

---

Doctoral

Science

---

2012-4

## Molecular Genetic and DNA Methylation Profiling of Chronic Lymphocytic Leukaemia: a Focus on Divergent Prognostic Subgroups and Subsets

Nicola Cahill  
*Technological University Dublin*

Follow this and additional works at: <https://arrow.tudublin.ie/sciendoc>



Part of the [Life Sciences Commons](#)

---

### Recommended Citation

Cahill, N. (2012). *Molecular Genetic and DNA Methylation Profiling of Chronic Lymphocytic Leukaemia: a Focus on Divergent Prognostic Subgroups and Subsets*. Doctoral Thesis. Technological University Dublin. doi:10.21427/D7V01B

This Theses, Ph.D is brought to you for free and open access by the Science at ARROW@TU Dublin. It has been accepted for inclusion in Doctoral by an authorized administrator of ARROW@TU Dublin. For more information, please contact [arrow.admin@tudublin.ie](mailto:arrow.admin@tudublin.ie), [aisling.coyne@tudublin.ie](mailto:aisling.coyne@tudublin.ie), [vera.kilshaw@tudublin.ie](mailto:vera.kilshaw@tudublin.ie).

**Molecular Genetic and DNA Methylation  
profiling of Chronic Lymphocytic Leukaemia**

*A focus on divergent prognostic Subgroups and Subsets*

**By**

**Nicola Cahill**

**For award of PhD**

**April 2012**



Department of Biological Sciences,  
Dublin Institute of Technology,  
Kevin Street,  
Dublin 8

Supervisors

Dr. Fergus Ryan  
Professor Richard Rosenquist  
Dr. Larry Mansouri

*Dedicated to my family*

# LIST OF PAPERS

This thesis is based on the following Papers, which are referred to in the text by their Roman numerals.

- I** Cahill N, Sutton LA, Jansson M, Murray F, Mansouri L, Gunnarsson R, Ryan F, Smedby KE, Geisler C, Juliusson G and Rosenquist R. (2012) IGHV3-21 gene frequency in a Swedish cohort of newly diagnosed chronic lymphocytic leukemia patients. *Clinical Lymphoma Myeloma and Leukemia* (Accepted)
- II** Marincevic M\*, Cahill N\*, Gunnarsson R\*, Isaksson A, Mansouri M, Göransson H, Rasmusson M, Jansson M, Ryan F, Karlsson K, Adami HO, Davi F, Jurlander J, Juliusson G, Stamatopoulos K, Rosenquist R. (2010) High-density screening reveals a different spectrum of genomic aberrations in chronic lymphocytic leukemia patients with "stereotyped" IGHV3-21 and IGHV4-34 B cell receptors. *Haematologica*, **95(9): 1519-25**
- III** Kanduri M\*, Cahill N\*, Göransson H, Enström C, Ryan F, Isaksson A, Rosenquist R. (2010) Differential genome-wide array-based methylation profiles in prognostic subsets of chronic lymphocytic leukemia. *Blood*, **115(2):296-305**
- IV** Cahill N, Bergh AC, Göransson-Kultima H, Mansouri L, Isaksson A, Ryan F, Smedby KE, Sundström C, Juliusson G, Rosén A and Rosenquist R. 450K DNA methylation analysis of chronic lymphocytic leukaemia reveals global methylation to be relatively stable over time and remarkably similar in cells derived from resting and proliferative compartments. *Manuscript*

\* Equal first authors

Reprints were made with the permission of the publishers

# ABSTRACT

Advancements in prognostication have improved the subdivision of chronic lymphocytic leukaemia (CLL) into diverse prognostic subgroups. In CLL, IGHV unmutated and IGHV3-21 genes are associated with a poor-prognosis, conversely, IGHV mutated genes with a favourable outcome. The finding of multiple CLL subsets expressing ‘stereotyped’ B-cell receptors (BCRs) has suggested a role for antigen(s) in leukemogenesis. Patients belonging to certain stereotyped subsets share clinical and biological characteristics, yet limited knowledge exists regarding the genetic and epigenetic events that may influence their clinical behaviour. This thesis aimed to, further investigate Swedish IGHV3-21-utilising patients, screen for genetic and DNA methylation events in CLL subgroups/subsets and study DNA methylation over time and within different CLL compartments.

In paper I, IGHV gene sequencing of 337 CLL patients from a Swedish population-based cohort revealed a lower (6.5%) IGHV3-21 frequency relative to previous Swedish hospital-based studies (10.1-12.7%). Interestingly, this frequency remained higher compared to other Western CLL (2.6-4.1%) hospital-based cohorts. Furthermore, we confirmed the poor-outcome for IGHV3-21 patients to be independent of mutational and stereotypy status.

In paper II, genomic events in stereotyped IGHV3-21-subset #2, IGHV4-34-subset #4 and subset #16 and their non-stereotyped counterparts were investigated via SNP arrays (n=101). Subset #2 and non-subset #2 carried a higher frequency of

events compared to subset #4. A high frequency of del(11q) was evident in IGHV3-21 patients particularly subset #2 cases, which may partially explain their poor-prognosis. In contrast, the lower prevalence of aberrations and absence of poor-prognostic alterations may reflect the inherent low-proliferative disease seen in subset #4 cases.

In papers III and IV, differential methylation profiles in IGHV mutated and IGHV unmutated patients were identified using DNA-methylation microarrays. CLL prognostic genes (*CLLU1*, *LPL*), tumor-suppressor genes (TSGs) (*ABI3*, *WISP3*) and genes belonging to TGF- $\beta$  and NF $\kappa$ B/ TNFR1 pathways were differentially methylated between the subgroups. Additionally, the re-expression of methylated TSGs by use of methyl and deacetyl inhibitors was demonstrated. Interestingly, analysis of patient-paired diagnostic/follow-up samples and patient-matched lymph node (LN) and peripheral blood (PB) cases revealed global DNA methylation to be relatively stable over time and remarkably similar within the different compartments. Altogether, this thesis provides insight into the aberrant genomic and DNA methylation events in divergent CLL subgroups. Moreover this thesis helps distinguish the extent to which DNA methylation changes with respect to time and microenvironment in CLL.

# DECLARATION

I certify that this thesis which I now submit for examination for the award of PhD, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for another award in any other third level institution.

The work reported on in this thesis conforms to the principles and requirements of the DIT's guidelines for ethics in research.

DIT has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be acknowledged.

Signature

A handwritten signature in black ink, appearing to read "Nicola Costantini". The signature is written in a cursive style with a horizontal line underneath.

Date

# ACKNOWLEDGEMENTS

This work was carried out at the Department of Immunology Genetics and Pathology, Rudbeck Laboratory, Uppsala University in association with the School of Biological Sciences, Dublin Institute of Technology, Ireland. Financial support was provided by the Nordic Cancer Union, the Swedish Cancer Society, the Swedish Medical Council and ABBEST Ireland.

First of all I would like to sincerely thank my supervisors Dr. Fergus Ryan and Prof Richard Rosenquist. Thanks for taking a chance on me and for giving me the opportunity undertake such rewarding projects.

Thank you to my most recent co-supervisor Larry Mansouri. You have always encouraged me whenever I needed it. Thanks for all the assistance, words of advice and suggestions I really appreciate it!

To my lab lads and lassies, both past and present, Meena, thanks for allowing me to collaborate with you at the start of my studies, I will never forget our most interesting discussions!!!. Fiona, thanks for keeping me going, making me laugh albeit sometimes at my own expense!! You're all right for a south-sider! Maria! Spiggelty wiggelty!!! Thanks for taking an interest in using Irish slang you are only 'savage' at it now! Marie, thanks for your good humour and sound advice!!! Rebeqa! I really enjoyed working with you on the SNP Paper, thank you for keeping me focused! Lesley! Thanks for everything!! You truly are a great friend and a great agony aunt!!.



Thanks for taking me in when I first came to Uppsala and keeping me forever entertained. Mi, thanks a million for all the fun times and laughs!! To Arifin, Ja, Ingrid, Anna, Djana, Cecilia and Andreas thanks for making the lab such a great place to work.

I would also like to thank Christer Sundstrom for all his help in finding samples and for “donating” the occasional few millilitres of blood!!! Special thanks to our Swedish collaborators Lotta Bergh and Anders Rosen!! Thanks to our other European collaborators. Kostas Stamatopolous, Fred Davi, Christian Geisler and Gunnar Juliusson for all their contributions! Sincere thanks to Hanna Göransson Kultima and Isaksson for all their hard bio-informatic work.

To all the Rudbeck/Akademiska/Ångstrom/BMC gang!! Where do I start.... you know who all you trouble makers are....but just in case others don't know, I feel its my duty to warn them....Jelena, Lucy, Johan S, Chris, Johan K, Sara K, Jess, Ammar, Paul, Demet, and Daniel and Tatjana. Thanks for keeping me sane!!!! More importantly thanks for making my 4 years here enjoyable. I will miss all the parties, laughs and friendship!

To Antonia, Charlotte, Prason, Lotta S, Pernilla, Helena JW, Fredrik Ö, Chandra's army- (Tanmoy, Deb, Matthieu, Sanhita Gaurav), Tobias B, Grzegorz, Viktoria A Spyros, Xiang, Umash, Jimmy, Sara L, Yanara, Lothar, Marie H, Sonia, the dutch dudes, Eva H, Annika H, the genome centre girlies, the teaching crew-Viktoria, Angie, Madhu, Rachel, Lena Å and Hamid ... big thanks for making Rudbeck a cool place to work!!!

I would like to thank Joe Vaughan, Brid Ann Ryan and Paddy Mc Hale (DIT) for helping me and encouraging me over my undergraduate and post graduate years! Also thank you to Joanne Richards of Baylor college of Medicine.

To my Irish mates, Dave, Eileen, Sarah, Gav, Lynne, Cliona and Fats!! Thanks for coming over to visit me and staying friends with the NERD!!!

Big thanks to all my Swedish buddies!!!

To my Swedish family the “Janssons”. Thanks for taking me in as one of your own. You have truly looked after me!!! Special thanks to pappa. Tack för alla sms.!!!

To my family, thanks for always being there and never doubting me! Thanks for your never ending support even if you still don’t understand what exactly what I have been doing over here!!!..... NO... dad I am not still writing my essay for school... this is it... what you think?’’

Last but by no means least...Mattias! I know I might regret putting this in writing...but you have the patience of a saint. I can simply say that without you I wouldn’t have made it.!! Thanks for everything xxx

## ABBREVIATIONS

|               |   |
|---------------|---|
| 5'aza         | 5-azacytidine   |
| <i>ABI3</i>   | ABI family member 3   |
| aCGH          | Array comparative genomic hybridisation                       |
| <i>ADORA3</i> | Adenosine A3 receptor   |
| AID           | Activation induced deaminase                                  |
| AKT           | V-akt murine thymoma viral oncogene homolog                   |
| AML           | Acute myeloid leukaemia                                       |
| AMP           | Adenosine monophosphate                                       |
| ANGPT2        | Angiopoietin 2  |
| ATM           | Ataxia telangiectasia mutated                                 |
| BAC           | Bacterial artificial chromosome                               |
| BCL10         | B-cell CLL/lymphoma 10  |
| BCL2          | B-cell CLL/lymphoma 2   |
| BCR           | B cell receptor   |
| BMSCs         | Bone marrow stromal cells                                     |
| C             | Constant  |
| CARD15        | Caspase recruitment domain family, member 15                  |
| CCL22         | C-C motif chemokine 22  |
| CD            | Cluster of differentiation                                    |
| CD40L         | CD40 ligand   |
| CDK           | Cyclin dependent kinase                                       |
| CDR           | Complementarity determining region                            |
| CGH           | Comparative genomic hybridization                             |
| CHARM         | Comprehensive high-throughput arrays for relative methylation |
| CLL           | Chronic lymphocytic leukaemia                                 |
| CLLU1         | Chronic lymphocytic leukaemia up-regulated gene 1             |
| CMV           | Cytomegalovirus   |
| CNA           | Copy number alteration  |
| CNAT          | Copy number analysis tool                                     |
| CNN-LOH       | Copy number neutral loss of heterozygosity                    |
| CNV           | Copy number variation   |
| CpG           | Cytosine-phosphate-guanine                                    |
| CXCR4         | C-X-C chemokine receptor 4                                    |
| D             | Diversity   |
| DAC           | 5-aza 2'-deoxycytidine  |
| DAPK1         | Death-associated protein kinase 1                             |
| ddNTPs        | Dideoxy nucleotide triphosphates                              |
| DLEU7         | Deleted in lymphocytic leukaemia 7                            |
| DLEU1         | Deleted in lymphocytic leukaemia 1                            |
| DNA           | Deoxyribonucleic acid   |
| DNA-PK        | DNA dependent protein kinase                                  |
| DNMT          | DNA methyltransferase   |

|        |   |
|--------|---|
| DNMT3B | DNA methyltransferase 3                                 |
| DNMT3L | DNA methyltransferase 3 ligand                          |
| EBV    | Epstein Barr virus                                      |
| ERK    | Extracellular signal related kinase                     |
| ESCs   | Embryonic stem cells                                    |
| EXO    | Exonuclease   |
| EZH2   | Enhancer of zeste homolog 2                             |
| FAS    | TNF receptor superfamily, member 6                      |
| FISH   | Flourescent in-situ hybridisation                       |
| FOXE1  | Forkhead Box E1   |
| FWR    | Framework region  |
| GC     | Germinal center   |
| GLRB   | Glycine receptor Beta                                   |
| GRM7   | Glutamate receptor metabotropic 7                       |
| H3K27  | Histone 3 lysine 27                                     |
| H3K4   | Histone 3 lysine 4                                      |
| H3K9   | Histone 3 lysine 9                                      |
| HELP   | HpaII tiny fragment Enrichment by Ligation-mediated PCR |
| HIP1R  | Huntingtin interacting protein 1 related                |
| HPLC   | High pressure liquid chromatography                     |
| IBTK   | Inhibitor of Bruton agammaglobulinemia tyrosine kinase  |
| IG     | Immunoglobulin  |
| IGH    | Immunoglobulin heavy                                    |
| IGHD   | Immunoglobulin heavy diversity                          |
| IGHJ   | Immunoglobulin heavy joining                            |
| IGHV   | Immunoglobulin heavy variable                           |
| IGK    | Immunoglobulin kappa                                    |
| IGL    | Immunoglobulin lambda                                   |
| IL17RC | Interleukin 17 receptor C                               |
| IL2    | Interleukin 2   |
| IPF1   | Pancreatic and duodenal homeobox 1                      |
| IWCLL  | International workshop of chronic lymphocytic leukaemia |
| J      | Joining   |
| JAIRD2 | Jumonji- and ARID-domain-containing protein             |
| LC     | Light chain   |
| LEF1   | Lymphoid enhancer binding factor 1                      |
| LINES  | Long Interspersed Nuclear Elements                      |
| LMO2   | LIM domain only 2                                       |
| LN     | Lymph node  |
| LPL    | Lipoprotein lipase                                      |
| LPS    | Lipo-polysaccharide                                     |
| MAPK   | Mitogen activated protein kinase                        |
| MBD2   | Methyl-CpG-binding domain protein 2                     |
| MBL    | Monoclonal B lymphocytosis                              |
| Mbp    | Mega basepair   |
| MCL-1  | Induced myeloid leukaemia cell differentiation protein  |
| MDM2   | Mdm2 p53 binding protein homolog (mouse)                |

|              |   |
|--------------|---|
| MDR          | Minimal deleted region                                    |
| me/ me2/ me3 | mono/di/tri methylation                                   |
| MEACS        | Non-muscle myosin heavy chain IIA exposed apoptotic cells |
| MECP2        | Methyl CpG binding protein 2                              |
| MeDIP        | Methyl DNA immunoprecipitation                            |
| MI           | Methylation index   |
| miRNA        | MicroRNA  |
| MLL          | Mixed-lineage leukaemia                                   |
| MRE          | methyl sensitive restriction enzyme                       |
| MSP          | Methylation specific PCR                                  |
| MYCN         | V-myc myelocytomatosis viral related oncogene             |
| MYD88        | Myeoid differentiation primary response gene              |
| MYF6         | Myogenic factor 6   |
| MZ           | Marginal zone   |
| n            | Non-templated   |
| NFkB         | Nuclear factor kappa B                                    |
| NGFR         | Neural growth factor receptor                             |
| NHL          | Non-Hodgkin lymphoma                                      |
| NK           | Natural killer  |
| NLCs         | Nurse like cells  |
| NOTCH1       | Notch homologue 1   |
| OS           | Overall survival  |
| p            | Palidromic  |
| P2RY14       | Purinergic receptor P2Y                                   |
| PAK5         | p21 protein (Cdc42/Rac)-activated kinase 7                |
| PB           | Peripheral blood  |
| PBMC         | Peripheral blood mononuclear cells                        |
| PC           | Proliferation center                                      |
| PCG7         | Primordial germ cell 7                                    |
| PDI          | Protein disulfide isomerase                               |
| PI3K         | Phosphoinositide-3-kinase                                 |
| PLD1         | Phospholipase D1  |
| PPP1R3A      | Protein phosphatase 1 regulatory (inhibitor) subunit 3A   |
| PRAME        | Preferentially expressed antigen in melanoma              |
| PREs         | Polycomb response elements                                |
| PRF1         | Perforin 1  |
| RAG          | Recombination activating gene                             |
| RASGRP3      | RAS guanyl releasing protein 3                            |
| RE           | Restriction enzyme  |
| RLGS         | Restriction landmark genome scanning                      |
| RQ-PCR       | Real time quantitative PCR                                |
| RSS          | Recombination signal sequences                            |
| SAHA         | Suberoylanilide hydroxamic acid                           |
| SAP          | Shrimp alkaline phosphatase                               |
| SDF1         | Stromal cell derived factor 1                             |
| SF3B1        | Splicing factor 3b, subunit 1                             |
| SHM          | Somatic hypermutation                                     |
| SLC22A18     | Solute carrier family 22, member 18                       |
| SLL          | Small lymphocytic lymphoma                                |

|         |  |
|---------|--|
| SNP     | Single nucleotide polymorphism                   |
| Sp1     | Specificity factor 1                             |
| T       | Thymine  |
| TBR     | Eomesodermin                                     |
| TBX3    | T-box 3  |
| TdT     | Terminal deoxynucleotidyl transferase            |
| TNF     | tumour necrosis factor                           |
| TP53    | Tumour protein p53 gene                          |
| Trx     | Thioredoxin                                      |
| TSA     | Tricostatin A                                    |
| TSG     | Tumour suppressor gene                           |
| TSS     | Transcription start site                         |
| TTT     | Time to treatment                                |
| TWIST-2 | Twist homolog 2                                  |
| USP     | Unmethylation specific PCR                       |
| UTR     | Untranslated region                              |
| V       | Variable   |
| VCAM1   | Vascular cell adhesion molecule 1                |
| VHL     | von Hippel-Lindau                                |
| VLA4    | Very late antigen 4                              |
| WHO     | World health organisation                        |
| WISP3   | WNT1 inducible signaling pathway protein 3       |
| XPO1    | Exportin 1                                       |
| ZAP-70  | Zeta-chain (TCR) associated protein kinase 70kDa |
| ZNF 540 | Zinc finger protein 540                          |

# CONTENTS

|                        |       |
|------------------------|-------|
| Title .....            | (i)   |
| Dedication .....       | (ii)  |
| List of papers .....   | (iii) |
| Abstract .....         | (iv)  |
| Declaration .....      | (vi)  |
| Acknowledgements ..... | (vii) |
| Abbreviations .....    | (x)   |
| Contents .....         | (xiv) |

|  |    |
|--|----|
| 1) Introduction.....   | 1  |
| 1.1 Chronic lymphocytic leukaemia .....  | 2  |
| 1.1 <i>i</i> Epidemiology and aetiology.....   | 2  |
| 1.1 <i>ii</i> Diagnosis .....  | 3  |
| 1.1 <i>iii</i> Staging .....   | 4  |
| 1.1 <i>iv</i> Infectious complications in CLL.....   | 5  |
| 1.1 <i>v</i> Treatment.....  | 5  |
| 1.2 Normal B cells and B cell development .....  | 7  |
| 1.2 <i>i</i> Structure of the immunoglobulin molecule.....   | 8  |
| 1.2 <i>ii</i> VDJ recombination.....   | 10 |
| 1.2 <i>iii</i> The germinal center reaction and somatic hypermutation .....                        | 12 |
| 1.2 <i>iv</i> Immunoglobulin diversity .....   | 14 |
| 1.3 The CLL cell and CLL pathogenesis .....  | 14 |
| 1.3 <i>i</i> CLL cell of origin.....   | 15 |
| 1.3 <i>ii</i> The microenvironment in CLL.....   | 17 |
| 1.3 <i>iii</i> Evidence of antigen(s) involvement in CLL pathogenesis .....                        | 21 |
| 1.3 <i>iv</i> The molecular nature of antigen(s) in CLL .....                                      | 22 |
| 1.4 IGHV3-21 and IGHV4-34 CLL subgroups: IG sequences,<br>antigens and clinical implications ..... | 23 |
| 1.4 <i>i</i> IGHV3-21 utilising patients .....   | 23 |
| 1.4 <i>ii</i> IGHV4-34 utilising patients .....  | 26 |
| 1.5 Prognostic markers in CLL.....   | 29 |
| 1.5 <i>i</i> IGHV gene mutational status .....   | 29 |
| 1.5 <i>ii</i> Protein based markers: CD38 and ZAP70 expression .....                               | 30 |
| 1.5 <i>iii</i> RNA based markers: <i>LPL</i> and <i>CLLUI</i> expression.....                      | 31 |
| 1.5 <i>iv</i> Genomic alterations .....  | 32 |
| 1.6 Genomic profiling technologies.....  | 42 |
| 1.7 DNA methylation.....   | 45 |

|  |     |
|--|-----|
| 1.7 i Establishment and maintenance of DNA methylation .....   | 46  |
| 1.7 ii DNA methylation regulation .....  | 48  |
| 1.7 iii DNA methylation in tumorigenesis.....  | 51  |
| 1.7 iv Hypermethylation .....  | 52  |
| 1.7 v Hypomethylation .....  | 53  |
| 1.7 vi DNA methylation and mutagenesis.....  | 54  |
| 1.7 vii Role of DNA methylation in CLL.....  | 54  |
| 1.7 viii Disruption of DNA methylation regulatory mechanisms in<br>tumourigenesis .....                        | 56  |
| 1.7 ix Use of DNA methylation markers in clinical routine.....   | 57  |
| 1.8 DNA methylation technologies.....  | 59  |
| 1.9 Epigenetic based treatment .....   | 64  |
| 2) Thesis Aims .....   | 65  |
| 3) Materials & Methods .....   | 66  |
| 3.1 Patient samples and patient characteristics .....  | 66  |
| 3.2 Ethical Approval .....   | 66  |
| 3.3 IG gene sequencing.....  | 68  |
| 3.4 250K Affymetrix SNP array .....  | 69  |
| 3.5 Illumina Infinium HumanMethylation27 and<br>HumanMethylation450 BeadChip array .....                       | 71  |
| 3.6 Bioinformatic analysis of the Illumina human27<br>and 450 DNA methylation array data .....                 | 74  |
| 3.7 Methylation-specific PCR analysis.....   | 75  |
| 3.8 Real-time quantitative PCR analysis .....  | 76  |
| 3.9 Bisulfite sequencing.....  | 76  |
| 3.10 5-aza-2'-deoxycytidine and Trichostatin A treatment .....   | 77  |
| 3.11 Pyrosequencing.....   | 77  |
| 3.12 Statistical analysis.....   | 78  |
| 4) Results & Discussion .....  | 80  |
| 4.1 Paper I: IGHV3-21 gene frequency in a Swedish population-based cohort.....                                 | 80  |
| 4.2 Paper II: High density array screening for genomic events in stereotyped<br>subsets of CLL .....           | 85  |
| 4.3 Paper III: Array-based methylation profiling in prognostic subsets of CLL ....                             | 90  |
| 4.4 Paper IV: DNA methylation screening of CLL samples over time and within<br>different CLL compartments..... | 95  |
| 5) Conclusions.....  | 103 |
| 5.1 Paper I.....   | 103 |
| 5.2 Paper II.....  | 103 |
| 5.3 Paper III .....  | 104 |
| 5.4 Paper IV .....   | 105 |
| 6) References.....   | 106 |
| 7) Papers I-IV .....   | 106 |



# 1: INTRODUCTION

Chronic lymphocytic leukaemia (CLL) is a malignant clonal disorder characterised by the accumulation of mature, functionally impaired neoplastic CD5<sup>+</sup> B cells. Originating in the bone marrow these cells slowly progress to infiltrate the peripheral blood, lymph nodes and spleen. At both the clinical and biological level, CLL is highly heterogeneous in nature. This heterogeneity gives rise to highly divergent disease courses, ranging from an asymptomatic disease in some patients to an aggressive and rapidly fatal condition in others.<sup>1</sup> Determination of the events contributing to CLL heterogeneity are key to identifying new prognostic markers that in turn will lead to better subdivision of CLL patients, allowing better implementation of optimal treatment regimes.

