of the *HLA-DQB1* locus. Higher resolution SNP typing of this region identified additional SNPs in complete LD with rs10484561 and rs7755224 within a 100kb region of high LD containing *HLA-DQB1* as well as *HLA-DQA1* just downstream of *HLA-DRB1*.

One of these SNPs (rs6457614) is a tag SNP for *HLA-DQB1*0501*. Typing of this and two additional tag SNPs (markers for the haplotype most frequently containing *HLA-DQB1*0501*) revealed that the *HLA-DRB1*0101 / HLA-DQB1*0501 / HLA-DQA1*0101 / rs104845561* haplotype associates with risk of FL ($P = 2.3 \times 10^{-4}$).

These two studies provide robust data as to the relevance of two segments within the HLA region that include rs6457327 or rs10484561, respectively, on chromosome 6p to risk of FL but the relevance of these loci to clinical outcome remains unknown.

1.11 The MDM2-p53 axis

Although SNPs rs6457327 and rs10484561 are strongly associated with FL risk, their functional relevance has not been determined. Conversely, SNPs influencing the MDM2-p53 axis, a critical component of cell-cycle control, have a role in several cancers but their contribution to FL is unknown.

The p53 protein is the product of the *TP53* gene and it maintains genomic integrity through initiation of DNA repair and acts as an inhibitory check-point to control cell-cycle progression and appropriately select cells for apoptosis (Levine *et al.* 1991; Ko and Prives 1996; Vogelstein *et al.* 2000; Vousden 2000; Ryan *et al.* 2001). It acts as a transcription factor, the functional unit being a tetramer (Farmer *et al.* 1992), and induces the expression of several genes which mediate these processes (Jin and Levine 2001). In recognition of its central role in preventing cell-cycle disruption, a hallmark of cancer (Elledge 1996), p53 has been described as the *"guardian of the genome"* (Lane 1992). Indeed, loss of its tumour suppressor activity has long been recognised in malignant cells (Hollstein *et al.* 1991).

Various cellular stresses (including, with particular relevance to cancer, DNA damage) activate p53 through at least three intra-cellular pathways (Vogelstein *et al.* 2000) which lead to stabilisation and activation of the protein (Giaccia and Kastan 1998). It then induces DNA repair mechanisms whilst the cell is held in growth arrest (Tanaka *et al.* 2000) or can induce apoptosis if the DNA damage is severe (Oda *et al.* 2000a; Oda *et al.* 2000b). Both processes thereby prevent the replication of damaged DNA in daughter cells and are therefore pivotal in the early recognition and elimination of potential tumorigenic events.

Not surprisingly, this maintenance of genome integrity is tightly regulated and MDM2, the human homologue of the Murine Double Minute 2 protein (Mdm2), is a principle inhibitor of p53 function (Oliner *et al.* 1992). MDM2 binds p53 (Momand *et al.* 1992) inducing transit of this complex from the nucleus into the cytoplasm (Freedman and Levine 1998; Roth *et al.* 1998; Lain *et al.* 1999; Tao and Levine 1999), so preventing the DNA binding necessary for its role as a transcription factor. Furthermore, MDM2 targets p53 for proteasomal degradation via its action as a protein ubiquitin ligase (Honda *et al.* 1997) and blocks p53 activation (Chen *et al.* 1993; Oliner *et al.* 1993; Picksley *et al.* 1994). More recently, the critical need for Mdm2 to suppress p53 has been demonstrated in mouse models, where those lacking Mdm2 had spontaneous p53 activity and developed rapid and irreversible widespread apoptosis in bone marrow and intestine leading to loss of weight, diarrhoea, aplastic anaemia and death (Ringshausen *et al.* 2006).

For optimal function p53 must escape the inhibitory action of MDM2, and this is possible through two processes. Firstly, phosphorylation of p53 by a series of protein kinases, including ATM (Banin *et al.* 1998; Canman *et al.* 1998), ATR (Canman *et al.* 1998), Chk1 (Shieh *et al.* 2000) and Chk2 (Hirao *et al.* 2000; Shieh *et al.* 2000) which form part of a molecular detection system for DNA damage, disrupts the binding site of MDM2. Secondly, other proteins (including Pim1) are recruited to form a p53 complex which induces a conformational change in p53 that further hinders MDM2 binding (Zacchi *et al.* 2002). However, p53 itself has a positive influence on MDM2 expression (Barak *et al.* 1993; Wu *et al.* 1993) providing an auto-regulatory negative feedback loop to prevent its uncontrolled activation.

This intricate homeostatic control of p53 function mediated through MDM2 yields low levels of p53 in the normal intra-cellular environment that can rise rapidly following cell stresses, which include genotoxic insults causing DNA damage. This has led to this role of MDM2 being succinctly summarised as, "*keeping p53 checkpoint function in check*" (Momand *et al.* 2000).

1.11.1 MDM2 in cancer

The importance of MDM2 homologues in tumour suppression has been demonstrated in animal models. Over-expression of Mdm2 in a mouse cell line

correlated with increased tumour formation (Fakharzadeh *et al.* 1991). Furthermore, a reduction of Mdm2 level by only a fifth in mice leads to a reduction in adenoma formation (Mendrysa *et al.* 2006) whereas increased expression leads to increased tumour formation (Jones *et al.* 1998). Human cancer cell lines may over-express MDM2 (Oliner *et al.* 1992) and one such cell line modulated to express a two fold increase in MDM2 shows reduced p53 function as compared to the unadulterated line (Ohkubo *et al.* 2006). Importantly, in primary human tumours MDM2 can be over-expressed and this may be associated with disease progression, resistance to therapy and survival (Cordon-Cardo *et al.* 1994; Freedman *et al.* 1999; Onel and Cordon-Cardo 2004).

The alterations in MDM2 levels (and their impact in cancer) may arise as a consequence of disruption of its genetic locus. In mice, *Mdm2* haplo-insufficiency leads to enhanced p53 function with increased sensitivity to DNA damage, lower transformation potential (Terzian *et al.* 2007) and delayed onset of lymphoma (Alt *et al.* 2003). In humans the *MDM2* gene is located on chromosome region 12q12-14 and amplification of this locus occurs in a range of tumours (Momand *et al.* 1998) and may be associated with increased expression of the MDM2 protein (Oliner *et al.* 1992; Leach *et al.* 1993; Reifenberger *et al.* 1993; Florenes *et al.* 1994) notably in cases without *TP53* mutations (Florenes *et al.* 1994). This is important in demonstrating the relevance of MDM2; as such mutations can prevent p53 feedback induction of MDM2 expression.

However, in one form of haematological malignancy at least (acute myeloid leukaemia), increased expression of MDM2 can occur without gene amplification suggesting alternative contributory mechanisms (Bueso-Ramos *et al.* 1993). It seems unlikely that mutation of the *MDM2* locus is such a mechanism as it is a rare event in cancer with only 3 single base pair substitutions identified in a series of 792 tumours in the Catalogue of Somatic Mutations in Cancer (COSMIC) database maintained by the Wellcome Trust Sanger Institute, Cambridge (Forbes *et al.* 2008), contrasting with the frequency of *TP53* mutations in cancer. As regards *MDM2* in FL, its locus is amplified in 52% of t-FL cases but not in the preceding FL counterparts (Hough *et al.* 2001) and MDM2 mRNA is increased between 10 and 45% in cases of t-FL compared to pre-transformation FL samples (Davies *et al.* 2005) suggesting that *MDM2* and its expression may be relevant in the transformation of FL.

1.11.2 *MDM2* SNP 309

These acquired genetic changes demonstrate a role for the MDM2 protein in cancer including, notably, FL. However, germline genomic attributes including polymorphic variation may influence gene expression (Stranger *et al.* 2007a; Stranger *et al.* 2007b; Shastry 2009) and a particularly important study by Bond and colleagues (Bond *et al.* 2004) identified that a SNP (T>G) located at nucleotide 309 in the first intron of *MDM2* occurs in a region utilised by p53 to activate *MDM2* transcription (Zauberman *et al.* 1995; Ries *et al.* 2000). The minor (G) allele was predicted to increase the length of a binding site for the transcription factor Sp1 and the study went on to demonstrate that Sp1 binding was enhanced by the presence of this variant. Furthermore, MDM2 levels correlated with the G allele and Sp1 binding at this site and a reduction in the stability of p53 was demonstrated in the presence of this allele (Bond *et al.* 2004).

This SNP, rs2279744, is termed *MDM2* SNP 309 and subsequent investigation revealed that the presence of the G allele associates with higher MDM2 protein levels and a lower apoptotic response than homozygosity for the T allele in surveys of cell lines (Bond *et al.* 2005) with homozygosity for the G allele associated with elevated MDM2 expression as compared to TT genotype (Hong *et al.* 2005). In the same year, the G allele was found to associate with inhibition of the transactivation of p53 dependent genes, as cell lines homozygous for this allele showed p53 binding to promoters of target genes but not transactivating them and this was due to MDM2 being co-associated with p53 (Arva *et al.* 2005). In human subjects *MDM2* SNP 309, is associated with worse prognosis in a variety of solid cancers including gastric carcinoma (Ohmiya, Taguchi et al. 2006) and oesophageal squamous cell carcinoma (Hong *et al.* 2005) where it is associated with increased risk of developing poorly differentiated carcinoma and advanced disease. Furthermore, a recent report has shown that homozygosity for the G allele predicts increased risk of death and earlier disease progression in the squamous cell carcinoma subgroup of oesophageal cancer (Cescon *et al.* 2009) and, significantly, the GG genotype correlates with earlier onset of *de novo* diffuse large B-cell lymphoma (DLBCL) in adult females (Bond *et al.* 2006). The *MDM2* SNP 309 G allele may also exert a dominant effect in tumour development as demonstrated in cases from families with the Li-Fraumeni syndrome

(who have heterozygous germline *TP53* mutations) in which presence of the G allele (as either TG or GG genotypes) predicted earlier age at first tumour development (Bond *et al.* 2004).

1.11.3 *TP53* in cancer

TP53 is the most common mutational target in cancer (Vogelstein *et al.* 2000) with over 26000 mutations described so far in the International Agency for Research on Cancer (IARC) *TP53* database (version R15, November 2010), available at website http://www-p53.iarc.fr (Petitjean *et al.* 2007). The *TP53* mutations typically translate in to deleterious amino acid changes and are often accompanied by loss of the second allele (Soussi *et al.* 2000). The multitude of acquired mutations described at the *TP53* locus (which also affect p53 function) can associate with poor outcome in a variety of solid tumours (Pricolo *et al.* 1996; Hardingham *et al.* 1998; Kandioler *et al.* 2002; Stander *et al.* 2004) and haematological cancers (Stirewalt *et al.* 2001) including DLBCL (Young *et al.* 2007). This reflects the importance of p53 disruption in cancer which is evident in FL where (as detailed earlier) *TP53* mutations confer a poor prognosis (O'Shea *et al.* 2008) and are more common at disease progression or transformation (Lo Coco *et al.* 1993; Sander *et al.* 1993; Davies *et al.* 2005).

1.11.4 *TP53* Arg72Pro

The frequency of *TP53* mutations in cancer and the importance of p53 in cell cycle control suggest that germline variation might influence p53 function and the development or progression of tumours. Indeed, one well characterised SNP, rs1042522 also termed *TP53* Arg72Pro, involves a non-synonymous substitution (G>C) at transcript nucleotide 466 of exon 4 in *TP53* (Matlashewski *et al.* 1987). This effects an amino acid change (arginine to proline) at residue 72, which modifies the functional properties of the p53 protein by conferring reduced potential to induce apoptosis or suppress cell transformation (Thomas *et al.* 1999; Marin *et al.* 2000; Dumont *et al.* 2003), and has been associated with increased risk of solid tumours including lung cancer (Zhang *et al.* 2006) in which a multiplicative effect with *MDM2* SNP 309 is observed. Another report identifies that an earlier onset of tumours in patients who carry a Li-Fraumeni germline *TP53* mutation occurs in cases

with the *MDM2* SNP 309 minor allele (G) and this effect may actually be amplified by presence of the *TP5*3 Arg72Pro major allele (G) (Bougeard *et al.* 2006). Reports also suggest the SNP may be important in the clinical course of cancer including cases of squamous cell lung carcinoma in which homozygosity for the minor allele (C) is associated with worse outcome (Wang *et al.* 1999).

However, despite the correlation of MDM2 expression and *TP53* mutation with FL outcome and the association of both *MDM2* SNP 309 and *TP53* Arg72Pro in risk and outcome of other cancers, the role of these SNPs in FL remains unknown.

1.12 Aims and objectives

In light of the familial influence and diverse cytogenetic events, this project addressed the contribution of inherited and acquired genetic factors to FL and its transformation.

The first phase of study aimed to assess the contribution of germline polymorphic variation influencing the MDM2-p53 axis (through *MDM2* SNP 309 and *TP53* Arg72Pro) and occurring in the HLA region (at rs6457327 and rs10484561), to FL and its transformation. The second phase set out to identify important acquired genetic factors influencing transformation and the third phase sought to model the evolutionary basis of transformation. Specific objectives were the clinical correlation of the SNP genotypes in germline DNA from a large series of FL cases, the high resolution genetic profiling of a series of sequential FL and t-FL samples and the subsequent molecular mapping of FL / t-FL evolution.

Chapter 2 Materials and Methods

Current regulations concerning health and safety at work were adhered to throughout the course of this study.

2.1 Ethical Considerations

Ethical approval was gained from East London and the City Health Authority (HA) Local Research Ethics Committee (LREC) with the following approval numbers:

> 05/Q0605/140 "The impact of the tissue microenvironment and immune system on haematological malignancies" (with a substantial amendment in 2007 to include: t(14;18) monitoring and genetic analysis of haematological malignancies); and
>
> 06/Q0605/69 "Clonal evolution of indolent and aggressive lymphoma arises from a common B-cell ancestor".

The clinical information used during the course of the study was available from the clinical database maintained within the Department of Medical-Oncology at Barts and the London NHS Trust. To preserve patient confidentiality, each patient was assigned a unique identifier code under which relevant clinical information was supplied by the Data Manager. Only the database management team could access information that might identify patients.

2.2 Patient Samples

Relevant patient identifier codes were then used by the Tissue Storage Manager to identify appropriate and available patient samples for the study. These included those involved by lymphoma and were typically obtained from lymph node biopsies but occasionally from biopsies of skin, spleen or breast. Further patient samples not involved by lymphoma represented patient germline samples and were obtained from biopsies of bone marrow, peripheral blood, saliva, tonsil and (reactive) lymph nodes not involved by lymphoma. All study patients were required to have a diagnosis of FL and a cohort who all subsequently developed transformation to DLBCL and had

sequential samples available were also included. These diagnoses were confirmed prior to entry on the clinical database by expert histopathologists (Dr A Norton, Dr M Calaminici and Dr H Rizvi) through a weekly multi-disciplinary team meeting (MDM) held at Barts and the London NHS Trust.

2.3 Cell lines

Cell lines analysed included B-NHL, EBV transformed B-cell lymphoblastoid and myeloid cell lines which were obtained from Cell Services, Cancer Research UK (CRUK), Clare Hall Laboratories, South Mimms, Herts, UK. They were maintained as suspensions in RPMI-1640® (Gibco) nutrient medium plus 10% foetal calf serum and 1% penicillin / streptomycin at 37°C in sterile culture flasks in a humidified incubator with 5% CO_2 (CB150; Binder, D-78532, Tuttingen, Germany). Cell concentration and percentage viability was determined using Trypan Blue 0.4% (Sigma) staining and manual cell counting on a haemocytometer under phase microscopy (Wilovert®, Helmut-Hund GmBH, D-6630 Wetzlar 21, Germany) prior to passage every 48 to 72 hours.

2.4 DNA

2.4.1 Genomic DNA extraction

Patient DNA was obtained from tissue biopsy samples, cell suspensions, peripheral blood, bone marrow aspirates or saliva. Whole tissue from biopsy samples was used to make the cell suspensions prior to being snap frozen on day of biopsy procedure and stored in liquid nitrogen. Subsequently, the frozen tissue was removed and powdered using a pestle and mortar or frozen cell suspensions thawed at 37°C and centrifuged to obtain a white cell pellet. For peripheral blood and bone marrow samples, red cell lysis was initially performed through the addition of 10× volume of ddH_2O prior to pelleting by centrifugation. The powdered sample or pellets were then incubated overnight at 55°C in digestion buffer (1mL per ~1×10^8 cells) with 0.1mg/mL proteinase K. DNA was then extracted using the phenol / chloroform method (Gross-Bellard *et al.* 1973; Albarino and Romanowski 1994). Briefly, an equal volume of phenol was added to the digestion mix, spun at 14000rpm for 10

minutes at 4°C and the top (aqueous) layer removed. This was added to an equal volume of 50:50 phenol / chloroform mix, spun again, the aqueous layer removed and added to an equal volume of chloroform and spun once more. The aqueous layer was the removed and DNA precipitated by adding 1/10 volume 3M sodium acetate (pH 5.2) and 2× volume 100% ethanol. The DNA was pelleted by centrifugation before being washed in 70% ethanol and air dried prior to re-suspension in TE buffer and storage at 4°C. In certain cases, saliva was collected as a source of germline (non-tumour) DNA using the Oragene DNA collection kit (DNA Genotek Inc., Ottawa, ON, Canada) according to the manufacturer's instructions.

2.4.2 DNA quantity and quality determination

These stock solutions of DNA were used to prepare working concentrations of DNA. To determine adequate quantity and quality of DNA, the light absorbance of stock solutions relative to the suspension medium (TE or tris) was determined using a NanoDrop® ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The sample absorbance of light at a wavelength of 260nm generates an optical density (OD) value (OD_{260}) which was multiplied by a factor of 50 to establish DNA quantity (ng/µl). OD at 280nm (OD_{280}) was used to generate the OD_{260}/OD_{280} ratio with values 1.80 - 2.00 indicating high purity (low protein contamination) of the DNA sample. A further measure of OD at 230nm (OD_{230}) enabled the OD_{260}/OD_{230} ratio to be established as a secondary measure of DNA quality (with high purity values > 2.0).

2.4.3 Conventional (qualitative) polymerase chain reaction

Conventional polymerase chain reaction (PCR) was performed in a volume of 20 to 50µl in thin walled MicroAmp 0.2ml PCR tubes (Applied Biosystems, Foster City, CA, USA) or Thermo-Start 24× and 96× 0.2ml well PCR plates (Thermo Scientific) on a PTC-225 Peltier Thermal Cycler (MJ Research, San Francisco, CA, USA) over 30 to 40 cycles (with more amplification cycles performed particularly if DNA template levels were low). Reaction mixes were assembled on ice using dedicated work spaces, pipettes and filter tips. Each mix typically contained 1× PCR buffer with 1.5mM $MgCl_2$, 1 unit of DNA *Taq* polymerase (Promega), deoxyribonucleotide

triphosphates (dNTPs) at 500µM, forward and reverse primers at 400 - 800nM (concentration determined after primer optimisation) and 10 to 50ng DNA. Cycling conditions consisted of an initial denaturation step at 94°C for 3 minutes followed by cycles of 94°C (denaturation), 54-66°C (primer annealing) and 72°C (extension) each for 30 - 60 seconds and a final step of at 72°C for 10 minutes to ensure full extension and deoxynucleotide addition at the 3'-terminal A overhang.

Prior to each aspect of the study, specific reaction concentrations and annealing temperatures were optimised for respective primer combinations and the predicted amplicon length for each PCR reaction determined the extension time (based on a *Taq* polymerase rate of approximately 1000 bp per minute). Primers used are detailed in ***Appendix* 2 Table 1**.

In the latter stages of the study, the HotStarTaq *Plus* DNA polymerase kit (Qiagen, Valencia, CA, USA) was used. This is a master-mix (consisting of buffer, hot start DNA polymerase, dNTPs) to which are added the respective primers and DNA template as well as a gel loading dye. This facilitated PCR on lower DNA concentrations, which helped to maximise use of samples with limited DNA stock.

Success of the PCR reactions was assessed through agarose gel electrophoresis of the PCR products followed by inspection of photographs of the gels (as described in section **2.4.8** below).

2.4.4 Allele specific oligonucleotide primers and probes

The study required allele specific oligonucleotide (ASO) primers and probes for PCR mediated detection of acquired somatic hypermutation (SHM) sites in the immunoglobulin heavy chain variable (*IGH-V*) region, or DNA mutations, in certain FL and t-FL samples. Detailed inspection of *IGH-V* sequences was performed in order to identify combinations of SHM sites which were best suited as templates for primer design. Several ASOs were designed manually (using the general principles detailed by Lowe and colleagues (Lowe *et al.* 1990)) for each target and, after optimisation, the most specific (lack of non-specific amplification) and sensitive (most easily visualised PCR product) combinations were preferentially used in the study.

Fluorescently labelled ssDNA oligonucleotide probes (which bind the DNA segment between primers) facilitated SNP genotyping (section **2.4.5**) or DNA quantitation (section **2.4.7**) and were designed using Primer Express® software v3.0 (Applied Biosystems).

2.4.5 Allelic discrimination (AD) PCR

A real-time PCR allelic discrimination multiplexed endpoint assay (AD-PCR) with the ABI PRISM™ 7700 or 7900HT Real-Time PCR Sequence Detection Systems (Applied Biosystems) was used to determine germline SNP genotypes. Reactions were performed at least in duplicate on a MicroAmp® Optical 96 well (0.1ml) reaction plate with three positive controls (one for each possible SNP genotype) and a no DNA template control included in each AD-PCR run. Each reaction included either:

1× TaqMan® Universal PCR Master Mix (Applied Biosystems) [which incorporates the hot-start AmpliTaq Gold® DNA polymerase to limit non-specific DNA amplification, optimized buffer, dNTPs and a passive internal fluorescence reference (ROX™ dye)], 10ng DNA and the respective oligonucleotide primers (at 300 to 900 nM) and fluorogenic probes (at 150 to 200 nM) in a final volume of 25µl; or

custom made Taqman® Genotyping Assays (incorporating primers and probes) and 1× Taqman® Genotyping Master Mix (Applied Biosystems) [which is a more recent and further optimized master mix specific for SNP genotyping], with 2 to 10 ng of genomic DNA in a final volume of 10µl.

Probes specific to each SNP allele were labeled at the 5' end with either 6-carboxyfluorescein (FAM™) or proprietary VIC® fluorescent reporter dyes and at the 3' end with tetramethyl-6-carboxyrhodamine (TAMRA™), a non-fluorescent quencher. The 5' nuclease activity of the DNA polymerase resulted in cleavage of the probe during PCR elongation causing separation of the 5' reporter dye from the 3' quencher. This generated an increase in fluorescence from the reporter dye specific to its respective SNP allele. Cycling conditions consisted of an initial AmpliTaq Gold® activation step of 95°C for 10 minutes then 40 cycles of 95°C

(denaturation) for 15 seconds followed by 60°C (primer annealing and extension) for 60 seconds.

Interpretation of reaction end-point AD-PCR results was performed using the Sequence Detection Systems (SDS) software version 2.3 (Applied Biosystems) with real-time PCR curves generated during the AD-PCR runs used to determine genotypes in the cases deemed undetermined by end-point analysis.

2.4.6 Control samples for establishing SNP risk associations

To examine risk associations between SNPs and FL in the United Kingdom (UK) population, a case-control analysis was performed. Individuals with available genotype data from the Wellcome Trust Case Control Consortium 2 (WTCCC2) 1958 birth cohort (BC58) formed control populations for the respective SNPs. Further details regarding the WTCCC are available on the website http://www.wtccc.org.uk.

The BC58 cohort includes blood samples obtained from 3000 individual volunteers. The (germline) DNA from these samples had been applied to two genome-wide SNP micro-array chips, the Affymetrix SNP 6.0 (Affymetrix, Santa Clara, CA) and Illumina 1.2M (Illumina, San Diego, CA) platforms, enabling the genotypes of over one million SNPs to be determined for each case. These individuals were originally included in the second of four Perinatal Mortality Surveys (the others occurring in 1946, 1970 and 2000/01) which enrolled all the people born in the week 3 - 9 March 1958 in England, Scotland and Wales and was designed initially to examine social and obstetric factors associated with still birth and neonatal death, with its first results reported five decades ago (Butler 1961). They remain in long-term follow-up as part of the larger longitudinal National Child Development Study (NCDS) initiative and were approached between 2002 and 2004 in a collaboration funded by the Medical Research Council (MRC), involving the second project of the WTCCC (WTCCC2), for blood sampling and storage of their DNA for analysis. The individuals who provided these samples form the WTCCC2 BC58 cohort.

For the current study, filtering was applied to the WTCCC2 BC58 results to enable control populations to be constructed for each target SNP. Cases in which the

respective SNP genotypes were not determined were excluded, furthermore individuals who were population outliers or showed significant relatedness were also excluded. These groups are only truly representative of the UK population born in 1958 who stayed alive until 2002 - 2004, but they do represent a valid and close approximation to the current UK population. Importantly, the NCDS cohorts were augmented during the course of the longitudinal studies with immigrants who were born in the same time periods and so at least in part reflect the differing ethnic / racial distributions present in the UK. (This is significant as differing SNPs can have marked differences in allele frequencies within different human populations, as described in Chapter 1.)

2.4.7 Quantitative PCR

Real-time quantitative PCR (qPCR) was employed for the determination of t-FL DNA levels in preceding FL tumour samples. Reactions were performed in triplicate on a MicroAmp® plate in a final volume of 25µl including TaqMan® Universal PCR Master Mix (Applied Biosystems), *IGH-VH* region primers (400nM) and 5' FAM™ labelled probe (200nM) with 500ng DNA. A control gene (β_2M) was used to normalise the level of amplification and was assayed in parallel but using 240nM of primers and 80nM of probe. Reaction conditions were as indicated in **Table 2.1**. Working FL and t-FL DNA sample concentrations were 100ng/µl with ten-fold dilutions (10^{-1} to 10^{-5}) of the t-FL samples prepared in pooled peripheral blood mononuclear cell (MNC) DNA derived from eight healthy individuals to enable estimation of the level of any t-FL detected in the FL sample. Reactions were performed on the same 7700 / 7900HT systems as used for AD-PCR with results interpreted using SDS v.2.3 software.

ASOs (section **2.4.4**) were designed on SHM sites that were unique to the respective FL and t-FL major clones identified by homo-duplex band analysis (as described in section **2.4.9** below). A median of 5 different primers was designed for each t-FL major *IGH-VH* clone sequence. Sensitivity of different primer / probe combinations was determined by analysis of serial ten-fold dilutions (from 10^{-2} to 10^{-5}) of t-FL DNA diluted in MNC DNA. Only primer / probe combinations optimised to a detection sensitivity of 10^{-4} were used, with two different combinations applied in

Table 2.1 qPCR conditions

Temperature* (°C)	Time	Cycles[#]
50	2 min	
95	10 min	
95	15 s	×50
58	1 min	

[#]To increase the resolution for detection of low level targets, amplification was performed over 50 cycles (as compared to 40 cycles which are typically used in qPCR).

Table 2.2 12% polyacrylamide gel

Constituent	Volume (ml)
30% Acrylamide (29:1 acrylamide:bisacrylamide)	40
H_2O	38.33
5× TBE	20
10% ammonium persulphate	1.67
*TEMED**	*0.083*
Total	100

*TEMED added just prior to pouring of the gel to facilitate acrylamide polymerisation.

separate reactions for each FL / t-FL pair analysed. Assays were performed in triplicate and specificity of reactions was indicated by no amplification in normal MNC derived DNA.

2.4.8 Agarose gel electrophoresis

DNA fragments were separated by agarose gel electrophoresis. Gels were made using agarose (Life Technologies, Paisley) dissolved in 1× TBE buffer to 2% by microwave melting with the volume made determined by the number of wells needed. After partial cooling either ethidium bromide (final concentration 0.5 μg/ml) or 1× GelRed (Biotium, Hayward, CA) solutions were added to the gel. The gel mix was poured into appropriately sized gel cast trays with combs applied to create sample wells. Samples mixed with 1× gel loading buffer were added to individual wells with either 1 kb (Gibco, Paisley, UK) or 100 bp (Promega, Madison, WI, USA) DNA ladders in a separate well as a size reference. Electrophoresis was performed with 2 to 10 Volts per cm applied across the gel in 1× TBE running buffer within an electrophoresis tank (Jencons Scientific Limited, Leighton Buzzard, UK). DNA was visualised by transillumination under ultraviolet (UV) light and photographed.

2.4.9 Homo-duplex analysis

For DNA homo- and hetero-duplex formation (to determine the major clones by *IGH-V* SHM pattern within FL and t-FL samples), PCR products were incubated at 94°C for 5 minutes (denaturation) then 50°C for 1 hour (prolonged re-annealing). Following this, samples were promptly placed on ice, mixed with 1× gel loading buffer and applied to individual wells of a non-denaturing 12% polyacrylamide gel (made as indicated in **Table 2.2** on the preceding page) with the 1 kb DNA ladder included as a size reference.

Electrophoresis was then performed with 20mA applied to the gel in TBE running buffer for 14 hours at 4°C in a DV-2020 vertical electrophoresis tank (Scotlab Ltd., Luton, UK). The gel was then immersed for 30 minutes in TBE buffer containing ethidium bromide (2.5μg/ml) so that homo- and hetero-duplex bands could be viewed under UV transillumination. Small fragments (approximately 2 × 4 mm in

surface dimensions) containing the homo-duplex bands (representing the major tumour clones) were then excised from the gel and placed in separate 1.5ml micro-centrifuge tubes prior to extraction of DNA.

2.4.10 Gel Extraction

DNA was extracted from gel fragments using the QIAquick® Gel Extraction Kit (Qiagen, Düsseldorf, Germany) followed according to the manufacturer's instructions. This system was used as it recovers purified DNA ≥70 bp and none of the predicted homo-duplex DNA complexes were smaller than this in size. The DNA products were then re-amplified by PCR, visualised through agarose gel electrophoresis and directly sequenced (following the method described below in section **2.4.12**).

2.4.11 Cloning

For characterisation of differing *IGH-V* SHM patterns in tumour sub-clones, PCR products were cloned. The TOPO TA Cloning® system (Invitrogen, Paisley, UK) was applied according to the manufacturer's protocol. The technology utilises a non-template-dependent terminal transferase activity of Taq polymerase that adds a single deoxyadenosine (A) to the 3' ends of PCR products. Topoisomerase I generates specific breaks in the dsDNA of the plasmid vector creating single overhanging 3' deoxythymidine (T) residues which then allows PCR products to ligate with the vector. Consequently, single PCR amplicons can be incorporated in to different plasmid vectors which are then transformed into *E. coli*. Following bacterial colonisation, individual colonies contain multiple copies of the DNA sequence (established from a single PCR amplicon). These can be amplified by PCR and their sequences compared to investigate any heterogeneity within the initial PCR product.

A 4µl aliquot of fresh PCR product was incubated with 1µl salt solution (200 mM NaCl and 10mM $MgCl_2$) and 1µl pCR®2.1-TOPO® vector at room temperature for 5 minutes. A 2µl aliquot of this reaction was then added to a vial of One Shot® chemically competent *E. coli* cells and incubated on ice for 30 minutes. The cells were then "heat-shocked" for 30 seconds at 42°C, placed on ice and 250µl of S.O.C. medium added before shaking the tube at 37°C for one hour. Either 50µl or 100µl

from each transformation was spread on pre-warmed agar plates containing ampicillin (50mg/mL) and X-gal (20mg/mL) with colonisation then performed by incubation at 37°C overnight. The pCR®2.1-TOPO® vector contains an ampicillin resistance motif facilitating ampicillin mediated selection of only *E. coli* successfully transformed with the vector; it also contains a *LacZα* motif coding for β-galactosidase (which cleaves X-gal creating a blue product) and M13 forward and reverse primer sites. The vector's cloning site lies within the *LacZα* motif so successful incorporation of a DNA product disrupts transcription of this locus leading to white colonies. Individual white colonies were then directly used as DNA templates for PCR amplification using M13 forward (CATTTTGCTGCCGGTC) and M13 reverse (GTCCTTTGTCGATACTG) primers.

2.4.12 Direct sequencing

PCR products were directly sequenced using the dideoxy chain termination method (Sanger *et al.* 1977). Oligonucleotide primers and unincorporated deoxynucleotides were firstly removed by clean-up using the Montage® spin filter column protocol (Millipore, Billerica, MA, USA) according to the manufacturer's instructions or through an exonuclease I and shrimp alkaline phosphatase (SAP) protocol (ExoSap), as detailed in **Table 2.3**.

Sequencing was then performed in a reaction volume of 10 or 20µl including BigDye Terminator® Master Mix version 3.1 (Applied Biosystems) [which incorporates buffer, fluorescently labelled DNA chain terminators, deoxynucleotides and AmpliTaq DNA polymerase], the forward or reverse primer from the preceding PCR at a final concentration of 160nM and 1 to 2µl of cleaned PCR product. Reaction conditions included an initial enzyme activation step at 96°C for 1 minute, then 26 cycles of 96°C for 30 seconds (denaturation), 50°C for 15 seconds (annealing) and 60°C for 4 minutes (extension).

Unincorporated dye terminators were removed using the DyeEx™ 2.0 Spin column kit (Qiagen), as per the manufacturer's instructions, and each reaction then vacuum dried in a DNA 120 SpeedVac® Concentrator (Savant Instruments Inc, Farmingdale, NY, USA). The dried DNA products were stored at -20°C and protected from light until their transfer for sequence reading by capillary electrophoresis on a 3730xl

Table 2.3 ExoSap protocol

Reagents for ExoSap master-mix	**Volume (µL)**	
ddH2O	179	
Exonuclease (20U/µl)	1	
SAP (1U/µl)	20	
Total	*200*	
Reaction mix	**Volume (µL)**	
ExoSap master-mix	2	
PCR product	2.5	
Total	*4.5*	
Conditions	**Temperature**	**Time**
Reaction	37°C	15 minutes
Enzyme inactivation	80°C	15 minutes
	4°C	hold

DNA Analyser (Applied Biosystems) in a service provided by Dr Graham Clark, Equipment Park, Cancer Research UK (CRUK), 44 Lincoln's Inn Fields, London. When high volume sequencing was performed, sequence reaction products were left in solution on 96 well plates and magnetic bead capture of DNA was performed by the Equipment Park using the Biomek® FX liquid handling robot (Beckman Coulter, Brea, CA, USA) immediately prior to sequence reading.

Sequence results were analysed and interpreted using the Lasergene® Editseq and SeqMan® software packages (DNAStar Inc., Madison, WI, USA) in combination with 4Peaks software (available at http://mekentosj.com/science/4peaks/). Visual comparison of sequences was performed against reference sequences obtained from the human genome release 18 on the Ensembl genome browser (available at http://www.ensembl.org). For analysis of *IGH-V* region sequences, comparisons were made to reference sequences obtained from the ImMunoGeneTics IMGT/V-QUEST system (http://imgt.cines.fr).

2.5 Genome-wide DNA copy number assay

A detailed protocol for the assay was downloaded from www.affymetrix.com. Two protocols are available for the assay (Genome-Wide SNP 6.0 Assay for 48 or 96 Samples and Affymetrix® Cytogenetics Copy Number Assay) the difference between the two being the number of samples processed in parallel and the method used to pool and purify PCR products. The Cytogenetics Copy Number Assay was followed in this study as it is tailored to lower volume sample throughput and determination of copy number changes (as opposed to the former protocol which is preferably used in high-throughput genome-wide association studies [GWAS]).

2.5.1 Genomic DNA preparation

To ensure success of the assay, high quality genomic DNA free from potential inhibitors of enzymatic processes and of sufficient quantity (500ng) to ensure an adequate level of PCR amplification was required.

All DNA samples were subject to electrophoresis on 2% agarose gels. Samples that failed to reveal a band of approximately 10 to 20 kb in size (representing intact, high

quality genomic DNA) or those that showed smearing (indicating DNA degradation) were excluded at this stage as indicted in **Figure 2.1**.

Concentrations of stock DNA samples were determined by the Nanodrop® system and working dilutions of each were prepared by addition of the appropriate volume of ddH_2O to achieve DNA concentrations of 50ng/μL in a volume of at least 10μL. High DNA quality in working dilutions was ensured by OD_{260}/OD_{280} ratios of ~1.80.

The DNA preparation, reagent storage and reaction mix preparation for the restriction enzyme digest, adaptor ligation and PCR steps of the assay occurred in a pre-PCR room to minimise risk of cross-contamination of genomic DNA by PCR products. The digest and ligation reactions were performed on a dedicated pre-PCR PTC-220 (DYAD™) Peltier Thermal Cycler (MJ Research). All reagents used in enzymatic stages were thawed on ice and the respective master-mixes prepared on ice. Although the protocol could be interrupted with storage of samples at -20°C after the digestion, ligation, PCR, pooling / purification and labelling steps this was avoided wherever possible, as is recommended by the manufacturer.

2.5.2 Restriction enzyme digestion

Master-mixes for two restriction enzymes, NspI and StyI (**Table 2.4**), were prepared on ice. For each sample, two separate DNA aliquots of 250ng in 5μL were prepared in parallel by addition of 14.75μL of one or other master-mix and mixed gently. The preparations were incubated at 37°C for 2 hours then 65°C for 20 minutes.

2.5.3 Adaptor Ligation

Ligation master-mixes were prepared on ice for NspI and StyI (**Table 2.5**) and 5.25μL of the respective mix was added to the digested samples and mixed by pipetting. The 25μL reaction volume was incubated at 16°C for 3 hours then 70°C for 20 minutes.

Immediately following this step, each DNA ligation sample was diluted by addition of 75μL ddH_2O giving a sample volume of 100μL.

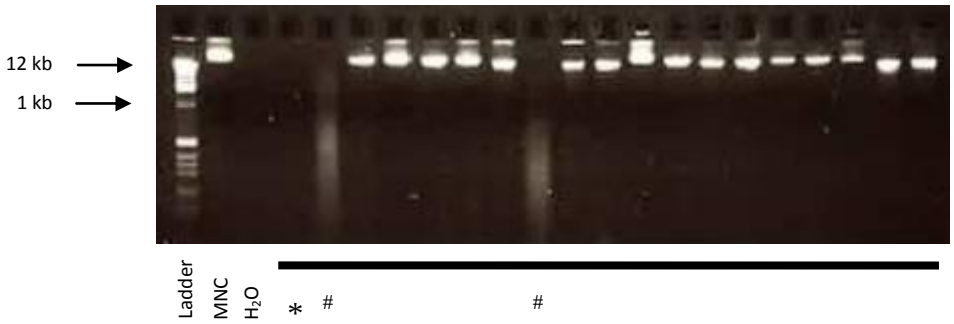


Figure 2.1 Genomic DNA quality assessment by agarose gel electrophoresis

Patient DNA samples (n = 19) are indicated by the black bar. The majority (n = 16) show a band equivalent to molecular size >12 kb. Three samples show suboptimal DNA quality. One sample (*) had no detectable intact DNA and two other samples ($^{\#}$) showed smears indicating DNA degradation.

Table 2.4 Digestion master-mix for NspI or StyI

Reagents	**Volume (μl)**
NEbuffer 2 or NEbuffer 3 (10×)	2
BSA (100×, 10mg/ml)	0.2
NspI or StyI (10U/μl)	1
H_2O	11.55
Total	14.75

Table 2.5 Ligation master-mix for NspI or StyI

Reagents	**Volume (μl)**
NspI adaptor or StyI adaptor (50μM)	0.75
T4 DNA ligase buffer	2.5
T4 DNA ligase (400U/μl)	2
Total	5.25

2.5.4 PCR

PCR reactions were performed in quadruplicate for the NspI and in triplicate for the StyI diluted ligated samples. PCR master-mix was prepared on ice as described in **Table 2.6**. A 10µl aliquot of the diluted ligated DNA sample from either the NspI or StyI processes was added to a fresh tube and mixed with 90µl of the appropriate PCR master-mix. The seven reaction mixes gave a combined PCR reaction volume of 700µl for each DNA sample which was designed to ensure sufficient PCR product was generated to enable adequate DNA hybridisation to each array. Reactions were performed on a PTC-225 (Tetrad™) Peltier Thermal Cycler (MJ Research) under the conditions given in **Table 2.7**.

To ensure consistent PCR results, a 3µl aliquot from each PCR reaction was run on a 2% agarose gel at 120V for 30 minutes. Verification that the average product distribution was between 250 to 1100 bp was then performed by comparison with DNA ladder size markers on inspection of a gel photograph (as illustrated in **Figure 2.2)**.

2.5.5 PCR product pooling and purification

The PCR products from the seven reactions performed for each DNA sample were pooled together in one capped 2.0mL micro-centrifuge tube. A 100µl aliquot of AMPure® magnetic beads (Agencourt Bioscience, Beverly, MA, USA) was added to each pooled 700µl sample and mixed by gentle inversion 10 times. Following incubation at room temperature for 10 minutes, the samples were centrifuged at 16100rcf for 3 minutes and then each tube was placed in a DynaMag™-2 magnetic stand (Invitrogen). The supernatant was carefully removed without disturbing the bead pellet which formed.

Pellets were washed by the addition of 1.5ml 70% EtOH to each tube which was then shaken at 2000rpm for 2 minutes in a MixMate desktop shaker (Eppendorf, Hamburg, Germany). The sample tubes were then re-applied to the magnetic stand and the supernatant removed without disturbing the bead pellet. Each tube was centrifuged at 16100rcf for 30 seconds, placed back in the magnetic stand and the remaining small volume of liquid removed from the base of the tube without

Table 2.6 PCR master-mix

Reagents	**Volume (µl)**
Water	39.5
Titanium Taq PCR buffer (10×) (Clontech)	10
dNTP (10mM)	14
Betaine (5M)	20
Primer 002 (100µM)	4.5
Titanium Taq (50×) (Clontech)	2
Total	90

Table 2.7 PCR conditions

Temperature (°C)	**Time**	**Cycles**
94	3 min	
94	30 s	×30
60	30 s	
68	15 s	
68	7 min	
4	*hold overnight*	

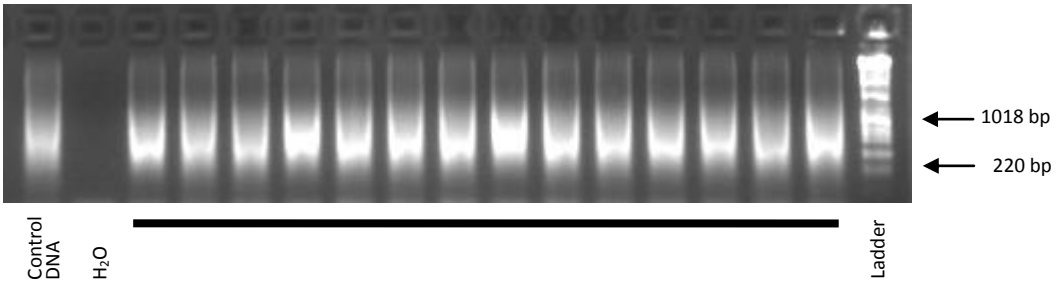


Figure 2.2 PCR product visualization

An example demonstrating that PCR product distribution was between 250 to 1100 bp. The black bar indicates test sample PCR products.

disturbing the pellet. The tubes were then left uncapped at room temperature for 15 minutes to enable evaporation of any residual EtOH.

DNA was eluted from the magnetic bead by addition of 55µL Tris-buffer to each tube which were then vortexed for 1 minute and shaken for 10 minutes at 2000rpm on the MixMate to re-suspend the beads. Following this, the tubes were centrifuged at 16100rcf for 30 minutes and then placed in the magnetic stand for 5 minutes. Taking care not to disturb the bead pellet, 47µL of the eluted sample was removed.

2.5.6 Quantitation

A 2µL aliquot of each purified sample was added to 18µL of ddH_2O in separate wells of a 96 0.2mL well plate and mixed by pipetting, giving a 1/10 dilution. Sample quality and quantity were determined using the Nanodrop® system. An OD_{260}/OD_{280} value between 1.8 and 2.0 was acceptable and a DNA yield of 450 to 700 ng/µL was required (representing a concentration in the purified sample of 4.5 to 7 µg/µL).

2.5.7 Fragmentation

A 5µl aliquot of 10× fragmentation buffer was added to each purified sample. Care was taken to minimise handling of the fragmentation enzyme (DNase I) tube in view of its extreme temperature sensitivity during preparation of the fragmentation master-mix. Enzyme concentration could vary by batch so the volumes of DNase I and H_2O in the master-mix were adjusted to give a final DNase I concentration of 0.1 U/µL as indicated in **Table 2.8**. The water and fragmentation buffer were mixed and allowed to cool on ice for 5 minutes prior to addition of the enzyme. After briefly vortexing (3 × 1 second) followed by spin down for 30 seconds, 5µL of the master-mix was added to the sample with fragmentation buffer giving a reaction volume of 55µL. These samples were incubated at 37°C for 35 minutes then 95°C for 15 minutes to terminate the enzyme activity. The labelling step was immediately commenced and, while the labelling reaction was underway, 2µl of each fragmentation reaction were run on 2% agarose gel by electrophoresis at 120V for 20 minutes to confirm fragment sizes were < 180 bp (**Figure 2.3**).

Table 2.8 Fragmentation master-mix

Reagents	**Volume (µL)**
Water	4.5 - x
10× Fragmentation buffer	0.5
Fragmentation reagent (DNase I)	x
Total	5

N.B. x = [0.1/y]*5 (where y = fragmentation reagent concentration; range 2 to 3 U/µL)

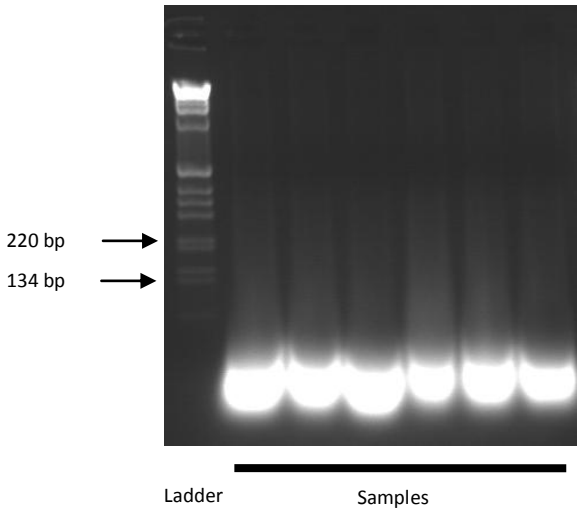


Figure 2.3 Visualisation of fragmentation products

Pooled and fragmented PCR products from 6 samples run on agarose gel demonstrating fragment sizes <180 bp.

2.5.8 Labeling

The labeling master-mix was prepared as indicated in **Table 2.9**, with 19.5µL added to the remaining fragmented sample which was then incubated at 37°C for 4 hours then 95°C for 15 minutes.

2.5.9 Hybridisation mix

A hybridisation master-mix was prepared as detailed in **Table 2.10** and 190µL added to each labelled sample (72.5µL) in a fresh 1.5mL micro-centrifuge tube. These samples were then passed to the designated Affymetrix micro-array operator in our institution (Mrs T Chaplin) who performed hybridisation, washing, staining and scanning of the array chips.

2.5.10 Hybridisation, washing, staining and scanning

Each sample of hybridisation mix and labelled DNA was incubated at 95°C for 10 minutes and then 49°C, at this stage 200µl of the sample mix was removed and loaded into individual arrays pre-warmed to room temperature. After sealing the loading ports on the chips, they were placed in a GeneChip® Hybridisation Oven 640 (Affymetrix) and rotated at 50°C, 60rpm for 16 to 18 hours to enable DNA hybridisation onto the array. After extraction of the hybridisation mix, each array was washed and stained in the GeneChip® Fluidics Station 450 (Affymetrix). An initial wash performed with 6× SSPE, 0.01% Tween 20 buffer at 25°C was followed by a stringent wash with 0.6× SSPE, 0.01% Tween 20 buffer at 45°C. A three step staining process was then performed with, firstly, 10µg/mL streptavidin R-phycoerythrin (SAPE) (Molecular Probes, Paisley, UK) applied to the chip to bind the biotin labelled DNA fragments that had been hybridised to the probes on the array. Following a further wash with 6× SSPE, 0.01% Tween 20 buffer at 25°C, signal amplification was enabled by addition of 5µg/mL biotinylated goat anti-streptavidin antibody (Vector Laboratories, Burlingame, CA, USA) and further 10µg/mL SAPE to each array. The three staining steps were performed in a 6× SSPE, 0.01% Tween 20, 1× Denhardt's solution buffer at 25°C and were followed by a final wash with 6× SSPE, 0.01% Tween 20 buffer at 30°C. The arrays were scanned in a

Table 2.9 Labelling master-mix

Reagents	**Volume (µL)**
Terminal deoxynucleotidyl transferase (TdT) buffer (5×)	2
DNA labeling reagent (30mM)	14
TdT (30U/µL)	3.5
Total	19.5

Table 2.10 Hybridisation master-mix

Reagents	**Volume (µL)**
MES (12×; 1.25M)	12
Denhardt's solution (50×)	13
EDTA (0.5M)	3
HS-DNA (10mg/mL)	3
OCR, 0100	2
Human Cot-1 DNA® (1mg/mL)	3
Tween-20 (3%)	1
DMSO (100%)	13
TMACL (5M)	140
Total	190

GeneChip® Scanner 3000 7G (Affymetrix) and signal intensity values for each hybridised probe were obtained using Affymetrix GeneChip® Command Console (AGCC) software.

2.5.11 Data Preparation for Analysis

Using AGCC, the signal intensities from each probe were adjusted for background signal intensities on the array to create normalised image files (.cel files). Three quality control (QC) steps were performed with Genotyping Console™ on the .cel file data:

i. An estimate of the efficiency of signal generation from all probes on the array was provided by measuring the Call Rate on a sub-set of 3022 SNPs. A level >97% indicated a satisfactory performance of the assay from the initial DNA digestion step through to the array scanning;

ii. The accuracy of discrimination of the three allele combinations for each SNP was estimated by the Contrast QC algorithm (CQC) which is a better indicator of genotyping performance for samples than the Call Rate alone. Values >0.4 were satisfactory provided no more than 10% of sample values <1.0 and mean value across the sample set >1.7 (Affymetrix White Paper: Quality Control Assessment in Genotyping Console™; September 30, 2008 Revision 1); and

iii. The $\log_2$ratio of target sample to reference sample normalised probe signal provided the relative copy number, for the target sample, at the DNA location complementary to the probe. The Median Absolute Pairwise Difference (MAPD) was calculated as the median of all values produced from determining the differences between $\log_2$ratios for each adjacent pair of probes (in terms of genomic distance) on the array. MAPD therefore estimated variability in the $\log_2$ratio and reflected accuracy of copy number estimation with values <0.4 considered satisfactory (Affymetrix White Paper: Median of the Absolute Values of all Pairwise Differences and Quality Control on the Affymetrix Genome-Wide Human SNP Array 6.0; February 11, 2008 Revision 1).

Samples with QC values outside these limits were likely to provide less reliable results and factors contributing to poorer performance of the assay include suboptimal quality of initial DNA sample, non-adherence to protocol and variation in reagents including the hybridisation cocktail. A further limitation was presence of germline DNA in tumour samples as this would also generate signals by hybridisation to probes. As genotyping calls are determined from ratios of one allele signal to the other, the presence of germline signals in combination with target (tumour) signals can generate a combined signal that is outside the limits utilised by the software algorithms for genotype allocation. Genotype is then not assigned and the probe signal is classed as a "no call" (as are any failed probe hybridisations). Similarly, as copy number loss or gain is a ratio of target to reference, the degree of these changes in a tumour target can be underestimated by the effect of co-occurring normal copy number signal from germline tissue within the sample that effectively dilutes the signal of tumour DNA. Two study aspects help to limit these issues:

i. samples with at least 70% tumour content (<30% germline tissue) are preferably used as the software algorithms can most robustly determine genotyping and copy number changes above this level; and
ii. where possible, target samples are compared to paired *separate* germline samples from the same case thereby maximising detection of features specific to the target DNA

The .cel file data was exported in to the PARTEK® Genomics Suite (Partek Incorporated, St Louis, MI, USA), a software package in which genotype and DNA copy number could be determined. A second platform, the Genome Orientated Laboratory File (GOLF) software, developed and maintained by Professor Bryan Young within this institution (and available at https://bioinformatics.cancerresearchuk.org/~cazier01/Golf.html), was also used in data analysis. Datasets compatible with GOLF were established by applying the Affymetrix Genotyping Console™ Software Birdseed algorithm to the .cel file data. This assigned genotype calls (.chp file) and signal intensities for copy number state determination (.cn.chp file). Data was extracted from these files as tab delimited text (in .txt and .si files respectively) and exported into GOLF.

2.5.12 Analysis workflow

DNA copy number changes were determined as a relative change of the target sample (tumour DNA) signal intensity compared to the corresponding signal intensity of a paired baseline reference (germline DNA) using $\log_2$ ratios (as indicated in section **2.5.11**). The germline baseline consisted of DNA extracted from bone marrow / peripheral blood / saliva samples that were uninvolved by lymphoma. For cases with no germline DNA available, a pooled reference, constructed by taking the median values of each probe signal across unpaired germline samples (n = 17), served as a baseline. Direct comparison of t-FL to FL was not performed as this could create misleading CN estimates as these are established relative to baseline so, for example, a loss in FL could be interpreted as gain in t-FL (as illustrated in ***Appendix 1*** **Figure 1**). Consequently, FL and t-FL were each assessed relative to the germline baseline. This established separate FL and t-FL CN estimates that could then be compared so establishing the *actual* (rather than relative) differences between samples.

2.5.13 Segmentation

The Partek Genomics Suite includes two algorithms to determine copy number. The first is based on the statistical hidden Markov model (HMM) and is best used on a homogenous cell population with anticipated integer value copy numbers. However, by the nature of FL / t-FL there are both normal and tumour cells within samples which can provide heterogeneity to copy number values. As it can account (at least in part) for this variation a second algorithm (termed segmentation), which compares adjacent regions using differences in the $\log_2$ values to determine aberrations in copy number, has been developed and was applied here. Segmentation has the additional benefit that it uses a two sample t-test to determine whether there is a difference between the mean $\log_2$ ratio for adjacent regions allowing identification of adjacent segments with similar copy number.

Varying parameters contribute to determining each segment. Firstly, as the number of consecutive genomic markers determines the size of a segment, resolution was maximised by defining the smallest segments as those with the minimum number of genomic markers (n = 10) as recommended by the segmentation algorithm. The

median spacing of markers on the array is 700bp giving a resolution of around 7000bp with this approach (although this varies throughout the genome due to differing probe densities and spacings on the array). A range of spread above or below the baseline of the $\log_2$ ratios (typically 0.3) was set as the threshold for allocation of "gain" or "loss" to a possible copy number change in the sample under investigation. As the samples included a degree of germline tissue, the relative copy number (section **2.5.12**) could be underestimated and in order to minimise this, the range was varied on a case-by-case basis. For 30 / 31 cases, the values were 0.2 to 0.3 with a single case requiring a range of 0.13 due to a higher proportion of normal tissue. A one-sided one sample t-test was performed by the Partek genomics suite to establish if segments were significantly above or below the range and a stringent *P* value ($<1 \times 10^{-4}$) was used to minimise the inclusion of false positives.

However, each sample contains within in it signal intensities attributable to background "noise" (non-specific signal) which may cause false positives to be included in copy number estimates. Consequently, reducing the range to increase detection of copy number changes can reduce specificity of the analysis by increasing the number of false positives. The "signal to noise" coefficient (the ratio of mean difference of the signals of two neighbouring segments to the standard deviation of the signals from the whole chromosome) can limit this effect and so a stringent cut-off level (0.8) was applied to the data set to limit the number of these false positives. A guide to this level is provided by the segmentation copy number results; if on inspection of the traces clear false positives are being included the signal to noise coefficient can be raised until these are (largely) eliminated.

2.5.14 Exclusions

As a subset of cases did not have paired germline DNA available, the copy number estimates for the respective tumours could include artefact changes at the sites of known polymorphic copy number variations (CNVs) (Iafrate *et al.* 2004; Redon *et al.* 2006; McCarroll *et al.* 2008) present in samples within the DNA germline pool (but not necessarily the investigated sample). Consequently, known CNVs (as previously determined on the SNP 6.0 array (McCarroll *et al.* 2008) were excluded from the copy number estimates of these samples as it was not possible to determine

if apparent gains or losses at these sites were real or were false positives due to CNVs in the baseline pool (analogous to the erroneous CN estimates that could occur by comparing t-FL directly to FL).

FL is a germinal centre derived B-cell malignancy and the tumour cells have undergone the normal B-cell process of V(D)J recombination at *IGH*, *IGK* +/- *IGL* (as described in Chapter 1) which can be used to indicate clonality of the B cell population (as shown for all 109 FL cases examined in a recent report (Evans *et al.* 2007)). As a consequence, these loci were also excluded from the copy number determination as CNA losses here might represent this normal B-cell process (which excludes segments of the respective genes). The T cell receptor loci, *TCRA / TCRB / TCRG / TCRD,* were also excluded for two reasons: firstly, in polyclonal T cells of healthy individuals a large proportion of those with TCRγδ show restriction to Vγ9–Jγ1.2 and Vδ2–Jδ1 (Triebel *et al.* 1988; Breit *et al.* 1994) so any infiltrating T-cells within biopsies might create a loss at γ (chromosome 7p14.1) or δ (chromosome 14q11.2) loci and, secondly, TCR re-arrangements may be detected in 10 - 13% of FL (Szczepanski *et al.* 1998; Evans *et al.* 2007).

2.6 RNA

Total RNA was extracted for gene expression analysis. The extraction method preferentially collected RNA >70bp, enriching for mRNA. Random hexamers were used in reverse-transcription to give maximal conversion of RNA to cDNA as both spliced and unspliced variants were assessed.

2.6.1 RNA extraction

The RNeasy® plus Mini kit (Qiagen, Hilden, Germany), which involves a series of steps performed in spin columns by centrifugation, was used for extraction of RNA from primary sample cell suspensions and cell line cultures according to the manufacturer's instructions. Briefly, around 3×10^6 cells were lysed in buffer containing guanidine-isothiocyanate to inactivate RNases then homogenized through a QIAshredder (Qiagen) spin column. After centrifugation in a genomic DNA eliminator column, RNA was captured in a silica-membrane then purified through

spin column washes and finally eluted in 30µl RNase-free H_2O. RNA was then used immediately or stored at -80°C.

Quality and quantity of the sample was determined using the NanoDrop® system (as described in section **2.4.2**), but using an RNA mode with the absorbance at 260nm multiplied by a factor of 40 to give RNA quantity and satisfactory sample quality indicated by $OD_{260/280}$ ratio ~2.0 and $OD_{260/230}$ ratio1.8 - 2.2.

2.6.2 RT-PCR: two step process

Two-step RT-PCR assessment of RNA expression involved first-strand cDNA generation with the M-MLV Reverse Transcriptase RNase H Minus, Point Mutant kit (Promega, Madison, WI, USA) according to the manufacturers protocol. This involved incubation of the total RNA preparation in the RNA random hexamer annealing mix (as detailed in **Table 2.11**) at 70°C for 5 minutes followed immediately by cooling on ice for 5 minutes. The reverse-transcription (RT) mix (**Table 2.12**) was then added to the RNA preparation with incubation performed for one hour at 42°C followed by a final step at 95°C for 5 minutes to inactivate the RT reaction. Second strand cDNA synthesis and PCR amplification were then performed in combination by using the HotStarTaq *Plus* DNA polymerase kit (as described in section **2.4.3**).

Table 2.11 First Strand cDNA synthesis: RNA random hexamer annealing mix

Reagents	**Volume (µL)**
Total RNA* sample	x
RNase free H_2O	10.5 - x
Random hexamers (50µM)	1.5
Total	12

*1µg of RNA was used (but differing sample concentrations meant volume [x] could vary)

Table 2.12 First strand cDNA synthesis: reverse-transcription mix

Reagents	**Volume (µL)**
5× M-MLV RT reaction buffer	6
dNTPs (10mM)	10
M-MLV RT (H-)	2
Total	18

Chapter 3 Germline factors

3.1 Introduction

The increasing appreciation that germline genetic factors contribute to cancer both through conferring predisposition to its development and by impacting on the clinical outcome of established malignancy led to the first phase of this project.

The advent of GWAS that can identify germline polymorphic variation, in the form of SNPs, with clinical associations has been a significant advance in the study of inherited determinants of disease. As SNPs can influence genetic regulation through various mechanisms including the modulation of transcription factor binding sites and generation of amino acid substitutions, they might influence cancer development and progression. A growing body of evidence suggests that this indeed the case and, in FL, several studies have investigated risk of developing disease particularly with regard to immune-gene related SNPs as detailed in Chapter 1. However, far less is known as to the effect of such polymorphic variation on the outcome of FL.

To address this further, four SNPs which might be relevant to the clinical course of FL, as indicated by their potential biological effects and known clinical associations, were included in the study. Two of the SNPs, rs2279744 (known as *MDM2* SNP 309) (Bond *et al.* 2004) and rs1042522 (known as *TP53* Arg72Pro) (Matlashewski *et al.* 1987), are well characterised functional SNPs that can modulate the MDM2-p53 axis, a critical tumour suppressor pathway. *TP53* is one of the most important tumour suppressor genes and its protein product (tumour protein 53; p53) has a critical role in cell-cycle progression, initiation of DNA repair and the appropriate selection of apoptosis (Levine *et al.* 1991; Ko and Prives 1996; Vogelstein *et al.* 2000; Vousden 2000; Ryan *et al.* 2001). As maintenance of genome integrity is highly regulated, p53 activity is tightly controlled and is suppressed by MDM2, its principal inhibitor (Oliner *et al.* 1992), until various cellular stresses, including DNA damage, lead to upregulation of its activated form (Giaccia and Kastan 1998; Vogelstein *et al.* 2000). However, although acquired defects in the MDM2-TP53 axis are important in FL (Davies *et al.* 2005; O'Shea *et al.* 2008) and both *MDM2* SNP 309 and *TP53* Arg72Pro are known to impact on risk and outcome of other cancers, the role of

these SNPs in FL remains unclear.

Two further SNPs, rs6457327 and rs10484561, were recently identified as robust risk predictors in the first genome-wide association studies (GWAS) of FL. Both are within the immune gene rich human leukocyte antigen (HLA) region on chromosome 6p21 with rs6457327 in a 26 kb region of high linkage disequilibrium (LD) that contains only one coding region *C6orf15* (Skibola *et al.* 2009) and rs10484561, upstream of *HLA-DQB1,* located in an independent 100 kb region of high LD (Conde *et al.* 2010). The GWAS used a three stage process to provide robust validation of the risk associations of both SNPs with FL but the relationship of rs6457327 and rs10484561 to FL outcome has not yet been examined.

3.2 Aims and objectives

The role of disruption to the MDM2-p53 axis in FL and the known contribution of *MDM2* SNP 309 and *TP53* Arg72Pro to the outcome of other cancers, together with the risk associations of rs6457327 and rs10484561 with FL and the importance of the immune response to the clinical outcome of this disease, led to this study to determine the relevance of these four SNPs to the clinical course of FL and its transformation.

Study objectives included examination of the respective SNPs' genotypes in a large series of DNA samples from patients diagnosed with FL and comparison of genotype frequencies to non-FL cases so that associations with disease risk could be examined. The associations of genotypes with both patient presentation characteristics and clinical outcome variables were established; and significant clinical outcome associations were addressed in a subset of cases that had originally presented to this institution and so had extensive, comprehensive follow-up data. Finally, the potential molecular basis of positive associations was explored through assessment of a linked genetic locus.

Data generated from the two phases of this chapter have subsequently been published over the course of this thesis (Wrench *et al.* 2009; Wrench *et al.* 2011).

3.3 Materials and Methods

3.3.1 Case selection

Study cases (n = 231) were selected retrospectively from the haemato-oncology database maintained at St. Bartholomew's Hospital. Cases were included in this study if the following criteria were met:

1. Prior histological diagnosis of FL
2. One tissue sample *not* involved by tumour taken at some point over the disease course
3. Sufficient amount of previously extracted DNA available in the tissue archive at St Bartholomew's Hospital from the uninvolved sample OR sufficient quantity of uninvolved tissue available for DNA extraction
4. Satisfactory DNA quality (Section **2.4.2**)

3.3.2 Samples

The uninvolved DNA samples represented germline and were used to determine the constitutional SNP genotypes of each case. The sample tissue sources included bone marrow or peripheral blood samples (n = 226) and rare reactive lymph node samples (n = 5), which were uninvolved by lymphoma and obtained throughout the course of disease from patients diagnosed with FL between 1977 and 2005 and managed at Bart's and the London NHS Trust.

In addition, DNA from diagnostic FL samples obtained at this institution (n = 50) and from FL samples (n = 12) examined in a previous investigation conducted by this institution in collaboration with the LLMPP (O'Shea *et al.* 2009) were included in the study.

3.3.3 Controls

For determination of risk, SNP genotype frequencies were compared to published control populations including those derived from previous reports and the HapMap

project ("The International HapMap Project" 2003). Potential risk associations were confirmed and validated by comparison of SNP genotype frequencies for the FL study population against genotype frequencies of controls derived from the WTCCC2 BC58 cohort which is described in detail in section **2.4.6**.

3.3.4 SNP Genotyping

A real-time polymerase chain reaction allelic discrimination multiplexed endpoint assay protocol (AD-PCR) determined the genotypes of *MDM2* SNP 309, *TP53* Arg72Pro, rs6457327 and rs10484561 for the study samples. For *MDM2* SNP 309 and *TP53* Arg72Pro, oligonucleotide primers were designed to amplify 109 bp and 74 bp amplicons incorporating *MDM2* SNP 309 and *TP53* Arg72Pro, respectively. Oligonucleotide probes labeled with either FAM or VIC fluorescent reporter dyes were designed for each allele of the respective SNPs. Details of the respective primers and probes used for the two SNPs are shown in ***Appendix 2*** **Table 1**. For rs6457327 and rs10484561, Taqman® SNP Genotyping custom assays C__11197333_10 and AHHR6XS (Applied Biosystems) were employed, respectively. These assays had been custom made and optimized to give maximal discrimination of genotypes and represent an important advance in this technique. Full details of reaction mixes and conditions as well as interpretation of results are provided in Section **2.4.5**.

3.3.5 Direct sequencing

Confirmation of SNP genotypes on a 10% subset of the AD-PCR study cohort and gene mutation screening were performed using direct sequencing as described in Section **2.4.12**. The oligonucleotide primers employed are shown in ***Appendix 2*** **Table 1**.

3.3.6 FL characteristics for correlation with SNP genotypes

To examine the relationships between SNPs and disease presentation characteristics, parameters including age, gender, stage, haemoglobin (Hb), lactate dehydrogenase (LDH), hydroxybutyrate dehydrogenase (HBD), number of nodal sites involved by

disease, follicular lymphoma international prognostic index (FLIPI) risk groups, tumour grade (1, 2 or 3a), B symptoms, performance status and number of extra-nodal disease sites were included in the study.

For correlation of SNP genotypes with the course of disease both management and outcome parameters were assessed and included: initial management, date of and response to first therapy, dates of first relapse and transformation and date of death or last follow-up. Associations were validated through multivariate analysis performed on a subset of cases (n = 130) which had been diagnosed at St Bartholomew's Hospital and so had prolonged follow-up data available.

3.3.7 Statistical validation of FL risk associations

Single marker association tests were conducted using the one degree of freedom (df) allelic chi-squared test, 2df genotypic test and 1df association test under an additive model (Cochran-Armitage trend test) implemented in the PLINK 1.07 software (http://pngu.mgh.harvard.edu/~purcell/plink/). PLINK v1.07 was also used to test deviations from Hardy-Weinberg equilibrium in controls. Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated for the minor allele and for the variant allele carriers (homozygous or heterozygous versus homozygous common allele carriers) by median-unbiased estimation using the mid-p method from the epitools R package (http://cran.r-project.org/web/packages/epitools/index.html).

3.3.8 Statistical validation of clinical outcome associations

Associations with diagnostic parameters were determined using the Wilcoxon-Mann-Whitney test or analysis of variance (ANOVA) for continuous variables and either Chi-squared or Fisher's exact tests for categorical data. Initial response to therapy was examined by the Kruskal-Wallis test which allowed the ordering in the response to treatment categories to be considered when assessing the importance of the SNP genotypes. Survival analysis was performed using the Kaplan-Meier survival function with log rank test for equality of survival (censoring at death or last follow-up).

Univariate associations were examined by multivariable analyses. For association with categorical data, potential confounders were investigated by Mantel-Haenszel analysis and those which caused a ≥10% change in the crude odds ratio (OR) were then assessed by logistic regression. Any confirmed confounders (those causing a ≥10% change to the estimated effect as indicated by crude odds ratios in the logistic regression analysis) were included in a final model incorporating both crude and adjusted ORs. For association with survival data, potential confounders which gave a ≥10% change to the estimate of effect indicated by altering the crude hazard ratio (HR) were included in a Cox regression analysis. Potential confounders identified by clinical hypotheses were also included in the regression analyses.

Analyses were performed using the STATA 11.1 (Stata Corporation, College Station, TX, USA) and GraphPad Prism version 5.03 (GraphPad Software Inc., La Jolla, CA) computer software packages. Results were considered statistically significant when $P < 0.05$.

3.3.9 RNA

RNA expression was determined for an additional series of available samples including 16 FL and 2 t-FL tumour samples (lymph nodes), 1 benign lymph node and 4 tonsils. Six B-cell NHL cell lines (Su-DHL4, RL, Su-DHL6, K422, SC1 and DoHH2) and 2 lymphoblastoid cell lines (HRC57 and NC-NC) were also examined as well as two myeloid cell lines (K562 and U937) which were included as non-lymphoid haematological tissue. The extraction of RNA and its subsequent amplification by RT-PCR were performed as described in section **2.6**. The primers used for PCR of cDNA are indicated in ***Appendix 2*** **Table 1**.

3.4 Results

This study was performed in two stages. Firstly, *MDM2* SNP 309 and *TP53* Arg72Pro were genotyped to determine the relevance in FL of these SNPs which may have functional influence on a major tumour suppressor axis. During the course of the study, the first GWAS in FL were reported and these identified rs6457327 and rs1048456 as risk determinants which were then examined here for their influence on the clinical characteristics of FL.

3.5 *MDM2* SNP 309 and *TP53* Arg72Pro

3.5.1 Patient and sample characteristics

A total of 226 cases were included with a median age at diagnosis of 46 years (range 21 – 85). The source of DNA was bone marrow (n = 207), peripheral blood (n = 14) or lymph node (n = 5). Median follow-up of the patients remaining alive was 9.5 years (range 1.4 to 30.6) and other presentation clinical characteristics are summarised in **Table 3.1**.

3.5.2 *MDM2* SNP 309 and *TP53* Arg72Pro do not predict FL risk

An example of the AD-PCR genotyping results is illustrated in ***Appendix 3*** **Figure 1**. Direct sequencing of the validation set (n = 24) revealed genotypes that were in full concordance with those determined by AD-PCR. The distributions and relative frequencies of the 226 genotypes for both *MDM2* SNP 309 and *TP53* Arg72Pro across the study population are indicated in **Table 3.2**. These were not significantly different when compared to genotypes in previously reported control groups (taken from populations with similar demographics to the study cohort) or to data available from the HapMap project. This suggests that *MDM2* SNP 309 and *TP53* Arg72Pro do not predict risk of FL.