Over the past twenty years, studies of the B cell receptor (BCR) molecule on CLL cells have denoted this molecule as a key player in CLL leukemogenesis and prognostication. For instance, the mutation status of the immuno-globulin heavy chain variable region gene (IGHV) has been established as one of the most reliable prognostic markers in CLL.<sup>2,3</sup> Moreover, studies characterising the molecular structure of immunoglobulin (IG) gene re-arrangements support the notion that certain CLL subsets sharing a distinct clinical outcome present with highly similar IG rearrangements that may recognise a common antigen and confer a growth advantage to the CLL clone.<sup>4-6</sup> In certain cases, other factors such as recurrent genomic alterations are known to contribute to CLL pathogenesis,<sup>7</sup> however knowledge of

aberrant DNA methylation events in this disease is limited. Hence, this thesis will focus on the relationship between the BCR structure, global genomic and DNA methylation events that may contribute to CLL leukemogenesis in different prognostic subgroups of CLL patients.

## **1.1: Chronic lymphocytic leukaemia**

### *1.1 i: Epidemiology and aetiology*

Today, CLL represents ~30-40% of all leukaemias and is the most commonly encountered adult leukaemia in the Western world.<sup>8</sup> Demonstrating a gender bias, the incidence of CLL is almost twice that in men as in women.<sup>9,10</sup> In Sweden alone, ~500 cases are diagnosed annually.<sup>11</sup> The median age of CLL diagnosis is ~70 years, however a third of patients are younger than 60 years of age at diagnosis.<sup>10,11</sup> First-degree relatives of patients with CLL can have a 7-8.5 fold increased risk of developing CLL and a 2-fold increase for other lymphoproliferative diseases, suggesting a shared genetic component in familial CLL cases.<sup>12-15</sup>

To date, the aetiology of CLL is still largely unknown, however a number of linkage studies have implicated some susceptibility loci in CLL. For instance, recent investigations have identified susceptibility loci at 2q13 2q37.1, 6p25.3, 11q24.1, 15q23 and 19q13.32 thus providing the first evidence for the existence of common, low-penetrance susceptibility loci in sporadic CLL.<sup>16-18</sup> Monoclonal B lymphocytosis (MBL), a condition characterised by the presence of  $<5.0 \times 10^9$  clonal B cells per litre of peripheral blood (PB), is now suspected to represent the pre-clinical stage of

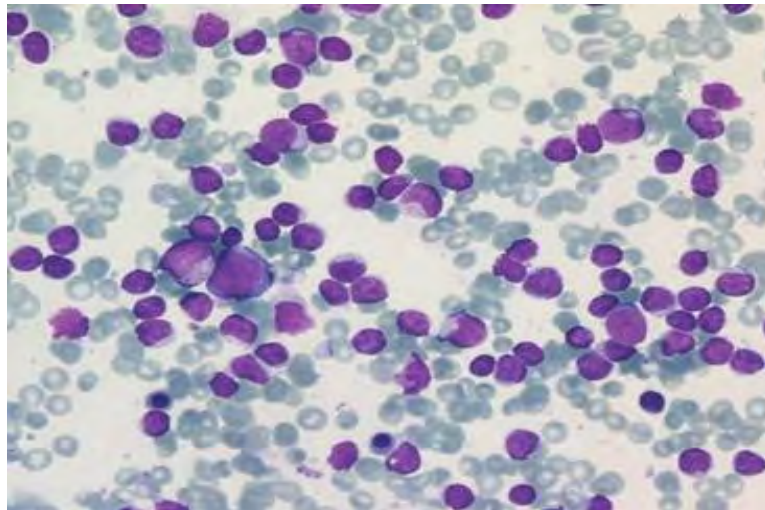
CLL. MBL is common among the general population, occurring in 3-5% of individuals over the age of 50 years.<sup>19-22</sup> MBL can be categorised into three immunophenotypes; i) CLL-like MBL, CD5 positive with low CD20 expression, ii) double positive CD5/CD20 MBL, resembling atypical CLL and iii) CD5 negative MBL.<sup>20,22</sup> CLL-like MBL is the most frequent MBL immunophenotype encountered. Nevertheless, it must be kept in mind, that in the vast majority of MBL cases CLL will not ultimately emerge.<sup>21</sup> Unlike other forms of leukaemia, only suggestive relationships between exposure to pesticides and certain solvents such as benzene have been indicated in CLL aetiology, thus the effect of hazardous environmental substances warrants further investigation.<sup>23,24</sup>

CLL is thought to be driven by a multistep pathogenic process. This involves the evolution of CLL cells over time, gaining genomic and epigenetic alterations along their evolutionary path to become increasingly tumourigenic. At each step, CLL is driven by the combined contribution of intrinsic CLL defects with tumour promoting extrinsic microenvironmental factors to finally form fully fledged CLL. However, the specific nature and exact timing of initiating/transforming events in CLL pathogenesis remain elusive.<sup>25-27</sup> Some of these elements contributing to CLL pathogenesis will be discussed in detail within the following sections.

### 1.1 *ii*: Diagnosis

The most characteristic feature of CLL is a peripheral blood B lymphocytosis of  $>5.0 \times 10^9/L$ .<sup>28</sup> Morphologically, CLL lymphocytes are smaller than normal B lymphocytes (Figure 1A). Furthermore, CLL cells appear to be more fragile resulting in the formation of characteristic "smudge" cells on a blood film preparation.<sup>29</sup>

Interestingly, some reports claim that the amount of smudge cells found in routine analysis may in fact be an independent prognostic factor for CLL.<sup>30</sup> Immunophenotypically, CLL monoclonal B cells typically express CD5, CD19, CD20, CD23 and faint levels of surface Ig (sIg).<sup>28,31</sup> The clinical features of CLL such as lymphadenopathy and splenomegaly are generally indicative of later disease stages where these symptoms arise due to the accumulation of lymphocytes in the bone marrow (BM), spleen, lymph nodes (LN) and liver.<sup>9,32,33</sup> Hypogammaglobulinemia is also common in late CLL disease and is associated with an increased susceptibility to infection.<sup>34</sup>



**Figure 1A.** May-Grünwald-Giemsa staining of CLL cells in the bone marrow.

### 1.1 *iii*: Staging

In clinical practice, two staging systems are used at present, Rai and Binet staging, however both of these systems have a limited ability to predict the clinical course at an early stage of the disease. This is due to the fact that these systems are dependent on the appearance of clinical symptoms which often occur late in disease.<sup>35,36</sup> At

diagnosis, the majority of CLL patients are asymptomatic and it is only through routine testing that CLL is identified in these cases. Approximately 30-50% of these patients with an indolent CLL at diagnosis will progress, however predicting which patients will progress is uncertain when using such clinical staging systems.<sup>37</sup> Therefore, the need for new highly sensitive, specific and easily attainable prognostic markers is warranted.

### 1.1 iv: Infectious complications in CLL

Infectious complications play a major role in the mortality of CLL patients. These complications come secondary to the immune defects associated with primary CLL disease and secondary to immunosuppression caused by certain treatments. In fact, about 80% of CLL patients will experience infectious complications at some point during their disease that can severely reduce patient quality of life. Moreover, 50-60% of patients will die due to infection. Some examples of commonly contracted infections are bacteremias and pneumonia caused by *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Legionella pneumophila*, *Staphylococcus aureus*, *Salmonella*, and others. Herpes viruses are also commonly encountered in CLL patients.<sup>34</sup> For instance, Epstein-Barr virus (EBV) can be a severe complication in some CLL cases, moreover it is thought to be involved in the transformation of CLL to aggressive Richters syndrome.<sup>38-40</sup>

### 1.1 v: Treatment

Currently, no outright cure for CLL exists, however various treatment strategies are employed. Asymptomatic CLL patients undergo a watch and wait approach, where

treatment is only administered upon evidence of progressive or symptomatic disease.<sup>28,41</sup> The co-morbidity and fitness status of patients is taken into consideration prior to the selection of a particular treatment regime.<sup>41</sup> Patients with active disease, that is patients with B symptoms, cytopenias and or complications from lymphadenopathy, hepatomegaly, splenomegaly and or a lymphocyte doubling time of <6 months are considered for treatment. Treatment with chlorambucil, an alkylating agent has been considered a front line therapy for CLL for many decades and remains an appropriate treatment today particularly in frail CLL patients.<sup>41</sup> In the recent past, fludarabine, a purine analogue in combination with cyclophosphamide, an alkylating agent was deemed the standard first line therapy for CLL patients.<sup>42,43</sup> Together fludarabine, and cyclophosphamide, work jointly to inhibit DNA synthesis and initiate cell death.<sup>44</sup> In some instances CLL patients may become refractory to such combination chemotherapy treatment.<sup>45</sup> In these cases, the additional use of rituximab a monoclonal anti-CD20 antibody can improve the survival for such patients.<sup>41</sup> Hence fludarabine in combination with cyclophosphamide and rituximab (FCR) is now considered the first line treatment of choice for fit CLL patients with active disease.<sup>41</sup>

Choice of therapy not only depends on disease stage and patient fitness and performance, it depends on the patient's cytogenetic profile. New evidence demonstrates that patients with deletion 11q can overcome the poor prognosis traditionally attributed to this genomic alteration when treated with FCR. More specifically, del(11q) patients treated with FCR have shown high rates of response, survival and relapse free survival.<sup>46</sup> In contrast, patients carrying del(17p) or *p53* mutation have compromised function of their DNA damage response pathway hence,

treatment of such patients with conventional chemotherapy agents such as fludarabine and cyclophosphamide that work through this pathway will not respond.<sup>41</sup> Even with FCR therapy these patients are shown to have a short progression free survival. In these cases, the use of anti-CD52 (Alemtuzumab) antibodies can be effective.<sup>47</sup> High dose methyl-prednisolone in combination with alemtuzumab has also shown to be effective in aggressive CLL patients with p53 defects.<sup>48</sup> Today, the only available potentially curative treatment is the use of allogenic stem cell transplantation; however this is solely considered for younger patients with unfavourable prognostic markers.<sup>41,49</sup>

## **1.2: Normal B cells and B cell development**

In order to comprehend CLL pathogenesis, it is fundamental to understand the function of normal B cells and the processes of normal B cell development, IG gene rearrangement and B cell communication with antigen. B cells are the main cellular components of the adaptive immune response that augment humoral immunity through the production of antibodies. These cells function as key immune surveyors that recognize foreign antigens and generate specific antibody responses that are adapted to aid the elimination of specific invading pathogens. In order to produce antibodies that specifically bind antigens, the B cell must undergo a process of differentiation, a hierarchical pathway of development where each step is characterised by a specific pattern of cell surface markers.<sup>50</sup> B cell maturation occurs in the bone marrow where stimulation from stromal cells and cytokines<sup>51,52</sup> induce the differentiation of pro B cells expressing CD43, CD19 and CD10 from stem cells. Further maturation to the pre B cell stage denotes the earliest cell type synthesising a

detectable IG, the cytoplasmic  $\mu$  heavy chain. This  $\mu$  heavy chain associates with surrogate light chains and  $Ig\alpha$  and  $Ig\beta$  signal transducer molecules to form the pre BCR.<sup>50,53</sup>

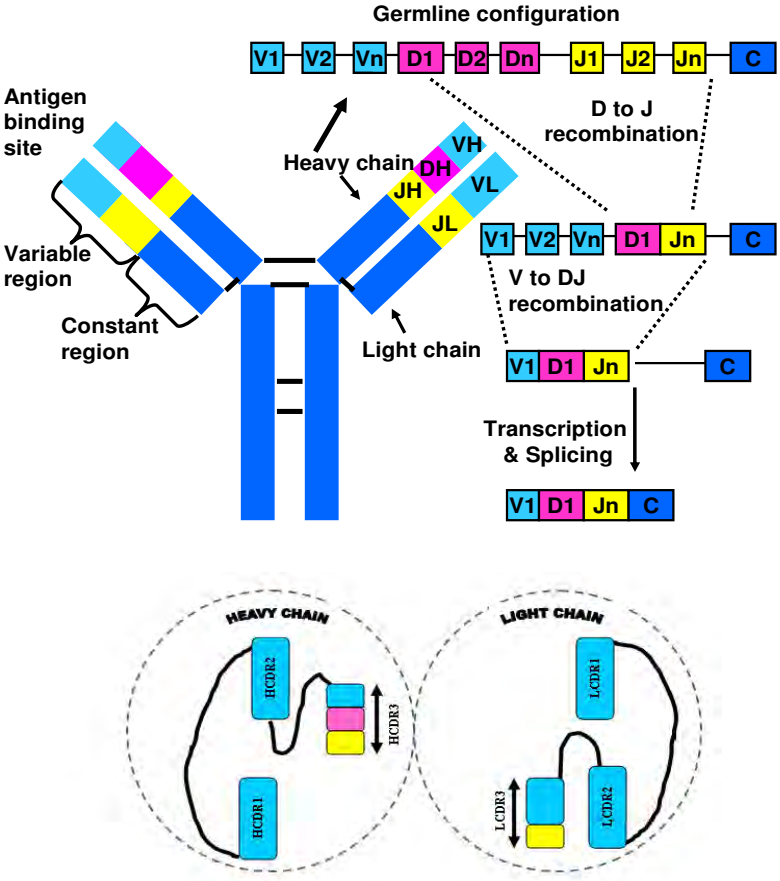
The next stage involves the production of the kappa or lambda light chains. These light chains complex with  $\mu$  heavy chains to produce IgM on the surface of the B cell at the immature B lymphocyte stage. At this stage, these cells do not respond to antigen, in fact on encountering antigen these cells may enter apoptosis or an anergic state. The transition from immature to mature cells is marked by the co-expression of  $\mu$  and  $\delta$  heavy chains producing membrane bound IgM and IgD, and this is accompanied by mature functional competence. Subsequently, fully mature B cells enter the circulation and lymphoid organs where they await antigen activation. Upon antigen encounter, the naïve B cell initiates cell signaling resulting in growth and proliferation of the B cell. This creates an amplified clone of plasma cells that secrete the antigen-specific Ig. Furthermore, following activation, B cells leave a lasting impression in the form of memory cells. These cells persist in circulation to produce a more rapid immune response should they encounter future challenges by the same antigen.<sup>50,53</sup>

## 1.2 *i*: Structure of the immunoglobulin molecule

The IG molecule is a membrane protein complex that is composed of two IG heavy chains and two IG light chains (Figure 1B).<sup>54</sup> Fundamental to the generation of a diverse IG repertoire is the variable (V) region of the IG molecule. This region is generated through a distinct process of gene recombination events at both the IG heavy chain (IGH) and light chain (LC) loci known as VDJ recombination (see



below). The V region is characterized by conserved framework regions (FWR) integrated among highly variable complementarity determining regions (CDRs), CDR1, CDR2 and CDR3. In particular, the CDR3 is the most hypervariable region of the IG molecule that is formed through the joining of three IG gene segments, the V, diversity (D) and joining (J) gene segments (Figure. 1B).<sup>54,55</sup> Thus, the CDR3 serves as the main player in determining IG specificity at the antigen binding groove.<sup>50</sup> As such, IG constant regions (C) do not partake in antigen specificity; instead they constitute the isotype of the IG and determine the effector function of the IG. These regions may bind complement and receptor proteins to augment processes such as lysis, opsonisation and cell de-granulation.<sup>50</sup>



**Figure 1B.** Immunoglobulin structure, VDJ recombination of the IGH locus and a representation of the heavy and light chain CDR3.

## 1.2 *ii*: VDJ recombination

As mentioned, the intricate process of VDJ recombination is a complex mechanism involving a series of sequential rearrangements that play key roles in IG diversity.<sup>50,56</sup> In short, the V region of the heavy chain is created by assembling distinct variable (IGHV), diversity (IGHD) and joining (IGHJ) gene segments. There are 123-129 IGHV gene segments available for recombination, however only 38-46 of these are functional gene segments. Moreover, 23 IGHD and 6 IGHJ functional gene segments as well as 9 constant (IGHC) gene segments are available for rearrangement (Figure 1B).<sup>54,57,58</sup> Rearrangement of the IGH locus commences at the pro B cell stage. On maturing to the pre B cell stage, the rearranged IGH complexes with a  $\mu$  constant region to form IgM on the B cell surface.<sup>50</sup>

In addition, rearrangement of the kappa (IGK) and lambda (IGL) light chain loci initiates during the pre-B cell stage, however only one specificity will be presented (kappa or lambda) on the cell surface of the mature B cell. Unlike rearrangement events at the IGH locus, only the joining of distinct V and J gene segments occurs at the IGK and IGL loci. In total, 36-40 functional IGKV, 5 IGKJ and one IGKC genes are available for recombination. At the IGL locus, 29-33 functional IGLV genes, 4-5 IGLJ and 4-5 IGLC genes are available for rearrangement. Due to the lack of D genes at the light chain loci, the level of diversity is less than that observed at the IGH locus.<sup>50,54,58</sup>

In short, the VDJ recombination mechanism is initiated by double stranded DNA breaks introduced by two lymphocyte specific recombination activating gene (RAG) enzymes, RAG 1 and RAG2.<sup>59-61</sup> These enzymes work by forming a complex that

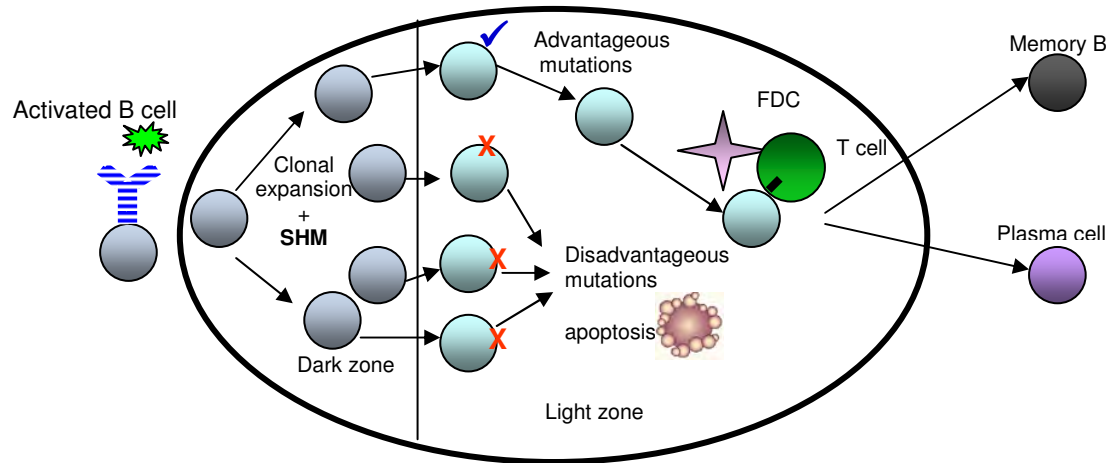
specifically recognises recombination signal sequences (RSS) that flank the coding V, D and J segments. These RSSs are made up of 3 distinct elements; a conserved heptamer and a nonamer separated by a spacer element of 12bp or 23bp.<sup>57,59,60</sup> The length of this spacer element plays a key role in driving recombination specificity.<sup>62</sup> This is explained by the 12/23 spacer rule that states that only pairs of dissimilar spacer RSSs are efficiently recombined, so that spacers of 12 nucleotides will only be recombined with spacers containing 23 nucleotides.<sup>56,63</sup> In accordance with the 12/23 spacer recombination rule, a D gene segment will first combine with a J gene segment forming a DJ complex. Following DJ joining a V gene will join to the DJ segment forming a VDJ complex.<sup>50</sup>

After the above recombination events have taken place, DNA repair mechanisms are employed to mend DNA breaks. DNA-dependent protein kinase complexes (DNA-PK) are recruited to the broken DNA ends.<sup>62</sup> Here, the DNA-PK recruits several other proteins and DNA polymerases to the free ends. On aligning of the two free DNA ends, the DNA-PK complex recruits the enzyme terminal deoxynucleotidyl transferase (TdT).<sup>64</sup> TdT works to add random non-templated (n) nucleotides to the free DNA ends thus generating further IG diversity. Alternatively, deletion of nucleotides through exonuclease activity can provide a further layer of diversity at these junctional regions. Finally, DNA polymerases may insert additional nucleotides as needed to make the two ends compatible for joining. A ligase IV enzyme finally links DNA strands on opposite ends of the break to each other, completing the joining process.<sup>59,64</sup>

### 1.2 *iii* : The germinal center reaction and somatic hypermutation

Secondary lymphoid organs such as the lymph node and spleen provide unique micro-environmental niches to allow naïve B cells the chance to encounter antigen, proliferate and produce specific antibodies in a process termed the germinal centre (GC) reaction (Figure 1C). Within these lymphoid organs are highly specialized GC follicles that provide niches encompassing a network of antigen presenting cells, co-stimulator cells and cytokines. The GC reaction is initiated when the naïve B cell encounters antigen in the extra follicular spaces. Upon activation of the cell it moves to T helper cell zones. On recognizing the antigen, T helper cells produce CD40L which binds to the B cell CD40 receptor to promote B cell activation.<sup>65</sup> This induces the B cell to transform into a highly proliferative centroblast within the dark zone of the GC follicle.<sup>65</sup> In an effort to generate antibodies with a higher affinity to the initiating antigen, the process of somatic hypermutation (SHM) refines the sequences of the BCRs by introducing single nucleotide mutations randomly into the IG genes.<sup>66-68</sup> In fact, SHM is a highly specialized process that occurs at a rate  $10^6$  times higher than spontaneous mutation.<sup>66,69</sup> At this stage, B cells acquiring disadvantageous SHM patterns will die by apoptosis.<sup>70</sup> However, those that acquire favourable antibody mutations move to the light zone of the follicle and are termed noncycling centrocytes.<sup>71</sup> These centrocytes bind their cognate antigen with the help of follicular dendritic cells and T cells to differentiate into plasma or memory B cells.<sup>65,72</sup> Although GCs are considered the principal sites capable of sustaining SHM, they are not thought to be the sole sites of such activity. It is thought that SHM may occur outside of the GC and independently of T cell help.<sup>73-76</sup> For instance, it has

been suggested that B cells contained within the extra follicular marginal zone (MZ) of the spleen can also undergo SHM.<sup>73-76</sup>



**Figure 1C.** The germinal centre reaction

SHM represents a second mode of IG diversification after VDJ recombination. The process of SHM allows a further increase in IG diversity and the production of IGs with a higher specificity,<sup>77</sup> however, knowledge of the exact mechanisms governing SHM are incomplete. SHM is known to involve the activation of induced cytidine deaminase (AID), an enzyme that works to deaminate cytosine bases to uracil in single stranded DNA creating a U:G mismatch.<sup>77</sup> The mismatch is recognised by the cell's DNA repair enzymes that excise the uracil base, however, error prone polymerases that work to fill in this gap result in the formation of mutations at this site.<sup>77</sup> These mutations are mainly substitution mutations; however insertions and deletions also occur. Mutations are more commonly encountered in the CDR sequences where certain hotspots within these regions appear to be targeted more often to other regions.<sup>78-82</sup>

## 1.2 *iv*: Immunoglobulin diversity

As mentioned, the vast variety of unique IG molecules making up the human Ig repertoire is partially given by the intricate processes of VDJ recombination, SHM and non-templated (n) and palindromic (p) nucleotide addition/excision at junctional borders. In fact, due to these processes, the likelihood of finding two identical BCRs is negligible. More specifically, when considering all permutations of V, D and J genes at the IGH locus, the likelihood of the same VDJ recombination event presenting in normal healthy persons is 1 in 6348. Taking into account the chances that individuals carrying the same heavy chain carry the same kappa or lambda rearrangement the probability further diminishes to approximately 1 in 2.3 million. Finally, SHM events and the enzymatic addition and or removal of n and p nucleotides at junctional regions further reduces the probability of two healthy persons expressing identical BCRs to  $2.3 \times 10^{12}$ .<sup>58</sup>

## 1.3: The CLL cell and CLL pathogenesis

On a blood film, CLL cells appear as small mature B lymphocytes containing a dense nuclear structure, partially aggregated chromatin and a narrow rim of cytoplasm.<sup>28</sup> However, CLL cells have intrinsic BCR signalling defects that are suggested to be characterized in some cases by truncated CD79b Ig domains,<sup>83</sup> defective assembly,<sup>84</sup> glycosylation and folding of  $\mu$  and CD79a chains.<sup>85</sup> In other cases, somatic mutations in the cytoplasmic domain of CD79b have also been described.<sup>85,86</sup> These impairments may result in a low surface Ig expression and an inability to engage in effective signal transduction, jeopardising the ability of the CLL cell to undergo

apoptosis. Moreover, the high expression of the anti-apoptotic BCL2 family described in CLL<sup>87</sup> and the up-regulation of the cell cycle arrest protein p27<sup>KIP</sup> may aid the accumulation of the CLL clone.<sup>88</sup> Further propagated by continuous CLL cell proliferation in lymph nodes,<sup>89</sup> CLL cells thrive and out-compete normal B cells.

Once released to the peripheral blood, it is thought that CLL cells continually recirculate to the secondary lymphoid tissues, to receive growth and survival promotional signals from accessory cells, cytokines and antigens contained within the microenvironment. One theory proposes genetic and epigenetic alterations to act as early initiating/driving events in CLL transformation.<sup>26,90-93</sup> In this model, microenvironmental extrinsic signals merely provide a support mechanism to the CLL cells, whereby antigen(s) is proposed to administer a chronic stimulation to BCR receptive CLL cells, propelling the clone along its tumorigenic path. In an opposing theory, antigen is suggested to act as the key culprit in CLL initiation, whereby long lived CLL cells acquire pathogenic genomic and epigenetic alterations as secondary hits over time. Despite the lack of knowledge about the order and sequence of such tumorigenic events, the proposed interplay between genomics, epigenomics and microenvironment suggests a complex multistep tumorigenesis model in CLL. However, key to this model is assessing the CLL cell of origin which to date remains elusive.<sup>25,27,94</sup>

### 1.3 *i*: CLL cell of origin

Over the past two decades the perception of the candidate cell(s) of origin for CLL has changed. Initially it was believed that a naive CD5<sup>+</sup> B cell was the originator of CLL. However, the discovery of divergent CLL subgroups (see prognostic markers

section), IGHV mutated CLL, characterised by SHM of their IGHV genes and IGHV unmutated CLL, characterised by the lack of SHM at their IGHV genes altered this opinion. This finding lead to the hypothesis that unmutated CLL cells were derived from CD5<sup>+</sup> naïve pre GC B cells, whereas mutated CLL derived from antigen experienced post GC memory B cells.<sup>2,3</sup> In the recent past, it has been suspected that both mutated and unmutated CLL cases derive from antigen experienced B cells, as CLL cells were shown to have a similar gene expression profile to that of antigen experienced memory B cells.<sup>95</sup> The candidate single cell origin of CLL is considered to be the marginal zone (MZ) B cell. The MZ B cell shares many CLL cell features such as, an active membrane phenotype, and BCRs encoded by either mutated or unmutated genes. Moreover, MZ B cells are also shown to demonstrate both poly and autoreactivity. Finally like CLL cells, these MZ B cells can respond to T cell dependent and independent antigens.<sup>25,75,94</sup>

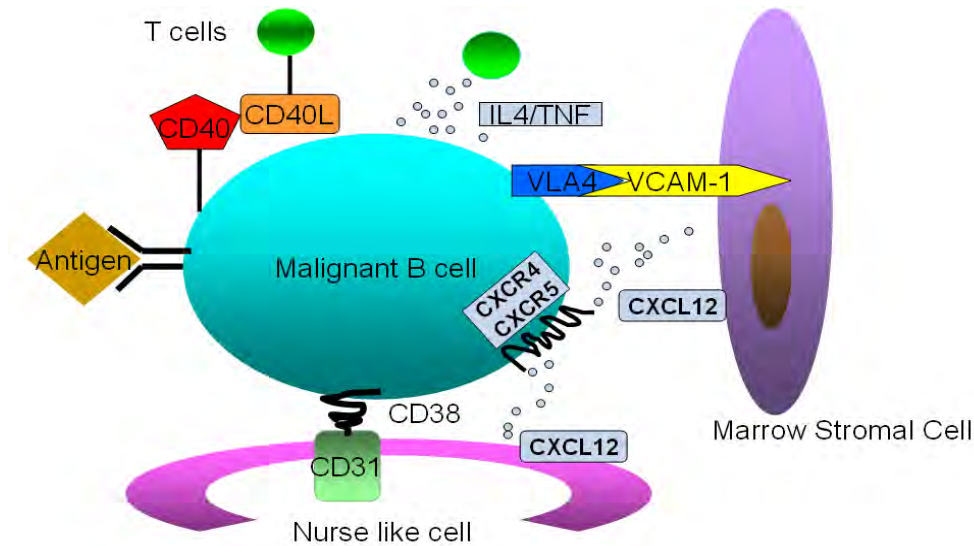
Nowadays however, the notion of a single cell of origin is being challenged.<sup>27</sup> Given the fact that IGHV mutated and unmutated CLL have previously been shown to have rather similar gene expression profiles, it is now suspected that perhaps the activated state of CLL cells may mask larger gene expression differences between these subgroups.<sup>27</sup> If this is the case, then conceivably it is more likely that mutated and unmutated CLL derive from two independent cell origins. Since CLL leukemogenesis is considered to follow a long stepwise process perhaps spanning several steps in B cell development, the prospect arises that transformation can occur anytime during B cell maturation once IG heavy chain is expressed.<sup>27</sup> Theoretically, this opens up the possibility that CLL could be derived from multiple cell precursors.<sup>27</sup> Today, it has



been suggested that the propensity to generate CLL clonal B cells is already decided at the hematopoietic stem cell stage.<sup>96,97</sup>

### 1.3 *ii*: The microenvironment in CLL

*In vitro* CLL cell culturing studies have long implicated the microenvironment to play a crucial role in CLL survival. CLL cells in culture rapidly undergo apoptosis in the absence of vital survival factors which are now known to occupy niches within complex in-vivo micro-environmental networks.<sup>98-102</sup> These networks involve active molecule cross-talk between CLL cells, accessory cells, activated T cells and a medley of soluble pro-survival chemokines and cytokines (Figure 1D). These molecules communicate through ligand-receptor and adhesion molecule interactions to activate autocrine and exocrine signaling in CLL cells.<sup>103-106</sup> Overall, microenvironmental interactions of CLL cells seem to mainly mirror those of normal lymphocytes.<sup>103,107</sup> However, certain distinct microenvironmental interactions controlling CLL migration do exist that are not seen in normal B cells.<sup>108</sup> In addition to providing support to the CLL cell, the microenvironment is said to provide physical protection from cytotoxic drug damage, thus contributing to treatment resistance.<sup>109</sup> CLL cell interactions within the bone marrow, peripheral blood and lymph node microenvironment are discussed below.



**Figure 1D.** The CLL cell microenvironment (Adapted from Burger *et al* 2009)<sup>110</sup>

### Bone marrow

In the bone marrow, CLL cells come into contact with bone marrow stromal cells (BMSCs).<sup>100,101</sup> *In vitro*, CLL cells co-cultured with BMSCs have been shown to live longer.<sup>100,101</sup> This longevity is partially attributed to the interactions of BMSC CD54 and CD106 integrins with CLL cell CD49d/CD11a or CD11b/CD18 integrins. This interaction is suggested to bring about protection from spontaneous apoptosis by increasing anti-apoptotic signals such as *BCL2*.<sup>100,111,112</sup> Accompanying survival promoting events include the release of stromal cell chemokines such as, stromal cell derived factor 1 (SDF-1) otherwise known as CXCL12. SDF-1 production is constitutively produced by BM stromal cells in CLL patients.<sup>113</sup> In response to SDF-1 release, CLL cells migrate towards these growth supporting accessory cells via their highly expressed CXCLR4 receptors (Figure 1D).<sup>113,114</sup> Once recruited, the stromal

cells adhere to CLL cells via VCAM-1 to VLA4 integrin interaction to subsequently activate anti-apoptosis pathways.<sup>100</sup>

### **Peripheral Blood**

Unlike mononuclear cells in the peripheral blood of normal individuals, some cells in CLL patients regularly morph into large nurse like cells (NLCs) that adhere to leukemic cells *in vitro*.<sup>99</sup> Similar to BMSCs, in-vitro co-culturing experiments have indicated the importance of NLC derived SDF-1 in the prevention of apoptosis through its interaction with CXCLR4 on CLL cells (Figure 1D). It is thought that perhaps NLCs help counteract apoptosis through the activation of MAP kinase ERK1/2 pathways that increase the expression of anti-apoptotic proteins like MCL-1.<sup>115-117</sup>

### **The lymph node**

The lymph node represents a specialized micro-environmental niche for CLL cells since it is the proposed main site of antigen encounter, BCR signaling and proliferation.<sup>105,118,119</sup> In the past, CLL was solely considered to be an accumulative disease driven by defects in apoptosis. Nowadays it is seen as a disease afflicted by a concomitant increase in proliferation of the leukemic clone. Up to 1% of the CLL clone is now known to proliferate daily, within specialized pseudo-follicle structures known as proliferation centers (PCs).<sup>89,119</sup> Large para-immunoblasts and pro-lymphocytes within these structures highly express the proliferation marker Ki-67 compared to smaller lymphocytes in the vicinity of the follicle.<sup>119</sup> These large cells are also known to have constitutive activation of the NFκB pathway, thus promoting cell proliferation through tight regulation of *NFκB* target genes such as *BCL-2* and

*MCL-1*.<sup>119-121</sup> In support of these findings, CLL cells isolated from LNs have been recently shown to have an increased expression of active BCR signaling and proliferation genes relative to cells from the blood or bone marrow.<sup>118</sup>

In PCs, subpopulations of T cells, stromal cells, macrophages and endothelial cells provide a cocktail of signaling mediators. Normally, the B cell receives survival signals via T cell CD40-CD40L stimulation. In CLL, increased numbers of T cells concentrate round PCs in response to B cell CCL22 release.<sup>119,122,123</sup> Together, secretion of T cell derived cytokines like IL-4 and accessory cell chemokines such as SDF-1 support the expansion of CLL clones by up regulating anti-apoptotic SURVIVIN and BCL2 regulators.<sup>99,123</sup>

LN derived CLL cells are known to harbor higher levels of CD38, compared to PB derived CLL cells.<sup>124,125</sup> Additionally, in-vitro studies show an increase in CD38 expression upon CLL activation<sup>126</sup> thus proposing CD38 positivity as a characteristic of the CLL cell-LN microenvironment interaction. CD38+ cells have been reported to interact with CD31 on stromal cells to promote CLL proliferation.<sup>127,128</sup> In fact, CD38+ cells are said to proliferate ~2X as fast as CD38- cells *in vivo*.<sup>129</sup> It is thought that CD38+ cells gain an enhanced migratory ability to home to favourable microenvironments due to an elevated responsiveness to the CD31 stromal ligand. Alternatively, given the strong association between CD38 positivity and increased CD31+ vascularisation in lymph nodes, it is further proposed that the microenvironment itself may induce CD38 expression.<sup>130,131</sup> Given its function as an accessory BCR signaling molecule, it is suggested that CD38+ cells can strongly transduce signals from their BCR, unlike their negative counterparts.<sup>132,133</sup> Altogether

these circumstances have culminated to suggest that CD38+ cells constitute the fraction of cells primed to proliferate in CLL.

### 1.3 *iii*: Evidence of antigen(s) involvement in CLL pathogenesis

Currently, the exact role of antigen(s) in CLL leukemogenesis is unknown. Deciphering whether antigen(s) play vital a part in CLL initiation by providing a proliferative stress or whether they contribute to CLL progression through chronic stimulation of the CLL clone is an active area of current research. Despite these uncertainties pertaining to time of antigen engagement in CLL pathogenesis, a plethora of circumstantial evidence supports the involvement antigen at some stage of CLL development. This evidence stems from the finding that compared to the normal repertoire; CLL shows a biased use of certain IGHV genes, namely IGHV1-69, IGHV3-7, IGHV3-23, IGHV3-21 and IGHV4-34.<sup>3,134,135</sup> Additionally, IG sequence analysis has described the presence of multiple CLL subsets with closely homologous or stereotyped BCRs.<sup>4-6,134,136-138</sup> More specifically, these stereotyped subsets are characterised by an amino acid identity of  $\geq 60\%$  in the heavy chain CDR3 and restricted usage of similar IGHV-D-J genes and light-chain genes. In fact, later reports have evidenced that ~30% of CLL patients belong to subsets (>100 defined today) carrying 'stereotyped' CDR3 sequences on their heavy and light chains.<sup>5,6</sup> Considering the likelihood of finding two B cell clones with identical BCRs is almost negligible, these findings are highly suggestive of antigen(s) involvement in leukemogenesis.<sup>4-6,136,137</sup>

### 1.3 iv: The molecular nature of antigen(s) in CLL

The specific nature of the antigen(s) involved in CLL is largely unknown. It is suspected that foreign antigens,<sup>139</sup> auto-antigens<sup>94</sup> and super-antigens<sup>6</sup> or combinations thereof play a role in CLL leukemogenesis. The reactivity of CLL cells to antigen stimulation has been shown to differ depending on the mutational status of the BCR. For instance, unmutated BCRs have been shown to be polyreactive against such molecules as bacterial lipo-polysaccharides (LPS). Moreover, unmutated BCRs have demonstrated autoreactivity against insulin and DNA molecules.<sup>139,140</sup> In contrast, mutated BCRs appear to lose their polyreactivity upon acquiring mutations through SHM.<sup>139</sup>

CLL BCRs including several stereotyped cases, have demonstrated homology to anti-DNA, anti-rheumatoid factor, and anti-cardiolipin auto-antibodies.<sup>5</sup> Moreover, BCR-stereotyped CLL cells have been shown to bind intracellular autoantigen such as cytoskeletal proteins vimentin and oxidised low density lipoproteins.<sup>141,142</sup> More recently, auto-antigen in the form of antigenic motifs on apoptotic blebs from dead cells have been implicated in CLL leukemogenesis. It is thought that the BCR can recognise these motifs leading to self stimulation and expansion of the CLL clone.<sup>141,142</sup> For example, unmutated CLL monoclonal antibodies have been shown to specifically bind non-muscle myosin heavy chain IIA (MYHIIA9) exposed apoptotic cells (MEACS). This event has been particularly evident in ‘subset #6’ patients expressing IGHV1-69/IGHD3-16/J3 genes.<sup>143</sup> In light of the latter finding, it has been suggested that CLL clones may arise from dual functional B cells that maintain

their ability as scavengers of apoptotic residues, whilst sustaining their ability to bind conserved bacterial cell motifs.<sup>142,144</sup>

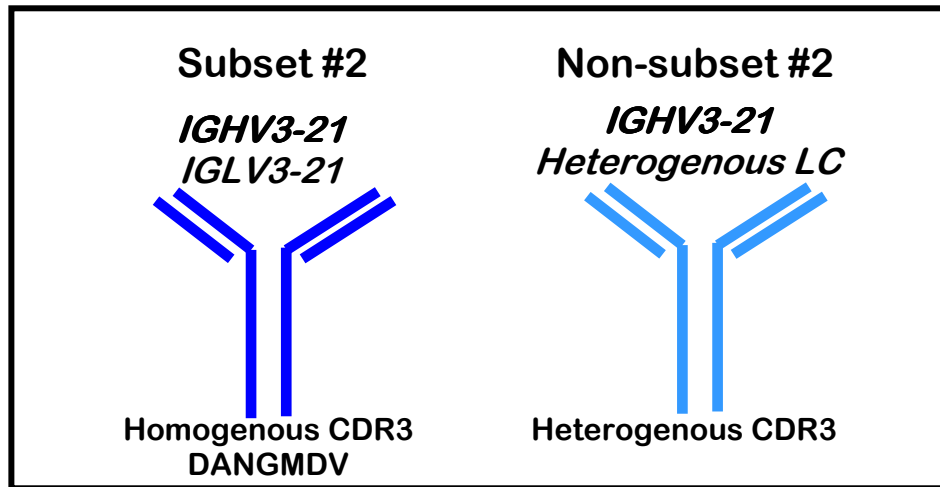
Interestingly, self antigens have been noted to bind certain IGHV3 and IGHV4 BCRs outside of the CDR3 region, instead binding to IG FWRs. Similarly, bacterial superantigens have been shown to bind FWR1, FWR3 and CDR2 regions of some IGHV3 BCRs.<sup>139,145,146</sup> Details of these proposed superantigens are discussed below in relation to the IGHV gene family in which they are suspected to have a role.