3.5.3 SNP genotype groups for correlation with clinical outcome

Both *MDM2* SNP 309 and *TP53* Arg72Pro can negatively impact on p53 function and so have potential for biological synergy through disruption of the MDM2-p53 axis. In view of this and their reported multiplicative relevance in cancer development (Hong *et al.* 2005; Zhang *et al.* 2006), the combined influence of *MDM2* SNP 309 and *TP53* Arg72Pro on FL outcome was determined in addition to the individual effects of each SNP.

Cross-tabulation of the respective SNP genotypes was performed which indicated there were no cases homozygous for both minor (*MDM2* SNP 309 G or *TP53* Arg72Pro C) alleles (***Appendix 3*** **Table 1**) and less than ten cases in each group with

Table 3.1 *MDM2* / *TP53* SNP genotypes & FL presenting features

Characteristic	Subtype	n (%)*	*MDM2* SNP 309		*P*	*TP53* Arg 72 Pro		*P*
			TT	TG+GG		GG	GC+CC	
Age (years)	Median (range)	46 (21 - 85)			0.80			0.77
	<60 years	181 (80)	73	108		83	98	
	≥60 years	45 (20)	26	19		22	23	
Sex	Male	124 (55)	55	69	0.89	57	67	0.89
	Female	102 (45)	44	58		48	54	
Stage	1 / 2	48 (25)	17	31	0.18	22	26	1.00
	3 / 4	146 (75)	69	77		67	79	
	missing	*32*						
B symptoms	Absent	89 (77)	44	44	0.66	40	48	0.51
	Present	27 (23)	12	15		10	17	
	missing	*111*						
PS	0 - 1	74 (93)	39	35	1.00	30	44	1.00
	2 - 4	6 (7)	3	3		2	4	
	missing	*146*						
Extranodal	0 - 1	65 (80)	33	32	0.78	24	41	0.40
	2 - 4	16 (20)	7	9		8	8	
	missing	*145*						
Nodal sites	0-4	45 (59)	23	22	0.82	19	26	0.64
	5-8	31 (41)	17	14		11	20	
	missing	*150*						
LDH/HBD[#] (iu)	≤ULN	74 (77)	35	41	1.00	37	39	0.08
	>ULN	22 (23)	9	11		5	15	
	missing	*130*						
Hb (g/dL)	Median (IQR)	13.6 (12.6 - 14.7)			0.18			0.92
	≥12g/l	91 (84)	42	49		40	51	
	<12g/l	17 (16)	7	10		6	11	
	missing	*118*						

*number (percentage) unless indicated as median and range
[#]LDH performed from the year 1989 onwards and HBD performed up to 1989
PS indicates performance status as per the ECOG criteria (Oken *et al.* 1982); LDH, lactate dehydrogenase; HBD, hydroxybutyrate dehydrogenase; iu, international units; Hb, haemoglobin; g/dL, grams per decilitre; ULN, upper limit of normal; IQR: inter-quartile range.

Table 3.2 *MDM2* SNP 309 and *TP53* Arg72Pro case control study

SNP	Genotype	FL Cases		Published controls*		MAF cases/controls	*P*
		Number	*Frequency (%)*	Number	*Frequency (%)*		
	TT	99	*44*	26	*48*		
MDM2 SNP 309	TG	98	*43*	22	*40*	0.34 / 0.32	0.82
	GG	29	*13*	7	*12*		
	GG	105	*46*	87	*50*		
TP53 Arg72Pro	GC	105	*46*	76	*44*	0.30 / 0.28	0.43
	CC	16	*7*	10	*6*		

*As per previous reports for *MDM2* SNP 309 (Bond *et al.* 2005) and *TP53* Arg72Pro (Pierce *et al.* 2000).
MAF, minor allele frequency.
MAF calculated using numbers from genotypes = [(2*homozygous minor genotype) + heterozygous genotype] / [total cases*2].
Both SNPs were in Hardy-Weinberg equilibrium in this cohort of FL cases ($P > 0.1$ for each SNP), which suggests they are not under selective pressure supporting the lack of differences in genotype frequencies when compared to controls.

homozygosity for one minor allele. However, there were satisfactory numbers in the four groups generated by cross-tabulation of cases who were minor allele carriers (either heterozygous cases or those homozygous for the minor allele) and cases who were homozygous for major alleles (as, respectively, TG+GG and TT for *MDM2* SNP 309; GC+CC and GG for *TP53* Arg72Pro), as summarised in ***Appendix 3*** **Table 2**. Consequently, these genotypes were assessed for the respective SNPs and the four genotype combinations from this cross-tabulation (*MDM2* SNP 309 TT with *TP53* Arg72Pro GG; *MDM2* SNP 309 TT with *TP53* Arg72Pro GC+CC; *MDM2* SNP 309 TG+GG with *TP53* Arg72Pro GG; and *MDM2* SNP 309 TG+GG with *TP53* Arg72Pro GC+CC) were examined regarding the combined effect of *MDM2* SNP 309 and *TP53* Arg72Pro in FL. Such groupings also enable the relevance of minor (variant) alleles (*MDM2* SNP 309 G or *TP53* Arg72Pro C) to be determined.

3.5.4 *MDM2* and *TP53* SNPs do not predict FL presentation features

There were no significant associations with age, gender, presentation stage, presence of B symptoms, performance status, extra-nodal disease, number of nodal sites involved, LDH or HBD, and haemoglobin at FL diagnosis for *MDM2* SNP 309 genotypes and *TP53* Arg72Pro genotypes (**Table 3.1**) or the four combinations of *MDM2* SNP 309 and *TP53* Arg72Pro genotypes (**Table 3.3**), $P \geq 0.18$.

3.5.5 *MDM2* and *TP53* SNPs do not predict clinical outcome of FL

Similarly, examination of the relationships between these SNP genotypes and clinical outcome revealed there was no association with response to first therapy for these genotypes (also $P \geq 0.18$) and, in particular, there was no association with risk of transformation (illustrated in **Table 3.4**), $P \geq 0.62$. Furthermore, there was no evidence for association of these genotype combinations with the survival outcomes of time to disease transformation (TTT), time to first relapse and overall survival (OS), $P \geq 0.17$. Examples for TTT and OS are illustrated in **Figure 3.1** and **Figure 3.2**, respectively.

Table 3.3 *MDM2* / *TP53* SNP combinations & FL presentation characteristics

Characteristic	Subtype	*MDM2* SNP 309 & *TP53* Arg72Pro genotypes				*P*
		Neither rare	Rare *TP53*	Rare *MDM2*	Both rare	
Age (years)	Median (range)					
	<60 years	35	38	48	60	0.79
	≥60 years	10	16	12	7	
Sex	Male	27	28	30	39	0.67
	Female	18	26	30	28	
Stage	1 – 2	11	6	11	20	0.13
	3 – 4	30	39	37	40	
	missing = 32					
B symptoms	Absent	21	23	19	25	0.65
	Present	6	6	4	11	
	missing =111					
PS	0 – 1	18	21	12	23	- †
	2 – 4	1	2	1	2	
	missing = 146					
Extranodal	0 – 1	13	20	11	21	0.49
	≥ 2	5	2	3	6	
	missing = 145					
Nodal sites	≤ 4	12	11	7	15	0.89
	≥ 5	7	10	4	10	
	missing = 150					
LDH/HBD (iu)	≤ULN	18	17	19	22	0.29
	>ULN	2	7	3	8	
	missing = 130					
Hb (g/dL)	Median (IQ range)					0.91
	≥12g/l	17	25	23	26	
	<12g/l	4	3	2	8	
	missing = 118					

†Too few values were present in the subsets for PS to satisfy Chi square criteria. On a restricted analysis of homozygosity for both major alleles (Neither rare) versus any minor allele present (Rare *TP53* + Rare *MDM2* + Both rare) by Fisher's exact test, there was no significant association.
Neither rare indicates *MDM2* SNP 309 TT with *TP53* Arg72Pro GG; Rare MDM2, *MDM2* SNP 309 TG or GG with *TP53* Arg72Pro GG; Rare TP53, *MDM2* SNP 309 TT and *TP53* Arg72Pro GC or CC; and Both rare, *MDM2* SNP 309 GT or GG and *TP53* Arg72Pro GC or CC genotypes.
Abbreviations are as indicated for **Table 3.1**.

Table 3.4 *MDM2* & *TP53* SNP combinations and transformation risk

SNP	Genotype	Transformed	Never transformed	*P*
MDM2 SNP 309	TT	33 (33)	66 (67)	0.89
	GT or GG	44 (35)	83 (65)	
MDM2 SNP 309 / *TP53* Arg72Pro[#]	Neither rare	13 (29)	32 (71)	0.86
	Rare TP53	20 (37)	34 (63)	
	Rare MDM2	21 (35)	39 (65)	
	Both rare	23 (34)	44 (66)	
TP53 Arg72Pro	GG	34 (32)	71 (68)	0.62
	CC or CG	43 (36)	78 (64)	
Total cases		*77(34)*	*149(66)*	

Parentheses include percentages transformed or never transformed for each genotype. The total number of cases which transformed or never transformed in the cohort are indicted for comparison.
[#]Genotypes are as detailed in **Table 3.3**.

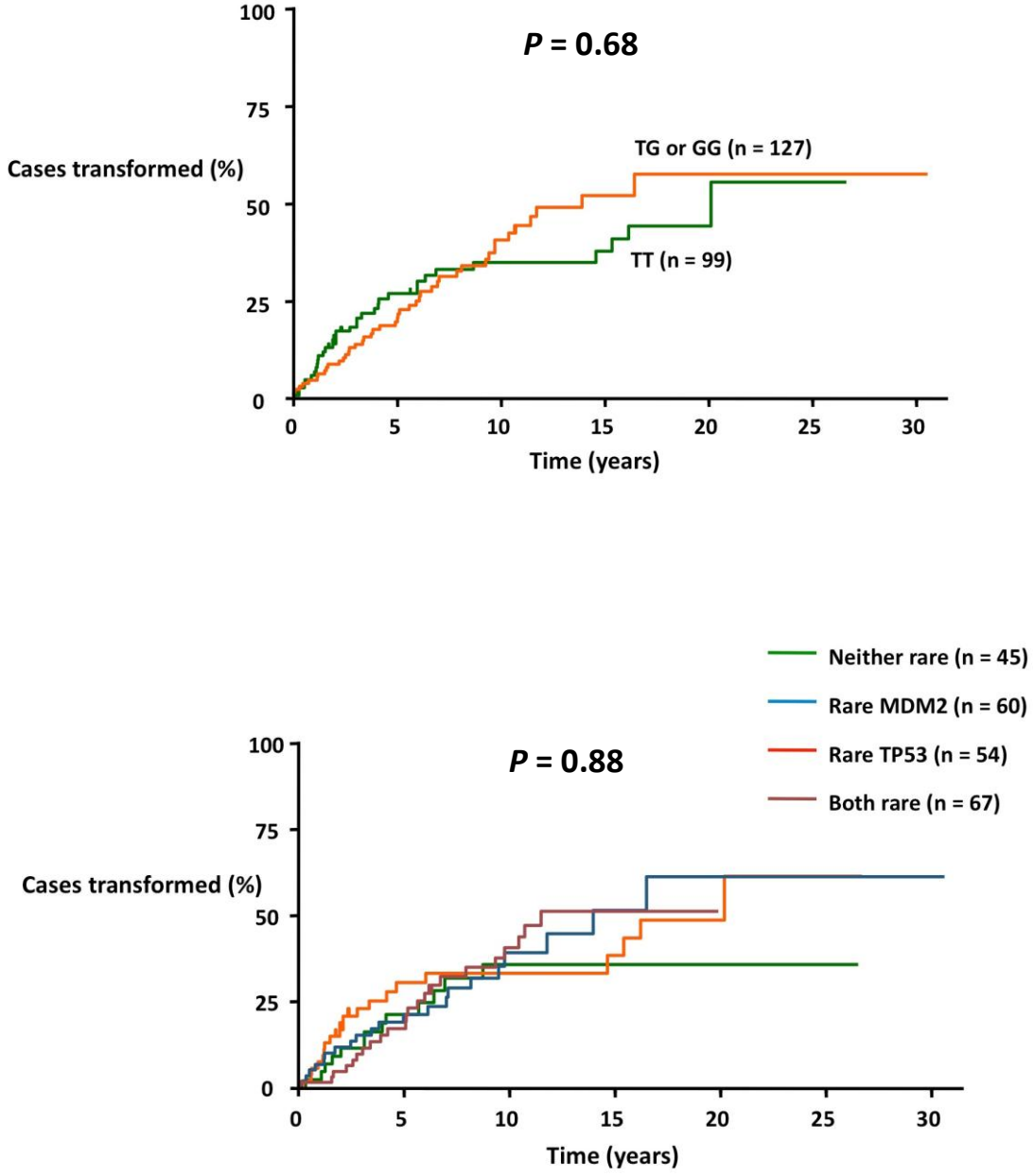


Figure 3.1 ***MDM2*** **SNP309 /** ***TP53*****Arg72Pro and time to transformation**

Example Kaplan-Meier plots and log rank test statistics for survival from diagnosis of FL to transformation (censored at last known follow-up +/- death). Top, *MDM2* SNP 309 restricted genotype groups (as labelled). Bottom, combined *MDM2* SNP 309 and *TP53* Arg72Pro restricted genotype groups (as indicated to the right); these groups are as detailed in **Table 3.3**.

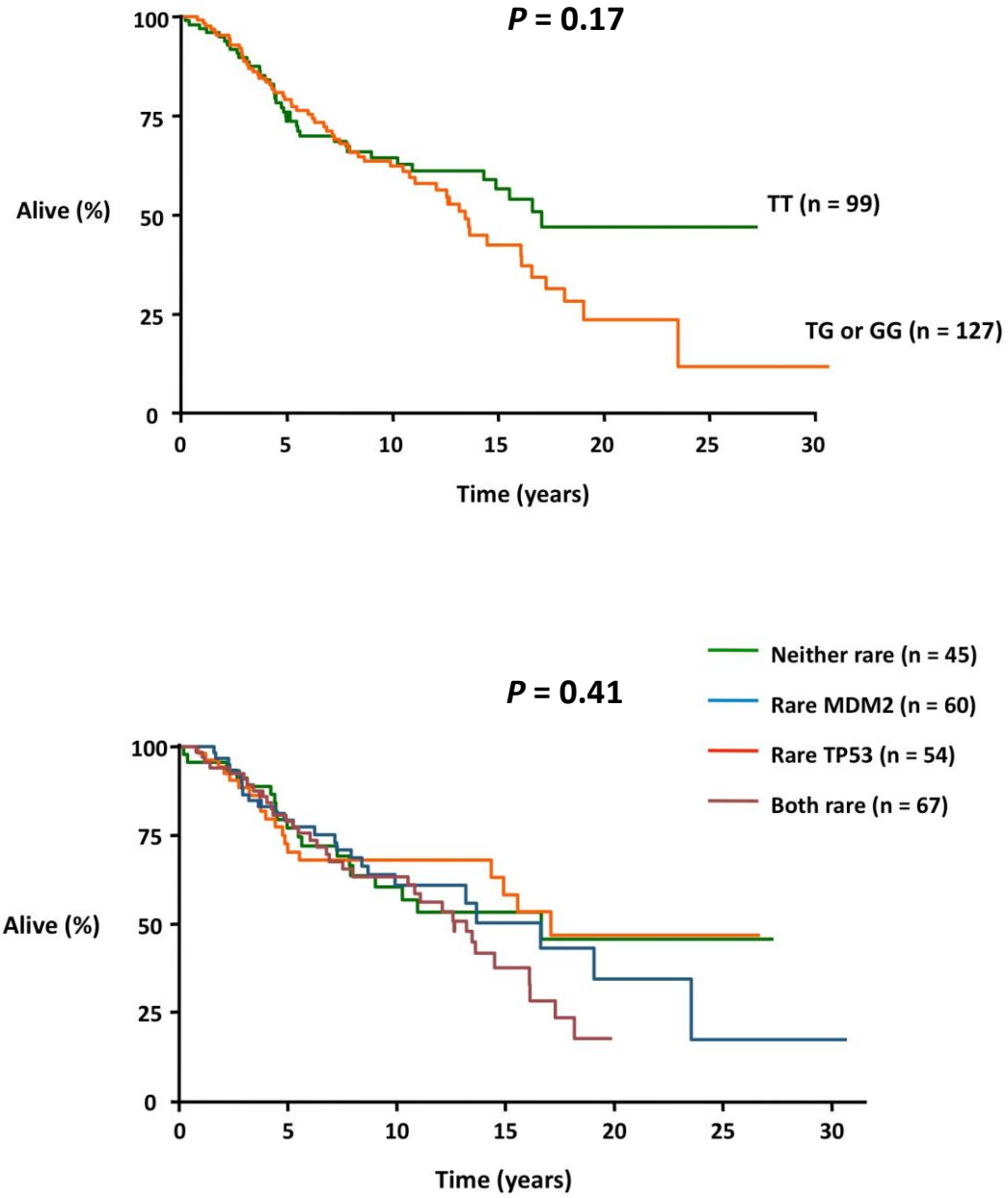


Figure 3.2 OS for *MDM2* SNP 309 or *MDM2* / *TP53* SNP combinations

Kaplan-Meier plots and log rank test statistics for overall survival (OS) from diagnosis of FL. The top and bottom of the figure as well as the combination genotype groups are as indicated in **Figure 3.1** with the exception that censored subjects are those alive at last follow-up.

3.5.6 *MDM2* SNP 309 does not predict t-FL in younger females

There were no differences in the median age at diagnosis for the *MDM2* SNP 309 and *TP53* Arg53Pro genotypes (***Appendix 3* Table 3**). Importantly, the *MDM2* SNP 309 minor (G) allele shows association with earlier onset of DLBCL in women under 51 years of age (Bond *et al.* 2006) and this gender specific effect may be due to the SNP variant modulating the known influence of oestrogen receptor alpha on *MDM2* expression (Phelps *et al.* 2003). Consequently, the relationship of this SNP to age at FL diagnosis and its transformation (as DLBCL) was examined in females who were less than 51 years old at time of FL diagnosis (n = 66). When assessing either the cases homozygous for the major T allele versus the minor allele carriers or the three separate genotypes there were no associations with either age at FL diagnosis and age at transformation (**Table 3.5**) or risk of transformation (**Table 3.6**) in this subgroup and, furthermore, there were no associations observed for the group of females who developed transformation < 51-y-o (**Table 3.5**), $P \geq 0.11$.

Table 3.5 *MDM2* SNP 309 and age at FL / t-FL diagnosis in females <51 years

MDM2 SNP 309 Genotype	Females < 51 years at FL diagnosis (n = 66)				Females < 51 years at t-FL (n = 17)	
	Median age at FL	*P*	Median age at t-FL	*P*	Median age at t-FL	*P*
TT	41		45		43	
TG	44	0.11	49	0.14	45	0.29
GG	44		45		42	

Also, there were no associations ($P \geq 0.15$) between the two genotype groups TT v TG+GG and any of the three median ages at FL or t-FL.

Table 3.6 *MDM2* SNP 309 and t-FL risk for females <51 years at FL diagnosis

MDM2 SNP 309 Genotype	Transformed	Never transformed	*P**
TT	9	18	
TG	8	21	0.17
GG	6	4	

In addition, there was no association ($P = 1.0$) between the two genotype groups TT v TG+GG and risk of t-FL in females < 51-y-o.

3.6 rs6457327 and rs10484561

The rs6457327 and rs10484561 SNPs are located on chromosome 6p and although no functional consequences have been described, both are found in independent small sections of high linkage disequilibrium (as illustrated in ***Appendix 4*** **Figure 1** and ***Appendix 4*** **Figure 2**, respectively). These regions contain genes which might therefore be relevant and this, in combination with the risk associations of both SNPs (Skibola *et al.* 2009; Conde *et al.* 2010), led to further examination of these two loci in FL.

3.6.1 Patient and sample characteristics

This investigation was performed following the *MDM2* SNP 309 / *TP53* Arg72Pro study. Due to new sample availability and exhaustion of some previous samples, a slightly revised cohort (n = 218) of FL cases was examined. For this cohort, genotyping of rs6457327 and rs10484561 was performed on germline DNA samples obtained from either bone marrow (n = 203) or peripheral blood (n = 15) and examples of the results plots are illustrated in ***Appendix 4*** **Figure 3**.

3.6.2 rs6457327 and rs10484561 predict risk of FL in the UK population

The genotype frequencies of rs6457327 and rs10484561 for the 218 FL cases were similar to those reported in the cases included in the GWAS studies (Skibola *et al.* 2009; Conde *et al.* 2010). The rs6457327 minor allele (A) appeared less commonly in FL (with a frequency of 0.32) than in the control population of the initial report by Skibola and colleagues (Skibola *et al.* 2009); and the rs10484561 minor allele (G) was enriched in FL (with a frequency of 0.21) compared to the controls in the follow-up report by the same group (Conde *et al.* 2010).

To further address the risk association of these SNPs in the UK population, the rs6457327 and rs10484561 genotypes of the study cohort were compared to those of the WTCCC BC58 derived control populations (**Table 3.7**). Importantly, no departures from Hardy-Weinberg equilibrium were detected among these controls, with $P > 0.80$ for each set used in the comparisons for rs6457327 (n = 2691) and

Table 3.7 SNPs rs10484561 and rs6457327 and risk of developing FL in the UK

SNP	Genotype	Genotype count		MAF cases/controls	HWE controls	Variant allele OR (95% CI)‡	Allelic OR (95% CI) ‡	Allelic *P*	Trend *P*	Genotypic *P*
		Cases*	Controls†							
	GG	5	34							
rs10484561	GT	82	548	0.21/0.11	0.92	2.40 (1.80-3.19)	2.07 (1.61-2.63)	3.50×10^{-9}	2.74×10^{-9}	6.34×10^{-9}
	TT	131	2102							
	AA	19	408							
rs6457327	AC	103	1272	0.32/0.39	0.81	0.77 (0.58-1.01)	0.75 (0.61-0.93)	0.01	0.01	0.02
	CC	96	1011							

MAF indicates minor allele frequency, HWE, Hardy-Weinberg equilibrium; OR, odds ratio; and CI, confidence interval.
*FL cases from St Bartholomew's Hospital (n = 218 for both SNPs).
†Controls from WTCCC2 BC58 (as described in Chapter 2). Both control populations maintain Hardy-Weinberg equilibrium.
‡ORs calculated for the variant allele carriers (homozygous or heterozygous) versus homozygous common allele carriers (Variant allele OR) and for minor versus major allele (Allelic OR). Furthermore, these risk associations do not appear restricted to younger onset of disease (***Appendix 4*** **Table 1**).

rs10484561 (n = 2684) genotypes, respectively. On examination of the three genotype combinations for each SNP in FL cases compared to BC58 controls, both rs6457327 (P = 0.02) and rs10484561 ($P = 6.34 \times 10^{-9}$) were associated with risk of FL in the UK population. Furthermore, a significant difference was observed between the minor allele frequency (MAF) for rs6457327 (A allele) of 0.32 and 0.39 [P = 0.01, OR 0.75 (95% CI 0.61 - 0.93)] and for rs10484561 (G allele) of 0.21 and 0.11 [$P = 3.50 \times 10^{-9}$, OR 2.07 (1.61 - 2.63)] in the FL cases and BC58 control sets respectively (**Table 3.7**). These findings confirm the GWAS results and indicate that both SNPs associate with risk of FL in the UK.

3.6.3 Correlating rs6457327 and rs10484561 with clinical outcome of FL

Following examination of the correlation between rs10484561 or rs6457327 and clinical outcome parameters in FL for the whole cohort of cases (n = 218), only associations between rs6457327 and risk of transformation (P = 0.01) or time to transformation (P = 0.01) were observed. To confirm these findings, clinical outcome associations were determined for the subset of cases (n = 130) initially diagnosed at Barts and the London NHS Trust which consequently had prolonged clinical follow-up data.

It was important to confirm that there were no differences in genotype and allele frequencies for each SNP in the 130 cases evaluated for clinical outcome compared to the remaining 88 cases from the case-control study. This was indeed the case, as illustrated in (***Appendix 4*** **Table 2**), indicating the sub-population assessed for clinical associations is representative of the whole FL cohort as regards these SNPs. As there were limited numbers of cases homozygous for the variant (minor) alleles (3 and 9 cases for genotypes GG and AA of rs10484561 and rs6457327 respectively) analysis of clinical associations was performed on restricted genotype groups similar to those described in the *MDM2* / *TP53* study. These were minor allele carriers and cases homozygous for the major allele which gave comparisons of rs10484561 genotypes GG+GT versus TT and rs6457327 genotypes AA+AC versus CC.

Presentation characteristics and subsequent management / events for the cohort of 130 cases are summarised in **Table 3.8**. The median age at diagnosis of 52 years is as

Table 3.8 Correlating rs10484561 / rs6457327 with FL clinical features

Characteristic	Subgroup	n*	$P^{\dagger}$	
			rs10484561	rs6457327
Presentation				
Age at Dx, years		52 (26-87)[¶]		
	<60	95 (73)	.59	.73
	≥60	35 (27)		
Sex	Male	72 (55)	.81	.29
	Female	58 (45)		
Stage	1-2	33 (26)	.77	.51
	3-4	96 (74)		
	Missing	1		
LDH / HBD, iu[‡]	≤ULN	94 (82)	.74	.91
	>ULN	21 (18)		
	missing	15		
Haemoglobin, g/L	≥120	105 (83)	.80	.66
	<120	22 (17)		
	Missing	3		
Nodal sites	≤4	61 (58)	.13	.90
	≥5	45 (42)		
	Missing	24		
FLIPI group[§]	Low	39 (31)	.53	.90
	low / int	11 (9)		
	Int	37 (29)		
	int / high	11 (9)		
	High	28 (22)		
	Missing	4		
Histological grade	1	75 (64)	.87	.82
	2	31 (27)		
	3	11 (9)		
	Missing	13		
Management				
Initial management	Observation	51 (39)	.54	.24
	RT alone	12 (9)		
	single chemo	51 (39)		
	comb chemo	15 (12)		
	Other	1 (1)		
Response to 1st Tx[||]	CR	30 (27)	.25	.93
	CR(u)	39 (35)		
	PR	26 (23)		
	Stable	4 (3)		
	Progression	14 (12)		
	Missing	17		
Events				
RFS, months	Relapse	30 (7-180)[#]	.40	.17
TTT, months	transformation	30 (2-242)[#]	.38	**.01**
Transformation	Occurred	45 (35)	.60	**.006**
OS, years	death	12 (8-20)**	.19	.97

The significant univariate associations are highlighted in bold.
*Number (percentage of non-missing) unless indicated.
†Associations for GG+GT v TT and AA+AC v CC genotypes of SNPs rs10484561 and rs6457327 respectively.
‡ULN range for LDH = 480 iu and ULN range for HBD = 140 iu. Cases had either LDH or HBD results.
§For FLIPI risk group determination, 11 cases had FLIPI scores of ≤2 and 11 had scores of ≥2 but none of these 22 cases could be further categorized in to one of the three FLIPI risk groups due to missing data so are indicated here as low / int (score ≤2) and int / high (score ≥2).
||As per the "Report of an International Workshop to Standardize Response Criteria for Non-Hodgkin's Lymphomas" (Cheson *et al.* 1999).
¶Median (range) for all cases (n = 130).
#Median (range) for those cases who relapsed (n = 91) or transformed (n = 45).
**Median (inter-quartile range) established from the reverse Kaplan-Meier test on the whole population (n = 130).
Dx indicates diagnosis; LDH, lactate dehydrogenase; HBD, hydroxybutyrate dehydrogenase; iu, international units; g/L, grams per litre; FLIPI, follicular lymphoma international prognostic index; Tx, treatment; RFS, relapse free survival; TTT, time to transformation; OS, overall survival; ULN, upper limit of normal range; int, intermediate; RT, radiotherapy; comb, combination; chemo, chemotherapy; CR, complete response; CR(u), complete response unconfirmed; and PR, partial response.

expected in a tertiary referral cancer centre. Furthermore, the gender, grade, stage, remaining parameters contributing to the FLIPI scoring system (Hb, LDH/HBD, number of nodal sites involved by disease) and the proportion of cases in each FLIPI risk group are similar to those previously described in FL (Solal-Celigny *et al.* 2004). An initial "watch and wait" policy of observation (expectant management) was adopted in 51/130 (39%) of cases with 12 cases (9%) never receiving any treatment through the course of their disease. Complete or partial responses, determined according to the recommendations of the "Report of an International Workshop to Standardize Response Criteria for Non-Hodgkin's Lymphomas" (Cheson *et al.* 1999), were achieved in 95 of 113 assessable cases (84%) including 30 (27%) with complete response (CR). One third (45/130) of cases experienced disease transformation at a median time of 2.5 years (range 0.2 – 20.1) from diagnosis, similar to recent findings both from this institution and elsewhere (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008). In the 91 cases whose disease relapsed, median time to first relapse was 2.5 years. The estimated median overall survival for all 130 cases (censoring at last follow-up) was 12 years and the 67 surviving patients were followed up for a median of 9.3 years (range 1.6 – 30.5) with 94% seen for > 5 years.

3.6.4 rs6457327 predicts risk of and time to transformation of FL

Although SNP rs10484561 showed the strongest association with risk of FL, it showed no associations with any of the presentation characteristics ($P \geq 0.13$) and clinical outcomes ($P \geq 0.19$) tested (**Table 3.8**).

In comparison, rs6457327 AA+AC genotype showed significant associations with both risk of disease transformation $P = 0.006$, OR 2.95 (95% CI = 1.35 - 6.48), as indicated in **Table 3.8** and **Table 3.9**, and time to transformation $P = 0.01$, HR 2.25 (95% CI 1.16 - 4.36) (**Table 3.8** and **Figure 3.3**). Transformation occurrence was more than doubled in carriers of the minor allele (**Table 3.9**). These associations were maintained on comparison of the three separate rs6457327 genotypes (AA, AC and CC) for both risk of transformation, $P = 0.02$ and time to transformation, $P = 0.02$. Furthermore, on comparison of individual alleles there was a significant association of allele A, as compared to C, with risk of transformation, $P = 0.03$; OR 1.88 (1.10 - 3.23).

Table 3.9 SNP rs6457327 genotypes and risk of transformation

rs6457327 genotype	Transformed		*P*
	Yes	No	
AA + AC	33 (45)	41 (55)	.006
CC	12 (21)	44 (79)	
Total	*45 (35)*	*85 (65)*	

Percentages for each genotype are included in parentheses.

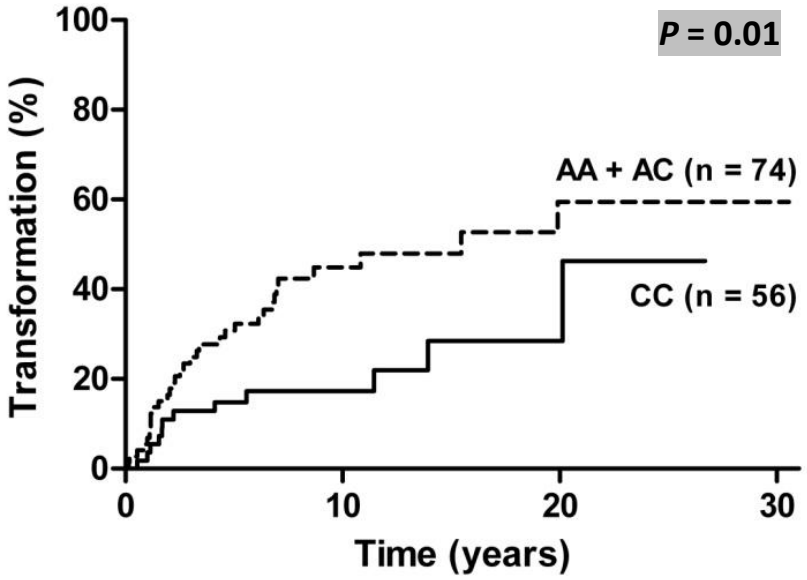


Figure 3.3 SNP rs6457327 genotypes and time to transformation

Kaplan-Meier plots with log rank test statistic *P* value of time to transformation for the rs6457327 AA+AC or CC genotype groups from the study cases diagnosed at Barts and the London NHS Trust (n = 130).

Although no other associations with rs6457327 were seen for any of the presentation or outcome parameters (**Table 3.8**), assessment of these against transformation yielded significant associations for FLIPI risk group and initial management with risk of transformation ($P \leq 0.02$); and for response to first therapy, FLIPI risk group and age at diagnosis with time to transformation ($P \leq 0.03$), indicating these might have confounding effects on the associations of rs6457327 with transformation.

These associations concur with previous reports examining risk factors for transformation (Bastion *et al.* 1997; Gine *et al.* 2006; Montoto *et al.* 2007b) although the association with initial management reported in a previous study from this institution, which included 89/130 cases from the current study (Montoto *et al.* 2007b), was not reproduced in a subsequent study from Vancouver (Al-Tourah *et al.* 2008).

Further potential confounders were identified as follows:

i. rs10484561 genotype was included for both risk of and time to transformation despite lack of association with either outcome as this SNP (although not in linkage disequilibrium with rs6457327) is located within 1.6 Mb on the same chromosome arm, 6p;

ii. response to 1st treatment was included as a potential confounder for risk of transformation with $P = 0.10$ (Fisher's exact test) in view of its previously reported association with transformation (Bastion *et al.* 1997) and its association here with time to transformation; and

iii. initial management was included as a potential confounder for time to transformation [with crude HR = 3.89 (0.90 - 16.77), $P = .07$ for observation] in view of a previous report of initial observation associating with subsequent transformation (Montoto *et al.* 2007b) and its association here with risk of transformation.

The potential confounders, as summarised in **Table 3.10,** were then taken into multivariate analyses assessing the associations of rs6457327 with risk of transformation and time to transformation.

Table 3.10 Potential confounders included in multivariate analyses

	Risk of transformation	Time to transformation
Potential confounders	FLIPI risk group *P* = 0.02 Initial management *P* = 0.02 rs10484561 genotype *P* = ns Response to 1st Treatment *P* = ns	Response to 1st Treatment HR 4.44 (1.70 – 11.55) *P* < 0.01 for progressive disease FLIPI risk group HR 2.95 (1.16 – 7.50) *P* = 0.02 for intermediate / high Age at FL Diagnosis HR 2.07 (1.06 – 4.02) *P* = 0.03 for age ≥60 years rs10484561 genotype *P* = ns Initial management *P* = ns

HR indicates hazard ratio; and ns, not significant.

The approach for the association between rs6457327 and *risk* of transformation utilized Mantel-Haenszel analysis and this revealed that FLIPI risk group, initial management and response to 1st treatment caused ≥ 10% change in the crude odds ratio for this association. Consequently, these were considered probable strong confounders and were therefore incorporated in to logistic regression analysis. This identified FLIPI risk group and response to 1st treatment as confounders and so these were included in the final adjusted model for the risk of transformation.

For the association between rs6457327 and *time* to transformation, the five potential confounders detailed in **Table 3.10** were incorporated in to multivariate analysis using Cox regression. None of the potential confounders altered the Cox regression hazard ratio by ≥ 10% so the final adjusted model includes only rs6457327 genotype.

Importantly, the two final adjusted models (**Table 3.11**) revealed that for risk of transformation only rs6457327 AA+AC genotype [$P < .01$, adjusted OR 5.48 (1.94 - 15.51)] and progression after first therapy [$P = .01$, adjusted OR 8.63 (1.56 - 47.74)] remained predictive and only rs6457327 genotype retained its predictive value for time to transformation [$P = .02$, HR 2.25 (1.16 - 4.36)].

3.6.5 *C6orf15* near rs6457327 is not a promising candidate in FL

The initial GWAS study by Skibola and colleagues (Skibola *et al.* 2009) revealed that the 26kb region of high linkage disequilibrium (LD) marked by rs6457327 contained a single open reading frame (*C6orf15*). In view of the significant association between rs6457327 and FL outcome (transformation), analysis of *C6orf15* was performed for acquired aberrations that might influence the clinical course of FL.

Direct sequencing of *C6orf15* in DNA extracted from a series of 50 diagnostic FL tumour samples revealed no mutational events. Subsequently, the expression of *C6orf15* for both malignant and germline lymphoid tissue was determined using reverse transcriptase (RT)-PCR, as the single probe for this locus on current gene expression arrays (GeneChip® Human Genome U133 and U133 Plus 2 gene expression arrays; Affymetrix) has shown only limited efficacy in other FL series (Dave *et al.* 2004). However, expression of the spliced *C6orf15* transcript was not

Table 3.11 Final adjusted models for rs6457327 associations with transformation

	Risk of transformation		Time to transformation	
	Adjusted OR (95% CI)	*P*	Adjusted HR (95% CI)	*P*
rs6457327 genotype				
AA + AC	5.48 (1.94 - 15.51)	**<0.01**	2.25 (1.16 - 4.36)	**0.02**
CC	1		1	
FLIPI score risk group				
High	0.86 (0.24 - 3.09)	0.81		
Int / High	4.88 (0.78 - 30.60)	0.09		
Int	0.30 (0.07 - 1.17)	0.08	n.a.	n.a.
Low / Int	3.87 (0.51 - 29.22)	0.19		
Low	1			
Response to first treatment				
Progression	8.63 (1.56 - 47.74)	**0.01**		
Stable	-	-		
PR	1.37 (0.33 - 5.73)	0.67	n.a.	n.a.
CR(u)	2.17 (0.61 - 7.71)	0.23		
CR	1			

The models were generated following multivariate analysis incorporating potential confounders using logistic regression for risk of transformation and Cox regression for time to transformation. There were no cases that had stable disease as response to 1st treatment which subsequently developed disease transformation so no odds ratio data is available. Significant associations are highlighted in bold.
OR indicates odds ratio; CI, confidence interval; HR, hazard ratio; Int, intermediate; PR, partial response; CR(u), complete response unconfirmed; CR, complete response; and n.a., not applicable.

detected in malignant B-cell lines and primary tumour samples (n = 24) nor in a reactive lymph node, 2 lymphoblastoid B-cell lines (and 2 myeloid cell lines) whereas in 3 of 4 tonsil samples expression was observed (**Figure 3.4**). Neither was expression of the un-spliced transcript detected in any of the primary tumour, nor other, samples used in this study (illustrated for three FL samples in **Figure 3.5**). This absence of expression was confirmed by using three different primer combinations (***Appendix 2*** **Table 1**) on the samples (including a previously published primer pair used to demonstrate expression of the un-spliced transcript in a single lymphoblastoid cell line (Skibola *et al.* 2009)). Consequently, although *C6orf15* is unlikely to have a role in FL, the correlation between rs6457327 and transformation suggests the HLA region does contain factors important in FL outcome. Furthermore, these might help explain the selective basis of chromosome 6p gain or aUPD that are detected in more than a quarter of FL cases (Cheung *et al.* 2009b; O'Shea *et al.* 2009).

3.7 Discussion

Germline factors are increasingly recognised as having relevance in FL. Several targeted studies have investigated the role of specific SNPs in disease development (Skibola *et al.* 2007) and suggest that major and minor alleles from several SNPs confer differing lifetime risks of FL. Importantly, in the first GWAS of FL, Skibola and colleagues demonstrated a risk locus on chromosome 6p (Skibola *et al.* 2009). However, less is known as to the role of germline determinants in established disease.

The MDM2-p53 axis is critical to cell cycle control and, although SNPs may influence the function of this axis and are important in some malignancies, little is known as to their role in FL. To address this, the current study established the genotypes of *MDM2* SNP 309 and *TP53* Arg72Pro in a series of 226 FL cases. The tissue used had no evidence of significant involvement by lymphoma and so extracted DNA represented germline sequence. This is important as tumour DNA may acquire changes (such as loss of heterozygosity and mutation) which could lead to incorrect germline SNP genotype assignment. Of course, acquired changes may themselves have clinical relevance and are investigated in Chapters 4 and 5.

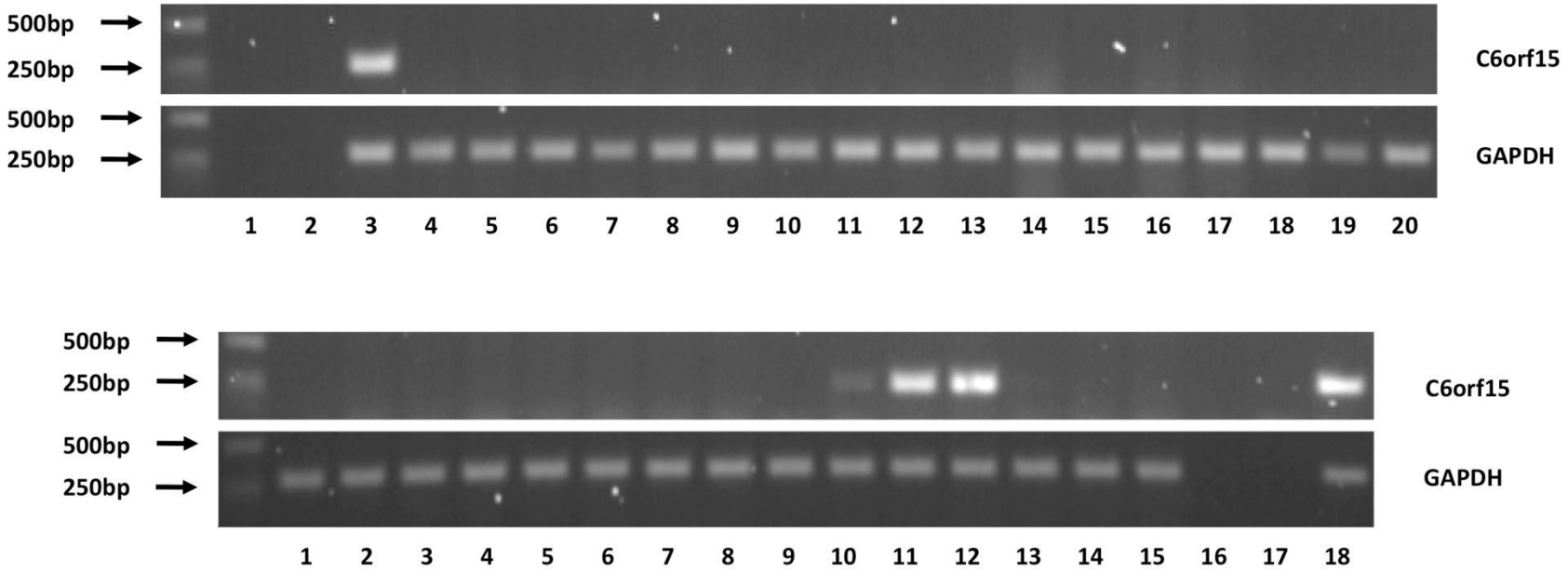


Figure 3.4 RT-PCR amplification of *C6orf15*

Top, primary tumour samples and bottom, cell lines and non-malignant tissue. *C6orf15* and *GAPDH* RT-PCR products are shown as indicated to the right. PCR reactions included oligonucleotide primers designed over splice sites in *C6orf15* and *GAPDH*. A commercial size marker (Promega) is in the left most lanes with band sizes indicated.

Top, lane 1: pooled MNC DNA (no amplification PCR negative control); lane 2: H_2O (no template PCR negative control); lane 3: cDNA from commercially prepared tonsil RNA (positive control); lanes 4 – 18: FL sample cDNA; lanes 19 and 20: t-FL sample DNA. (The FL and t-FL samples in lanes 18 and 19 respectively are from the same patient.) Bottom, lanes 1-6: B-NHL cell line cDNA [lane 1: Su-DHL4; lane 2: RL; lane3: Su-DHL6; lane 4: K422; lane 5: SC1; lane 6: DoHH2]; lanes 7 and 8: lymphoblastoid cell line cDNA [lane 7: HRC57; lane 8: Nc-Nc]; lane 9: reactive lymph node cDNA; lanes 10 to 13: non-malignant tonsil cDNA; lanes 14 and 15 myeloid cell line cDNA [lane 14: K562; lane 15: U937]; lane 16: pooled MNC DNA (no amplification PCR negative control); lane 17: H_2O (no template PCR negative control); lane 18: cDNA from commercially prepared tonsil RNA (positive control).

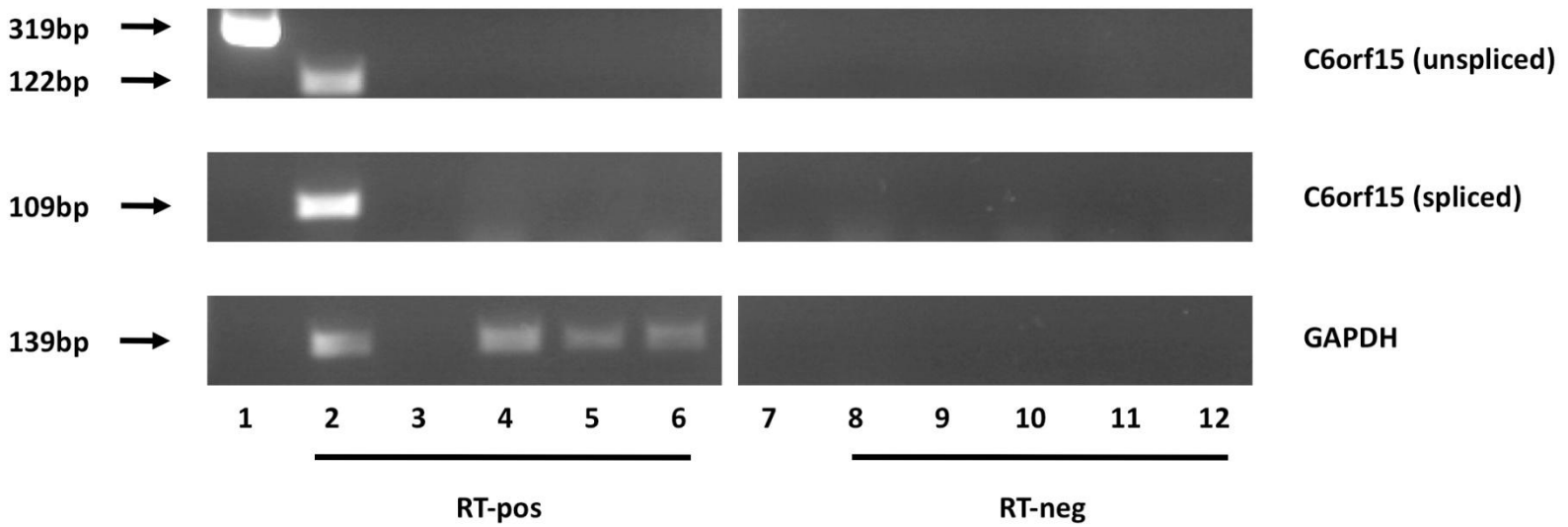


Figure 3.5 RT-PCR amplification of unspliced *C6orf15* transcript

Top, the two images include products obtained from PCR using one primer in exon1 and the other in exon 2 of *C6orf15* so that unspliced cDNA could be detected. The MNC genomic DNA control (1) amplifies a product which includes the single 197 bp intron within *C6orf15* and so represents the expected size of unspliced cDNA. Middle, for comparison the two images show PCR products from the same samples but amplified using one primer sited over the *C6orf15* splice site and so products are specific to the spliced cDNA. Bottom, the two images show PCR products obtained using *GAPDH* cDNA specific primers.

Lane 1: pooled MNC DNA (positive or no amplification PCR control for unspliced *C6orf15* or spliced *C6orf15*/*GAPDH*, respectively); lane 2: cDNA from commercially prepared tonsil RNA; lane 3: H_2O; lanes 4 – 6 cDNA from primary FL samples; lane 7: H_2O (negative PCR control); lanes 8 – 12 are the PCR products obtained from the RT-negative controls for the cDNA samples used in lanes 2 to 6. *RT indicates reverse transcriptase; pos, positive; and neg, negative.*

A challenge in the genotyping process was the number of cases classified as undetermined by the SPS software programme (using an end-point result) and so required inspection of real-time PCR curves to enable genotype determination. This was required for fewer than 10% of the *TP53* Arg72Pro results but just under 20% of the *MDM2* SNP 309 results and reflects the less distinct clustering of genotypes observed in the latter AD-PCR plots. This did not invalidate the results as curves clearly distinguished genotypes in these cases but it did add an extra interpretation step for almost 50 cases. Use of independently optimised commercial genotyping assays may help with this issue, so these were applied to genotype rs6457327 and rs10484561 which limited the need to assess real-time curves to less than 6% of cases.

With a median age younger than that typically seen in FL, the study population is likely to be affected by a degree of selection bias, reflecting this institute's status as a tertiary referral centre. This might limit the study's ability to determine the impact of SNP genotype on *later* age of FL onset. However, reports suggest that minor alleles can predispose to *earlier* onset of disease (Bond *et al.* 2004; Bond *et al.* 2006) a particularly important clinical facet which could be addressed in this cohort.

3.7.1 *MDM2* SNP 309 and *TP53* Arg72Pro

In this regard, there was no association with age at FL onset for *MDM2* SNP 309 or *TP53* Arg72Pro or the combined genotypes. Furthermore, for both *MDM2* SNP 309 and *TP53* Arg72Pro the similarity of allele frequencies to published controls (Pierce *et al.* 2000; Bond *et al.* 2005) and their concordance with HWE (suggesting stability of the SNP alleles within the study population) indicate that neither SNP predicts FL risk.

The rarity of the variant (minor) alleles (*MDM2* SNP 309 G and *TP53* Arg72Pro C) limited the significance which could be attributed to any impact of their homozygosity. Consequently, analysis was restricted to two groups: homozygosity for the major allele and presence of the minor allele. This enabled any dominant effect attributable to the minor allele, as in other malignancies (Bond *et al.* 2004), to be examined. However, *MDM2* SNP 309 and *TP53* Arg72Pro genotypes did not predict diagnostic features of FL and there were no associations with its clinical

course suggesting that the potential influence of different genotypes does not translate into poorer outcome. Importantly, although i) MDM2 expression may be modulated by *MDM2* SNP 309, ii) MDM2 over-expression is associated with t-FL and iii) earlier age of DLBCL onset is associated with *MDM2* SNP 309 (at least in younger females) no association between *MDM2* SNP 309 and either risk of or time to FL transformation was observed. In view of the lack of outcome associations, multivariate analysis was not performed and the FLIPI prognostic risk groups were not determined.

aUPD affecting chromosome 12q15 (the site of *MDM2*) occurs in 10% of all FL (O'Shea *et al.* 2009) and more frequently (19%) in FL that transforms (Fitzgibbon *et al.* 2007). As the minor allele (G) of *MDM2* SNP 309 can influence MDM2 expression (Bond *et al.* 2004), when rendered homozygous it might associate with increased tumour MDM2 protein and help explain the selective basis of 12q aUPD in FL. However, when examining a subset of study cases (n = 14) also included in a previous investigation (Davies *et al.* 2005), there was no difference in tumour MDM2 protein level across TT (n = 6) and GG (n = 3) genotypes. This is consistent with mantle cell lymphoma (MCL) where there is no association between tumour *MDM2* SNP 309 genotype and MDM2 expression (Hartmann *et al.* 2007).

The role of MDM2 SNP 309 in other cancers remains unclear, as reports of risk (Wilkening *et al.* 2007; Chua *et al.* 2010)) and outcome (Gryshchenko *et al.* 2008; Zenz *et al.* 2008) have been inconsistent. However, the varying significance in different cancers may in part be explained by a recent study that suggests *MDM2* SNP 309 genotype may have a tissue specific influence on *MDM2* mRNA expression and tumour occurrence (Post *et al.* 2010). Nevertheless, although genomic lesions targeting the MDM2-p53 axis are an important feature of FL, this study has demonstrated that *MDM2* SNP 309 and *TP53* Arg72Pro do not appear to be significant in FL.

3.7.2 rs10484561 and rs6457327

The second section of this study (n = 218) establishes that SNPs rs10484561 and rs6457327 on chromosome 6p both predict risk of FL in the UK population. The data confirms the recent GWAS (Skibola *et al.* 2009), which demonstrated that the

rs6457327 minor allele confers a protective effect against FL, and supports a follow-up study (Conde *et al.* 2010) that identified rs10484561 as a major susceptibility locus for this disease. Significantly, the current study adds to these observations by demonstrating for the first time that rs6457327 predicts clinical outcome of FL, as genotypes containing the minor (A) allele associate with a higher risk of transformation and shorter time to transformation, in an extensively followed-up population. The single other factor that could be confirmed as an independent predictor of t-FL in this study was progression (as response to initial therapy) and this correlation has been identified in other reports (Bastion *et al.* 1997; Al-Tourah *et al.* 2008).

The correlation of rs6457327 with both risk of FL and its subsequent transformation, meant that the single open-reading frame *C6orf15* (also termed *STG*) that is in high LD with this SNP (Skibola *et al.* 2009) was prioritised as a candidate gene for testing. Although limited data exists as to the role of this gene, it is expressed in skin, taste-buds and tonsils (Neira *et al.* 2001; Sanchez *et al.* 2004). More recently, its mouse homologue (*emprin*) has been shown to code for an extra-cellular matrix (ECM) protein (Manabe *et al.* 2008). This role may be relevant in cell-cell associations, which is of interest in FL due to the frequently advanced stage of disease at presentation (indicating a propensity to early metastatic processes) and importance of the surrounding immune micro-environment to outcome (Dave *et al.* 2004; Byers *et al.* 2008).

C6orf15 consists of two short exons so the entire gene could be sequenced directly. However, no mutations were identified from a cohort (n = 50) of diagnostic FL samples that included 15 cases which transformed and 12 cases with aUPD affecting the region of chromosome 6p containing rs6457327 (O'Shea *et al.* 2009). This is important as aUPD may preferentially select for mutations that confer a competitive advantage to the cell). Although its expression has been demonstrated in whole blood and a single lymphoid cell line (Skibola *et al.* 2009), transcripts were not detected in primary FL / t-FL samples, DNA from pooled peripheral blood mononuclear cells taken from healthy volunteers, a reactive lymph node sample and eight (consisting of six malignant and two virally transformed, non-malignant) lymphoid cell lines in the current study. Expression of the spliced transcript was observed in tonsillar tissue which is in keeping with previous reports (Neira *et al.* 2001; Sanchez *et al.* 2004).

Consequently, these data suggest it is unlikely that disruption of *C6orf15* in FL cells underlies the correlation of rs6457327 with transformation and its role will not be examined further by our group.

Other genetic loci linked to rs10484561 or rs6457327 might have a role in FL. In this regard, the location of rs10484561 near *HLA-DRB1* and *HLA-DQB1* in a region of high LD (***Appendix 4* Figure 2**) may help explain its significant risk association with FL. However, sufficient samples were not available for correlation of 4 digit HLA-genotypes with risk of FL (as recently demonstrated by Wang and colleagues (Wang *et al.* 2010)) or its clinical outcome. Although tag SNPs have been used to demonstrate an association between *HLA-DQB1* and FL risk (Conde *et al.* 2010), this locus was not included in the study by Wang so is incompletely characterised regarding FL risk, and its effect on clinical outcome is unknown. This should be addressed in future studies investigating germline determinants of FL, in particular whether differences between germline and tumour HLA haplotypes, for example through LOH, are relevant. Furthermore, although *C6orf15* is the only gene in the risk locus of high linkage disequilibrium (LD) (indicated by $R^2 > 0.8$) surrounding rs6457327 (***Appendix 4* Figure 1**), several candidate genes lie more distally in an area of moderate LD ($R^2 > 0.5$). Two of these, *POU5F1* and *TCF19* are transcription factors which may influence embryonic stem cell pluripotency and cell cycle progression, respectively, and *POU5F1* has been implicated in the clinical outcome of other cancers (Chang *et al.* 2008; Saigusa *et al.* 2009; Wang *et al.* 2009).

Although rs6457327 is a germline factor, acquired changes affecting this locus might contribute to the association with clinical outcome. Indeed, LOH disrupting chromosome 6p, principally by aUPD as illustrated in **Figure 3.6**, is a recurrent finding in FL tumours (O'Shea *et al.* 2009) and so might show bias for one of the rs6457327 alleles. However, a preliminary investigation of tumour samples (n = 12) known to have aUPD of this locus revealed rs6457327 AA (33%) and GG (67%) genotypes occurred at similar frequencies to allele frequencies in germline controls (**Table 3.7**) suggesting there is no selection for either allele by aUPD.

As described in Chapter 1, over thirty studies have reported SNPs associated with FL risk. However, only the recent GWAS were validated in multiple patient cohorts and the current study is the first to independently confirm the risk associations of SNPs

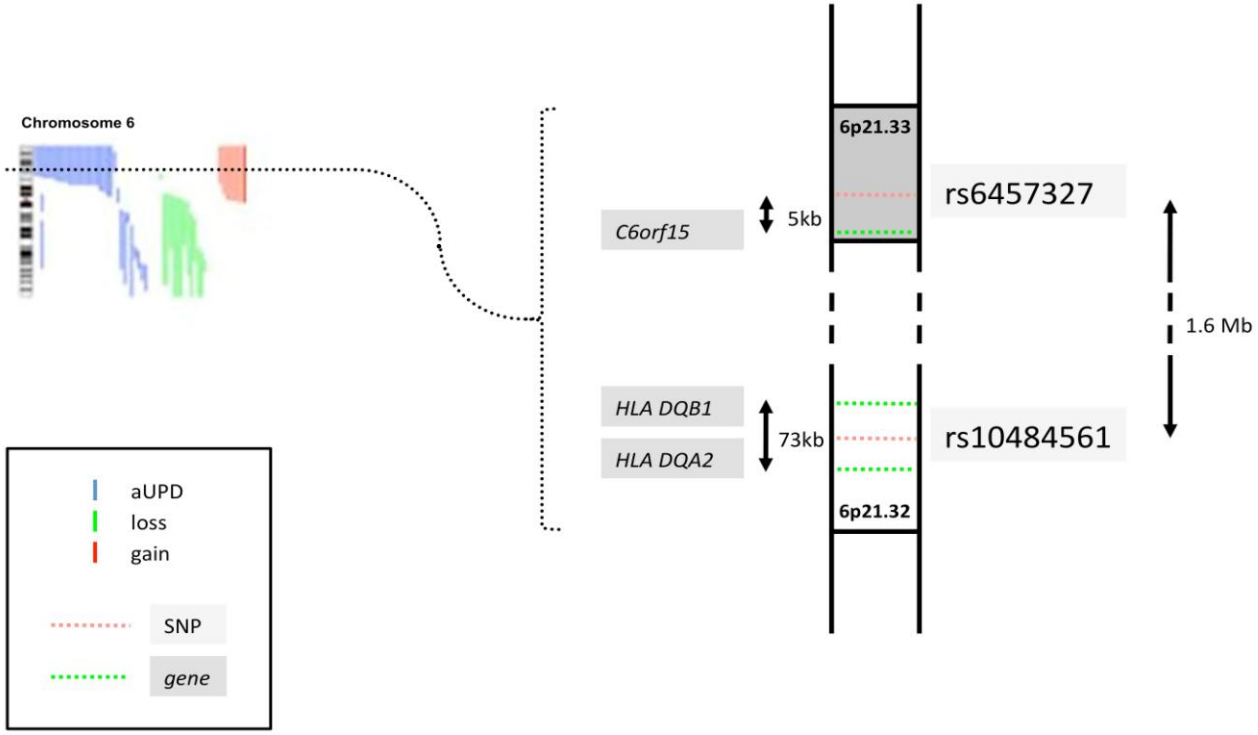


Figure 3.6 rs6457327 & rs10484561 in a region affected by aUPD or gain

On the left is a vertical G-banding map of chromosome 6 with associated regions of aUPD and gain (and loss) affecting chromosome 6 illustrated in a figure adapted from the recent study by O'Shea and colleagues (O'Shea *et al*. 2009). On the right is a schematic representation of the positions of rs6453727 and rs10484561 on chromosome 6p with the nearby genes illustrated.
aUPD indicates acquired uniparental disomy; kb, kilo base pairs; Mb, million base pairs.

rs10484561 and rs6457327 identified in those investigations. Moreover, SNP effects on disease course have generally not been tested, so fewer associations with FL outcome have been reported (Weng and Levy 2003; Jardin *et al.* 2005; Cerhan *et al.* 2007b; Hohaus *et al.* 2007; Racila *et al.* 2008; Han *et al.* 2010). Indeed, at the time of writing, there is only a single report of a SNP (within intron 1 of *BCL6* on chromosome 3q27) that associates with transformation (Jardin *et al.* 2005) and this was in an FL cohort assessed prior to the FLIPI era. The majority of these studies have included fewer cases than this thesis project (Weng and Levy 2003; Jardin *et al.* 2005; Hohaus *et al.* 2007; Han *et al.* 2010), although one study assessed a similar number (n = 133) regarding the impact of C1q SNPs specifically in the outcome of rituximab treated FL (the majority at relapse) (Racila *et al.* 2008) and a larger study assessed SNPs in immune related genes for 278 FL cases and correlated these with clinical and demographic parameters (Cerhan *et al.* 2007b) but neither of these reports included FLIPI scoring. Importantly, the inclusion of this independently validated prognostic index (Solal-Celigny *et al.* 2004; van de Schans *et al.* 2009) in the current study provides an FL population assessed according to current clinical practice so gives particular credibility to the disease outcome correlations.

As a robust clinical correlation between rs6457327 and t-FL was observed, this should help focus molecular study of transformation which remains a major cause of morbidity and mortality in FL. Moreover, the association provides further evidence as to the importance of the immune gene rich HLA region on chromosome 6p. This builds on previous reports in to the relevance of both acquired (Cheung *et al.* 2009b; O'Shea *et al.* 2009) and germline variation (Wang *et al.* 2010) at this locus in FL and the known importance of the immune response to FL outcome (Dave *et al.* 2004). Although the underlying mechanism of the rs6457327 association with t-FL remains to be determined, it represents a step forward in characterising inherited predictors of clinical outcome which, together with recently described acquired predictors (O'Shea *et al.* 2008; Cheung *et al.* 2009b; O'Shea *et al.* 2009; Cheung *et al.* 2010b), represent a growing pool of molecular markers in FL. Prospective validation is needed before their suitability can be established for incorporation into clinical-molecular prognostic (and predictive) indices which could help optimise and tailor patient management.

Chapter 4 Acquired genetic events and t-FL

4.1 Introduction

The transformation of FL is an event marked by clinical and radiological indicators suggestive of disease progression that typically lead to sampling of suspicious lesions. Histological detection of aggressive lymphoma (usually DLBCL) then provides a formal diagnosis of t-FL. Transformation represents a significant event as it confers a worse prognosis on FL (**Figure 1.3**), survival from t-FL is particularly poor (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008) and its management is often complicated by previous treatments. As t-FL is not reliably predicted by current profiling methods, a more comprehensive understanding of the underlying processes will be an important aspect of addressing the major clinical challenge posed by this event.

As described in Chapter 1 and recently reviewed by Relander and colleagues (Relander *et al.* 2010), t-FL is accompanied by genomic changes that are not detectable in FL. However, the changes reported are diverse and mirror the marked clinical heterogeneity of transformation which may develop throughout the course of FL (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008). Indeed, the risk of t-FL is maintained at around 3% per year (Al-Tourah *et al.* 2008) which, with the median age at FL diagnosis of around 60 years, hints that most FL have a propensity to transform. This observation is supported by the demonstration of transformation in most FL cases examined at autopsy (Garvin *et al.* 1983).

So, if FL has an intrinsic *potential* for transformation then why do only certain cases manifest this clinically over the course of their disease? The identification in Chapter 3 that germline polymorphic variation associates with transformation, together with reports that secondary genetic changes present at FL diagnosis can predict this event (Cheung *et al.* 2009b; O'Shea *et al.* 2009), supports a model in which the combined influence of inherited and acquired predisposing genetic variation contributes to subsequent t-FL development. However, as not all cases with these risk factors go on to present with transformation it is clear that additional changes are necessary.

Indeed, genetic studies comparing sequential FL / t-FL samples have identified that disruption of certain genes is more prevalent in t-FL (Lo Coco *et al.* 1993; Matolcsy *et al.* 1996; Pinyol *et al.* 1998; Lossos and Levy 2000; Lossos *et al.* 2002). However, as this occurs in only a subset of cases such events cannot drive the majority of disease transformation, and their specific role remains unclear. Other studies have determined higher level chromosomal changes in t-FL (Hough *et al.* 2001; Martinez-Climent *et al.* 2003; Fitzgibbon *et al.* 2007) but, in part due to the limits of resolution, these remain a diverse group of aberrations that have only rarely identified potential key underlying events. Moreover, gene expression profiling indicates that transformation may develop through differing functional modes (Lossos *et al.* 2002; Elenitoba-Johnson *et al.* 2003; Davies *et al.* 2007), paralleling this genetic diversity.

Recent advances in genetic profiling techniques include genome-wide interrogation by micro-arrays containing many thousands of DNA probes, which are included on current platforms at a median genomic spacing of 700 bp (as compared to the 300 kb spacing used in initial studies of FL (Fitzgibbon *et al.* 2007)). This now enables high resolution mapping of acquired DNA alterations in tumour genomes. By examining sequential FL / t-FL samples in combination with matched germline DNA, this approach could enable better characterisation of the genetic changes associated with t-FL and help focus on important disease mechanisms. Furthermore, it might reveal events in FL that are suitable for further examination in large homogenously managed case cohorts to determine their clinical relevance in the prediction and management of transformation.

4.2 Aims and Objectives

This study sought to determine DNA copy number aberrations in FL / t-FL as well as examine gene mutation status and further characterise the genetic evolution of transformation.

To facilitate this, analysis was restricted to germline, FL and t-FL DNA samples obtained sequentially from a series of patients with prolonged clinical follow-up. Specific objectives included high resolution genome-wide DNA copy number determination for these samples using the Affymetrix® genome-wide SNP 6.0 micro-

array; establishing the clinical significance of copy number changes; characterising mutational events in FL / t-FL; and mapping the genetic evolution of transformation using copy number or mutation profiles. The data generated from this study has since been presented at the 52nd annual meeting of the American Society of Hematology (Wrench *et al.* 2011).

4.3 Materials and Methods

4.3.1 Patients and samples

Cases with sequential FL / t-FL samples were included for genome-wide array analysis based on the availability of tissue samples. Each had at least one (earliest available) FL and a subsequent t-FL sample with additional FL / t-FL and germline samples included where possible. In total 90 samples from 31 cases were examined along with the WSU-NHL cell line. In a further cohort of 56 presentation samples, mutation frequencies at FL diagnosis were assessed. An additional set of DNA samples (n = 31) was included from the study performed in our institution for the LLMPP (O'Shea *et al.* 2009).

4.3.2 DNA

DNA was extracted and stored as previously described (section 2.4.1).

4.3.3 PCR

Standard (qualitative) PCR was performed as detailed in section **2.4.3** using either *Taq* polymerase (Promega) with proprietary buffer (final concentration 1×), dNTPs (500µM), $MgCl_2$ (1.5mM), primers (400nM) with 10 ng DNA in reaction volume of 20µl or (later in the project) HotStarTaq Plus Mastermix with 2.2 ng DNA in a final volume of 10µl. PCR products were then visualised following agarose gel electrophoresis as described in section **2.4.8**.

4.3.4 Sequencing

Following visualisation of distinct bands, the PCR products were purified using the ExoSap protocol (**Table 2.3**) then directly sequenced using BigDye® Terminator chemistry (section **2.4.10**). The primers used in PCR and sequencing are given in ***Appendix 2*** **Table 1**.

4.3.5 DNA micro-arrays

High resolution DNA copy number assays were performed on the sequential FL / t-FL samples using 0.5µg of high quality DNA (section **2.5.1**). To enable successful generation of adequate hybridisation signals on the micro-arrays, PCR amplification was performed following DNA restriction enzyme digestion and PCR-adaptor ligation at the restriction sites (sections **2.5.2** to **2.5.4**). PCR products were then pooled and purified (section **2.5.5**) before being fragmented using DNase I, with the fragments then biotin-labelled ahead of hybridisation to the micro-array (sections **2.5.7** to **2.5.9**). After an overnight incubation, hybridisation signals were established by application of phycoerythrin labelled streptavidin (sections **2.5.10**). Quality control steps were included at the DNA preparation (section **2.5.1**), PCR / quantitation (**2.5.5** and **2.5.6**) and fragmentation stages (**2.5.7**).

4.3.6 Interpretation of copy number aberrations

Raw data files (.cel files) were established from the hybridisation signals generated from the micro-arrays using the methods and quality control (QC) steps detailed in section **2.5.11**. Interpretation and analysis of the data was performed using the Partek® Genomics Suite and an institutional software package: GOLF available at https://bioinformatics.cancerresearchuk.org/~cazier01/Golf.html (as described in sections **2.5.12** to **2.5.14**). Genomic coordinates and gene annotations were applied according to human genome version 18 (hg18), which corresponds to the National Center for Biotechnology Information (NCBI) genome build 36 on which the Affymetrix micro-arrays used in this study had been designed.

Summary plots of DNA copy number aberrations (CNAs) were generated using an on-line interface maintained by Dr J-B Cazier, The Wellcome Trust Centre for

Human Genetics, University of Oxford and accessible at http://www.well.ox.ac.uk/~jcazier/GWA_View.html. This platform was also used to determine the genomic coordinates of "minimum common regions" (MCRs) of CN gain or loss that were common to more than one sample and so might contain genes relevant in the progression of FL / t-FL.

The number of CNAs in FL and t-FL was compared using two-way ANOVA and recurrent minimum common regions (MCRs) of copy number gain or loss were compared to disease outcome parameters using Kaplan-Meier plots and the log-rank test statistic. CNAs and MCRs were then examined for their frequency, gene associations (with reference to hg18 using the genome browser website www.ensembl.org), clinical outcome association and pattern of acquisition.

The copy number estimates obtained through Partek® analysis for each sample were directly viewed in GOLF which enabled selection of appropriate raw copy number dot-plots for use in illustrative figures of results.

4.4 Results

4.4.1 Patient samples for copy number profiling

For genome-wide examination of DNA copy number aberrations in FL / t-FL, 90 samples were available from 31 cases. Each case had FL and t-FL biopsies, 19 having germline samples. Further samples included additional FL biopsies subsequent to FL but prior to transformation (n=6), relapse of t-FL (n=1) and FL recurrence after t-FL (n=2). The majority of samples (n = 66) where obtained from lymph nodes, with the remainder sourced from bone marrow (n = 15), peripheral blood (n = 2), skin (n = 2), buccal mucosa (n = 2), breast (n = 1), abdominal mass (n = 1) or spleen (n = 1) (**Table 4.1**). The time-points at which study samples were taken relative to the individual disease course for respective cases are indicated in **Figure 4.1**; the initial FL samples were the earliest available and occurred at diagnosis (n = 7) and pre-treatment (n = 8) with remainder from episodes of FL progression or relapse after treatment.