## **1.4: IGHV3-21 and IGHV4-34 CLL subgroups: IG sequences, antigens and clinical implications**

### **1.4 *i*: IGHV3-21 utilising patients**

IGHV3-21 patients have been shown to demonstrate a poor clinical outcome and aggressive disease regardless of their IGHV mutation status.<sup>4,135,137,147</sup> In recent years, studies of IGHV3-21 CLL have reported an abundance of evidence linking antigen involvement to IGHV3-21 pathogenesis.<sup>4,6,135,137,148</sup> For instance, approximately 50% of CLL patients using the IGHV3-21 gene demonstrate stereotyped BCRs and these are denoted as subset #2 patients. More specifically, subset #2 IGHV3-21 patients carry stereotyped BCRs consisting of a conserved 9 amino acid long HCDR3 sequence (Figure 1E).<sup>137,148-150</sup> In a large proportion of cases, the amino acid sequence of these stereotyped HCDR3 regions are identical (ARDANGMDV). However, in some cases sequences may differ merely by 1-3 amino acids at different positions. Furthermore, in the majority of subset #2 cases, it is not possible to assign any D gene

usage.<sup>137,147,148,150</sup> Interestingly, IGHV3-21 stereotyped patients show restricted expression of the IGLV3-21 gene.<sup>137,147,148</sup> Altogether, these findings indicate the recognition of a common antigen in IGHV3-21 subset #2 cases.<sup>4-6,137,147,148</sup>



**Figure 1E.** A representation characterising IGHV3-21 stereotyped and non-stereotyped BCRs

Non-stereotyped IGHV3-21 (non-subset #2) patients have a heterogeneous CDR3 and light chain gene usage (Figure 1E). Both stereotyped and non-stereotyped BCRs have been reported to share an equally poor overall survival similar to that displayed by other IGHV unmutated cases (see prognostic markers section). Nevertheless, IGHV3-21 stereotypy has also been suggested to influence clinical behaviour. For instance, patients carrying stereotyped CDR3s are proposed to have a more progressive disease compared to non-stereotyped patients.<sup>5,149</sup> Supporting this proposal is the finding that IGHV3-21 patients with homologous HCDR3s more frequently express high levels of the poor prognostic markers CD38 and ZAP70.<sup>5,147,150</sup>



IGHV3-21 sequences are predominantly border-line mutated. This finding can be partially explained by the under-targeting of SHM across all regions of IGHV3-21 sequences compared to other IGHV3 subgroup genes.<sup>6</sup> Interestingly, IGHV3-21 subset #2 patients have been shown to have lower targeting of mutations across all IG regions except the HCDR2 compared to non-subset 2 IGHV3-21 sequences. That said, several recurrent amino acid changes have been observed among subset #2 cases. Remarkably, a serine deletion at HCDR2 codon 59 is commonly detected in stereotyped IGHV3-21 CLL sequences.<sup>6</sup> This deletion appears to be both a CLL and subset biased mutation. The finding of specific recurrent mutations in IGHV3-21 CLL further strengthens the role of antigen selection in CLL.

Interestingly, IGHV3-21 patients and in particular stereotyped patients, showed a strong tendency to retain germline configuration in the binding motif for Staphylococcal protein A, a suspected superantigen in CLL.<sup>6</sup> However, the biologic and clinical implications of this finding (if any) remain unknown. Additionally, mutated IGHV3-21 subset #2 cases have been shown to bind the intracellular auto-antigen cofilin-1, an actin binding protein found localise with molecular complexes on the cell surface at apoptosis.<sup>142</sup>

Finally, a geographic bias has been attributed to this poor prognostic IGHV3-21 gene. For instance, Scandinavia has demonstrated the highest IGHV3-21 gene frequency (10-13%)<sup>4,135,137,150</sup> compared to other Southern European countries and North America (~3-4%).<sup>5,151,152</sup> These findings suggest that aside from geographical bias, differences in ethnicity and/or environment may account for the varying frequency of the IGHV3-21 gene seen globally. Nevertheless, it is also possible that biases in

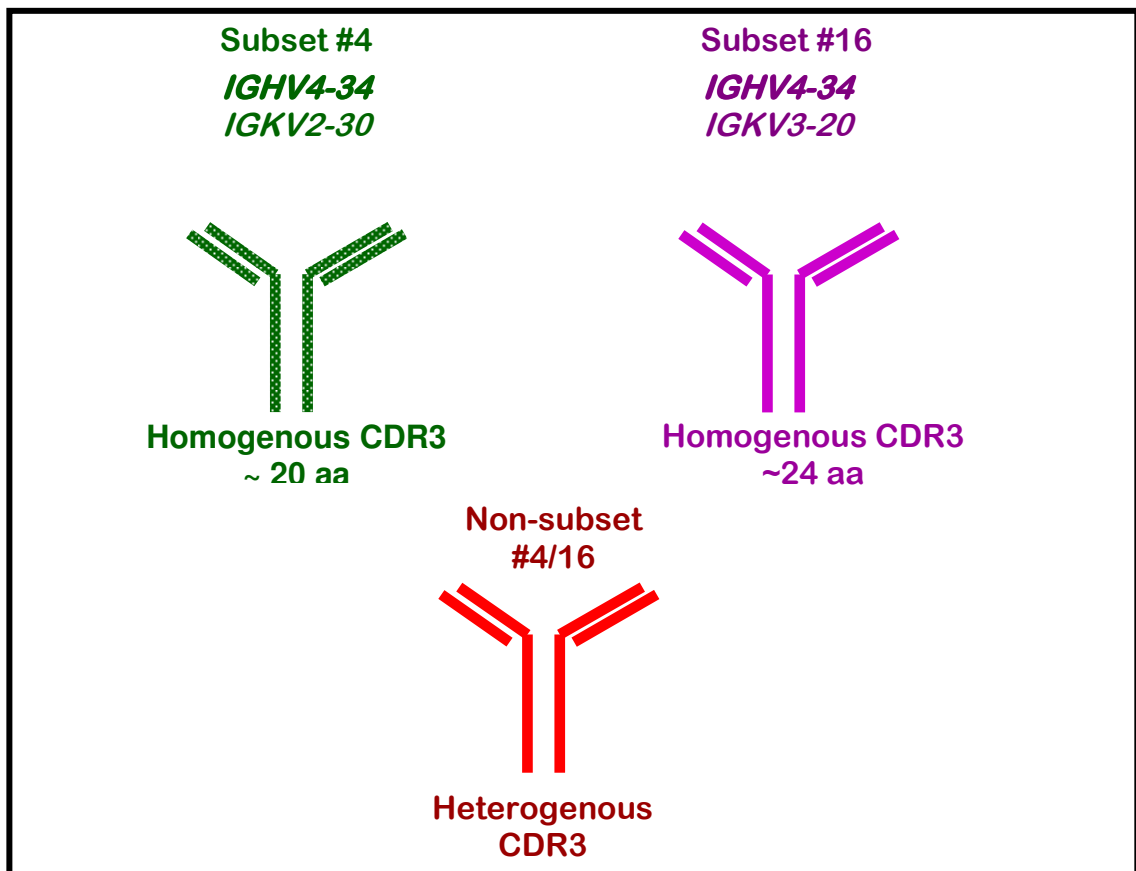
sample selection can contribute to the varying IGHV3-21 gene frequencies detected around the globe.<sup>4,5,135,137,147,149-155</sup>

#### 1.4 *ii*: IGHV4-34 utilising patients

IGHV4-34 gene usage is seen in approximately 8-10% of CLL patients. Two CLL subsets carrying stereotyped IGHV4-34 BCRs namely subset #4 and subset #16 have been described in CLL. Subset #4 is the most common stereotyped subset occurring at an overall frequency of ~1%. This subset is characterised by a 20 amino acid long CDR3 sequence and restricted usage of the IGKV2-30 light chain (Figure 1F). Furthermore, these patients almost exclusively carry IGHV mutated genes. Remarkably, these restricted biological features are associated with an indolent disease course compared to non-stereotyped IGHV4-34 CLL patients displaying non-restricted IG features (Figure 1F).<sup>5,6</sup> The indolent disease course noted for these patients, may in part be explained by the fact that these cases are associated with low expression of the poor prognostic CD38 marker. Interestingly, subset #4 patients have a low median age at diagnosis.<sup>5</sup> Moreover, these patients demonstrate a potential association with persistent infection by common herpes viruses such as EBV and CMV.<sup>156</sup> Recently, subset #4 cases have been shown to display extensive intraclonal diversification suggesting a role of ongoing active antigen stimulation in these cases.<sup>157</sup>

The second IGHV4-34 stereotyped subset, subset #16 is present at an overall frequency of 0.3%. This subset is characterised by mutated IGHV genes, a restricted IGKV3-20 light chain usage and a 24 amino acid long homologous CDR3 sequence

(Figure 1F). However, as this subset is not commonly encountered little is known about the clinical outcome for these patients.<sup>5</sup>



**Figure 1F.** A representation characterising IGHV4-34 stereotyped and non-stereotyped BCRs

In the germline state IGHV4-34 BCRs appear to be inherently autoreactive and may bind to auto-antigens. However, since IG sequencing studies have revealed IGHV4-34 BCRs to be mainly mutated,<sup>6</sup> it has been suggested that through SHM, IGHV4-34 BCRs can alleviate their auto-reactivity.<sup>158</sup> More specifically, stereotyped IGHV4-34 cases demonstrate stereotyped amino acid sequence changes at their IGHV genes and these changes are largely subset biased. Stereotyped IGHV4-34 subsets #4 and #16 have distinctive SHM distribution patterns in their HCDRs and HFRs compared to

their non-stereotyped counterparts. For instance, subsets #4 and #16 carrying basic lysine residues within their HCDR3 regions also incur stereotyped mutations in the HCDR1 region resulting in the introduction of negatively charged residues. The latter mutations are said to eliminate the potential DNA binding properties given by the positively charged HCDR3 region. Furthermore, it is noted that in stereotyped IGHV4-34 subsets the HFR1 motif conferring anti I/i reactivity is rather conserved compared to non-subset cases. Hence, these subsets hold the potential to bind the I/i blood group antigen or the B cell isoform of CD45 containing the N-acetyllactosamine antigenic determinant. Hence, through SHM, IGHV4-34 utilising B cells lose their auto-reactivity transitioning to a more safe state so that they can persist within the functional IG repertoire.<sup>6,94</sup>

In summary, the proposed stepwise model of CLL leukemogenesis suggests that the survival of the preleukemic clone depends on the interaction of CLL cells with the pro-survival microenvironment. Antigen(s) residing in this microenvironment are thought to provide transient or chronic stimulation to the CLL clone. Under this model, antigen stimulation, aberrant genetic alterations and epimutations work together in a stepwise fashion to create a thriving CLL clone. Transforming and progressive events may occur at currently unknown discrete B cell developmental stages or randomly depending on the type, efficiency and potency of promotional microenvironmental signals. A summary of the genetic and epigenetic alterations seen in CLL will be described within the following sections.

## 1.5: Prognostic markers in CLL

### 1.5 *i*: IGHV gene mutational status

In 1999, two independent studies demonstrated that the IGHV mutational status could categorise CLL into divergent clinical subgroups. Remarkably, the IGHV mutation status allowed the subdivision of CLL into two disease entities where patients with unmutated IGHV genes were shown to experience an inferior prognosis while, patients with mutated IGHV genes demonstrated a superior prognosis.<sup>2,3</sup> More specifically, patients with IGHV unmutated genes have a more aggressive condition, including evidence of advanced progressive disease, atypical peripheral blood cell morphology, adverse cytogenetic features, clonal evolution, and resistance to therapy, than those with mutated IGHV genes.<sup>2,3</sup> As mentioned, an exception to this rule is the presence of the IGHV3-21 gene (see the IGHV3-21 utilising patients section).

Today it is widely accepted that a 98% identity cut-off is best at making the clinically relevant distinction between mutated and unmutated CLL. CLL patients who carry IGHV genes with  $\geq 98\%$  identity to germline are considered unmutated whereas patients with  $< 98\%$  identity to germline are denoted as mutated.<sup>2,3,159,160</sup> Nevertheless, applying this arbitrary mathematical cut-off to a biological event must be interpreted with caution, particularly when dealing with borderline cases. Despite the advantages of this robust prognostic marker, the methods used to obtain IGHV gene sequences can be laborious, thus the search for more accessible prognostic markers conducive to routine procedures are well sought after.

### 1.5 *ii*: Protein based markers: CD38 and ZAP70 expression

CD38 expression was the first prognostic factor found to correlate with IGHV mutational status, where high expression was shown to be associated with unmutated genes and a poor prognosis.<sup>2,124,159,161</sup> However, it has been evidenced that this relationship does not always hold true.<sup>159,162,163</sup> Another issue preventing the use of CD38 expression as a substitute marker for mutational status is the lack of a definitive threshold delineating what percentage of the CLL clone should possess elevated CD38 expression in order to assign CD38 positivity.<sup>161,164</sup> Moreover, CD38 has been reported to vary over time.<sup>124,162</sup> Nevertheless, CD38 expression is considered an independent prognostic marker, albeit not the strongest.<sup>165</sup>

Similarly, ZAP70 expression shows a good correlation to IGHV mutational status. In general, the majority of mutated cases are shown to be ZAP70 negative, whereas unmutated cases are mainly shown to be positive.<sup>166</sup> Nevertheless, Rassenti *et al* demonstrate that approximately 23% of patients show discrepant results that do not correlate with mutational status.<sup>167</sup> Discordant cases may frequently present with other poor prognostic features such as del(17p), del(11q) and/or IGHV3-21 expression. For instance, 17p and 11q deleted patients have been found to have a low expression whereas IGHV3-21 patients have been shown to have a high expression.<sup>168</sup> Interestingly, Rassenti *et al* have further shown ZAP70 expression to be a better predictor of time to progression compared to IGHV mutational status. Another advantage of ZAP70 measurement is the fact that ZAP-70 levels also appear to remain stable over time.<sup>167</sup>

ZAP70 measurement holds many advantages in that it can be easily measured by readily available methods such as flow cytometry at the protein level and RQ-PCR at the gene level<sup>166,169</sup> ZAP70 is expressed by other immune cells such as natural killer (NK) and T cells which may influence ZAP70 expression analysis. Therefore, it is required that CLL tumour B cells be purified away from such cells prior to ZAP70 expression analysis. Hence, the inconvenience of having to perform cell purification prior to analysis makes measurement of ZAP70 expression non conducive for use in clinical routine.

### 1.5 *iii*: RNA based markers: *LPL* and *CLU1* expression

Today, Lipoprotein lipase (*LPL*) gene expression serves as an independent prognostic marker in CLL.<sup>170</sup> Over a decade ago, gene expression studies identified *LPL* to be differentially expressed in IGHV mutated and unmutated CLL.<sup>95,171</sup> Although not highly expressed in CLL, *LPL* expression is higher in IGHV unmutated patients relative to IGHV mutated CLL patients.<sup>170</sup> Expression further correlates with other negative prognosticators such as high CD38, ZAP-70 expression and poor genetic alterations.<sup>172</sup> Predictive of a shorter treatment free survival and in most cases a shorter overall survival, *LPL* expression in some instances has been deemed as a superior independent prognosticator compared to IGHV mutational status.<sup>170,173</sup> Moreover, its prognostic value is maintained independently of clinical stage.<sup>172-175</sup> Very low *LPL* expression in normal peripheral blood mononuclear cells (PBMCs)<sup>174,176</sup> permits the use of non-sorted PBMCs from CLL patients for analysis. Altogether, quantification of *LPL* expression by RQ-PCR may offer a cheaper, less labour intensive method to identify prognostic subgroups of CLL patients.

Differential display screening first identified CLL upregulated gene 1 (*CLLUI*) to be differentially expressed in CLL. *CLLUI* is highly expressed in IGHV unmutated patients relative to mutated patients<sup>177,178</sup> and is associated with other poor risk characteristics namely advanced clinical stage, poor cytogenetics and high CD38 and ZAP70 expression.<sup>173,178,179</sup> Like *LPL*, high *CLLUI* expression correlates to a shorter overall survival and treatment free survival.<sup>173,178,180</sup> Recently, the value of *CLLUI* expression in IGHV mutated patients has revealed important prognostic information, indicating those IGHV mutated patients with high *CLLUI* expression to have a worse prognosis relative to mutated patients with low expression.<sup>179</sup> Additionally, *CLLUI* expression appears to be exclusive to CLL cells and is stable over time.<sup>180</sup> Hence, determination of *CLLUI* expression levels can be used as a reliable indicator of tumour burden and is suggested to complement currently used techniques for minimal residual disease (MRD) monitoring in CLL patients post therapy.<sup>181</sup>

### 1.5 iv: Genomic alterations

CLL is not defined by a single recurrent genomic alteration but rather by genomic instability that frequently culminates into chromosomal aberrations which are present in up to 80% of patients. Today, fluorescent *in-situ* hybridization (FISH) techniques are implemented to detect the most common chromosomal alterations namely del(13q), del(11q), del(17p) and trisomy 12, which are of important prognostic significance in CLL (Table 1A).<sup>7</sup>



### **Deletion 13q**

Deletion at chromosome 13q14 is the most frequently encountered alteration affecting up to 55-60% of CLL patients (Table 1A).<sup>7</sup> Deletion of 13q most often occurs as the sole alteration and is associated with a favourable prognosis and indolent disease course.<sup>7</sup> However, the prognostic significance of mono-allelic versus bi-allelic deletions of 13q still remains controversial. Recently, a study revealed that a similar time to initial treatment and overall survival was attained in both mono- and bi-allelic del(13q) CLL patient groups.<sup>182</sup> However, conflicting reports have suggested that bi-allelic disease is associated with a tendency for better overall survival compared to mono-allelic disease.<sup>183</sup> That said, a study by Chena *et al*<sup>184</sup> has shown bi-allelic del(13q) patients to have a shorter treatment free survival relative to heterozygous patients. Moreover, copy number neutral loss of heterozygosity (CNN-LOH) events involving the loss of one allele and the subsequent duplication of the alternative allele have been reported in CLL.<sup>183</sup> Interestingly, these events have been reported on 13q, particularly in patients showing bi-allelic deletions of 13q.<sup>183</sup> These observations suggest that the recurrence of CNN-LOH at chromosome 13 is an important mechanism in CLL biology.

It is further argued that detection of del(13q) is not fully sufficient to define a CLL with a good prognosis. For instance, the size of the clone carrying this alteration may influence the disease course. Recent reports reveal that patients harboring a higher percentage of del(13q) cells among their CLL clone have a worse clinical outcome compared to those with a lesser percentage of cells.<sup>182,185</sup> Further, evidence is given by recent SNP profiling studies demonstrating the heterogeneity of 13q deletions in

terms of alteration size, the genes they cover and the location of their breakpoints. For instance, in most 13q deletion cases a minimally deleted region (MDR) covering *DLEU1*, *DLEU2*, *miRNA-15a* and *miRNA-16-1* genes is noted.<sup>186-188</sup> Interestingly, a study by Ouillette *et al* found that patients with del(13q) covering the MDR and the *RB* gene positioned beyond the MDR region were associated with a higher Rai stage compared to del(13q) patients not including the *RB* gene. Moreover, patients with deletions covering the MDR in addition to the *RB* gene were more likely to be treated. Hence, it appears that genes outside of the MDR play key roles in hampering patient outcome.<sup>189</sup>

In del(13q) CLL, it is suspected that deregulation of *miRNA-15a/miRNA-16-1* and hence, deregulation of its downstream target the anti-apoptotic *BCL2* gene, may confer a selective cumulative advantage.<sup>186</sup> The roles of *DLEU1* and 2 are less clear, however it is known that *DLEU2* expression provides the primary transcript for production of *miRNA-15a/miRNA-16-1*.<sup>190</sup> Like *miRNA-15a/miRNA-16-1*, the *DLEU7* gene is also positioned within the MDR in 13q deleted patients. *DLEU7* plays a key function as a potent NFkB inhibitor by interacting with members of the tumour necrosis factor (TNF) receptor family.<sup>191</sup> Moreover, studies on lung cancer cell lines have shown *DLEU7* expression to result in an increase in apoptosis.<sup>191</sup> These observations suggest that loss of *DLEU7* may work in an additive manner with the loss of *miRNA-15a/miRNA-16-1* to augment the pathogenesis of del(13q) CLL.<sup>191</sup>

A recent mouse study has provided *in vivo* evidence suggesting that the *DLEU2/miRNA-15a/miRNA-16-1* locus controls B cell expansion by modulating proliferation rather than affecting survival via the anti-apoptotic *BCL2* gene.<sup>190</sup> This

study suggests that the function of this locus may be to tightly regulate cell cycle entry in response to antigen. This is supported by the observation that *DLEU2/miRNA-15a/miRNA-16-1* expression was down-regulated in proliferating B cells within germinal centers, thus establishing a tumor suppressor role for this locus. Moreover, deletion of the *DLEU2/miR-15a/16-1* cluster in mice was shown to cause the development of indolent B cell-autonomous, clonal lymphoproliferative disorders akin to the spectrum of CLL-associated phenotypes observed in humans. In addition, this study demonstrated that *miRNA-15a/miRNA16-1*-deletion in human and mouse leads to an accelerated B cell lymphoproliferative state that arises through modulating the expression of genes controlling cell-cycle progression.<sup>190</sup>

### **Deletion 11q**

Deletion of 11q is the second most common alteration in CLL (12-18%) and is associated with an aggressive disease course and a poor clinical outcome (Table 1A).<sup>7</sup> These patients are often characterized with abdominal or mediastinal lymphadenopathy.<sup>192</sup> Molecular characterization of del(11q) has identified a minimally deleted region of 2-3Mb that includes the *ATM* gene. In the majority of cases, *ATM* a tumour suppressor gene is considered the most likely candidate gene in del(11q) CLL, since it is known to play a vital role in activating p53 and augmenting the DNA damage response. Due to the integral role of *ATM* in maintaining genome stability, mutation of *ATM* is strongly suggested to be involved in CLL pathogenesis.<sup>192-194</sup> Furthermore, approximately one third of 11q deleted CLL patients display a somatic or germline *ATM* gene mutation.<sup>195,196</sup>

Similar to del(13q) alterations, a marked complexity in the size and structure of 11q deletions has been shown. For instance, a deletion of 11q can be discontinuous which is indicative of complex genomic events taking place.<sup>183,197</sup> In some cases 11q deletions do not cover the *ATM* gene thus indicating the importance of alternative genes in the critical region such as *RDX*. In other instances, genes outside of this region such as the *MLL* gene may play a role in certain 11q deleted CLL cases.<sup>192,194</sup> Recently, SNP analysis has demonstrated a MDR of 0.68Mbp at 11q22.3 encompassing the following genes, *CUL5*, *RAB39*, *ACAT1*, *NPAT* and *KDELC*.<sup>198</sup> However, a more recent study describes a new MDR for 11q22.3 excluding the *RAB39 gene*.<sup>199</sup>

Deletion of *miR-34b/miR-34c* is also thought to play a role in del(11q) CLL pathogenesis. Given the important function of micro RNA in guiding cell proliferation and survival it is interesting to speculate that loss of *miR-34b/miR-34c* may contribute to the aggressive disease state seen in del(11q) patients. In fact one study demonstrates *miR-34b-5p* to target the *TCL1* oncogene which may contribute to *TCL1* overexpression seen in CLL.<sup>200</sup> Surprisingly, a conserved gene in the critical region of 11q22-23 called *ARHGAP20* is shown to be substantially up regulated in 11q deleted cases relative to patients with no genomic alterations and trisomy 12 patients. Interestingly, 13q deleted patients also have a higher expression of this gene.<sup>201</sup> This finding supports a functional link between 11q and 13q deletions in CLL. This notion is further endorsed by the finding that both MDRs in 13q and 11q are syntenic to a single region in the zebra fish on chromosome 9.<sup>202</sup>

## **Trisomy 12**

Today trisomy 12 is considered the third most common genetic alteration in CLL occurring in 11-16% of patients (Table 1A).<sup>7</sup> Trisomy 12 is caused by the duplication of one of its chromosome pairs and is associated with an intermediate survival and shorter time to treatment.<sup>7,203</sup> Little is known of the pathogenesis behind trisomy 12 CLL, however, chromosome banding studies have identified a partial trisomy 12q13-q15 in a minority of patients covering the *MDM2* gene, a known gene deregulated in CLL and a negative regulator of p53.<sup>204-207</sup> Another candidate gene known to be up regulated in CLL is the disease specific *CLLUI* gene located at 12q22, however *CLLUI* expression was not found to relate to trisomy 12.<sup>177</sup> Nevertheless, new gene expression studies have provided further insight into trisomy 12 CLL pathogenesis. Recently, four genes have been shown to be significantly associated with trisomy 12 CLL, including overexpression of the *HIP1R* gene at 12q24 and the *MYF6* gene at 12q21. In contrast, the *P2RY14* gene at 3q21-q25 and *CD200* at 3q12-q13 were shown to be under-expressed in trisomy 12 patients.<sup>208</sup>

## **Deletion 17p**

Deletion of 17p, covering the *TP53* gene, occurs in up to 10% of CLL patients and is associated with a particularly poor survival and short time to treatment (Table 1A).<sup>7</sup> Moreover, deletion of 17p is commonly observed to co-exist with other unfavorable prognostic markers such as unmutated IGHV genes, high ZAP70 and CD38 expression.<sup>159</sup> Additionally, del(17p) patients experience chemo-refractoriness as chemotherapeutic reagents often work by inducing apoptosis or impeding the cell cycle via the p53 pathway.<sup>209</sup> Nevertheless, reagents which utilize p53 independent

mechanisms such as alemtuzumab, an anti-CD52 monoclonal antibody can be implemented in these cases.<sup>28</sup> Recent studies have shown that the majority of 17p deleted patients present with *TP53* mutations on their other allele, however it was also shown that a subset of CLL patients carry *TP53* mutation without a deletion of 17p (Table 1A).<sup>210-212</sup> In fact, patients solely carrying a 17p deletion or *TP53* mutation alone have an equally poor overall survival similar to those patients carrying both del(17p) and *TP53* mutation. Thus, *TP53* mutational status is an independent prognostic indicator that is important for steering treatment choices.<sup>210-213</sup>

### **Deletion 6q**

Today, approximately 6-7% of patients acquire del(6q) (Table 1A).<sup>214,215</sup> However, it is considered a progression marker as it is mainly detected in CLL patients at follow up.<sup>216</sup> Deletions are most frequently encountered at 6q21, however deletions at 6q27 have also been detected.<sup>214-218</sup> No definitive tumour suppressor genes have been accounted for in del(6q) CLL pathogenesis. The presence of del(6q) has been associated with an increased white blood cell count and in some cases a more extensive lymphadenopathy. Despite these associations, it does not appear to be a negative prognostic indicator in CLL. That said, some studies have associated 6q deletions with a shorter treatment free survival.<sup>215,219,220</sup>

**Table 1A.** Summary of Genomic alterations in CLL.

| Alteration              | Del(17p)    | TP53 mutation | del(11q)   | Trisomy 12   | del(13q)         | Deletion 6q |
|-------------------------|-------------|---------------|------------|--------------|------------------|-------------|
| Frequency               | 5-10%       | 1-5%          | 12-18%     | 11-16%       | 55%              | 6-7%        |
| Candidate gene(s)       | <i>TP53</i> | <i>TP53</i>   | <i>ATM</i> | unknown      | <i>miR-15/16</i> | unknown     |
| Clinical outcome        | poor        | poor          | poor       | intermediate | good             | progression |
| Use in clinical routine | yes         | no            | yes        | yes          | yes              | no          |

### Translocations in CLL

In the past, cumbersome classical cytogenetic methods used to detect translocations in CLL were often unsuccessful (see the Genomic profiling technologies section). Hence, the frequency of CLL translocations was grossly underestimated at this time. Today, new and improved cytogenetic analysis has allowed the detection of translocations in up to 30% of CLL patients, however these are largely non-recurrent events.<sup>203,221</sup> The presence of balanced and unbalanced translocations are linked to an unfavourable clinical outcome in CLL.<sup>221</sup> Interestingly, recurring breakpoints in CLL translocations are shown to cluster in regions affected in other lymphomas, such as 1(q21q25), 1(p32p36), 2(p11p13), 6(p11p12), 6(p21p25), 13q14, 14q32, and 18q21.<sup>222</sup> These regions are suspected to contain genes involved in CLL leukemogenesis. For instance the t(14;19)(q32;q13) translocation involves the juxtaposition of the transcriptional co-activator *BCL3* gene to the *IGH* locus causing an increase in *BCL3* expression.<sup>223,224</sup>

### **Genomic complexity**

Recently, a correlation between genetic complexity and clinical outcome has been described in CLL.<sup>225</sup> For instance, a recent SNP study showed CLL patients carrying more than two alterations to have a poorer overall survival.<sup>225</sup> Similarly, patients carrying a higher number of larger sized alterations have been associated with a worse clinical outcome compared to those with smaller alterations.<sup>183</sup> Additionally, the genomes of patients carrying deletions of 11q and 17p are shown to be more complex. For instance, a recent study demonstrated that patients with four or more alterations greater than 5Mbs carried del(17p) or del(11q). Hence, due to the loss of vital genes integral to genome stability, the genomes of 11q and particularly 17p deleted patients are more complex and thus more likely to acquire further alterations compared to other CLL patients.<sup>183</sup>

### **Novel alterations**

Reports have indicated del(14q) as a recurrent alteration in CLL occurring in ~1.5% of patients.<sup>226</sup> It has been shown that this aberration most often occurs in IGHV unmutated patients.<sup>226</sup> Furthermore, del(14q) patients were shown to have a high occurrence of trisomy 12. Deletion of 14q was associated with an intermediate survival and shorter time to treatment although low sample numbers were investigated. Deletions of 14q are heterogeneous in size; however, recurrent breakpoints have been documented.<sup>226</sup> Due to the frequency of these alterations and their association with disease outcome, it is widely debated whether or not to include deletions of 14q to the established FISH panel for detection of known recurrent alterations in CLL.



Through the use of new microarray studies, some novel alterations such as gain of 2p, gain of 8q and loss of 4p, 8p and 22q have been identified,<sup>183,227-229</sup> however no consensus exists regarding the frequency of these novel alterations. This is likely due to discrepancies in the array platforms employed and the study populations investigated. Interestingly, gain of 2p covering the *MYCN* oncogene was recently found in up to 28% of untreated Binet stage B/C patients.<sup>230</sup> These patients also experienced an increased *MYCN* mRNA expression, suggesting *MYCN* to have a pathogenic role in these CLL cases.<sup>230</sup> Additionally, a recent SNP-array study showed this alteration to occur frequently in del(11q) patients.<sup>183</sup> In patients with a loss at chromosome 22q, the *PRAME* gene has been implicated as the candidate gene of interest.<sup>228</sup> Finally, dic(17;18)(p11.2;p11.2) a dicentric chromosome alteration has been described as a novel recurring abnormality in CLL. This aberration is associated with a complex karyotype, progressive disease, unmutated IGHV genes and an early age at diagnosis relative to other CLL patients.<sup>231</sup>

Today, a small percentage of trisomy 12 patients 1.6-5% are known to carry a coexisting trisomy on chromosome 19.<sup>232-234</sup> More recent studies have shown this genetic profile to occur exclusively in isotype switched sIg positive CLL. Interestingly, this profile has been associated with the use of IGHV mutated genes and high CD38 expression. Furthermore, it has recently been discovered that these cases are often found to carry trisomy 18.<sup>235</sup> Notably, trisomy 12 has been reported to be associated with CLL patients expressing IgG-positive stereotyped unmutated IGHV4-39/IGKV1-39 (1D-39) BCRs.<sup>236</sup> In some instances, these cases have been shown to be associated with the t(14;19)(q32;q13) translocation. As mentioned above (in the translocations in CLL section), t(14;19)(q32;q13) causes the *BCL-3* proto-

oncogene to become juxtaposed next to the IGH locus resulting in overexpression of the BCL-3 protein and deregulation of NF- $\kappa$ B.<sup>177,223,237</sup>

With the advent of whole genome sequencing technologies a number of novel mutations have recently been discovered in CLL. Studies have identified repeated mutations in *XPO1*, a nuclear transport factor gene in ~2.4% of patients<sup>238,239</sup> and *NOTCH1*, a ligand activated transcription factor gene in ~12.2% of CLL cases. *NOTCH1* mutations are more frequently found in IGHV-unmutated/ZAP70+ cases compared to IGHV-mutated/ZAP70- patients. More specifically, *NOTCH1* mutations are reported to occur in 18-20% of IGHV-unmutated/ZAP70+ patients and 4-7% of mutated/ZAP70 cases.<sup>239,240</sup> In contrast, a frequency of 4% in IGHV-unmutated/ZAP70+ and 1.5% in mutated/ZAP70 cases has also been reported. *NOTCH1* mutations in CLL have been frequently found to occur in trisomy 12 patients. Altogether, these findings suggest *NOTCH1* activation to play a role in IGHV-unmutated/ZAP70+ trisomy 12 CLL.<sup>238</sup> In addition to these genes, recurrent mutations of *MYD88* and *SF3B1* have been identified in 10% and 15% of CLL cases respectively. Interestingly, *SF3B1* mutations are shown to occur in CLL patients carrying deletion 11q.<sup>241</sup>

## **1.6: Genomic profiling technologies**

Karyotyping ascertains gross chromosomal features such as chromosome number, size and structure. Conventional cytogenetic staining techniques such as G, Q and R-banding pinpoint the chromosomal regions affected by genomic alterations, however, only large alterations (approximately from 5-10 Mb) across chromosome metaphase

spreads are analysed.<sup>242</sup> This method requires cumbersome cell culturing which has often lead to a low percentage of successful karyotypes for CLL cells, given their inherent inability to survive and proliferate *ex-vivo*.<sup>243,244</sup> However, today the use of growth stimulants such as CpG oligonucleotides, IL-2 and CD40L, have allowed better detection rates in CLL.<sup>243,244</sup>

Today, in clinical practice, interphase FISH is instead the gold standard method of choice for CLL, since it has the ability to detect smaller DNA target sequences (probe size starting at 100-1000bp) without cumbersome culturing methods.<sup>245-247</sup> FISH involves the hybridization of a panel of fluorescently labelled region specific probes to complementary DNA sequences at specific chromosomal regions. In CLL, interphase FISH can detect clinically relevant and recurrent lesions in up to 80% of CLL patients.<sup>7</sup> The advantage of FISH relies on its ability to detect alterations even in samples with a low percentage of tumour cells. Other advantages include its capacity to detect homo- and heterozygous deletions and gains, as well as complex chromosomal translocations and rearrangements. That said, FISH is limited in its ability to detect aberrations that lie outside of the probe specific regions, making global screening difficult. Techniques such as spectral karyotyping and multiplex FISH simultaneously employ multiple probes sets in order to asses all chromosomes. However, these technologies are costly and labour intensive and in some instances limited by difficulties in probe design to certain chromosomal regions.<sup>245,247,248</sup>

Unlike FISH, comparative genomic hybridisation (CGH) has the ability to readily examine the entire genome for alterations (approximately 5-10 Mbp in size). CGH involves the quantitative hybridization of differentially labeled normal and tumour

DNA to normal metaphase chromosomes. On the other hand, CGH has a low sensitivity, requiring a high percentage of cells to have a particular alteration in order to be detectable. Furthermore, such alterations as inversions and translocations can not be identified.<sup>249</sup>

Today, the need for genome wide, high resolution scanning has culminated in the development of array technologies and sequencing platforms allowing broader characterization of known alterations and novel smaller genetic lesions contributing to complexity. Array-comparative genomic hybridization (aCGH) and single nucleotide polymorphism (SNP) array technologies investigate global genomic events with the exception of repetitive regions, without the requirement of cumbersome metaphase chromosome preparations. Using aCGH, copy number alterations (CNAs) at a level of <100 kb can be detected.

In general, aCGH utilises the same principals as those of conventional CGH except that instead of employing metaphase chromosomes as targets, thousands of short DNA probes representing the genome are arrayed onto a chip.<sup>250</sup> aCGH designed using bacterial artificial chromosomes (BAC) tend to hold a lower power of resolution<sup>251</sup> compared to aCGH oligo-nucleotide arrays.<sup>252</sup> Oligo aCGH permit a higher resolution to BAC arrays since a higher number of probes made up of 60-mer oligonucleotides are hybridised to the chip. Although a reliable method, aCGH is limited by its inability to detect CNN-LOH events. In contrast, SNP array platforms allow for the simultaneous detection of CNN-LOH events and CNAs at high resolution down to ~10-25kb in size depending on the array.<sup>253</sup>

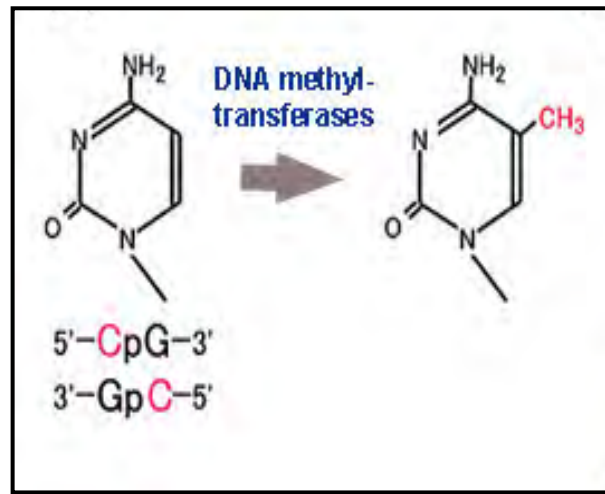
SNP arrays are constructed using 25-30-mer oligonucleotide probes containing specific SNPs. Each SNP position across the genome possesses two allelic variants on the array, SNP A and SNP B. Today up to  $10^6$  probes can be spotted on the array. After hybridising the sample of interest to the array, SNP copy number is determined from the fluorescent signals arising from SNP A and B alleles. Normalising the SNP data to a previously analysed healthy individual reference set determines sample/reference ratios indicating the presence of chromosome alterations. One crux of such SNP and array-CGH technologies is that they do not detect alterations that occur in a low percentage of tumour cells.<sup>254</sup> Failure to detect alterations also depends on the platform and the algorithm used. For instance, some algorithms do not tolerate normal cell contamination >20%, whereas some algorithms are adjusted to account for normal cell interference.<sup>254</sup>

In the case of CLL, alternative optimization measures such as B cell enrichment techniques help resolve the issue of normal cell contamination and low tumour percentage.<sup>254</sup> Recently, the Affymetrix 250K SNP array technology used in this thesis, have been correlated to FISH analysis in CLL patients. Remarkably, this array showed a 98.5% correlation to gold standard FISH analysis. On this basis, this array was suggested for use in clinical routine. Furthermore, this array enabled detection of the same prognostically relevant lesions as higher resolution arrays such as the Affymetrix 6.0 at lower cost.<sup>254</sup>

## 1.7: DNA methylation

### 1.7 *i*: Establishment and maintenance of DNA methylation

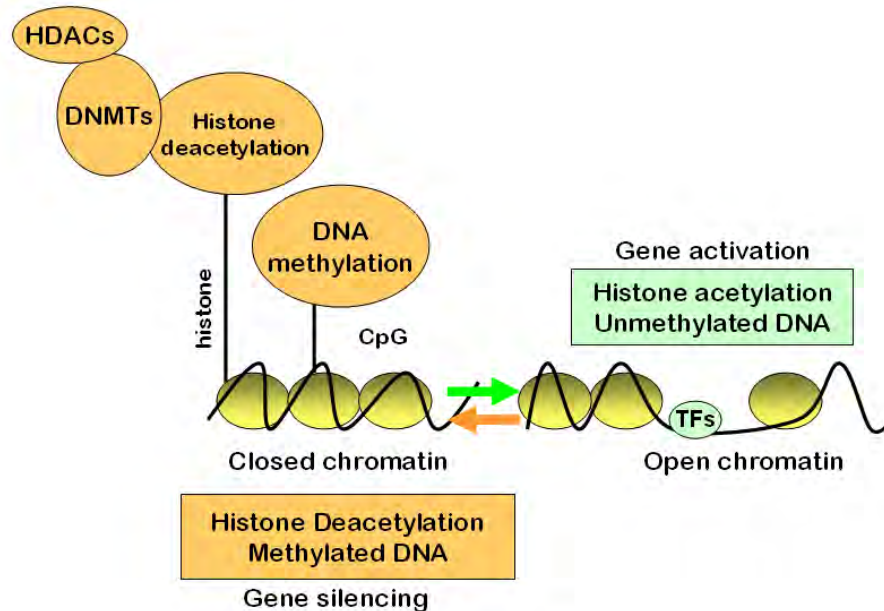
DNA methylation is a heritable epigenetic mark that involves the non-random addition of a methyl group to the fifth position of a cytosine ring within CpG dinucleotides (Figure 1G).



**Figure 1G.** DNA Methylation. Representation of the addition of a methyl group to the fifth position of the cytosine pyrimidine ring.

DNA methylation is accompanied by concomitant histone modifications that play key roles in the control of gene activity and nuclear structure.<sup>255</sup> For instance, DNA methylation and repressive histone marks create a condensed genomic structure that impedes the access or inhibits the binding of the transcription machinery resulting in gene inactivation. Conversely, the presence of activating histone marks and the lack of DNA methylation promotes an open chromatin structure supporting transcription and gene activation (Figure 1H). Together, these actions function to regulate normal

cell differentiation, germline/tissue specific gene expression, genomic imprinting, embryogenesis and transposon gene silencing.<sup>256-260</sup>



**Figure 1H.** Diagram representing the interplay of histone modifications and DNA methylation in the control of gene activity and nuclear structure. TFs: transcription factors.