For this cohort of FL patients, the median age at diagnosis was 46 years (range 22-71) and, with median follow-up of patients who remained alive (n = 8) of 10 years

Table 4.1 Study cases and samples

Study Number	Sample source	Sample type
1	LN	FL1
	LN	t-FL
2	BM	Germline
	LN	FL1
	LN	t-FL
4	LN	FL1
	LN	t-FL
5	LN	FL1
	LN	t-FL
6	BM	Germline
	LN	FL1
	LN	t-FL
7	Skin	FL1
	LN	t-FL
8	BM	Germline
	LN	FL1
	LN	t-FL
9	PB	Germline
	LN	FL1
	LN	t-FL
	LN	t-FL2
10	LN	FL1
	LN	t-FL
	LN	FL post t-FL
11	LN	FL1
	LN	FL2
	LN	t-FL
12	LN	FL1
	LN	t-FL
13	BM	Germline
	LN	FL1
	LN	t-FL
15	BM	Germline
	LN	FL1
	LN	t-FL
16	BM	Germline
	LN	FL1
	LN	t-FL
17	BM	Germline
	LN	FL1
	LN	t-FL
18	LN	FL1
	LN	t-FL
19	BM	Germline
	LN	FL1
	LN	t-FL
20	Buccal	Germline
	LN	FL1
	Abdo mass	t-FL
21	LN	FL1
	LN	t-FL
	Breast	FL post t-FL
22	LN	FL1
	LN	t-FL
23	BM	Germline
	LN	FL1
	LN	FL2
	Spleen	t-FL
24	BM	Germline
	Skin	FL1
	LN	t-FL
25	LN	FL1
	LN	t-FL
26	BM	Germline
	LN	FL1
	LN	t-FL
27	LN	FL1
	LN	t-FL
28	BM	Germline
	LN	FL1
	LN	t-FL
30	BM	Germline
	LN	FL1
	LN	t-FL
31	PB	Germline
	LN	FL1
	LN	t-FL
32	Buccal	Germline
	LN	FL1
	LN	FL2
	LN	FL3
	LN	FL4
	LN	t-FL
33	BM	Germline
	LN	FL1
	LN	FL2
	LN	t-FL
34	BM	Germline
	LN	FL1
	LN	t-FL

LN indicates lymph node; BM, remission bone marrow; PB, remission peripheral blood; Buccal, oral epithelium via saliva; Abdo, abdominal;, FL1, earliest available FL sample; FL2, 2nd available FL sample; FL3, 3rd available FL sample; FL4, 4th available FL sample; t-FL, earliest available t-FL sample; t-FL2, 2nd available t-FL sample (at relapse of transformation); FL post t-FL, earliest available relapse FL sample after t-FL.

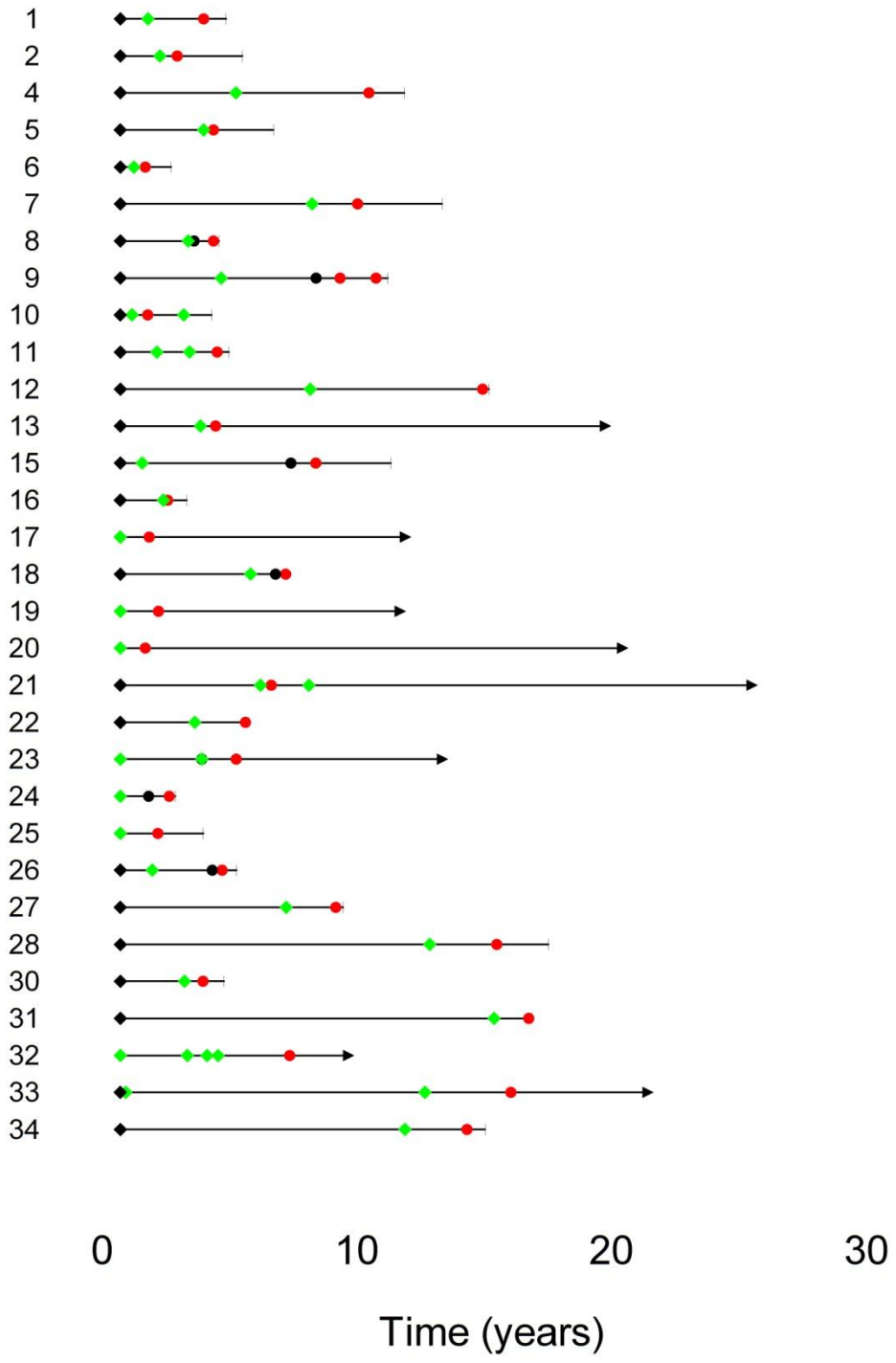


Figure 4.1 Study sample time-lines

Illustrated for each of the 31 study patients are the time points at which the sequential samples were obtained. Patient numbers are to the left of the figure. For each patient the time-line commences at FL diagnosis (left hand diamond). FL samples are illustrated by green diamonds and t-FL samples by red spots. If the diagnostic FL sample was not available for this study the left hand diamond is black. The left most spot represents t-FL diagnosis (also coloured black if the sample was not available for this study.

(2-19), the median time to transformation was 4 years (1-16) which is similar to recent large series (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008)). Cause of death was available for 22 / 23 patients who died, with lymphoma (n = 17) and treatment (n = 4) constituting 95% of deaths and one case dying from a second malignancy.

4.4.2 Micro-array quality control

Quality control (QC) data were satisfactory and are presented in ***Appendix 5*** **Table 1**. The median call rate of 98.99% (range 96.59 to 99.61), with 85 / 90 (94%) samples having rates >97%, indicated good overall applicability of the assay from DNA preparation to micro-array probe hybridisation signal generation. Contrast QC was satisfactory at a median of 1.94 (range 0.58 to 3.18) with only 6 samples (7%) < 1.0. The MAPD results had a median of 0.29 (range 0.23 to 0.40) with only one sample above 0.4. A final QC step was performed in GOLF by comparing the signals of 100 randomly selected SNPs in each of the 90 samples. This generated a dendrogram (hierarchical tree) of the relatedness of all samples and is illustrated in **Figure 4.2**. It reveals that, for each case, the respective samples cluster together indicating that the samples did indeed originate from the same patient.

4.4.3 Copy number aberrations in FL / t-FL

In total, 1114 of CNAs were identified across all 71 tumour samples including 511 CN gains and 603 CN losses. The median number of CNAs per initial FL sample was 11 (range 1 - 43) and per t-FL sample was 16 (range 0 to 56); with gains at a median of 5 (0 to 13) and 8 (0 to 33) and losses at a median of 5 (0 to 39) and 7 (0 to 30) in FL and t-FL respectively. All samples except one t-FL (Case 23) showed CNAs of at least one chromosome. CN losses occurred on every chromosome and only chromosome 4, in FL, was unaffected by CN gains.

Four cases had additional FL samples available which showed heterogeneity in their CNA acquisition compared to the preceding FL: for case 11 the second FL had more (n = 14) CNAs compared to the initial FL (n = 6); for case 23 the second sample had fewer CNAs (n = 3) as compared to the initial sample (n = 41); in the four sequential

Figure 4.2 Confirmation of sample origin

Dendrogram generated in the GOLF software package by comparing genotypes of 100 randomly selected SNPs across the 31 study cases. The respective samples from each case cluster together, illustrated by alternating blue / red, indicating they are highly related and do indeed arise from the same patient. Differences between germline and tumour samples for each case give the divergence seen to the left of the plot.

FL samples available for case 32 the CNAs numbered 5, 2, 2 and 5 respectively; and the second FL had slightly fewer CNAs (n = 2) than the preceding FL (n = 4) for case 33.

The two FL samples (for cases 10 and 21) occurring post-t-FL showed more CNAs compared to the preceding FL (33 versus 24 and 23 versus 16 for cases 10 and 21, respectively) but not the t-FL (which had 32 and 56 CNAs respectively). The sole additional t-FL sample (in case 9) showed a marked increase in the number of CNAs (n = 39) compared to the preceding FL (n = 6) and t-FL (n = 5) samples.

To compare FL and subsequent t-FL, the sets of 31 earliest FL and corresponding earliest t-FL samples were examined and no differences were observed in the overall frequencies of gains or losses or total CNAs between the FL and t-FL sets (**Figure 4.3**). The complete set of CNAs for the earliest FL samples and earliest t-FL samples are illustrated **Figure 4.4** and **Figure 4.5**, respectively. The minimum common regions (MCRs) of gain or loss (defined as the common areas affected by losses or gains that overlap in different samples) that were present in at least 20% cases ($n \geq 7$) were used to identify the sites most frequently subject to CNAs. In FL, MCR gains (n = 15) and losses (n = 11) were identified across 9 chromosomes (**Table 4.2**) and in t-FL MCR gains (n = 20) and losses (n = 7) were identified across 10 chromosomes (**Table 4.3**). The most frequent CNAs were loss 1p36 and gains 7p, 7q, 18p, 18q in FL samples ($n \geq 10$); and loss 1p36 and gains 2p15, 18p, 18q, Xq in t-FL samples ($n \geq 10$).

4.4.4 Clinical outcome associations

To examine whether these recurrent CNAs might predict clinical outcome, their associations with overall survival (OS) and time to transformation (TTT) from FL diagnosis (FL CNAs) and survival from transformation (SFT) (t-FL CNAs) were determined. Of the 57 MCR gains and losses, gain 2p15-16 in FL samples (n = 7), predicted overall survival (**Figure 4.6**) with a median survival of 5.2 years compared to 12.9 years for those cases without this CNA. No cases with this gain remained alive beyond 10.9 years whereas over half of cases (13/24) without the gain survived for more than 11.1 years post FL diagnosis. The MCR of gain 2p15-16 is just over 2 Mb in size and contains 13 genes, of these *REL* and *BCL11A* are established

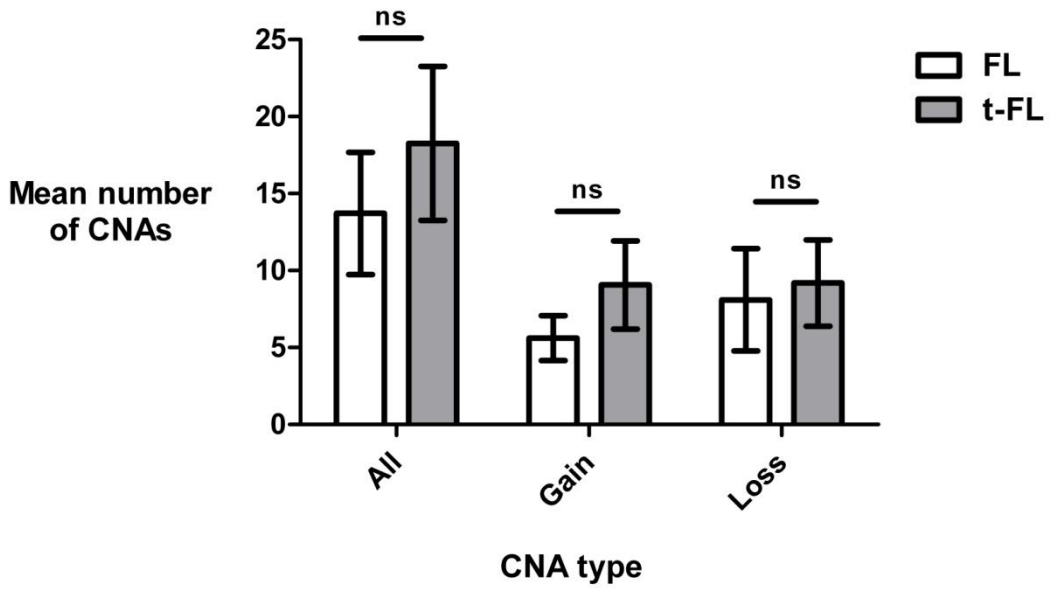


Figure 4.3 Comparison of CNAs in FL and t-FL

The mean per case (vertical bars) with 95% confidence intervals (vertical lines with caps) is illustrated for the number of CNAs (all gains and losses, gains only and losses only) across the 31 samples in the earliest FL and earliest t-FL sets. There were no significant differences between the groups.
ns indicates P value is not significant.

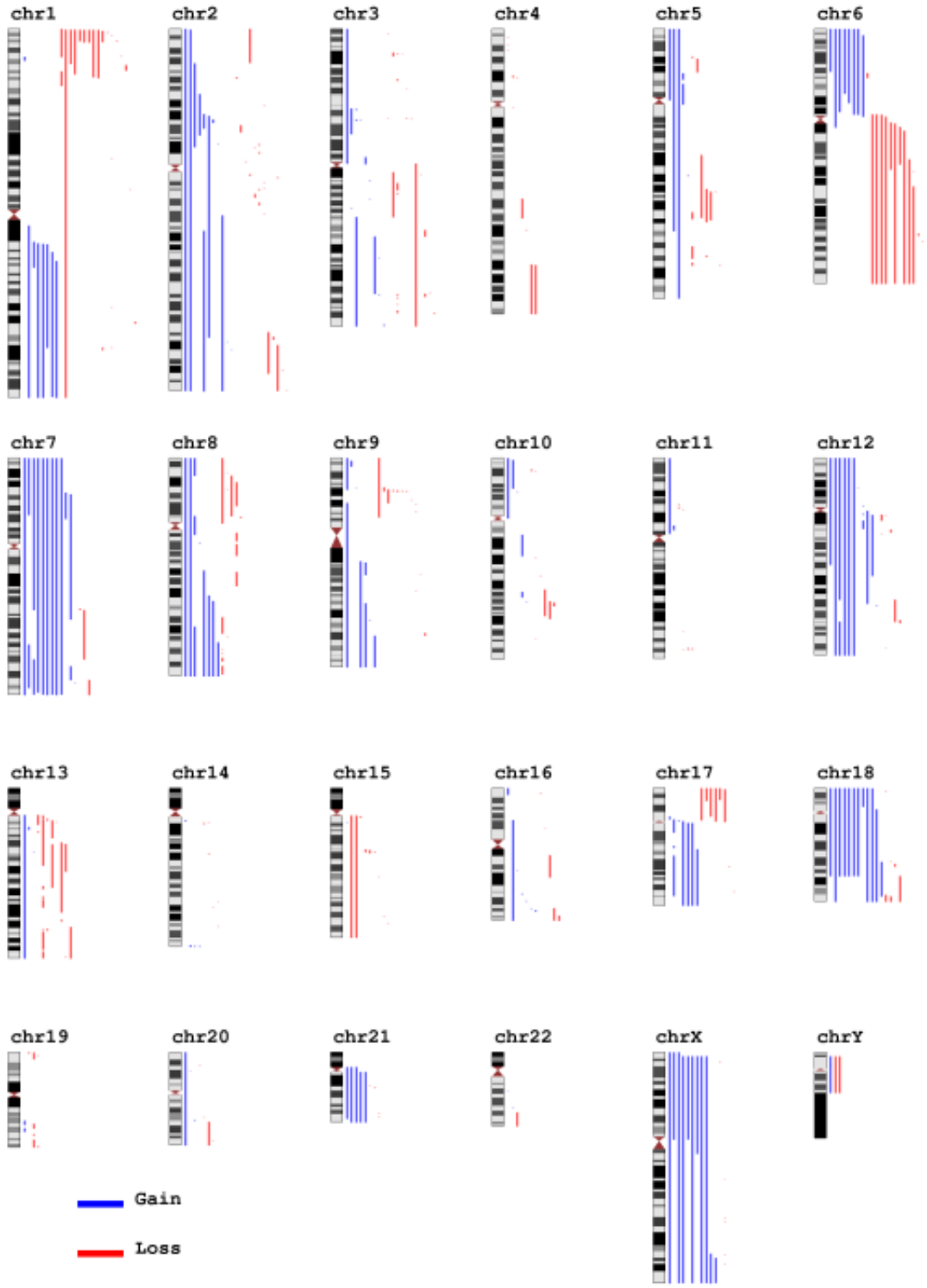


Figure 4.4 Summary of all CN gains and losses in initial FL

Gains and losses are illustrated as vertical bars positioned by their start and end co-ordinates with respect to the G-band representation of the chromosome to the left in each plot. All co-ordinates and mapping are done with respect to human genome mapping hg18 (NCBI build 36).

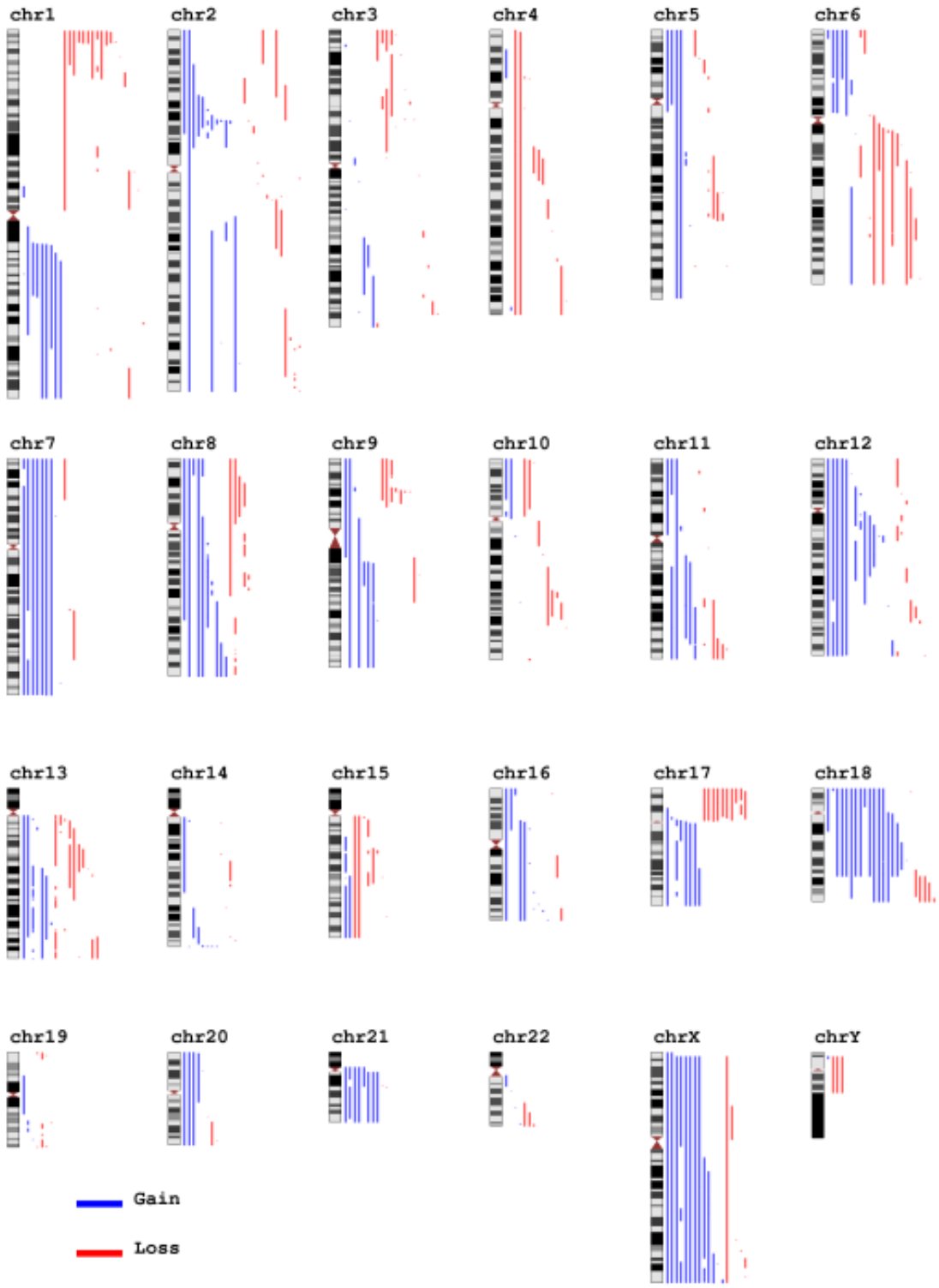


Figure 4.5 Summary of all CN gains and losses in t-FL

The plot layout is as described for **Figure 4.4**.

Table 4.2 Minimum common regions of CN gain or loss in initial FL

Chr	Start	End	Length	Type	Cyto-band	No. Genes	Candidate genes	Cases	Case numbers involved
1	2430024	2617647	187623	Loss	1p36.32	4	*PANK4, HES5, TNFRSF14, MMEL1*	11	5, 7, 10, 13, 15, 16, 20, 21, 27, 33, 34
1	4058258	4071424	13166	Loss	1p36.32	0	-	11	5, 10, 13, 15, 16, 20, 21, 23, 27, 33, 34
1	7688809	7769768	80959	Loss	1p36.23	3	*CAMTA1, VAMP3, PER3*	11	1, 5, 10, 13, 15, 16, 20, 21, 27, 33, 34
1	8353401	8432517	79116	Loss	1p36.23	1	*RERE*	10	5, 10, 13, 15, 16, 20, 21, 27, 33, 34
1	18449498	18476791	27293	Loss	1p36.13	1	*IGSF21*	7	10, 15, 16, 20, 21, 27, 31
1	155604487	159964392	4359905	Gain	1q23.1 - 1q23.3	>50	*CD1, NHLH1*	7	6, 7, 10, 13, 16, 22, 27
2	60841274	62878002	2036728	Gain	2p16.1 - 2p15	13	*REL*	7	5, 10, 15, 16, 18, 22, 30
6	4345108	28648043	24302935	Gain	6p25.1 - 6p22.1	>100	*DEK, JARID2, ID4, SOX4*	8	6, 15, 16, 17, 21, 22, 23, 27
6	68316017	87472057	19156040	Loss	6q12 - 6q14.3	50	*IBTK, TTK*	8	7, 15, 20, 22, 23, 27, 33, 34
6	95662927	95709520	46593	Loss	6q16.1	0	-	8	4, 13, 15, 20, 22, 23, 27, 33
6	137342963	137728730	385767	Loss	6q23.3	3	*IFNGR1, IL20RA, IL22RA2*	9	4, 13, 15, 16, 20, 22, 23, 27, 33
6	138023953	138459950	435997	Loss	6q23.3	2	*PERP, TNFAIP3*	9	4, 13, 15, 16, 20, 22, 23, 27, 33
6	142365293	142454020	88727	Loss	6q24.1	1	*NMBR*	9	4, 5, 13, 15, 20, 22, 23, 27, 33
7	24327356	38203836	13876480	Gain	7p15.3 - 7p14.1	68	*CREB5*	11	1, 5, 6, 8, 10, 15, 17, 19, 22, 27, 32
7	139605151	148810234	9205083	Gain	7q34 - 7q36.1	67	*BRAF, EZH2, (TRB)*	10	1, 5, 6, 10, 15, 17, 19, 22, 27, 32
7	150603168	150622306	19138	Gain	7q36.1	1	*SMARCD3*	10	1, 5, 6, 10, 15, 17, 19, 22, 27, 31
8	123643836	146264219	22620383	Gain	8q24.13 - 8q24.3	>100	*MYC, PTK2, MAPK15, NFKBIL2*	7	1, 4, 5, 10, 11, 22, 31
9	21891747	22197146	305399	Loss	9p21.3	3	*CDKN2A, CDKN2B, MTAP*	7	2, 8, 11, 18, 26, 28, 31
12	20125482	20129995	4513	Gain	12p12.2	0	-	7	1, 5, 13, 15, 18, 22, 32
12	31760458	32007644	247186	Gain	12p11.21	1	*AMN1*	7	1, 5, 10, 15, 18, 22, 32
12	37536490	37970109	433619	Gain	12q12	2	*CPNE8, KIF21A*	8	1, 5, 10, 12, 15, 18, 22, 32
12	41955589	47430170	5474581	Gain	12q12 - 12q13.12	32	*ARID2, HDAC7*	9	1, 5, 10, 12, 15, 18, 20, 22, 32
18	148226	12395119	12246893	Gain	18p11.32 - 18p11.21	42	*USP14, YES1*	10	4, 6, 9, 12, 16, 19, 25, 26, 28, 31
18	49444874	58857002	9412128	Gain	18q21.2 - 18q21.33	31	*TNFRSF11A, MALT1*	11	4, 6, 8, 9, 12, 16, 25, 26, 28, 30, 31
X	2713391	58319642	55606251	Gain	Xp22.33 - Xp11.1	>100	*UTX, JARID1C*	9	6, 13, 15, 18, 22, 25, 26, 27, 32
X	137893124	154578240	16685116	Gain	Xq27.1 - Xq28	>100	*IKBKG,IRAK1*	8	12, 13, 15, 18, 20, 25, 27, 32

Shown are the genomic start / end chromosome coordinates, length and cyto-band location of the MCRs of copy number (CN) loss or gain in the set of 31 initial FL samples. The number of genes within each MCR and examples of possible candidates for involvement in FL / t-FL (identified by their known function + / - disruption in other malignancies) are indicated together with the amount and identity of the cases involved. *Chr indicates chromosome*.

Table 4.3 Minimum common regions of CN gain or loss in t-FL

Chr	Start	End	Length	Type	Cyto-band	No. Genes	Candidate genes	Cases	Case numbers involved
1	1406566	6202339	4795773	Loss	1p36.32 - 1p36.31	37	*CDC2L2, TNFRSF14, TP73*	11	5, 10, 11, 13, 15, 16, 19, 20, 26, 27, 34
1	7688809	7982611	293802	Loss	1p36.23	6	*CAMTA1, VAMP3, PER3, UTS1, TNFRSF9, PARK7*	11	1, 5, 10, 13, 15, 16, 19, 20, 26, 27, 34
1	163839348	164031598	192250	Gain	1q24.1	3	*MGST3, ALDH9A1, TMCO1*	9	1,10, 13, 16, 17, 21, 22, 26, 27
2	60899425	61552674	653249	Gain	2p15	5	*REL, PUS10, PEX13, AHSA2, USP34*	11	1, 4, 5, 10, 15, 16, 17, 18, 21, 25, 30
6	87472057	91385530	3913473	Loss	6q14.3 - 6q15	24	*CASP8AP2, MAP3K7*	7	1, 13, 17, 20, 21, 24, 34
6	137018030	138821596	1803566	Loss	6q23.3	12	*IFNGR1, IL20RA, IL22RA2, PERP, TNFAIP3*	7	1, 4, 8, 13, 20, 24, 25
7	138342	6660808	6522466	Gain	7p22.3 - 7p22.1	59	*CARD11, FAM*	7	1, 5, 10, 15, 17, 19, 27
7	18357112	101876473	83519361	Gain	7p21.1-7q22.1	>100	*HDAC9, CREB5, IGFBP1/3, EGFR*	7	1, 5, 10, 15, 17, 19, 27
7	150603168	150622306	19138	Gain	7q36.1	1	*SMARCD3*	8	1, 5, 10, 15, 17, 19, 27, 31
9	21813210	22146338	333128	Loss	9p21.3	3	*CDKN2A, CDKN2B, MTAP*	7	7, 10, 18, 26, 27, 28, 31
12	23344385	24396333	1051948	Gain	12p12.1	1	*SOX5*	7	1, 5, 6, 10, 15, 18, 32
12	37536490	37970109	433619	Gain	12q12	2	*CPNE8, KIF21A*	8	1, 5, 6, 10, 12, 15, 18, 32
12	44391098	45110683	719585	Gain	12q12 - 12q13.11	4	*ARID2, SFRS2IP, SLC38A1, SLC38A2*	10	1, 5, 6, 10, 12, 15, 18, 20, 22, 32
12	51714508	51844412	129904	Gain	12q13.13	6	*CSAD, EIF4B, IGFBP6, SOAT2, SPRYD3, TENC1*	10	1, 5, 6, 7, 12, 18, 20, 21, 22, 32
12	55116404	56037731	921327	Gain	12q13.2 - 12q13.3	26	*STAT6*	10	1, 5, 6, 7, 12, 15, 18, 20, 22, 32
17	1874025	7888604	6014579	Loss	17p13.3 - 17p13.1	>100	*TP53, CD68, JMJD3, TNFSF12*	10	4, 6, 8, 9, 13, 15, 17, 20, 21, 24
17	10881798	17567464	6685666	Loss	17p13.1 - 17p11.2	40	*TNFRSF13B (TACI), RASD1*	9	4, 6, 8, 9, 13, 15, 20, 21, 24
18	1543	1567856	1566313	Gain	18p11.32	9	*USP14, YES1*	13	4, 8, 9, 12, 15, 16, 17, 19, 24, 25, 26, 28, 31
18	23335186	28890957	5555771	Gain	18q12.1	16	*MEP1B*	14	6, 8, 9, 12, 13, 15, 16, 17, 24, 25, 26, 28, 30, 31
18	48387638	48415217	27579	Gain	18q21.2	1	*DCC*	15	6, 8, 9, 10, 12, 13, 15, 16, 17, 22, 24, 26, 28, 30, 31
21	13266376	18159067	4892691	Gain	21p11.1 - 21q21.1	10	*USP25*	8	1, 12, 13, 16, 17, 18, 30, 34
21	23133880	44590433	21456553	Gain	21q21.2 - 21q22.3	>100	*ERG, ETS2, DYRK1A*	7	1, 12, 13, 16, 17, 18, 30
X	2713391	63173777	60460386	Gain	Xp22.33 - Xq11.1	>100	*UTX, JARID1C*	8	1, 6, 12, 13, 15, 18, 27, 32
X	104841149	113186400	8345251	Gain	Xq22.3 - Xq23	42	*IRS4, MYCL2*	10	1, 6, 12, 13, 15, 18, 21, 27, 28, 32
X	135045986	138303981	3257995	Gain	Xq26.3 - Xq27.1	12	*CD40LG, FGF13*	10	6, 12, 13, 15, 18, 20, 21, 27, 28, 32
X	142106349	142197820	91471	Gain	Xq27.3	0	-	11	1, 4, 6, 12, 13, 15, 18, 20, 21, 27, 32
X	152485294	154292605	1807311	Gain	Xq28	61	*FUNDC2*	12	1, 6, 12, 13, 15, 18, 20, 21, 24, 27, 28, 32

The table layout is as indicated for **Table 4.2**.

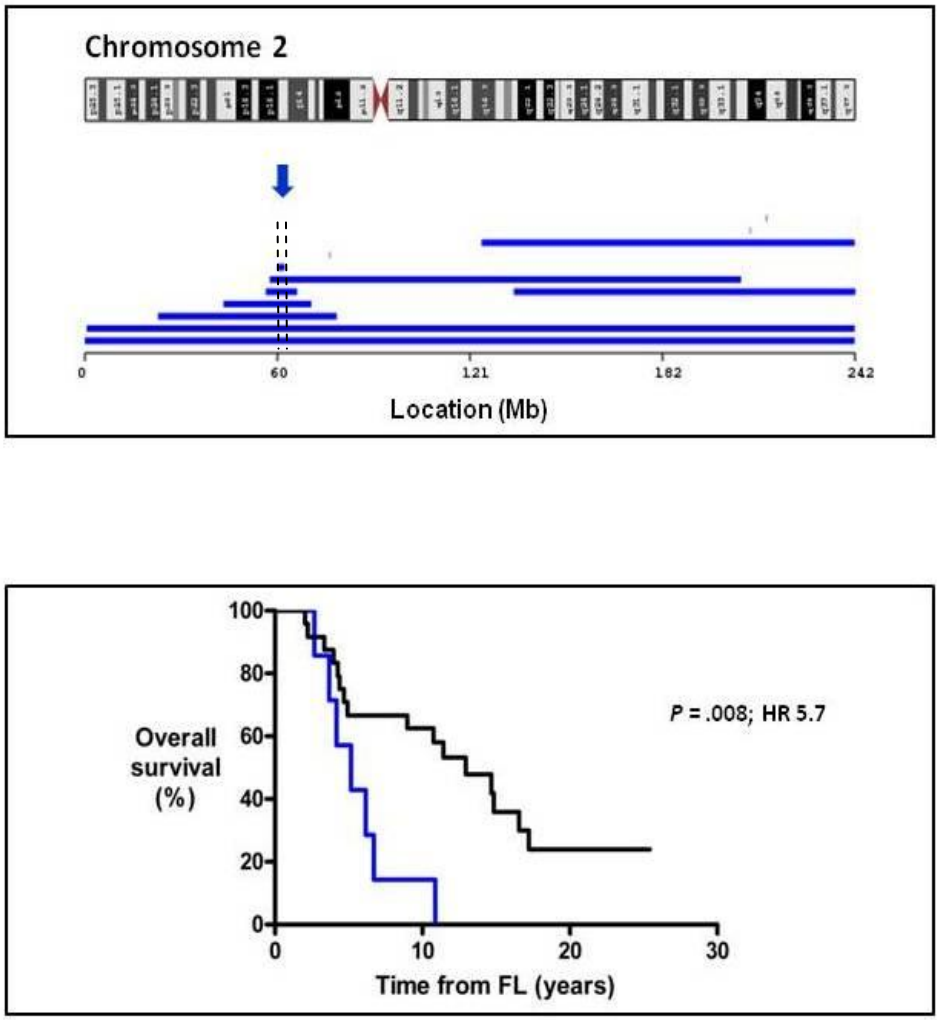


Figure 4.6 Gain 2p16.1-15 predicts OS from survival in FL that transforms

Top, the minimum common region of gain at chromosome 2p16.1-15 (marked by the vertical dashed lines and extending from 60.8 to 62.8 Mb) in 7 / 31 FL cases is indicated by the blue arrow. The horizontal blue bars represent regions of CN gain in different cases (with one case having two regions of separate gain). Chromosome cytobands and genomic positions are indicated. Bottom, Kaplan-Meier plots for the seven cases with gain at 2p16.1-15 (in blue) and the remaining cases (n = 24) all with normal copy number at this location (in black); the log rank test statistic and hazard ratio (HR) are given on the right.

oncogenes in NHL (Satterwhite *et al.* 2001) and, specifically, FL (Lu *et al.* 1991; Goff *et al.* 2000). No associations were observed with TTT or survival from transformation for gain 2p15-16 in FL or t-FL, respectively.

No other associations were observed between CNAs and the survival outcomes, although all cases with gains of either 8q24.13-q24.3 (n = 7) or 18q21.2-q21.33 (n = 11) in the initial FL sample or loss 9p21.3 in the initial FL sample (n = 7) or t-FL sample (n = 7) had died.

4.4.5 CNAs emerging at t-FL

Disruption of genes by CNAs occurring at t-FL might contribute to the transformation process and determining which MCRs of loss or gain are enriched at t-FL may help identify these potential driver events. Eight such MCRs were identified by comparing the t-FL and FL CNA sets on a chromosome-by-chromosome basis (**Table 4.4)**, four of these contained multiple genes (≥37) over a minimum 1.7Mb size and one was a micro-CNA (<30kb) but contained no known coding regions. The remaining three MCRs (**Figure 4.7**), which ranged from 0.3 to 1.4 Mb in size and contained less than ten genes (suggesting these sites might be preferentially targeted in transformation), are detailed as follows:

i. loss 11q24.2 which includes two putative tumour suppressors: *EI24* (coding a protein which has a role in TP53 mediated apoptosis (Gu *et al.* 2000), binds Bcl-2 (Zhao *et al.* 2005) and when subject to loss may confer therapy resistance (Mork *et al.* 2007)) and *CHEK1* (protein product of which inhibits CDC25 by phosphorylation leading to CDK-cyclin complex phosphorylation and cell cycle arrest (Sorensen *et al.* 2003)). Although only 4 cases were affected by this loss, they showed shorter OS from FL with no cases surviving more than 2.5 years after transformation (***Appendix 5* Figure 1**);

ii. gain 13q31.3 containing two genes: *GPC5* (coding a cell surface proteoglycan which may have a role in cell division and control of growth and which is over-expressed in lymphoma cell lines (Yu *et al.* 2003) and solid tumours (Williamson *et al.* 2007)) and *MIRHG1* (a potential oncogene (He *et al.* 2005; Olive *et al.* 2010)) encompassing the miR17-92 cluster,

Table 4.4 MCRs of loss or gain enriched in t-FL compared to FL

Chr	Start	End	Length	Type	Cyto-band	No. genes	Candidate genes[§]	t-FL cases involved	t-FL (FL) samples
2	36982238	44981413	7999175	Loss	2p22.2-p21	41	*SOS1, CEBPZ,*	1, 15, 21	3 (0)
3	3621942	3650872	28930	Loss	3p26.2	0	-	1, 4, 15, 17, 21	5 (1)
3	48089538	49873624	1784086	Loss	3p21.31	52	*TRAIP, CDC25A*	4, 17, 20, 21	4 (0)
4	86320359	96154196	9833837	Loss	4q21.23 - q22.3	37	*MAPK10*	4, 7, 9, 15, 21	5 (0)
11	99607925	120648330	21040405	Gain	11q22.1 - q23.3	>100	*MLL, ATM, IL10RA, CBL, POU2AF1 (BOB1), BCL9L*	4, 6, 21, 22, 24	5 (0)
11	124517442	125193868	676426	Loss	11q24.2	8	*EI24, CHEK1*	4, 6, 10, 22	4 (0)
13	90007394	91500384	1492990	Gain	13q31.3	2	*GPC5, MIRHG1*[#]	6, 8, 10, 18, 21, 22	6 (1)
16	27077108	27413252	336144	Gain	16p12.1	5	*IL4-R, IL21-R , NSMCE1, GTF3C1, JMJD5*	1, 7, 15, 17, 18	5 (1)

The table layout is as described for **Table 4.2**. Included here are minimum common regions (MCRs) of gain / loss that occur in at least 3 additional t-FL samples compared to the preceding FL sample and are either absent from FL or present in only one FL sample.
[§]Candidate genes were determined for each region based on their known (or predicted) function + / - their known association with other cancers.
[#]This region encompasses a cluster of micro-RNAs (also termed miR17-92) which includes mirR-17, mirR-18a, mirR-19a, mirR-19b-1, mirR-20a, mirR-92a-1.

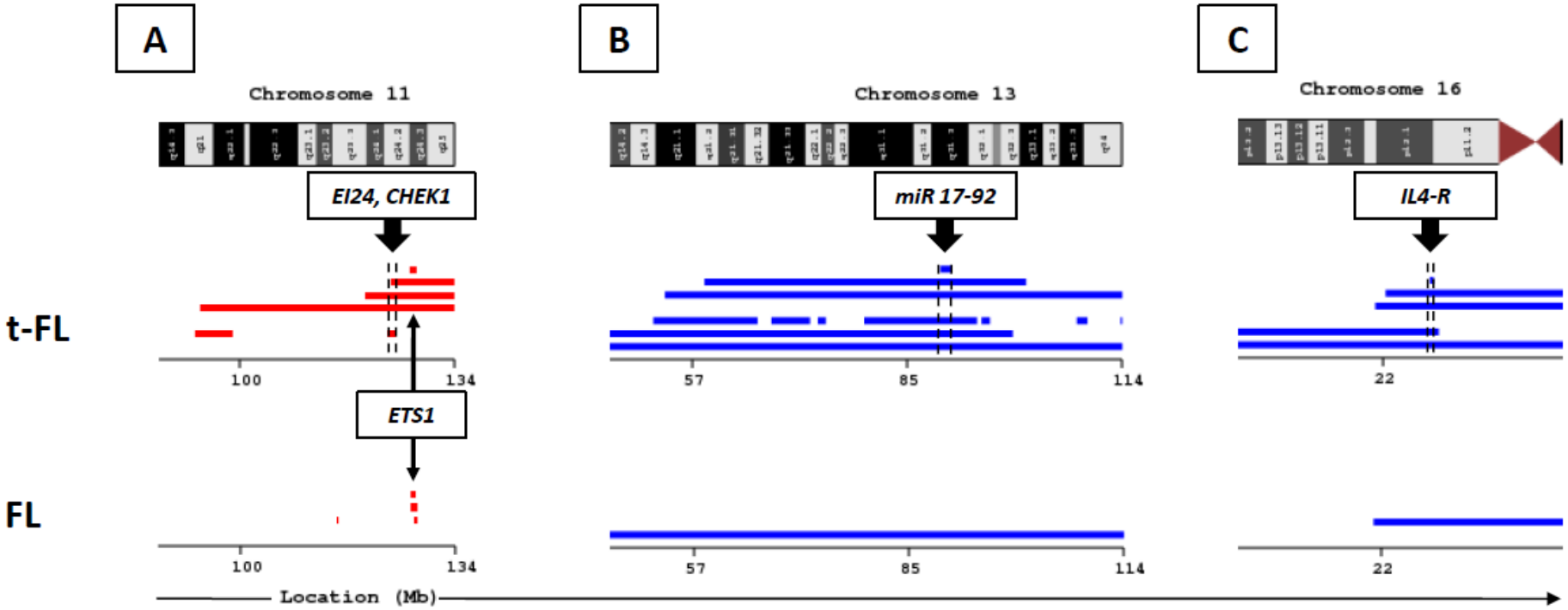


Figure 4.7 MCRs enriched in t-FL and loss *ETS1*

Examples of minimum common regions (MCRs) of copy number aberrations enriched at transformation including: A, loss chromosome 11q; B, gain chromosome 13q; and C, gain chromosome 16p. Candidate genes in the MCRs are indicated. Also shown (A) is a recurrent region of loss affecting the *ETS1* locus. The figure layout is as described for **Figure 4.6** (top) with the addition that losses are indicated by horizontal red bars.

encoding at least 6 microRNAs, a locus frequently amplified in *de novo* DLBCL (Ota *et al.* 2004) and follicular lymphoma (Neat *et al.* 2001); and

iii. gain 16p12.1 that harbours *IL4-R*, a potentially relevant locus in view of the recently described elevated IL4 expression in the micro-environment in primary FL samples (Calvo *et al.* 2008) and increased expression of IL4Rα previously observed in FL cells versus normal GC B cells (Husson *et al.* 2002).

4.4.6 CNAs affecting single genes

CNAs that disrupt single genes may specifically represent events that contribute to disease. MCRs including only one gene (detailed in **Table 4.2** and **Table 4.3**) occurred for five genes in FL: *RERE* (loss 1p36.23), *IGSF21* (Loss 1p36.13), *NMBR* (Loss 6q24.1), *SMARCD3* (Gain 7q36.1) and *AMN1* (Gain 12p11.21); and three genes in t-FL: *SMARCD3* (Gain 7q36.1), *SOX5* (Gain 12p12.1) and *DCC* (Gain 18q21.2), with the differences between the two tumour types reflecting the heterogeneity of CNAs present in FL and t-FL. Furthermore, to date none of these genes have been associated with FL or its transformation.

Expanding the examination to all CNAs revealed additional single gene events including a loss that includes the entirety of *ETS1* on chromosome 11q24.3 in FL (n = 3) and t-FL (n = 4 cases) in an MCR that is < 500 kb, as highlighted in **Figure 4.7**. Single cases demonstrated CNAs disrupting single genes, with a loss including 30 / 31 exons (with only the 3' exon unaffected) from *CREBBP* on chromosome 16p13.3 occurring in both FL and t-FL samples (**Figure 4.8**) from case 26; a 10kb loss occurred in ALK (Case 33; both FL and t-FL); MSI2 (which regulates RNA expression) was disrupted by loss 17q22 in case 16 (FL); both SLCO3A1 (15q26.1) and NDUFV3 (21q22.3) were deleted in case 23 (FL) and EBF1 (coding a transcription factor with a role in B-cell lineage commitment) was included in a 125 kb deletion at 5q33.3 in both FL and t-FL from case 26. Although not a single gene deletion, a 123 kb MCR of loss targeted the *CDKN2A / CDKN2B* locus in both FL and t-FL for 7 cases; a region known to be disrupted in FL as described in Chapter 1.

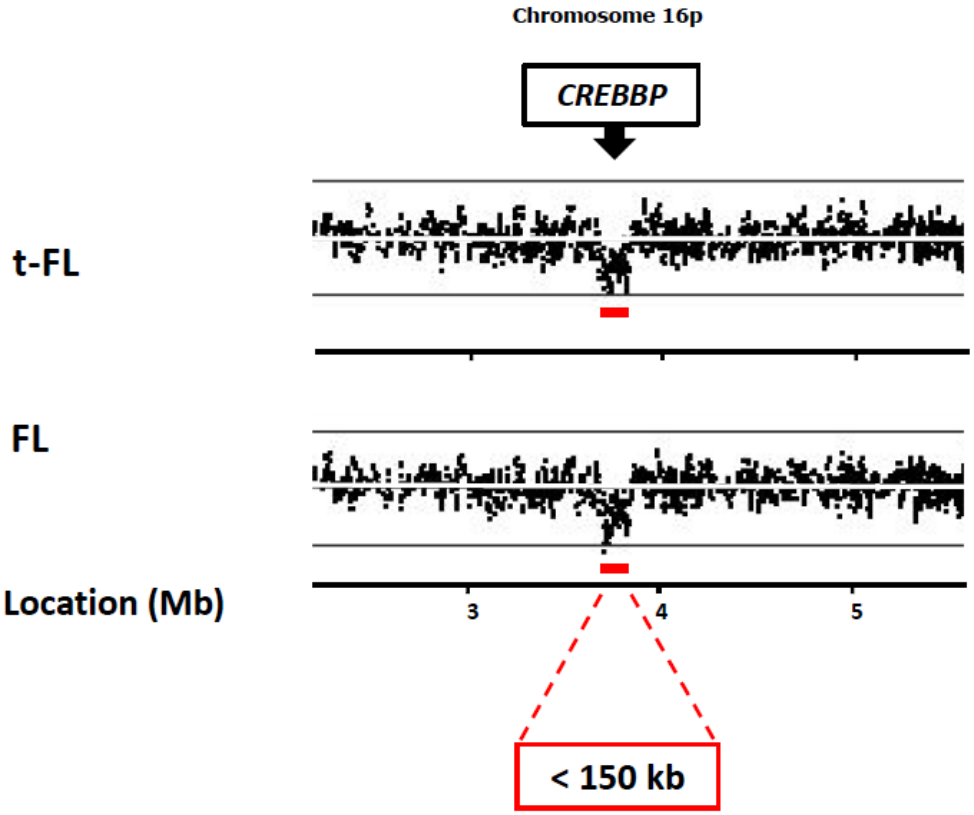


Figure 4.8 CN loss of *CREBBP*

The plots represent the estimated CN of each tumour sample for Case 26 for a ~3.5Mb section of chromosome 16p (obtained by comparison to germline signal intensities and expressed as $\log_2$ratios) for t-FL (top) and FL (bottom). Each black dot represents the $\log_2$ratio for an individual probe: the horizontal white line represents a ratio = 0 (the majority of probes cluster around this line indicating the same signal intensity in germline and tumour identifying that tumour is 2n for the majority of the section of 16p illustrated here); the two thin horizontal black lines above and below represent $\log_2$ratios of 1 and -1 respectively (a ratio of -1 indicates the relative CN in the tumour is half of that in germline and so here is 1n). The red bar highlights a small loss which disrupts *CREBBP*. Scales underneath each plot indicate genomic location.
Mb indicates megabase; and kb, kilobase.

The identification of loss targeting *CREBBP* is particularly relevant as this gene lies in a region of chromosome 16p that is subject to recurrent aUPD in FL (O'Shea *et al.* 2009) and predicts worse PFS. The gene product has a role in co-activation of transcription factors and cellular growth control as well as co-activation of p53 (Avantaggiati *et al.* 1997; Gu *et al.* 1997; Lill *et al.* 1997). As dominant negative CREBBP mutants suppress p53-dependent CDKN1A expression (Scolnick *et al.* 1997) this indicates these variants could attenuate the tumour suppressor activity of p53, a mechanism contributing to many malignancies (as discussed in Chapter 1). Furthermore, the following clinical associations suggest a role for disruption of *CREBBP* in cancer, including FL:

loss or mutation of *CREBBP* is associated with the Rubinstein-Taybi syndrome (Petrij *et al.* 1995) which includes a propensity to malignancy;

translocations involving *CREBBP* occur in AML (Borrow *et al.* 1996; Giles *et al.* 1997) and secondary malignancy (Rowley *et al.* 1997); and

the *CREBBP* locus was involved in a t(16;18)(p13;q21.3) described in one case of FL (Mahmoodi *et al.* 2004).

To investigate whether the loss of *CREBBP* observed in case 26 identifies a target that is disrupted in FL, mutational analysis was performed. As *CREBBP* is a large, gene consisting of 31 exons which would require at least 39 PCR amplicons to enable conventional sequencing of the full coding region, a search was performed for existing mutations in the Sanger Institute Catalogue Of Somatic Mutations In Cancer web site, http://www.sanger.ac.uk/cosmic (Forbes *et al.* 2008). This recorded three mutations in the coding regions of *CREBBP* in a search done on 23 February 2010 (as of January 2011 the record has been expanded to 10 missense substitution or deletion mutations across ovarian, lung and skin cancers).

These included two missense single base substitutions at position 248 (A>C; N83T) and 2809 (C>T; P937S) and a single synonymous replacement at position 5271 (C>T; G1757G) in exons 2, 14 and 31 respectively. As mutations in other genes can cluster in “hotspots” (for example those found in *TP53* cluster to exons 5 to 8 (Petitjean *et al.* 2007)), the three corresponding exons were selected for mutational analysis in a cohort (n = 19) of the LLMPP DNA samples known to have LOH

(which may render mutations homozygous thereby aiding their detection) at the *CREBBP* locus on chromosome 16p13.1. These were used to rationalise use of limited DNA in the sequential FL / t-FL set, and were sourced from a previous study performed in our institution (O'Shea *et al.* 2009). However, mutations were not detected in any of the 60 amplicons.

4.4.7 Gain chromosome 7 and *EZH2* mutation

The finding of CN gains affecting chromosome 7 in FL ($n \geq 10$; **Table 4.2**) and t-FL ($n \geq 7$; **Table 4.3**) supports previous studies which have identified gain 7 occurring in at least 18% of FL cases (Bloomfield *et al.* 1983; Avet-Loiseau *et al.* 1997; Horsman *et al.* 2001; Mohamed *et al.* 2001). Indeed this was the most frequent chromosomal gain (20%) in a large series ($n = 336$) of FL in a report which also suggested this was an early event in disease evolution (Hoglund *et al.* 2004). Moreover, gain 7 may have a role in subsequent disease progression / transformation (Bernell *et al.* 1998; Horsman *et al.* 2001; Boonstra *et al.* 2003). However, until recently, it was not clear which genes affected by chromosome 7 gain might be relevant in FL.

In a report first presented at the American Society of Hematology annual meeting, New Orleans 2009, Morin and colleagues (Morin *et al.* 2009) identified that *EZH2* may be one such target. This is a histone methyltransferase gene, the protein product of which forms part of the polycomb repressive complex 2 (PRC2) which functions to trimethylate histone H3 at lysine 27 creating a repressive mark (H3K27me3) associated with transcriptionally silent chromatin (Cao *et al.* 2002; Kuzmichev *et al.* 2002; Kirmizis *et al.* 2004). The study identified that recurrent, non-synonymous single nucleotide substitution mutations affecting codon 641 in the catalytic domain of *EZH2* occurred in germinal centre lymphomas including 7.2% of FL cases (Morin *et al.* 2010a).

To investigate these mutations in transformation and whether they might be under selective pressure through amplification of the *EZH2* locus at chromosome 7q36.1, mutational analysis was performed. In total, 9 / 31 (29%) of the sequential cases demonstrated mutations in codon 641 of *EZH2*. These were all non-synonymous single nucleotide substitution mutations (**Figure 4.9**) which, in six cases, were

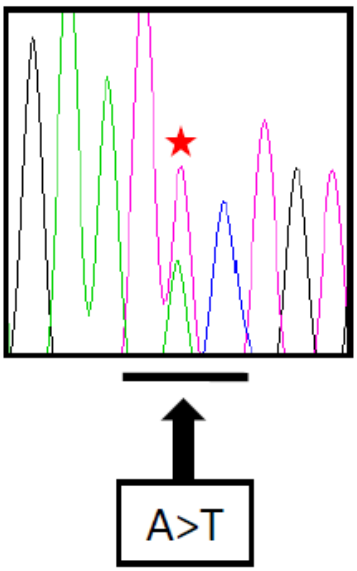


Figure 4.9 *EZH2* mutation

Illustrated is an example of one single nucleotide substitution mutation (red star) affecting codon 641 (black bar) of *EZH2*. The electropherogram shows double the level of mutated allele as compared to normal and is consistent with amplification of chromosome 7 in this case (Patient 1; **Table 4.5**) involving the mutated allele.

present in both FL and t-FL samples (for one case this included four FL samples); two cases showed mutations in t-FL that could not be detected in FL and one case showed a mutation only in FL and not the subsequent t-FL (**Table 4.5**). In six cases (67%), the mutations were accompanied by CN gains that included the *EZH2* locus; for four of these cases both mutation and CN gain were present in FL and t-FL samples, for one (case 17) the CN gain was present in both samples but the mutation was only detected in t-FL whereas in case 33 the mutation was present in all FL and t-FL samples but CN gain only affected the initial FL sample. None of the mutations were accompanied by CN loss of the *EZH2* locus. Five of the mutated cases had germline samples available which confirmed that the mutations were somatically acquired.

In t-FL and FL, 8 (26%) and 7 (23%) of the 31 cases, respectively, were mutated which suggests that the rate might be higher (in this population selected on the basis of transformation) than that determined for a large series of FL (n = 251) in the report described above (Morin *et al.* 2010a). To examine these mutations in an independent, non-selected FL cohort a series of diagnostic samples (n = 56) from our institution were screened. In 6 cases (10.7%), *EZH2* codon 641 single base substitution mutations were detected that similarly encoded amino acid replacements: asparagine (n = 3), phenylalanine (n = 2) and histidine (n = 1) with the mutation frequency no different to that identified by Morin and colleagues (Fisher's exact test, $P = 0.41$).

The doubling of mutation rate in FL from the 31 sequential cases (23%) is not significantly different to that in the 56 samples in diagnostic series (10.7%), $P = 0.12$, which might, in part, reflect sample size limitations. Furthermore, in the diagnostic set of 56 samples, transformation developed in 2 / 6 (33%) mutated and 18 / 50 (36%) un-mutated cases; and this as well as OS and TTT comparisons between these groups were not significantly different. However, as the FL samples in the sequential cases are predominantly from disease episodes occurring after the initial diagnosis, and mutations can develop with transformation (Cases 17 and 20) indicating they can be acquired later in the course of disease, it is possible that such disruption of *EZH2* may contribute to progression / transformation.

Table 4.5 *EZH2* mutations in FL / t-FL

Patient^	Sample	Mutation[$]	Codon 641 change*	Germline[#]	Variant amino acid	Chr 7 copy number change including *EZH2* locus	
						Loss	Gain
1	FL1	+	TA/TC	n.a.	Phe	-	+
	t-FL	+				-	+
5	FL1	+	TA/TC	n.a.	Phe	-	+
	t-FL	+				-	+
7	FL1	+	TA/CC	n.a.	Ser	-	-
	t-FL	-				-	-
12	FL1	+	T/CAC	n.a.	His	-	-
	t-FL	+				-	-
15	FL1	+	TA/TC	Normal	Phe	-	+
	t-FL	+				-	+
17	FL1	-	T/AAC	Normal	Asp	-	+
	t-FL	+				-	+
19	FL1	+	TA/TC	Normal	Phe	-	+
	t-FL	+				-	+
20	FL1	-	TA/TC	Normal	Phe	-	-
	t-FL	+				-	-
32	FL1	+	TA/TC	Normal	Phe	-	+
	FL2	+				-	-
	FL3	+				-	-
	FL4	+				-	-
	t-FL	+				-	-

^Patient number in *italics* = only FL or **bold** = only t-FL sample mutated. [$]For each patient the same mutation was found in all affected samples. *All mutations identified within *EZH2* codon 641 were non-synonymous single base substitutions. Mutations are indicated by an underline with the left-hand replaced by the right-hand base. [#]Germline DNA was analysed where available. *+indicates present; -, absent; n.a., not available; Phe, phenylalanine; Ser, serine; His, histidine; and Asp, asparagine.*

4.4.8 Candidate genes identified by loss of chromosome 1p36

Loss of the telomeric aspect of chromosome 1p occurs in a variety of cancers (Ragnarsson *et al.* 1999) and its loss has been reported in FL (Horsman *et al.* 2001; Hoglund *et al.* 2004; Zhang *et al.* 2004). Ross and colleagues have since identified that LOH of chromosome 1p36 through allelic loss or aUPD occurs in up to 50% of FL (Ross *et al.* 2007). Subsequently, aUPD 1p36 has been shown to predict OS in FL (O'Shea *et al.* 2009) and an array-CGH determined 11.4 Mb loss at 1p36 found to predict both OS and risk of transformation (Cheung *et al.* 2009b). This strongly suggests that 1p36 contains a genetic locus which, when disrupted, contributes to FL outcome.

4.4.9 *CDC2L2:* a candidate tumour suppressor gene on 1p36

In the current study, loss of 1p36 that affected genes was present in four MCRs found in FL samples and two found in t-FL. The MCRs ranged from 27 to 81kb and 293 kb to 4.8 Mb in size, in FL or t-FL respectively, and occurred in up to 11 / 31 cases (35%). To address this further the Wellcome Trust Sanger Institute Cancer Genome Project (CGP) website, http://www.sanger.ac.uk/cgi-bin/genetics/CGP was examined. This includes over 800 cell lines that have been assayed for LOH and CNAs; of these 127 are haematological with 15 sourced from B-NHL cases, four of these contain t(14;18) and one, WSU-NHL (derived from the pleural effusion of a progressive, refractory nodular histiocytic B-NHL which may have transformed), has an apparent homozygous loss of chromosome 1 at around 1 to 2 Mb (within 1p36).

A small, homozygous region of loss extending from 1.60 to 1.66 Mbps was identified within chromosome 1p36.33 in DNA from WSU-NHL (**Figure 4.10**) and this was confirmed by PCR (**Figure 4.11**). This 60 kb region was included within losses affecting both the FL (n = 10) and t-FL (n = 11) samples and contained two genes. One of these, *CDC2L2* (*CDK11A*), has a role in the negative regulation of normal cell cycle progression and may function as a tumour suppressor gene as it is deleted or altered frequently in neuroblastoma (Lahti *et al.* 1994) and its single allele loss facilitates skin carcinogenesis in mice (Chandramouli *et al.* 2007). Furthermore, its protein product down-regulates BCL2 (Yun *et al.* 2007) and enhances apoptosis

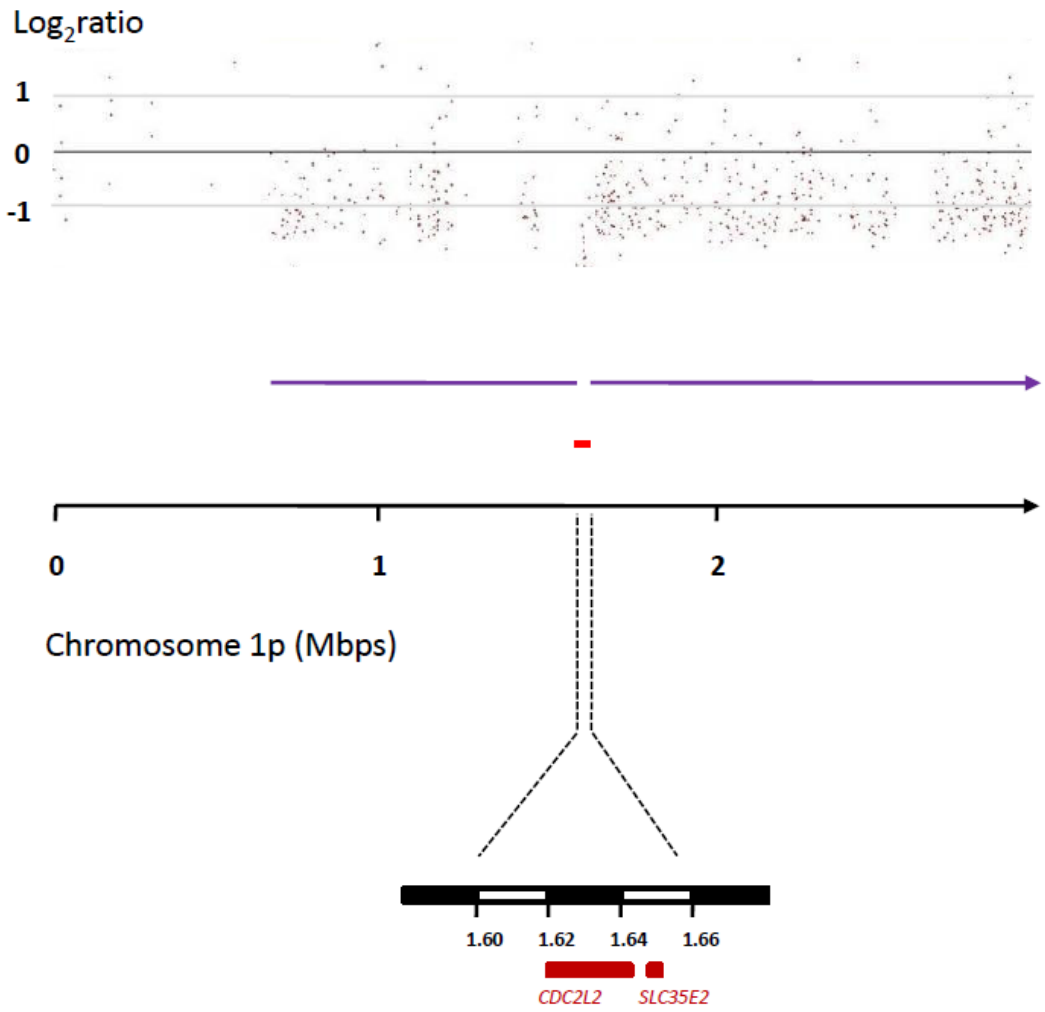


Figure 4.10 Chromosome 1p loss in WSU-NHL cell line

Illustration of homozygous loss within a larger area of hemizygous loss on chromosome 1p in WSU-NHL. Top, copy number plot with each dot representing one probe (from the Affymetrix GenomeWide SNP6.0 array) and its signal intensity for the WSU-NHL cell line relative to that of a germline pool of non-malignant human DNA plotted as $\log_2$ratio, as described in **Figure 4.8.** Middle, the majority of the plot shows copy number of 1 highlighted by the interrupted purple line; the small red line highlights a region of homozygous loss extending from 1.60 to 1.66 Mbps. Bottom, the two known genes located within this region are indicated: *CDC2L2* (*CDK11a*) has a role in cell cycle control and *SLC35E2* is a predicted solute transporter gene.
Mbps indicates megabase pairs

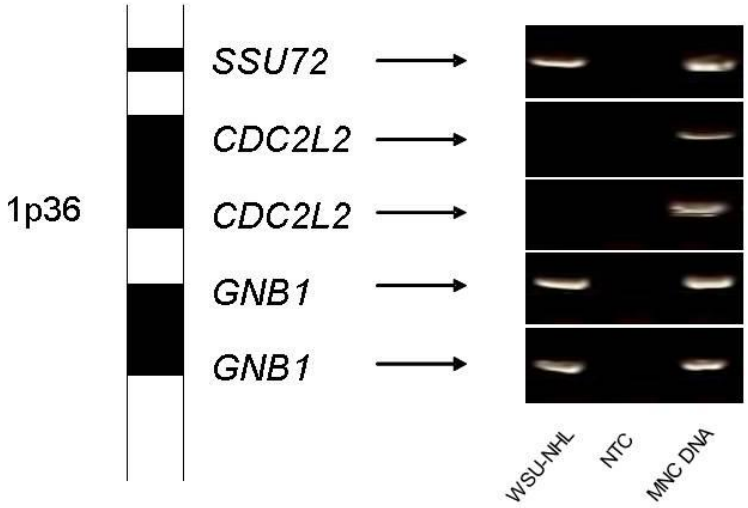


Figure 4.11 Confirming *CDC2L2* homozygous loss

Illustration of the approach taken to confirm the *CDC2L2* homozygous deletion in WSU-NHL and to demonstrate that the loss seen on the micro-array was not an artefact due to CNVs in the germline baseline pool of normal DNA (analogous to the effect of CN changes in FL if t-FL were compared directly to FL as illustrated in ***Appendix 1*** **Figure 1**). Left, schematic representation of genes (black boxes) within chromosome region 1p36 including *CDC2L2* and two neighbouring genes (not to scale); the telomeric gene (*SSU72*) is uppermost. Right, agarose gel electrophoresis images of PCR products from amplification of regions of the respective genes demonstrates lack of amplification of *CDC2L2* (illustrated here for the first and last exons) in WSU-NHL DNA confirming the microarray findings of a homozygous loss at this locus.
NTC indicates no template control; MNC, pooled mononuclear cells.

of HCC cells (Cai *et al.* 2002). Consequently, FL cases (n = 12; four of which subsequently transformed) from the LLMPP series (O'Shea *et al.* 2009) were screened for *CDC2L2* mutations. These represented a good screening cohort as, in each case, 1p36 was affected by aUPD which would render mutations homozygous.

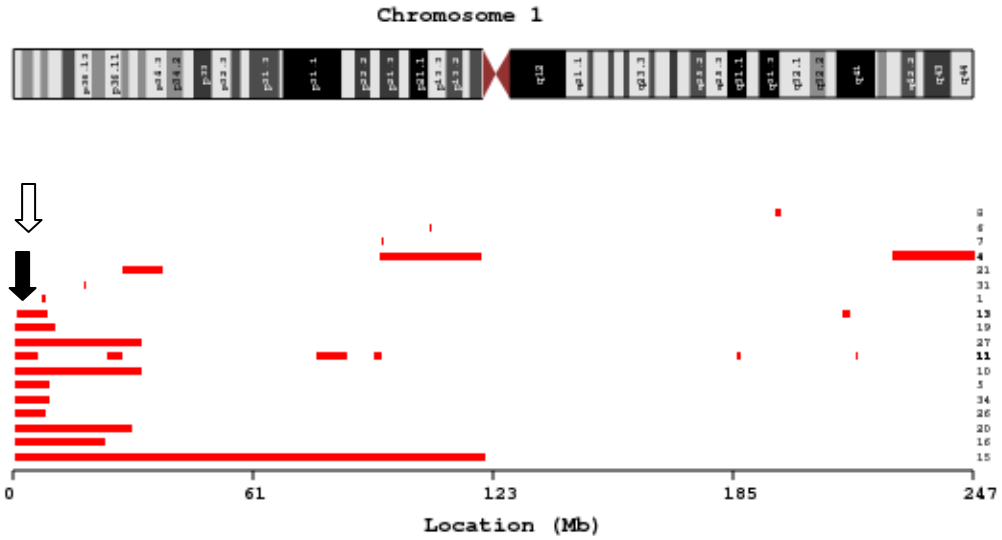
All 20 exons in *CDC2L2* were amenable to sequencing as 15 amplicons. As *CDC2L1* has a high degree of homology to *CDC2L2*, unique primers for respective amplicons could not always be designed. Consequently, forward and reverse sequencing was performed upfront with nucleotides common to both determining a consensus sequence. However, no *CDC2L2* mutations were detected across all exons in the FL cases examined.

4.4.10 *TNFRSF14* mutations occur frequently in FL and t-FL

An elegant study, first presented at the American Society of Hematology annual meeting in December 2009, and recently published (Cheung *et al.* 2010b), identified that *TNFRSF14* (located less than 1Mb centromeric to *CDC2L2* as shown in **Figure 4.12**) was mutated in 18% FL cases at diagnosis. The gene product of *TNFRSF14* is expressed on the surface of many cell types including both B and T lymphocytes and can bind different ligands which may activate effector function, although it can function in a co-stimulatory or co-inhibitory manner and this is dependent on the ligand bound (Cai and Freeman 2009).

As this gene is located within the MCR losses in FL (n = 11) and t-FL (n = 11) cases at sizes 190kb and 4.8Mb, respectively (**Table 4.2**, **Table 4.3** and **Figure 4.12**), the mutation status of *TNFRSF14* was determined for the FL / t-FL sample set to examine the relationship with transformation and loss 1p36. All coding regions of the eight exons were assessed and 12 mutations were detected in 10 / 31 (32.3%) cases across the first six exons (**Table 4.6**). In half of these cases, the same mutation was present in all FL and t-FL samples; two cases (7 and 26) showed mutations only in t-FL (case 7 had two distinct mutational events); case 33 had one mutation found only in the earliest FL and a second mutation found only in the subsequent FL2 and t-FL samples (indicating that ongoing acquisition of different *TNFRSF14* mutations by different disease sub-clones can occur over the course of FL / t-FL) whereas the

A



B

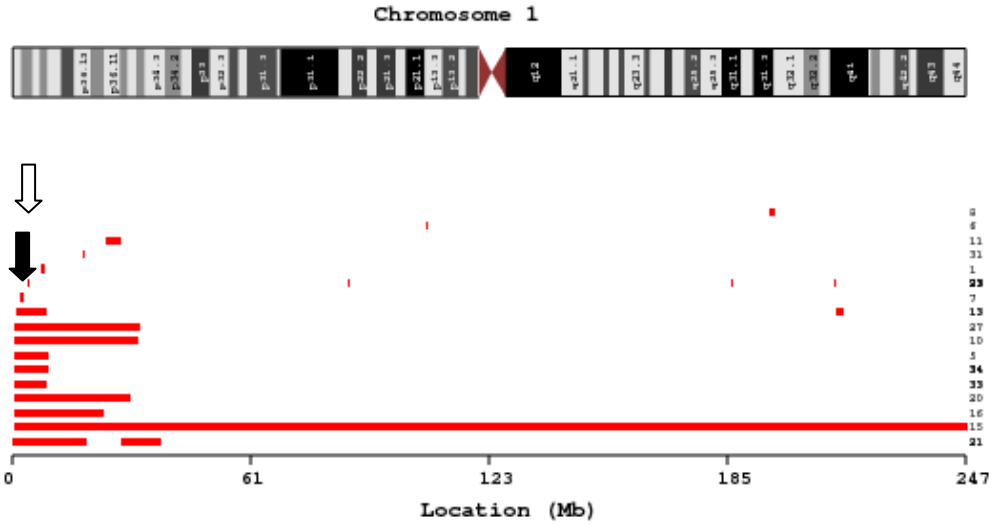


Figure 4.12 *CDC2L2* and *TNFRSF14* are candidate genes within loss 1p36

The figure layouts are similar to that detailed for **Figure 4.6** (top) with the addition that red bars indicate region of loss. The positions of *CDC2L2* (black filled arrow) and *TNFRSF14* (white filled arrow) within the common regions of chromosome 1p loss are indicated for t-FL (A) and FL (B) samples.