During early mammalian development, a global de-methylation event occurs during pre-implantation. Once implanted, the embryo undergoes a wave of *de novo* methylation.<sup>261-263</sup> *De novo* methylation patterns are established through the actions of DNA methyl transferases DNMT3A and B.<sup>264</sup> The above actions lead to wide spread global DNA methylation with the exception of CpG islands. Once established, this global bimodal methylation pattern is faithfully maintained through the actions of the ‘maintenance’ DNA methyltransferase DNMT1 enzyme. This enzyme is recruited by the replication machinery to hemi-methylated DNA where it catalyses the transfer of methyl groups to newly synthesized DNA strands after DNA replication.<sup>256,265</sup>

During mammalian development, the bimodal DNA methylation pattern encounters further *de novo* DNA methylating and de-methylating events in order to bring about cell differentiation patterns. Unlike DNA methylation, mechanisms of DNA demethylation in mammals remain largely unexplored. That said, observations of global de-methylation during embryogenesis have pointed to both passive and active modes of action. Passive demethylation is said to occur when DNA methylation pathways are inactivated and DNA methylation is diluted after subsequent replication cycles.<sup>261</sup> Active DNA demethylation in mammals is thought to involve AID induced deamination of methyl cytosine to thymine.<sup>266</sup> This is followed by a base excision repair by DNA glycosylase enzymes of the T/G mismatch.<sup>267,268</sup> Alternatively, *in vitro* studies have shown that DNMTs 3a and 3B can also deaminate 5-methyl cytosine.<sup>269</sup> Imprinted regions are sheltered from active global demethylation as DNA methylation maintenance proteins such as PGC7 and DNMT1 protect them.<sup>258,270,271</sup> Moreover, it is also hypothesized that de-methylation may occur in somatic cells at sequence specific sites.<sup>261,269</sup>

### 1.7 ii: DNA methylation regulation

The success of epigenetic marks governing gene activity and nuclear structure depends on the strict regulation of the epigenetic machinery. DNMT enzymes, histone modification enzymes, DNA-methyl-binding proteins, non-coding RNAs, polycomb complexes amongst others work together to target specific genes for activation or repression. To date, the order and sequence in which these regulators crosstalk to co-ordinate epigenetic marks at specific targets is not fully elucidated. Nevertheless, clues from normal cell development and differentiation provide some



indication as to how DNA methylation may be regulated. In development, histone marks are thought to play a role in establishing DNA methylation. Here, they act as platforms permitting the recruitment, or inhibition of, the DNA methylation machinery.<sup>255,272</sup> In contrast, DNA methylation is considered to play a role in maintaining histone modifications. Methylated DNA bound by methyl binding proteins (MECP2 and MBD2) is said to act as a platform guiding the reassembly of disrupted histone modifications post cell replication.<sup>255,273</sup>

### **Protection of CpG islands during development**

Bimodal DNA methylation patterns set up during development are thought to be directed through specific histone modifications. For instance, unmethylated H3K4 histones are associated with DNA methylation in many cell types.<sup>274,275</sup> At unmethylated H3K4 histones, it is proposed that DNA methylation is permitted through DNMT3L which acts as a bridge to summon active DNMT3A and 3B enzymes to H3.<sup>272,276</sup> H3K4 methylation is said to be regulated by components of the transcription machinery such as RNA polymerase II. Localized mainly at CpG islands, RNA polymerase II works to recruit histone methyltransferases to H3K4.<sup>276</sup> It is thought that repressive methylation marks at H3K4 inhibit the recruitment of *de novo* DNA methylating enzymes, keeping CpG islands free of DNA methylation repression. Alternatively, mouse studies indicate that CpG islands may escape *de novo* methylation through the action of cis acting sequences such as Sp1 sites that bind the Sp1 transcription factors and through unknown mechanisms hinder DNA methylation at these sites.<sup>277</sup>

### **Epigenetic regulation of cell differentiation.**

In embryonic stem cells, differentiation requires pluripotency genes to be silenced. In their active pluripotent state, these genes are found within an open chromatin structure bound by acetylated H3 and H4, methylated H3K4 and unmethylated DNA marks. Upon cell differentiation, gene silencing is thought to involve prompt repression by protein repressor molecules.<sup>255,278</sup> This action is said to signal the recruitment of a histone methyltransferase G9 protein and histone deacetylase enzymes, promoting the formation of a more compact chromatin.<sup>279,280</sup> The actions of G9 results in the methylation of H3K9, however at the same time by unknown means, H3K4 is demethylated. H3K9 methylation acts as a platform allowing the binding of heterochromatin proteins such as HP1. Additionally, G9 summons *de novo* DNA methyltransferases DNMT3A and 3B to promote DNA methylation resulting in the subsequent formation of condensed chromatin at these pluripotency genes.<sup>272,279,280</sup>

### **Polycomb target genes in cell type specific DNA methylation**

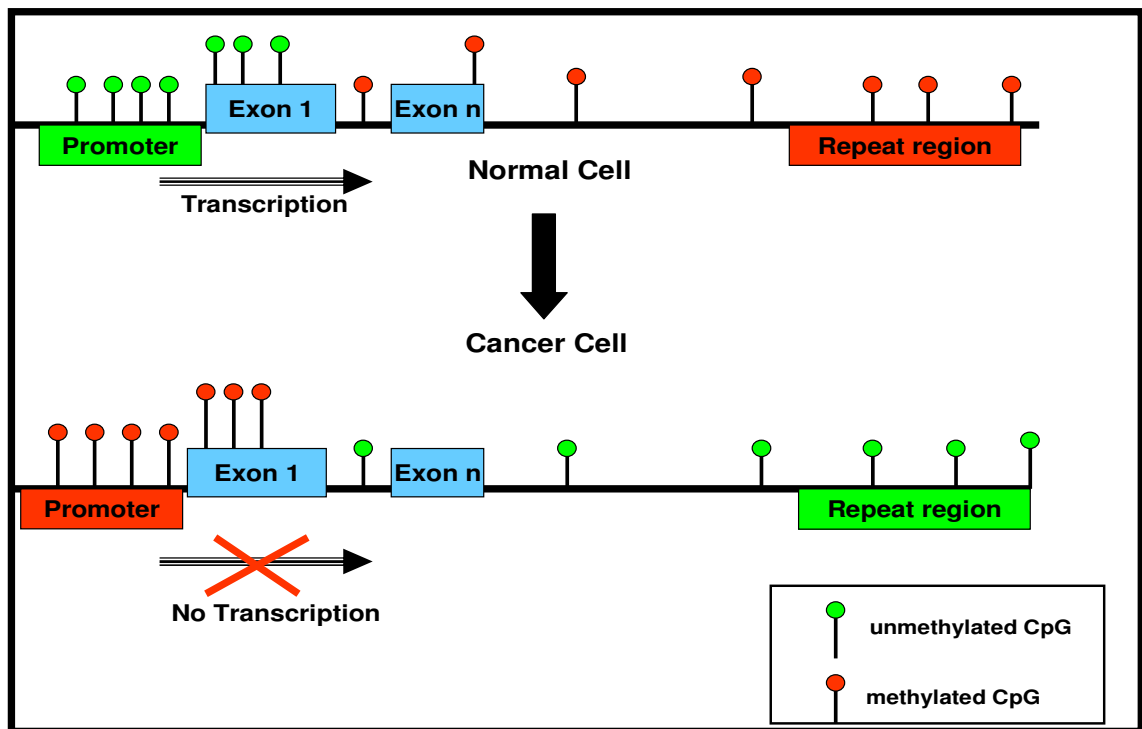
Polycomb-multi-protein repressive complexes silence polycomb target genes vital for mammalian development.<sup>281</sup> More specifically, polycomb target genes encode transcription factors important for embryogenesis, cell lineage specification, growth and proliferation.<sup>282</sup> It is thought that polycomb silencing of these developmental genes may be overcome during germline differentiation. In embryonic stem cells (ESCs), it is proposed that histones carrying easily reversible, bivalent repressive (H3K27me3) and activating (H3K4me3) marks sit at these polycomb target genes, allowing them to actively switch between active and repressive states depending on the demands of the differentiating cell.<sup>283</sup>

Interestingly, at the DNA level, the majority of these target genes stay unmethylated during development.<sup>275</sup> This is thought to be partially attributed to the fact that these genes are mainly affiliated to CpG island promoters that are largely safeguarded from DNA methylation during development. In certain instances, such as the differentiation of ESCs to neural precursors, DNA methylation of polycomb target genes has been described.<sup>275</sup> DNA methylation of these genes is thought to be mediated by polycomb complexes. Polycomb response elements (PREs), regulatory sequence specific elements, are proposed to enable the binding of polycomb (PRCs) subunits, marking the gene for repression.<sup>284</sup> Binding of EZH2, the histone methyltransferase subunit of the PRC2 complex, is known to catalyse the methylation of H3K27. *In vitro* studies indicate EZH2 to interact with the DNMTs to mediate DNA methylation of polycomb target genes.<sup>285</sup>

### 1.7 *iii*: DNA methylation in tumorigenesis

Today, it is widely accepted that alterations to the DNA methylome and changes to histone modification patterns contribute to cancer pathogenesis. In particular, DNA methylation of vital genes involved in key cellular pathways is now recognized as one of the hallmark mechanisms of aberrant gene silencing (Figure 1I).<sup>286</sup> Moreover, hypomethylation of instable repetitive genomic regions is further thought to contribute to cancer (Figure 1I). Unlike genetic lesions, such as mutations and deletions, DNA methylation provokes gene silencing without changes to the DNA sequence.<sup>287</sup> In some instances, aberrant DNA methylation is suggested as an early event in cancer occurring at the stem cell stage. However, other arguments support aberrant DNA methylation as a late event, in tumorigenesis. Many of the DNA

methylation changes found in cancer are said to merely reflect the patho-biology of long lived progressive tumour cells and that these changes are mostly indicative of passenger rather than driver epigenetic events.<sup>288</sup> Overall, altered DNA methylation is one component in a very complex epigenetic network that co-operates with intrinsic cellular defects and tumour promoting extrinsic factors to aid carcinogenesis.



**Figure 1I.** Illustration of DNA methylation in normal and cancer cells. (Adapted from Baylin 2005)<sup>289</sup>

### 1.7 iv: Hypermethylation

As previously described, globally, DNA is largely methylated with targeted regional hypomethylation of most CpG islands in normal cells (Figure 1I). Vast DNA methylation covering repetitive genomic regions and transposable elements ensures the maintenance of genomic stability.<sup>260</sup> On the other hand, hypomethylation of CpG

islands allows the transcription of vital genes governing normal cell growth, proliferation, differentiation and tissue specific gene expression.<sup>257,290</sup> In cancer, the opposite DNA methylation pattern ensues where targeted promoter hypermethylation and global hypomethylation work in synergy with anomalous histone modifications and interaction proteins to disrupt gene regulation.<sup>256,287,291-293</sup> Pivotal to carcinogenesis is the role of promoter DNA methylation and transcriptional silencing of vital tumour suppressor genes.<sup>92,287,292</sup> This largely involves methylation of cytosines within densely populated CpG regions known as CpG islands spanning the 5' end of many genes within their promoters. Nevertheless, new evidence suggests that methylation of CpG island shores and shelf regions of low CpG density located outside the vicinity of traditional CpG islands are fundamental to gene silencing.<sup>294,295</sup> For instance, Doi *et al* have demonstrated a number of overlapping differentially methylated loci involved in normal tissue programming, epigenetic reprogramming to pluripotency and aberrant programming of cancer cells.<sup>294</sup> Interestingly, these loci were shown to be predominantly contained within CpG island shores. These findings suggest that epigenetic changes encountered in cancer are likely due to the reprogramming of the normal pattern of tissue specific differentiation and highlight the importance of CpG island shores in regulating transcription of specific loci.<sup>294,295</sup>

## 1.7 v: Hypomethylation

In cancer, aberrant global hypomethylation has been proposed to inappropriately activate possible oncogenic genes,<sup>260</sup> little evidence of this has been documented. More importantly, hypomethylation of gene poor repetitive DNA sequences that account for up to 30% of the genome are suspected to contribute to genome instability.<sup>296-298</sup>

Furthermore, it has recently been described in mature B cell lymphomas that unlike global hypo-methylation previously implicated in cancer, gene specific hypo-methylation can fraternize with local DNA hypermethylation during tumorigenesis.<sup>299</sup>

### 1.7 vi: DNA methylation and mutagenesis

DNA methylation can further influence genome stability through the introduction of mutations at CpG sites. Methylated cytosines are subjected to a higher rate of mutagenesis relative to unmethylated cytosines.<sup>300</sup> This is thought to occur due to an increased propensity of spontaneous deamination at methylated cytosines. These mutations account for a large number of known point mutations contributing to genetic disease and are said to account for up to 30% of all germline point mutations.<sup>301,302</sup> Moreover CpG sites in the coding regions of *TP53* are methylated in all human tissues.<sup>303</sup> Approximately 25% of all inactivating mutations in cancer in general can be attributed to transitions at CpG dinucleotides.<sup>302</sup>

### 1.7 vii: Role of DNA methylation in CLL

Like other cancers, the genome of CLL patients is hypomethylated compared to PB mononuclear cells from healthy normal controls.<sup>304-306</sup> Until recent times, epigenetics was explored solely on a gene to gene basis. Nevertheless, this candidate gene approach identified a number of DNA methylated promoters contributing to CLL pathogenesis. For instance, *TWIST-2*, a transcription factor and well known silencer of p53 function in other malignancies, was shown to be methylated predominantly in IGHV mutated patients.<sup>307</sup> Similarly, silencing of *ZAP70* through methylation has

been strongly correlated to IGHV mutational status.<sup>308</sup> Additionally, loss or reduced expression of *DAPK1* through promoter methylation has been shown in a large proportion of CLL.<sup>309</sup> In mutated CLL, DNA methylation of *CD38* and *BTG4* has been correlated to a favourable outcome, whereas DNA methylation of *HOXA4* has been affiliated to poor outcome.<sup>310</sup>

More recently, a study by Rush *et al* has demonstrated the genome-wide methylation profiles of ten CLL patients.<sup>311</sup> Using restriction landmark genome scanning (RLGS), a restriction enzyme, 2D gel electrophoresis method, only a limited number of CpG sites (~3,000) were analysed across the whole genome. Interestingly, between 2.5–8.1% of the CpG islands were shown to be aberrantly methylated in CLL samples compared with normal controls. Moreover, 193 aberrantly methylated sequences were identified, of which 93% had CpG island characteristics and 90% showed homology to genes or expressed sequences. Many of the genes identified were transcription factors such as *DERMO1*, *FOXE1*, *TBX3*, and *IPF1*. Other loci known to play roles in the nervous system such as *TBR1*, *GLRB*, and *PAK5* were also identified, however their role in CLL if any is still unknown. Finally, the *GRM7* gene, a gene functioning in the possible inhibition of cyclic AMP signaling in the induction of apoptosis was identified as methylated in CLL patients.<sup>311</sup>

With the advent of higher resolution microarray and sequencing technologies, overall DNA methylation scores have been indicated as strong independent predictors in CLL progression.<sup>312</sup> Patients with a higher genome-wide DNA methylation index relative to their age expected index have been associated with an earlier indication for therapy and Rai stage IV disease.<sup>312</sup> On the other hand, patients displaying a lower

index with respect to age have clinically stable, Rai stage 0-1 disease.<sup>312</sup> Additionally these new technologies have allowed the interrogation of repetitive sequences in CLL patients. Here, aberrant hypomethylation at repeat sequences including ALU, LINES and SAT $\alpha$ , is shown to be more marked in 17p deleted cases. In addition, a lower SAT $\alpha$  methylation level has been shown to act as an independent marker predicting a shorter time to first treatment in CLL.<sup>313</sup> Furthermore, the methylation profiles of CLL patients are shown to segregate according to *CD38* expression.<sup>310</sup>

### 1.7 *viii*: Disruption of DNA methylation regulatory mechanisms in tumorigenesis

In cancer, the epigenetic machinery fall victim to destruction and disorganisation. The exact mechanisms involved in epigenetic deregulation are largely unknown. Genomic loss or mutation to the DNA methyltransferases and other regulatory molecules are suspected to contribute to epigenomic disarray. For instance, somatic mutations decreasing *de novo* DNMT3A activity has been found in up to 20% of acute myeloid leukaemia (AML) patients.<sup>314,315</sup> Moreover, mutations in polycomb repressive complex components have also been described. For example, mutations at tyrosine 641 in the SET methyltransferase domain of EZH2, are shown to change the substrate specificity of EZH2 for various methylation states at H3K27 in non-Hodgkin lymphoma (NHL).<sup>316</sup> For instance, the Y641C mutation abrogates the enzymatic activity of EZH2 against unmethylated and mono methylated H3K27. Remarkably, this mutation allows this enzyme to carry out better catalysation and formation of trimethylated H3K27 compared to the wild type enzyme.<sup>316</sup> Additionally, in myeloid disorders, inactivating mutations of EZH2 have also been reported.<sup>317</sup>



Altered expression levels and improper activation of the epigenetic machinery are also considered to be attributed to epigenetic disorder. For instance high expression of DNA methyltransferases are thought to be involved in the pathogenesis of some cancer cell types. Other studies have noted the over-expression of EZH2 to be associated with cancer progression. One study by Chen *et al* suggests the over expression of cyclin dependent kinases (CDKs) in cancer to aberrantly phosphorylate and activate EZH2.<sup>318</sup> Other *in vitro* prostate cell line studies suggest that increases in EZH2 expression are due to genomic loss of miRNA 101.<sup>319</sup> Interestingly, in MYC induced lymphoma, MYC has been implicated to stimulate *EZH2* expression by repression of its negative regulator miR-26a, thus contributing to lymphomagenesis, therefore a regulatory role for miRNAs in the epigenetic control of gene regulation has been proposed.<sup>320</sup>

### 1.7 ix: Use of DNA methylation markers in clinical routine

In recent years, the plethora of genome wide DNA methylation studies have noted differentially methylated gene patterns in different cancers.<sup>299,321-323</sup> In some instances, these methylation changes constitute tumour and tissue-specific changes that contribute to tumorigenesis,<sup>294,295</sup> thus marking these genes as possible candidate prognostic markers and therapeutic targets in cancer. Despite the volume of these differentially methylated candidates, very few have graduated to be true bona fide prognosticators or therapeutic targets. One reason for this is the inability to decipher true tumour specific differentially methylated genes from normal tissue and patient specific changes.<sup>324</sup> For instance, changes attributed to such factors as age, sex,<sup>325</sup> hormones (androgens)<sup>326</sup> and diet complicate matters since increasing age and poor

folate diets leading to insufficient *in vivo* methyl stores can aid hypomethylation events.<sup>327</sup> Furthermore, changes attributed to patient treatment regimes and disease progression may mask true tumour specific changes.

Another barrier preventing the possible use of such candidate markers in clinical routine is the question of methylation stability over time. To date, little evidence exists on this issue, since many studies are based on snap shot analysis of different individuals and diseases.<sup>328</sup> One study investigating bladder cancer metachronous tumours from two individuals, has shown DNA methylation to be relatively stable over time both within the individual and between patients.<sup>329</sup> In contrast, a study by Bjornsson *et al*<sup>328</sup> focusing on absolute inter- and intra-individual DNA methylation changes have identified substantial inter-individual age variation and considerable differences occurring over time within individuals that would otherwise be missed by group averaging. Interestingly, these changes were shown to be clustered within family members. Hence, these findings support the age related loss of normal methylation patterns as a mechanism for the late onset of disease and further highlight the possible influence of genetic background on DNA methylation changes.<sup>324,328</sup> That said, a study of DNA methylation in 40 monozygotic twins found 14 to have a discordant degree of total DNA methylation, particularly those over 28 years of age.<sup>330</sup> Thus, the latter finding is indicative of environmental factors that are already known to alter DNA methylation. While issues of stability and specificity prevail and since the clinically relevant level of DNA methylation conducive to target gene inactivation is unknown, use of differential DNA methylation changes as candidate prognostic markers or treatment targets remain to be seen.

## 1.8. DNA methylation technologies

DNA methylation quantification methods have evolved rapidly over the past two decades. Initially, global DNA methylation levels were estimated using high performance liquid chromatography (HPLC).<sup>331,332</sup> This method, although highly quantitative, failed to give information on the exact sequences afflicted with aberrant DNA methylation. The need for region specific analysis, provoked the development of cytosine methylation enrichment techniques employing methylation sensitive restriction enzyme (MRE) digestion,<sup>333</sup> *Mcr*BC methylation dependent enzyme fractionation,<sup>334</sup> 5'-methylcytosine antibody immuno-precipitation,<sup>298</sup> methyl-binding protein capture methods<sup>335</sup> and chemical bisulfite treatment approaches.<sup>336</sup> Applying enriched methylated DNA to various microarray and sequencing technologies has allowed the interrogation of DNA methylation at individual regions, particularly within CpG islands and gene promoters. More recently, advancing technologies have enabled a more global view, uncovering the DNA methylation status of non CpG island sites and repetitive sequences throughout the genome.<sup>337</sup>

Early sequence specific genome wide DNA methylation investigations utilised the (RLGS) platform, an MRE based 2D gel method.<sup>338,339</sup> This technology explored the whole genome on a limited scale, deciphering methylation events contained only within MRE recognition sites within a limited number of genes as reviewed in Fouse *et al.*<sup>337</sup> Plagued by false positive and negative results originating from incomplete digestion the true reflection of DNA methylation at these sites is often hampered, as reviewed in Esteller *et al.*<sup>297</sup> Considering the disadvantages of RLGS and the fact that it only assess ~3,000 CpG sites,<sup>311,338</sup> the need for more reliable, higher resolution

analysis paved the way for the use of high fidelity bisulfite treatment in combination with microarray technology.

The onset of bisulfite conversion revolutionised DNA methylation analysis. This allowed the reproducible change of unmethylated cytosines to be deaminated to uracil, leaving methylated cytosines unchanged as they are refractory to treatment.<sup>336</sup> However, bisulfite treatment reduces the sequence complexity of the genome complicating its application to oligonucleotide arrays. Nevertheless, today bisulfite treatment is considered the gold standard method of cytosine DNA methylation enrichment<sup>336</sup> for many high throughput DNA methylation sequencing platforms such as shotgun bisulfite sequencing<sup>340</sup> and DNA methylation microarray Illumina human27 and human450 technologies. The 27K array allows the investigation of ~28,000 CpG sites covering ~14,000 genes, whereas the 450K array covers ~ 485,000 sites across the whole genome. These arrays, particularly the latter permit, the interrogation of CpG sites both within gene promoters and outside of their coding regions.<sup>341</sup>

In brief, microarray platforms interrogating DNA methylation at pre-determined CpG sites<sup>342-344</sup> generally offer less coverage ( $10^3$ - $10^6$ ) of the DNA methylome relative to DNA methylation sequencing methods and next generation technologies ( $10^6$ - $10^7$ ).<sup>337,345,346</sup> Nevertheless, the microarray comprehensive high throughput arrays for relative methylation (CHARM) tiling array provides a high CpG coverage similar to that of some sequencing technologies  $\sim 10^6$ .<sup>334</sup> In general, microarray platforms based on the hybridization of bisulfite treated and MRE prepared DNA can provide single CpG site resolution. In contrast, microarray platforms utilising immunoprecipitated or

methyl binding protein enrichment, quantify DNA methylation across a region of ~1000 base pairs. Hence, the latter provide higher coverage of the genome yet at a poorer resolution (as reviewed in Fouse *et al*).<sup>337</sup> Such techniques as MeDIP and MIRA-Chip rely on the faithful precipitation of DNA methylated sites using specific antibodies against 5'methylcytosine or methyl CpG binding proteins which bind methylated cytosines. Enriched DNA is then hybridized to a genomic microarray.<sup>298,335</sup> A summary of the advantages, disadvantages and characteristics of some of the available DNA methylation platforms is described below in Table 1B and 1C.