Table 4.6 *TNFRSF14* mutations in FL / t-FL

Exon region	Patient[ψ]	Sample	Mutation	Mutation type	Start position[+]	Mutation effect	Germline	Chr 1 CNA inc. *TNFRSF14*	
								gain	loss
1	**7**	FL1	-	del 24 bp[‡]	between 297 and 299[†]	TSS deletion	n.a.^	-	+
		t-FL1	+					-	-
	17	FL1	+	G>A[#]	302	TSS disruption (ATG > ATA)	normal	-	-
		t-FL1	+					-	-
	22	FL1	+	G>A	1 bp intron 1-2	Alters splice region	n.a.^^	-	-
		t-FL1	-					-	-
	32	FL1	+	T>A	2 bp intron 1-2	Alters splice region	normal	-	-
		FL2	+					-	-
		FL3	+					-	-
		FL4	+					-	-
		t-FL1	+					-	-
2	5*	FL1	+	G>A	67	Non-synonymous: Glu > Lys	n.a.	-	+
		t-FL1	+					-	+
3	**33**	FL1	-	G>A	109	Non-synonymous: Cys > Tyr	normal	-	+
		FL2	+					-	-
		t-FL1	+					-	-
4	19	FL1	+	C>A	119	Premature stop codon	normal	-	-
		t-FL1	+					-	+
	23	FL1	+	ins TGTGTGA	55-56	Frameshift	normal	-	-
		FL2	+					-	-
		t-FL1	-					-	-
	33	FL1	+	C>T	111	Non-synonymous: Arg > Cys	normal	-	+
		FL2	-					-	-
		t-FL1	-					-	-
5	26	FL1	-	ins C	between 34 and 41[$]	Frameshift	normal	-	-
		t-FL1	+					-	+
6	**7**	FL1	-	ins A	between 4 and 6[§]	Frameshift	n.a.^	-	+
		t-FL1	+					-	-
	9	FL1	+	del G	60 or 61[¥]	Frameshift	normal	-	-
		t-FL1	+					-	-
		t-FL2	+					-	+

Table 4.6 *TNFRSF14* mutations in FL / t-FL (continued).

[ψ]Patients highlighted in **bold** have two mutations (affecting different exons in each case.).
[+]bp position in respective exon unless indicated.
‡due to different possible start / end combinations, there are three possible 24 bp losses from GGCATGGAGCCTCCTGGAGACTGGGG including either the left-hand two, the inner two or the right-hand two G bases along with the other 22 bases.
[†]There are differing potential start (and end) sites of deletion reflecting runs of G bases at start and end sites. The three possible start sites to the loss are 297-8, 298-9, 299-300 bp.
^ mutation in t-FL only (not in FL)
[#]t-FL homozygous
^^ mutation in FL1 only (not t-FL1)
*The single bp substitution in exon 2 of case 5 could be a rare non-synonymous SNP variant. However, no SNP has been identified from dbSNP build 131. This is the only one of the twelve mutations which has potential to be a germline variation. (The other patients who lack germline have at least one tumour sample without mutation which shows germline sequence.)
$ can be any of position 34-5, 35-6, 36-7, 37-8, 38-9, 39-40, 40-1 bp due to run of six Cs in germline and ins C
§ position 5 in germline is A so ins A either between 4-5 or 5-6 bps
¥positions 60 and 61 bp both G in germline so del single G either 60 or 61 bp
Chr indicates chromosome; CNA, copy number alteration; -, absent; +, present; del, deletion; ins, insertion; bp, base pair; TSS, translation start site; n.a., not available; Glu, glutamic acid; Lys, lysine; Cys, cysteine; Tyr, tyrosine; Arg, arginine.

remaining two cases (22 and 23) showed mutations only in FL samples but not the subsequent t-FL, giving a frequency in transformed samples of 26% (n = 8).

The mutations included single nucleotide substitutions (n = 7), single nucleotide insertions (n = 2), a 7bp insertion, a single nucleotide deletion and a 24 bp deletion. Consistent with a role as a tumour suppressor gene, the mutations all had potential to disrupt translation / splicing / protein sequence through disruption of the translation start site / ATG initiation codon (n = 2), amino-acid substitutions (n = 3); frameshifts (n = 4) and direct creation of stop codon (n = 1). The two remaining mutations were located 1 and 2 bp in to intron 1 - 2 and had potential to disrupt exon:intron splicing.

The mutations in the eight cases with germline DNA available were confirmed as acquired (examples are given in **Figure 4.13**) and of those remaining, three were present in only one of the FL or t-FL samples. The final mutation (a single nucleotide substitution at position 67 of exon 2 in case 5) was present in both FL and t-FL and, without germline DNA for comparison, a SNP could not be excluded although this is unlikely as such a variant has so far not been reported (http://www.ncbi.nlm.nih.gov/projects/SNP/) and the encoded amino acid change would replace glutamic acid (acidic) with lysine (basic) and is predicted to be potentially damaging to protein function using the PolyPhen website, http://coot.embl.de/PolyPhen/ (Ramensky *et al.* 2002). When the respective FL or t-FL mutated versus un-mutated cases were compared, no associations were demonstrated regarding time to transformation or OS in FL nor survival from t-FL.

Mutational frequency was then determined at FL presentation by screening samples from the diagnostic FL set (and was possible for 54 / 56 cases). A total of 17 mutations each in separate cases (a frequency of 31%) and again affecting the first 6 exons were detected. The majority (n = 13) constituted single nucleotide substitutions (three of which disrupted the ATG initiation codon); three were deletions including two single nucleotide losses and one 28 bp loss; and one involved a single nucleotide insertion. Germline DNA was available and confirmed the mutations for 6 cases whereas for four single nucleotide substitutions (two synonymous and two occurring at the exon:intron junction) the possibility that they represent undescribed SNPs could not be excluded (in a similar manner to case 5 from the sequential set above). All events had the potential to disrupt initiation of

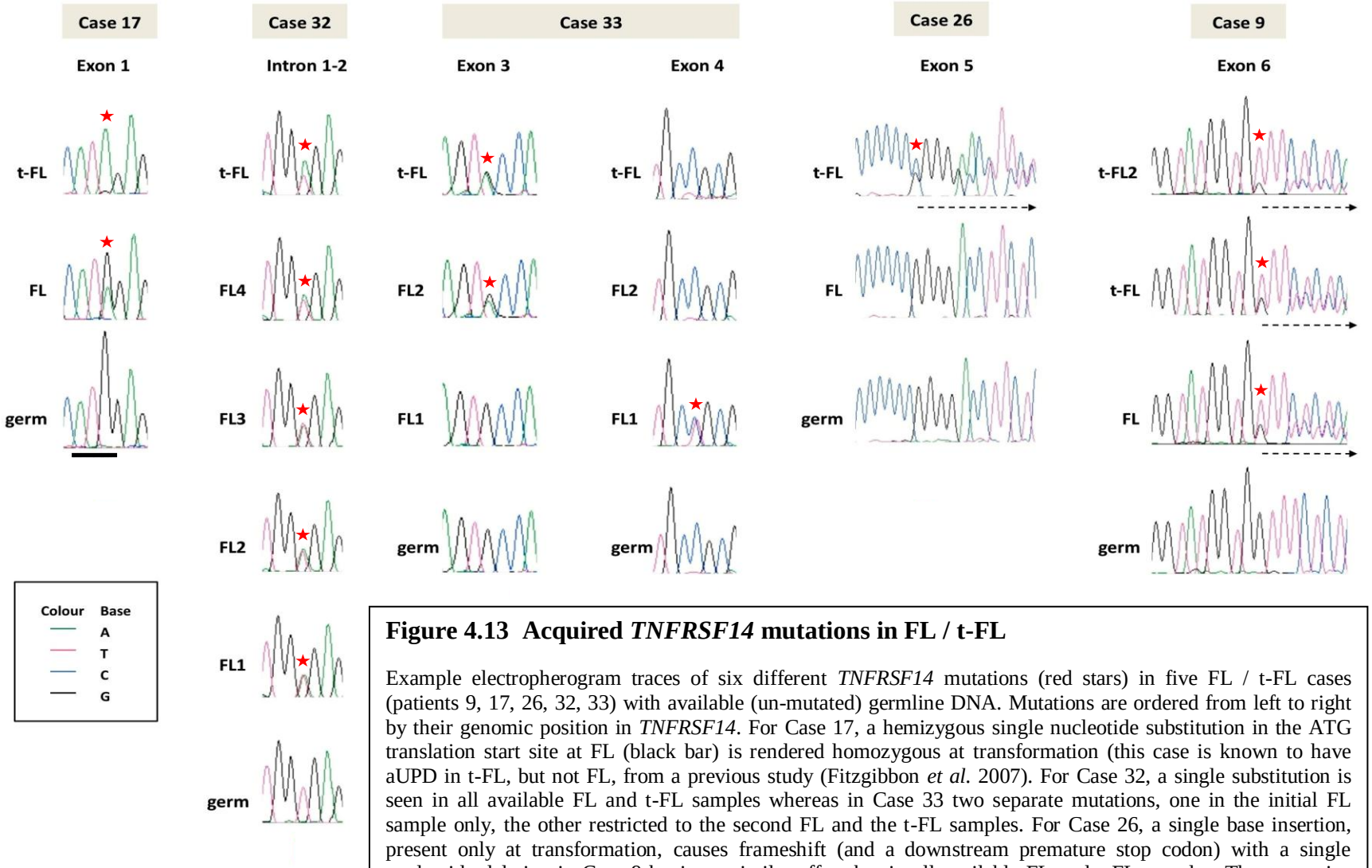


Figure 4.13 Acquired *TNFRSF14* mutations in FL / t-FL

Example electropherogram traces of six different *TNFRSF14* mutations (red stars) in five FL / t-FL cases (patients 9, 17, 26, 32, 33) with available (un-mutated) germline DNA. Mutations are ordered from left to right by their genomic position in *TNFRSF14*. For Case 17, a hemizygous single nucleotide substitution in the ATG translation start site at FL (black bar) is rendered homozygous at transformation (this case is known to have aUPD in t-FL, but not FL, from a previous study (Fitzgibbon *et al.* 2007). For Case 32, a single substitution is seen in all available FL and t-FL samples whereas in Case 33 two separate mutations, one in the initial FL sample only, the other restricted to the second FL and the t-FL samples. For Case 26, a single base insertion, present only at transformation, causes frameshift (and a downstream premature stop codon) with a single nucleotide deletion in Case 9 having a similar effect but in all available FL and t-FL samples. The respective mutation sites are indicated below the case numbers and sample origin is indicated to the left of each trace (with numbers corresponding to disease episodes). *germ indicates germline.*

translation, protein sequence or intron splicing as, once again, confirmed mutations were detected at the start (within the initial 5 bp) of introns. The risk of transformation was no different for mutated (41%) versus un-mutated (32%) cases, *P* = 0.55, and, similarly, no significant associations were demonstrated for time to transformation and overall survival, contrasting with the recently published report by Cheung and colleagues (Cheung *et al.* 2010b). However, on exclusion of the four possible SNPs and the three early intronic events, the revised mutation rate (including only non-synonymous exon mutations) of 10 / 54 cases (19%) is similar to that study.

4.4.11 t-FL associated mutations may occur in FL sub-clones

Of the four *TNFRSF14* mutations demonstrated only at transformation in three FL / t-FL cases (7, 26 and 33; **Table 4.6**), mutation specific forward and reverse ASO primers were generated for two (24 bp del in exon 1 and ins A in exon 6, both from Case 7). These enabled FL samples to be examined for presence of (low-level) t-FL clones and were applied in separate PCR reactions to the respective FL and t-FL samples. This confirmed the mutations were present in the t-FL samples. Significantly, in the preceding FL samples for which conventional sequencing had not detected the mutations, the targeted PCR confirmed they were indeed present but at lower levels than in the subsequent t-FL, as illustrated (for ins A) in **Figure 4.14**, which indicates that (in this case) the t-FL clone is present in FL almost two years before presenting clinically as transformation.

4.4.12 Non-sequential mutations suggest divergent evolution in FL / t-FL

Three cases (22, 23 and 33) contained mutations present in FL but absent from subsequent samples (**Table 4.6**); similar to the scenario observed for the *EZH2* mutation in Case 7 (section **4.4.7**). In case 22, a single nucleotide substitution (G>A) occurred at the first position of intron 1-2 in FL only and in case 23 a 7bp insertion within exon 4 was identified in the two available FL samples but could not be detected at subsequent transformation. For case 33 (which had normal germline DNA available) a non-synonymous C>T substitution (coding Arg>Cys) was present at position 111 of exon 4 in the initial FL sample but neither the subsequent FL2 or

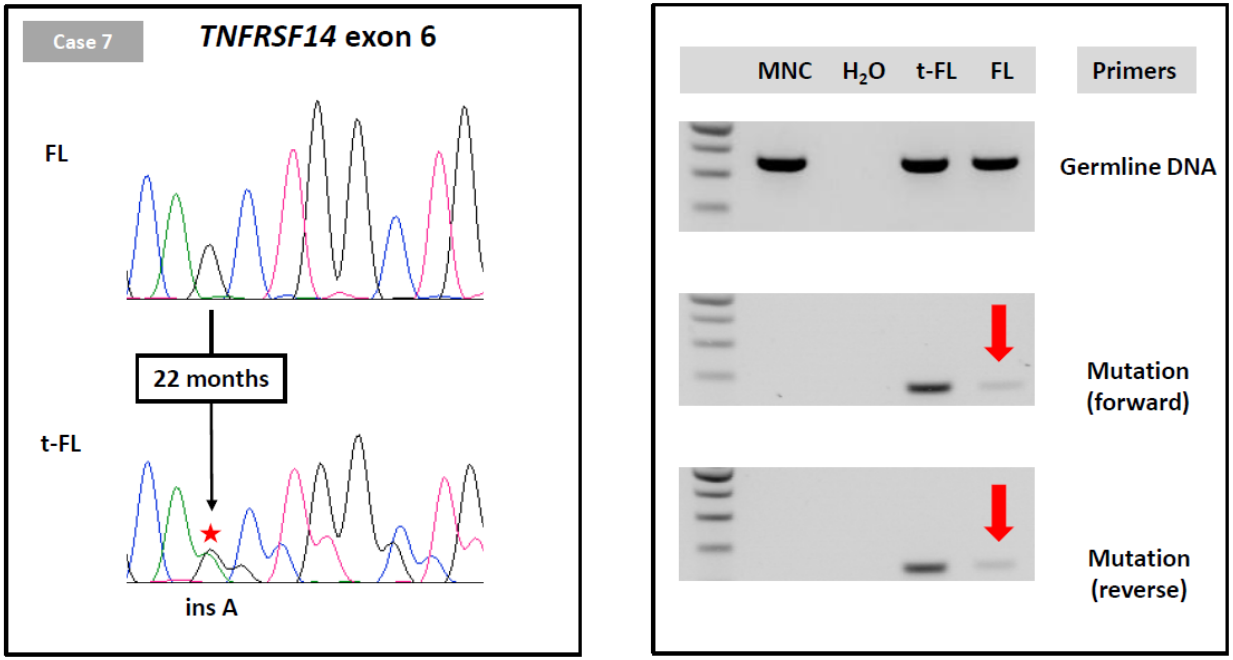


Figure 4.14 Detection of t-FL associated mutations in preceding FL

Left, electropherogram traces demonstrating germline sequence (with no detectable mutation) in FL and the presence of a mutation (insertion A) in exon 6 of *TNFRSF14* causing frameshift in t-FL occurring almost two years later, as illustrated by the red star. Right, agarose gel electrophoresis of mutation specific PCR with primers designed on germline exon 6 sequences used (top) as a control and showing amplification in both tumour samples and the normal DNA; forward (middle) and reverse (bottom) primers designed on the mutation detected only in t-FL and applied to the same samples confirm its absence in normal (MNC) DNA and its presence in t-FL, as expected, but also reveal a low level amplification in the preceding FL (red arrow). This suggests the t-FL clone is present, in FL, almost two years before presenting clinically as disease.

t-FL samples. In the same case, a second mutation (the non-synonymous G>A substitution coding Cys>Tyr) occurred in FL2 and t-FL but was not detected in the initial FL.

Consequently, these cases demonstrate that future FL / t-FL episodes may not arise directly from the preceding FL but, rather, a more complex evolutionary pattern occurs. Using case 33 as an example, a divergent model of disease development can be proposed in which a common origin acquires differing *TNFRSF14* mutations present in either the initial FL or the subsequent relapse (FL2) and transformation (**Figure 4.15**). In particular, a common origin, as opposed to two independent FL episodes, is indicated by the presence of chromosome 10q loss (identified by the earlier micro-array study) in all three samples (FL, FL2 and t-FL). As over a decade elapses between FL and FL2, this indicates the persistence of a long-lived clone marking a common origin to both episodes. For case 22, the CNAs identified by the micro-array study also support a common origin as, in addition to those present in FL but absent at t-FL, and those occurring only in t-FL, other gains and losses are common to both FL and t-FL.

4.4.13 Divergent evolution of FL / t-FL from a common progenitor cell

In light of the divergent evolution suggested for four cases by the *EZH2* and *TNFRSF14* mutational analysis, the distributions of micro-array determined CN changes for the respective samples from each of the 31 sequential FL / t-FL cases were compared. This demonstrated that in 22 / 31 cases (71%) development arose by divergent evolution from a common origin, based on certain CNAs present in FL being absent from subsequent t-FL. The remaining cases showed either sequential acquired CN changes in subsequent disease episodes (n = 7), indicating direct evolution of later disease from the initial FL, or had insufficient changes (n = 2) to define the evolutionary pattern (***Appendix 5* Table 2**).

All samples in each case were shown to be clonally related as follows:

> all except three cases (22, 23 and 24) had some CNAs that were present in each respective sample;

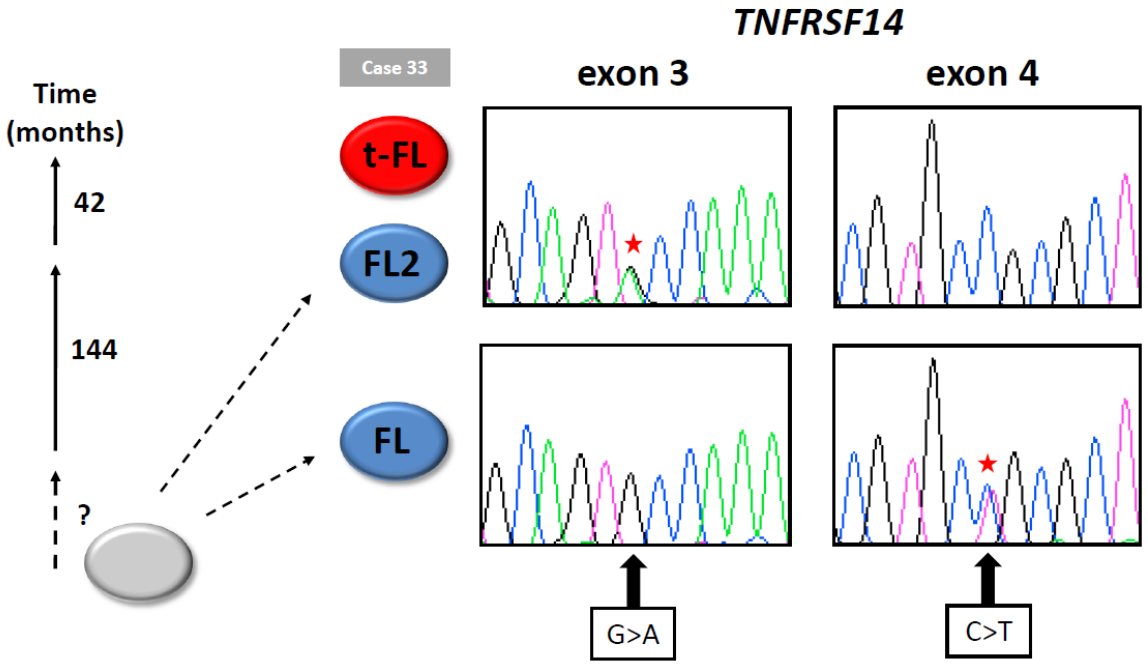


Figure 4.15 Non-sequential mutations suggest divergent evolution of FL / t-FL

Illustrated for the three samples (FL, FL2 and t-FL) available for Case 33 are electropherogram traces of two *TNFRSF14* mutations (indicated by the red stars). In exon 4, a mutation (C>T) present in FL, but absent from future FL2 and t-FL, indicates that these disease episodes did not develop directly from the FL clone. Furthermore, a second mutation (G>A) appears at FL2 and persists at transformation. As the earlier micro-array study indicates that loss chromosome 10q occurs in all three samples, this leads to a divergent model of development, illustrated on the left, in which a common origin (grey oval) that contains loss 10q acquires differing *TNFRSF14* mutations present in either the initial FL or the subsequent relapse (FL2) and transformation.
A time-line for disease episodes is given to the left, with an unknown (?) time elapsing between development of the common precursor, its acquisition of the exon 4 mutation and clinical presentation as FL. The long time course (12 years) between FL and FL2 indicates the persistence of a long-lived clone marking a common disease origin (that lacks the exon 4 mutation and acquires the exon 3 mutation at some point prior to FL2 presentation). The exon 4 mutation demonstrates the divergent evolution of disease clones and the additional exon three mutation reflects ongoing acquisition of different *TNFRSF14* mutations by different disease clones over the course of FL / t-FL.

cases 22 and 24 were previously shown to have the same *IGH-BCL2* rearrangements in their respective FL and t-FL samples (Jenner *et al.* 2002); and

case 23 shows the same *IGH-V* segment selection (demonstrated later in chapter 5, **Table 5.2**) in the FL and t-FL samples

which indicates that no disease episodes occurring after the initial FL represent entirely new, spontaneously arising disease clones.

Consequently, these data suggest that a putative common progenitor cell (CPC) that has acquired some genetic changes may give rise to FL or t-FL through subsequent non-sequential, divergent genetic evolution. As an example, in case 33, loss 10q is common to all cases, with loss 1p36 and the *TNFRSF14* exon 4 mutation occurring only in FL whereas the *TNFSRF14* exon 3 mutation occurs only in FL2 and t-FL; suggesting that a CPC has acquired loss 10q and goes on to acquire the other genetic changes as the respective disease episodes develop.

Recurrent copy number changes in cases with divergent evolution might represent early events when present in both FL and t-FL. The CN changes in these cases are indicated in ***Appendix 5*** **Figure 2** and recurrent MCRs (n = 29), present in all samples from each case and occurring in at least 3 / 22 (>10%) cases, are summarised in **Table 4.7**. Three of these involve only one gene and, in particular, gain 7q36.1 includes only *SMARCD3* the product of which contributes to regulation of gene transcription through chromatin remodelling and is part of a chromatin remodelling complex that regulates the self-renewal / proliferation of neuronal stem cells (Lessard *et al.* 2007). As these can give rise to diverse progeny, which at least in part mirrors the putative CPC, *SMARCD3* represents a further candidate gene in FL / t-FL development.

Finally, the CPC model of evolution not only applies to FL followed by transformation but can also apply to sequential t-FL samples (as illustrated for Case 9 in **Figure 4.16** and **Figure 4.17**) and FL that relapses after t-FL (Case 21; **Figure 4.18** and **Figure 4.19**), with these two scenarios reflecting variation in the hierarchy of progenitor development. In case 9, a less primitive sub-clone gives rise to t-FL, but undergoes clonal divergence in giving rise to relapse of this disease (t-FL2). In

Table 4.7 CPC common MCRs in all samples of 3 or more cases

Chr	Start	End	Length	Type	Cytoband	No. Cases	No. Samples	No. Genes	Candidate genes	Cases
1	18449498	18476791	27293	Loss	1p36.13	4	9	1	*IGSF21*	10, 15, 27, 31
1	24306170	27853506	3547336	Loss	1p36.11	4	10	60	*RUNX3, CD52*	10, 11, 15, 27
1	28523347	32332437	3809090	Loss	1p35.3-35.1	4	10	33	*PUM1*	10, 15, 21, 27
1	155604487	179419537	23815050	Gain	1q23.1-25.3	3	7	>100	*ABL2*	10, 22 ,27
2	60646978	66775162	6128184	Gain	2p16.1-14	3	7	26	*REL, MEIS1*	10, 15, 18
2	135357790	141512804	6155014	Gain	2q21.3-22.1	3	8	16	*CXCR4*	8, 15, 18
2	209792047	209825550	33503	Gain	2q34	3	9	0	*MAP2* <200kb downstream	8, 18, 25
2	238257900	239413568	1155668	Gain	2q37.3	3	8	14	*HES6, TRAF3IP1*	8, 10, 18
7	138342	101876473	101738131	Gain	7p22.3-q22.1	3	7	>100	*CARD11, FAM, HDAC9, CREB5, IGFBP1/3, EGFR*	10, 15, 27
7	150603168	150622306	19138	Gain	7q36.1	3	8	1	*SMARCD3*	15, 27, 31
8	112962509	113314553	352044	Gain	8q23.3	3	10	1	*CSMD3*	10, 22, 31
8	132846688	146264219	13417531	Gain	8q24.22-q24.3	3	10	99	*MAPK15, NFKBIL2, PTK2*	4, 10, 31
9	21813210	22146338	333128	Loss	9p21.3	4	8	3	*CDKN2A, CDKN2B, MTAP*	18, 26, 28, 31
12	37536490	37970109	433619	Gain	12q12	5	15	2	*CPNE8, KIF21A*	10, 12, 15, 18, 32
12	44391098	45110683	719585	Gain	12q12-q13.11	6	16	4	*ARID2, SFRS2IP, SLC38A1, SLC38A2*	10, 12, 15, 18, 22, 32
12	98797775	98984451	186676	Gain	12q23.1	3	10	2	*ANKS1B, UHRF1BP1L*	4, 19, 32
15	18973525	19685239	711714	Loss	15q11.2	3	7	1	*POTEC*	10, 15, 31
15	41134331	42863544	1729213	Loss	15q15.2-q21.1	3	7	31	*TP53BP1, CCNDBP1, B2M, TRIM69*	6, 10, 15
15	59476426	59482274	5848	Loss	15q22.2	3	8	0	*RORA* <200kb upstream	7, 10, 15
17	52368	7883476	7831108	Loss	17p13.3-p13.1	3	7	>100	*TP53, CD68, JMJD3, TNFSF12*	4, 8, 21
17	23383257	25111724	1728467	Gain	17q11.2	3	6	40	*TNFAIP1*	8, 18, 31
17	38777643	39890078	1112435	Gain	17q21.31	3	6	27	*ETV4, HDAC5, GRN*	8, 18, 31
17	45304401	61464987	16160586	Gain	17q21.33-q24.1	3	6	>100	*CD79B, SMARCD2,*	8, 18, 31
17	69521471	72216370	2694899	Gain	17q25.1-q25.2	3	6	76	*CDK3*	8, 18, 31
18	148226	1567856	1419630	Gain	18p11.32	7	16	9	*YES1*	4, 9, 12, 25, 26, 28, 31
18	16864653	28890957	12026304	Gain	18q11.1-q12.1	7	17	42	*GATA6*	6, 9, 12, 25, 26, 28, 31
18	49444874	58857002	9412128	Gain	18q21.2-q21.33	6	16	31	*MALT1, TNFRSF11A*	6, 8, 9, 12, 26, 28
X	6838911	58319642	51480731	Gain	Xp22.31-p11.1	5	14	>100	*UTX, JARID1C*	6, 15, 18, 27, 32
X	137893124	154578240	16685116	Gain	Xq27.1-28	5	14	>100	*IKBKG, IRAK1*	12, 15, 18, 27, 32

The table layout is as indicated for **Table 4.2**.

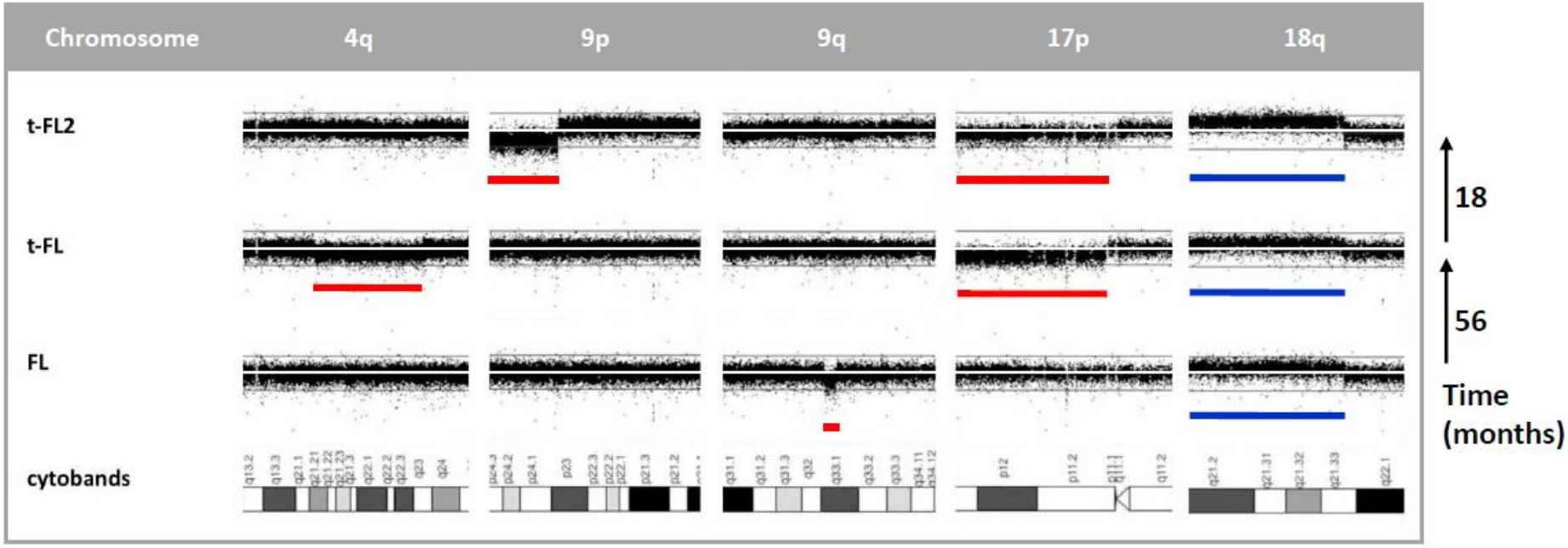


Figure 4.16 Divergent evolution giving rise to relapsed t-FL

Shown are fifteen CN plots (as described for **Figure 4.8**) from Case 9. There are five plots each for FL, t-FL and subsequent relapsed t-FL samples (labelled left), for five separate chromosome arms (indicated top) and corresponding cytobands (indicated bottom). Horizontal bars highlight the CN losses (red) and gains (blue). All cases have the same gain 18q and this in combination with loss 9q in FL only suggests divergent evolution of FLl and subsequent t-FL from a CPC containing gain 18q. A t-FL specific loss on 17p occurs but loss 4q in only the initial t-FL suggests divergence of t-FL from a precursor containing both gain 18q and loss 17p and this then acquires loss 9p at relapse of transformation (t-FL2). Time between samples is indicated to the right. The gain on chromosome 18 includes 18p, only gain 18q is illustrated here.

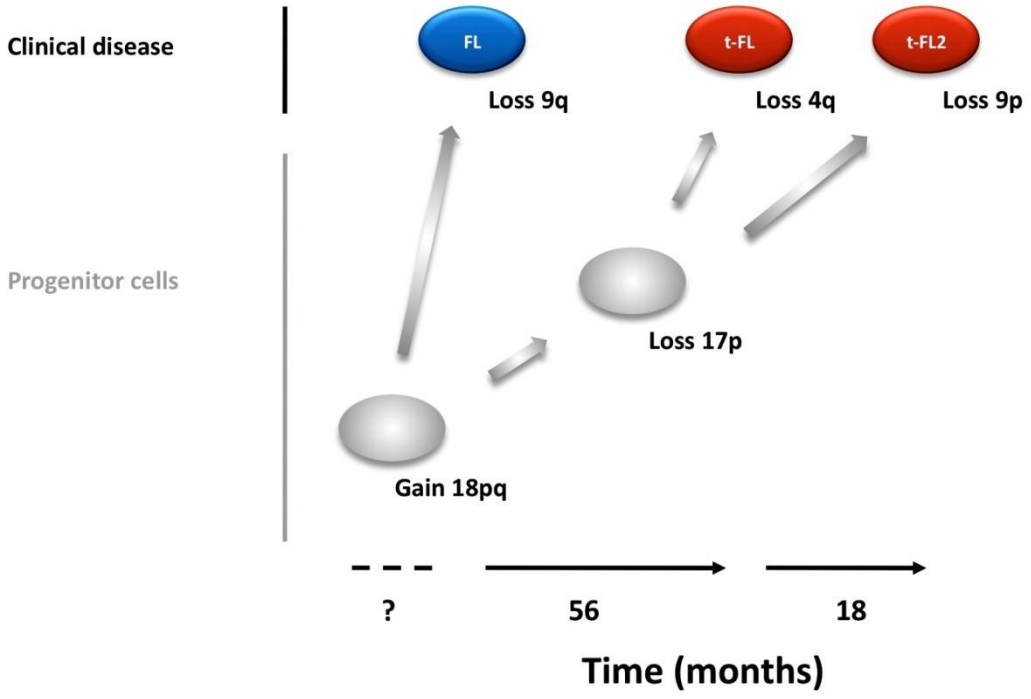


Figure 4.17 Model of divergent evolution of relapsed t-FL

A model of disease evolution for Case 9 developed from the CNAs illustrated in **Figure 4.16**. FL then t-FL / relapsed t-FL (t-FL2) are illustrated as blue and red ovals respectively. Putative progenitor stages are illustrated as grey ovals. The CNAs occurring at each stage are shown below the ovals. The more primitive progenitor stage gives rise to FL and a less primitive progenitor stage by divergent evolution, subsequently the less primitive progenitor stage gives rise to t-FL and its clinical relapse by divergent evolution. The time course between clinical episodes (bottom) indicates progenitor cells can persist for many months before disease presents clinically.

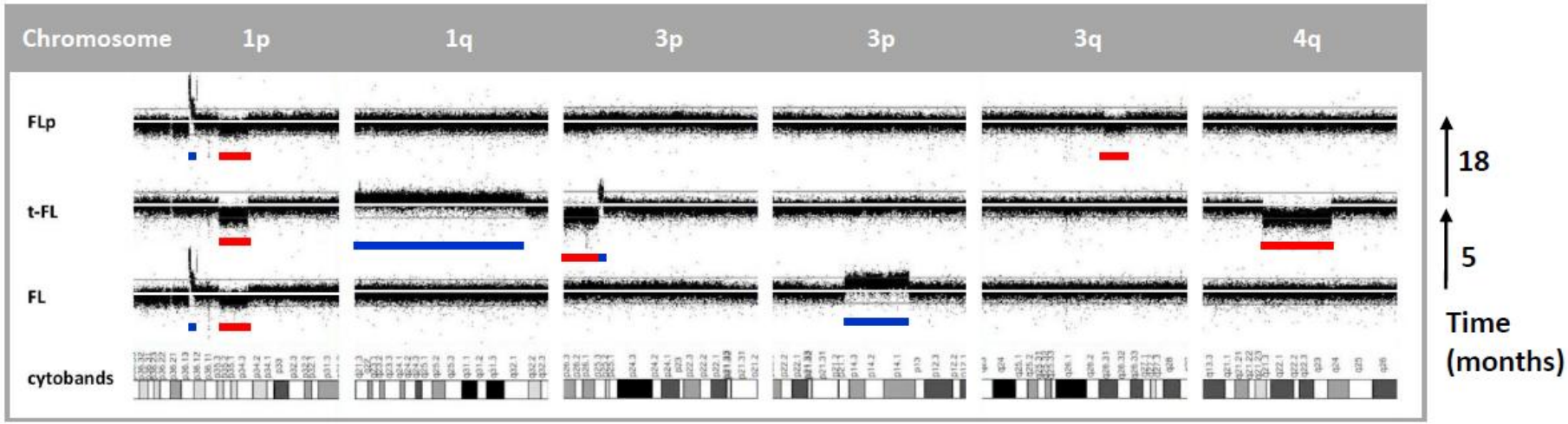


Figure 4.18 Divergent evolution giving rise to relapse of FL after transformation

The figure layout is as described for **Figure 4.16** with the addition that there are six CN plots for each sample, and these correspond to FL, t-FL and FL relapse post-transformation (FLp) from Case 21. All cases have the same loss 1p and this in combination with gain 3p in FL only suggests divergent evolution of FL and subsequent t-FL from a CPC containing loss 1p. CNAs specific to t-FL (gain 1q, loss and gain 3p and loss 4q) indicate FLp does not arise directly from t-FL. The gain 1p common to both FL and FLp suggests they share an origin less primitive than that which gives rise to transformation. The mutually exclusive loss 3q (FLp) and gain 3p (FL) indicate non-sequential evolution of FL / FLp.

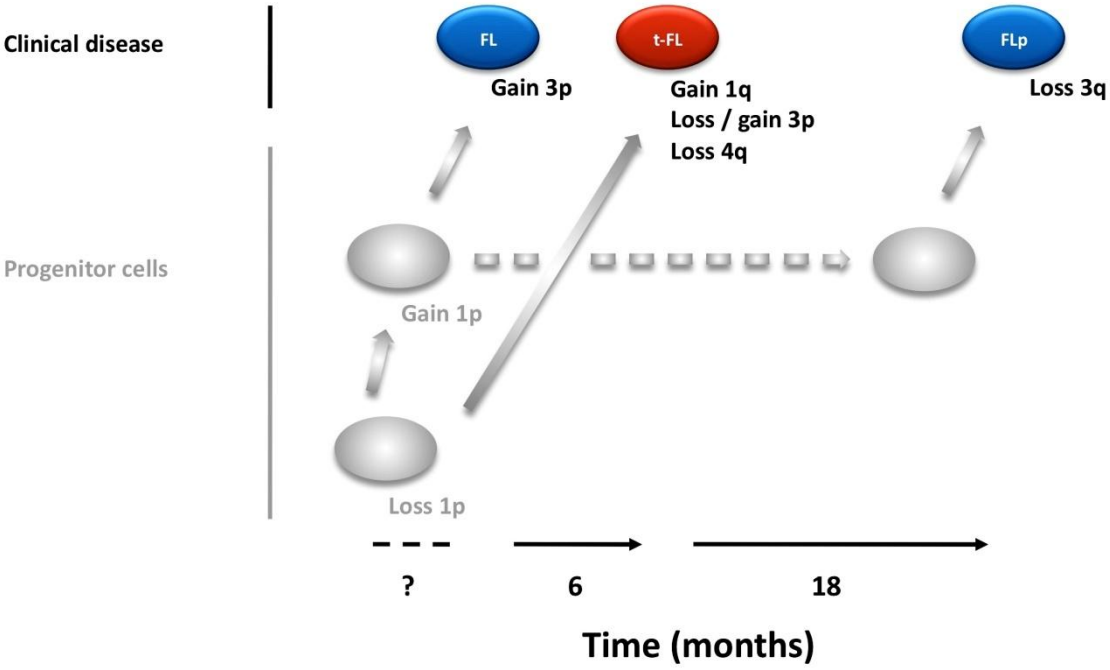


Figure 4.19 Model of divergent evolution of FL relapse after transformation

The figure layout is as described for **Figure 4.17**, except that samples are FL, t-FL and FLp (Case 21) and CNAs are those from **Figure 4.18**. The time course between clinical disease events indicates that both the more primitive and less primitive progenitors can persist for months.

case 21, a less primitive sub-clone gives rise to initial FL (FL1) and subsequent FL relapse after transformation (FLp), with t-FL arising from a more primitive sub-clone. This case raises the possibility that different (long-lived) progenitor sub-clones can co-exist, with temporal evolution of these progenitors not necessarily reflecting the temporal occurrence of clinical disease.

4.5 Discussion

Following the demonstration that germline variation is important in FL outcome (Chapter 3), the investigation described in this chapter has examined the role of acquired genetic changes in FL and its transformation. Using a series of sequential FL and t-FL samples from cases with comprehensive clinical follow-up, high resolution genome wide copy number profiling was performed to determine tumour associated CNAs. Those that either occurred recurrently or appeared enriched at transformation or targeted single genes were then used to determine candidate regions that might be relevant to disease development or progression in FL / t-FL. The inclusion of germline DNA was an important aspect of the study in both identifying disease associated CNAs and revealing somatically acquired mutations.

4.5.1 DNA copy number changes identify candidate regions in FL / t-FL

CNA losses and gains occurred in every sample, consistent with previous a large karyotyping study in which 97% FL cases harboured chromosome changes in addition to t(14;18) (Horsman *et al.* 2001). The recurrent CNAs in FL (**Table 4.2**) or t-FL (**Table 4.3**) occur at similar chromosome regions to those described in other reports in FL (using karyotyping (Johansson *et al.* 1995; Horsman *et al.* 2001; Hoglund *et al.* 2004); CGH (Bentz *et al.* 1996; Avet-Loiseau *et al.* 1997); or array-CGH (Cheung *et al.* 2009b) and FL / t-FL (using array-CGH (Martinez-Climent *et al.* 2003)) that have examined chromosome imbalances in FL / t-FL, with the high resolution of the Affymetrix SNP 6.0 genome-wide array used in the current study facilitating the refinement of certain recurrent CNAs to smaller (sub-megabase) MCRs, for example the losses at chromosome 1p36, as well as identifying those targeting single genes.

Previous reports have identified gain 2p (including the proto-oncogene *REL*) in FL (Ross *et al.* 2007) and FL / t-FL (Goff *et al.* 2000) and the survival association shown here suggests this gain may have a clinical impact, at least in FL that is destined to transform. However, although none of the other CNA MCRs predicted outcome, this association requires examination in a larger cohort to determine whether this effect is independent from other, clinical, predictors.

The focal loss of the majority of *CREBBP* (on chromosome 16p13.3) in Case 26 in combination with the association of 16p aUPD in FL and PFS (O'Shea *et al.* 2009) and the disruption of *CREBBP* in other malignancies (Borrow *et al.* 1996; Giles *et al.* 1997; Rowley *et al.* 1997; Mahmoodi *et al.* 2004) led to investigation of this locus here. Although no mutations were identified in a rationalised study of this large gene, this locus does appear to be relevant in FL as, subsequently, Pasqualucci and colleagues have recently demonstrated using whole exome sequencing that *CREBBP* is disrupted by genomic deletions and / or somatic point mutations in 32% FL (and 29% *de novo* DLBCL) (Pasqualucci *et al.* 2011). Furthermore, similar to findings from a recent report in FL (n = 25) in which both BAC array-CGH and the Affymetrix 500K SNP array were applied to samples to provide increased resolution to the genomic analysis (Cheung *et al.* 2010a), other single genes were also disrupted by CN changes in the current investigation. In addition, chromosomal breakpoints occurred through potentially important genes such as *KDM2B* (a histone demethylase) on chromosome 12 and *PRIM2* (DNA primase polypeptide 2 which has a role in DNA replication) on chromosome 6p. The location of such breaks may either directly impact on gene function or may represent translocations with potential to create hybrid fusion gene products. Consequently, these and the other single gene events, particularly the loss disrupting *ETS1*, represent candidates for future study. In this regard, recent presentations made at the American Society of Hematology annual meeting, Orlando, December 2010 identify that, in up to 89% of FL cases, recurrent inactivating mutations affect *MLL2*, which codes a histone methyltransferase involved in the epigenetic control of transcription (Mendez-Lago *et al.* 2010; Morin *et al.* 2010b). Importantly, this follows recent methylation studies indicating that epigenetic influences have a role in FL (Martin-Subero *et al.* 2009a; O'Riain *et al.* 2009). Moreover, *MLL2* is located on chromosome 12q so the mutations offer a

selective basis for the recurrent 12q aUPD observed in other FL series (Fitzgibbon *et al.* 2007; O'Shea *et al.* 2009).

Of CNAs potentially enriched at transformation, the MCR of gain 16p, including *IL4-R*, is of particular interest in view of the known relevance of micro-environment immune response in FL outcome (Dave *et al.* 2004). This follows the demonstration that expression of IL4 protein (a cytokine with roles in T cell differentiation and B cell activation by T_H cells) is increased in the micro-environment of FL tumours (Calvo *et al.* 2008) and that the IL4-R protein is over-expressed in FL cells (Husson *et al.* 2002), raising the possibility that gain *IL4-R* might predispose to its increased protein expression which, through interaction with IL4 present in the FL micro-environment, somehow favours transformation.

4.5.2 *EZH2* and *TNFRSF14* mutations in FL / t-FL

During the course of this study, the discovery of somatic mutations affecting *EZH2* (chromosome 7q.36.1) in GC lymphomas (including FL) was reported (Morin *et al.* 2010a). As gain 7 occurred in almost one third of the sequential FL / t-FL cohort, *EZH2* was examined in this cohort and mutations were identified in a subset of these cases with some restricted to transformation. Although *EZH2* mutations at FL diagnosis do not appear to confer increased risk of developing transformation, their development later in the course of disease in some cases suggests they may have a driver role in progression / transformation. The subsequent demonstration that heterozygous codon 641 point mutations are gain of function and act in a dominant manner (Sneeringer *et al.* 2010; Yap *et al.* 2011) suggests that the 7q CN gain, including the *EZH2* locus, observed in ⅔ of the mutated FL / t-FL cases is selecting for this oncogenic event (**Table 4.5**).

Disruption of chromosome 1p36 by aUPD (O'Shea *et al.* 2009) or CN loss (Cheung *et al.* 2009b) has clinical significance in FL. CN loss occurred frequently in the sequential FL / t-FL series with sub-megabase MCRs identified. Two genes were investigated: *CDC2L2* showed no mutations despite its involvement in recurring CN losses (subsequently a single non-synonymous mutation G>C at cDNA position 145 in one case of lung cancer has been reported (Ding *et al.* 2008)) whereas *TNFRSF14* was mutated frequently. The investigation of *TNFRSF14* mutations followed a

presentation at the American Society of Hematology Annual Conference, New Orleans, 2009 (Cheung *et al.* 2009a), subsequently reported in November 2010 (Cheung *et al.* 2010b), which identified exonic, non-synonymous mutations in *TNFRSF14*. This was a large study including over 200 FL cases and identified that *TNFRSF14* mutations at diagnosis predict OS in the era of rituximab therapy, although this association could not be reproduced here this is probably an effect of the limited population size of cases with available diagnostic FL samples (n = 56).

It is of significance that additional mutations, occurring within the first five bases of introns within *TNFRSF14* and confirmed by comparison to germline DNA, are demonstrated in the current study suggesting that the actual mutation rate of this gene in FL may be higher than that initially reported (Cheung *et al.* 2010b). All five intronic splice site mutations identified in this study occurred in 5' donor splice sites and, using the website http://www.cbs.dtu.dk/services/NetGene2/, each mutation was predicted to abolish the donor splice site function and disrupt exon:intron splicing. Analogous point mutations in 5' splice sites of other genes typically induce exon skipping and their exclusion from gene transcripts but can lead to inclusion of intronic regions and early termination of the transcript both of which can lead to truncated protein products and reduced expression (Antonarakis *et al.* 2002) and can occur in different cancers (Brose *et al.* 2004; Bianchi *et al.* 2011). Future work should determine the consequences of *TNFRSF14* splice site mutations on transcription and translation and whether these associate with gain or loss of function. This should help further unravel the selective basis of such mutations in FL given the role of ligand binding by TNFRSF14 (also termed HVEM) on the cell surface of normal B cells and their subsequent activation but the down-regulation of TNFRSF14 in GC B cells (Duhen *et al.* 2004), the origin of FL.

The *TNFRSF14* mutations are also detected only at t-FL and this suggests they may have relevance as drivers in the process of transformation, whereas the lack of association between mutations at FL diagnosis and transformation risk suggests these mutations do not predispose to t-FL. However, the demonstration that apparent transformation specific mutations are indeed detectable, at low level, in preceding FL suggests they may be early events in the transformation process. Furthermore, these mutations might occur in long-lived disease sub-clones which exist many months before clinical presentation as transformation.

4.5.3 Evolution of FL / t-FL may occur from a common progenitor cell

Having provided confirmation of *EZH2* and *TNFRSF14* mutation frequencies in FL, this study went on to demonstrate that acquired, disease specific mutations can occur non-sequentially in FL / t-FL as, for both *TNFRSF14* and *EZH2* mutations, there were specific cases in which mutations were found only in FL (being absent from subsequent t-FL). This suggested that t-FL may not develop directly from the (principal) FL disease clone and on inspection of the CNAs for each sample across all chromosomes, analogous modes of FL / t-FL development were determined in 71% cases. These are consistent with reports from prior to this study (Martinez-Climent *et al.* 2003; Fitzgibbon *et al.* 2007) and others published during the course of this study (d'Amore *et al.* 2008; Johnson *et al.* 2008b; Ruminy *et al.* 2008; Eide *et al.* 2010) which identify that genetic changes present in FL can be absent from relapse or transformation; and helps explain the similar levels of CNAs in the initial FL sample set compared to the initial t-FL set. The presence of shared CNAs in such cases supports the existence of a common progenitor cell (CPC) that is the origin to different disease episodes through acquisition of genetic events that are specific to respective disease episodes. In certain cases, the time-course of this non-sequential (divergent) indirect clonal evolution of FL and its transformation indicates that the putative CPC is also long-lived, giving rise to different disease episodes many months, even years, apart. Although the CNAs provided a means by which the evolutionary pathway of most FL / t-FL cases could be inferred and these changes may harbour (or cause) genetic events that contribute to these tumours, they may also represent passenger events that occur *secondary* to the genomic instability that is a hall-mark of malignant cells. Consequently, it is an important finding that specific potential driver events, namely the focal mutations of *EZH2* or *TNFRSF14* that occur independently in case 7 or cases 22 / 23 / 33 respectively, are similarly suggestive of divergent evolution.

The CNAs that occur in all samples of a case with evolution from a putative CPC, and occur recurrently in different cases, may represent important early events in development of a putative CPC. In this cohort there are four MCRs of loss or gain affecting only one gene (1p36.13, 7q36.1, 8q23.3, 15q11.2) and a further four MCRs that affect only 2 to 4 genes (9p21.3, 12q12, 12q12-13.11, 12q23.1). In particular,

gain 1p36.13 includes *SMARCD3*, the product of which has a role in chromatin remodelling and in regulation of neuronal stem cells and merits further study in the context of a putative CPC in FL / t-FL.

Furthermore, the cases with more than just single FL and t-FL samples provide additional insights in to the complexity of disease evolution, as follows:

a) a "transformation-primed" progenitor clone can persist for at least 18 months before giving rise to t-FL relapse (**Figure 4.17**);

b) less primitive progenitor clones can give rise to either FL (**Figure 4.19**) or t-FL (**Figure 4.17**);

c) these clones can survive the therapy applied at initial transformation (**Figure 4.17** and **Figure 4.19**); and

d) the sequence of genomic CN change acquisition in the progenitor clones themselves may be divergent and not mirror the clinical course of disease (**Figure 4.19**).

Significantly, the demonstration of divergent evolution in FL / t-FL has important clinical implications regarding the process of disease recurrence as it suggests this may arise from a potential long-lived progenitor (that is resistant to conventional therapies). To examine this further, the concept of the putative CPC origin was addressed in more detail in a subset of cases and this investigation is described in the following chapter.

Chapter 5 Evolution of FL and its transformation

5.1 Introduction

The conventional model of FL transformation is one of progressive acquisition of additional genetic events that promote expansion of an affected FL sub-clone which eventually presents clinically as t-FL. This is supported by the demonstration that sequential FL / t-FL tumours are clonally related and that t-FL may arise from a single FL cell (Zelenetz *et al.* 1991). Furthermore, as described in Chapter 1, several subsequent studies provide evidence supporting this model.

However, recurrent secondary genetic changes in t(14;18) containing tumours suggest that this model may be an over-simplification (Horsman *et al.* 2001; Hoglund *et al.* 2004; de Jong 2005) with the findings from chapter 4 adding further complexity as, in certain cases, copy number changes or gene mutations present in FL are absent from subsequent t-FL. As certain changes remain common to both tumours and t-FL can acquire other genetic changes, this suggests that FL and t-FL can develop from a common precursor by acquisition of independent genetic events in a pattern of non-sequential divergent disease evolution.

This follows the proposal, over a decade ago, by Matolcsy and colleagues of a common malignant ancestor model for the evolution of FL. This was based on a non-sequential clonal relationship of *Immunoglobulin* V_HDJ_H sequence variability, in which sequence changes present in a case of FL were absent from subsequent t-FL (Matolcsy *et al.* 1999). Furthermore, reports soon indicated that FL relapse might arise by evolution from pre-existing minor sub-clones (Aarts *et al.* 2000; Aarts *et al.* 2001). Subsequently, support for the model has been provided by the identification (through genome-wide array-CGH performed on a series of samples from FL that relapsed or transformed) of DNA copy number losses or gains in 9 / 12 FL samples that were absent from the subsequent disease episodes (Martinez-Climent *et al.* 2003). More recently, analogous results were provided by the demonstration that aUPD may occur non-sequentially in the majority (10 / 18) of FL / t-FL cases (Fitzgibbon *et al.* 2007). These data suggest that a long-lived precursor may give rise to temporally and genetically distinct disease episodes. Clinical support for such a

model is provided by the demonstration that t-FL arising 11 years post-procedure in the recipient of an allo-HSCT (performed for high risk AML) was derived from donor cells. Significantly, the tumour contained an identical *BCL2-MBR* re-arrangement to that found in FL which developed in the donor 3 years post-donation (and which subsequently transformed). This suggests that a long-lived progenitor cell, which was passed from donor to recipient, gave rise to both lymphomas (Hart *et al.* 2007). These data support a model, as illustrated in Chapter 4, indicating that FL / t-FL may arise from common progenitor cells (CPCs) that can survive for months or even years in patients with FL.

To examine such evolution of FL / t-FL, it is important to have genetic markers that can distinguish differing disease clones including potential progenitor cells. As described in Chapter 1, FL originates from lymphoid germinal centres where SHM occurs in B cells. Furthermore, malignant clonal expansion in FL occurs after the initiation of SHM (Zelenetz *et al.* 1992) and ongoing SHM occurs within the tumour (Zelenetz *et al.* 1992; Ottensmeier *et al.* 1998). Consequently, SHM patterns can be utilised as markers of FL / t-FL clones (after sites of SHM have been established from tumour derived *IGH-V* sequences). Comparison of the SHM patterns in sequential FL / t-FL samples should then determine a pattern of the clonal evolution of disease progression and transformation. Furthermore, unique sites of SHM could be used to generate allele specific oligonucleotides enabling interrogation of FL tumours for the presence of t-FL clones and potential progenitor clones. By using a series of sequential FL / t-FL sample sets, this approach could provide important new insights in to the evolution of FL and its transformation.

5.2 Aims and Objectives

The suggestion that t-FL may not develop directly from preceding FL, and the possibility of examining sub-clones of these diseases using *IGH-V* SHM patterns, led to this study exploring the evolution FL transformation.

Study objectives included determination of *IGH-V* usage in sequential FL and t-FL samples, mapping the SHM patterns at the respective VH loci, interrogating FL samples with t-FL clone specific primers / probes to determine whether t-FL could be detected in preceding FL samples and (in cases showing divergent evolution)

identifying the SHM sequence of putative common precursor cells then confirming its presence in the respective FL and t-FL sample pairs.

5.3 Materials and Methods

5.3.1 Patients and samples

The same cases with sequential FL and t-FL samples that had been investigated in Chapter 4 were included in this study. Due to limitations of DNA availability, only a subset consisting of 18 cases could be examined. In parallel, samples from a father and son allo-HSCT donor / recipient pair who each subsequently developed FL / t-FL, (provided by Drs Jason Hart and Raymond Lai from the Departments of Medicine and Pathology, University of Alberta and Cross Cancer Institute, Edmonton, AB, Canada) were examined.

5.3.2 DNA preparation

DNA was extracted and stored as previously described (section **2.4.1**).

5.3.3 PCR detection of *IGH-VH* segment usage and SHM sites

The BIOMED consortium's VH segment consensus primers, as described in Chapter 1 and detailed in ***Appendix 6*** **Table 1**, were used to determine clonality of samples. Amplification was performed with either Amplitaq® Gold (Applied Biosystems), Platinum Taq® (Invitrogen) or AccuPrime™ Pfx (Invitrogen) DNA Polymerases (1.25units) in a final volume of 50µl containing the respective proprietary buffers (final concentration 1×), dNTPs (500µM), $MgCl_2$ (1.5mM), primers (400nM) and 100ng of DNA as further described in section **2.4.3**. To confirm sites of SHM and exclude single base changes generated by DNA polymerase template reading errors, PCR reactions were performed initially with Amplitaq® Gold then repeated with a second enzyme.

5.3.4 Homo-hetero-duplex analysis

Homo-hetero-duplex (HH) analysis and gel extraction of the *IGH-VH* PCR products, performed as described in sections **2.4.9** and **2.4.10**. The homo-duplex bands, representing the major tumour clones, were used to:

1. Identify VH segment usage by the tumour clone
2. Determine SHM patterns of the major clones from sequential FL / t-FL samples by PCR re-amplification of the extracted homo-duplex band using the appropriate VH primer and then sequencing (as in section **5.3.5**).
3. Design allele specific oligonucleotides on SHM sites specific to t-FL (as indicated in section **2.4.4**)

5.3.5 Determining SHM sites in tumour sub-clones

To map the evolution of SHM patterns within FL and t-FL sub-clones, PCR products from *IGH-VH* reactions were cloned using the TOPO TA Cloning® system and separate colonies were selected as templates for PCR amplification as described in section **2.4.11**.

5.3.6 Sequencing

Purified PCR products following cloning or HH analysis were directly sequenced using BigDye® Terminator chemistry as described in section **2.4.12**.

5.3.7 SHM sequences

The *IGH-VH* region sequences obtained from homo-duplexes and single colonies were then compared to germline sequence through the IMGT/V-QUEST programme available at http://www.imgt.org (Brochet *et al.* 2008) to determine SHM sites. The sites of SHM for each homo-duplex product were compared for each set of FL / t-FL sequential samples to determine the pattern of evolution. SHM sites from multiple clones from each tumour sample were also compared. This enabled the construction of genealogical trees. These chart the acquisition of SHM sites by different sub-

clones present within respective tumours and present them as an evolutionary hierarchy (examples are presented later in section **5.4.3**). Moreover, after excluding SHM sites unique to either FL or t-FL, the remaining SHM sites (present in both tumour samples) provided a mark that might help detect a putative common progenitor cell (CPC).

5.3.8 qPCR

To investigate the time-course of transformation, FL samples were examined for presence of t-FL clones. qPCR was performed using ASO primers and probes as described in section **2.4.7**.

5.4 Results

5.4.1 Patient samples

A total of 43 lymph node biopsy samples from 20 patients were analysed. They included FL and t-FL samples for each of the 18 cases with sequential samples, with an additional FL sample available for each of 5 cases, and one sample each from the HSCT donor / recipient pair.

Patient and sample characteristics for the series of cases with sequential sample sets are summarised in **Table 5.1**. The median age at FL diagnosis was 44 years (range 22 to 70) with 10 cases (56%) male. Available FL tumour grades at diagnosis (17 / 18 cases) were 1 (n = 12) or 2 (n = 5). One case had no detectable *IGH-BCL2* translocation, 12 cases had rearrangements involving the *IGH* MBR region and the remaining 5 involved *IGH* mcr or atypical regions, a distribution similar to previous studies (Tsujimoto *et al.* 1985a; Cleary *et al.* 1986a; Ngan *et al.* 1989; Buchonnet *et al.* 2002). FL samples were taken prior to treatment in 7 cases (3 available from diagnosis, 4 at progression after initial "watch and wait" expectant management), at first or second relapse in 9 cases and at later stages of disease for the remaining 2 cases. There was a median of 4 treatment courses (range 0 to 13) prior to t-FL and the median time to transformation from FL diagnosis for this cohort was 45.9 months (range 12.1 to 174.9), similar to that observed in two recent large series of FL patients (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008).

Table 5.1 Characteristics of the cases (n = 18) with sequential FL / t-FL samples

Case*	Gender	Age	Grade	Sample		TTT (months)	Tx prior to t-FL
				FL[§]	t-FL[#]		
1	F	41	2	Rel1	1	40.2	3
2	F	70	2	Prog	1	27.5	0
4	F	43	1	Rel1	1	120.0	4
5	M	45	2	Rel2	1	44.9	7
6	M	53	1	Prog	1	12.1	2
7	M	22	1	Rel2	1	114.5	5
8	M	67	2	Prog	1	35.7	1
9	F	47	1	Prog	1	94.5	7
11	M	46	1	Rel1	1	46.8	4
12	F	30	1	Rel7	1	174.9	13
15	M	30	1	Rel1	2	82.4	5
16	M	61	2	Rel2	1	22.8	3
20	F	26	1	Diag	1	12.1	1
21	F	32	NA	Rel2	1	72.9	4
23	F	42	1	Diag	1	39.3	0
25	M	45	1	Diag	1	18.2	2
26	M	40	1	Rel1	1	49.2	7
27	M	60	1	Rel2	1	104.0	4

*For five cases (indicated in bold) second FL samples were included in the study: in cases 26 and 27 the second FL sample was obtained from FL relapse prior to t-FL whereas in cases 2, 5 and 21 the sample was obtained at FL relapse after t-FL. Of the 18 cases, 15 have died and the median follow-up of the 3 remaining cases is 20.2 years (range 13.0 to 25.4).

[§]Earliest FL sample in the study.

[#]For t-FL samples, 1 indicates the first episode of transformation and 2 the second.

TTT indicates time to transformation from diagnosis of FL; Tx, a separate course of treatment (radiotherapy / chemotherapy / immunotherapy or combinations of these together); F, female; M, male; Rel, relapse; Prog, progression; Diag, diagnosis; and NA, not available.

5.4.2 *IGH-V* usage indicates common origin of FL and t-FL samples

The *IGH-V* segment (VH) usage for each of the 18 paired cases, as determined by *IGH-V* PCR then direct sequencing of homo-duplex products, is detailed in **Table 5.2** with examples of homo-duplex bands and the subsequent PCR products obtained after amplification of DNA from the extracted bands provided in **Figure 5.1** and **Figure 5.2**, respectively. A homo-duplex band (major clone) was obtained for all samples, with FL and t-FL samples from respective cases showing the same VH segment usage which indicated they were clonally related. The majority of cases (n = 13) showed VH3 usage, 4 used VH4 and, in a single case, the VH5 segment had been selected. This pattern of VH segment usage is in keeping with those described in previous reports for normal B cells (Schroeder *et al.* 1987; Mortuza *et al.* 2001) and malignant B cells (Pritsch *et al.* 1999; Camacho *et al.* 2003) including FL cells (Hummel *et al.* 1994). In the majority of FL / t-FL cases (n=17) the *IGH-V* sequences contained no stop codons nor insertions / deletions with accompanying frame-shifts and were potentially functional; in the remaining instance (case 11) the selected VH5 segment was partially deleted in both FL and t-FL samples.

5.4.3 SHM in major (HH) clones indicates divergent evolution can occur

Comparison of SHM sites in the homo-duplex sequences from paired FL / t-FL samples indicated that in the majority of cases (n = 12) sequential additive acquisition of SHM sites in the progression from FL to t-FL (indicating direct evolution as illustrated in ***Appendix 6*** **Figure 1**) did *not* occur. Rather, indirect evolution (**Figure 5.3**) could be demonstrated with common SHM sites found in both FL and t-FL but additional SHM sites found in FL that were absent from t-FL. This indicated that t-FL arose by divergent evolution from an earlier clone, which contained fewer SHM sites than occurred in FL, representing a putative common progenitor cell (CPC) origin for both tumours. In these 12 cases the median number of tumour specific SHMs acquired by the major clones of the FL and t-FL samples were 4 (range 1 to 9) and 3 (range 0 to 22), respectively, with a median homology of the CPC sequence to germline of 89.78 % (range 83.47 [Case 6] to 94.49 [Case 12]). The evolution patterns determined from the homo-duplex sequence results for all 18

Table 5.2 FL / t-FL *IGH-V* segment usage and evolution on HH analysis

Case number*	*IGH-V* allele[#]	Evolution pattern		Tumour specific SHMs[$]		CPC homology to germline (%)
		SHM	*CN*	FL	t-FL	
1	*IGH-V4-34*01/*02*	CPC	*CPC*	5	12	87.35
2	*IGH-V3-11*01*	DIRECT	*DIRECT*	-	-	-
4	*IGH-V3-48*03*	CPC	*CPC*	4	3	90.73
5	*IGH-V3-48*03*	DIRECT	*np*	-	-	-
6	*IGH-V3-21*01*	CPC	*CPC*	4	3	83.47
7	*IGH-V4-61*01*	CPC	*CPC*	9	22	94.42
8	*IGH-V4-59*01*	CPC	*CPC*	9	0	93.51
9	*IGH-V3-23*01/*03/*04/*05*	DIRECT	*CPC*	-	-	-
11	*IGH-V5-51*01*	DIRECT	*CPC*	-	-	-
12	*IGH-V3-15*01/*02*	CPC	*CPC*	4	3	94.49
15	*IGH-V4-61*02*	CPC	*CPC*	4	12	89.24
16	*IGH-V3-23*01/*02/*04*	CPC	*CPC*	1	3	87.10
20	*IGH-V3-23*01/*04*	DIRECT	*DIRECT*	-	-	-
21	*IGH-V3-53*01*	DIRECT	*CPC*	-	-	-
23	*IGH-V3-48*01*	CPC	*CPC*	4	1	90.73
25	*IGH-V3-30*01/-03*01*	CPC	*CPC*	4	7	90.32
26	*IGH-V3-11*01*	CPC[§]	*CPC*	6	4	87.90
27	*IGH-V3-23*01/04*	CPC	*CPC*	3	0	85.89

The *IGH-V* segment usage for the FL and t-FL samples from each respective case was the same. The evolution pattern of FL to t-FL suggested by copy-number (CN) changes (from chapter 4) is shown for comparison to that determined by SHM analysis.
For all cases (n = 12) with indirect (CPC) evolution by SHM analysis, this pattern was confirmed by CN analysis.
In three cases (9, 11 and 21) with apparent direct evolution according to SHM analysis (underlined), the copy number analysis provided additional evidence that indicated evolution had occurred divergently from a putative CPC.
[*]Cases with second FL samples analysed are highlighted in bold (as detailed in **Table 5.1**). For the copy number analysis performed in chapter 4, the second FL samples from cases 2, 5 and 26 had not been available.
[#]*IGH-V* allele used by the major tumour clones. Nomenclature as per the IGMT database (Brochet *et al*. 2008).
[$]For the twelve CPC cases, the number of SHMs acquired by the CPC in subsequent FL and t-FL major clones are indicated.
[§]In case 26, the SHM analysis revealed that FL showed divergent evolution as compared to both relapsed FL and subsequent t-FL but t-FL showed direct evolution from FL relapse (***Appendix 6* Figure 2**).
IGH-V: immunoglobulin heavy chain variable region; SHM, somatic hypermutation; CN, DNA copy number; CPC: common progenitor cell; and np, not possible to determine whether DIRECT or CPC.

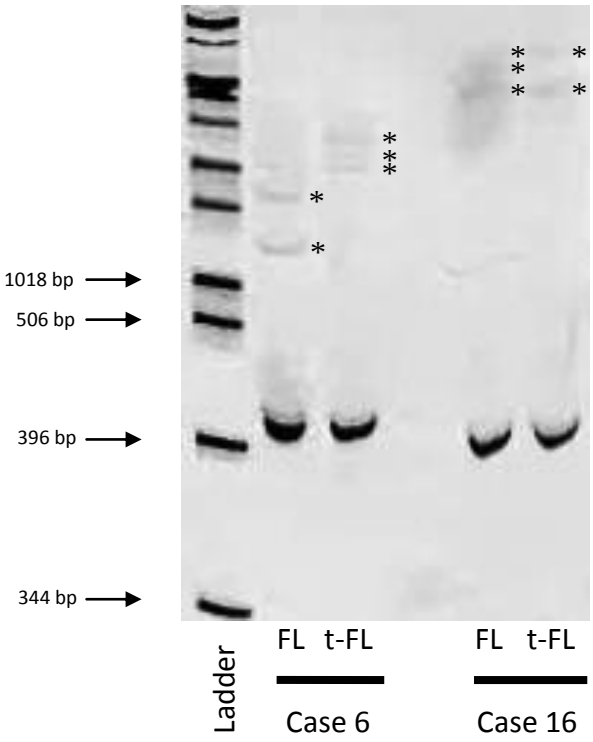


Figure 5.1 Homo-hetero-duplex analysis

DNA from paired FL and t-FL samples from two cases (6 and 16) is shown. Following PCR specific to the *IGH-V* region usage (VH3 for both cases in this example) the PCR products were subject to homo-hetero-duplex formation. The products of this are observed here following electrophoresis at 20mA for 16 hours on a polyacrylamide gel. The image is inverted to maximise the visualisation of bands. Prominent homo-duplex bands are clearly visible for each sample at around 400bp in size. More slowly migrating hetero-duplex bands are seen as fainter bands occurring earlier in the gel (*). The prominent bands represent the major (most highly represented) SHM sequence pattern of the *IGH-V* region present in the tumour DNA samples.

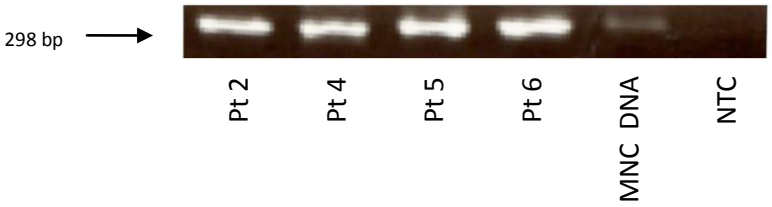
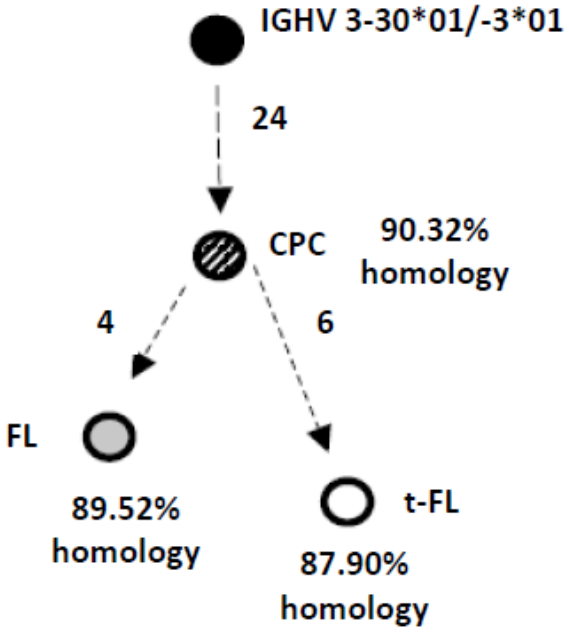


Figure 5.2 PCR of homo-duplex products

Four of the follicular lymphoma samples which were subject to hetero-duplex analysis are illustrated. All four cases involve VH3 segment usage. The homoduplex bands were then amplified by PCR and the products sequenced. Only 10ng of MNC DNA were used in the MNC DNA control lane. MNC pools contain DNA from multiple differing B cell sub-clones: those with VH3 segment usage generate the band seen here for the MNC DNA used as a control.
MNC: mononuclear cell; NTC: no template control.



Carlotti *et al*. 2009

Figure 5.3 Homo-duplex derived SHM sequences from sequential FL / t-FL

Schematic summary of SHM analysis on homo-duplex major clones from the FL and t-FL samples of one patient (Case 25) illustrating the acquisition of SHMs from germline through CPC on to FL and t-FL. The numbers beside arrows indicate the amount of SHM sites acquired between each successive step in the ontogeny. The numbers for FL (4) and t-FL (6) arrows represent mutually exclusive, tumour specific SHMs. The percentage homology of tumour *IGH-V* sequences to the germline *IGH-V* sequence (circle filled with black) is provided for both FL (circle filled with grey) and t-FL sequences (unfilled circle). The sites of SHM common to both samples (n = 24) provide a common *IGH-V* sequence (represented by the circle filled with hatching), which has higher homology to germline and so represents a sequence that is intermediate between germline and tumour. This then acquires mutually exclusive additional sites of SHM found in either FL (n = 4) or t-FL (n = 6), which mark divergence of the two separate tumour clones from this intermediary.

cases and the tumour specific SHMs and germline homology for the CPC cases are provided in **Table 5.2**.

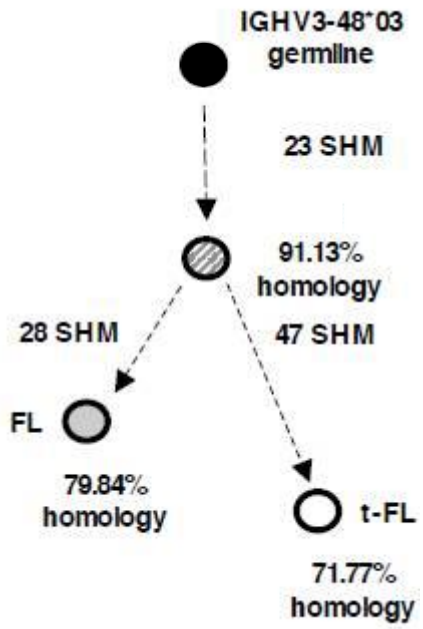
For the father and son cases, one FL and one t-FL sample were available respectively. The *IGH-V* segment usage in the homo-duplex bands of both samples was the same (IGHV3-48*03), indicating these samples were clonally related. On determination of SHM sites from the homo-duplex bands, 23 common to both the FL sample (father) and t-FL sample (son) were identified with additional tumour specific SHM sites present in the FL (n = 28) and t-FL (n = 47) samples (**Figure 5.4**). This suggested that the tumours present in father and son had undergone divergent evolution (in an analogous manner to the 12 sequential cases) from a common origin marked by the 23 shared SHM sites (and the common *BCL2-IGH* fusion site previously identified (Hart *et al.* 2007)) giving a 91.1% homology of the *IGH-V* sequence in the CPC to germline.

5.4.4 Cloning confirms divergent evolution can occur

To further investigate the CPC model of tumour development, six of the twelve sequential cases identified through HH analysis as having patterns compatible with divergent acquisition of tumour specific SHM in FL and t-FL (as shown in **Table 5.2**) were selected, based on DNA availability, for examination by cloning. An additional sequential case showing direct evolution and the father / son set were also included in this study.

Following re-amplification by PCR (**Figure 5.5**), the *IGH-V* sequences derived from multiple colonies from the cloning assay of each sample were compared to enable mapping of SHM in FL and t-FL through generation of genealogical trees specific to each case. An example is given in **Figure 5.6** for the same case illustrated in **Figure 5.3**. The results for all six putative CPC cases (and the additional case showing direct, sequential SHM acquisition from FL to t-FL) are shown in **Table 5.3**. A similar number of SHM sites were detected by cloning as compared to HH analysis across both tumour types for all seven cases.

A median of 15 (range 9 to 25) and 9 (3 to 18) unique *IGH-V* sequences were identified by comparison of SHM patterns from a median of 28 (range 17 to 47) and



Carlotti *et al*. 2009

Figure 5.4 Homo-duplex derived SHM sequences from father and son tumours

Schematic summary of SHM analysis on homo-duplex major clones from the father (FL) and son (t-FL) tumour samples. The symbols and labelling are as described for **Figure 5.3** and, similar to that example, an intermediate *IGH-V* sequence can be defined (containing 23 common SHM sites) which undergoes divergent acquisition of tumour specific SHM sites in FL (n = 28) and t-FL (n = 47).

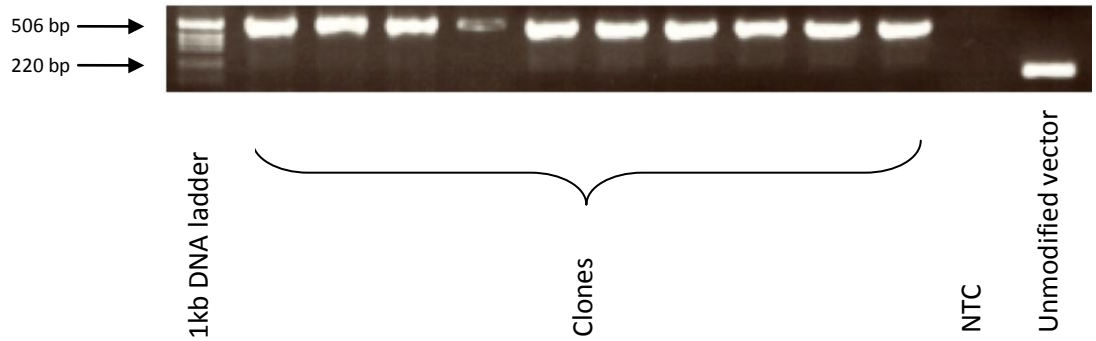


Figure 5.5 PCR of cloning products

An example of the bands generated by PCR applied to the products of cloning from the pCR®2.1-TOPO® kit. The unmodified vector has no DNA insert. Each clone has a larger PCR product which indicates incorporation of a DNA insert (each of similar size as they are derived from the same VH segment).
Clones indicates PCR products of individual white colonies (the pCR®2.1-TOPO® vector with incorporated IGH-V product) obtained by cloning; NTC, no template control (H_20); and unmodified vector, PCR product of a single blue colony (the naive pCR®2.1-TOPO® vector without IGVH PCR product incorporation) obtained by cloning.

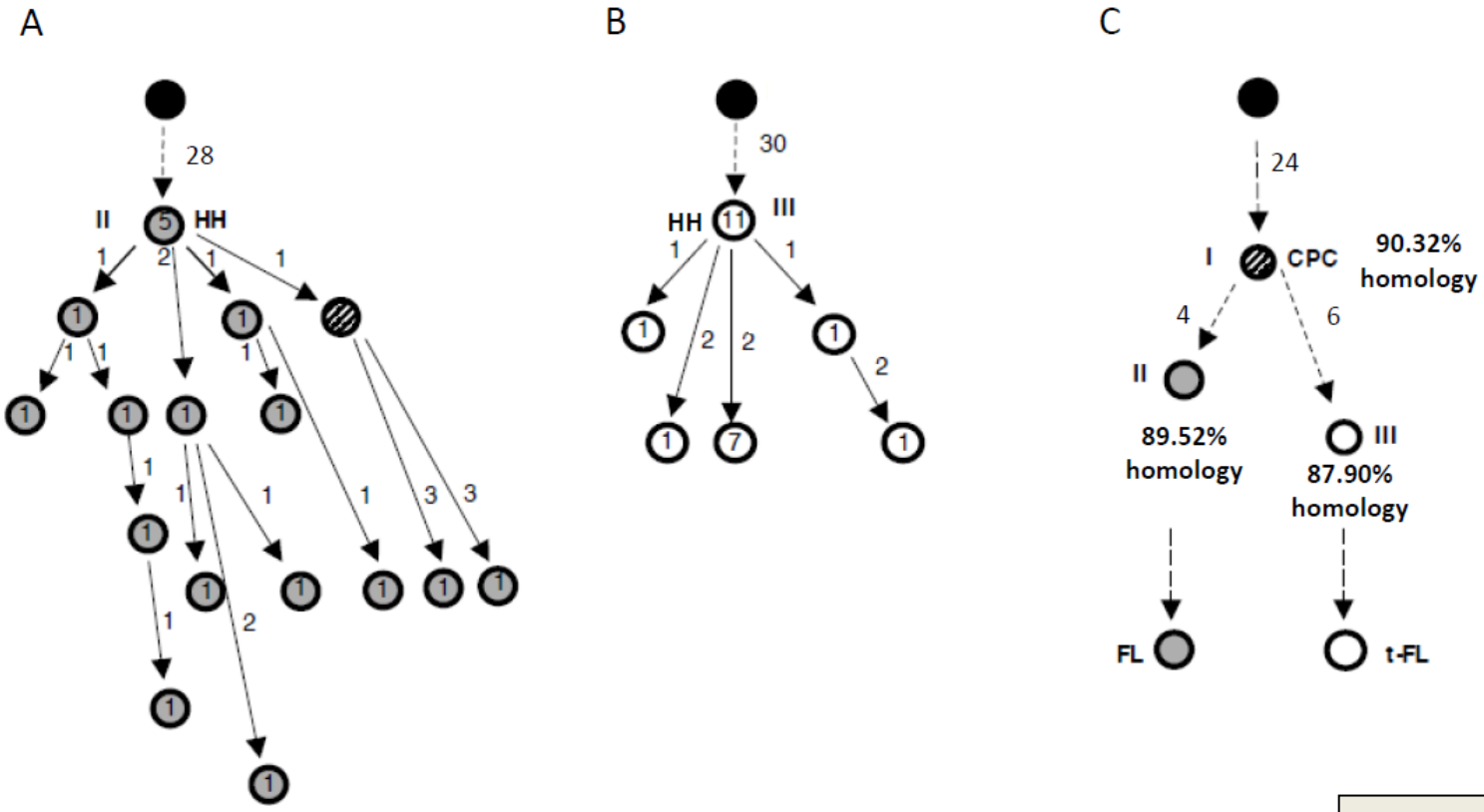


Figure 5.6 Cloning *IGH-V* sequences indicates FL / t-FL evolves from a CPC

Illustrated are genealogical trees summarising cloning results from the same case (25) illustrated for HH analysis in **Figure 5.3**. A: FL, B: t-FL, and C: combined summary of FL and t-FL analysis. The circles represent the same as indicated in **Figure 5.3** with the addition that for A and B each grey, white or hatch filled circle represents a unique *IGH-V* sequence (clone) with acquired SHM. Digits in the circles represent the number of each respective clone identified. Numbers beside arrows represent the number of SHM sites acquired between successive stages in the hierarchy. The major clones identified by HH analysis are indicated. I, II and III indicate levels in the hierarchy with comparison of sequences in II (FL) and III (t-FL) enabling prediction of a common SHM pattern present in the putative CPC (I).

Table 5.3 Comparison of HH and cloning in evolution determination

Case	FL				t-FL				Evolution pattern	
	Clones		SHM		Clones		SHM			
	Total[§]	Distinct[#]	Cloning range* (rate)	*HH*	Total[§]	Distinct[#]	Cloning range* (rate)	*HH*	Cloning	*HH*
1^	34	9	36-37 (0.11-0.11)	*36*	35	10	41-45 (0.12-0.14)	*43*	CPC	*CPC*
2	26	15	39-48 (0.13-0.15)	*44*	36	10	48-53 (0.15-0.17)	*48*	Direct	*Direct*
6	28	25	45-51 (0.14-0.16)	*46*	38	7	45-46 (0.14-0.15)	*45*	CPC	*CPC*
12	47	19	20-24 (0.06-0.07)	*22*	35	18	19-22 (0.06-0.07)	*21*	CPC	*CPC*
15	20	12	29-41 (0.09-0.12)	*31*	24	9	39-41 (0.12-0.12)	*39*	Direct	*CPC*
25	20	15	28-32 (0.08-0.10)	*28*	23	6	30-33 (0.09-0.10)	*30*	CPC	*CPC*
27	17	14	39-43 (0.12-0.13)	*42*	18	3	39-43 (0.12-0.13)	*39*	CPC	*CPC*

Shown are summaries of SHM analysis performed by the cloning analysis for FL and t-FL. The number of SHM sites and the evolution pattern in the major (homo-duplex) clones are indicated in italics for comparison. All HH results lie within the range determined by the cloning analysis. Only case 15 showed discordance in the evolution pattern determined by the two approaches, however, the presence of an earlier common precursor could not be excluded by the cloning analysis.

[§]The total number individual colonies sampled from which clone sequences were generated.

[#]The number of distinct clone sequences identified from all the clones sampled.

*The range of SHM sites per clone in the distinct clone sequences; rate indicates number of SHM sites per bp of PCR product [calculated as SHM sites identified in a PCR product / product length (bp)]. *HH indicates homo-hetero-duplex analysis.*

35 (18 to 38) single colony derived tumour sub-clones from the FL and t-FL samples, respectively (**Table 5.3**). Using the Wilcoxon matched pair signed rank test, a significant difference, $P = .04$, was demonstrated between the levels of clones identified in the FL and t-FL samples.

In 6 / 7 sequential cases the mode of SHM acquisition concurred with the HH analysis. In a single instance (Case 15), cloning indicated sequential acquisition of SHM by t-FL from FL in contrast to the divergent model suggested by HH analysis. However, the data for this case did not exclude the presence of a common intermediate (with few SHM sites) that was not detected by cloning but which could support divergence of SHM patterns for this case.

In the father and son HSCT donor / recipient pair, cloning supported the HH evidence for a divergent pattern of evolution from a CPC. From 36 colonies, the father's tumour showed 18 distinct clone sequences (SHM range 51 to 53) with 32 colonies obtained from the son's tumour sample revealing fewer distinct clones (8) but a higher rate of SHM (range 70 to 73). The total number of HH identified SHMs for the father (51) and son (70) (**Figure 5.4**) tumour samples lie within the respective cloning ranges.

In summary, the genealogical trees generated here from the cloning data are consistent with previous reports of ongoing SHM in FL (Zelenetz *et al.* 1992; Ottensmeier *et al.* 1998) and confirm that divergent evolution of tumour clones from a putative CPC can occur across multiple cases.

5.4.5 t-FL clones are detectable in preceding FL tumours

In chapter 4, for a single case, the t-FL clone could be detected (at low level) in the preceding FL sample many months prior to clinical transformation. This indicates that a long-lived transformation primed sub-clone can co-exist with the (major) FL clone. As the putative CPC might give rise to both FL and t-FL, determining whether the tumour clones co-exist would provide insights in to their temporal evolution. This is important as long-lived transformation primed clones may survive prior therapies and treatment resistance is a significant factor complicating disease management.

By using SHM sites unique to major t-FL clone *IGH-V* sequences, as illustrated in **Figure 5.7 (A)**, ASO primers and probes specific to the t-FL sequence could be designed (**Figure 5.8**) so that FL samples could be examined for the presence of t-FL clones. In contrast to the acquired gene mutations in chapter 4 that occur only in a subset of cases, SHMs affecting *IGH-V* are common in FL / t-FL so a more comprehensive screen could be performed. In view of the DNA quantity required for the assays, eight cases could be included in this study which consisted of the seven sequential cases used in cloning and case 26. Ten FL samples were analysed as cases 26 and 27 each had a second FL sample available which preceded t-FL. Each qPCR assay of the FL samples was performed in duplicate using two different t-FL specific primer / probe combinations and these produced similar results for all ten assays (**Table 5.4**).

Importantly, the t-FL clone was detected in 5 / 10 FL samples (as illustrated for case 12 in **Figure 5.9**) at levels ranging from $10^{-1}/10^{-2}$ to 10^{-4} relative to the quantity of the clone in t-FL samples (**Table 5.4**). Moreover, the t-FL clone could be detected in FL up to 2¼ years before clinical transformation occurred. In case 26, the t-FL clone was detected only in the more recent of two FL samples suggesting either it was below the detection threshold in the earlier FL sample or it had developed at some point between the two episodes of FL. Divergent evolution from a putative CPC had been suggested for four of these cases but the fifth (case 2) had shown a direct pattern of evolution (**Table 5.2**). The eight months between FL and t-FL samples in this case indicates that t-FL primed clones that arise directly from FL (as opposed to a CPC) may also co-exist with FL for some time prior to clinical presentation as transformation.

5.4.6 Detection of a CPC in both FL and t-FL samples

Following the demonstration that t-FL clones can be detected in FL, this experiment sought to determine whether the putative CPC could be detected in tumour samples and so provide a direct link between divergent FL / t-FL clones. ASO primers specific to the imputed *IGH-V* sequence of the CPC were designed as illustrated in **Figure 5.7 (B)**. This sequence contained only SHM sites present in both FL and t-FL samples but not the tumour specific sites. The ASOs were applied in nested (4

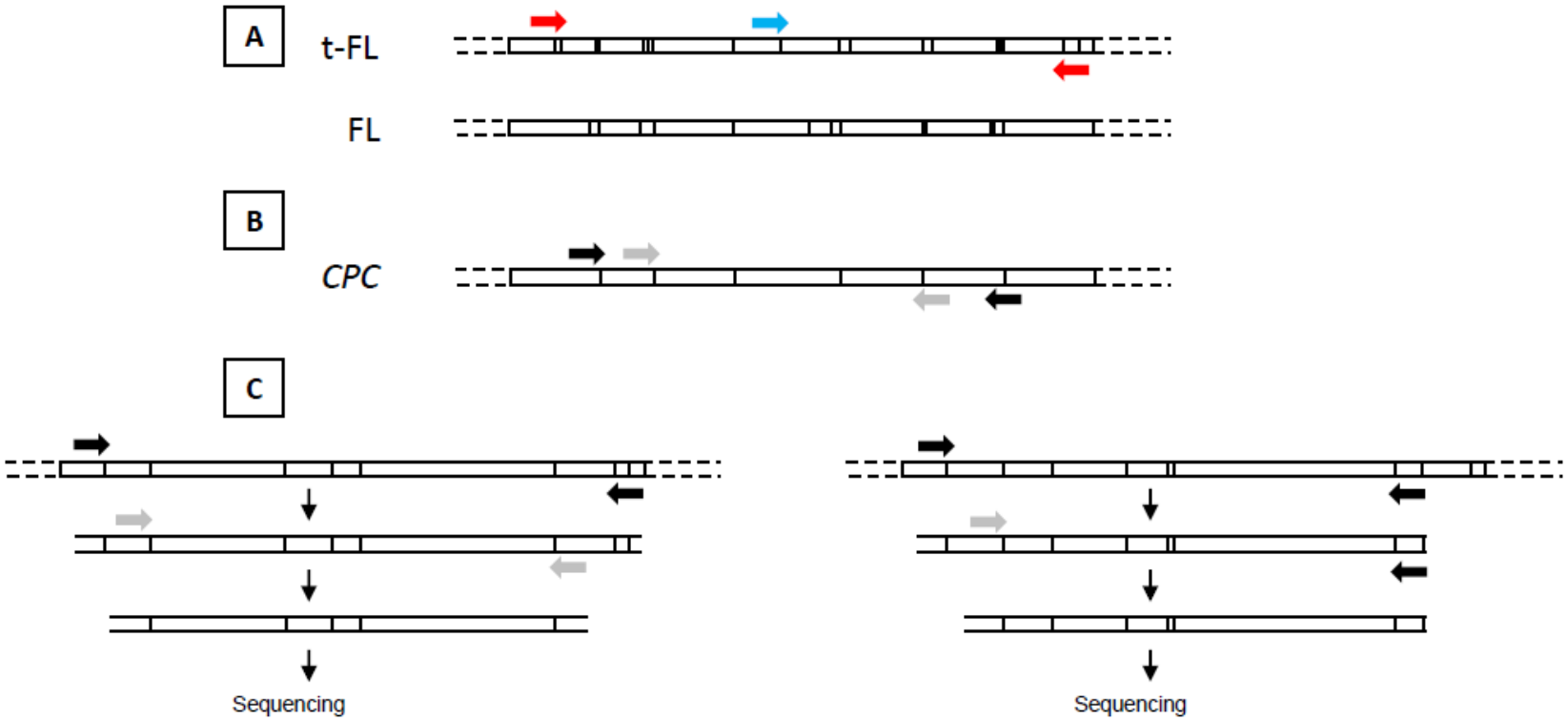


Figure 5.7 Allele specific oligonucleotide design and (hemi-) nested PCR

Schematic representations of the *IGH-V* region with sites of SHM indicated by vertical bars. A: examples of SHM patterns in t-FL and FL *IGV-H* sequences with t-FL specific primers (red arrows) and probe (blue arrow) designed on SHM sites unique to t-FL. B: example of imputed SHM sites in CPC *IGH-V* sequence; forward and reverse primers specific to the CPC sequence (black and grey arrows) are designed to overlap both common SHM sites (incorporated in the primer) and FL / t-FL specific SHM sites (germline sequence used in the primer) thereby ensuring specificity to only the CPC *IGH-V* sequence. C: examples of the two rounds of PCR followed by sequencing of reaction products for, left, nested PCR of CPC sequence using four CPC specific primers; and right, hemi-nested PCR using three CPC specific primers.

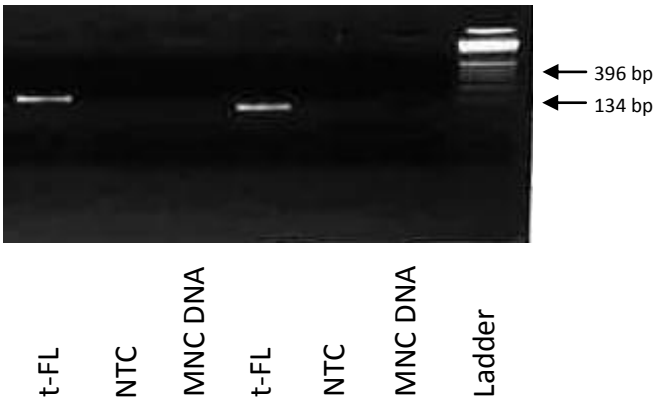


Figure 5.8 Optimisation of t-FL major clone specific primers

An example agarose gel visualisation of the PCR products obtained using two sets of t-FL *IGH-V* sequence specific primers (giving different amplicon sizes) for the t-FL sample of one case. Specificity of the primers for t-FL is indicated by the lack of amplification when MNC DNA is used as a reaction template.

Table 5.4 Identification of t-FL clones in preceding FL by qPCR

Patient	Sample[¥]	Primer site[†]		Primers / sample*	Sensitivity[#]	Time from FL sample to t-FL (months)	Amplification Level[§]
		Forward	Reverse				
1	FL	CDR1-FR2	CDR2-FR3	4	10^{-4}	26.8	10^{-2} / 10^{-3}
	FL	CDR1	FR3		10^{-4}		10^{-2} / 10^{-3}
2	FL	FR2-CDR2	FR3	3	10^{-4}	8.4	~10^{-3}
	FL	FR2-CDR2	FR3		10^{-4}		~10^{-3}
6	FL	FR1-CDR1	FR3	4	10^{-4}	5.5	~10^{-4}
	FL	FR1-CDR1	FR3		10^{-4}		~10^{-4}
12	FL	FR2-CDR2	FR3-CDR3	4	10^{-4}	76.8	No amp
	FL	FR2-CDR2	FR3-CDR3		10^{-4}		No amp
15[$]	FL	CDR1-FR2	FR3	3	10^{-4}	72.0	No amp
	FL	CDR1-FR2	FR3		10^{-4}		No amp
25	FL	FR2-CDR2	FR3	3	10^{-4}	18.2	No amp
	FL	FR2-CDR2	CDR3		10^{-4}		No amp
	FL2	FR2-CDR2	FR3	3	10^{-4}	8.5	No amp
	FL2	FR2-CDR2	FR3		10^{-4}		No amp
26	FL	FR1-CDR1	FR3	3	10^{-4}	33.8	No amp
	FL	FR1-CDR1	FR3		10^{-4}		No amp
	FL2	FR1-CDR1	FR3	3	10^{-4}	19.8	10^{-1} / 10^{-2}
	FL2	FR1-CDR1	FR3		10^{-4}		10^{-1} / 10^{-2}
27	FL	FR2-CDR2	CDR3	4	10^{-4}	25.4	10^{-1} / 10^{-2}
	FL	FR2-CDR2	CDR3		10^{-4}		10^{-1} / 10^{-2}

Results are given in duplicate for each patient sample as two qPCRs using different primers and probe combinations were used for each sample. All duplicates showed concordant results.

[¥]Cases 25 and 26 had second FL samples (FL2) available from subsequent disease episodes prior to transformation.

[†]*IGH-V* region where the t-FL clone specific primer was designed

*Total number of primers applied to each sample in the two qPCRs.

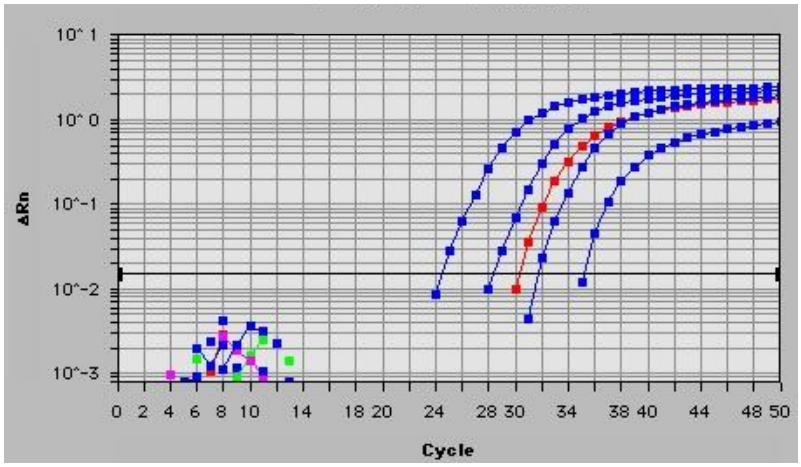
[#]Maximal sensitivity detected for each primers + probe combination.

[§]Relative amplification level of the major t-FL clone within FL sample compared to t-FL sample.

[$]For case 15, two different probes were used, one in each qPCR.

CDR and FR indicate cluster determination region and framework region respectively as described in Chapter 1; and No amp indicates no amplification detected.

A



B

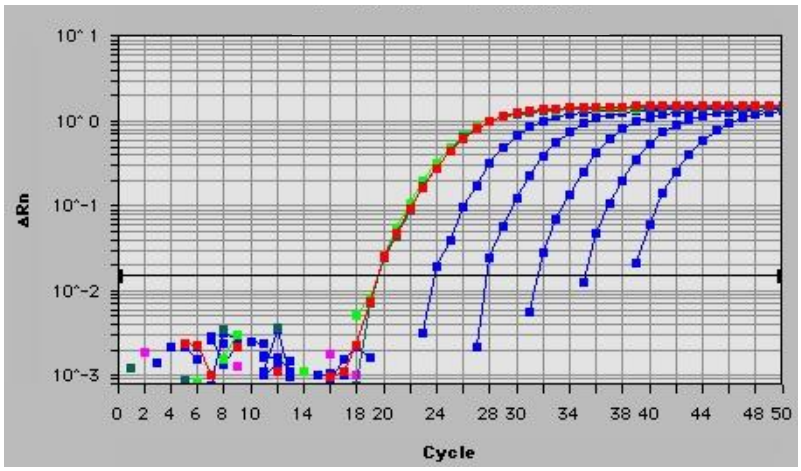


Figure 5.9 qPCR of t-FL clone in preceding FL

Two qPCR plots (for case 1) demonstrating amplification of (A) the t-FL clone in FL and (B) β_2M. Shown in blue are dilutions of a 100ng/µl preparation of t-FL DNA (from left to right: A, 10^{-1} to 10^{-4} and B, 10^{-1} to 10^{-5}). To preserve t-FL DNA, the 100ng/µl working solution was not run. In red is the FL sample DNA (at 100ng/µl) with a no template control (H_20) in purple and germline DNA (MNC DNA at 100ng/µl) in green. Reactions were performed in triplicate but for clarity single reactions are illustrated. The relative level of t-FL clone in the FL sample shown (between 10^{-2} and 10^{-3}) indicates that it constitutes <1% of the FL tumour. The similarity of amplification between t-FL clone and β_2M assays for the t-FL tenfold dilutions indicates that the t-FL clone constitutes virtually the entirety of the t-FL tumour sample. (β_2M amplification confirms similar total DNA levels in the FL, t-FL and MNC DNA samples.)

primers) or hemi-nested (3 primers) PCR, **Figure 5.7 (C)**, to FL / t-FL samples that had shown divergent evolution. This approach provided a means of:

i. minimising contamination of the second PCR product by products of unintended primer binding;

ii. amplifying low levels of template

and was possible in 6 / 7 putative CPC cases examined by qPCR in the previous section. For (hemi-) nested PCR products providing a single band on agarose gel electrophoresis, the product was then cloned (as described in section **2.4.11**) and sequenced.

Significantly, sequencing confirmed detection of the predicted CPC *IGH-V* sequence for cases 12 and 26. In a further patient (Case 1), a common sequence containing fewer SHM sites (representing an earlier common intermediate clone) was detected (**Table 5.5**). Consequently, these data demonstrate that CPC clones can be detected in tumours by using SHM profiles as specific markers and they can co-exist with both FL and t-FL. Moreover, the detection of the CPC in samples obtained over six years apart (Case 12) provides important evidence for its considerable longevity.

5.5 Discussion

This study demonstrates that evolution of t-FL from FL may occur by non-sequential ("indirect") clonal evolution from a putative common origin (a common progenitor cell; CPC), as illustrated in ***Appendix 6*** **Figure 3**. In 12 / 18 (67%) cases with sequential FL and t-FL samples, and in the donor / recipient pair, divergent evolution of tumours was identified by comparison of SHMs at their respective *IGH-V* segments. To examine further the clonal development in these cases, precursor *IGH-V* sequences were defined by the presence of only those SHMs common to both tumours. These sequences had between 83.5% and 94.5% homology with the respective VH segment germline sequence and served as genetic markers for the putative CPC enabling its detection in both FL and t-FL. Similarly, transformation specific SHMs enabled detection of t-FL clones in preceding FL.

Considerable longevity could be attributed to the CPC clones, as they were detected in sequential FL and t-FL samples taken up to 6½ years apart. Furthermore, indirect

Table 5.5 Identification of CPC sequence in FL and t-FL samples.

Case	Sample	Primers					Clones screened[¥]	CPC detected
		Number[$]	Forward		Reverse			
			Region[†]	SHM[§]	Region[†]	SHM[§]		
1	FL	4	FR1-CDR1	2	FR3	3	31	earlier**
	t-FL						-	-
6	FL	4	FR2-CDR2	6	CDR3-N1-N2	N1, N2	11	no
	t-FL						11	no
12	FL	3	CDR2-FR3	2	CDR3-N1	5	18	YES (16)*
	t-FL						39	YES (3)*
15	FL	3	FR2-CDR2	4	N2	N2	22	no
	t-FL						-	-
25^	FL	3	FR3	7	FR3	3	7	no
	t-FL						7	no
25^	FL	4	FR3	7	N2	N2	7	no
	t-FL						7	no
26	FL	3	CDR1-FR2	5	N1	N1	11	YES (5)*
	t-FL						11	YES (4)*

Summarised is the approach for the detection of CPC clones in FL and t-FL using clone-specific primers designed on unique *IGH-V* sequences identified through SHM analysis of homo-duplex products. In cases 1 and 15, nested and hemi-nested PCR on the respective t-FL samples was not successful so cloning results are not available. In two cases, 12 and 26, the CPC clone was successfully detected in both FL and t-FL samples.

[$]The number of primers used across the two reactions of nested (n = 4) or hemi-nested (n = 3) PCR.

[†]*IGH-V* region where the t-FL clone specific primer was designed.

[§]Total number of SHM sites represented on the combination of forward (or reverse) primers used.

[¥]The total number of individual colonies sampled from which clone sequences were generated.

**In case 1, although the putative precursor sequence was not detected, a sequence with fewer SHM sites (representing an earlier precursor) was detectable in the FL sample.

*The number of clones in which the CPC was detected is indicated in parentheses; the remainder of the clones screened yielded a more mature sequence attributable to non-specific priming.

^For case 25, reverse primers on FR3 and on DH-JH (N2) regions could be designed, enabling one nested and one hemi-nested PCR for this case.

N1and N2 indicate clone specific mutations by nucleotide insertion in the VH-DH and DH-JH junctional regions, respectively.

experimental support for the CPC model is provided by the demonstration here that a transplantable (Hart *et al.* 2007) long-lived lymphoma-initiating cell can give rise to clonally related, but distinct, disease in separate individuals. Specifically, FL presenting with transformation over one decade post-HSCT in the recipient arose by divergent clonal evolution from a long-lived CPC that was transplanted from the donor who themselves developed FL over three years post-procedure.

These data support and extend the hypothesis of FL development proposed by Matolcsy and colleagues (Matolcsy *et al.* 1999), although the SHM patterns used then and here remain only a surrogate marker for the malignant clone. However, validation for the concept of a common cell of origin in FL comes from three recent reports which identified that divergence of acquired cytogenetic aberrations occurs in sequential FL samples (d'Amore *et al.* 2008) and that, in 40 to 60% of cases, subsequent biopsies lack alterations found in the original FL karyotype (Johnson *et al.* 2008b; Eide *et al.* 2010). A further study that assessed mutations of the *IGH* S(mu) region, which precede (but do not necessarily lead to) class switch recombination, provided evidence that FL relapse could arise through the evolution of pre-existing minor sub-clones and that these could have fewer mutations and represent earlier sub-populations of disease (Ruminy *et al.* 2008).

The use of a second technique (cloning) provided similar rates of SHM, and discrimination of FL / t-FL evolution, compared to the initial HH analysis supporting the applicability of HH in mapping disease evolution. It also provided further evidence for the known oligoclonality of FL (Zhu *et al.* 1994; Ottensmeier *et al.* 1998; Bognar *et al.* 2005). Of course, HH only yields the most prevalent (major) tumour clone and so cannot directly address any other sub-clones (up to 25 and 18 in FL and t-FL, respectively) nevertheless, it could be used to impute CPC clones which were subsequently detected in tumour samples.

Similar levels of subsequent SHMs were acquired by the CPC in the sequential FL / t-FL samples (with medians of 4 and 3 respectively) but considerably more were acquired by the tumours in the donor / recipient pair (28 and 47 in the major clones respectively). This variation may reflect contrasting development of the CPC in the two scenarios as, in the donor / recipient pair, the CPC has migrated to the BM whereas in the sequential cases the CPC may never have left the germinal centre.

Furthermore, the length of time (>10 years) between transplant and development of tumour in the recipient may have contributed to the high number of SHMs in this sample.

A lower degree of oligoclonality observed by cloning in t-FL as compared to FL (P = .04) is likely to reflect the prominent outgrowth of a single disease clone. The detection of t-FL clones within earlier FL samples obtained over 2 years prior to transformation is not contradictory to this scenario, but reflects a sub-clone which has a specific SHM pattern that (at some future time-point) acquires changes conferring aggressive phenotype. This then becomes by far the dominant clone in contrast to FL where less rapidly proliferating sub-clones may develop a more diverse spectrum of SHM sub-clones.

Although the SHM patterns provide tumour genetic markers (valuable in quantitating t-FL and CPC clones), these reflect the activity of a normal B-cell process occurring in malignant cells. Consequently, the SHMs may not completely map all tumour sub-clones. This was evident for cases 9, 11 and 21 where SHM determination suggested direct evolution of t-FL from FL whereas the pattern of acquired CN changes (determined in chapter 4) indicated that evolution was in fact divergent from an inferred CPC. It is possible in these cases that, firstly, additional tumour specific SHMs were not detected in FL and t-FL clones or, alternatively, a CPC containing a certain level of SHM then acquired divergent CN changes, perhaps driving either FL or t-FL disease evolution, with subsequent SHMs in the t-FL clone creating a pattern of direct evolution by SHM but indirect evolution by CN changes.

What contributes to selection of t-FL clones is not clear, although specific genetic changes which predispose to transformation are present at FL diagnosis and certain genetic events or expression profiles mark transformation, as described in Chapter 1. Here, clones with the same SHM pattern as t-FL are present in FL samples from patients managed in a variety of ways: an expectant “watch and wait” policy (case 2), local radiotherapy (case 26) or chemotherapy (cases 1, 6 and 27). This suggests that selective pressures leading to further outgrowth of these t-FL clones, detectable far before clinical transformation, can include mechanisms other than the development of drug resistance (a factor potentially relevant to the increased risk of transformation observed in FL that has required multiple treatments (Al-Tourah *et al.* 2008)).

Another important factor in FL is the role of the microenvironment immune response which has a significant impact on disease outcome (Dave *et al.* 2004). A mechanism for interaction between FL cells and the microenvironment is mediated by the SHM process which introduces novel glycosylation motifs at the *IGH-V* locus in the majority of FL (Zhu *et al.* 2002; McCann *et al.* 2008). These are tumour specific and can result in oligomannosylation of surface immunoglobulin molecules on FL cells (McCann *et al.* 2008) which may bind lectins within the surrounding microenvironment (Coelho *et al.* 2010), a process which potentially supports the tumour cell. In view of this, *IGH-V* sequences in the sequential FL / t-FL sample cohort were assessed for novel glycosylation sites. These were demonstrated in the majority of cases (n = 14) with 13 / 14 being conserved in both FL and t-FL (***Appendix 6* Table 2**) and their presence in 9 CPC sequences suggests they may play an early role in disease development while differences between t-FL and FL (as observed in one case) might reflect a more specific role in transformation.

The precise origin and role of CPCs in FL / t-FL remains uncertain but the presence of specific SHM signatures suggests they are derived from germinal centre B cells (a conclusion supported by recent gene-expression data indicating the precursor to t-FL derives from a germinal centre B-cell (Gentles *et al.* 2009)). Subsequent development of tumour may occur in different contexts: firstly, as t-FL clones are detectable in FL and the CPC clone is detectable in both tumours, the CPC, FL and t-FL clones may co-exist within lymph nodes; alternatively, the CPC may reside in a pre-lymphoma (normal) follicle and give rise to FL and t-FL clones respectively which then form independent tumours (in which the CPC is, therefore, not detectable). Moreover, as in the donor / recipient pair, the CPC may migrate from its lymph node of origin to the bone marrow where it is supported until migrating to a secondary lymphoid site where it can initiate disease. The donor / recipient model also provides direct evidence that the CPC can give rise to tumour following transplantation in to a different host indicating the CPC has innate tumour initiating properties. Analogous xenograft repopulating models have been used to examine other cancer-initiating cells with their existence first demonstrated over a decade ago in acute myeloid leukaemia (Lapidot *et al.* 1994; Bonnet and Dick 1997). Since then cancer-initiating cells have been demonstrated in a variety of other cancers including both solid tumours (Al-Hajj *et al.* 2003; Singh *et al.* 2004; O'Brien *et al.* 2007; Ricci-

Vitiani *et al.* 2007; Schatton *et al.* 2008) and other haematological malignancies (Cobaleda *et al.* 2000; Bhatia *et al.* 2003; Jamieson *et al.* 2004; Matsui *et al.* 2008). Although their precise origin and nature may differ; a recurrent feature shared between different cancer initiating cells is a stem cell-like phenotype including the ability for self-renewal (Reya *et al.* 2001; Visvader 2011).

In this regard, the report by Gentles and colleagues cited above (Gentles *et al.* 2009) also demonstrates that a gene transcription signature (based on functional modules) predicts transformation and provides evidence for a stem cell-like contribution to t-FL that is compatible with both the direct and indirect routes of FL / t-FL evolution. This model suggests that the expression of an embryonic stem cell-like gene expression program in certain germinal centre FL or FL-progenitor clones may prime these cells to subsequently transform. Furthermore, the indirect model of recurrence can be applied to other B-cell malignancies with one report indicating that divergent evolution of a long-lived progenitor cell may give rise to relapse of childhood B-cell acute lymphoblastic leukaemia (Mullighan *et al.* 2008), an aggressive tumour, and so parallels the findings here in t-FL.

In summary, the demonstration of a long-lived CPC, that may undergo divergent evolution to either FL or t-FL, builds on previous models of FL by revealing heterogeneous routes (direct or indirect) of t-FL development and explains reports demonstrating acquired chromosomal changes present at diagnosis can be lost at disease recurrence (Martinez-Climent *et al.* 2003; Fitzgibbon *et al.* 2007; d'Amore *et al.* 2008; Johnson *et al.* 2008b; Eide *et al.* 2010). Such a model can also help explain the observation that the total number of SHMs in FL cases can be similar at progression when compared to prior disease episodes (Aarts *et al.* 2001; Halldorsdottir *et al.* 2008) (as demonstrated here for sequential FL and t-FL in **Table 5.3**) despite the known ongoing SHM process within FL (Zhu *et al.* 1994; Ottensmeier *et al.* 1998; Bognar *et al.* 2005), because subsequent episodes arise from the same precursor. In addition, the inclusion of additional (FL2) samples demonstrates that the conventional model of direct evolution to t-FL, through outgrowth of an FL clone, can co-exist with that of the CPC in individual cases. An example of this occurs in case 26, where t-FL evolves directly from FL2 but this and the initial FL arise by divergence from a CPC (***Appendix 6*** **Figure 2**). Furthermore, FL2 occurring after t-FL by direct evolution from the FL clone (in case 2) is

evidence that, despite eradication of t-FL, the original FL clone can persist and subsequently present clinically as relapse post transformation.

Consequently, the clinical pattern of FL as a progression of multiple relapses and transformation is not always reflected by sequential acquisition of molecular hits but instead may occur by divergent evolution from a CPC. This model provides new insights in FL / t-FL and efforts are now needed to fully characterise CPCs as these could represent attractive targets for therapy. Furthermore, an important aspect of future studies addressing the complexity of disease development will be to determine the relationship between the CPC and the long-lived t(14;18) containing atypical B cells, found in healthy individuals (Limpens *et al.* 1995; Summers *et al.* 2001; Schuler *et al.* 2009), which undergo population expansion in GCs as FL-like clones (Roulland *et al.* 2006).

Chapter 6 Discussion

Follicular lymphoma (FL) is the most common form of indolent non-Hodgkin lymphoma. Although its incidence is far less than that of the most common solid cancers (namely lung, breast, colon and prostate), the prolonged relapsing / remitting course of the disease over a median survival of around one decade confers a significant population disease burden. Recent improvements in survival are, however, encouraging (Fisher *et al.* 2005; Swenson *et al.* 2005; Liu *et al.* 2006) and these probably reflect a combination of better supportive care and advances in the therapeutic options available to clinicians, particularly the addition of rituximab immunotherapy to combination chemotherapy (as discussed in Chapter 1).

However, it is not clear whether these impact on aggressive transformation, an event which confers a worse prognosis (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008) and has a persistent risk throughout the course of FL (Al-Tourah *et al.* 2008). Furthermore, the management of t-FL is difficult as it cannot be reliably predicted and is often complicated by previous therapies, advanced age and patient co-morbidities. An important aspect of addressing this significant clinical challenge is understanding the mechanisms of disease transformation, as this should ultimately help improve both t-FL prediction and management. Consequently, the current study set out to identify novel genetic events in transformation.

The genetic aspects of FL / t-FL are complex as although t(14;18) is present in around 85% of cases (and is important as it leads to up-regulation of the anti-apoptotic BCL2 protein) it is insufficient in isolation to cause disease, is typically accompanied by other genetic changes and cases lacking this translocation may even form a distinct clinico-pathological subtype of this disease (Katzenberger *et al.* 2009). Moreover, reports identifying familial FL cases (Marco *et al.* 1999; Last *et al.* 2000) and increased risk of FL in first degree relatives of cases (Goldin *et al.* 2009) suggest that inherited factors have a role. This has been confirmed in numerous studies assessing germline predisposition to FL development most notably in the first genome-wide association study (GWAS) in FL (Skibola *et al.* 2009; Conde *et al.* 2010) but few reports have addressed the influence of germline genetic variation on

FL outcome (Weng and Levy 2003; Jardin *et al.* 2005; Cerhan *et al.* 2007b; Hohaus *et al.* 2007; Racila *et al.* 2008; Han *et al.* 2010).

Importantly, the current study addressed this for two SNPs (rs10484561 and rs6457327) and confirmed their association with FL risk, with rs6457327 also predicting both risk of and time to transformation (as indicated in Chapter 3). This is an important finding as the effect was independent of other (clinical) variables which are known to associate with transformation. The influence of rs6457327 genotype was investigated by examining the expression of *C6orf15*, the only coding region in high linkage disequilibrium with this SNP. However, this was not a promising candidate as the transcript was not detected in any of a series of FL and t-FL samples and *C6orf15* was not mutated in FL at diagnosis. As the protein product of the murine homologue of *C6orf15*, termed *emprin*, has recently been identified as a component of the extra-cellular matrix (ECM) (Manabe *et al.* 2008) and *C6orf15* is expressed in skin (Oka *et al.* 1999; Sanchez *et al.* 2004), this raises the possibility that the expression detected in tonsils might be a consequence of either the extensive connective or epithelial tissue contained within their structure (as opposed to expression by infiltrating lymphocytes). This might then help explain the lack of expression detected in other lymphoid material included in the current study.

As rs6457327 is found within the HLA region and immune response profiles associate with FL outcome (Dave *et al.* 2004), germline variation at this SNP might influence the immune response to FL. Furthermore, as part of an ongoing departmental study in to the role of the micro-environment in FL, 56 of the study cases have already been included in FL tissue micro-arrays. Immuno-histochemical staining of the tumour samples has been performed for CD4, CD8, FOXP3 and CD68 / CD163 representing infiltrating T_H1, T_H2, T_{reg} lymphocytes and macrophages, respectively. Interestingly, the number of FOXP3 positive cells with an intra-follicular location was significantly increased ($P = .02$; two-tailed unpaired t test) in FL tumour samples from the rs6457327 AA+AC cases (n = 32) compared to the rs6457327 CC cases (n = 21), with 3/56 cases lacking data. This follows a report (Farinha *et al.* 2010) in which a “follicular” distribution (characterized by an intrafollicular and perifollicular predominance) of FOXP3 positive cells in FL biopsy samples, as opposed to a "diffuse" pattern (in which the cells were diffusely distributed with no clear relationship to the follicles) was associated with

transformation risk in a multivariate analysis (P = .004, HR = 3.9). Consequently, investigation of rs6457327 has now been incorporated in to the ongoing micro-environment study in this department.

It might be expected that SNPs which confer increased cancer risk also confer a worse clinical outcome and this has been demonstrated in other cancers (Grochola *et al.* 2009; Cizmarikova *et al.* 2010). However, in the current study, the minor (A) allele of rs6457327 associates with reduced risk of developing FL but in established cases confers an increase in risk of transformation (a deleterious clinical outcome). This is reminiscent of a recent study in colorectal cancer which demonstrates that the variant (G) allele of SNP rs1799977 in *MLH1* confers increased risk of this disease but associates with better clinical outcome (lower risk of vascular invasion, distant metastases and recurrence) (Nejda *et al.* 2009). Potential explanations for these scenarios include i) differing alleles of the same SNP being contained within different risk and outcome inherited genetic haplotypes or ii) that risk and outcome determinants occur together within a haplotype containing one SNP allele but the influence of associated determinants is exerted at different times.

In addition to marking risk alleles, SNPs may have direct functional consequences (Thomas *et al.* 1999; Bond *et al.* 2004). Therefore, it is noteworthy that the Encyclopedia of DNA elements consortium's website, http://genome.ucsc.edu/cgi-bin/hgGateway?org=human ("The ENCODE Project" 2004; Rosenbloom *et al.* 2010) reveals both rs10484561 and rs6457327 are located near or within putative regulatory regions. A genomic site particularly sensitive to DNase I activity (a feature associated with regulatory regions) is found around 3kb telomeric to rs10484561 whereas rs6457327 is found within a potential DNase I hypersensitivity region. The latter is associated with modification of H3K4Me1, an enhancer and promoter associated histone mark, providing further evidence that this region has a regulatory function. rs6457327 is also located less than 5kb centromeric from a potential enhancer region (as well as being less than 5kb telomeric from the *C6orf15* coding locus). Consequently, both SNPs might be implicated in control of gene transcription through modulation of these putative regulatory regions.

The prolonged clinical follow-up of the cases within the germline investigation is an important aspect of this study. From the cohort (n = 130) used in determining clinical

outcome associations for rs6457327, the 67 surviving patients were followed up for a median of 9.3 years (range 1.6 – 30.5) with 94% seen for >5 years and 48% for >10 years and such prolonged follow-up is important as FL typically has a long disease course. This provides robust clinical outcome data with a transformation frequency of 35% (45 / 130 cases) similar to that reported in large series of biopsy proven and clinically determined transformation (Al-Tourah *et al.* 2008) or biopsy proven transformation (Montoto *et al.* 2007b). Furthermore, it enabled identification of two cases in which transformation presented in the third decade after FL diagnosis (**Figure 3.3**), supporting the ongoing risk of transformation demonstrated in a recent report (Al-Tourah *et al.* 2008) and extends it to over 20 years.

The significant relationship between transformation and overall survival reported for the two recent large FL series above (Montoto *et al.* 2007b; Al-Tourah *et al.* 2008) was maintained in this cohort, $P = .0004$ (Wrench *et al.* 2011). However, the rs6457327 genotype association with transformation did not translate to prediction of survival, $P = .97$ (**Table 3.8**). This may reflect the contribution of several different factors to the transformation and survival of FL as, firstly, this SNP is one of a large number of heterogeneous factors linked with transformation (Bernstein and Burack 2009) and, secondly, there is emerging data (Davies *et al.* 2007; Gentles *et al.* 2009) suggesting there are potentially different early and late mechanisms contributing to transformation. It may be that the degree of influence for each individual factor towards transformation or survival differs, and this could account for the seemingly discrepant rs6457327 result (as well as that for a marker previously reported to predict transformation but not OS (Jardin *et al.* 2005)).

Two well described SNPs influencing the MDM2-p53 axis, *MDM2* SNP 309 and *TP53* Arg72Pro, were also examined but they showed no associations with either clinical outcome or risk of FL. Although such correlations do occur for both SNPs in other cancers, these effects are often not consistent and at times are contradictory (Wilkening *et al.* 2007; Gryshchenko *et al.* 2008; Zenz *et al.* 2008; Chua *et al.* 2010).

However, these SNPs might have future relevance in FL by affecting responses to novel therapies such as the nutlins, small molecule inhibitors which disrupt MDM2 binding to p53 (Vassilev *et al.* 2004). These might relieve p53 inhibition in cancer where the MDM2-p53 axis is disrupted thereby limiting malignant cell proliferation.

Significantly, these compounds show activity in a range of B-NHL including MCL (Tabe *et al.* 2009), primary effusion lymphoma (Sarek *et al.* 2007) and Burkitt lymphoma (Renouf *et al.* 2009) as well as CLL (Coll-Mulet *et al.* 2006; Kojima *et al.* 2006). A structurally similar molecule, MI-319, has recently shown activity against a t(14;18) containing NHL cell line *in vitro* and in murine models (Mohammad *et al.* 2009). Should experimental molecules such as these become part of the treatment armament for FL in future, disease responses may be affected by polymorphic variation affecting *MDM2* and its expression in FL cells, such as that at *MDM2* SNP 309.

The germline investigations followed a targeted approach in which the SNPs examined were selected based on their known characteristics or disease associations. However, new methods are opening up the entire genome to simultaneous investigation. In particular, the development of genome-wide micro-array based techniques represents a significant advance in the ability to examine genetic variation in both normal and disease states. They provide high resolution combined with the ability to examine multiple samples in parallel. In studies of the human genome they are revealing novel information on germline genomic variants including SNPs (McCarroll *et al.* 2008) and copy number polymorphisms (Iafrate *et al.* 2004; Redon *et al.* 2006; McCarroll *et al.* 2008). Furthermore, SNPs can be used as genomic tags in large GWAS, which have assessed how such variation may predispose to disease such as cancer, including FL (Skibola *et al.* 2009; Conde *et al.* 2010).

The most recent tools can determine copy numbers and genotypes for both SNP and non-SNP loci providing information on CNAs and LOH over a large number of genomic loci at a high resolution: every ~700 bp for the array used in this study. Consequently, regions of genomic change that are beyond the resolution of earlier techniques such as conventional karyotyping, CGH, and array-CGH, could be detected down to around 7 kb in size (using segments with at least 10 consecutive probes) which was evident in detection of CN loss affecting *CREBBP* (150 kb) and *CDC2L2* (60kb) or gain affecting *SMARCD3* (16 kb), as detailed in Chapter 5. During the course of the study both array-CGH and SNP-array platforms with even higher resolution have been developed and these represent important developments that should help identify the critical genetic events in cancers.

A particular important advance has been the development of genome-wide sequencing. This has enabled the identification of novel gene mutations in whole exome sequencing of malignancies including AML (Ley *et al.* 2008; Mardis *et al.* 2009) and subsequently solid cancers including melanoma and lung cancer (Lee *et al.* 2010). The application of this technique to FL samples is similarly revealing novel mutations which include those reported in *EZH2* (Morin *et al.* 2010a) and recently described events affecting *MLL2*, *MEF2B* and *CREBBP* (Mendez-Lago *et al.* 2010; Morin *et al.* 2010b; Pasqualucci *et al.* 2011) as discussed in Chapter 4. This technique can also identify a variety of other genomic aberrations including breakpoints, micro-insertions and deletions, cryptic translocations and inversions as well as disease subclones and so it represents a powerful tool for the discovery of driver events as well as determinants of treatment resistance and disease progression. Moreover, as this technique can now establish the human genome sequence in a matter of hours it may ultimately enable the production of personalised cancer genetic signatures that could be applied as a routine clinical adjunct to enhance disease management by facilitating individual, tailored treatment regimes.

The identification by such sequencing of *EZH2* mutations in FL (Morin *et al.* 2010a) which, as suggested in Chapter 4, may be enriched in t-FL is of particular interest as the EZH2 protein product forms part of the polycomb repressor complex 2 (PRC2). This complex also includes the SUZ12, EED and RbAp46 proteins and functions as a histone methylase facilitating the trimethylation of histone H3 Lysine 27 (H3K27me3) (Cao *et al.* 2002; Kuzmichev *et al.* 2002; Kirmizis *et al.* 2004). The *EZH2* mutations are now known to confer gain of function specific for the di- to trimethylation step (Sneeringer *et al.* 2010; Yap *et al.* 2011) which is important as H3K27me3 creates a transcriptionally repressive histone mark and is associated with gene methylation in cancer (Schlesinger *et al.* 2007). Moreover, recent studies, performed by our group for FL / t-FL (O'Riain *et al.* 2009) and elsewhere for DLBCL (Martin-Subero *et al.* 2009b), have identified that the gene set targeted for transcriptional repression by methylation is enriched for genes previously shown to be targets of PRC2 mediated repression in embryonic stem cells (Lee *et al.* 2006b). This suggests that a stem cell profile might be important in FL development (and could be influenced by somatically acquired mutations such as those in *EZH2*). It is therefore significant that a gene transcription signature, based on functional modules

and which predicts transformation of FL, also provides evidence for a stem cell-like origin to t-FL (Gentles *et al.* 2009). Overall, these data provide a tentative model in which disruption of PRC2 mediated chromatin remodelling, that targets genes specifically regulated in stem cells, may represent an early event in FL.

Although cancer initiating cells have now been described in several malignancies, including a lymphoid tumour (B-ALL) (Cox *et al.* 2004; Notta *et al.* 2011), these are yet to be described in FL. However, the data from chapters 4 and 5 indicates that FL / t-FL may arise by divergent evolution of a common progenitor cell (CPC) and supports earlier (Matolcsy *et al.* 1999; Martinez-Climent *et al.* 2003; Fitzgibbon *et al.* 2007) and more recent (Ruminy *et al.* 2008; Eide *et al.* 2010) reports detailing non-sequential patterns of clonal evolution in sequential FL samples. As the CPC has undergone somatic hyper-mutation of the *IGH-V* gene this suggests that, like its progeny, it arises from the germinal centre, a concept also proposed in the report by Gentles and colleagues (Gentles *et al.* 2009). Furthermore, the evolution of FL and then t-FL from a CPC by non-sequential, indirect evolution many months or even years apart indicates the attributes of CPCs include longevity, chemotherapy resistance and the ability to generate distinct episodes of divergent disease (either FL or DLBCL / t-FL). Indeed, this may help explain the related clinical scenarios of discordant, composite and downgrade lymphoma as well as the relapse of FL after t-FL. Importantly, the CPC attributes are reflected in normal haematopoietic stem cells which can give rise to differing cell types, can repopulate the bone marrow after high-dose chemotherapy and, through self-renewal, are long-lived. Consequently, determining whether stem cell-like molecular profiles contribute to the CPC, as recently proposed by Martinez-Climent and colleagues (Martinez-Climent *et al.* 2010), could provide valuable insights in to the processes that drive evolution of FL / t-FL.

The relative importance of inherited genomic variation, host micro-environment immune response and tumour genetic or epigenetic events to FL is unclear. However, molecular outcome predictors have been derived from acquired tumour genetic events (O'Shea *et al.* 2008; Cheung *et al.* 2009b; O'Shea *et al.* 2009; Cheung *et al.* 2010b) and host microenvironment immune profiles (Dave *et al.* 2004). Indeed, correlation between these processes has been reported (O'Shea *et al.* 2008) and this supports a dynamic interplay between tumour cells and the microenvironment in FL.

In this regard, the mutations of *TNFRSF14* examined in Chapter 4 are particularly relevant as its gene product has a role in signalling between B cells and other immune cells (Cai and Freeman 2009) and mutations are known to disrupt its ligand binding (Compaan *et al.* 2005). Similarly, gain affecting the *IL4-R* locus in t-FL (detailed in the same chapter) provides another potential bridge between tumour genetics and the microenvironment and follows studies in FL showing over-expression of IL4-R (Calvo *et al.* 2008) and IL4 (Husson *et al.* 2002) in tumour and immune cells, respectively. Furthermore, SHMs that create novel glycosylation motifs in the Ig genes of FL cells lead to Ig molecules on the cell surface that bind lectins in the surrounding microenvironment (Coelho *et al.* 2010), and these are interactions which might support tumour cells. Consequently these, and the association of rs6457327 genotype with FOXP3 positive tumour T cells detailed in chapter 3, represent good candidates for further study as the pathways underlying genetic-microenvironment correlations in FL may be targets for therapeutic disruption.

This study has identified germline and acquired genetic disease outcome predictors with regard to FL and its transformation. Together with the recently described acquired predictors (O'Shea *et al.* 2008; Cheung *et al.* 2009b; O'Shea *et al.* 2009; Cheung *et al.* 2010b), these represent a growing pool of molecular markers in FL. Moreover, although the use of micro-array profiling has provided valuable insights in to FL and t-FL (Lossos *et al.* 2002; Dave *et al.* 2004; Glas *et al.* 2005; Davies *et al.* 2007; Glas *et al.* 2007), including prognostic signatures (Dave *et al.* 2004), it requires complex research techniques currently beyond the remit of most clinical centres. Nevertheless, progress has been made through the validation of microarray derived RNA expression profiles using RT-PCR (Sakhinia *et al.* 2007; Byers *et al.* 2008) and this represents a step towards routine application of molecular profiling. However, to achieve these aims, such research tools should be adapted into approaches amenable for use on formalin fixed paraffin embedded material, the current standard preparation of tissue biopsies. Moreover, integration of both germline and acquired molecular determinants in such profiling should be considered in view of the robust correlation shown here between SNP rs6453727 and t-FL.

In a similar way to the FLIPI clinical prognostic tools (Solal-Celigny *et al.* 2004; Federico *et al.* 2009), molecular markers are yet to routinely influence patient

management and determining whether this can be achieved requires their validation in prospective clinical studies. This would pave the way for assembly of clinical-molecular prognostic algorithms as well as predictors of treatment response and transformation that might aid rationalised, targeted therapy; a credible goal given the plethora of available and emerging treatments and the grave outcome of transformation.

The recent therapeutic advances and improvements in outcome provide important developments in FL. Whether these herald a change in the natural history of transformation, which remains a persistent clinical challenge more than eighty years since the first description of FL (Brill *et al.* 1925; Symmers 1927) is unclear. However, the demonstration here that a CPC can give rise to distinct disease episodes by divergent clonal evolution indicates that the clinical pattern of FL, as a progression of multiple relapses and transformation, is not always reflected by sequential acquisition of molecular hits. Furthermore, this model augurs well for the formal identification of an FL progenitor cell and the subsequent eradication of a lymphoma initiating cell pool, enabling better treatment of this disease.

Bibliography

Aarts, W. M., R. J. Bende, J. G. Bossenbroek, S. T. Pals and C. J. van Noesel (2001). "Variable heavy-chain gene analysis of follicular lymphomas: subclone selection rather than clonal evolution over time." Blood **98**(1): 238-40.

Aarts, W. M., R. J. Bende, E. J. Steenbergen, P. M. Kluin, E. C. Ooms, S. T. Pals and C. J. van Noesel (2000). "Variable heavy chain gene analysis of follicular lymphomas: correlation between heavy chain isotype expression and somatic mutation load." Blood **95**(9): 2922-9.

Agopian, J., J. M. Navarro, A. C. Gac, Y. Lecluse, M. Briand, P. Grenot, P. Gauduchon, P. Ruminy, P. Lebailly, B. Nadel and S. Roulland (2009). "Agricultural pesticide exposure and the molecular connection to lymphomagenesis." J Exp Med **206**(7): 1473-83.

Al-Hajj, M., M. S. Wicha, A. Benito-Hernandez, S. J. Morrison and M. F. Clarke (2003). "Prospective identification of tumorigenic breast cancer cells." Proc Natl Acad Sci U S A **100**(7): 3983-8.

Al-Tourah, A. J., K. K. Gill, M. Chhanabhai, P. J. Hoskins, R. J. Klasa, K. J. Savage, L. H. Sehn, T. N. Shenkier, R. D. Gascoyne and J. M. Connors (2008). "Population-based analysis of incidence and outcome of transformed non-Hodgkin's lymphoma." J Clin Oncol **26**(32): 5165-9.

Albarino, C. G. and V. Romanowski (1994). "Phenol extraction revisited: a rapid method for the isolation and preservation of human genomic DNA from whole blood." Mol Cell Probes **8**(5): 423-7.

Alexander, D. D., P. J. Mink, H. O. Adami, E. T. Chang, P. Cole, J. S. Mandel and D. Trichopoulos (2007). "The non-Hodgkin lymphomas: a review of the epidemiologic literature." Int J Cancer **120 Suppl 12**: 1-39.

Allen, C. D., T. Okada and J. G. Cyster (2007). "Germinal-center organization and cellular dynamics." Immunity **27**(2): 190-202.

Alnemri, E. S., N. M. Robertson, T. F. Fernandes, C. M. Croce and G. Litwack (1992). "Overexpressed full-length human BCL2 extends the survival of baculovirus-infected Sf9 insect cells." Proc Natl Acad Sci U S A **89**(16): 7295-9.

Alt, J. R., T. C. Greiner, J. L. Cleveland and C. M. Eischen (2003). "Mdm2 haplo-insufficiency profoundly inhibits Myc-induced lymphomagenesis." Embo J **22**(6): 1442-50.

Altekruse, S. F., Kosary C. L., Krapcho M., Neyman N., Aminou R., Waldron W., Ruhl J., Howlader N., Tatalovich Z., Cho H., Mariotto A., Eisner M. P., Lewis D. R., Cronin K., Chen H. S., Feuer E. J., Stinchcomb D. G. and Edwards B. K. (2009). SEER Cancer Statistics Review, 1975-2007. National Cancer Institute, Bethesda.

Altieri, A., J. L. Bermejo and K. Hemminki (2005). "Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: the Swedish Family-Cancer Database." Blood **106**(2): 668-72.

Alvaro, T., M. Lejeune, M. T. Salvado, C. Lopez, J. Jaen, R. Bosch and L. E. Pons (2006). "Immunohistochemical patterns of reactive microenvironment are associated with clinicobiologic behavior in follicular lymphoma patients." J Clin Oncol **24**(34): 5350-7.

Anderson, J. R., J. O. Armitage and D. D. Weisenburger (1998). "Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project." Ann Oncol **9**(7): 717-20.

Antonarakis, S. E., M. Krawczak and D. N. Cooper (2002). Part 1 Basic Concepts in Cancer Genetics: The Nature and Mechanisms of Human Gene Mutation. The Genetic Basis of Human Cancer. B. Vogelstein and K. W. Kinzler. New York, McGraw Hill

Arakawa, H., J. Hauschild and J. M. Buerstedde (2002). "Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion." Science **295**(5558): 1301-6.

Ardeshna, K. M., Qian W., Smith P., Warden J., Stevens L., Pocock C. F. E., Miall F., Cunningham D., Davies J., Walewski J., Burhan Ferhanoglu A., Bradstock K. and Linch D. C. (2010). "An Intergroup Randomised Trial of Rituximab Versus a Watch and Wait Strategy In Patients with Stage II, III, IV, Asymptomatic, Non-Bulky Follicular Lymphoma (Grades 1, 2 and 3a). A Preliminary Analysis." Blood (ASH Annual Meeting Abstracts), Nov 2010; 116: #6.

Ardeshna, K. M., P. Smith, A. Norton, B. W. Hancock, P. J. Hoskin, K. A. MacLennan, R. E. Marcus, A. Jelliffe, G. Vaughan, Hudson and D. C. Linch (2003). "Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial." Lancet **362**(9383): 516-22.

Arva, N. C., T. R. Gopen, K. E. Talbott, L. E. Campbell, A. Chicas, D. E. White, G. L. Bond, A. J. Levine and J. Bargonetti (2005). "A chromatin-associated and transcriptionally inactive p53-Mdm2 complex occurs in mdm2 SNP309 homozygous cells." J Biol Chem **280**(29): 26776-87.

Avantaggiati, M. L., V. Ogryzko, K. Gardner, A. Giordano, A. S. Levine and K. Kelly (1997). "Recruitment of p300/CBP in p53-dependent signal pathways." Cell **89**(7): 1175-84.

Avet-Loiseau, H., M. Vigier, A. Moreau, M. P. Mellerin, F. Gaillard, J. L. Harousseau, R. Bataille and N. Milpied (1997). "Comparative genomic hybridization detects genomic abnormalities in 80% of follicular lymphomas." Br J Haematol **97**(1): 119-22.

Baldo, P., M. Rupolo, A. Compagnoni, R. Lazzarini, A. Bearz, R. Cannizzaro, S. Spazzapan, I. Truccolo and L. Moja (2010). "Interferon-alpha for maintenance of follicular lymphoma." Cochrane Database Syst Rev(1): CD004629.

Banin, S., L. Moyal, S. Shieh, Y. Taya, C. W. Anderson, L. Chessa, N. I. Smorodinsky, C. Prives, Y. Reiss, Y. Shiloh and Y. Ziv (1998). "Enhanced phosphorylation of p53 by ATM in response to DNA damage." Science **281**(5383): 1674-7.

Barak, Y., T. Juven, R. Haffner and M. Oren (1993). "mdm2 expression is induced by wild type p53 activity." Embo J **12**(2): 461-8.

Barrans, S., S. Crouch, A. Smith, K. Turner, R. Owen, R. Patmore, E. Roman and A. Jack (2010). "Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab." J Clin Oncol **28**(20): 3360-5.

Bartlett, N. L., M. Rizeq, R. F. Dorfman, J. Halpern and S. J. Horning (1994). "Follicular large-cell lymphoma: intermediate or low grade?" J Clin Oncol **12**(7): 1349-57.

Bastion, Y., C. Sebban, F. Berger, P. Felman, G. Salles, C. Dumontet, P. A. Bryon and B. Coiffier (1997). "Incidence, predictive factors, and outcome of lymphoma transformation in follicular lymphoma patients." J Clin Oncol **15**(4): 1587-94.

Bentz, M., C. A. Werner, H. Dohner, S. Joos, T. F. Barth, R. Siebert, M. Schroder, S. Stilgenbauer, K. Fischer, P. Moller and P. Lichter (1996). "High incidence of chromosomal imbalances and gene amplifications in the classical follicular variant of follicle center lymphoma." Blood **88**(4): 1437-44.

Berek, C., A. Berger and M. Apel (1991). "Maturation of the immune response in germinal centers." Cell **67**(6): 1121-9.

Berglund, M., G. Enblad, U. Thunberg, R. M. Amini, C. Sundstrom, G. Roos, M. Erlanson, R. Rosenquist, C. Larsson and S. Lagercrantz (2007). "Genomic imbalances during transformation from follicular lymphoma to diffuse large B-cell lymphoma." Mod Pathol **20**(1): 63-75.

Bergman, Y. and H. Cedar (2004). "A stepwise epigenetic process controls immunoglobulin allelic exclusion." Nat Rev Immunol **4**(10): 753-61.

Bernell, P., B. Jacobsson, J. Liliemark, V. Hjalmar, I. Arvidsson and R. Hast (1998). "Gain of chromosome 7 marks the progression from indolent to aggressive follicle centre lymphoma and is a common finding in patients with diffuse large B-cell lymphoma: a study by FISH." Br J Haematol **101**(3): 487-91.

Bernstein, S. H. and W. R. Burack (2009). "The incidence, natural history, biology, and treatment of transformed lymphomas." Hematology Am Soc Hematol Educ Program: 532-41.

Bhatia, R., M. Holtz, N. Niu, R. Gray, D. S. Snyder, C. L. Sawyers, D. A. Arber, M. L. Slovak and S. J. Forman (2003). "Persistence of malignant hematopoietic progenitors in chronic myelogenous leukemia patients in complete cytogenetic remission following imatinib mesylate treatment." Blood **101**(12): 4701-7.

Bianchi, F., M. Raponi, F. Piva, A. Viel, I. Bearzi, E. Galizia, R. Bracci, L. Belvederesi, C. Loretelli, C. Brugiati, F. Corradini, D. Baralle and R. Cellerino (2011). "An intronic mutation in MLH1 associated with familial colon and breast cancer." Fam Cancer **10**(1): 27-35.

Bloomfield, C. D., D. C. Arthur, G. Frizzera, E. G. Levine, B. A. Peterson and K. J. Gajl-Peczalska (1983). "Nonrandom chromosome abnormalities in lymphoma." Cancer Res **43**(6): 2975-84.

Bognar, A., B. Csernus, C. Bodor, L. Reiniger, A. Szepesi, E. Toth, L. Kopper and A. Matolcsy (2005). "Clonal selection in the bone marrow involvement of follicular lymphoma." Leukemia **19**(9): 1656-62.

Bond, G. L., K. M. Hirshfield, T. Kirchhoff, G. Alexe, E. E. Bond, H. Robins, F. Bartel, H. Taubert, P. Wuerl, W. Hait, D. Toppmeyer, K. Offit and A. J. Levine (2006). "MDM2 SNP309 accelerates tumor formation in a gender-specific and hormone-dependent manner." Cancer Res **66**(10): 5104-10.