**Table 1B:** Summary of DNA methylation microarray platforms.

| Microarray Technology                         | Pre-Processing              | CpG site Coverage/Resolution                           | Advantages  | Disadvantages  |
|---|-----------------------------|--|---|--|
| Illumina Human27                              | Bisulfite conversion        | ~27,000/<br>~1 bp                                      | Cheap<br>Reliable<br>Single bp resolution                                     | Low coverage<br>Promoter biased<br>No assessment of repeat regions                               |
| Illumina Human450                             | Bisulfite conversion        | ~450,000/<br>~1 bp                                     | Reliable<br>Single bp resolution  | Low coverage<br>Inability to assess repeat regions   |
| HELP-Chip Promoter array                      | MRE digestion               | 10 <sup>5</sup> /~1000 bp                              | High genome coverage  | Low resolution<br>Limited to MRE sites   |
| MIRA-Chip (CGI array) (Global tiling array)   | Sonication or MRE/MB column | 10 <sup>5</sup> /~1000bp<br>10 <sup>6</sup> /~1000 bp  | High genome coverage<br>No need for ssDNA                                     | Low resolution<br>Limited to MRE sites   |
| MeDIP-Chip Promoter array Global tiling array | Sonication/ IP              | 10 <sup>5</sup> /~1000 bp<br>10 <sup>6</sup> /~1000 bp | High coverage<br>Not limited by restriction sites                             | Low resolution<br>Enrichment bias towards high density CpG sites<br>Needs ssDNA                  |
| CHARM tiling array                            | McrBC digestion             | 10 <sup>6</sup> /~1000 bp                              | High coverage<br>Less biased towards high density CpG sites compared to MeDIP | No single site resolution<br>Limited to McrBC sites<br>Can not interrogate low density CpG sites |

IP: immunoprecipitation, MRE: methylation sensitive restriction enzyme, CGI: CpG island, MB column: methyl binding protein column, CHARM comprehensive high throughput arrays for relative methylation. HELP- *HpaII* tiny fragment enrichment by ligation mediated PCR. MeDIP- methylated DNA immunoprecipitation, MIRA: methylated CpG island recovery assay, ssDNA: single stranded DNA. Adapted from Fouse *et al* 2010.<sup>337</sup>

**Table 1C:** Summary of DNA methylation sequencing platforms

| Sequencing Platforms         | Pre-processing        | CpG site Coverage / Resolution          | Advantages   | Disadvantages  |
|------------------------------|-----------------------|---|--|--|
| Methyl-seq /Help-seq         | MRE digest            | 10 <sup>5</sup> -10 <sup>6</sup> /~1 bp | High coverage<br>Interrogates repeat sequences<br>Single bp resolution | Limited to MRE sites                                     |
| Shotgun Bisulfite sequencing | Bisulfite /sonication | 10 <sup>7</sup> /~1 bp                  | High coverage<br>Interrogates repeat sequences<br>Single bp resolution | Difficulties in assembling bisulfite converted sequences |
| MeDIP-seq                    | Sonication/ IP        | 10 <sup>7</sup> /<br>200-300 bp         | High coverage<br>Interrogates repeat sequences                         | No single site resolution<br>Complicated to analyse      |

IP: immunoprecipitation, MRE: methylation sensitive restriction enzyme, HELP- *HpaII* tiny fragment enrichment by ligation mediated PCR, MeDIP- methylated DNA immunoprecipitation, MIRA Adapted from Fouse *et al* 2010.<sup>337</sup>

Although all the aforementioned technologies offer useful information, interpretation of observations from these methods often need to be independently reconfirmed with alternative methods. Semi-quantitative methods such as methylation specific PCR (MSP)<sup>347</sup> and bisulfite sequencing are often employed.<sup>348</sup> MSP, although easy to implement, is restricted in that it solely interrogates the CpG sites covered by the primer. Unlike MSP, bisulfite sequencing is a labour intensive technique. That said, bisulfite sequencing is much more advantageous than MSP since it readily interrogates CpG sites contained within the amplified region determined by the primers. Fully quantitative techniques such as pyrosequencing, can overcome some of the disadvantages of the former methods. Specifically, pyrosequencing easily allows the quantification of DNA methylation across a chosen site, often containing a number of CpG sites with minimal DNA input.<sup>349,350</sup>

## 1.9: Epigenetic based treatment

Epigenetic defects are heritable changes that do not affect DNA sequences; hence their effects are potentially reversible through the use of methyl inhibitors and histone deacetylation drugs. Today, the use of hypo-methylating drugs 5-azacytidine (5'aza) and its more potent counterpart 5-aza-2'-deoxycytidine (DAC) has been approved for the treatment of myelodysplastic syndrome patients. The use of histone deacetylase inhibitors also represents another group of potential drugs that can be used in epigenetic cancer therapy. Recently the first one of its type, suberoylanilide hydroxamic acid (SAHA), has been approved for the treatment of cutaneous T cell lymphoma.<sup>351</sup>



## 2) THESIS AIMS

The overall aim of this thesis was to screen for genetic and DNA-methylation events in divergent CLL prognostic subgroups and investigate DNA methylation over time and within different CLL compartments. The specific aims were as follows:

- To determine the frequency of IGHV3-21 gene usage among a Swedish population-based cohort using Sanger sequencing and assess the impact of IGV3-21 stereotypy status on clinical outcome. (Paper I)
- To investigate genomic alterations in CLL patients expressing stereotyped and non-stereotyped IGHV3-21 and IGHV4-34 BCRs using 250K SNP arrays. (Paper II)
- To study DNA methylation profiling in prognostically diverse CLL subsets using 27K DNA methylation arrays. (Paper III)
- To evaluate the contribution of time and microenvironment on the global DNA methylation profiles of CLL patients, by assaying patient-matched diagnostic and follow-up samples and patient-matched LN and PB samples on 450K DNA methylation arrays. (Paper IV)

## 3: MATERIALS & METHODS

### 3.1: Patient samples and patient characteristics

In Paper I, peripheral blood (PB) was obtained from 337 patients at diagnosis including, 325 CLL and 12 CLL/small lymphocytic leukaemia (SLL) cases from the Swedish cohort of a Scandinavian population-based case-control study called SCALE (Scandinavian Lymphoma Etiology).<sup>352</sup>

In Paper II, tumour material from 101 CLL patients was obtained from four cooperating institutes in France ( $n=15$ ), Greece ( $n=25$ ), Denmark ( $n=4$ ) and Sweden ( $n=57$ ). Samples mainly originated from peripheral blood ( $n=88$ ), although samples from spleen ( $n=5$ ) and bone marrow ( $n=8$ ) were also included. In total, 42 IGHV3-21 samples, encompassing 29 subset #2 and 13 non-subset #2 cases, and 59 IGHV4-34 samples, comprising of 17 subset #4, 8 subset #16 and 34 non-subset #4/16 cases were analysed. Subsets were defined according to Stamatopoulos *et al* and Murray *et al*.<sup>5,6</sup> A CLL sample set ( $n=203$ ) from a recent array study was used for comparison purposes.<sup>183</sup>

In Paper III, 23 CLL samples (6 IGHV mutated, 7 IGHV unmutated and 10 IGHV3-21) were obtained from the Biobank at the Department of Pathology, Uppsala University Hospital, Sweden. All samples were derived from peripheral blood. For validation experiments, an independent set of 10 and 50 CLL samples were analysed by MSP-PCR and RQ-PCR respectively. For control purposes, one peripheral blood

mononuclear cell (PBMC) sample, one CD19 sorted B-cell sample and one negative whole-genome amplified DNA control were additionally analysed. For comparison purposes, 2 CLL EBV transformed cell lines (I83 derived from a mutated CLL patient and HG3 derived from an unmutated CLL patient) were included for analysis.

In Paper IV, a total of 36 PB CLL samples obtained from the Swedish cohort of the SCALE study<sup>352</sup> were chosen for analysis. To study the DNA methylation profiles of IGHV mutated and unmutated subgroups, sorted PB samples derived from 9 IGHV mutated and 9 unmutated patients at diagnosis were analysed. Sorted PB samples derived from these same patients were taken at follow-up to investigate DNA methylation over time. (Patient-paired samples: IGHV mutated cases n=9 cases at diagnosis and n=9 cases at follow-up and IGHV unmutated cases n=9 cases at diagnosis and n= 9 cases at follow-up) In addition, a separate set of 10 patient-matched PB and LN samples collected from biobanks at Uppsala University and Linköping University hospitals were analysed. PB from three age-matched normal individuals and one whole-genome amplified DNA derived from PB were included as controls.

All samples were diagnosed according to the WHO classification and the International Workshop on Chronic Lymphocytic Leukaemia criteria (IWCLL).<sup>28</sup> In Paper I, samples with highly variable tumour load were included for analysis. In Papers II and III, only samples with  $\geq 70\%$  tumour cell content were included for analyses. In Paper IV, all samples were CD5+/CD19+ sorted. Clinical and molecular data characterising Binet stage, IGHV mutational status, treatment and survival status, and chromosomal alterations are detailed in Papers I-IV.

### **3.2: Ethical Approval**

All samples and clinical data included in the research described herein were collected following strict ethical guidelines. Written informed consent was obtained according to the Declaration of Helsinki. All personal data was de-personalised prior to analysis and publication to maintain the privacy of the participant. Traceability procedures were put in place allowing full traceability of patient samples and clinical data should the participant wish to withdraw from such studies at any point. Each study was approved by the local Ethics Review Committee from all collaborating institutions mentioned in the relevant studies. Registration numbers of the ethical review board decision from Uppsala University and the Karolinska Institute Sweden are as follows Ups:01-082 and 2006/964-31/2.

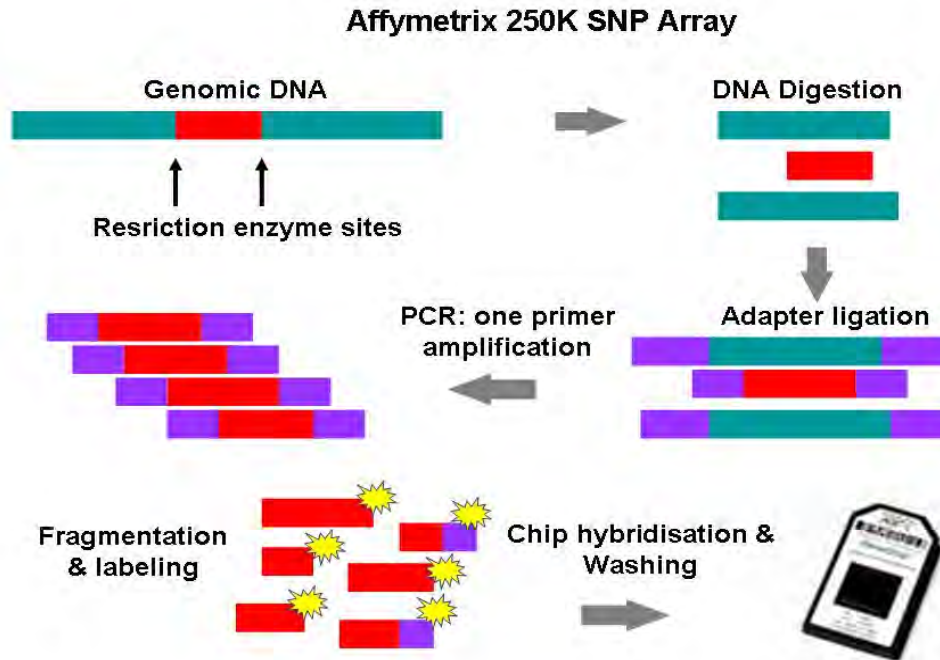
### **3.3: IG gene sequencing**

In all Papers, CLL patients were characterized according to IG gene usage, mutational status and the presence of stereotyped and non-stereotyped receptor sequences.<sup>5,6</sup> These characteristics were detected through PCR amplification of gDNA or cDNA using IGHV gene family specific framework 1 or leader primers.<sup>353</sup> Consensus primers for the kappa/lambda FR1 and IGKJ/IGLJ genes<sup>6,147,149</sup> were used to amplify light chain IG rearrangements. Following amplification, most clonal products were sequenced using the Big Dye terminator cycle sequencing reaction kit (Applied Biosystems, Foster City, CA). Where direct sequencing was unsuccessful, gel purification (Qiagen, Venlo, Netherlands) and/or cloning prior to sequencing was performed. Sequences were then aligned in the IMGT database using the IMGT/V-

QUEST online tool (<http://www.imgt.org>).<sup>354</sup> IGHV genes were denoted unmutated where a  $\geq 98\%$  identity to the corresponding germline gene was determined. Sequences with  $< 98\%$  identity were considered mutated. In Paper I, IGHV unmutated cases were further classified into truly unmutated (100% identity to germline), minimally (99-99.9% identity) and borderline mutated (98-98.9% identity). Stereotypy status was determined based on established molecular IG sequence criteria.<sup>5,6</sup>

### **3.4: 250K Affymetrix SNP array**

In Paper II, 250K *NspI* SNP-arrays were applied to study CNAs and CNN-LOH events in IGHV3-21 and IGHV4-34 stereotyped and non-stereotyped samples. Array experiments were performed according to the standard protocols for Affymetrix GeneChip<sup>®</sup> Mapping *NspI*-250K arrays (Figure 3A) (Gene Chip Mapping 500K Assay Manual (P/N 701930 Rev2), Affymetrix Inc, Santa Clara, CA, USA). Subsequent to array scanning using the GeneChip<sup>®</sup> Scanner 3000 7G, probe level normalisation and SNP calling were accomplished using Affymetrix GeneChip<sup>®</sup> Genotyping Analysis Software (GTYPE) 4.1.



**Figure 3A** The 250K Affymetrix SNP array contains 262,000 SNP probes covering the whole genome. In short, the tumor DNA is enzymatically fragmented using the *NspI* restriction enzyme and adapters are added allowing PCR amplification of the fragments. After amplification the fragments are further fragmented and end labeled. Following hybridization of these labeled fragments to the array, copy number is estimated based on the ratio of fluorescence signals detecting gains and losses. (Adapted from affymetrix.com)<sup>355</sup>

Single sample QC was performed using the Dynamic model algorithm and genotype calls were made using the BRLMM algorithm.<sup>356</sup> Quality control measures were given as a neighbor score. This score was calculated by averaging the euclidean distances between the  $\log_2$  ratios of 5 consecutive SNPs. A neighbor score of  $\leq 0.4$  was applied as cut-off for sample inclusion. Log transformation of the data producing  $\log_2$ -ratios allowed for a more symmetrical data distribution. Subsequently, copy-number normalization was performed using the Copy Number Analysis Tool (CNAT) 4.0.1 and normal samples (n=82) previously run at the Uppsala Array Platform were used as a reference set for comparison.

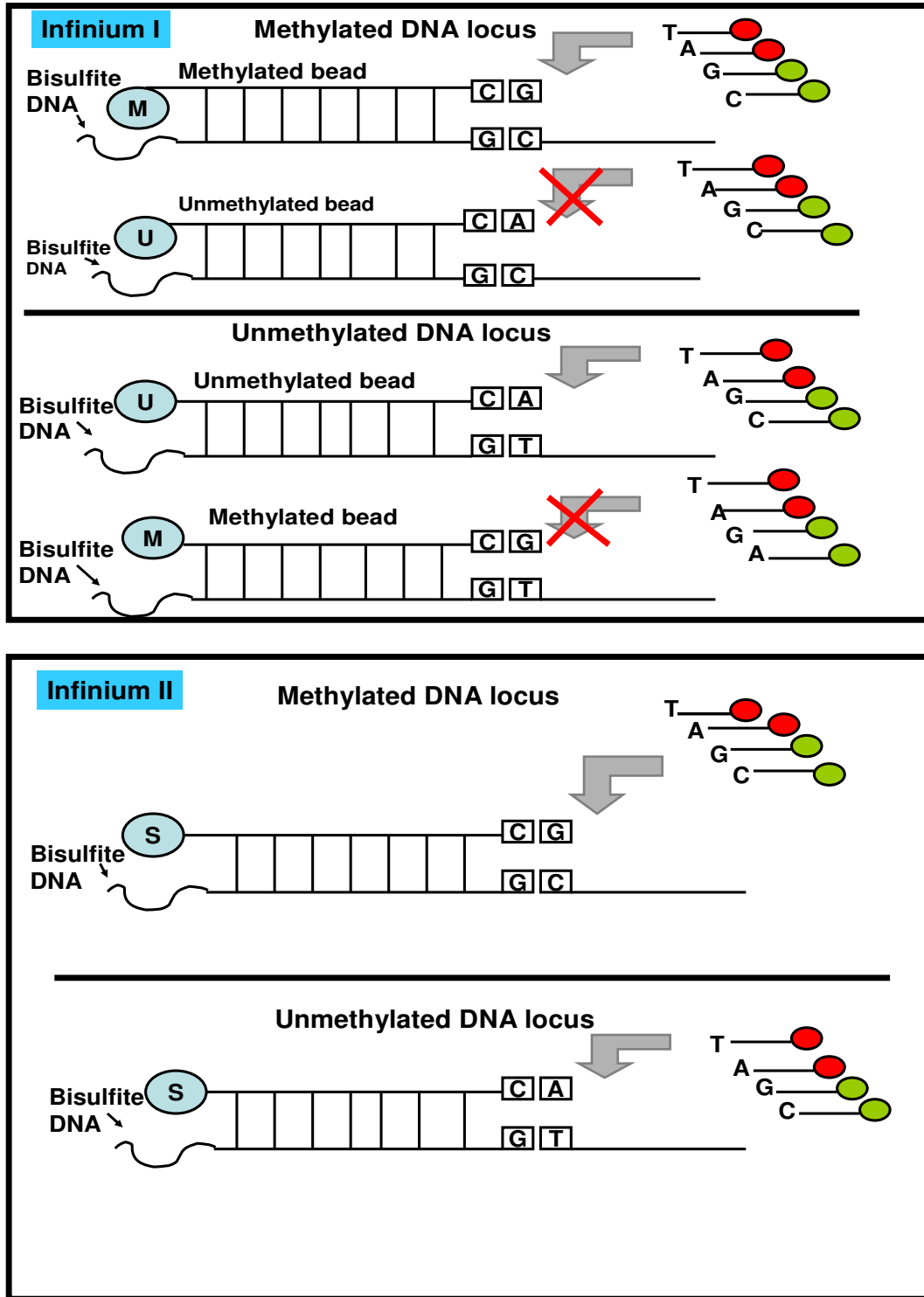
Copy-number analysis using the BioDiscovery Nexus Copy Number 3.0 software (BioDiscovery, El Segundo CA, USA) operating the Rank Segmentation algorithm was employed. CNAs were defined on significance thresholds (p-value) of  $1 \times 10^{-6}$  and a  $\log_2$ ratio cut-off at  $\pm 0.2$  for regions sized 200-500 kbp and  $\pm 0.15$  for regions  $> 500$  kbp, respectively. These settings were defined through previously described real-time quantitative PCR (RQ-PCR) validation experiments.<sup>183</sup> To exclude polymorphic regions, CNAs that overlapped with  $> 50\%$  of a known CNV region were not considered as true CNAs and excluded from analysis. CNN-LOH analysis was performed using SNP array-data, taking into account the fraction of normal cells indicated from the flow-cytometry data.<sup>357</sup> CNN-LOH regions larger than 3 Mbp with less than 50% overlap to CNVs were acknowledged as legitimate CNN-LOH events.

### **3.5: Illumina Infinium HumanMethylation27 and HumanMethylation450 BeadChip array**

In Paper III, the HumanMethylation27 array (Illumina, San Diego, USA) was applied to study the differences in methylation profiles between mutated, unmutated and IGHV3-21 CLL patients. In Paper IV, the HumanMethylation450 array was employed to study genome wide methylation differences between 1) IGHV mutated and unmutated subgroups, 2) paired diagnostic and follow-up CLL samples and 3) paired matched PB and LN CLL samples. All samples underwent bi-sulfite conversion of genomic DNA according to the standard EZ DNA Methylation Kit protocol. (Zymo Research, Orange, CA, USA).

Following conversion, the DNA was whole-genome amplified, enzymatically fragmented, precipitated, resuspended and hybridized over night to locus specific oligonucleotide primers. The 27K array based on the Infinium I assay design employs locus specific oligonucleotide primers that are linked to individual bead types (Figure 3B). Two bead types correspond to each CpG locus – one to the methylated cytosine (C) and the other to the unmethylated thymine (T) state. Following hybridization, single base primer extension allows the detection of C or T nucleotides through a fluorescent signal that is measured using the Illumina BeadStation GX scanner. The intensities of the unmethylated and methylated bead types are measured and a DNA methylation value, described as the Beta values or methylation index (MI), is recorded for each locus via BeadStudio software. Unlike the 27K array, the 450K array employs both Infinium I and Infinium II technologies (Figure 3B). The Infinium II design utilizes a single bead type. In this case, the methylation status is calculated after hybridisation at the single base extension phase. The intensity of the fluorescent signal emitted from this extension is scanned and measured. Since the 450K array elicits the use of both Infinium I and II assay designs, the array can increase genome wide coverage across CpG sites compared to the limited coverage of the 27K array. In both array types, the methylation status for each detected CpG can range between 0 (completely unmethylated) to 1 (completely methylated). MI cut-offs of  $\geq 0.7$  and  $\leq 0.4$  (average of all MI values in each group) were considered methylated and unmethylated, respectively.