Bond, G. L., W. Hu, E. E. Bond, H. Robins, S. G. Lutzker, N. C. Arva, J. Bargonetti, F. Bartel, H. Taubert, P. Wuerl, K. Onel, L. Yip, S. J. Hwang, L. C. Strong, G. Lozano and A. J. Levine (2004). "A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans." Cell **119**(5): 591-602.

Bond, G. L., W. Hu and A. Levine (2005). "A single nucleotide polymorphism in the MDM2 gene: from a molecular and cellular explanation to clinical effect." Cancer Res **65**(13): 5481-4.

Bonnet, D. and J. E. Dick (1997). "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." Nat Med **3**(7): 730-7.

Boonstra, R., A. Bosga-Bouwer, M. Mastik, E. Haralambieva, J. Conradie, E. van den Berg, A. van den Berg and S. Poppema (2003). "Identification of chromosomal copy number changes associated with transformation of follicular lymphoma to diffuse large B-cell lymphoma." Hum Pathol **34**(9): 915-23.

Borrow, J., V. P. Stanton, Jr., J. M. Andresen, R. Becher, F. G. Behm, R. S. Chaganti, C. I. Civin, C. Disteche, I. Dube, A. M. Frischauf, D. Horsman, F. Mitelman, S. Volinia, A. E. Watmore and D. E. Housman (1996). "The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein." Nat Genet **14**(1): 33-41.

Bosga-Bouwer, A. G., G. W. van Imhoff, R. Boonstra, A. van der Veen, E. Haralambieva, A. van den Berg, B. de Jong, V. Krause, M. C. Palmer, R. Coupland, P. M. Kluin, E. van den Berg and S. Poppema (2003). "Follicular lymphoma grade 3B includes 3 cytogenetically defined subgroups with primary t(14;18), 3q27, or other translocations: t(14;18) and 3q27 are mutually exclusive." Blood **101**(3): 1149-54.

Bougeard, G., S. Baert-Desurmont, I. Tournier, S. Vasseur, C. Martin, L. Brugieres, A. Chompret, B. Bressac-de Paillerets, D. Stoppa-Lyonnet, C. Bonaiti-Pellie and T. Frebourg (2006). "Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome." J Med Genet **43**(6): 531-3.

Bracci, P. M., C. F. Skibola, L. Conde, E. Halperin, T. Lightfoot, A. Smith, R. A. Paynter, D. R. Skibola, L. Agana, E. Roman, E. Kane and J. K. Wiencke (2010). "Chemokine polymorphisms and lymphoma: a pooled analysis." Leuk Lymphoma **51**(3): 497-506.

Brack, C., M. Hirama, R. Lenhard-Schuller and S. Tonegawa (1978). "A complete immunoglobulin gene is created by somatic recombination." Cell **15**(1): 1-14.

Breit, T. M., I. L. Wolvers-Tettero and J. J. van Dongen (1994). "Unique selection determinant in polyclonal V delta 2-J delta 1 junctional regions of human peripheral gamma delta T lymphocytes." J Immunol **152**(6): 2860-4.

Brice, P., Y. Bastion, E. Lepage, N. Brousse, C. Haioun, P. Moreau, N. Straetmans, H. Tilly, I. Tabah and P. Solal-Celigny (1997). "Comparison in low-tumor-burden follicular lymphomas between an initial no-treatment policy, prednimustine, or interferon alfa: a randomized study from the Groupe d'Etude des Lymphomes Folliculaires. Groupe d'Etude des Lymphomes de l'Adulte." J Clin Oncol **15**(3): 1110-7.

Brill, N. E., G. Baehr and N. Rosenthal (1925). "Generalized giant follicle hyperplasia of lymph nodes and spleen. A hitherto undescribed type." JAMA **84**: 668-671.

Brochet, X., M. P. Lefranc and V. Giudicelli (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**(Web Server issue): W503-8.

Brose, M. S., P. Volpe, K. Paul, J. E. Stopfer, T. A. Colligon, K. A. Calzone and B. L. Weber (2004). "Characterization of two novel BRCA1 germ-line mutations involving splice donor sites." Genet Test **8**(2): 133-8.

Buchonnet, G., F. Jardin, N. Jean, P. Bertrand, F. Parmentier, S. Tison, S. Lepretre, N. Contentin, P. Lenain, A. Stamatoullas-Bastard, H. Tilly and C. Bastard (2002). "Distribution of BCL2 breakpoints in follicular lymphoma and correlation with clinical features: specific subtypes or same disease?" Leukemia **16**(9): 1852-6.

Buchonnet, G., P. Lenain, P. Ruminy, S. Lepretre, A. Stamatoullas, F. Parmentier, F. Jardin, C. Duval, H. Tilly and C. Bastard (2000). "Characterisation of BCL2-JH rearrangements in follicular lymphoma: PCR detection of 3' BCL2 breakpoints and evidence of a new cluster." Leukemia **14**(9): 1563-9.

Bueso-Ramos, C. E., Y. Yang, E. deLeon, P. McCown, S. A. Stass and M. Albitar (1993). "The human MDM-2 oncogene is overexpressed in leukemias." Blood **82**(9): 2617-23.

Burger, J. A., P. Ghia, A. Rosenwald and F. Caligaris-Cappio (2009). "The microenvironment in mature B-cell malignancies: a target for new treatment strategies." Blood **114**(16): 3367-75.

Butler, N. R. (1961). "National survey of perinatal mortality: first results." British Medical Journal **1**(5235): 1313 - 1315.

Byers, R. J., E. Sakhinia, P. Joseph, C. Glennie, J. A. Hoyland, L. P. Menasce, J. A. Radford and T. Illidge (2008). "Clinical quantitation of immune signature in follicular lymphoma by RT-PCR-based gene expression profiling." Blood **111**(9): 4764-70.

Cai, G. and G. J. Freeman (2009). "The CD160, BTLA, LIGHT/HVEM pathway: a bidirectional switch regulating T-cell activation." Immunol Rev **229**(1): 244-58.

Cai, M. M., S. W. Zhang, S. Zhang, S. Chen, J. Yan, X. Y. Zhu, Y. Hu, C. Chen and J. X. Gu (2002). "Different effects of p58PITSLRE on the apoptosis induced by etoposide, cycloheximide and serum-withdrawal in human hepatocarcinoma cells." Mol Cell Biochem **238**(1-2): 49-55.

Calvo, K. R., B. Dabir, A. Kovach, C. Devor, R. Bandle, A. Bond, J. H. Shih and E. S. Jaffe (2008). "IL-4 protein expression and basal activation of Erk in vivo in follicular lymphoma." Blood **112**(9): 3818-26.

Camacho, F. I., P. Algara, A. Rodriguez, E. Ruiz-Ballesteros, M. Mollejo, N. Martinez, J. A. Martinez-Climent, M. Gonzalez, M. Mateo, A. Caleo, M. Sanchez-Beato, J. Menarguez, J. Garcia-Conde, F. Sole, E. Campo and M. A. Piris (2003). "Molecular heterogeneity in MCL defined by the use of specific VH genes and the frequency of somatic mutations." Blood **101**(10): 4042-6.

Canman, C. E., D. S. Lim, K. A. Cimprich, Y. Taya, K. Tamai, K. Sakaguchi, E. Appella, M. B. Kastan and J. D. Siliciano (1998). "Activation of the ATM kinase by ionizing radiation and phosphorylation of p53." Science **281**(5383): 1677-9.

Cao, R., L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones and Y. Zhang (2002). "Role of histone H3 lysine 27 methylation in Polycomb-group silencing." Science **298**(5595): 1039-43.

Carreras, J., A. Lopez-Guillermo, G. Roncador, N. Villamor, L. Colomo, A. Martinez, R. Hamoudi, W. J. Howat, E. Montserrat and E. Campo (2009). "High numbers of tumor-infiltrating programmed cell death 1-positive

regulatory lymphocytes are associated with improved overall survival in follicular lymphoma." J Clin Oncol **27**(9): 1470-6.

Casali, P. and H. Zan (2004). "Class switching and Myc translocation: how does DNA break?" Nat Immunol **5**(11): 1101-3.

Casey, R., P. Brennan, N. Becker, P. Boffetta, P. Cocco, E. Domingo-Domenech, L. Foretova, A. Nieters, S. de Sanjose, A. Staines, M. Vornanen and M. Maynadie (2006). "Influence of familial cancer history on lymphoid neoplasms risk validated in the large European case-control study epilymph." Eur J Cancer **42**(15): 2570-6.

Castillo, J. J., S. Dalia and S. K. Pascual (2010). "Association between red blood cell transfusions and development of non-Hodgkin lymphoma: a meta-analysis of observational studies." Blood **116**(16): 2897-907.

Cattoretti, G., C. C. Chang, K. Cechova, J. Zhang, B. H. Ye, B. Falini, D. C. Louie, K. Offit, R. S. Chaganti and R. Dalla-Favera (1995). "BCL-6 protein is expressed in germinal-center B cells." Blood **86**(1): 45-53.

Cerhan, J. R., S. M. Ansell, Z. S. Fredericksen, N. E. Kay, M. Liebow, T. G. Call, A. Dogan, J. M. Cunningham, A. H. Wang, W. Liu-Mares, W. R. Macon, D. Jelinek, T. E. Witzig, T. M. Habermann and S. L. Slager (2007a). "Genetic variation in 1253 immune and inflammation genes and risk of non-Hodgkin lymphoma." Blood **110**(13): 4455-63.

Cerhan, J. R., W. Liu-Mares, Z. S. Fredericksen, A. J. Novak, J. M. Cunningham, N. E. Kay, A. Dogan, M. Liebow, A. H. Wang, T. G. Call, T. M. Habermann, S. M. Ansell and S. L. Slager (2008). "Genetic variation in tumor necrosis factor and the nuclear factor-kappaB canonical pathway and risk of non-Hodgkin's lymphoma." Cancer Epidemiol Biomarkers Prev **17**(11): 3161-9.

Cerhan, J. R., C. M. Vachon, T. M. Habermann, S. M. Ansell, T. E. Witzig, P. J. Kurtin, C. A. Janney, W. Zheng, J. D. Potter, T. A. Sellers and A. R. Folsom (2002). "Hormone replacement therapy and risk of non-hodgkin lymphoma and chronic lymphocytic leukemia." Cancer Epidemiol Biomarkers Prev **11**(11): 1466-71.

Cerhan, J. R., S. Wang, M. J. Maurer, S. M. Ansell, S. M. Geyer, W. Cozen, L. M. Morton, S. Davis, R. K. Severson, N. Rothman, C. F. Lynch, S. Wacholder, S. J. Chanock, T. M. Habermann and P. Hartge (2007b). "Prognostic significance of host immune gene polymorphisms in follicular lymphoma survival." Blood **109**(12): 5439-46.

Cescon, D. W., P. A. Bradbury, K. Asomaning, J. Hopkins, R. Zhai, W. Zhou, Z. Wang, M. Kulke, L. Su, C. Ma, W. Xu, A. L. Marshall, R. S. Heist, J. C. Wain, T. J. Lynch, Jr., D. C. Christiani and G. Liu (2009). "p53 Arg72Pro and MDM2 T309G polymorphisms, histology, and esophageal cancer prognosis." Clin Cancer Res **15**(9): 3103-9.

Chandramouli, A., J. Shi, Y. Feng, H. Holubec, R. M. Shanas, A. K. Bhattacharyya, W. Zheng and M. A. Nelson (2007). "Haploinsufficiency of the cdc2l gene contributes to skin cancer development in mice." Carcinogenesis **28**(9): 2028-35.

Chang, C. C., G. S. Shieh, P. Wu, C. C. Lin, A. L. Shiau and C. L. Wu (2008). "Oct-3/4 expression reflects tumor progression and regulates motility of bladder cancer cells." Cancer Res **68**(15): 6281-91.

Chang, E. T., K. E. Smedby, S. M. Zhang, H. Hjalgrim, M. Melbye, A. Ost, B. Glimelius, A. Wolk and H. O. Adami (2005). "Dietary factors and risk of

non-hodgkin lymphoma in men and women." Cancer Epidemiol Biomarkers Prev **14**(2): 512-20.

Chatterjee, N., P. Hartge, J. R. Cerhan, W. Cozen, S. Davis, N. Ishibe, J. Colt, L. Goldin and R. K. Severson (2004). "Risk of non-Hodgkin's lymphoma and family history of lymphatic, hematologic, and other cancers." Cancer Epidemiol Biomarkers Prev **13**(9): 1415-21.

Chen, J., V. Marechal and A. J. Levine (1993). "Mapping of the p53 and mdm-2 interaction domains." Mol Cell Biol **13**(7): 4107-14.

Cheson, B. D., S. J. Horning, B. Coiffier, M. A. Shipp, R. I. Fisher, J. M. Connors, T. A. Lister, J. Vose, A. Grillo-Lopez, A. Hagenbeek, F. Cabanillas, D. Klippensten, W. Hiddemann, R. Castellino, N. L. Harris, J. O. Armitage, W. Carter, R. Hoppe and G. P. Canellos (1999). "Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group." J Clin Oncol **17**(4): 1244 - 1253.

Cheung, K.-J., Johnson NA, Affleck JG, Severson T, Steidl C, Ben-Neriah S, Schein J, Morin RD, Moore R, Shah SP, Qian H, Paul JE, Telenius A, Lai B, Relander T, Lam W, Savage K, Connors JM, Brown C, Marra MA, Gascoyne RD and Horsman DE. (2009a). "TNFRSF14 is Mutated in a Subset of Follicular Lymphoma and Correlated with Inferior Prognosis.": Blood (ASH Annual Meeting Abstracts), Nov 2009; 114: #1919.

Cheung, K. J., A. Delaney, S. Ben-Neriah, J. Schein, T. Lee, S. P. Shah, D. Cheung, N. A. Johnson, A. J. Mungall, A. Telenius, B. Lai, M. Boyle, J. M. Connors, R. D. Gascoyne, M. A. Marra and D. E. Horsman (2010a). "High resolution analysis of follicular lymphoma genomes reveals somatic recurrent sites of copy-neutral loss of heterozygosity and copy number alterations that target single genes." Genes Chromosomes Cancer **49**(8): 669-81.

Cheung, K. J., N. A. Johnson, J. G. Affleck, T. Severson, C. Steidl, S. Ben-Neriah, J. Schein, R. D. Morin, R. Moore, S. P. Shah, H. Qian, J. E. Paul, A. Telenius, T. Relander, W. Lam, K. Savage, J. M. Connors, C. Brown, M. A. Marra, R. D. Gascoyne and D. E. Horsman (2010b). "Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis." Cancer Res **70**(22): 9166-74.

Cheung, K. J., S. P. Shah, C. Steidl, N. Johnson, T. Relander, A. Telenius, B. Lai, K. P. Murphy, W. Lam, A. J. Al-Tourah, J. M. Connors, R. T. Ng, R. D. Gascoyne and D. E. Horsman (2009b). "Genome-wide profiling of follicular lymphoma by array comparative genomic hybridization reveals prognostically significant DNA copy number imbalances." Blood **113**(1): 137-48.

Chiu, B. C., B. J. Dave, A. Blair, S. M. Gapstur, S. H. Zahm and D. D. Weisenburger (2006). "Agricultural pesticide use and risk of t(14;18)-defined subtypes of non-Hodgkin lymphoma." Blood **108**(4): 1363-9.

Christie, L., N. Kernohan, D. Levison, M. Sales, J. Cunningham, K. Gillespie, P. Batstone, D. Meiklejohn and J. Goodlad (2008). "C-MYC translocation in t(14;18) positive follicular lymphoma at presentation: An adverse prognostic indicator?" Leuk Lymphoma **49**(3): 470-6.

Chua, H. W., D. Ng, S. Choo, S. S. Lum, H. Li, L. Y. Soh, K. Sabapathy and A. Seow (2010). "Effect of MDM2 SNP309 and p53 codon 72 polymorphisms on lung cancer risk and survival among non-smoking Chinese women in Singapore." BMC Cancer **10**: 88.

Cizmarikova, M., M. Wagnerova, L. Schonova, V. Habalova, A. Kohut, A. Linkova, M. Sarissky, J. Mojzis, L. Mirossay and A. Mirossay (2010). "MDR1 (C3435T) polymorphism: relation to the risk of breast cancer and therapeutic outcome." Pharmacogenomics J **10**(1): 62-9.

Cleary, M. L., N. Galili and J. Sklar (1986a). "Detection of a second t(14;18) breakpoint cluster region in human follicular lymphomas." J Exp Med **164**(1): 315-20.

Cleary, M. L., N. Galili, M. Trela, R. Levy and J. Sklar (1988). "Single cell origin of bigenotypic and biphenotypic B cell proliferations in human follicular lymphomas." J Exp Med **167**(2): 582-97.

Cleary, M. L., T. C. Meeker, S. Levy, E. Lee, M. Trela, J. Sklar and R. Levy (1986b). "Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma." Cell **44**(1): 97-106.

Cleary, M. L., S. D. Smith and J. Sklar (1986c). "Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation." Cell **47**(1): 19-28.

"A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project." (1997). Blood **89**(11): 3909-18.

Co, M. S., D. A. Scheinberg, N. M. Avdalovic, K. McGraw, M. Vasquez, P. C. Caron and C. Queen (1993). "Genetically engineered deglycosylation of the variable domain increases the affinity of an anti-CD33 monoclonal antibody." Mol Immunol **30**(15): 1361-7.

Cobaleda, C., N. Gutierrez-Cianca, J. Perez-Losada, T. Flores, R. Garcia-Sanz, M. Gonzalez and I. Sanchez-Garcia (2000). "A primitive hematopoietic cell is the target for the leukemic transformation in human philadelphia-positive acute lymphoblastic leukemia." Blood **95**(3): 1007-13.

Coelho, V., S. Krysov, A. M. Ghaemmaghami, M. Emara, K. N. Potter, P. Johnson, G. Packham, L. Martinez-Pomares and F. K. Stevenson (2010). "Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins." Proc Natl Acad Sci U S A **107**(43): 18587-92.

Coll-Mulet, L., D. Iglesias-Serret, A. F. Santidrian, A. M. Cosialls, M. de Frias, E. Castano, C. Campas, M. Barragan, A. F. de Sevilla, A. Domingo, L. T. Vassilev, G. Pons and J. Gil (2006). "MDM2 antagonists activate p53 and synergize with genotoxic drugs in B-cell chronic lymphocytic leukemia cells." Blood **107**(10): 4109-14.

Coloma, M. J., R. K. Trinh, A. R. Martinez and S. L. Morrison (1999). "Position effects of variable region carbohydrate on the affinity and in vivo behavior of an anti-(1-->6) dextran antibody." J Immunol **162**(4): 2162-70.

Compaan, D. M., L. C. Gonzalez, I. Tom, K. M. Loyet, D. Eaton and S. G. Hymowitz (2005). "Attenuating lymphocyte activity: the crystal structure of the BTLA-HVEM complex." J Biol Chem **280**(47): 39553-61.

"Complete sequence and gene map of a human major histocompatibility complex." (1999). Nature **401**(6756): 921-3.

Conde, L., E. Halperin, N. K. Akers, K. M. Brown, K. E. Smedby, N. Rothman, A. Nieters, S. L. Slager, A. Brooks-Wilson, L. Agana, J. Riby, J. Liu, H. O. Adami, H. Darabi, H. Hjalgrim, H. Q. Low, K. Humphreys, M. Melbye, E. T. Chang, B. Glimelius, W. Cozen, S. Davis, P. Hartge, L. M. Morton, M.

Schenk, S. S. Wang, B. Armstrong, A. Kricker, S. Milliken, M. P. Purdue, C. M. Vajdic, P. Boyle, Q. Lan, S. H. Zahm, Y. Zhang, T. Zheng, N. Becker, Y. Benavente, P. Boffetta, P. Brennan, K. Butterbach, P. Cocco, L. Foretova, M. Maynadie, S. de Sanjose, A. Staines, J. J. Spinelli, S. J. Achenbach, T. G. Call, N. J. Camp, M. Glenn, N. E. Caporaso, J. R. Cerhan, J. M. Cunningham, L. R. Goldin, C. A. Hanson, N. E. Kay, M. C. Lanasa, J. F. Leis, G. E. Marti, K. G. Rabe, L. Z. Rassenti, L. G. Spector, S. S. Strom, C. M. Vachon, J. B. Weinberg, E. A. Holly, S. Chanock, M. T. Smith, P. M. Bracci and C. F. Skibola (2010). "Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32." Nat Genet **42**(8): 661-4.

Cordon-Cardo, C., E. Latres, M. Drobnjak, M. R. Oliva, D. Pollack, J. M. Woodruff, V. Marechal, J. Chen, M. F. Brennan and A. J. Levine (1994). "Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas." Cancer Res **54**(3): 794-9.

Cox, C. V., R. S. Evely, A. Oakhill, D. H. Pamphilon, N. J. Goulden and A. Blair (2004). "Characterization of acute lymphoblastic leukemia progenitor cells." Blood **104**(9): 2919-25.

Czuczman, M. S., A. J. Grillo-Lopez, C. A. White, M. Saleh, L. Gordon, A. F. LoBuglio, C. Jonas, D. Klippenstein, B. Dallaire and C. Varns (1999). "Treatment of patients with low-grade B-cell lymphoma with the combination of chimeric anti-CD20 monoclonal antibody and CHOP chemotherapy." J Clin Oncol **17**(1): 268-76.

d'Amore, F., E. Chan, J. Iqbal, H. Geng, K. Young, L. Xiao, M. M. Hess, W. G. Sanger, L. Smith, C. Wiuf, O. Hagberg, K. Fu, W. C. Chan and B. J. Dave (2008). "Clonal evolution in t(14;18)-positive follicular lymphoma, evidence for multiple common pathways, and frequent parallel clonal evolution." Clin Cancer Res **14**(22): 7180-7.

Dave, S. S., G. Wright, B. Tan, A. Rosenwald, R. D. Gascoyne, W. C. Chan, R. I. Fisher, R. M. Braziel, L. M. Rimsza, T. M. Grogan, T. P. Miller, M. LeBlanc, T. C. Greiner, D. D. Weisenburger, J. C. Lynch, J. Vose, J. O. Armitage, E. B. Smeland, S. Kvaloy, H. Holte, J. Delabie, J. M. Connors, P. M. Lansdorp, Q. Ouyang, T. A. Lister, A. J. Davies, A. J. Norton, H. K. Muller-Hermelink, G. Ott, E. Campo, E. Montserrat, W. H. Wilson, E. S. Jaffe, R. Simon, L. Yang, J. Powell, H. Zhao, N. Goldschmidt, M. Chiorazzi and L. M. Staudt (2004). "Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells." N Engl J Med **351**(21): 2159-69.

Davies, A. J., A. M. Lee, C. Taylor, A. J. Clear, L. K. Goff, S. Iqbal, D. Cuthbert-Heavens, M. Calaminici, A. J. Norton, T. A. Lister and J. Fitzgibbon (2005). "A limited role for TP53 mutation in the transformation of follicular lymphoma to diffuse large B-cell lymphoma." Leukemia **19**(8): 1459-65.

Davies, A. J., A. Z. Rohatiner, S. Howell, K. E. Britton, S. E. Owens, I. N. Micallef, D. P. Deakin, B. M. Carrington, J. A. Lawrance, S. Vinnicombe, S. J. Mather, J. Clayton, R. Foley, H. Jan, S. Kroll, M. Harris, J. Amess, A. J. Norton, T. A. Lister and J. A. Radford (2004). "Tositumomab and iodine I 131 tositumomab for recurrent indolent and transformed B-cell non-Hodgkin's lymphoma." J Clin Oncol **22**(8): 1469-79.

Davies, A. J., A. Rosenwald, G. Wright, A. Lee, K. W. Last, D. D. Weisenburger, W. C. Chan, J. Delabie, R. M. Braziel, E. Campo, R. D. Gascoyne, E. S. Jaffe, K. Muller-Hermelink, G. Ott, M. Calaminici, A. J. Norton, L. K. Goff, J. Fitzgibbon, L. M. Staudt and T. Andrew Lister (2007). "Transformation of

follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms." Br J Haematol **136**(2): 286-93.

de Bakker, P. I., G. McVean, P. C. Sabeti, M. M. Miretti, T. Green, J. Marchini, X. Ke, A. J. Monsuur, P. Whittaker, M. Delgado, J. Morrison, A. Richardson, E. C. Walsh, X. Gao, L. Galver, J. Hart, D. A. Hafler, M. Pericak-Vance, J. A. Todd, M. J. Daly, J. Trowsdale, C. Wijmenga, T. J. Vyse, S. Beck, S. S. Murray, M. Carrington, S. Gregory, P. Deloukas and J. D. Rioux (2006). "A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC." Nat Genet **38**(10): 1166-72.

de Jong, D. (2005). "Molecular pathogenesis of follicular lymphoma: a cross talk of genetic and immunologic factors." J Clin Oncol **23**(26): 6358-63.

de Jong, D., A. Koster, A. Hagenbeek, J. Raemaekers, D. Veldhuizen, S. Heisterkamp, J. P. de Boer and M. van Glabbeke (2009). "Impact of the tumor microenvironment on prognosis in follicular lymphoma is dependent on specific treatment protocols." Haematologica **94**(1): 70-7.

De Jong, D., B. M. Voetdijk, G. C. Beverstock, G. J. van Ommen, R. Willemze and P. M. Kluin (1988). "Activation of the c-myc oncogene in a precursor-B-cell blast crisis of follicular lymphoma, presenting as composite lymphoma." N Engl J Med **318**(21): 1373-8.

Decaudin, D., E. Lepage, N. Brousse, P. Brice, J. L. Harousseau, K. Belhadj, H. Tilly, L. Michaux, S. Cheze, B. Coiffier and P. Solal-Celigny (1999). "Low-grade stage III-IV follicular lymphoma: multivariate analysis of prognostic factors in 484 patients--a study of the groupe d'Etude des lymphomes de l'Adulte." J Clin Oncol **17**(8): 2499-505.

Deconinck, E., C. Foussard, N. Milpied, P. Bertrand, P. Michenet, P. Cornillet-LeFebvre, M. Escoffre-Barbe, H. Maisonneuve, V. Delwail, R. Gressin, E. Legouffe, J. P. Vilque, B. Desablens, J. Jaubert, J. F. Ramee, A. Jenabian, A. Thyss, A. Le Pourhiet-Le Mevel, P. Travade, R. Delepine and P. Colombat (2005). "High-dose therapy followed by autologous purged stem-cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: a randomized multicenter study by GOELAMS." Blood **105**(10): 3817-23.

DeWolf, W. C., P. H. Lange, R. Shepherd, S. Martin-Alosco and E. J. Yunis (1981). "Association of HLA and renal cell carcinoma." Hum Immunol **2**(1): 41-4.

Ding, L., G. Getz, D. A. Wheeler, E. R. Mardis, M. D. McLellan, K. Cibulskis, C. Sougnez, H. Greulich, D. M. Muzny, M. B. Morgan, L. Fulton, R. S. Fulton, Q. Zhang, M. C. Wendl, M. S. Lawrence, D. E. Larson, K. Chen, D. J. Dooling, A. Sabo, A. C. Hawes, H. Shen, S. N. Jhangiani, L. R. Lewis, O. Hall, Y. Zhu, T. Mathew, Y. Ren, J. Yao, S. E. Scherer, K. Clerc, G. A. Metcalf, B. Ng, A. Milosavljevic, M. L. Gonzalez-Garay, J. R. Osborne, R. Meyer, X. Shi, Y. Tang, D. C. Koboldt, L. Lin, R. Abbott, T. L. Miner, C. Pohl, G. Fewell, C. Haipek, H. Schmidt, B. H. Dunford-Shore, A. Kraja, S. D. Crosby, C. S. Sawyer, T. Vickery, S. Sander, J. Robinson, W. Winckler, J. Baldwin, L. R. Chirieac, A. Dutt, T. Fennell, M. Hanna, B. E. Johnson, R. C. Onofrio, R. K. Thomas, G. Tonon, B. A. Weir, X. Zhao, L. Ziaugra, M. C. Zody, T. Giordano, M. B. Orringer, J. A. Roth, M. R. Spitz, Wistuba, II, B. Ozenberger, P. J. Good, A. C. Chang, D. G. Beer, M. A. Watson, M. Ladanyi, S. Broderick, A. Yoshizawa, W. D. Travis, W. Pao, M. A. Province, G. M. Weinstock, H. E. Varmus, S. B. Gabriel, E. S. Lander, R. A. Gibbs, M.

Meyerson and R. K. Wilson (2008). "Somatic mutations affect key pathways in lung adenocarcinoma." Nature **455**(7216): 1069-75.

Dolken, G., G. Illerhaus, C. Hirt and R. Mertelsmann (1996). "BCL-2/JH rearrangements in circulating B cells of healthy blood donors and patients with nonmalignant diseases." J Clin Oncol **14**(4): 1333-44.

Doll, R. and A. B. Hill (1964). "Mortality in Relation to Smoking: Ten Years' Observations of British Doctors." Br Med J **1**(5395): 1399-410.

Dreyling, M. (2010). "Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up." Ann Oncol **21 Suppl 5**: v181-3.

Duhen, T., C. Pasero, F. Mallet, B. Barbarat, D. Olive and R. T. Costello (2004). "LIGHT costimulates CD40 triggering and induces immunoglobulin secretion; a novel key partner in T cell-dependent B cell terminal differentiation." Eur J Immunol **34**(12): 3534-41.

Dumont, P., J. I. Leu, A. C. Della Pietra, 3rd, D. L. George and M. Murphy (2003). "The codon 72 polymorphic variants of p53 have markedly different apoptotic potential." Nat Genet **33**(3): 357-65.

Durbin, R. M., G. R. Abecasis, D. L. Altshuler, A. Auton, L. D. Brooks, R. A. Gibbs, M. E. Hurles and G. A. McVean (2010). "A map of human genome variation from population-scale sequencing." Nature **467**(7319): 1061-73.

Eide, M. B., K. Liestol, O. C. Lingjaerde, M. E. Hystad, S. H. Kresse, L. Meza-Zepeda, O. Myklebost, G. Troen, H. V. Aamot, H. Holte, E. B. Smeland and J. Delabie (2010). "Genomic alterations reveal potential for higher grade transformation in follicular lymphoma and confirm parallel evolution of tumor cell clones." Blood **116**(9): 1489-97.

Elenitoba-Johnson, K. S., R. D. Gascoyne, M. S. Lim, M. Chhanabai, E. S. Jaffe and M. Raffeld (1998). "Homozygous deletions at chromosome 9p21 involving p16 and p15 are associated with histologic progression in follicle center lymphoma." Blood **91**(12): 4677-85.

Elenitoba-Johnson, K. S., S. D. Jenson, R. T. Abbott, R. A. Palais, S. D. Bohling, Z. Lin, S. Tripp, P. J. Shami, L. Y. Wang, R. W. Coupland, R. Buckstein, B. Perez-Ordonez, S. L. Perkins, I. D. Dube and M. S. Lim (2003). "Involvement of multiple signaling pathways in follicular lymphoma transformation: p38-mitogen-activated protein kinase as a target for therapy." Proc Natl Acad Sci U S A **100**(12): 7259-64.

Elledge, S. J. (1996). "Cell cycle checkpoints: preventing an identity crisis." Science **274**(5293): 1664-72.

"The ENCODE (ENCyclopedia Of DNA Elements) Project." (2004). Science **306**(5696): 636-40.

Evans, P. A., C. Pott, P. J. Groenen, G. Salles, F. Davi, F. Berger, J. F. Garcia, J. H. van Krieken, S. Pals, P. Kluin, E. Schuuring, M. Spaargaren, E. Boone, D. Gonzalez, B. Martinez, R. Villuendas, P. Gameiro, T. C. Diss, K. Mills, G. J. Morgan, G. I. Carter, B. J. Milner, D. Pearson, M. Hummel, W. Jung, M. Ott, D. Canioni, K. Beldjord, C. Bastard, M. H. Delfau-Larue, J. J. van Dongen, T. J. Molina and J. Cabecadas (2007). "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 **21**(2): 207-14.

Fakharzadeh, S. S., S. P. Trusko and D. L. George (1991). "Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumor cell line." Embo J **10**(6): 1565-9.

Farinha, P., A. Al-Tourah, K. Gill, R. Klasa, J. M. Connors and R. D. Gascoyne (2010). "The architectural pattern of FOXP3-positive T cells in follicular lymphoma is an independent predictor of survival and histologic transformation." Blood **115**(2): 289-95.

Farinha, P., H. Masoudi, B. F. Skinnider, K. Shumansky, J. J. Spinelli, K. Gill, R. Klasa, N. Voss, J. M. Connors and R. D. Gascoyne (2005). "Analysis of multiple biomarkers shows that lymphoma-associated macrophage (LAM) content is an independent predictor of survival in follicular lymphoma (FL)." Blood **106**(6): 2169-74.

Farmer, G., J. Bargonetti, H. Zhu, P. Friedman, R. Prywes and C. Prives (1992). "Wild-type p53 activates transcription in vitro." Nature **358**(6381): 83-6.

Fearon, E. R. and B. Vogelstein (1990). "A genetic model for colorectal tumorigenesis." Cell **61**(5): 759-67.

Federico, M., M. Bellei, L. Marcheselli, S. Luminari, A. Lopez-Guillermo, U. Vitolo, B. Pro, S. Pileri, A. Pulsoni, P. Soubeyran, S. Cortelazzo, G. Martinelli, M. Martelli, L. Rigacci, L. Arcaini, F. Di Raimondo, F. Merli, E. Sabattini, P. McLaughlin and P. Solal-Celigny (2009). "Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project." J Clin Oncol **27**(27): 4555-62.

Federico, M., U. Vitolo, P. L. Zinzani, T. Chisesi, V. Clo, G. Bellesi, M. Magagnoli, M. Liberati, C. Boccomini, P. Niscola, V. Pavone, A. Cuneo, G. Santini, M. Brugiatelli, L. Baldini, L. Rigacci and L. Resegotti (2000). "Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. Intergruppo Italiano Linfomi." Blood **95**(3): 783-9.

Fisher, R. I., M. LeBlanc, O. W. Press, D. G. Maloney, J. M. Unger and T. P. Miller (2005). "New treatment options have changed the survival of patients with follicular lymphoma." J Clin Oncol **23**(33): 8447-52.

Fisher, S. G. and R. I. Fisher (2004). "The epidemiology of non-Hodgkin's lymphoma." Oncogene **23**(38): 6524-34.

Fitzgibbon, J., S. Iqbal, A. Davies, D. O'Shea, E. Carlotti, T. Chaplin, J. Matthews, M. Raghavan, A. Norton, T. A. Lister and B. D. Young (2007). "Genome-wide detection of recurring sites of uniparental disomy in follicular and transformed follicular lymphoma." Leukemia **21**(7): 1514-20.

Florenes, V. A., G. M. Maelandsmo, A. Forus, A. Andreassen, O. Myklebost and O. Fodstad (1994). "MDM2 gene amplification and transcript levels in human sarcomas: relationship to TP53 gene status." J Natl Cancer Inst **86**(17): 1297-302.

Forbes, S. A., G. Bhamra, S. Bamford, E. Dawson, C. Kok, J. Clements, A. Menzies, J. W. Teague, P. A. Futreal and M. R. Stratton (2008). "The Catalogue of Somatic Mutations in Cancer (COSMIC)." Curr Protoc Hum Genet **Chapter 10**: Unit 10 11.

Forstpointner, R., M. Unterhalt, M. Dreyling, H. P. Bock, R. Repp, H. Wandt, C. Pott, J. F. Seymour, B. Metzner, A. Hanel, T. Lehmann, F. Hartmann, H. Einsele and W. Hiddemann (2006). "Maintenance therapy with rituximab leads to a significant prolongation of response duration after salvage therapy with a combination of rituximab, fludarabine, cyclophosphamide, and

mitoxantrone (R-FCM) in patients with recurring and refractory follicular and mantle cell lymphomas: Results of a prospective randomized study of the German Low Grade Lymphoma Study Group (GLSG)." Blood **108**(13): 4003-8.

Frazer, K. A., D. G. Ballinger, D. R. Cox, D. A. Hinds, L. L. Stuve, R. A. Gibbs, J. W. Belmont, A. Boudreau, P. Hardenbol, S. M. Leal, S. Pasternak, D. A. Wheeler, T. D. Willis, F. Yu, H. Yang, C. Zeng, Y. Gao, H. Hu, W. Hu, C. Li, W. Lin, S. Liu, H. Pan, X. Tang, J. Wang, W. Wang, J. Yu, B. Zhang, Q. Zhang, H. Zhao, J. Zhou, S. B. Gabriel, R. Barry, B. Blumenstiel, A. Camargo, M. Defelice, M. Faggart, M. Goyette, S. Gupta, J. Moore, H. Nguyen, R. C. Onofrio, M. Parkin, J. Roy, E. Stahl, E. Winchester, L. Ziaugra, D. Altshuler, Y. Shen, Z. Yao, W. Huang, X. Chu, Y. He, L. Jin, Y. Liu, W. Sun, H. Wang, Y. Wang, X. Xiong, L. Xu, M. M. Waye, S. K. Tsui, H. Xue, J. T. Wong, L. M. Galver, J. B. Fan, K. Gunderson, S. S. Murray, A. R. Oliphant, M. S. Chee, A. Montpetit, F. Chagnon, V. Ferretti, M. Leboeuf, J. F. Olivier, M. S. Phillips, S. Roumy, C. Sallee, A. Verner, T. J. Hudson, P. Y. Kwok, D. Cai, D. C. Koboldt, R. D. Miller, L. Pawlikowska, P. Taillon-Miller, M. Xiao, L. C. Tsui, W. Mak, Y. Q. Song, P. K. Tam, Y. Nakamura, T. Kawaguchi, T. Kitamoto, T. Morizono, A. Nagashima, Y. Ohnishi, A. Sekine, T. Tanaka, T. Tsunoda, P. Deloukas, C. P. Bird, M. Delgado, E. T. Dermitzakis, R. Gwilliam, S. Hunt, J. Morrison, D. Powell, B. E. Stranger, P. Whittaker, D. R. Bentley, M. J. Daly, P. I. de Bakker, J. Barrett, Y. R. Chretien, J. Maller, S. McCarroll, N. Patterson, I. Pe'er, A. Price, S. Purcell, D. J. Richter, P. Sabeti, R. Saxena, S. F. Schaffner, P. C. Sham, P. Varilly, L. D. Stein, L. Krishnan, A. V. Smith, M. K. Tello-Ruiz, G. A. Thorisson, A. Chakravarti, P. E. Chen, D. J. Cutler, C. S. Kashuk, S. Lin, G. R. Abecasis, W. Guan, Y. Li, H. M. Munro, Z. S. Qin, D. J. Thomas, G. McVean, A. Auton, L. Bottolo, N. Cardin, S. Eyheramendy, C. Freeman, J. Marchini, S. Myers, C. Spencer, M. Stephens, P. Donnelly, L. R. Cardon, G. Clarke, D. M. Evans, A. P. Morris, B. S. Weir, J. C. Mullikin, S. T. Sherry, M. Feolo, A. Skol, H. Zhang, I. Matsuda, Y. Fukushima, D. R. Macer, E. Suda, C. N. Rotimi, C. A. Adebamowo, I. Ajayi, T. Aniagwu, P. A. Marshall, C. Nkwodimmah, C. D. Royal, M. F. Leppert, M. Dixon, A. Peiffer, R. Qiu, A. Kent, K. Kato, N. Niikawa, I. F. Adewole, B. M. Knoppers, M. W. Foster, E. W. Clayton, J. Watkin, D. Muzny, L. Nazareth, E. Sodergren, G. M. Weinstock, I. Yakub, B. W. Birren, R. K. Wilson, L. L. Fulton, J. Rogers, J. Burton, N. P. Carter, C. M. Clee, M. Griffiths, M. C. Jones, K. McLay, R. W. Plumb, M. T. Ross, S. K. Sims, D. L. Willey, Z. Chen, H. Han, L. Kang, M. Godbout, J. C. Wallenburg, P. L'Archeveque, G. Bellemare, K. Saeki, D. An, H. Fu, Q. Li, Z. Wang, R. Wang, A. L. Holden, L. D. Brooks, J. E. McEwen, M. S. Guyer, V. O. Wang, J. L. Peterson, M. Shi, J. Spiegel, L. M. Sung, L. F. Zacharia, F. S. Collins, K. Kennedy, R. Jamieson and J. Stewart (2007). "A second generation human haplotype map of over 3.1 million SNPs." Nature **449**(7164): 851-61.

Freedman, D. A. and A. J. Levine (1998). "Nuclear export is required for degradation of endogenous p53 by MDM2 and human papillomavirus E6." Mol Cell Biol **18**(12): 7288-93.

Freedman, D. A., L. Wu and A. J. Levine (1999). "Functions of the MDM2 oncoprotein." Cell Mol Life Sci **55**(1): 96-107.

Friedberg, J. W., M. D. Taylor, J. R. Cerhan, C. R. Flowers, H. Dillon, C. M. Farber, E. S. Rogers, J. D. Hainsworth, E. K. Wong, J. M. Vose, A. D. Zelenetz and B. K. Link (2009). "Follicular lymphoma in the United States: first report of the national LymphoCare study." J Clin Oncol **27**(8): 1202-8.

Friedman, G. D. and C. P. Quesenberry, Jr. (1999). "Spousal concordance for cancer incidence: A cohort study." Cancer **86**(11): 2413-9.

Gall, E. A. and T. B. Mallory (1942). "Malignant Lymphoma: A Clinico-Pathologic Survey of 618 Cases." Am J Pathol **18**(3): 381-429.

Gallagher, C. J., W. M. Gregory, A. E. Jones, A. G. Stansfeld, M. A. Richards, H. S. Dhaliwal, J. S. Malpas and T. A. Lister (1986). "Follicular lymphoma: prognostic factors for response and survival." J Clin Oncol **4**(10): 1470-80.

Garvin, A. J., R. M. Simon, C. K. Osborne, J. Merrill, R. C. Young and C. W. Berard (1983). "An autopsy study of histologic progression in non-Hodgkin's lymphomas. 192 cases from the National Cancer Institute." Cancer **52**(3): 393-8.

"Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." (2007). Nature **447**(7145): 661-78.

Gentles, A. J., A. A. Alizadeh, S. I. Lee, J. H. Myklebust, C. M. Shachaf, B. Shahbaba, R. Levy, D. Koller and S. K. Plevritis (2009). "A pluripotency signature predicts histologic transformation and influences survival in follicular lymphoma patients." Blood **114**(15): 3158-66.

Ghielmini, M., S. F. Schmitz, S. B. Cogliatti, G. Pichert, J. Hummerjohann, U. Waltzer, M. F. Fey, D. C. Betticher, G. Martinelli, F. Peccatori, U. Hess, E. Zucca, R. Stupp, T. Kovacsovics, C. Helg, A. Lohri, M. Bargetzi, D. Vorobiof and T. Cerny (2004). "Prolonged treatment with rituximab in patients with follicular lymphoma significantly increases event-free survival and response duration compared with the standard weekly x 4 schedule." Blood **103**(12): 4416-23.

Giaccia, A. J. and M. B. Kastan (1998). "The complexity of p53 modulation: emerging patterns from divergent signals." Genes Dev **12**(19): 2973-83.

Giles, R. H., J. G. Dauwerse, C. Higgins, F. Petrij, J. W. Wessels, G. C. Beverstock, H. Dohner, M. Jotterand-Bellomo, J. H. Falkenburg, R. M. Slater, G. J. van Ommen, A. Hagemeijer, B. A. van der Reijden and M. H. Breuning (1997). "Detection of CBP rearrangements in acute myelogenous leukemia with t(8;16)." Leukemia **11**(12): 2087-96.

Gine, E., S. Montoto, F. Bosch, L. Arenillas, S. Mercadal, N. Villamor, A. Martinez, L. Colomo, E. Campo, E. Montserrat and A. Lopez-Guillermo (2006). "The Follicular Lymphoma International Prognostic Index (FLIPI) and the histological subtype are the most important factors to predict histological transformation in follicular lymphoma." Ann Oncol **17**(10): 1539-45.

Glas, A. M., M. J. Kersten, L. J. Delahaye, A. T. Witteveen, R. E. Kibbelaar, A. Velds, L. F. Wessels, P. Joosten, R. M. Kerkhoven, R. Bernards, J. H. van Krieken, P. M. Kluin, L. J. van't Veer and D. de Jong (2005). "Gene expression profiling in follicular lymphoma to assess clinical aggressiveness and to guide the choice of treatment." Blood **105**(1): 301-7.

Glas, A. M., L. Knoops, L. Delahaye, M. J. Kersten, R. E. Kibbelaar, L. A. Wessels, R. van Laar, J. H. van Krieken, J. W. Baars, J. Raemaekers, P. M. Kluin, L. J. van't Veer and D. de Jong (2007). "Gene-expression and immunohistochemical study of specific T-cell subsets and accessory cell

types in the transformation and prognosis of follicular lymphoma." J Clin Oncol **25**(4): 390-8.

Glick, J. H., J. M. Barnes, E. Z. Ezdinli, C. W. Berard, E. L. Orlow and J. M. Bennett (1981). "Nodular mixed lymphoma: results of a randomized trial failing to confirm prolonged disease-free survival with COPP chemotherapy." Blood **58**(5): 920-5.

Goff, L. K., M. J. Neat, C. R. Crawley, L. Jones, E. Jones, T. A. Lister and R. K. Gupta (2000). "The use of real-time quantitative polymerase chain reaction and comparative genomic hybridization to identify amplification of the REL gene in follicular lymphoma." Br J Haematol **111**(2): 618-25.

Goldin, L. R., M. Bjorkholm, S. Y. Kristinsson, I. Turesson and O. Landgren (2009). "Highly increased familial risks for specific lymphoma subtypes." Br J Haematol **146**(1): 91-4.

Gribben, J. G., D. Neuberg, A. S. Freedman, C. D. Gimmi, K. W. Pesek, M. Barber, L. Saporito, S. D. Woo, F. Coral, N. Spector and et al. (1993). "Detection by polymerase chain reaction of residual cells with the bcl-2 translocation is associated with increased risk of relapse after autologous bone marrow transplantation for B-cell lymphoma." Blood **81**(12): 3449-57.

Grochola, L. F., A. Vazquez, E. E. Bond, P. Wurl, H. Taubert, T. H. Muller, A. J. Levine and G. L. Bond (2009). "Recent natural selection identifies a genetic variant in a regulatory subunit of protein phosphatase 2A that associates with altered cancer risk and survival." Clin Cancer Res **15**(19): 6301-8.

Gross-Bellard, M., P. Oudet and P. Chambon (1973). "Isolation of high-molecular-weight DNA from mammalian cells." Eur J Biochem **36**(1): 32-8.

Groves, F. D., M. S. Linet, L. B. Travis and S. S. Devesa (2000). "Cancer surveillance series: non-Hodgkin's lymphoma incidence by histologic subtype in the United States from 1978 through 1995." J Natl Cancer Inst **92**(15): 1240-51.

Gryshchenko, I., S. Hofbauer, M. Stoecher, P. T. Daniel, M. Steurer, A. Gaiger, K. Eigenberger, R. Greil and I. Tinhofer (2008). "MDM2 SNP309 is associated with poor outcome in B-cell chronic lymphocytic leukemia." J Clin Oncol **26**(14): 2252-7.

Gu, W., X. L. Shi and R. G. Roeder (1997). "Synergistic activation of transcription by CBP and p53." Nature **387**(6635): 819-23.

Gu, Z., D. J. Gilbert, V. A. Valentine, N. A. Jenkins, N. G. Copeland and G. P. Zambetti (2000). "The p53-inducible gene EI24/PIG8 localizes to human chromosome 11q23 and the proximal region of mouse chromosome 9." Cytogenet Cell Genet **89**(3-4): 230-3.

Gupta, M., S. R. Dillon, S. C. Ziesmer, A. L. Feldman, T. E. Witzig, S. M. Ansell, J. R. Cerhan and A. J. Novak (2009). "A proliferation-inducing ligand mediates follicular lymphoma B-cell proliferation and cyclin D1 expression through phosphatidylinositol 3-kinase-regulated mammalian target of rapamycin activation." Blood **113**(21): 5206-16.

Gyan, E., C. Foussard, P. Bertrand, P. Michenet, S. Le Gouill, C. Berthou, H. Maisonneuve, V. Delwail, R. Gressin, P. Quittet, J. P. Vilque, B. Desablens, J. Jaubert, J. F. Ramee, N. Arakelyan, A. Thyss, C. Molucon-Chabrot, R. Delepine, N. Milpied, P. Colombat and E. Deconinck (2009). "High-dose therapy followed by autologous purged stem cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular

lymphoma: a randomized multicenter study by the GOELAMS with final results after a median follow-up of 9 years." Blood **113**(5): 995-1001.

Hainsworth, J. D., S. Litchy, H. A. Burris, 3rd, D. C. Scullin, Jr., S. W. Corso, D. A. Yardley, L. Morrissey and F. A. Greco (2002). "Rituximab as first-line and maintenance therapy for patients with indolent non-hodgkin's lymphoma." J Clin Oncol **20**(20): 4261-7.

Halldorsdottir, A. M., M. Fruhwirth, A. Deutsch, A. Aigelsreiter, C. Beham-Schmid, B. A. Agnarsson, P. Neumeister and W. Richard Burack (2008). "Quantifying the role of aberrant somatic hypermutation in transformation of follicular lymphoma." Leuk Res **32**(7): 1015-21.

Han, X., T. Zheng, F. M. Foss, Q. Lan, T. R. Holford, N. Rothman, S. Ma and Y. Zhang (2010). "Genetic polymorphisms in the metabolic pathway and non-Hodgkin lymphoma survival." Am J Hematol **85**(1): 51-6.

Hans, C. P., D. D. Weisenburger, J. M. Vose, L. M. Hock, J. C. Lynch, P. Aoun, T. C. Greiner, W. C. Chan, R. G. Bociek, P. J. Bierman and J. O. Armitage (2003). "A significant diffuse component predicts for inferior survival in grade 3 follicular lymphoma, but cytologic subtypes do not predict survival." Blood **101**(6): 2363-7.

Hardingham, J. E., W. J. Butler, D. Roder, A. Dobrovic, R. B. Dymock, R. E. Sage and I. C. Roberts-Thomson (1998). "Somatic mutations, acetylator status, and prognosis in colorectal cancer." Gut **42**(5): 669-72.

Harris, N. L., E. S. Jaffe, H. Stein, P. M. Banks, J. K. Chan, M. L. Cleary, G. Delsol, C. De Wolf-Peeters, B. Falini, K. C. Gatter and et al. (1994). "A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group." Blood **84**(5): 1361-92.

Hart, J., A. R. Turner, L. Larratt, J. Russell, B. Franko, C. Frantz, T. Paonessa, A. Mansoor and R. Lai (2007). "Transmission of a follicular lymphoma by allogeneic bone marrow transplantation--evidence to support the existence of lymphoma progenitor cells." Br J Haematol **136**(1): 166-7.

Hartmann, E., V. Fernandez, H. Stoecklein, L. Hernandez, E. Campo and A. Rosenwald (2007). "Increased MDM2 expression is associated with inferior survival in mantle-cell lymphoma, but not related to the MDM2 SNP309." Haematologica **92**(4): 574-5.

He, L., J. M. Thomson, M. T. Hemann, E. Hernando-Monge, D. Mu, S. Goodson, S. Powers, C. Cordon-Cardo, S. W. Lowe, G. J. Hannon and S. M. Hammond (2005). "A microRNA polycistron as a potential human oncogene." Nature **435**(7043): 828-33.

Henderson, B. E., E. Louie, J. SooHoo Jing, P. Buell and M. B. Gardner (1976). "Risk factors associated with nasopharyngeal carcinoma." N Engl J Med **295**(20): 1101-6.

Herold, M., A. Haas, S. Srock, S. Neser, K. H. Al-Ali, A. Neubauer, G. Dolken, R. Naumann, W. Knauf, M. Freund, R. Rohrberg, K. Hoffken, A. Franke, T. Ittel, E. Kettner, U. Haak, U. Mey, C. Klinkenstein, M. Assmann and U. von Grunhagen (2007). "Rituximab added to first-line mitoxantrone, chlorambucil, and prednisolone chemotherapy followed by interferon maintenance prolongs survival in patients with advanced follicular lymphoma: an East German Study Group Hematology and Oncology Study." J Clin Oncol **25**(15): 1986-92.

Hicks, E. B., H. Rappaport and W. J. Winter (1956). "Follicular lymphoma; a re-evaluation of its position in the scheme of malignant lymphoma, based on a survey of 253 cases." Cancer **9**(4): 792-821.

Hiddemann, W., M. Kneba, M. Dreyling, N. Schmitz, E. Lengfelder, R. Schmits, M. Reiser, B. Metzner, H. Harder, S. Hegewisch-Becker, T. Fischer, M. Kropff, H. E. Reis, M. Freund, B. Wormann, R. Fuchs, M. Planker, J. Schimke, H. Eimermacher, L. Trumper, A. Aldaoud, R. Parwaresch and M. Unterhalt (2005). "Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group." Blood **106**(12): 3725-32.

Hillion, J., C. Mecucci, A. Aventin, D. Leroux, I. Wlodarska, H. Van Den Berghe and C. J. Larsen (1991). "A variant translocation t(2;18) in follicular lymphoma involves the 5' end of bcl-2 and Ig kappa light chain gene." Oncogene **6**(1): 169-72.

Hirao, A., Y. Y. Kong, S. Matsuoka, A. Wakeham, J. Ruland, H. Yoshida, D. Liu, S. J. Elledge and T. W. Mak (2000). "DNA damage-induced activation of p53 by the checkpoint kinase Chk2." Science **287**(5459): 1824-7.

Hochster, H., E. Weller, R. D. Gascoyne, T. M. Habermann, L. I. Gordon, T. Ryan, L. Zhang, N. Colocci, S. Frankel and S. J. Horning (2009). "Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: results of the randomized phase III ECOG1496 Study." J Clin Oncol **27**(10): 1607-14.

Hoffman, A. E., T. Zheng, R. G. Stevens, Y. Ba, Y. Zhang, D. Leaderer, C. Yi, T. R. Holford and Y. Zhu (2009). "Clock-cancer connection in non-Hodgkin's lymphoma: a genetic association study and pathway analysis of the circadian gene cryptochrome 2." Cancer Res **69**(8): 3605-13.

Hoglund, M., L. Sehn, J. M. Connors, R. D. Gascoyne, R. Siebert, T. Sall, F. Mitelman and D. E. Horsman (2004). "Identification of cytogenetic subgroups and karyotypic pathways of clonal evolution in follicular lymphomas." Genes Chromosomes Cancer **39**(3): 195-204.

Hohaus, S., G. Mansueto, G. Massini, F. D'Alo, M. Giachelia, M. Martini, L. M. Larocca, M. T. Voso and G. Leone (2007). "Glutathione-S-transferase genotypes influence prognosis in follicular non-Hodgkin's Lymphoma." Leuk Lymphoma **48**(3): 564-9.

Hollstein, M., D. Sidransky, B. Vogelstein and C. C. Harris (1991). "p53 mutations in human cancers." Science **253**(5015): 49-53.

Honda, R., H. Tanaka and H. Yasuda (1997). "Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53." FEBS Lett **420**(1): 25-7.

Hong, Y., X. Miao, X. Zhang, F. Ding, A. Luo, Y. Guo, W. Tan, Z. Liu and D. Lin (2005). "The role of P53 and MDM2 polymorphisms in the risk of esophageal squamous cell carcinoma." Cancer Res **65**(20): 9582-7.

Horning, S. J. and S. A. Rosenberg (1984). "The natural history of initially untreated low-grade non-Hodgkin's lymphomas." N Engl J Med **311**(23): 1471-5.

Horsman, D. E., J. M. Connors, T. Pantzar and R. D. Gascoyne (2001). "Analysis of secondary chromosomal alterations in 165 cases of follicular lymphoma with t(14;18)." Genes Chromosomes Cancer **30**(4): 375-82.

Hough, R. E., J. R. Goepel, H. E. Alcock, B. W. Hancock, P. C. Lorigan and D. W. Hammond (2001). "Copy number gain at 12q12-14 may be important in the transformation from follicular lymphoma to diffuse large B cell lymphoma." Br J Cancer **84**(4): 499-503.

Hozumi, N. and S. Tonegawa (1976). "Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions." Proc Natl Acad Sci U S A **73**(10): 3628-32.

Hummel, M., J. Tamaru, B. Kalvelage and H. Stein (1994). "Mantle cell (previously centrocytic) lymphomas express VH genes with no or very little somatic mutations like the physiologic cells of the follicle mantle." Blood **84**(2): 403-7.

Husson, H., E. G. Carideo, D. Neuberg, J. Schultze, O. Munoz, P. W. Marks, J. W. Donovan, A. C. Chillemi, P. O'Connell and A. S. Freedman (2002). "Gene expression profiling of follicular lymphoma and normal germinal center B cells using cDNA arrays." Blood **99**(1): 282-9.

Iafrate, A. J., L. Feuk, M. N. Rivera, M. L. Listewnik, P. K. Donahoe, Y. Qi, S. W. Scherer and C. Lee (2004). "Detection of large-scale variation in the human genome." Nat Genet **36**(9): 949-51.

Ingram, W., S. Devereux, E. P. Das-Gupta, N. H. Russell, A. P. Haynes, J. L. Byrne, B. E. Shaw, A. McMillan, J. Gonzalez, A. Ho, G. J. Mufti and A. Pagliuca (2008). "Outcome of BEAM-autologous and BEAM-alemtuzumab allogeneic transplantation in relapsed advanced stage follicular lymphoma." Br J Haematol **141**(2): 235-43.

"The International HapMap Consortium. The International HapMap Project." (2003). Nature **426**(6968): 789-96.

Ismail, S. I., N. A. Ababneh, Y. Khader, A. A. Abu-Khader and A. Awidi (2009). "Methylenetetrahydrofolate reductase genotype association with the risk of follicular lymphoma." Cancer Genet Cytogenet **195**(2): 120-4.

Jacob, J., G. Kelsoe, K. Rajewsky and U. Weiss (1991). "Intraclonal generation of antibody mutants in germinal centres." Nature **354**(6352): 389-92.

Jaffe, E. S., Harris N. L., Stein H. and Vardiman J.W., Eds. (2001). Pathology and Genetics of Tumours of the Haematopoietic and Lymphoid Tissues. IARC Press, Lyon.

Jamieson, C. H., L. E. Ailles, S. J. Dylla, M. Muijtjens, C. Jones, J. L. Zehnder, J. Gotlib, K. Li, M. G. Manz, A. Keating, C. L. Sawyers and I. L. Weissman (2004). "Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML." N Engl J Med **351**(7): 657-67.

Jardin, F., P. Ruminy, F. Parmentier, J. M. Picquenot, M. N. Courel, P. Bertrand, G. Buchonnet, H. Tilly and C. Bastard (2005). "Clinical and biological relevance of single-nucleotide polymorphisms and acquired somatic mutations of the BCL6 first intron in follicular lymphoma." Leukemia **19**(10): 1824-30.

Jenner, M. J., K. E. Summers, A. J. Norton, J. A. Amess, R. S. Arch, B. D. Young, T. A. Lister, J. Fitzgibbon and L. K. Goff (2002). "JH probe real-time quantitative polymerase chain reaction assay for Bcl-2/IgH rearrangements." Br J Haematol **118**(2): 550-8.

Jin, S. and A. J. Levine (2001). "The p53 functional circuit." J Cell Sci **114**(Pt 23): 4139-40.

Johansson, B., F. Mertens and F. Mitelman (1995). "Cytogenetic evolution patterns in non-Hodgkin's lymphoma." Blood **86**(10): 3905-14.

Johnson, A. D., R. E. Handsaker, S. L. Pulit, M. M. Nizzari, C. J. O'Donnell and P. I. de Bakker (2008a). "SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap." Bioinformatics **24**(24): 2938-9.

Johnson, N. A., A. Al-Tourah, C. J. Brown, J. M. Connors, R. D. Gascoyne and D. E. Horsman (2008b). "Prognostic significance of secondary cytogenetic alterations in follicular lymphomas." Genes Chromosomes Cancer **47**(12): 1038-48.

Johnson, N. A., Morin R., Severson T., Mungall A. J., Zhao Y., Schein J. E., Boyle M., Woolcock B., Moore R., Holt R. A., Horsman D. E., Connors J. M., Jones S. J., Hirst M., Marra M. A. and Gascoyne R. D. (2009a). "FAS Mutations in Follicular Lymphoma Are Rare but Associated with Aggressive Clinical Behaviour." Blood (ASH Annual Meeting Abstracts), Nov 2009; 114: #3967.

Johnson, N. A., K. J. Savage, O. Ludkovski, S. Ben-Neriah, R. Woods, C. Steidl, M. J. Dyer, R. Siebert, J. Kuruvilla, R. Klasa, J. M. Connors, R. D. Gascoyne and D. E. Horsman (2009b). "Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival." Blood **114**(11): 2273-9.

Johnson, P. W., A. Z. Rohatiner, J. S. Whelan, C. G. Price, S. Love, J. Lim, J. Matthews, A. J. Norton, J. A. Amess and T. A. Lister (1995). "Patterns of survival in patients with recurrent follicular lymphoma: a 20-year study from a single center." J Clin Oncol **13**(1): 140-7.

Jones, E. H., R. J. Biggar, F. K. Nkrumah and S. D. Lawler (1980). "Study of the HLA system in Burkitt's lymphoma." Hum Immunol **1**(3): 207-10.

Jones, E. H., R. J. Biggar, F. K. Nkrumah and S. D. Lawler (1985). "HLA-DR7 association with African Burkitt's lymphoma." Hum Immunol **13**(3): 211-7.

Jones, S. N., A. R. Hancock, H. Vogel, L. A. Donehower and A. Bradley (1998). "Overexpression of Mdm2 in mice reveals a p53-independent role for Mdm2 in tumorigenesis." Proc Natl Acad Sci U S A **95**(26): 15608-12.

Juszczynski, P., E. Kalinka, J. Bienvenu, G. Woszczek, M. Borowiec, T. Robak, M. Kowalski, E. Lech-Maranda, L. Baseggio, B. Coiffier, G. Salles and K. Warzocha (2002). "Human leukocyte antigens class II and tumor necrosis factor genetic polymorphisms are independent predictors of non-Hodgkin lymphoma outcome." Blood **100**(8): 3037-40.

Kallioniemi, A., O. P. Kallioniemi, D. Sudar, D. Rutovitz, J. W. Gray, F. Waldman and D. Pinkel (1992). "Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors." Science **258**(5083): 818-21.

Kaminski, M. S., A. D. Zelenetz, O. W. Press, M. Saleh, J. Leonard, L. Fehrenbacher, T. A. Lister, R. J. Stagg, G. F. Tidmarsh, S. Kroll, R. L. Wahl, S. J. Knox and J. M. Vose (2001). "Pivotal study of iodine I 131 tositumomab for chemotherapy-refractory low-grade or transformed low-grade B-cell non-Hodgkin's lymphomas." J Clin Oncol **19**(19): 3918-28.

Kandioler, D., R. Zwrtek, C. Ludwig, E. Janschek, M. Ploner, F. Hofbauer, I. Kuhrer, S. Kappel, F. Wrba, M. Horvath, J. Karner, K. Renner, M. Bergmann, J. Karner-Hanusch, R. Potter, R. Jakesz, B. Teleky and F. Herbst (2002). "TP53 genotype but not p53 immunohistochemical result predicts response to preoperative short-term radiotherapy in rectal cancer." Ann Surg **235**(4): 493-8.

Kane, E. V., D. Painter, E. Roman, J. Allan, G. Law and T. Lightfoot (2010). "Melanocortin 1 receptor (MC1R), pigmentary characteristics and sun

exposure: findings from a case-control study of diffuse large B-cell and follicular lymphoma." Cancer Epidemiol **34**(2): 136-41.

Karube, K., Y. Guo, J. Suzumiya, Y. Sugita, Y. Nomura, K. Yamamoto, K. Shimizu, S. Yoshida, H. Komatani, M. Takeshita, M. Kikuchi, N. Nakamura, O. Takasu, F. Arakawa, H. Tagawa, M. Seto and K. Ohshima (2007). "CD10-MUM1+ follicular lymphoma lacks BCL2 gene translocation and shows characteristic biologic and clinical features." Blood **109**(7): 3076-9.

Katzenberger, T., J. Kalla, E. Leich, H. Stocklein, E. Hartmann, S. Barnickel, S. Wessendorf, M. M. Ott, H. K. Muller-Hermelink, A. Rosenwald and G. Ott (2009). "A distinctive subtype of t(14;18)-negative nodal follicular non-Hodgkin lymphoma characterized by a predominantly diffuse growth pattern and deletions in the chromosomal region 1p36." Blood **113**(5): 1053-61.

Katzenberger, T., G. Ott, T. Klein, J. Kalla, H. K. Muller-Hermelink and M. M. Ott (2004). "Cytogenetic alterations affecting BCL6 are predominantly found in follicular lymphomas grade 3B with a diffuse large B-cell component." Am J Pathol **165**(2): 481-90.

Khouri, I. F., P. McLaughlin, R. M. Saliba, C. Hosing, M. Korbling, M. S. Lee, L. J. Medeiros, L. Fayad, F. Samaniego, A. Alousi, P. Anderlini, D. Couriel, M. de Lima, S. Giralt, S. S. Neelapu, N. T. Ueno, B. I. Samuels, F. Hagemeister, L. W. Kwak and R. E. Champlin (2008). "Eight-year experience with allogeneic stem cell transplantation for relapsed follicular lymphoma after nonmyeloablative conditioning with fludarabine, cyclophosphamide, and rituximab." Blood **111**(12): 5530-6.

Killian, J. K., S. Bilke, S. Davis, R. L. Walker, M. S. Killian, E. B. Jaeger, Y. Chen, J. Hipp, S. Pittaluga, M. Raffeld, R. Cornelison, W. I. Smith, Jr., M. Bibikova, J. B. Fan, M. R. Emmert-Buck, E. S. Jaffe and P. S. Meltzer (2009). "Large-scale profiling of archival lymph nodes reveals pervasive remodeling of the follicular lymphoma methylome." Cancer Res **69**(3): 758-64.

Kirmizis, A., S. M. Bartley, A. Kuzmichev, R. Margueron, D. Reinberg, R. Green and P. J. Farnham (2004). "Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27." Genes Dev **18**(13): 1592-605.

Klein, J. and A. Sato (2000a). "The HLA system. First of two parts." N Engl J Med **343**(10): 702-9.

Klein, J. and A. Sato (2000b). "The HLA system. Second of two parts." N Engl J Med **343**(11): 782-6.

Knudson, A. G., Jr. (1971). "Mutation and cancer: statistical study of retinoblastoma." Proc Natl Acad Sci U S A **68**(4): 820-3.

Knutsen, T. (1997). "Cytogenetic mechanisms in the pathogenesis and progression of follicular lymphoma." Cancer Surv **30**: 163-92.

Ko, L. J. and C. Prives (1996). "p53: puzzle and paradigm." Genes Dev **10**(9): 1054-72.

Kojima, K., M. Konopleva, T. McQueen, S. O'Brien, W. Plunkett and M. Andreeff (2006). "Mdm2 inhibitor Nutlin-3a induces p53-mediated apoptosis by transcription-dependent and transcription-independent mechanisms and may overcome Atm-mediated resistance to fludarabine in chronic lymphocytic leukemia." Blood **108**(3): 993-1000.

Kon, S., S. Levy and R. Levy (1987). "Retention of an idiotypic determinant in a human B-cell lymphoma undergoing immunoglobulin variable-region mutation." Proc Natl Acad Sci U S A **84**(14): 5053-7.

Kroft, S. H., R. Domiati-Saad, W. G. Finn, D. B. Dawson, B. Schnitzer, T. P. Singleton and C. W. Ross (2000). "Precursor B-lymphoblastic transformation of grade I follicle center lymphoma." Am J Clin Pathol **113**(3): 411-8.

Kuppers, R. and R. Dalla-Favera (2001). "Mechanisms of chromosomal translocations in B cell lymphomas." Oncogene **20**(40): 5580-94.

Kuppers, R., U. Klein, M. L. Hansmann and K. Rajewsky (1999). "Cellular origin of human B-cell lymphomas." N Engl J Med **341**(20): 1520-9.

Kuppers, R., M. Zhao, M. L. Hansmann and K. Rajewsky (1993). "Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections." EMBO J **12**(13): 4955-67.

Kuzmichev, A., K. Nishioka, H. Erdjument-Bromage, P. Tempst and D. Reinberg (2002). "Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein." Genes Dev **16**(22): 2893-905.

Kwak, L. W., M. J. Campbell, D. K. Czerwinski, S. Hart, R. A. Miller and R. Levy (1992). "Induction of immune responses in patients with B-cell lymphoma against the surface-immunoglobulin idiotype expressed by their tumors." N Engl J Med **327**(17): 1209-15.

Lahti, J. M., M. Valentine, J. Xiang, B. Jones, J. Amann, J. Grenet, G. Richmond, A. T. Look and V. J. Kidd (1994). "Alterations in the PITSLRE protein kinase gene complex on chromosome 1p36 in childhood neuroblastoma." Nat Genet **7**(3): 370-5.

Lai, R., D. A. Arber, K. L. Chang, C. S. Wilson and L. M. Weiss (1998). "Frequency of bcl-2 expression in non-Hodgkin's lymphoma: a study of 778 cases with comparison of marginal zone lymphoma and monocytoid B-cell hyperplasia." Mod Pathol **11**(9): 864-9.

Lain, S., D. Xirodimas and D. P. Lane (1999). "Accumulating active p53 in the nucleus by inhibition of nuclear export: a novel strategy to promote the p53 tumor suppressor function." Exp Cell Res **253**(2): 315-24.

Lan, Q., T. Zheng, S. Chanock, Y. Zhang, M. Shen, S. S. Wang, S. I. Berndt, S. H. Zahm, T. R. Holford, B. Leaderer, M. Yeager, R. Welch, D. Hosgood, P. Boyle and N. Rothman (2007). "Genetic variants in caspase genes and susceptibility to non-Hodgkin lymphoma." Carcinogenesis **28**(4): 823-7.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L.

Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y. J. Chen (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Lane, D. P. (1992). "Cancer. p53, guardian of the genome." Nature **358**(6381): 15-6.

Lapidot, T., C. Sirard, J. Vormoor, B. Murdoch, T. Hoang, J. Caceres-Cortes, M. Minden, B. Paterson, M. A. Caligiuri and J. E. Dick (1994). "A cell initiating human acute myeloid leukaemia after transplantation into SCID mice." Nature **367**(6464): 645-8.

Last, K. W., L. K. Goff, K. E. Summers, M. Neat, M. Jenner, C. Crawley, A. Z. Rohatiner, J. Fitzgibbon and T. A. Lister (2000). "Familial follicular lymphoma: a case report with molecular analysis." Br J Haematol **110**(3): 744-5.

Leach, F. S., T. Tokino, P. Meltzer, M. Burrell, J. D. Oliner, S. Smith, D. E. Hill, D. Sidransky, K. W. Kinzler and B. Vogelstein (1993). "p53 Mutation and MDM2 amplification in human soft tissue sarcomas." Cancer Res **53**(10 Suppl): 2231-4.

LeBrun, D. P., B. Y. Ngan, L. M. Weiss, P. Huie, R. A. Warnke and M. L. Cleary (1994). "The bcl-2 oncogene in Hodgkin's disease arising in the setting of follicular non-Hodgkin's lymphoma." Blood **83**(1): 223-30.

Lee, A. M., A. J. Clear, M. Calaminici, A. J. Davies, S. Jordan, F. MacDougall, J. Matthews, A. J. Norton, J. G. Gribben, T. A. Lister and L. K. Goff (2006a). "Number of CD4+ cells and location of forkhead box protein P3-positive

cells in diagnostic follicular lymphoma tissue microarrays correlates with outcome." J Clin Oncol **24**(31): 5052-9.

Lee, J. T., D. J. Innes, Jr. and M. E. Williams (1989). "Sequential bcl-2 and c-myc oncogene rearrangements associated with the clinical transformation of non-Hodgkin's lymphoma." J Clin Invest **84**(5): 1454-9.

Lee, K. M., Q. Lan, A. Kricker, M. P. Purdue, A. E. Grulich, C. M. Vajdic, J. Turner, D. Whitby, D. Kang, S. Chanock, N. Rothman and B. K. Armstrong (2007). "One-carbon metabolism gene polymorphisms and risk of non-Hodgkin lymphoma in Australia." Hum Genet **122**(5): 525-33.

Lee, T. I., R. G. Jenner, L. A. Boyer, M. G. Guenther, S. S. Levine, R. M. Kumar, B. Chevalier, S. E. Johnstone, M. F. Cole, K. Isono, H. Koseki, T. Fuchikami, K. Abe, H. L. Murray, J. P. Zucker, B. Yuan, G. W. Bell, E. Herbolsheimer, N. M. Hannett, K. Sun, D. T. Odom, A. P. Otte, T. L. Volkert, D. P. Bartel, D. A. Melton, D. K. Gifford, R. Jaenisch and R. A. Young (2006b). "Control of developmental regulators by Polycomb in human embryonic stem cells." Cell **125**(2): 301-13.

Lee, W., Z. Jiang, J. Liu, P. M. Haverty, Y. Guan, J. Stinson, P. Yue, Y. Zhang, K. P. Pant, D. Bhatt, C. Ha, S. Johnson, M. I. Kennemer, S. Mohan, I. Nazarenko, C. Watanabe, A. B. Sparks, D. S. Shames, R. Gentleman, F. J. de Sauvage, H. Stern, A. Pandita, D. G. Ballinger, R. Drmanac, Z. Modrusan, S. Seshagiri and Z. Zhang (2010). "The mutation spectrum revealed by paired genome sequences from a lung cancer patient." Nature **465**(7297): 473-7.

Lennert, K., H. Stein and E. Kaiserling (1975). "Cytological and functional criteria for the classification of malignant lymphomata." Br J Cancer Suppl **2**: 29-43.

Leonard, R. C., R. L. Hayward, R. J. Prescott and J. X. Wang (1991). "The identification of discrete prognostic groups in low grade non-Hodgkin's lymphoma. The Scotland and Newcastle Lymphoma Group Therapy Working Party." Ann Oncol **2**(9): 655-62.

Lessard, J., J. I. Wu, J. A. Ranish, M. Wan, M. M. Winslow, B. T. Staahl, H. Wu, R. Aebersold, I. A. Graef and G. R. Crabtree (2007). "An essential switch in subunit composition of a chromatin remodeling complex during neural development." Neuron **55**(2): 201-15.

Levine, A. J., J. Momand and C. A. Finlay (1991). "The p53 tumour suppressor gene." Nature **351**(6326): 453-6.

Ley, T. J., E. R. Mardis, L. Ding, B. Fulton, M. D. McLellan, K. Chen, D. Dooling, B. H. Dunford-Shore, S. McGrath, M. Hickenbotham, L. Cook, R. Abbott, D. E. Larson, D. C. Koboldt, C. Pohl, S. Smith, A. Hawkins, S. Abbott, D. Locke, L. W. Hillier, T. Miner, L. Fulton, V. Magrini, T. Wylie, J. Glasscock, J. Conyers, N. Sander, X. Shi, J. R. Osborne, P. Minx, D. Gordon, A. Chinwalla, Y. Zhao, R. E. Ries, J. E. Payton, P. Westervelt, M. H. Tomasson, M. Watson, J. Baty, J. Ivanovich, S. Heath, W. D. Shannon, R. Nagarajan, M. J. Walter, D. C. Link, T. A. Graubert, J. F. DiPersio and R. K. Wilson (2008). "DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome." Nature **456**(7218): 66-72.

Lieber, M. R., K. Yu and S. C. Raghavan (2006). "Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations." DNA Repair (Amst) **5**(9-10): 1234-45.

Lill, N. L., S. R. Grossman, D. Ginsberg, J. DeCaprio and D. M. Livingston (1997). "Binding and modulation of p53 by p300/CBP coactivators." Nature **387**(6635): 823-7.

Limpens, J., R. Stad, C. Vos, C. de Vlaam, D. de Jong, G. J. van Ommen, E. Schuuring and P. M. Kluin (1995). "Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals." Blood **85**(9): 2528-36.

Lin, P., R. Jetly, P. A. Lennon, L. V. Abruzzo, S. Prajapati and L. J. Medeiros (2008). "Translocation (18;22)(q21;q11) in B-cell lymphomas: a report of 4 cases and review of the literature." Hum Pathol **39**(11): 1664-72.

Lincz, L. F., F. E. Scorgie, I. Kerridge, R. Potts, A. Spencer and A. Enno (2003). "Methionine synthase genetic polymorphism MS A2756G alters susceptibility to follicular but not diffuse large B-cell non-Hodgkin's lymphoma or multiple myeloma." Br J Haematol **120**(6): 1051-4.

Linet, M. S. and L. M. Pottern (1992). "Familial aggregation of hematopoietic malignancies and risk of non-Hodgkin's lymphoma." Cancer Res **52**(19 Suppl): 5468s-5473s.

Lister, T. A., M. H. Cullen, M. E. Beard, R. L. Brearley, J. M. Whitehouse, P. F. Wrigley, A. G. Stansfeld, S. B. Sutcliffe, J. S. Malpas and D. Crowther (1978). "Comparison of combined and single-agent chemotherapy in non-Hodgkin's lymphoma of favourable histological type." Br Med J **1**(6112): 533-7.

Liu, Q., L. Fayad, F. Cabanillas, F. B. Hagemeister, G. D. Ayers, M. Hess, J. Romaguera, M. A. Rodriguez, A. M. Tsimberidou, S. Verstovsek, A. Younes, B. Pro, M. S. Lee, A. Ayala and P. McLaughlin (2006). "Improvement of overall and failure-free survival in stage IV follicular lymphoma: 25 years of treatment experience at The University of Texas M.D. Anderson Cancer Center." J Clin Oncol **24**(10): 1582-9.

Liu, Y. J., D. E. Joshua, G. T. Williams, C. A. Smith, J. Gordon and I. C. MacLennan (1989). "Mechanism of antigen-driven selection in germinal centres." Nature **342**(6252): 929-31.

Lo Coco, F., G. Gaidano, D. C. Louie, K. Offit, R. S. Chaganti and R. Dalla-Favera (1993). "p53 mutations are associated with histologic transformation of follicular lymphoma." Blood **82**(8): 2289-95.

Lopez-Guillermo, A., E. Montserrat, F. Bosch, M. J. Terol, E. Campo and C. Rozman (1994). "Applicability of the International Index for aggressive lymphomas to patients with low-grade lymphoma." J Clin Oncol **12**(7): 1343-8.

Lossos, I. S., A. A. Alizadeh, M. Diehn, R. Warnke, Y. Thorstenson, P. J. Oefner, P. O. Brown, D. Botstein and R. Levy (2002). "Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes." Proc Natl Acad Sci U S A **99**(13): 8886-91.

Lossos, I. S., C. D. Jones, J. L. Zehnder and R. Levy (2001). "A polymorphism in the BCL-6 gene is associated with follicle center lymphoma." Leuk Lymphoma **42**(6): 1343-50.

Lossos, I. S. and R. Levy (2000). "Higher-grade transformation of follicle center lymphoma is associated with somatic mutation of the 5' noncoding regulatory region of the BCL-6 gene." Blood **96**(2): 635-9.

Lowe, T., J. Sharefkin, S. Q. Yang and C. W. Dieffenbach (1990). "A computer program for selection of oligonucleotide primers for polymerase chain reactions." Nucleic Acids Res **18**(7): 1757-61.

Lu, D., J. D. Thompson, G. K. Gorski, N. R. Rice, M. G. Mayer and J. J. Yunis (1991). "Alterations at the rel locus in human lymphoma." Oncogene **6**(7): 1235-41.

Lukes, R. J. and R. D. Collins (1974). "Immunologic characterization of human malignant lymphomas." Cancer **34**(4 Suppl): suppl:1488-503.

MacLennan, I. C. (1994). "Germinal centers." Annu Rev Immunol **12**: 117-39.

MacManus, M. P. and R. T. Hoppe (1996). "Is radiotherapy curative for stage I and II low-grade follicular lymphoma? Results of a long-term follow-up study of patients treated at Stanford University." J Clin Oncol **14**(4): 1282-90.

Mahe, B., N. Milpied, M. P. Mellerin, P. Moreau, N. Morineau, M. Vigier and J. L. Harousseau (2003). "PCR detection of residual Bcl-2/IgH-positive cells after high-dose therapy with autologous stem cell transplantation is a prognostic factor for event-free survival in patients with low-grade follicular non-Hodgkin's lymphoma." Bone Marrow Transplant **31**(6): 467-73.

Mahmoodi, M., S. S. Tanev, H. H. Punnett, P. Crilley and J. S. Hou (2004). "Translocation (16;18)(p13;q21.3) in follicular lymphoma." Cancer Genet Cytogenet **154**(2): 160-2.

Manabe, R., K. Tsutsui, T. Yamada, M. Kimura, I. Nakano, C. Shimono, N. Sanzen, Y. Furutani, T. Fukuda, Y. Oguri, K. Shimamoto, D. Kiyozumi, Y. Sato, Y. Sado, H. Senoo, S. Yamashina, S. Fukuda, J. Kawai, N. Sugiura, K. Kimata, Y. Hayashizaki and K. Sekiguchi (2008). "Transcriptome-based systematic identification of extracellular matrix proteins." Proc Natl Acad Sci U S A **105**(35): 12849-54.

Mann, R. B. and C. W. Berard (1983). "Criteria for the cytologic subclassification of follicular lymphomas: a proposed alternative method." Hematol Oncol **1**(2): 187-92.

Marco, F., R. Manjon, C. Richard, F. Mazorra, A. Garcia-Valtuille, M. D. Delgado, M. L. Loyo, M. A. Cuadrado and A. Zubizarreta (1999). "Simultaneous occurrence of follicular lymphoma in two monozygotic twins." Br J Haematol **107**(2): 461-2.

Marcus, R., K. Imrie, A. Belch, D. Cunningham, E. Flores, J. Catalano, P. Solal-Celigny, F. Offner, J. Walewski, J. Raposo, A. Jack and P. Smith (2005). "CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma." Blood **105**(4): 1417-23.

Marcus, R., K. Imrie, P. Solal-Celigny, J. V. Catalano, A. Dmoszynska, J. C. Raposo, F. C. Offner, J. Gomez-Codina, A. Belch, D. Cunningham, E. Wassner-Fritsch and G. Stein (2008). "Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma." J Clin Oncol **26**(28): 4579-86.

Mardis, E. R., L. Ding, D. J. Dooling, D. E. Larson, M. D. McLellan, K. Chen, D. C. Koboldt, R. S. Fulton, K. D. Delehaunty, S. D. McGrath, L. A. Fulton, D. P. Locke, V. J. Magrini, R. M. Abbott, T. L. Vickery, J. S. Reed, J. S. Robinson, T. Wylie, S. M. Smith, L. Carmichael, J. M. Eldred, C. C. Harris, J. Walker, J. B. Peck, F. Du, A. F. Dukes, G. E. Sanderson, A. M. Brummett, E. Clark, J. F. McMichael, R. J. Meyer, J. K. Schindler, C. S. Pohl, J. W. Wallis, X. Shi, L. Lin, H. Schmidt, Y. Tang, C. Haipek, M. E. Wiechert, J. V. Ivy, J. Kalicki, G. Elliott, R. E. Ries, J. E. Payton, P. Westervelt, M. H. Tomasson, M. A. Watson, J. Baty, S. Heath, W. D. Shannon, R. Nagarajan, D. C. Link, M. J. Walter, T. A. Graubert, J. F. DiPersio, R. K. Wilson and T. J. Ley (2009).

"Recurring mutations found by sequencing an acute myeloid leukemia genome." N Engl J Med **361**(11): 1058-66.

Marin, M. C., C. A. Jost, L. A. Brooks, M. S. Irwin, J. O'Nions, J. A. Tidy, N. James, J. M. McGregor, C. A. Harwood, I. G. Yulug, K. H. Vousden, M. J. Allday, B. Gusterson, S. Ikawa, P. W. Hinds, T. Crook and W. G. Kaelin, Jr. (2000). "A common polymorphism acts as an intragenic modifier of mutant p53 behaviour." Nat Genet **25**(1): 47-54.

Marsh, S. G., E. D. Albert, W. F. Bodmer, R. E. Bontrop, B. Dupont, H. A. Erlich, M. Fernandez-Vina, D. E. Geraghty, R. Holdsworth, C. K. Hurley, M. Lau, K. W. Lee, B. Mach, M. Maiers, W. R. Mayr, C. R. Muller, P. Parham, E. W. Petersdorf, T. Sasazuki, J. L. Strominger, A. Svejgaard, P. I. Terasaki, J. M. Tiercy and J. Trowsdale (2010). "Nomenclature for factors of the HLA system, 2010." Tissue Antigens **75**(4): 291-455.

Martin-Subero, J. I., O. Ammerpohl, M. Bibikova, E. Wickham-Garcia, X. Agirre, S. Alvarez, M. Bruggemann, S. Bug, M. J. Calasanz, M. Deckert, M. Dreyling, M. Q. Du, J. Durig, M. J. Dyer, J. B. Fan, S. Gesk, M. L. Hansmann, L. Harder, S. Hartmann, W. Klapper, R. Kuppers, M. Montesinos-Rongen, I. Nagel, C. Pott, J. Richter, J. Roman-Gomez, M. Seifert, H. Stein, J. Suela, L. Trumper, I. Vater, F. Prosper, C. Haferlach, J. Cruz Cigudosa and R. Siebert (2009a). "A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms." PLoS One **4**(9): e6986.

Martin-Subero, J. I., M. Kreuz, M. Bibikova, S. Bentink, O. Ammerpohl, E. Wickham-Garcia, M. Rosolowski, J. Richter, L. Lopez-Serra, E. Ballestar, H. Berger, X. Agirre, H. W. Bernd, V. Calvanese, S. B. Cogliatti, H. G. Drexler, J. B. Fan, M. F. Fraga, M. L. Hansmann, M. Hummel, W. Klapper, B. Korn, R. Kuppers, R. A. Macleod, P. Moller, G. Ott, C. Pott, F. Prosper, A. Rosenwald, C. Schwaenen, D. Schubeler, M. Seifert, B. Sturzenhofecker, M. Weber, S. Wessendorf, M. Loeffler, L. Trumper, H. Stein, R. Spang, M. Esteller, D. Barker, D. Hasenclever and R. Siebert (2009b). "New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling." Blood **113**(11): 2488-97.

Martinelli, G., S. F. Schmitz, U. Utiger, T. Cerny, U. Hess, S. Bassi, E. Okkinga, R. Stupp, R. Stahel, M. Heizmann, D. Vorobiof, A. Lohri, P. Y. Dietrich, E. Zucca and M. Ghielmini (2010). "Long-term follow-up of patients with follicular lymphoma receiving single-agent rituximab at two different schedules in trial SAKK 35/98." J Clin Oncol **28**(29): 4480-4.

Martinez-Climent, J. A., A. A. Alizadeh, R. Segraves, D. Blesa, F. Rubio-Moscardo, D. G. Albertson, J. Garcia-Conde, M. J. Dyer, R. Levy, D. Pinkel and I. S. Lossos (2003). "Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations." Blood **101**(8): 3109-17.

Martinez-Climent, J. A., L. Fontan, R. D. Gascoyne, R. Siebert and F. Prosper (2010). "Lymphoma stem cells: enough evidence to support their existence?" Haematologica **95**(2): 293-302.

Matlashewski, G. J., S. Tuck, D. Pim, P. Lamb, J. Schneider and L. V. Crawford (1987). "Primary structure polymorphism at amino acid residue 72 of human p53." Mol Cell Biol **7**(2): 961-3.

Matolcsy, A., P. Casali, R. A. Warnke and D. M. Knowles (1996). "Morphologic transformation of follicular lymphoma is associated with somatic mutation of the translocated Bcl-2 gene." Blood **88**(10): 3937-44.

Matolcsy, A., E. J. Schattner, D. M. Knowles and P. Casali (1999). "Clonal evolution of B cells in transformation from low- to high-grade lymphoma." Eur J Immunol **29**(4): 1253-64.

Matsui, W., Q. Wang, J. P. Barber, S. Brennan, B. D. Smith, I. Borrello, I. McNiece, L. Lin, R. F. Ambinder, C. Peacock, D. N. Watkins, C. A. Huff and R. J. Jones (2008). "Clonogenic multiple myeloma progenitors, stem cell properties, and drug resistance." Cancer Res **68**(1): 190-7.

Matsuo, K., N. Hamajima, R. Suzuki, M. Ogura, Y. Kagami, H. Taji, T. Yasue, N. E. Mueller, S. Nakamura, M. Seto, Y. Morishima and K. Tajima (2004). "Methylenetetrahydrofolate reductase gene (MTHFR) polymorphisms and reduced risk of malignant lymphoma." Am J Hematol **77**(4): 351-7.

McCann, K. J., C. H. Ottensmeier, A. Callard, C. M. Radcliffe, D. J. Harvey, R. A. Dwek, P. M. Rudd, B. J. Sutton, P. Hobby and F. K. Stevenson (2008). "Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma." Mol Immunol **45**(6): 1567-72.

McCarroll, S. A., F. G. Kuruvilla, J. M. Korn, S. Cawley, J. Nemesh, A. Wysoker, M. H. Shapero, P. I. de Bakker, J. B. Maller, A. Kirby, A. L. Elliott, M. Parkin, E. Hubbell, T. Webster, R. Mei, J. Veitch, P. J. Collins, R. Handsaker, S. Lincoln, M. Nizzari, J. Blume, K. W. Jones, R. Rava, M. J. Daly, S. B. Gabriel and D. Altshuler (2008). "Integrated detection and population-genetic analysis of SNPs and copy number variation." Nat Genet **40**(10): 1166-74.

McDonnell, T. J. and S. J. Korsmeyer (1991). "Progression from lymphoid hyperplasia to high-grade malignant lymphoma in mice transgenic for the t(14; 18)." Nature **349**(6306): 254-6.

McLaughlin, P., A. J. Grillo-Lopez, B. K. Link, R. Levy, M. S. Czuczman, M. E. Williams, M. R. Heyman, I. Bence-Bruckler, C. A. White, F. Cabanillas, V. Jain, A. D. Ho, J. Lister, K. Wey, D. Shen and B. K. Dallaire (1998). "Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program." J Clin Oncol **16**(8): 2825-33.

Mendez-Lago, M., Morin R. D., Mungall A. J., Chan S., Chittaranjan S., Severson T., Goya R., Mungal A.J., Johnson N. A., Boyle M., Woolcock B., Zeng T., McDonald H., An J., Yakovenko O., Tam A., Zhao Y., Hirst M., Moore R., Schein J. E., Jones S. J., Horsman D. E., Gascoyne R. D., Connors J. M. and Marra M. A. (2010). "Mutations In MLL2 and MEF2B Genes In Follicular Lymphoma and Diffuse Large B-Cell Lymphoma." (Blood (ASH Annual Meeting Abstracts), Nov 2010; 116: #473.).

Mendrysa, S. M., K. A. O'Leary, M. K. McElwee, J. Michalowski, R. N. Eisenman, D. A. Powell and M. E. Perry (2006). "Tumor suppression and normal aging in mice with constitutively high p53 activity." Genes Dev **20**(1): 16-21.

"Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer." (2011). Mitelman F., Johansson B., Mertens F., (eds). Accessable at website http://cgap.nci.nih.gov/Chromosomes/Mitelman.

Mohamed, A. N., M. Palutke, L. Eisenberg and A. Al-Katib (2001). "Chromosomal analyses of 52 cases of follicular lymphoma with t(14;18), including blastic/blastoid variant." Cancer Genet Cytogenet **126**(1): 45-51.

Mohammad, R. M., J. Wu, A. S. Azmi, A. Aboukameel, A. Sosin, S. Wu, D. Yang, S. Wang and A. M. Al-Katib (2009). "An MDM2 antagonist (MI-319) restores p53 functions and increases the life span of orally treated follicular lymphoma bearing animals." Mol Cancer **8**: 115.

Moller, M. B., O. Nielsen and N. T. Pedersen (2002). "Frequent alteration of MDM2 and p53 in the molecular progression of recurring non-Hodgkin's lymphoma." Histopathology **41**(4): 322-30.

Momand, J., D. Jung, S. Wilczynski and J. Niland (1998). "The MDM2 gene amplification database." Nucleic Acids Res **26**(15): 3453-9.

Momand, J., H. H. Wu and G. Dasgupta (2000). "MDM2--master regulator of the p53 tumor suppressor protein." Gene **242**(1-2): 15-29.

Momand, J., G. P. Zambetti, D. C. Olson, D. George and A. J. Levine (1992). "The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation." Cell **69**(7): 1237-45.

Montoto, S., C. Canals, A. Z. Rohatiner, G. Taghipour, A. Sureda, N. Schmitz, C. Gisselbrecht, L. Fouillard, N. Milpied, C. Haioun, S. Slavin, E. Conde, C. Fruchart, A. Ferrant, V. Leblond, H. Tilly, T. A. Lister and A. H. Goldstone (2007a). "Long-term follow-up of high-dose treatment with autologous haematopoietic progenitor cell support in 693 patients with follicular lymphoma: an EBMT registry study." Leukemia **21**(11): 2324-31.

Montoto, S., A. J. Davies, J. Matthews, M. Calaminici, A. J. Norton, J. Amess, S. Vinnicombe, R. Waters, A. Z. Rohatiner and T. A. Lister (2007b). "Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma." J Clin Oncol **25**(17): 2426-33.

Morin, R. D., Johnson NA, Severson TM, Mungall AJ, An J, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, Shah S, Corbett R, Tam A, Varhol R, Smailus D, Moksa M, Zhao Y, Delaney A, Qian H, Birol I, Zhu H, Kimbara M, Shashkin P, Charlot JF, Tcherpakov M, Aparicio S, Schein J, Moore R, Holt R, Horsman DE, Connors JM, Jones S, Hirst M, Gascoyne RD and Marra MA (2009). "Tyrosine 641 of the EZH2 Oncogene Is Frequently Mutated in Follicular and Diffuse Large B-Cell Lymphomas of Germinal Centre Origin." Blood (ASH Annual Meeting Abstracts), Nov 2009; 114: #139

Morin, R. D., N. A. Johnson, T. M. Severson, A. J. Mungall, J. An, R. Goya, J. E. Paul, M. Boyle, B. W. Woolcock, F. Kuchenbauer, D. Yap, R. K. Humphries, O. L. Griffith, S. Shah, H. Zhu, M. Kimbara, P. Shashkin, J. F. Charlot, M. Tcherpakov, R. Corbett, A. Tam, R. Varhol, D. Smailus, M. Moksa, Y. Zhao, A. Delaney, H. Qian, I. Birol, J. Schein, R. Moore, R. Holt, D. E. Horsman, J. M. Connors, S. Jones, S. Aparicio, M. Hirst, R. D. Gascoyne and M. A. Marra (2010a). "Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin." Nat Genet **42**(2): 181-5.

Morin, R. D., Mendez-Lago M., Mungall A.J., Johnson N.A., Goya R., Severson T., Mungall K., An J., Yakovenko O., Jackman S., Krzywinski M., Griffith M., Chan S., Tam A., Smailus D., McDonald H., Moksa M., Boyle M., Woolcock B., Zeng T., Zhao Y., Holt R. A., Moore R., Schein J. E., Birol I., Horsman D. E., Jones S. J., Connors J. M., Hirst M., Gascoyne R. D. and Marra M. A. (2010b). "Identification of Genes Frequently Mutated In FL and DLBCL with Transcriptome, Genome and Exome Sequencing." Blood (ASH Annual Meeting Abstracts), Nov 2010; 116: #804.

Mork, C. N., D. V. Faller and R. A. Spanjaard (2007). "Loss of putative tumor suppressor EI24/PIG8 confers resistance to etoposide." FEBS Lett **581**(28): 5440-4.

Morschhauser, F., J. Radford, A. Van Hoof, U. Vitolo, P. Soubeyran, H. Tilly, P. C. Huijgens, A. Kolstad, F. d'Amore, M. Gonzalez Diaz, M. Petrini, C. Sebban, P. L. Zinzani, M. H. van Oers, W. van Putten, A. Bischof-Delaloye, A. Rohatiner, G. Salles, J. Kuhlmann and A. Hagenbeek (2008). "Phase III trial of consolidation therapy with yttrium-90-ibritumomab tiuxetan compared with no additional therapy after first remission in advanced follicular lymphoma." J Clin Oncol **26**(32): 5156-64.

Morton, L. M., P. Hartge, T. R. Holford, E. A. Holly, B. C. Chiu, P. Vineis, E. Stagnaro, E. V. Willett, S. Franceschi, C. La Vecchia, A. M. Hughes, W. Cozen, S. Davis, R. K. Severson, L. Bernstein, S. T. Mayne, F. R. Dee, J. R. Cerhan and T. Zheng (2005). "Cigarette smoking and risk of non-Hodgkin lymphoma: a pooled analysis from the International Lymphoma Epidemiology Consortium (interlymph)." Cancer Epidemiol Biomarkers Prev **14**(4): 925-33.

Morton, L. M., M. P. Purdue, T. Zheng, S. S. Wang, B. Armstrong, Y. Zhang, I. Menashe, N. Chatterjee, S. Davis, Q. Lan, C. M. Vajdic, R. K. Severson, T. R. Holford, A. Kricker, J. R. Cerhan, B. Leaderer, A. Grulich, M. Yeager, W. Cozen, S. Hoar Zahm, S. J. Chanock, N. Rothman and P. Hartge (2009). "Risk of non-Hodgkin lymphoma associated with germline variation in genes that regulate the cell cycle, apoptosis, and lymphocyte development." Cancer Epidemiol Biomarkers Prev **18**(4): 1259-70.

Morton, L. M., S. S. Wang, S. S. Devesa, P. Hartge, D. D. Weisenburger and M. S. Linet (2006). "Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001." Blood **107**(1): 265-76.

Mortuza, F. Y., I. M. Moreira, M. Papaioannou, P. Gameiro, L. A. Coyle, C. S. Gricks, P. Amlot, H. G. Prentice, A. Madrigal, A. V. Hoffbrand and L. Foroni (2001). "Immunoglobulin heavy-chain gene rearrangement in adult acute lymphoblastic leukemia reveals preferential usage of J(H)-proximal variable gene segments." Blood **97**(9): 2716-26.

Muller-Hermelink, H. K. and A. Marx (2000). "Towards a histogenetic classification of thymic epithelial tumours?" Histopathology **36**(5): 466-9.

Mullighan, C. G., L. A. Phillips, X. Su, J. Ma, C. B. Miller, S. A. Shurtleff and J. R. Downing (2008). "Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia." Science **322**(5906): 1377-80.

Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai and T. Honjo (2000). "Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme." Cell **102**(5): 553-63.

Muramatsu, M., V. S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N. O. Davidson and T. Honjo (1999). "Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells." J Biol Chem **274**(26): 18470-6.

Nagaoka, H., M. Muramatsu, N. Yamamura, K. Kinoshita and T. Honjo (2002). "Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin Smu region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation." J Exp Med **195**(4): 529-34.

Nakamura, M., M. Yasunami, H. Kondo, H. Horie, Y. Aiba, A. Komori, K. Migita, H. Yatsuhashi, M. Ito, S. Shimoda and H. Ishibashi (2010). "Analysis of HLA-DRB1 polymorphisms in Japanese patients with primary biliary cirrhosis (PBC): The HLA-DRB1polymorphism determines the relative risk of antinuclear antibodies for disease progression in PBC." Hepatol Res **40**(5): 494-504.

"National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a working formulation for clinical usage. The Non-Hodgkin's Lymphoma Pathologic Classification Project." (1982). Cancer **49**(10): 2112-35.

Neat, M. J., N. Foot, M. Jenner, L. Goff, K. Ashcroft, D. Burford, A. Dunham, A. Norton, T. A. Lister and J. Fitzgibbon (2001). "Localisation of a novel region of recurrent amplification in follicular lymphoma to an approximately 6.8 Mb region of 13q32-33." Genes Chromosomes Cancer **32**(3): 236-43.

Neira, M., V. Danilova, G. Hellekant and E. A. Azen (2001). "A new gene (rmSTG) specific for taste buds is found by laser capture microdissection." Mamm Genome **12**(1): 60-6.

Nejda, N., D. Iglesias, M. Moreno Azcoita, V. Medina Arana, J. J. Gonzalez-Aguilera and A. M. Fernandez-Peralta (2009). "A MLH1 polymorphism that increases cancer risk is associated with better outcome in sporadic colorectal cancer." Cancer Genet Cytogenet **193**(2): 71-7.

Ngan, B. Y., J. Nourse and M. L. Cleary (1989). "Detection of chromosomal translocation t(14;18) within the minor cluster region of bcl-2 by polymerase chain reaction and direct genomic sequencing of the enzymatically amplified DNA in follicular lymphomas." Blood **73**(7): 1759-62.

Niens, M., R. F. Jarrett, B. Hepkema, I. M. Nolte, A. Diepstra, M. Platteel, N. Kouprie, C. P. Delury, A. Gallagher, L. Visser, S. Poppema, G. J. te Meerman and A. van den Berg (2007). "HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma." Blood **110**(9): 3310-5.

Nieters, A., L. Beckmann, E. Deeg and N. Becker (2006). "Gene polymorphisms in Toll-like receptors, interleukin-10, and interleukin-10 receptor alpha and lymphoma risk." Genes Immun **7**(8): 615-24.

Nieters, A., P. M. Bracci, S. de Sanjose, N. Becker, M. Maynadie, Y. Benavente, L. Foretova, P. Cocco, A. Staines, E. A. Holly, P. Boffetta, P. Brennan and C. F. Skibola (2010). "A functional TNFRSF5 polymorphism and risk of non-Hodgkin lymphoma, a pooled analysis." Int J Cancer. Prepublished on June 14, 2010 as DOI 10.1002/ijc.25420.

"Nomenclature for factors of the HLA system, 1987." (1988). Vox Sang **55**(2): 119-26.

Notta, F., C. G. Mullighan, J. C. Wang, A. Poeppl, S. Doulatov, L. A. Phillips, J. Ma, M. D. Minden, J. R. Downing and J. E. Dick (2011). "Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells." Nature **469**(7330): 362-7.

Novak, A. J., S. L. Slager, Z. S. Fredericksen, A. H. Wang, M. M. Manske, S. Ziesmer, M. Liebow, W. R. Macon, S. R. Dillon, T. E. Witzig, J. R. Cerhan and S. M. Ansell (2009). "Genetic variation in B-cell-activating factor is associated with an increased risk of developing B-cell non-Hodgkin lymphoma." Cancer Res **69**(10): 4217-24.

Novik, K. L., J. J. Spinelli, A. C. Macarthur, K. Shumansky, P. Sipahimalani, S. Leach, A. Lai, J. M. Connors, R. D. Gascoyne, R. P. Gallagher and A. R. Brooks-Wilson (2007). "Genetic variation in H2AFX contributes to risk of non-Hodgkin lymphoma." Cancer Epidemiol Biomarkers Prev **16**(6): 1098-106.

Nowak, J., E. Kalinka-Warzocha, P. Juszczynski, P. Bilinski, R. Mika-Witkowska, M. Zajko, J. Bienvenu, B. Coiffier, G. Salles and K. Warzocha (2007). "Association of human leukocyte antigen ancestral haplotype 8.1 with adverse outcome of non-Hodgkin's lymphoma." Genes Chromosomes Cancer **46**(5): 500-7.

Nunez, G., M. Seto, S. Seremetis, D. Ferrero, F. Grignani, S. J. Korsmeyer and R. Dalla-Favera (1989). "Growth- and tumor-promoting effects of deregulated BCL2 in human B-lymphoblastoid cells." Proc Natl Acad Sci U S A **86**(12): 4589-93.

O'Brien, C. A., A. Pollett, S. Gallinger and J. E. Dick (2007). "A human colon cancer cell capable of initiating tumour growth in immunodeficient mice." Nature **445**(7123): 106-10.

O'Brien, M. E., P. Easterbrook, J. Powell, G. R. Blackledge, L. Jones, I. C. MacLennan and R. C. Leonard (1991). "The natural history of low grade non-Hodgkin's lymphoma and the impact of a no initial treatment policy on survival." Q J Med **80**(292): 651-60.

O'Riain, C., D. M. O'Shea, Y. Yang, R. Le Dieu, J. G. Gribben, K. Summers, J. Yeboah-Afari, L. Bhaw-Rosun, C. Fleischmann, C. A. Mein, T. Crook, P. Smith, G. Kelly, A. Rosenwald, G. Ott, E. Campo, L. M. Rimsza, E. B. Smeland, W. C. Chan, N. Johnson, R. D. Gascoyne, S. Reimer, R. M. Braziel, G. W. Wright, L. M. Staudt, T. A. Lister and J. Fitzgibbon (2009). "Array-based DNA methylation profiling in follicular lymphoma." Leukemia **23**(10): 1858-66.

O'Shea, D., C. O'Riain, M. Gupta, R. Waters, Y. Yang, D. Wrench, J. Gribben, A. Rosenwald, G. Ott, L. M. Rimsza, H. Holte, J. B. Cazier, N. A. Johnson, E. Campo, W. C. Chan, R. D. Gascoyne, B. D. Young, L. M. Staudt, T. A. Lister and J. Fitzgibbon (2009). "Regions of acquired uniparental disomy at diagnosis of follicular lymphoma are associated with both overall survival and risk of transformation." Blood **113**(10): 2298-301.

O'Shea, D., C. O'Riain, C. Taylor, R. Waters, E. Carlotti, F. Macdougall, J. Gribben, A. Rosenwald, G. Ott, L. M. Rimsza, E. B. Smeland, N. Johnson, E. Campo, T. C. Greiner, W. C. Chan, R. D. Gascoyne, G. Wright, L. M. Staudt, T. A. Lister and J. Fitzgibbon (2008). "The presence of TP53 mutation at diagnosis of follicular lymphoma identifies a high-risk group of patients with shortened time to disease progression and poorer overall survival." Blood **112**(8): 3126-9.

Oda, E., R. Ohki, H. Murasawa, J. Nemoto, T. Shibue, T. Yamashita, T. Tokino, T. Taniguchi and N. Tanaka (2000a). "Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis." Science **288**(5468): 1053-8.

Oda, K., H. Arakawa, T. Tanaka, K. Matsuda, C. Tanikawa, T. Mori, H. Nishimori, K. Tamai, T. Tokino, Y. Nakamura and Y. Taya (2000b). "p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53." Cell **102**(6): 849-62.

Offit, K., S. C. Jhanwar, M. Ladanyi, D. A. Filippa and R. S. Chaganti (1991). "Cytogenetic analysis of 434 consecutively ascertained specimens of non-Hodgkin's lymphoma: correlations between recurrent aberrations, histology, and exposure to cytotoxic treatment." Genes Chromosomes Cancer **3**(3): 189-201.

Ohkubo, S., T. Tanaka, Y. Taya, K. Kitazato and C. Prives (2006). "Excess HDM2 impacts cell cycle and apoptosis and has a selective effect on p53-dependent transcription." J Biol Chem **281**(25): 16943-50.

Oka, A., G. Tamiya, M. Tomizawa, M. Ota, Y. Katsuyama, S. Makino, T. Shiina, M. Yoshitome, M. Iizuka, Y. Sasao, K. Iwashita, Y. Kawakubo, J. Sugai, A. Ozawa, M. Ohkido, M. Kimura, S. Bahram and H. Inoko (1999). "Association analysis using refined microsatellite markers localizes a susceptibility locus for psoriasis vulgaris within a 111 kb segment telomeric to the HLA-C gene." Hum Mol Genet **8**(12): 2165-70.

Oken, M. M., R. H. Creech, D. C. Tormey, J. Horton, T. E. Davis, E. T. McFadden and P. P. Carbone (1982). "Toxicity and response criteria of the Eastern Cooperative Oncology Group." Am J Clin Oncol **5**(6): 649-55.

Oliner, J. D., K. W. Kinzler, P. S. Meltzer, D. L. George and B. Vogelstein (1992). "Amplification of a gene encoding a p53-associated protein in human sarcomas." Nature **358**(6381): 80-3.

Oliner, J. D., J. A. Pietenpol, S. Thiagalingam, J. Gyuris, K. W. Kinzler and B. Vogelstein (1993). "Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53." Nature **362**(6423): 857-60.

Olive, V., I. Jiang and L. He (2010). "mir-17-92, a cluster of miRNAs in the midst of the cancer network." Int J Biochem Cell Biol **42**(8): 1348-54.

Onel, K. and C. Cordon-Cardo (2004). "MDM2 and prognosis." Mol Cancer Res **2**(1): 1-8.

Ota, A., H. Tagawa, S. Karnan, S. Tsuzuki, A. Karpas, S. Kira, Y. Yoshida and M. Seto (2004). "Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma." Cancer Res **64**(9): 3087-95.

Otsuki, T., T. Yano, H. M. Clark, C. Bastard, J. P. Kerckaert, E. S. Jaffe and M. Raffeld (1995). "Analysis of LAZ3 (BCL-6) status in B-cell non-Hodgkin's lymphomas: results of rearrangement and gene expression studies and a mutational analysis of coding region sequences." Blood **85**(10): 2877-84.

Ott, G., T. Katzenberger, A. Lohr, S. Kindelberger, T. Rudiger, M. Wilhelm, J. Kalla, A. Rosenwald, J. G. Muller, M. M. Ott and H. K. Muller-Hermelink (2002). "Cytomorphologic, immunohistochemical, and cytogenetic profiles of follicular lymphoma: 2 types of follicular lymphoma grade 3." Blood **99**(10): 3806-12.

Ottensmeier, C. H., A. R. Thompsett, D. Zhu, B. S. Wilkins, J. W. Sweetenham and F. K. Stevenson (1998). "Analysis of VH genes in follicular and diffuse lymphoma shows ongoing somatic mutation and multiple isotype transcripts in early disease with changes during disease progression." Blood **91**(11): 4292-9.

Parham, P. and T. Ohta (1996). "Population biology of antigen presentation by MHC class I molecules." Science **272**(5258): 67-74.

Pasqualucci, L., G. Bhagat, M. Jankovic, M. Compagno, P. Smith, M. Muramatsu, T. Honjo, H. C. Morse, 3rd, M. C. Nussenzweig and R. Dalla-Favera (2008).

"AID is required for germinal center-derived lymphomagenesis." Nat Genet **40**(1): 108-12.

Pasqualucci, L., D. Dominguez-Sola, A. Chiarenza, G. Fabbri, A. Grunn, V. Trifonov, L. H. Kasper, S. Lerach, H. Tang, J. Ma, D. Rossi, A. Chadburn, V. V. Murty, C. G. Mullighan, G. Gaidano, R. Rabadan, P. K. Brindle and R. Dalla-Favera (2011). "Inactivating mutations of acetyltransferase genes in B-cell lymphoma." Nature **471**(7337): 189-95.

Patel, B., L. Rai, G. Buck, S. M. Richards, Y. Mortuza, W. Mitchell, G. Gerrard, A. V. Moorman, V. Duke, A. V. Hoffbrand, A. K. Fielding, A. H. Goldstone and L. Foroni (2010). "Minimal residual disease is a significant predictor of treatment failure in non T-lineage adult acute lymphoblastic leukaemia: final results of the international trial UKALL XII/ECOG2993." Br J Haematol **148**(1): 80-9.

Petitjean, A., E. Mathe, S. Kato, C. Ishioka, S. V. Tavtigian, P. Hainaut and M. Olivier (2007). "Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database." Hum Mutat **28**(6): 622-9.

Petrij, F., R. H. Giles, H. G. Dauwerse, J. J. Saris, R. C. Hennekam, M. Masuno, N. Tommerup, G. J. van Ommen, R. H. Goodman, D. J. Peters and et al. (1995). "Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP." Nature **376**(6538): 348-51.

Phelps, M., M. Darley, J. N. Primrose and J. P. Blaydes (2003). "p53-independent activation of the hdm2-P2 promoter through multiple transcription factor response elements results in elevated hdm2 expression in estrogen receptor alpha-positive breast cancer cells." Cancer Res **63**(10): 2616-23.

Picksley, S. M., B. Vojtesek, A. Sparks and D. P. Lane (1994). "Immunochemical analysis of the interaction of p53 with MDM2;--fine mapping of the MDM2 binding site on p53 using synthetic peptides." Oncogene **9**(9): 2523-9.

Pierce, L. M., L. Sivaraman, W. Chang, A. Lum, T. Donlon, A. Seifried, L. R. Wilkens, A. F. Lau and L. Le Marchand (2000). "Relationships of TP53 codon 72 and HRAS1 polymorphisms with lung cancer risk in an ethnically diverse population." Cancer Epidemiol Biomarkers Prev **9**(11): 1199-204.

Pinkel, D., R. Segraves, D. Sudar, S. Clark, I. Poole, D. Kowbel, C. Collins, W. L. Kuo, C. Chen, Y. Zhai, S. H. Dairkee, B. M. Ljung, J. W. Gray and D. G. Albertson (1998). "High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays." Nat Genet **20**(2): 207-11.

Pinyol, M., F. Cobo, S. Bea, P. Jares, I. Nayach, P. L. Fernandez, E. Montserrat, A. Cardesa and E. Campo (1998). "p16(INK4a) gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas." Blood **91**(8): 2977-84.

Portlock, C. S. and S. A. Rosenberg (1976). "Combination chemotherapy with cyclophosphamide, vincristine, and prednisone in advanced non-Hodgkin's lymphomas." Cancer **37**(3): 1275-82.

Portlock, C. S. and S. A. Rosenberg (1979). "No initial therapy for stage III and IV non-Hodgkin's lymphomas of favorable histologic types." Ann Intern Med **90**(1): 10-3.

Post, S. M., A. Quintas-Cardama, V. Pant, T. Iwakuma, A. Hamir, J. G. Jackson, D. R. Maccio, G. L. Bond, D. G. Johnson, A. J. Levine and G. Lozano (2010).

"A high-frequency regulatory polymorphism in the p53 pathway accelerates tumor development." Cancer Cell **18**(3): 220-30.

"A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project." (1993). N Engl J Med **329**(14): 987-94.

Pricolo, V. E., S. D. Finkelstein, T. T. Wu, G. Keller, A. Bakker, P. A. Swalsky and K. I. Bland (1996). "Prognostic value of TP53 and K-ras-2 mutational analysis in stage III carcinoma of the colon." Am J Surg **171**(1): 41-6.

Pritsch, O., X. Troussard, C. Magnac, F. R. Mauro, F. Davi, B. Payelle-Brogard, G. Dumas, M. Pulik, F. Clerget, F. Mandelli, N. Chiorazzi, H. W. Schroeder, Jr., M. Leporrier and G. Dighiero (1999). "VH gene usage by family members affected with chronic lymphocytic leukaemia." Br J Haematol **107**(3): 616-24.

Purdue, M. P., Q. Lan, A. Kricker, A. E. Grulich, C. M. Vajdic, J. Turner, D. Whitby, S. Chanock, N. Rothman and B. K. Armstrong (2007). "Polymorphisms in immune function genes and risk of non-Hodgkin lymphoma: findings from the New South Wales non-Hodgkin Lymphoma Study." Carcinogenesis **28**(3): 704-12.

Racila, E., B. K. Link, W. K. Weng, T. E. Witzig, S. Ansell, M. J. Maurer, J. Huang, C. Dahle, A. Halwani, R. Levy and G. J. Weiner (2008). "A polymorphism in the complement component C1qA correlates with prolonged response following rituximab therapy of follicular lymphoma." Clin Cancer Res **14**(20): 6697-703.

Raghavan, S. C., P. C. Swanson, X. Wu, C. L. Hsieh and M. R. Lieber (2004). "A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex." Nature **428**(6978): 88-93.

Ragnarsson, G., G. Eiriksdottir, J. T. Johannsdottir, J. G. Jonasson, V. Egilsson and S. Ingvarsson (1999). "Loss of heterozygosity at chromosome 1p in different solid human tumours: association with survival." Br J Cancer **79**(9-10): 1468-74.

Rajewsky, K. (1996). "Clonal selection and learning in the antibody system." Nature **381**(6585): 751-8.

Ramensky, V., P. Bork and S. Sunyaev (2002). "Human non-synonymous SNPs: server and survey." Nucleic Acids Res **30**(17): 3894-900.

Redon, R., S. Ishikawa, K. R. Fitch, L. Feuk, G. H. Perry, T. D. Andrews, H. Fiegler, M. H. Shapero, A. R. Carson, W. Chen, E. K. Cho, S. Dallaire, J. L. Freeman, J. R. Gonzalez, M. Gratacos, J. Huang, D. Kalaitzopoulos, D. Komura, J. R. MacDonald, C. R. Marshall, R. Mei, L. Montgomery, K. Nishimura, K. Okamura, F. Shen, M. J. Somerville, J. Tchinda, A. Valsesia, C. Woodwark, F. Yang, J. Zhang, T. Zerjal, L. Armengol, D. F. Conrad, X. Estivill, C. Tyler-Smith, N. P. Carter, H. Aburatani, C. Lee, K. W. Jones, S. W. Scherer and M. E. Hurles (2006). "Global variation in copy number in the human genome." Nature **444**(7118): 444-54.

Reed, J. C., M. Cuddy, T. Slabiak, C. M. Croce and P. C. Nowell (1988). "Oncogenic potential of bcl-2 demonstrated by gene transfer." Nature **336**(6196): 259-61.

Reifenberger, G., L. Liu, K. Ichimura, E. E. Schmidt and V. P. Collins (1993). "Amplification and overexpression of the MDM2 gene in a subset of human malignant gliomas without p53 mutations." Cancer Res **53**(12): 2736-9.

Relander, T., N. A. Johnson, P. Farinha, J. M. Connors, L. H. Sehn and R. D. Gascoyne (2010). "Prognostic factors in follicular lymphoma." J Clin Oncol **28**(17): 2902-13.

Renouf, B., E. Hollville, A. Pujals, C. Tetaud, J. Garibal and J. Wiels (2009). "Activation of p53 by MDM2 antagonists has differential apoptotic effects on Epstein-Barr virus (EBV)-positive and EBV-negative Burkitt's lymphoma cells." Leukemia **23**(9): 1557-63.

Reya, T., S. J. Morrison, M. F. Clarke and I. L. Weissman (2001). "Stem cells, cancer, and cancer stem cells." Nature **414**(6859): 105-11.

Ricci-Vitiani, L., D. G. Lombardi, E. Pilozzi, M. Biffoni, M. Todaro, C. Peschle and R. De Maria (2007). "Identification and expansion of human colon-cancer-initiating cells." Nature **445**(7123): 111-5.

Richards, M. A., W. M. Gregory, P. A. Hall, H. S. Dhaliwal, J. Fernandez, A. G. Stansfeld, A. E. Jones and T. A. Lister (1989). "Management of localized non-Hodgkin's lymphoma: the experience at St. Bartholomew's Hospital 1972-1985." Hematol Oncol **7**(1): 1-18.

Riemersma, S. A., E. S. Jordanova, G. W. Haasnoot, J. Drabbels, E. Schuuring, G. M. Schreuder and P. M. Kluin (2006). "The relationship between HLA class II polymorphisms and somatic deletions in testicular B cell lymphomas of Dutch patients." Hum Immunol **67**(4-5): 303-10.

Ries, S., C. Biederer, D. Woods, O. Shifman, S. Shirasawa, T. Sasazuki, M. McMahon, M. Oren and F. McCormick (2000). "Opposing effects of Ras on p53: transcriptional activation of mdm2 and induction of p19ARF." Cell **103**(2): 321-30.

Ringshausen, I., C. C. O'Shea, A. J. Finch, L. B. Swigart and G. I. Evan (2006). "Mdm2 is critically and continuously required to suppress lethal p53 activity in vivo." Cancer Cell **10**(6): 501-14.

"Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma." (2008). National Institute for Health and Clinical Excellence: London.

Rohatiner, A. Z., L. Nadler, A. J. Davies, J. Apostolidis, D. Neuberg, J. Matthews, J. G. Gribben, P. M. Mauch, T. A. Lister and A. S. Freedman (2007). "Myeloablative therapy with autologous bone marrow transplantation for follicular lymphoma at the time of second or subsequent remission: long-term follow-up." J Clin Oncol **25**(18): 2554-9.

Roix, J. J., P. G. McQueen, P. J. Munson, L. A. Parada and T. Misteli (2003). "Spatial proximity of translocation-prone gene loci in human lymphomas." Nat Genet **34**(3): 287-91.

Rollinson, S., H. Kesby and G. J. Morgan (2006). "Haplotypic variation in MRE11, RAD50 and NBS1 and risk of non-Hodgkin's lymphoma." Leuk Lymphoma **47**(12): 2567-83.

Romaguera, J. E., P. McLaughlin, L. North, D. Dixon, K. B. Silvermintz, L. A. Garnsey, W. S. Velasquez, F. B. Hagemeister and F. Cabanillas (1991). "Multivariate analysis of prognostic factors in stage IV follicular low-grade lymphoma: a risk model." J Clin Oncol **9**(5): 762-9.

Rosenbloom, K. R., T. R. Dreszer, M. Pheasant, G. P. Barber, L. R. Meyer, A. Pohl, B. J. Raney, T. Wang, A. S. Hinrichs, A. S. Zweig, P. A. Fujita, K. Learned, B. Rhead, K. E. Smith, R. M. Kuhn, D. Karolchik, D. Haussler and W. J. Kent (2010). "ENCODE whole-genome data in the UCSC Genome Browser." Nucleic Acids Res **38**(Database issue): D620-5.

Roses, A. D. (2000). "Pharmacogenetics and the practice of medicine." Nature **405**(6788): 857-65.

Ross, C. W., P. D. Ouillette, C. M. Saddler, K. A. Shedden and S. N. Malek (2007). "Comprehensive analysis of copy number and allele status identifies multiple chromosome defects underlying follicular lymphoma pathogenesis." Clin Cancer Res **13**(16): 4777-85.

Rossi, D., E. Berra, M. Cerri, C. Deambrogi, C. Barbieri, S. Franceschetti, M. Lunghi, A. Conconi, M. Paulli, A. Matolcsy, L. Pasqualucci, D. Capello and G. Gaidano (2006). "Aberrant somatic hypermutation in transformation of follicular lymphoma and chronic lymphocytic leukemia to diffuse large B-cell lymphoma." Haematologica **91**(10): 1405-9.

Roth, J., M. Dobbelstein, D. A. Freedman, T. Shenk and A. J. Levine (1998). "Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein." Embo J **17**(2): 554-64.

Rothman, N., C. F. Skibola, S. S. Wang, G. Morgan, Q. Lan, M. T. Smith, J. J. Spinelli, E. Willett, S. De Sanjose, P. Cocco, S. I. Berndt, P. Brennan, A. Brooks-Wilson, S. Wacholder, N. Becker, P. Hartge, T. Zheng, E. Roman, E. A. Holly, P. Boffetta, B. Armstrong, W. Cozen, M. Linet, F. X. Bosch, M. G. Ennas, T. R. Holford, R. P. Gallagher, S. Rollinson, P. M. Bracci, J. R. Cerhan, D. Whitby, P. S. Moore, B. Leaderer, A. Lai, C. Spink, S. Davis, R. Bosch, A. Scarpa, Y. Zhang, R. K. Severson, M. Yeager, S. Chanock and A. Nieters (2006). "Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium." Lancet Oncol **7**(1): 27-38.

Roulland, S., P. Lebailly, Y. Lecluse, M. Briand, D. Pottier and P. Gauduchon (2004). "Characterization of the t(14;18) BCL2-IGH translocation in farmers occupationally exposed to pesticides." Cancer Res **64**(6): 2264-9.

Roulland, S., J. M. Navarro, P. Grenot, M. Milili, J. Agopian, B. Montpellier, P. Gauduchon, P. Lebailly, C. Schiff and B. Nadel (2006). "Follicular lymphoma-like B cells in healthy individuals: a novel intermediate step in early lymphomagenesis." J Exp Med **203**(11): 2425-31.

Rowley, J. D., S. Reshmi, O. Sobulo, T. Musvee, J. Anastasi, S. Raimondi, N. R. Schneider, J. C. Barredo, E. S. Cantu, B. Schlegelberger, F. Behm, N. A. Doggett, J. Borrow and N. Zeleznik-Le (1997). "All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders." Blood **90**(2): 535-41.

Ruminy, P., F. Jardin, J. M. Picquenot, F. Parmentier, N. Contentin, G. Buchonnet, S. Tison, V. Rainville, H. Tilly and C. Bastard (2008). "S(mu) mutation patterns suggest different progression pathways in follicular lymphoma: early direct or late from FL progenitor cells." Blood **112**(5): 1951-9.

Rummel, M. J., Niederle N., Maschmeyer G., Banat A., von Gruenhagen U., Losem C., Heil G., Welslau M., Balser C., Kaiser U., Ballo H., Weidmann E., Duerk H. A., Kofahl-Krause D., Roller F., Barth J., Hoelzer D., Hinke A. and Brugger W. (2009). "Bendamustine Plus Rituximab Is Superior in Respect of Progression Free Survival and CR Rate When Compared to CHOP Plus Rituximab as First-Line Treatment of Patients with Advanced Follicular, Indolent, and Mantle Cell Lymphomas: Final Results of a Randomized Phase III Study of the StiL (Study Group Indolent Lymphomas, Germany)." Blood: (ASH Annual Meeting Abstracts) Nov 2009; 110: #405.

Ryan, K. M., A. C. Phillips and K. H. Vousden (2001). "Regulation and function of the p53 tumor suppressor protein." Curr Opin Cell Biol **13**(3): 332-7.

Saigusa, S., K. Tanaka, Y. Toiyama, T. Yokoe, Y. Okugawa, Y. Ioue, C. Miki and M. Kusunoki (2009). "Correlation of CD133, OCT4, and SOX2 in rectal cancer and their association with distant recurrence after chemoradiotherapy." Ann Surg Oncol **16**(12): 3488-98.

Sakhinia, E., C. Glennie, J. A. Hoyland, L. P. Menasce, G. Brady, C. Miller, J. A. Radford and R. J. Byers (2007). "Clinical quantitation of diagnostic and predictive gene expression levels in follicular and diffuse large B-cell lymphoma by RT-PCR gene expression profiling." Blood **109**(9): 3922-8.

Salles, G., N. Mounier, S. de Guibert, F. Morschhauser, C. Doyen, J. F. Rossi, C. Haioun, P. Brice, B. Mahe, R. Bouabdallah, B. Audhuy, C. Ferme, C. Dartigeas, P. Feugier, C. Sebban, L. Xerri and C. Foussard (2008). "Rituximab combined with chemotherapy and interferon in follicular lymphoma patients: results of the GELA-GOELAMS FL2000 study." Blood **112**(13): 4824-31.

Salles, G., J. F. Seymour, F. Offner, A. Lopez-Guillermo, D. Belada, L. Xerri, P. Feugier, R. Bouabdallah, J. V. Catalano, P. Brice, D. Caballero, C. Haioun, L. M. Pedersen, A. Delmer, D. Simpson, S. Leppa, P. Soubeyran, A. Hagenbeek, O. Casasnovas, T. Intragumtornchai, C. Ferme, M. G. da Silva, C. Sebban, A. Lister, J. A. Estell, G. Milone, A. Sonet, M. Mendila, B. Coiffier and H. Tilly (2011). "Rituximab maintenance for 2 years in patients with high tumour burden follicular lymphoma responding to rituximab plus chemotherapy (PRIMA): a phase 3, randomised controlled trial." Lancet **377**(9759): 42-51.

Sanchez, F., S. J. Holm, L. Mallbris, K. P. O'Brien and M. Stahle (2004). "STG does not associate with psoriasis in the Swedish population." Exp Dermatol **13**(7): 413-8.

Sander, C. A., T. Yano, H. M. Clark, C. Harris, D. L. Longo, E. S. Jaffe and M. Raffeld (1993). "p53 mutation is associated with progression in follicular lymphomas." Blood **82**(7): 1994-2004.

Sanger, F., S. Nicklen and A. R. Coulson (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-7.

Sarek, G., S. Kurki, J. Enback, G. Iotzova, J. Haas, P. Laakkonen, M. Laiho and P. M. Ojala (2007). "Reactivation of the p53 pathway as a treatment modality for KSHV-induced lymphomas." J Clin Invest **117**(4): 1019-28.

Satterwhite, E., T. Sonoki, T. G. Willis, L. Harder, R. Nowak, E. L. Arriola, H. Liu, H. P. Price, S. Gesk, D. Steinemann, B. Schlegelberger, D. G. Oscier, R. Siebert, P. W. Tucker and M. J. Dyer (2001). "The BCL11 gene family: involvement of BCL11A in lymphoid malignancies." Blood **98**(12): 3413-20.

Savage, K. J., N. A. Johnson, S. Ben-Neriah, J. M. Connors, L. H. Sehn, P. Farinha, D. E. Horsman and R. D. Gascoyne (2009). "MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy." Blood **114**(17): 3533-7.

Schatton, T., G. F. Murphy, N. Y. Frank, K. Yamaura, A. M. Waaga-Gasser, M. Gasser, Q. Zhan, S. Jordan, L. M. Duncan, C. Weishaupt, R. C. Fuhlbrigge, T. S. Kupper, M. H. Sayegh and M. H. Frank (2008). "Identification of cells initiating human melanomas." Nature **451**(7176): 345-9.

Schlesinger, Y., R. Straussman, I. Keshet, S. Farkash, M. Hecht, J. Zimmerman, E. Eden, Z. Yakhini, E. Ben-Shushan, B. E. Reubinoff, Y. Bergman, I. Simon

and H. Cedar (2007). "Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer." Nat Genet **39**(2): 232-6.

Schouten, H. C., W. Qian, S. Kvaloy, A. Porcellini, H. Hagberg, H. E. Johnson, J. K. Doorduijn, M. R. Sydes and G. Kvalheim (2003). "High-dose therapy improves progression-free survival and survival in relapsed follicular non-Hodgkin's lymphoma: results from the randomized European CUP trial." J Clin Oncol **21**(21): 3918-27.

Schroeder, H. W., Jr., J. L. Hillson and R. M. Perlmutter (1987). "Early restriction of the human antibody repertoire." Science **238**(4828): 791-3.

Schroeder, J. C., A. F. Olshan, R. Baric, G. A. Dent, C. R. Weinberg, B. Yount, J. R. Cerhan, C. F. Lynch, L. M. Schuman, P. E. Tolbert, N. Rothman, K. P. Cantor and A. Blair (2001). "Agricultural risk factors for t(14;18) subtypes of non-Hodgkin's lymphoma." Epidemiology **12**(6): 701-9.

Schuler, F., L. Dolken, C. Hirt, T. Kiefer, T. Berg, G. Fusch, K. Weitmann, W. Hoffmann, C. Fusch, S. Janz, C. S. Rabkin and G. Dolken (2009). "Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals." Int J Cancer **124**(4): 958-63.

Schulz, H., J. Bohlius, N. Skoetz, S. Trelle, T. Kober, M. Reiser, M. Dreyling, M. Herold, G. Schwarzer, M. Hallek and A. Engert (2007). "Chemotherapy plus Rituximab versus chemotherapy alone for B-cell non-Hodgkin's lymphoma." Cochrane Database Syst Rev(4): CD003805.

Schwaenen, C., A. Viardot, H. Berger, T. F. Barth, S. Bentink, H. Dohner, M. Enz, A. C. Feller, M. L. Hansmann, M. Hummel, H. A. Kestler, W. Klapper, M. Kreuz, D. Lenze, M. Loeffler, P. Moller, H. K. Muller-Hermelink, G. Ott, M. Rosolowski, A. Rosenwald, S. Ruf, R. Siebert, R. Spang, H. Stein, L. Truemper, P. Lichter, M. Bentz and S. Wessendorf (2009). "Microarray-based genomic profiling reveals novel genomic aberrations in follicular lymphoma which associate with patient survival and gene expression status." Genes Chromosomes Cancer **48**(1): 39-54.

Scolnick, D. M., N. H. Chehab, E. S. Stavridi, M. C. Lien, L. Caruso, E. Moran, S. L. Berger and T. D. Halazonetis (1997). "CREB-binding protein and p300/CBP-associated factor are transcriptional coactivators of the p53 tumor suppressor protein." Cancer Res **57**(17): 3693-6.

Shastry, B. S. (2009). "SNPs: impact on gene function and phenotype." Methods Mol Biol **578**: 3-22.

Shen, M., T. Zheng, Q. Lan, Y. Zhang, S. H. Zahm, S. S. Wang, T. R. Holford, B. Leaderer, M. Yeager, R. Welch, D. Kang, P. Boyle, B. Zhang, K. Zou, Y. Zhu, S. Chanock and N. Rothman (2006). "Polymorphisms in DNA repair genes and risk of non-Hodgkin lymphoma among women in Connecticut." Hum Genet **119**(6): 659-68.

Shia, J., J. Teruya-Feldstein, D. Pan, A. Hegde, D. S. Klimstra, R. S. Chaganti, J. Qin, C. S. Portlock and D. A. Filippa (2002). "Primary follicular lymphoma of the gastrointestinal tract: a clinical and pathologic study of 26 cases." Am J Surg Pathol **26**(2): 216-24.

Shieh, S. Y., J. Ahn, K. Tamai, Y. Taya and C. Prives (2000). "The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites." Genes Dev **14**(3): 289-300.

Shiina, T., H. Inoko and J. K. Kulski (2004). "An update of the HLA genomic region, locus information and disease associations: 2004." Tissue Antigens **64**(6): 631-49.

Shustik, J., M. Quinn, J. M. Connors, R. D. Gascoyne, B. Skinnider and L. H. Sehn (2010). "Follicular non-Hodgkin lymphoma grades 3A and 3B have a similar outcome and appear incurable with anthracycline-based therapy." Ann Oncol(e-pub ahead of print doi:10.1093/annonc/mdq574).

Siddiqui, M., K. Ristow, S. N. Markovic, T. E. Witzig, T. M. Habermann, J. P. Colgan, D. J. Inwards, W. L. White, S. M. Ansell, I. N. Micallef, P. B. Johnston, T. G. Call and L. F. Porrata (2006). "Absolute lymphocyte count predicts overall survival in follicular lymphomas." Br J Haematol **134**(6): 596-601.

Singh, S. K., C. Hawkins, I. D. Clarke, J. A. Squire, J. Bayani, T. Hide, R. M. Henkelman, M. D. Cusimano and P. B. Dirks (2004). "Identification of human brain tumour initiating cells." Nature **432**(7015): 396-401.

Skibola, C. F., P. M. Bracci, E. Halperin, L. Conde, D. W. Craig, L. Agana, K. Iyadurai, N. Becker, A. Brooks-Wilson, J. D. Curry, J. J. Spinelli, E. A. Holly, J. Riby, L. Zhang, A. Nieters, M. T. Smith and K. M. Brown (2009). "Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma." Nat Genet **41**(8): 873-5.

Skibola, C. F., P. M. Bracci, E. Halperin, A. Nieters, A. Hubbard, R. A. Paynter, D. R. Skibola, L. Agana, N. Becker, P. Tressler, M. S. Forrest, S. Sankararaman, L. Conde, E. A. Holly and M. T. Smith (2008a). "Polymorphisms in the estrogen receptor 1 and vitamin C and matrix metalloproteinase gene families are associated with susceptibility to lymphoma." PLoS One **3**(7): e2816.

Skibola, C. F., P. M. Bracci, A. Nieters, A. Brooks-Wilson, S. de Sanjose, A. M. Hughes, J. R. Cerhan, D. R. Skibola, M. Purdue, E. Kane, Q. Lan, L. Foretova, M. Schenk, J. J. Spinelli, S. L. Slager, A. J. De Roos, M. T. Smith, E. Roman, W. Cozen, P. Boffetta, A. Kricker, T. Zheng, T. Lightfoot, P. Cocco, Y. Benavente, Y. Zhang, P. Hartge, M. S. Linet, N. Becker, P. Brennan, L. Zhang, B. Armstrong, A. Smith, R. Shiao, A. J. Novak, M. Maynadie, S. J. Chanock, A. Staines, T. R. Holford, E. A. Holly, N. Rothman and S. S. Wang (2010). "Tumor necrosis factor (TNF) and lymphotoxin-alpha (LTA) polymorphisms and risk of non-Hodgkin lymphoma in the InterLymph Consortium." Am J Epidemiol **171**(3): 267-76.

Skibola, C. F., P. M. Bracci, R. A. Paynter, M. S. Forrest, L. Agana, T. Woodage, K. Guegler, M. T. Smith and E. A. Holly (2005a). "Polymorphisms and haplotypes in the cytochrome P450 17A1, prolactin, and catechol-O-methyltransferase genes and non-Hodgkin lymphoma risk." Cancer Epidemiol Biomarkers Prev **14**(10): 2391-401.

Skibola, C. F., J. D. Curry and A. Nieters (2007). "Genetic susceptibility to lymphoma." Haematologica **92**(7): 960-9.

Skibola, C. F., A. Nieters, P. M. Bracci, J. D. Curry, L. Agana, D. R. Skibola, A. Hubbard, N. Becker, M. T. Smith and E. A. Holly (2008b). "A functional TNFRSF5 gene variant is associated with risk of lymphoma." Blood **111**(8): 4348-54.

Skibola, D. R., M. T. Smith, P. M. Bracci, A. E. Hubbard, L. Agana, S. Chi and E. A. Holly (2005b). "Polymorphisms in ghrelin and neuropeptide Y genes are associated with non-Hodgkin lymphoma." Cancer Epidemiol Biomarkers Prev **14**(5): 1251-6.

Sneeringer, C. J., M. P. Scott, K. W. Kuntz, S. K. Knutson, R. M. Pollock, V. M. Richon and R. A. Copeland (2010). "Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas." Proc Natl Acad Sci U S A **107**(49): 20980-5.

Solal-Celigny, P., P. Roy, P. Colombat, J. White, J. O. Armitage, R. Arranz-Saez, W. Y. Au, M. Bellei, P. Brice, D. Caballero, B. Coiffier, E. Conde-Garcia, C. Doyen, M. Federico, R. I. Fisher, J. F. Garcia-Conde, C. Guglielmi, A. Hagenbeek, C. Haioun, M. LeBlanc, A. T. Lister, A. Lopez-Guillermo, P. McLaughlin, N. Milpied, P. Morel, N. Mounier, S. J. Proctor, A. Rohatiner, P. Smith, P. Soubeyran, H. Tilly, U. Vitolo, P. L. Zinzani, E. Zucca and E. Montserrat (2004). "Follicular lymphoma international prognostic index." Blood **104**(5): 1258-65.

Sorensen, C. S., R. G. Syljuasen, J. Falck, T. Schroeder, L. Ronnstrand, K. K. Khanna, B. B. Zhou, J. Bartek and J. Lukas (2003). "Chk1 regulates the S phase checkpoint by coupling the physiological turnover and ionizing radiation-induced accelerated proteolysis of Cdc25A." Cancer Cell **3**(3): 247-58.

Soubeyran, P., H. Eghbali, M. Trojani, F. Bonichon, P. Richaud and B. Hoerni (1996). "Is there any place for a wait-and-see policy in stage I0 follicular lymphoma? A study of 43 consecutive patients in a single center." Ann Oncol **7**(7): 713-8.

Soussi, T., K. Dehouche and C. Beroud (2000). "p53 website and analysis of p53 gene mutations in human cancer: forging a link between epidemiology and carcinogenesis." Hum Mutat **15**(1): 105-13.

Spink, C. F., L. J. Keen, F. K. Mensah, G. R. Law, J. L. Bidwell and G. J. Morgan (2006). "Association between non-Hodgkin lymphoma and haplotypes in the TNF region." Br J Haematol **133**(3): 293-300.

Stadler, Z. K., P. Thom, M. E. Robson, J. N. Weitzel, N. D. Kauff, K. E. Hurley, V. Devlin, B. Gold, R. J. Klein and K. Offit (2010). "Genome-wide association studies of cancer." J Clin Oncol **28**(27): 4255-67.

Stander, M., A. Peraud, B. Leroch and F. W. Kreth (2004). "Prognostic impact of TP53 mutation status for adult patients with supratentorial World Health Organization Grade II astrocytoma or oligoastrocytoma: a long-term analysis." Cancer **101**(5): 1028-35.

Steward, W. P., D. Crowther, L. J. McWilliam, J. M. Jones, D. P. Deakin, I. D. Todd, G. Blackledge, J. Wagstaff, J. H. Scarffe and M. Harris (1988). "Maintenance chlorambucil after CVP in the management of advanced stage, low-grade histologic type non-Hodgkin's lymphoma. A randomized prospective study with an assessment of prognostic factors." Cancer **61**(3): 441-7.

Stirewalt, D. L., K. J. Kopecky, S. Meshinchi, F. R. Appelbaum, M. L. Slovak, C. L. Willman and J. P. Radich (2001). "FLT3, RAS, and TP53 mutations in elderly patients with acute myeloid leukemia." Blood **97**(11): 3589-95.

Stranger, B. E., M. S. Forrest, M. Dunning, C. E. Ingle, C. Beazley, N. Thorne, R. Redon, C. P. Bird, A. de Grassi, C. Lee, C. Tyler-Smith, N. Carter, S. W. Scherer, S. Tavare, P. Deloukas, M. E. Hurles and E. T. Dermitzakis (2007a). "Relative impact of nucleotide and copy number variation on gene expression phenotypes." Science **315**(5813): 848-53.

Stranger, B. E., A. C. Nica, M. S. Forrest, A. Dimas, C. P. Bird, C. Beazley, C. E. Ingle, M. Dunning, P. Flicek, D. Koller, S. Montgomery, S. Tavare, P.

Deloukas and E. T. Dermitzakis (2007b). "Population genomics of human gene expression." Nat Genet **39**(10): 1217-24.

Sturm, R. A., D. L. Duffy, Z. Z. Zhao, F. P. Leite, M. S. Stark, N. K. Hayward, N. G. Martin and G. W. Montgomery (2008). "A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color." Am J Hum Genet **82**(2): 424-31.

Summers, K. E., L. K. Goff, A. G. Wilson, R. K. Gupta, T. A. Lister and J. Fitzgibbon (2001). "Frequency of the Bcl-2/IgH rearrangement in normal individuals: implications for the monitoring of disease in patients with follicular lymphoma." J Clin Oncol **19**(2): 420-4.

Swenson, W. T., J. E. Wooldridge, C. F. Lynch, V. L. Forman-Hoffman, E. Chrischilles and B. K. Link (2005). "Improved survival of follicular lymphoma patients in the United States." J Clin Oncol **23**(22): 5019-26.

Swerdlow, S. H. (2004). "Pediatric follicular lymphomas, marginal zone lymphomas, and marginal zone hyperplasia." Am J Clin Pathol **122 Suppl**: S98-109.

Swerdlow, S. H., Campo E., Harris N. L., Jaffe E. S., Pileri S. A., Stein H., Thiele J., Vardiman J. W., Eds. (2008). WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. IARC Press, Lyon.

Symmers, D. (1927). "Follicular lymphadenopathy with splenomegaly: a newly recognized disease of the lymphatic system." Archives of Pathology **3**: 816-820.

Szczepanski, T., A. W. Langerak, J. J. van Dongen and J. H. van Krieken (1998). "Lymphoma with multi-gene rearrangement on the level of immunoglobulin heavy chain, light chain, and T-cell receptor beta chain." Am J Hematol **59**(1): 99-100.

Tabe, Y., D. Sebasigari, L. Jin, M. Rudelius, T. Davies-Hill, K. Miyake, T. Miida, S. Pittaluga and M. Raffeld (2009). "MDM2 antagonist nutlin-3 displays antiproliferative and proapoptotic activity in mantle cell lymphoma." Clin Cancer Res **15**(3): 933-42.

Tanaka, H., H. Arakawa, T. Yamaguchi, K. Shiraishi, S. Fukuda, K. Matsui, Y. Takei and Y. Nakamura (2000). "A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage." Nature **404**(6773): 42-9.

Tao, W. and A. J. Levine (1999). "Nucleocytoplasmic shuttling of oncoprotein Hdm2 is required for Hdm2-mediated degradation of p53." Proc Natl Acad Sci U S A **96**(6): 3077-80.

Terzian, T., Y. Wang, C. S. Van Pelt, N. F. Box, E. L. Travis and G. Lozano (2007). "Haploinsufficiency of Mdm2 and Mdm4 in tumorigenesis and development." Mol Cell Biol **27**(15): 5479-85.

Thomas, M., A. Kalita, S. Labrecque, D. Pim, L. Banks and G. Matlashewski (1999). "Two polymorphic variants of wild-type p53 differ biochemically and biologically." Mol Cell Biol **19**(2): 1092-100.

Thomson, K. J., E. C. Morris, A. Bloor, G. Cook, D. Milligan, A. Parker, F. Clark, L. Yung, D. C. Linch, R. Chakraverty, K. S. Peggs and S. Mackinnon (2009). "Favorable long-term survival after reduced-intensity allogeneic transplantation for multiple-relapse aggressive non-Hodgkin's lymphoma." J Clin Oncol **27**(3): 426-32.

Thomson, K. J., E. C. Morris, D. Milligan, A. N. Parker, A. E. Hunter, G. Cook, A. J. Bloor, F. Clark, M. Kazmi, D. C. Linch, R. Chakraverty, K. S. Peggs and S. Mackinnon (2010). "T-cell-depleted reduced-intensity transplantation

followed by donor leukocyte infusions to promote graft-versus-lymphoma activity results in excellent long-term survival in patients with multiply relapsed follicular lymphoma." J Clin Oncol **28**(23): 3695-700.

Tilly, H., A. Rossi, A. Stamatoullas, B. Lenormand, C. Bigorgne, A. Kunlin, M. Monconduit and C. Bastard (1994). "Prognostic value of chromosomal abnormalities in follicular lymphoma." Blood **84**(4): 1043-9.

Tomblyn, M. R., M. Ewell, C. Bredeson, B. S. Kahl, S. A. Goodman, M. M. Horowitz, J. M. Vose, R. S. Negrin and G. G. Laport (2010). "Autologous versus Reduced-Intensity Allogeneic Hematopoietic Cell Transplantation for Patients with Chemosensitive Follicular Non-Hodgkin Lymphoma beyond First Complete Response or First Partial Response." Biol Blood Marrow Transplant **10.1016/j.bbmt.2010.11.004**.

Tonegawa, S. (1983). "Somatic generation of antibody diversity." Nature **302**(5909): 575-81.

Tough, I. M., W. M. Court Brown, A. G. Baikie, K. E. Buckton, D. G. Harnden, P. A. Jacobs, M. J. King and B. J. Mc (1961). "Cytogenetic studies in chronic myeloid leukaemia and acute leukaemia associated with monogolism." Lancet **1**(7174): 411-7.

Triebel, F., F. Faure, M. Graziani, S. Jitsukawa, M. P. Lefranc and T. Hercend (1988). "A unique V-J-C-rearranged gene encodes a gamma protein expressed on the majority of CD3+ T cell receptor-alpha/beta- circulating lymphocytes." J Exp Med **167**(2): 694-9.

Tsujimoto, Y., J. Cossman, E. Jaffe and C. M. Croce (1985a). "Involvement of the bcl-2 gene in human follicular lymphoma." Science **228**(4706): 1440-3.

Tsujimoto, Y. and C. M. Croce (1986). "Analysis of the structure, transcripts, and protein products of bcl-2, the gene involved in human follicular lymphoma." Proc Natl Acad Sci U S A **83**(14): 5214-8.

Tsujimoto, Y., L. R. Finger, J. Yunis, P. C. Nowell and C. M. Croce (1984). "Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation." Science **226**(4678): 1097-9.

Tsujimoto, Y., J. Gorham, J. Cossman, E. Jaffe and C. M. Croce (1985b). "The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining." Science **229**(4720): 1390-3.

van Besien, K. (2009). "Allogeneic stem cell transplantation in follicular lymphoma: recent progress and controversy." Hematology Am Soc Hematol Educ Program: 610-8.

van de Schans, S. A., E. W. Steyerberg, M. R. Nijziel, G. J. Creemers, M. L. Janssen-Heijnen and D. J. van Spronsen (2009). "Validation, revision and extension of the Follicular Lymphoma International Prognostic Index (FLIPI) in a population-based setting." Ann Oncol **20**(10): 1697-702.

van Dongen, J. J., A. W. Langerak, M. Bruggemann, P. A. Evans, M. Hummel, F. L. Lavender, E. Delabesse, F. Davi, E. Schuuring, R. Garcia-Sanz, J. H. van Krieken, J. Droese, D. Gonzalez, C. Bastard, H. E. White, M. Spaargaren, M. Gonzalez, A. Parreira, J. L. Smith, G. J. Morgan, M. Kneba and E. A. Macintyre (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): 2257-317.

van Oers, M. H., R. Klasa, R. E. Marcus, M. Wolf, E. Kimby, R. D. Gascoyne, A. Jack, M. Van't Veer, A. Vranovsky, H. Holte, M. van Glabbeke, I.

Teodorovic, C. Rozewicz and A. Hagenbeek (2006). "Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial." Blood **108**(10): 3295-301.

van Oers, M. H., M. Van Glabbeke, L. Giurgea, R. Klasa, R. E. Marcus, M. Wolf, E. Kimby, M. van t Veer, A. Vranovsky, H. Holte and A. Hagenbeek (2010). "Rituximab maintenance treatment of relapsed/resistant follicular non-Hodgkin's lymphoma: long-term outcome of the EORTC 20981 phase III randomized intergroup study." J Clin Oncol **28**(17): 2853-8.

van Zelm, M. C., M. van der Burg and J. J. van Dongen (2007). "Homeostatic and maturation-associated proliferation in the peripheral B-cell compartment." Cell Cycle **6**(23): 2890-5.

Vassilev, L. T., B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu (2004). "In vivo activation of the p53 pathway by small-molecule antagonists of MDM2." Science **303**(5659): 844-8.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S.

Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. Zhu (2001). "The sequence of the human genome." Science **291**(5507): 1304-51.

Viardot, A., P. Moller, J. Hogel, K. Werner, G. Mechtersheimer, A. D. Ho, G. Ott, T. F. Barth, R. Siebert, S. Gesk, B. Schlegelberger, H. Dohner and M. Bentz (2002). "Clinicopathologic correlations of genomic gains and losses in follicular lymphoma." J Clin Oncol **20**(23): 4523-30.

Villeneuve, S., L. Orsi, A. Monnereau, C. Berthou, P. Fenaux, G. Marit, P. Soubeyran, F. Huguet, N. Milpied, M. Leporrier, D. Hemon, X. Troussard and J. Clavel (2009). "Increased frequency of hematopoietic malignancies in relatives of patients with lymphoid neoplasms: a French case-control study." Int J Cancer **124**(5): 1188-95.

Villuendas, R., M. Sanchez-Beato, J. C. Martinez, A. I. Saez, B. Martinez-Delgado, J. F. Garcia, M. S. Mateo, L. Sanchez-Verde, J. Benitez, P. Martinez and M. A. Piris (1998). "Loss of p16/INK4A protein expression in non-Hodgkin's lymphomas is a frequent finding associated with tumor progression." Am J Pathol **153**(3): 887-97.

Visvader, J. E. (2011). "Cells of origin in cancer." Nature **469**(7330): 314-22.

Vogelstein, B. and K. W. Kinzler (2004). "Cancer genes and the pathways they control." Nat Med **10**(8): 789-99.

Vogelstein, B., D. Lane and A. J. Levine (2000). "Surfing the p53 network." Nature **408**(6810): 307-10.

Von Fliedner, V. E., H. Merica, M. Jeannet, C. Barras, A. Feldges, P. Imbach and M. Wyss (1983). "Evidence for HLA-linked susceptibility factors in childhood leukemia." Hum Immunol **8**(3): 183-93.

Vousden, K. H. (2000). "p53: death star." Cell **103**(5): 691-4.

Wang, Q., W. He, C. Lu, Z. Wang, J. Wang, K. E. Giercksky, J. M. Nesland and Z. Suo (2009). "Oct3/4 and Sox2 are significantly associated with an unfavorable clinical outcome in human esophageal squamous cell carcinoma." Anticancer Res **29**(4): 1233-41.

Wang, S. S., A. M. Abdou, L. M. Morton, R. Thomas, J. R. Cerhan, X. Gao, W. Cozen, N. Rothman, S. Davis, R. K. Severson, L. Bernstein, P. Hartge and M. Carrington (2010). "Human leukocyte antigen class I and II alleles in non-Hodgkin lymphoma etiology." Blood **115**(23): 4820-3.

Wang, S. S., J. R. Cerhan, P. Hartge, S. Davis, W. Cozen, R. K. Severson, N. Chatterjee, M. Yeager, S. J. Chanock and N. Rothman (2006a). "Common genetic variants in proinflammatory and other immunoregulatory genes and risk for non-Hodgkin lymphoma." Cancer Res **66**(19): 9771-80.

Wang, S. S., S. Davis, J. R. Cerhan, P. Hartge, R. K. Severson, W. Cozen, Q. Lan, R. Welch, S. J. Chanock and N. Rothman (2006b). "Polymorphisms in oxidative stress genes and risk for non-Hodgkin lymphoma." Carcinogenesis **27**(9): 1828-34.

Wang, Y. C., H. S. Lee, S. K. Chen, Y. Y. Chang and C. Y. Chen (1999). "Prognostic significance of p53 codon 72 polymorphism in lung carcinomas." Eur J Cancer **35**(2): 226-30.

Warzocha, K., P. Ribeiro, J. Bienvenu, P. Roy, C. Charlot, D. Rigal, B. Coiffier and G. Salles (1998). "Genetic polymorphisms in the tumor necrosis factor locus influence non-Hodgkin's lymphoma outcome." Blood **91**(10): 3574-81.

Weiss, L. M., R. A. Warnke, J. Sklar and M. L. Cleary (1987). "Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas." N Engl J Med **317**(19): 1185-9.

Weng, W. K. and R. Levy (2003). "Two immunoglobulin G fragment C receptor polymorphisms independently predict response to rituximab in patients with follicular lymphoma." J Clin Oncol **21**(21): 3940-7.

Wilkening, S., J. L. Bermejo and K. Hemminki (2007). "MDM2 SNP309 and cancer risk: a combined analysis." Carcinogenesis **28**(11): 2262-7.

Willemze, R., E. S. Jaffe, G. Burg, L. Cerroni, E. Berti, S. H. Swerdlow, E. Ralfkiaer, S. Chimenti, J. L. Diaz-Perez, L. M. Duncan, F. Grange, N. L. Harris, W. Kempf, H. Kerl, M. Kurrer, R. Knobler, N. Pimpinelli, C. Sander, M. Santucci, W. Sterry, M. H. Vermeer, J. Wechsler, S. Whittaker and C. J. Meijer (2005). "WHO-EORTC classification for cutaneous lymphomas." Blood **105**(10): 3768-85.

Willett, E. V., C. F. Skibola, P. Adamson, D. R. Skibola, G. J. Morgan, M. T. Smith and E. Roman (2005). "Non-Hodgkin's lymphoma, obesity and energy homeostasis polymorphisms." Br J Cancer **93**(7): 811-6.

Williamson, D., J. Selfe, T. Gordon, Y. J. Lu, K. Pritchard-Jones, K. Murai, P. Jones, P. Workman and J. Shipley (2007). "Role for amplification and expression of glypican-5 in rhabdomyosarcoma." Cancer Res **67**(1): 57-65.

Willis, T. G. and M. J. Dyer (2000). "The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies." Blood **96**(3): 808-22.

Wrench, D., P. Leighton, C. F. Skibola, L. Conde, J. B. Cazier, J. Matthews, S. Iqbal, E. Carlotti, C. Bodor, S. Montoto, M. Calaminici, J. G. Gribben, T. A. Lister and J. Fitzgibbon (2011). "SNP rs6457327 in the HLA region on chromosome 6p is predictive of the transformation of follicular lymphoma." Blood **117**(11): 3147-50.

Wrench, D., R. Waters, E. Carlotti, S. Iqbal, J. Matthews, M. Calaminici, J. Gribben, T. A. Lister and J. Fitzgibbon (2009). "Clinical relevance of MDM2 SNP 309 and TP53 Arg72Pro in follicular lymphoma." Haematologica **94**(1): 148-50.

Wu, X., J. H. Bayle, D. Olson and A. J. Levine (1993). "The p53-mdm-2 autoregulatory feedback loop." Genes Dev **7**(7A): 1126-32.

Yano, T., E. S. Jaffe, D. L. Longo and M. Raffeld (1992). "MYC rearrangements in histologically progressed follicular lymphomas." Blood **80**(3): 758-67.

Yap, D. B., J. Chu, T. Berg, M. Schapira, S. W. Cheng, A. Moradian, R. D. Morin, A. J. Mungall, B. Meissner, M. Boyle, V. E. Marquez, M. A. Marra, R. D. Gascoyne, R. K. Humphries, C. H. Arrowsmith, G. B. Morin and S. A. Aparicio (2011). "Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation." Blood **117**(8): 2451-9.

Young, K. H., D. D. Weisenburger, B. J. Dave, L. Smith, W. Sanger, J. Iqbal, E. Campo, J. Delabie, R. D. Gascoyne, G. Ott, L. Rimsza, H. K. Muller-Hermelink, E. S. Jaffe, A. Rosenwald, L. M. Staudt, W. C. Chan and T. C. Greiner (2007). "Mutations in the DNA-binding codons of TP53, which are

associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B-cell lymphoma." Blood **110**(13): 4396-405.

Young, R. C., D. L. Longo, E. Glatstein, D. C. Ihde, E. S. Jaffe and V. T. DeVita, Jr. (1988). "The treatment of indolent lymphomas: watchful waiting v aggressive combined modality treatment." Semin Hematol **25**(2 Suppl 2): 11-6.

Yu, W., J. Inoue, I. Imoto, Y. Matsuo, A. Karpas and J. Inazawa (2003). "GPC5 is a possible target for the 13q31-q32 amplification detected in lymphoma cell lines." J Hum Genet **48**(6): 331-5.

Yun, X., Y. Wu, L. Yao, H. Zong, Y. Hong, J. Jiang, J. Yang, Z. Zhang and J. Gu (2007). "CDK11(p58) protein kinase activity is associated with Bcl-2 down-regulation in pro-apoptosis pathway." Mol Cell Biochem **304**(1-2): 213-8.

Yunis, J. J., G. Frizzera, M. M. Oken, J. McKenna, A. Theologides and M. Arnesen (1987). "Multiple recurrent genomic defects in follicular lymphoma. A possible model for cancer." N Engl J Med **316**(2): 79-84.

Zacchi, P., M. Gostissa, T. Uchida, C. Salvagno, F. Avolio, S. Volinia, Z. Ronai, G. Blandino, C. Schneider and G. Del Sal (2002). "The prolyl isomerase Pin1 reveals a mechanism to control p53 functions after genotoxic insults." Nature **419**(6909): 853-7.

Zauberman, A., D. Flusberg, Y. Haupt, Y. Barak and M. Oren (1995). "A functional p53-responsive intronic promoter is contained within the human mdm2 gene." Nucleic Acids Res **23**(14): 2584-92.

Zelenetz, A. D., T. T. Chen and R. Levy (1991). "Histologic transformation of follicular lymphoma to diffuse lymphoma represents tumor progression by a single malignant B cell." J Exp Med **173**(1): 197-207.

Zelenetz, A. D., T. T. Chen and R. Levy (1992). "Clonal expansion in follicular lymphoma occurs subsequent to antigenic selection." J Exp Med **176**(4): 1137-48.

Zenz, T., S. Habe, A. Benner, D. Kienle, H. Dohner and S. Stilgenbauer (2008). "The MDM2 -309 T/G promoter single nucleotide polymorphism does not alter disease characteristics in chronic lymphocytic leukemia." Haematologica.

Zhang, X., S. Karnan, H. Tagawa, R. Suzuki, S. Tsuzuki, Y. Hosokawa, Y. Morishima, S. Nakamura and M. Seto (2004). "Comparison of genetic aberrations in CD10+ diffused large B-cell lymphoma and follicular lymphoma by comparative genomic hybridization and tissue-fluorescence in situ hybridization." Cancer Sci **95**(10): 809-14.

Zhang, X., X. Miao, Y. Guo, W. Tan, Y. Zhou, T. Sun, Y. Wang and D. Lin (2006). "Genetic polymorphisms in cell cycle regulatory genes MDM2 and TP53 are associated with susceptibility to lung cancer." Hum Mutat **27**(1): 110-7.

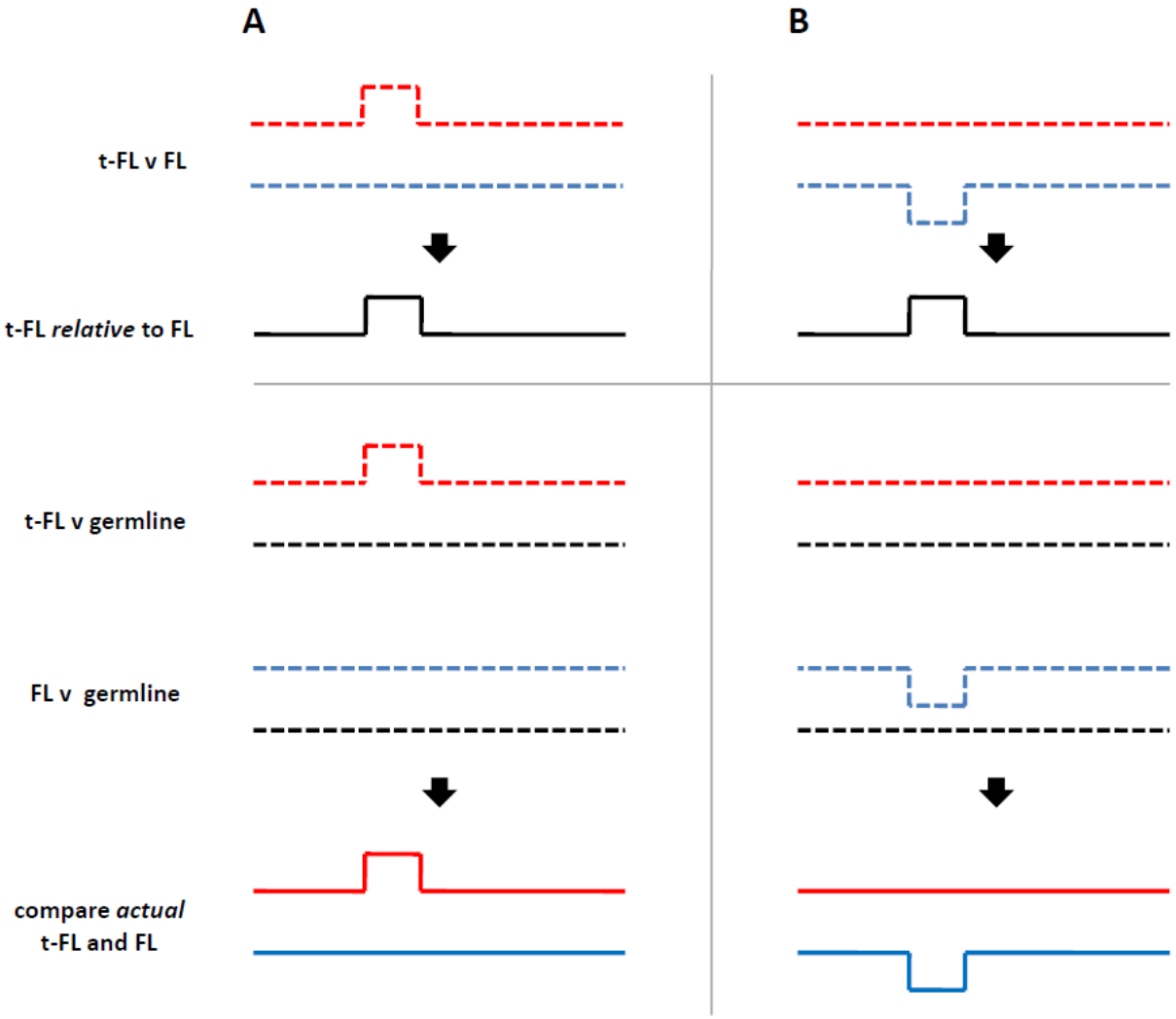
Zhao, X., R. E. Ayer, S. L. Davis, S. J. Ames, B. Florence, C. Torchinsky, J. S. Liou, L. Shen and R. A. Spanjaard (2005). "Apoptosis factor EI24/PIG8 is a novel endoplasmic reticulum-localized Bcl-2-binding protein which is associated with suppression of breast cancer invasiveness." Cancer Res **65**(6): 2125-9.

Zhu, D., R. E. Hawkins, T. J. Hamblin and F. K. Stevenson (1994). "Clonal history of a human follicular lymphoma as revealed in the immunoglobulin variable region genes." Br J Haematol **86**(3): 505-12.

Zhu, D., H. McCarthy, C. H. Ottensmeier, P. Johnson, T. J. Hamblin and F. K. Stevenson (2002). "Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma." Blood **99**(7): 2562-8.

Appendices

Appendix 1



Appendix 1 **Figure 1 Determining copy number changes in FL / t-FL**

Shown are schematics of copy number (CN) plots for FL and t-FL. Dashed lines indicate signal intensities over a section of one chromosome with red indicating t-FL, blue FL and black germline. Solid lines represent the CN plots produced by comparison of sample of interest (top dashed line) to reference sample (bottom dashed line) with black indicating t-FL against FL, red t-FL against germline and blue FL against germline. The necessity for generating CN plots for FL and t-FL against germline is indicated as for a gain in t-FL against normal FL (A, top) and normal t-FL against a loss in FL (B, top) the CN plots produced are the same. However, by instead generating CN plots for the same tumour samples against germline and then comparing the two CN plots the gain in t-FL (A, bottom) can be distinguished from the loss in FL (B, bottom).

Appendix 2

Appendix 2 Table 1 Oligonucleotide primers and probes

	Locus[#]		Oligonucleotide[§]	5' to 3' sequence
Chapter 3				
	MDM2 SNP 309		F	CGGGAGTTCAGGGTAAAGGT
			R	GCGCAGCGTTCACACTAG
			Allele probe 1	VIC-CTCCCGCGCCGAAG
			Allele probe 2	FAM-TCCCGCGCCGCAG
	TP53 Arg72Pro		F	GCGCAGCGTTCACACTAG
			R	CCAGATGAAGCTCCCAG
			Allele probe 1	VIC-TCCCCGCGTGGCC
			Allele probe 2	FAM-CTCCCCCCGTGGCC
	SNP rs6457327		F	TGAAGGCTGGTGACAGATGT
			R	TCTGTCCCTCCTTTCTCTGA
	SNP rs10484561		F	GCCATTTGAGTTCTTCTGTGA
			R	AGAAGCCAGGTGAGAACTTG
	C6orf15	exon 1 + 2 (proximal)	F	GAATGAGGGTGGAATGTGGA
			R	CACTGGAGAGGTAAGAGAGT
		exon 2 (distal)	F	ATCCTTGGCAGATGATGGCT
			R	TCAGCTTATGCCCAGACTCA
	C6orf15 cDNA	exon 1*	F	GGCCGCGTGGCAGGGAG
		exon 2*	R	AGTGGAGGAAGGTTGTCCGAGCTG
		splice site	F	GTCTTCATCTCCCAGGCC
		exon 2	R	CCAGTGGAGGAAGGTTGTC
		exon 2	F	ATCCTTGGCAGATGATGGCT
		exon 2	R	CCAGATTCCCTCAGGGTGT
Chapter 4				
	CREBBP	exon 2	F	TGGCAGTTGGAGAGCTGTA
			R	CTGGCTGCCTGTTTAGGCA
		exon 14	F	ACTCAGCCATCAACTCCTGT
			R	TGCTGGAGACGAGGTCTCA
		exon 31	F	TACGACCTCTGCATCAACTG
			R	AGTGCTTGGCGTGGTAGCA
	EZH2	exon	F	GTGTGCCCAATTACTGCCTT
			R	AGCATGCAAATCCACAAAC
	SSU72	exon 2	F	GTGTCTTAGTGTGACTGTGAT
			R	ACCAGGAATAGAAACCAAGCT
	GNB1	exon 3	F	TTTACTTTTCTCACCTGCTTCA
			R	AACAAGTCATCCTGTAAACCAT

Appendix 2 Table 1 (continued)

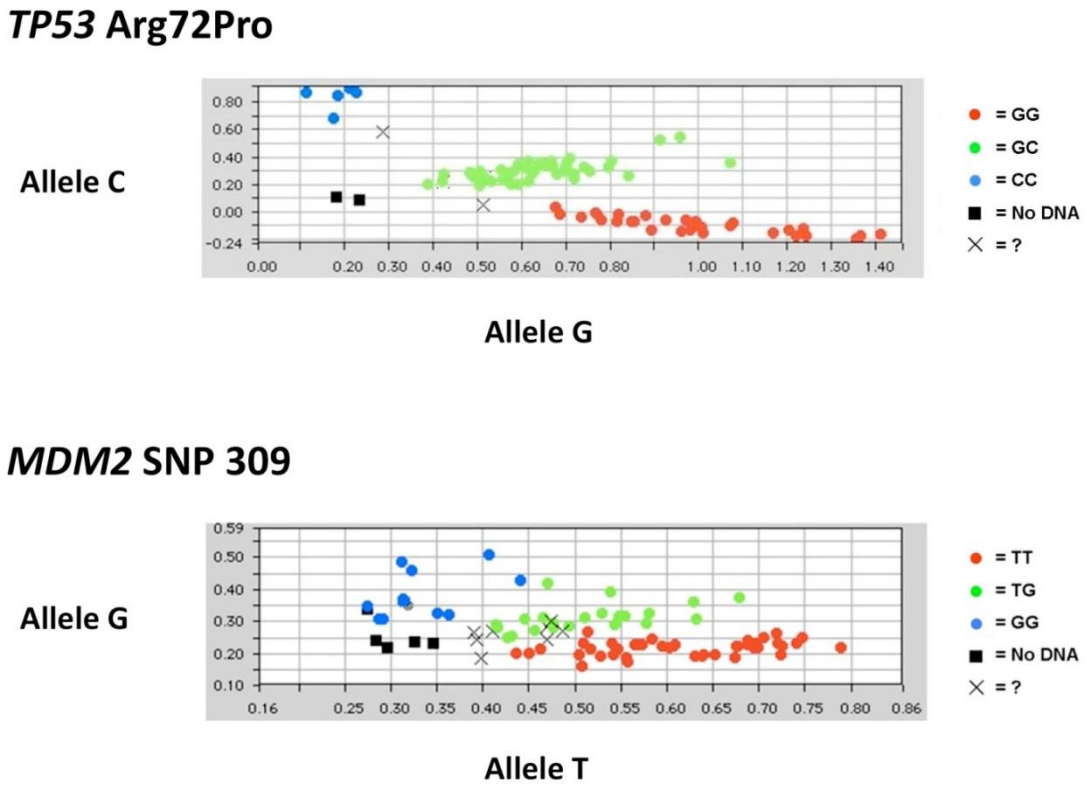
	Locus		Oligonucleotide[§]	5' to 3' sequence
Chapter 4				
	GNB1	exon 11	F	GGCACACAGGTGAGATCGA
			R	CTAGAAACCGTTAATGACAACT
	CDC2L2	exon 1	F	TTTGGAGTCCTGGACCTGA
			R	TTCCGTGCGGGATGAGACT
		exon 2	F	TGTGTCCTGATGTAGGCTGA
			R	CTTCACCGAAGAAGCGTTGT
		exon 3	F	TTTGGGATGTTGCCCAGGTT
			R	GCACACAGTTGTCACAGTGA
		exon 4	F	TTCTCAAGCATGACGCAGT
			R	CACACTTGAACATGATGTCA
		exon 5	F	TGAAGTGAGGTTTCTTTGCA
			R	TCTTCTTCATTGCTGTGACA
		exon 6	F	GAATTACGTAGAAAGCCAGCT
			R	AAAGCAAGCACCGGAACCAA
		exon 7	F	AGGAAGGCAGTGCTGCTAGT
			R	TTCAGCCAGCACATCTCTCA
		exons 8 + 9	F	CGTCTGTGGATGAGCTGGTT
			R	TAGTGAAGTCACAACACCTCA
		exon 10	F	TTCTTGGGAACACTTGTTGCT
			R	GACCGGCTGTCTTAGGAGA
		exon 11	F	AGGAGGAGGTTCATGATCCT
			R	TGGTAACTCCGACTGCCAAT
		exon 12	F	AAAGAACTGAGCTGCTCCCA
			R	CCAGTAGCTGCTCAGGTGAA
		exon 13 + 14	F	CCGTGACTTTGAGAATGCCT
			R	CACTTGTACGCAGACAGGA
		exons 15 + 16	F	GTCCTGTCTGCGTACAAGT
			R	GGAGTATTCCTAAGACGCCA
		exon 17 + 18	F	GCACTGGCCTTCCTGAACA
			R	GACTGGGAAGTCACCGCTA
		exons 19 + 20	F	TGGAGCACACGAGCACCT
			R	GAGGAGTTCCGAGTCTCAT
	TNFRSF14	exon 1	F	AGTGGACTGGAATGGTGCA
			R	TGGACGGAGTGGTGAGTG
		exon 2	F	CTGTCCTCTCTACCAGGCA
			R	TTCATCACACCCATGGCTGT

Appendix 2 Table 1 (continued)

Locus		Oligonucleotide[§]	5' to 3' sequence
Chapter 4			
	exon 2	F	CTGTCCTCTCTACCAGGCA
		R	TTCATCACACCCATGGCTGT
	exon 3	F	TAGCTGGTGTCTCCCTGCT
		R	AAGTGAGAGGTGAGACAGAG
	exon 4	F	TCCATGCTGGGTACCTCTG
		R	CCAAGGTTCTGTCTCACCA
	exon 5	F	GAAGCTCACAGACAAGCAGT
		R	CTTCAAGCCTTTCTGCTGGA
	exon 6	F	TCCCTGGGACCTGTCTTCA
		R	GGTGACAGAGCTCCAAGAG
	exon 7	F	GGGAGAAGCAGGAGTTGTG
		R	GACTCAGGACCCTCAGAGA
	exon 8	F	TAGCTCAGGAAAGAACCCAC
		R	AGGTGGACAGCCTCTTTC

[#]Primers for *SSU72* and *GNB1* were used as positive controls to confirm DNA amplification either side of the chromosome 1p36 deletion observed in the WSU-NHL cell line.
[§]Forward and reverse oligonucleotide primers are all designed on intronic sequences to give full exon coverage except for those used to sequence SNP *MDM2* SNP 309 (intronic) and *TP53 Arg72Pro* (designed in exon so as near to the SNP as possible); SNP rs6457327, SNP rs10484561 which are not in genes; and for examination of *C6orf15* expression using cDNA. Oligonucleotide probes are indicated for AD-PCR determination of *MDM2* SNP 309 and *TP53 Arg72Pro* genotypes. For rs6457327 and rs10484561 commercial prepared AD-PCR genotyping assays were used as detailed in **3.3.4**.
*These primers were utilised from the report by Skibola and colleagues (Skibola *et al.* 2009).
F indicates forward; and R reverse.

Appendix 3



***Appendix 3* Figure 1 AD-PCR plots for *TP53* Arg72Pro and *MDM2* SNP 309**

Illustrated are AD-PCR result plots with assigned genotypes (GG, GC, CC for *TP53* Arg72Pro; TT, TG, GG for *MDM2* SNP 309) indicated on the right. Those samples marked with a cross required subsequent inspection of the real-time PCR curves produced in the AD-PCR reaction for genotype assignment. Each axis corresponds to the signal intensity of the respective allele's oligonucleotide probe with those incorporating the FAM fluorescent reporter dye on the y axis and those incorporating the VIC fluorescent reporter dye on the x axis. Top, example of plot for one run of *TP53* Arg72Pro AD-PCR; bottom, example of plot for one run of *MDM2* SNP 309 AD-PCR. Units for each axis = normalized RFU (relative fluorescence units).

Appendix 3 Table 1 Cross-tabulation of all *MDM2* / *TP53* SNP genotypes

MDM2 SNP 309 genotype	*TP53* Arg72Pro genotype			*Total*
	GG	GC	CC	
TT	45 (20)	46 (20)	8 (4)	*99*
TG	50 (22)	40 (18)	8 (4)	*98*
GG	10 (4)	19 (8)	0 (0)	*29*
Total	*105*	*105*	*16*	*226*

Appendix 3 Table 2 Cross-tabulation of restricted *MDM2* / *TP53* SNP genotypes

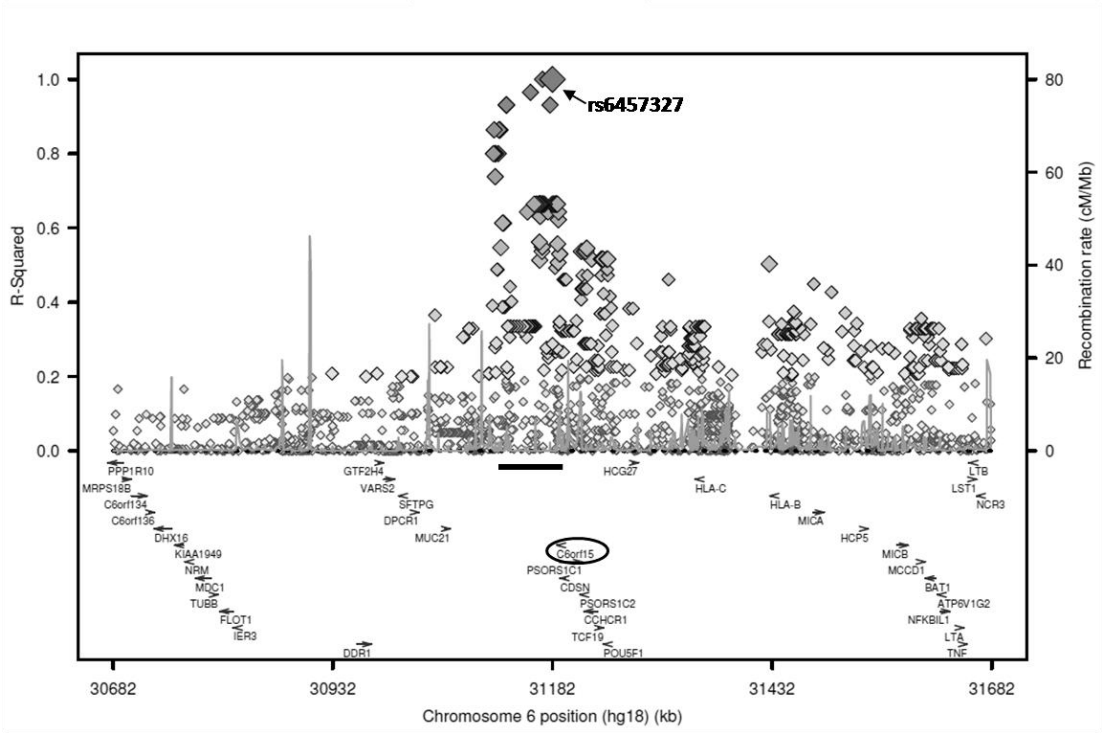
MDM2 SNP 309 genotype*	*TP53* Arg72Pro genotype*		*Total*
	GG	GC / CC	
TT	45 (20)	54 (24)	*99*
TG / GG	60 (26)	67 (30)	*127*
Total	*105*	*121*	*226*

Percentages are included within parentheses.
*Genotypes are restricted to homozygosity for major allele and presence of minor allele.

Appendix 3 Table 3 Median age at diagnosis for *MDM2* SNP 309 and *TP53* Arg72Pro

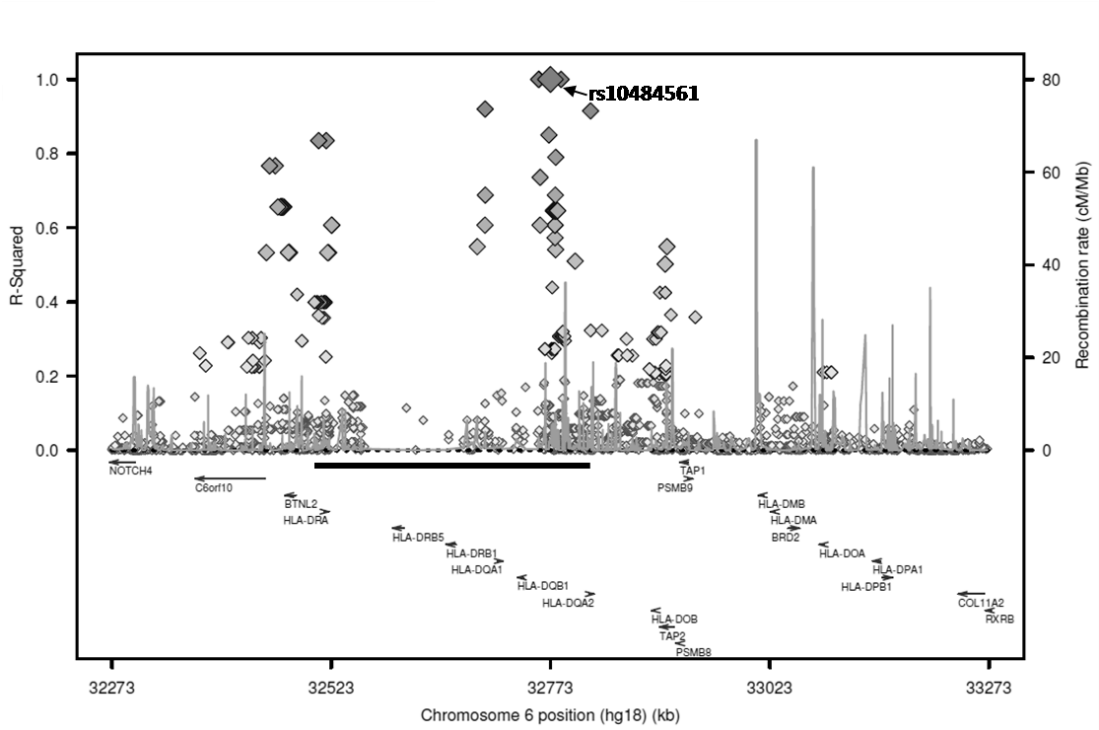
SNP	Genotype	Median age at FL diagnosis (range)	*P*
	TT	46 (21 - 84)	
MDM2 SNP 309	TG	47 (22 - 85)	0.91
	GG	47 (33 - 69)	
	GG	47 (21 - 84)	
TP53 Arg72Pro	GC	46 (30 - 79)	0.60
	CC	46 (30 - 74)	

Appendix 4



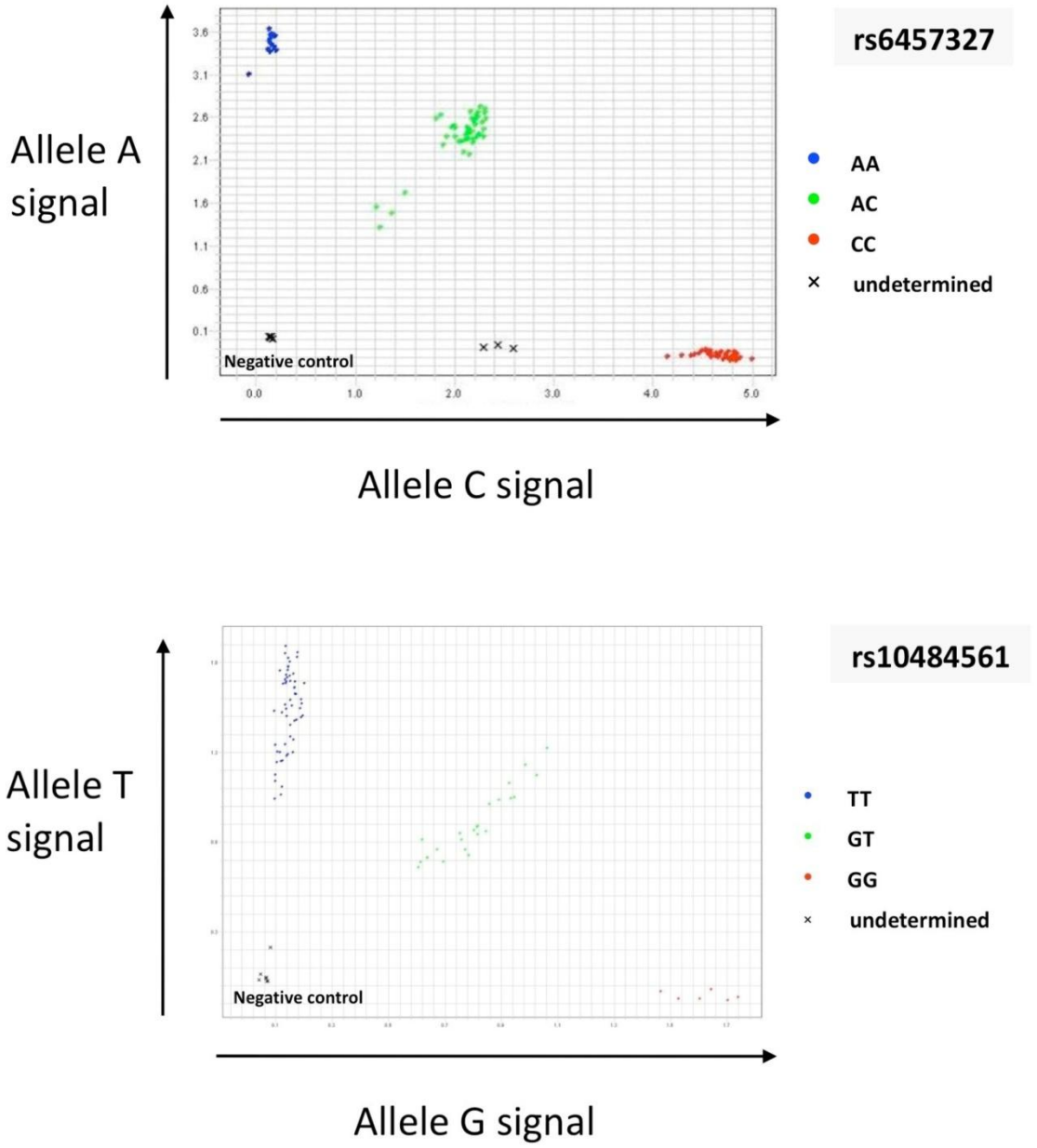
***Appendix 4* Figure 1 Linkage between rs6457327 and nearby SNPs and genes**

The plot was generated using the SNP Annotation and Proxy (SNAP) search (Johnson *et al.* 2008a) version 2.2, available at http://www.broadinstitute.org/mpg/snap/ldplot.php. Diamonds represent individual SNPs from the HapMap ("The International HapMap Project" 2003) and 1000 genome (Durbin *et al.* 2010) projects, with the largest diamond representing the SNP of interest. R-squared represents the degree of pair-wise linkage between SNPs and values > 0.8 represent high LD with limits indicated by the black bar. (Also illustrated is the recombination rate represented by the spikes in the plot.) Genes are indicated at the foot of the plot with *C6orf15* located at the centromeric aspect of the region of high LD indicated by the black oval. The study by Skibola and colleagues (Skibola *et al.* 2009) localized the region of high LD around rs6457327 to a smaller, 26 kb, region that still contained *C6orf15*.



Appendix 4 Figure 2 Linkage between rs10484561 and nearby SNPs

The layout and abbreviations are as indicated in ***Appendix 4* Figure 1**. rs10484561 is located 30kb centromeric to *HLA-DQB1* and 43kb telomeric to *HLA-DQA2*. Conde and colleagues (Conde *et al.* 2010) further localized the high LD surrounding rs10484561 to a 100 kb region.



Appendix 4 **Figure 3 AD-PCR plots of rs6457327 and rs10484561 genotyping results**

Illustrated are the plots obtained from interpretation of the fluorescence signals at the end of AD-PCR reactions for a cohort of samples typed for top, rs6457327 and bottom, rs10484561. For rs6457327 one sample in triplicate is "undetermined" (in addition to the negative control). The position of the three samples suggests CC genotype and this was confirmed by examination of the real-time qPCR curves generated during the AD-PCR reaction. For rs10484561, one sample in triplicate is undetermined with no signal (similar to the negative control). This was due to suboptimal quantity and quality of the dilution made from the stock DNA. When a new DNA dilution was reassessed, genotyping by AD-PCR was successful. For both plots the legend to the right indicates the genotypes of the respective SNP which correspond to the plot points. Units for each axis = normalized RFU (relative fluorescence units).

Appendix 4 Table 1 rs6457327 and rs10484561 genotypes and age at FL onset

SNP genotype	n	Median age in years (range)	Association
rs6457327			
AA+AC	74	51 (26 – 87)	*P* = 0.73
CC	65	52 (31 – 78)	
Total	*130*		
rs10484561			
GG+GT	59	51 (35 – 84)	*P* = 0.59
TT	71	53 (26 – 87)	
Total	*130*		

Appendix 4 Table 2 rs6457327 allele distributions in 130 and 88 case sets

Case set	Number of alleles			*P*
	A	C	*Total*	
"130"	83	177	*260*	0.84
"88"	58	118	*176*	
Total	*141*	*295*	*436*	

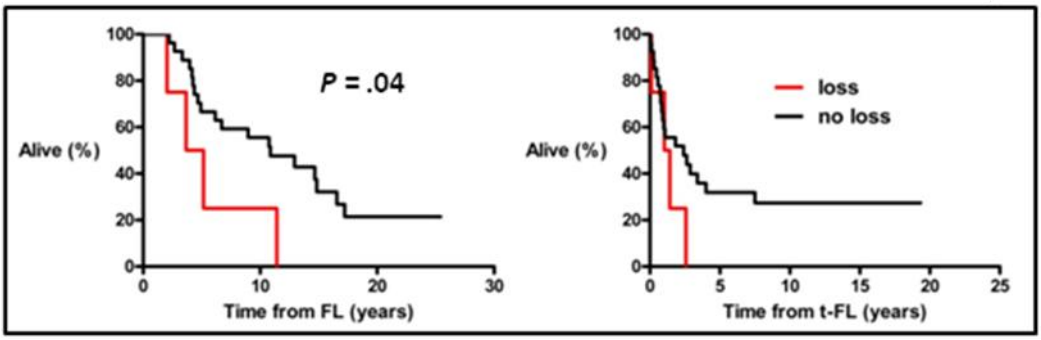
The "130" case set was used to determine clinical outcome associations and was drawn from the full cohort (n = 218). The "88" case set represents the remainder.

Appendix 5

Appendix 5 Table 1 Micro-array quality control

Sample	Call rate (%)	CQC	MAPD
1_F	99.01981	1.54	0.2930170
1_T	99.15756	2.15	0.2643848
2_G	99.31840	2.09	0.3020828
2_F	99.04686	2.64	0.2924905
2_T	99.34577	1.87	0.2876506
4_F	98.94682	1.43	0.3025794
4_T	98.24619	2.03	0.2746587
5_F	99.31224	2.24	0.2273867
5_T	98.70496	2.22	0.2546665
6_G	97.92727	1.13	0.3279828
6_F	98.29391	0.81	0.2904454
6_T	97.68443	1.13	0.3165253
7_F	99.49660	2.73	0.2632862
7_T	99.45483	2.15	0.2596367
8_G	98.55467	1.07	0.3157633
8_F	97.35989	0.58	0.3652553
8_T	96.58660	0.77	0.3472361
9_G	98.49553	2.22	0.3240928
9_F	99.18449	1.76	0.2928613
9_T1	99.05664	1.34	0.2831904
9_T2	97.93233	1.05	0.2918422
10_F	99.41042	2.59	0.2510170
10_T	97.62737	1.51	0.2674138
10_Fp	98.68066	2.24	0.2479891
11_F1	99.49221	2.62	0.3017525
11_F2	99.27992	2.00	0.2810147
11_T	99.07423	2.68	0.2768654
12_F	98.46838	1.76	0.2939214
12_T	98.26125	2.18	0.2981420
13_G	98.46661	1.73	0.3036524
13_F	98.37053	2.41	0.2512861
13_T	99.18559	2.29	0.2520840
15_G	99.17614	1.33	0.3125776
15_F	99.33269	1.62	0.2483445
15_T	97.07845	1.30	0.3108491
16_G	99.45263	3.18	0.3115786
16_F	97.92518	2.32	0.3272201
16_T	97.62990	2.03	0.3756202
17_G	98.11108	1.40	0.2789112
17_F	97.94717	1.54	0.3352565
17_T	96.66763	1.05	0.2838591
18_F	99.13789	1.19	0.2932741
18_T	99.31543	1.94	0.2720354
19_G	99.45944	1.95	0.2774876
19_F	96.83561	1.28	0.3585160
19_T	98.78543	0.85	0.3040063
20_G	99.56377	1.86	0.2901816
20_F	98.72914	2.16	0.2813541
20_T	98.49289	2.07	0.3115596
21_F	99.20252	2.16	0.3030906
21_T	98.34921	1.65	0.2881464
21_Fp	98.71936	1.58	0.2836809
22_F	97.16190	1.34	0.3092867
22_T	96.96171	0.94	0.3415176
23_G	99.61105	2.52	0.2806892
23_F1	99.04884	2.00	0.3153980
23_F2	99.33852	2.25	0.3068451
23_T	99.53278	2.61	0.2691558
24_G	99.58653	2.22	0.2653824
24_F	99.20451	2.18	0.3202963
24_T	98.12999	1.37	0.2984878
25_F	99.39700	2.09	0.2731955
25_T	99.20341	1.53	0.2669889
26_G	99.00794	1.84	0.2918555
26_F	99.31367	1.96	0.3031302
26_T	99.25903	1.53	0.2854231
27_F	99.05950	2.23	0.2697315
27_T	98.54588	1.76	0.2700410
28_G	99.06422	2.02	0.3252032
28_F	99.45923	2.67	0.2555414
28_T	99.42954	2.22	0.2393161
30_G	99.42316	2.25	0.2761419
30_F	98.88152	2.16	0.2905606
30_T	98.38889	2.23	0.2740566
31_G	96.71732	1.05	0.4027874
31_F	98.94242	2.30	0.2967747
31_T	98.86656	1.73	0.2877406
32_G	99.56180	2.37	0.2856510
32_F1	99.30344	1.94	0.3169082
32_F2	99.26112	2.21	0.2723175
32_F3	98.98606	1.71	0.2963441
32_F4	98.98299	1.96	0.2965111
32_T	99.14492	1.64	0.2838159
33_G	98.89185	1.16	0.2876970
33_F1	98.89097	1.48	0.2745235
33_F2	97.91837	1.63	0.3411722
33_T	97.42432	1.28	0.3606412
34_G	98.87656	0.97	0.3289226
34_F	99.54047	2.19	0.2629440
34_T	99.59258	2.21	0.2690359

Sample is the unique sample identifier consisting of the patient study number followed by either G (germline sample); F (FL sample); F1 / F2 / F3 / F4 (representing the first, second, third or fourth sample respectively in cases with more than one pre-transformation FL sample); T (t-FL sample); T2 (second t-FL sample); or Fp (FL sample from relapse after t-FL). CQC indicates contrast quality control; and MAPD median absolute pairwise difference. Sample (31_G) had a MAPD score just over the recommended screening threshold of 0.4 but was included as call rate, CQC and subsequent CN plots for case 31 were satisfactory. The calculation of Call rate, CQC and MAPD are as described in the main text.



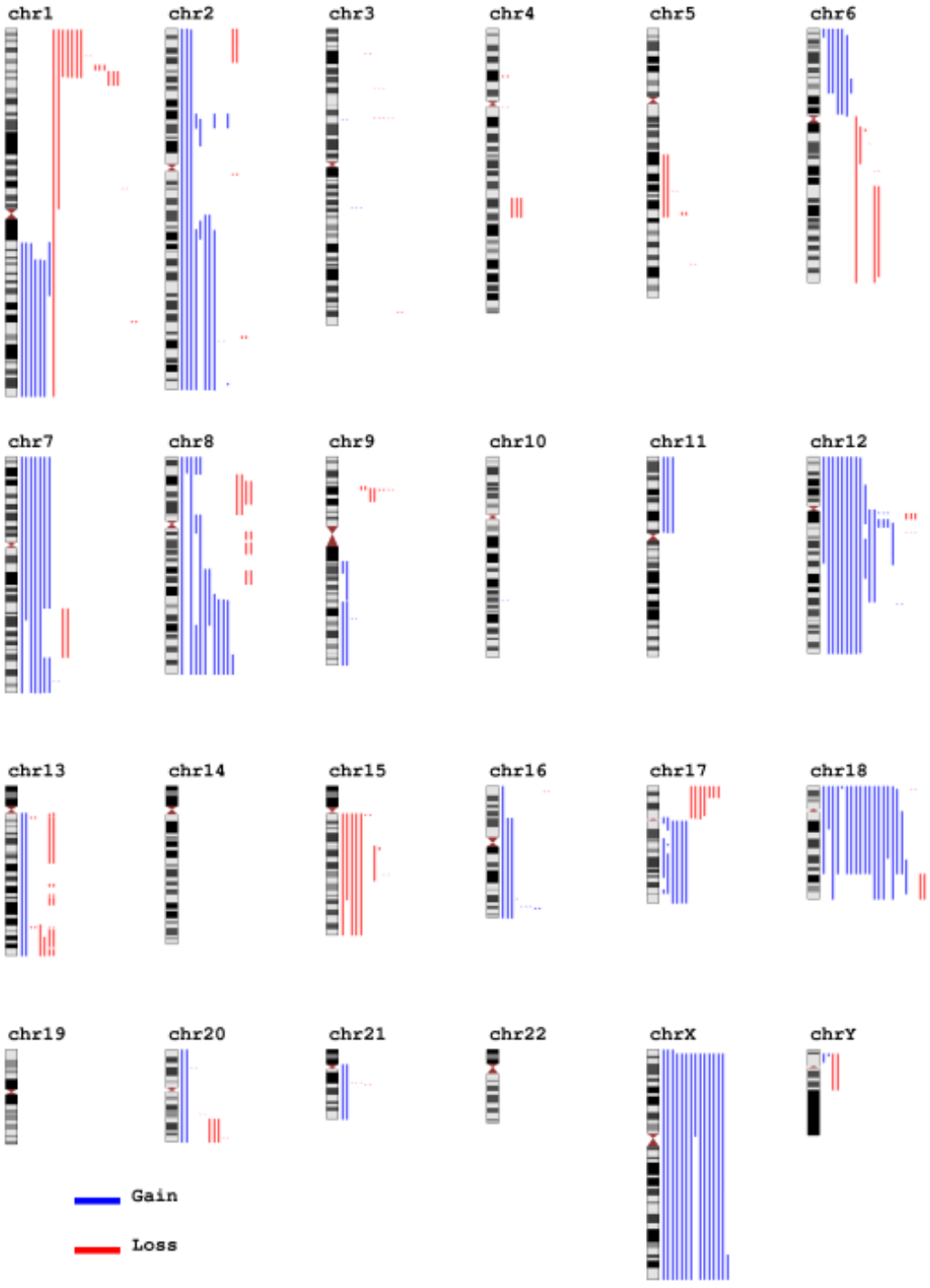
Appendix 5 Figure 1 Survival from FL or t-FL and *EI24 / CHEK1* loss

Kaplan-Meier curves comparing cases with *EI24 / CHEK1* locus loss on chromosome 11q (red line) against cases without loss (black line). Left, although not present in initial FL samples, loss *EI24 / CHEK1* associates with worse OS from FL diagnosis (although only 4 cases affected and a small patient cohort; n = 31). Right, no cases with loss *EI24 / CHEK1* at transformation survive more than 2.5 years from this event (which, although *P* = ns, may be a factor contributing to the difference in OS).

Appendix 5 Table 2 Determining the evolutionary route of FL / t-FL

Case	Evolution	Chromosome																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y
1	CPC																								
2	Direct																								
4	CPC																								
5	Either																								
6	CPC																								
7	CPC																								
8	CPC																								
9	CPC																								
10	CPC																								
11	CPC																								
12	CPC																								
13	Direct																								
15	CPC																								
16	CPC																								
17	Direct																								
18	CPC																								
19	Either																								
20	Direct																								
21	CPC																								
22	CPC																								
23	CPC																								
24	Direct																								
25	CPC																								
26	CPC																								
27	CPC																								
28	CPC																								
30	Direct																								
31	CPC																								
32	CPC																								
33	CPC																								
34	Direct																								

For the 31 study cases, the evolution mode is given based on summaries of the CN changes given for all chromosomes. Coloured boxes indicate CNAs are present on that chromosome in at least one sample from the respective case. Yellow indicates only the same chromosome CNAs are present across all available samples for the respective case; red indicates that (with or without common changes) sequentially acquired CNAs are present in later FL/ t-FL samples for the respective chromosome and case; and blue indicates that there is at least one non-sequential CNA for the respective chromosome and case. Consequently, presence of blue supports a CPC origin to disease evolution.



Appendix 5 Figure 2 Recurring CNAs in CPC cases

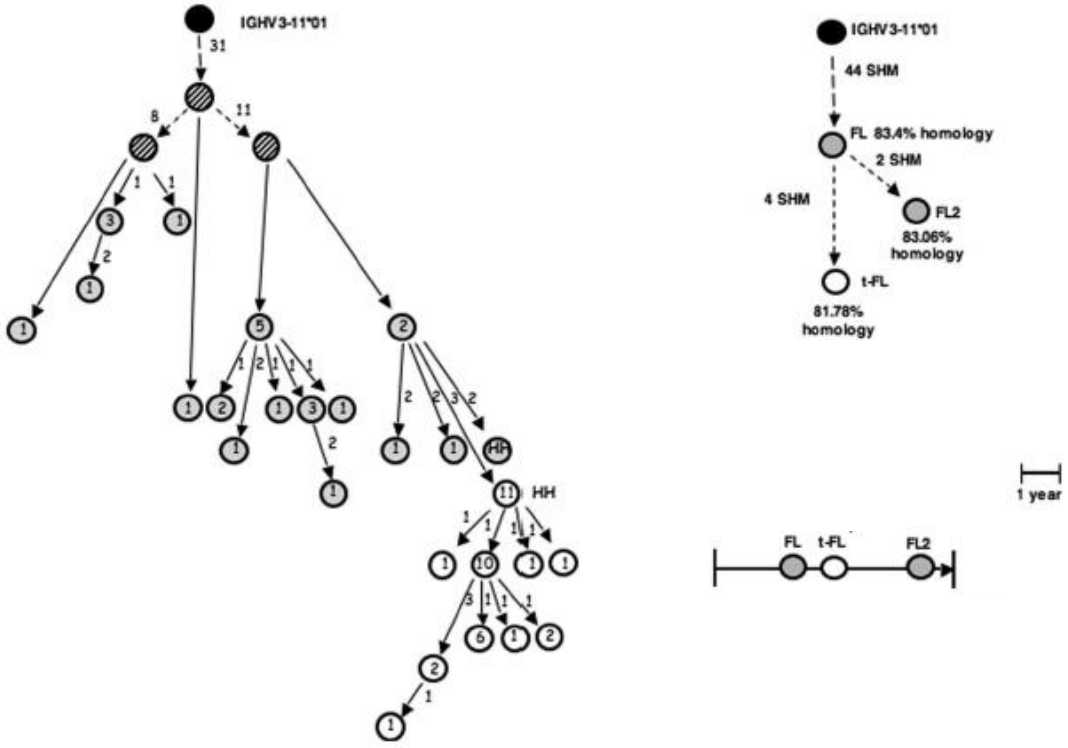
The plot layout is as described for **Figure 4.4**.

Appendix 6

Appendix 6 Table 1 *IGH-V* subgroup family primers

Forward primer		Reverse primer	
Name	**Sequence (5' to 3')**	**Name**	**Sequence (5' to 3')**
V_H1	GGCCTCAGTGAAGGTCTCCTGCAAG	J_H consensus	CTTACCTGAGGAGACGGTGACC
V_H2	GTCTGGTCCTACGCTGGTGAAACCC		
V_H3	CTGGGGGGTCCCTGAGACTCTCCTG		
V_H4	CTTCGGAGACCCTGTCCCTCACCTG		
V_H5	CGGGGAGTCTCTGAAGATCTCCTGT		
V_H6	TCGCAGACCCTCTCACTCACCTGTG		

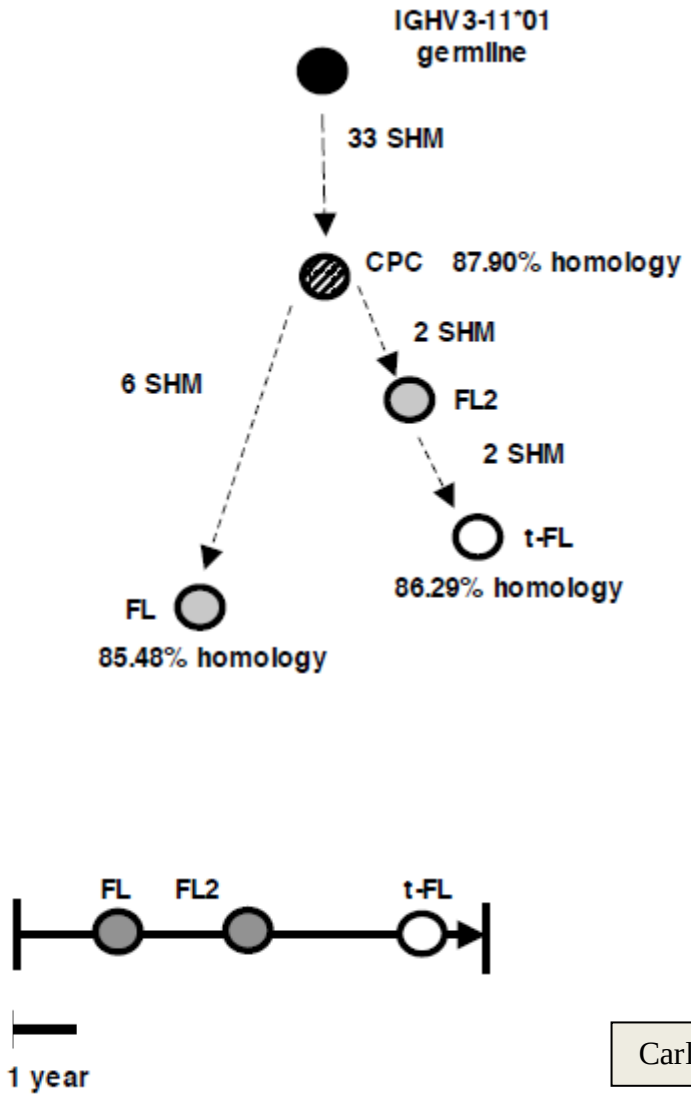
The BIOMED consortium designed primers on the three FR regions in VH segments. For FR1 there are six VH families which can be distinguished whereas for FR2 and FR3 there are seven. The six FR1 primers are used here as FR1 is situated farthest from the reverse primer in a consensus region of the J region common to all segments and this provides the largest PCR products so maximises the sensitivity for detecting SHM sites in the FL / t-FL samples.



Carlotti *et al*. 2009

Appendix 6 Figure 1 Direct evolution of t-FL from FL

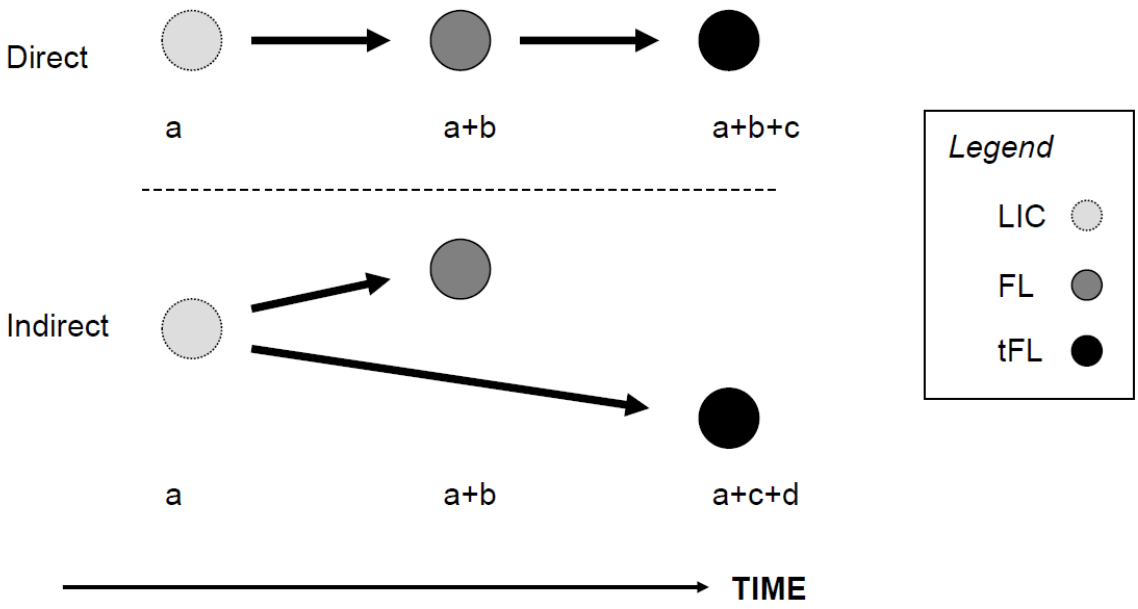
Shown are summaries of the SHM analysis for case 2. Symbols are as detailed in **Figure 5.3**. Left, genealogical tree constructed from cloning results shows that t-FL development is mirrored by sequential acquisition of SHM sites from FL to t-FL. Top right, HH major clones indicate that both FL2 and t-FL arise by direct evolution from FL. In combination with the time line of the three disease episodes, bottom right, this indicates that the FL clone persists after t-FL and gives rise to the FL2 clone (left hand vertical bar represents FL diagnosis and right hand vertical bar represents patient death).



Carlotti *et al*. 2009

***Appendix 6* Figure 2 Direct and indirect evolution may both occur in FL / t-FL**

Summary of *IGH-V* sequences and SHM sites determined for the major disease clone of case 26 by HH analysis. Symbols are as detailed in **Figure 5.3** and ***Appendix 6* Figure 1**. Top, divergent (indirect) evolution from a CPC gives rise to FL and FL relapse (FL2) whereas transformation arises by direct evolution from FL2. Bottom, time-line of the three disease episodes indicates the CPC is present for at least 18 months before FL2 presents.



***Appendix 6* Figure 3 Models of direct and indirect evolution**

A summary of the two routes of FL / t-FL evolution identified by SHM analysis. LIC indicates lymphoma initiating cell (such as an early t(14;18) containing cell).
For direct evolution, the sequential acquisition of different SHMs over time is represented by a, b and c accompanied by parallel evolution from the potential precursor to FL and on to t-FL.
In the case of indirect evolution, the non-sequential pattern of acquired genetic events is illustrated by a, b, c and d. A common event, a, is followed by b in FL. However, subsequent t-FL does not arise from FL but from an earlier clone which acquires c and d (but not b) in progression to transformation. For this model, the LIC represents a common progenitor cell (CPC) which gives rise to FL and t-FL by divergent evolution.

Appendix 6 Table 2 Novel glycosylation sites in CPC / FL / t-FL

Case	Clone[#]	Glycosylation sites			
		Number[§]	Region(s)	αα number	αα motif*
1	FL t-FL CPC	3	CDR2 / FR3 / CDR3	58 / 67 / 109	NHS / NYS / NNS
2	FL t-FL FL2	1	CDR1	30	NFS
4	FL t-FL CPC	1	CDR2	56	NIS
5	FL t-FL FL2	1	CDR3	108	NLS
6	FL t-FL CPC	1	CDR3	108	NLS
7	FL t-FL CPC	2 3 1	FR2 / FR3 FR1 / FR2 / FR3 FR3	39 / 69 25 / 39 / 69 69	NWT / NSS NVS / NWT / NSS NSS
8	FL t-FL CPC	0	-	-	-
9	FL t-FL	0	-	-	-
12	FL t-FL CPC	1	CDR2	59	NKT
15	FL t-FL CPC	0	-	-	-
16	FL t-FL CPC	4	CDR1 / CDR2 / CDR2 / CDR3	30 / 56 / 62 / 116	NIT / NIS / NKT / NCT
20	FL t-FL	1	FR2	55	NIS
21	FL t-FL FL2	1	FR2	39	NMT
23	FL t-FL CPC	1	CDR2	56	NIT
25^	FL t-FL CPC	1	CDR3	108	NFS
26	FL t-FL FL2 CPC	4	CDR1 / FR2 / CDR2 / FR3	30 / 39 / 56 / 85	NFS / NMS / NIS / NNS
27	FL t-FL CPC	3	FR3 / CDR3 / CDR3	83 / 108 / 112	NST / NQT / NKS
Father Son	FL t-FL CPC	1 0 0	CDR2 - -	58 - -	NQS - -

Glycosylation sites introduced by SHMs are indicated for samples from 17 / 18 sequential cases and the father / son pair. Case 11 is excluded as its DH-JH region is deleted. Protein region of the glycosylation site and amino acid (αα) position (number) of the predicted change are indicated. The sites consist of a three αα motif (N-x-S/T, where x is any αα except proline / aspartic acid / glutamic acid (Co *et al.* 1993; Coloma *et al.* 1999)) and are detailed. Sites were identified in 11 / 12 CPC sequences for the sequential cases. The remaining case (15) showed no sites in any samples. In the father / son pair a single site occurred only in the father's tumour. Case 7 was the sole case to show differing acquisition of additional glycosylation sites in FL and t-FL samples.

[#]The major clones identified by HH analysis.

[§]If one number that applies to all samples for the case.

**C indicates cysteine; F, phenylalanine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; Q, glutamine; S, serine; T, threonine; V, valine; and Y, tyrosine.*