**Figure 3B.** Diagram representing the principals of the HumanMethylation 27 and 450 BeadChip arrays. The 27K array is based solely on the infinium I design employing two bead types per locus, one for the methylated (M) and one for the unmethylated (UM) state. The 450K array is based on both Infinium I and II design. The Infinium II design uses only a single bead type (S). The methylated state is determined at base pair extension. (Adapted from Illumina.com)<sup>341</sup>

### **3.6: Bioinformatic analysis of the Illumina human27 and 450**

#### **DNA methylation array data**

To systematically identify differentially DNA methylated genes in paper III, bioinformatic analysis of the methylation data was carried out using the statistical computing language R (<http://www.r-project.org>) software. Similarly for Paper IV, bioinformatic analysis of the methylation data was conducted using the R based illumina methylation analyzer (IMA) package. In both studies, raw data was quantile normalized, arcsin transformed and an empirical Bayes moderated t-test was then applied using the 'limma' package to identify differential methylation patterns between the various sample comparisons. The p-values were adjusted using the method of Benjamini and Hochberg, and a level of  $p < 0.05$  was used as a cutoff.<sup>358</sup> In Paper III, to ensure only CpG sites with a large methylation differences were obtained, an average geometric difference in MI of 0.45 between the IGHV mutated and IGHV unmutated subgroups, 0.35 between the IGHV3-21 and IGHV mutated subgroups and 0.35 between IGHV3-21 and IGHV unmutated subgroups was applied.

In Paper IV, an average difference in MI of  $\geq 0.40$  was applied to decipher recurrent differentially methylated sites between the IGHV mutated and IGHV unmutated patients at diagnosis. To identify recurrent methylation differences between paired diagnostic and follow-up samples, and patient matched samples from different anatomical compartments, a paired t-test was adopted. Again, P values were adjusted according to the Benjamini and Hochberg method and a P value of  $< 0.05$  was applied as a cut-off. To determine recurrent differentially methylated sites across the paired

over time and compartment samples, an average difference in MI of  $>0.20$  was first applied. On recognising the low number of recurrent sites passing the  $>0.20$  threshold an average difference  $\leq 0.20$  was employed. In order to assess the likelihood of small methylation differences representing technical variation,  $R^2$  values were calculated for diagnostic/follow-up pairs and patient matched compartment pairs and compared to the  $R^2$  value reported by Illumina for technical replicates assayed on the 450K array ( $R^2 > 98$ ).

To assess non-recurrent intra-individual methylation differences between patient-paired samples and patient-matched compartment cases that would otherwise be masked by averaging of the data, intra-individual analysis based on non transformed data of the paired samples was conducted. Specifically, the number of sites differing by  $\geq 0.4$  in beta value (MI) between the individual diagnostic and follow-up pairs was employed. Sites differing by  $\geq 0.4$  in MI between the LN and PB patient-matched pairs was first applied to determine non recurrent differences across the compartments. Subsequently a filter of  $\geq 0.3$  in beta value was assessed.

### **3.7: Methylation-specific PCR analysis**

In Paper III, methylation array results were verified using methylation specific (MSP)-PCR for 4 differentially methylated genes. Methylated specific primers were designed according to Li and Dahiya selection criteria using the Methprimer design software (Bisearch, San Francisco, USA).<sup>359</sup> Standard PCR protocol was applied and bands were visualised on a 2% agarose gel.

### **3.8: Real-time quantitative PCR analysis**

In Paper III, to investigate the influence of methylation status on gene expression of specific genes, the expression of 4 differentially methylated genes were tested using RQ-PCR on 60 independent samples. Total RNA was extracted using the RNA extraction Kit (Qiagen, Venlo, Netherlands) and reverse transcription was performed using the MMLV-RT kit (Invitrogen, Carlsbad, CA) and random hexamers (Fermentas, Burlington, Canada) according to the manufacturer's protocol. RQ-PCR primers were designed using the Primer3 software (Broad institute, Boston, USA) ([www.wi.mit.edu/](http://www.wi.mit.edu/)). RQ-PCR analysis was performed using 2X SYBR green master mix (Fermentas, Burlington, Canada) and expression was analyzed using the Stratagene Mx 3005p (Stratagene, La Jolla, CA, USA) detection system and Max Pro QPCR software (Stratagene).

### **3.9: Bisulfite sequencing**

To confirm the methylation status of individual genes in Paper III, two genes underwent bi-sulfite sequencing. Bi-sulfite sequencing primers (BSP) were designed according to primer selection criteria as outlined by Li and Dahiya.<sup>359</sup> A touch down PCR was applied as described in Paper III. The PCR product was cloned using the Invitrogen TOPO TA cloning kit-2.1 TOPO vector (Invitrogen, Carlsbad, CA, USA), and 10-12 positive clones per sample were selected. The product was subsequently exonuclease/shrimp alkaline phosphatase (ExoI/SAP) purified and sequenced using the BigDye Terminator Cycle Sequencing Reaction Kit (Perkin-Elmer, Foster City, CA). Sequences were aligned and analyzed using the bisulfite sequencing web-based

tool BiQ Analyzer software (Max-Planck Institut für informatik, Saarbrücken, Germany).

### **3.10: 5-aza-2'-deoxycytidine and Trichostatin A treatment**

To test for the re-expression of methylated genes after methyl inhibitor and/or histone deacetylase inhibitor treatment, primary CLL cell samples and CLL cell lines were cultured in supplemented RPMI-1640 media as outlined in Paper III. Confluent cells were subdivided and cultured over 3 days in supplemented RPMI media treated with one of the following treatments: 1) medium containing 5-aza-2'-deoxycytidine (DAC) (5 $\mu$ M/L, Sigma Aldrich, St. Louis, MO) for 72 hours (medium was changed every 24 hours), 2) medium containing Trichostatin A (TSA) (500nM/lit, Sigma Aldrich, St. Louis, MO) for the last 24 hours, and 3) medium containing DAC for 48 hrs followed by addition of TSA for 24 hrs. Control cells were cultured in similar way with no drugs added.

### **3.11 Pyrosequencing**

In Paper IV, pyrosequencing was chosen to validate the methylation status of selected differentially methylated genes seen on the Illumina 450K methylation array. Pyrosequencing, a sequencing by synthesis method, gauges the primer directed and polymerase extended incorporation of specific nucleotides to a bisulfite converted DNA template. Pyrosequencing quantitatively measures the incorporation of nucleotides since the amount of pyrophosphate emitted is ultimately converted to a

proportional light signal. The degree of methylation is determined at each CpG position in the analysed sequence through the ratio of T and C.<sup>350</sup>

First, bisulfite converted DNA was PCR amplified using a forward and 5' biotin labeled reverse primer pair designed using the PyroMark<sup>TM</sup> software (Qiagen, Venlo, Netherlands). Primers were designed to cover the exact CpG sites identified on the 450k methylation array. Once amplified, PCR products were immobilized on streptavidin sepharose beads, denatured to form single strands and annealed to the sequencing primer (designed by PyroMark<sup>TM</sup> software). Separation and annealing steps were enabled by the PyroMark Q24 prep vacuum workstation. Following annealing, CpG site methylation analysis was performed on the PyroMark Q24 pyrosequencing instrument and interpreted on the PyroMark Q24 software (Qiagen, Venlo, Netherlands).

### **3.12. Statistical analysis**

In Paper I, the statistical significance of IGHV3-21 gene frequency differences incurred across our study and others was determined by a chi-squared test. Overall survival was measured from the diagnosis date until the date of death or last known follow-up. Time to treatment was calculated from the diagnosis date to the date of first known treatment. Kaplan-Meier analysis visualized OS and TTT within the different groups studied in the form of survival curves. A log rank test was employed to determine the statistical differences in OS and TTT between the studied groups.

In Paper II, statistical differences in the frequency and size of the alterations between the studied groups were calculated via the Chi-square test, one way ANOVA test and t-test where appropriate. In Paper III, statistical differences in expression as determined through RQ-PCR were calculated by a t-test and represented in box plot graphs. In Paper IV, statistical differences in the level of methylation determined by pyrosequencing were calculated using a t-test. All statistical analyses were performed using the Statistica 8.0 and 10.0 software (StatSoft, Tulsa, OK).

## **4: RESULTS & DISCUSSION**

### **4.1: Paper I: IGHV3-21 gene frequency in a Swedish population-based cohort**

A decade ago, the importance of IGHV3-21 gene usage as an independent prognostic marker was discovered.<sup>135</sup> In light of IGHV3-21 as a poor prognosticator irrespective of mutational status, a myriad of studies were then dedicated to investigate the prevalence and clinical outcome of this gene. Across European and North American CLL cohorts, CLL was established to display a biased IGHV gene repertoire with an over-representation of IGHV1-69, IGHV4-34, IGHV3-7, IGHV3-23 and IGHV3-21 genes.<sup>3,134,135,149</sup> Additionally, studies continually demonstrated both mutated and unmutated IGHV3-21 patients to experience a poor survival, similar to that of other unmutated CLL cases utilising alternative IGHV genes.<sup>137,150,152</sup> However, on comparing the predominance of this poor prognostic IGHV3-21 gene across different geographical regions, large variances in IGHV3-21 gene frequency were continually noted. Remarkably, Scandinavia has been shown to present with a high IGHV3-21 gene frequency (10-13%)<sup>4,135,137,150</sup> compared to other regions namely the USA and Southern Europe (3-4%) (Table 4.1 A).<sup>5,151,152</sup>

On recognising the geographical bias of IGHV3-21 and the particularly poor clinical outcome of these patients, attentions turned to the biological reasons behind such findings. Sequencing of IGHV3-21 gene rearrangements noted a large proportion of



cases to carry stereotyped HCDR3 sequences.<sup>5,135,137,147,149</sup> Moreover these stereotyped cases were shown to frequently present with restricted IGLV3-21 gene usage.<sup>137,149,150</sup> Stemming from the findings of IGHV3-21 over representation and IGHV3-21 BCR sequence restriction, the prospect of a specific antigen contributing to IGHV3-21 leukemogenesis was suggested.<sup>25,144,148,360</sup> Interestingly, some studies have reported stereotyped IGHV3-21 (subset #2) patients to have a shorter time to progression relative to their non stereotyped IGHV3-21 counterparts.<sup>149</sup> That said, other studies have not found such a correlation, thus no consensus exists regarding the clinical impact of IGHV3-21 stereotypy in CLL.

Since many of these above findings are based on hospital-based CLL cohorts and are subjected to large cohort selection bias, it was consequently of interest to determine the frequency of IGHV3-21 gene usage among a CLL Swedish population based cohort (n=337). Subsequently, we attempted to address the clinical impact of IGHV3-21 stereotypy and mutational status on CLL patient survival and progression.

This population-based study reports the Swedish IGHV3-21 frequency to be 6.5%. Although this frequency is lower compared to previous Scandinavian rates constructed from university hospital data (10-13%),<sup>4,135,137,150</sup> the frequency still remains higher relative to other large studies from the USA<sup>152</sup> and the Mediterranean (3-4%) (Table 4.1 A).<sup>5,151,152</sup> In light of this finding, the differences in frequency encountered between small hospital-based cohorts demonstrating an over-representation of IGHV3-21, such as Ireland (7.9%),<sup>361</sup> Britain (9.6%),<sup>154</sup> and the Ukraine (5.8%)<sup>153</sup> must be interpreted with caution, since discrepancies can be introduced by such factors as sample size and cohort selection bias. That said, despite

the increased tendency to select for more aggressive CLL cases in hospital-based studies, such reports from the Mediterranean and USA still find IGHV3-21 frequency to remain rather low (3-4%).<sup>5,151,152</sup> Despite the different nature of the CLL cohorts studied, in this population-based report versus our previous hospital-based study we reconfirmed that a true IGHV3-21 geographical bias exists in Sweden.

**Table 4.1 A** Comparison of IGHV3-21 gene frequency in large CLL cohorts from Scandinavia, Mediterranean and USA

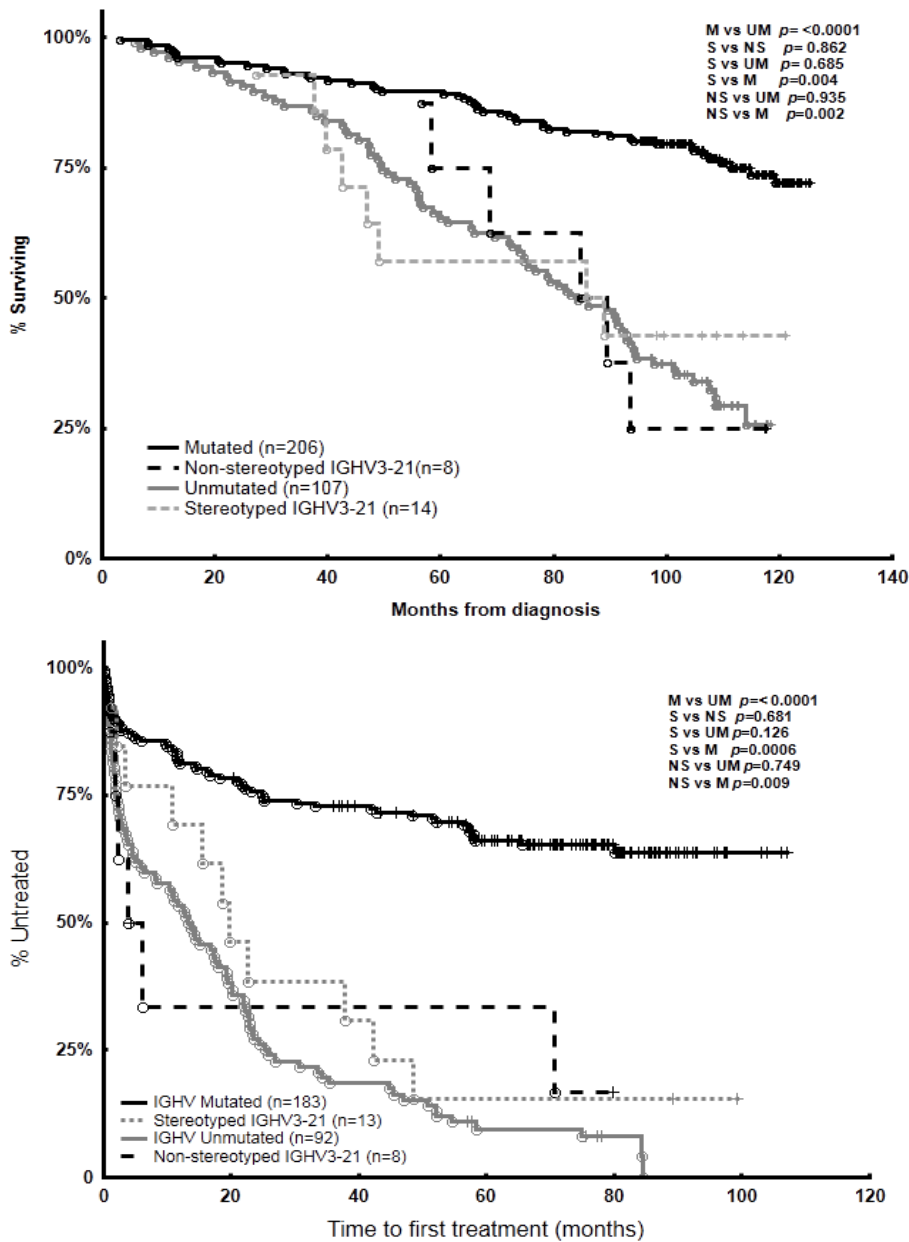
| IGHV3-21 cohort | Current Swedish population-based study | Hospital-based Scandinavian study <sup>4</sup> | Multi-centre Italian study <sup>151</sup> | Multi-centre Mediterranean Study <sup>5</sup> | USA Study <sup>152</sup> |
|-----------------|--|--|---|---|--------------------------|
| IGHV3-21        | <b>6.5 %</b>                           | <b>10.1 %</b>                                  | <b>4.1 %</b>                              | <b>3.5 %</b>                                  | <b>2.6%</b>              |

Other explanations for disparity in global IGHV3-21 frequency may be explained by differences in genetic background, influencing the expansion of the IGHV3-21 precursor CLL cells. Interestingly, Iranian CLL hospital based studies have shown no IGHV3-21 usage.<sup>362</sup> Alternatively, it is plausible that the IGHV3-21 geographic bias may reflect differences in the patient environment and the type of antigen exposure encountered. Given that we find a similar proportion of stereotyped cases in this report (14/22, 63.6%) as in other western CLL cohorts from the Mediterranean area (16/32, 50%),<sup>5</sup> Italy (32/58, 55.2%)<sup>151</sup> and the US (25/63, 39.9%),<sup>152</sup> the notion of a native antigen primarily encountered in Sweden impacting on IGHV3-21 gene frequency and leukemogenesis is less reasonable. On the other hand, non-stereotyped IGHV3-21 BCRs, given their heterogeneous molecular structure, may enable the binding of diverse antigens during leukemogenesis.

This study, like many previous, has established IGHV3-21 patients to experience a very poor overall survival (Figure 4.1 A).<sup>137,147,149-151</sup> However, the clinical impact of IGHV3-21 stereotypy is less clear. Given the suggestion of specific antigen involvement in stereotyped IGHV3-21 CLL and since IGHV3-21 stereotypy has been affiliated to disease progression,<sup>5,149,151</sup> it was of particular interest to determine its clinical impact on the disease. In a large collaborative European study incorporating 90 IGHV3-21 cases collected mainly from Sweden and Germany, no differences in survival were encountered between stereotyped IGHV3-21 patients and their non-stereotyped counterparts.<sup>150</sup> In support of these findings, studies from the US and Italy encompassing 63 and 37 IGHV3-21 cases respectively corroborated the disassociation between survival and stereotypy. As mentioned, smaller studies from the Mediterranean (n=16 and n=32) have shown IGHV3-21 cases with stereotyped receptors to undergo a shorter time to progression relative to non-stereotyped cases.<sup>5,149</sup> Perhaps these findings imply that the nature of the antigen may influence patient clinical behavior. That said, our study could not corroborate this finding, since no difference in time to treatment was seen between stereotyped and non-stereotyped IGHV3-21 patients (p=0.68) (Figure 4.1 A).

Again, the possible reason for this disparity between our study and others is perhaps explained by differences in cohort type and size. Alternatively, biological reasons such as a higher expression of the negative prognosticator CD38 and the higher occurrence of del(11q) documented in stereotyped relative to non-stereotyped IGHV3-21 cases may partially explain the poor disease course seen in these earlier studies.<sup>147</sup> Although this study showed a trend towards having a higher frequency of del(11q) in stereotyped IGHV3-21 patients, a higher expression of CD38 was not

observed in these cases. Nevertheless, other studies have implicated a high WNT-16 expression and expression of alternative negative indicators ZAP70, CD49d, and CD49b to contribute to the poor prognosis indicated in stereotyped IGHV3-21 patients.<sup>147</sup>



**Figure 4.1 A.** Kaplan Meier curve showing overall survival and time to treatment in IGHV3-21 patients.

In summary, using a population-based cohort we reconfirm the geographic bias in IGHV3-21 gene frequency previously identified in Scandinavia. Moreover, we determined that IGHV3-21 patients experience an inferior outcome independently of stereotypy. Regarding the prognostic role of stereotypy in IGHV3-21 CLL, further large collaborative studies are needed before a definite conclusion can be made since a large heterogeneity among published studies exists.

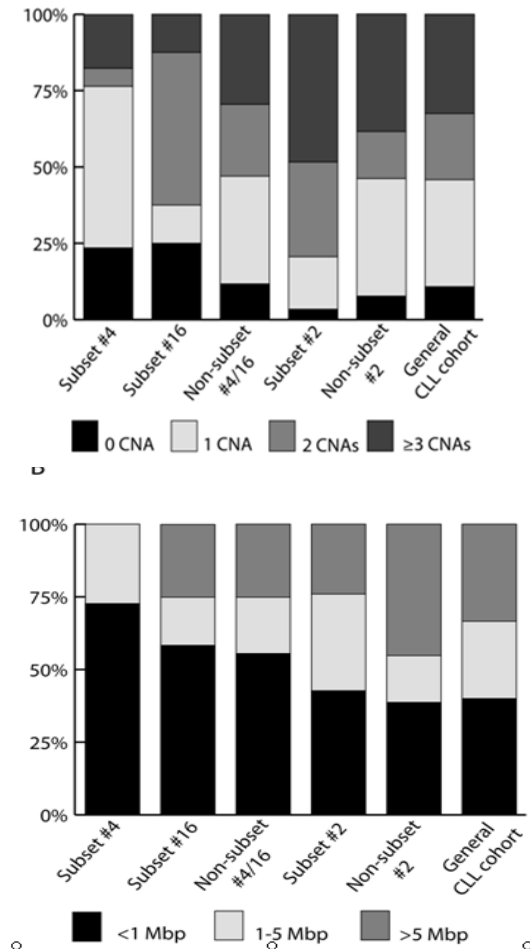
## **4.2: Paper II: High density array screening for genomic events in stereotyped subsets of CLL**

Although the advancement of new prognostic markers has allowed improved subdivision of CLL, the causes of CLL still remain unknown and there is yet no single genetic event indicated in all cases. Since the existence of multiple subsets of CLL expressing ‘stereotyped’ BCRs implies a role for antigen(s) in CLL,<sup>5,6</sup> and considering ‘stereotypy’ may influence CLL clinical course, it was of interest to determine if biases existed in the genomic profiles of stereotyped patients. As mentioned, both the stereotyped IGHV3-21 subset #2 and IGHV3-21 non-stereotyped counterparts are associated with a poor overall survival.<sup>137,150,152</sup> Moreover, some studies report subset #2 to be associated with a more progressive disease<sup>5,6,149</sup> relative to non stereotyped IGHV3-21, a finding we did not find evidence of among our own cohort of Swedish stereotyped IGHV3-21 patients in Paper I. On the other hand, stereotyped IGHV4-34 subset #4 is associated with an indolent disease course compared to its non-stereotyped counterparts.<sup>5,6</sup> Regarding IGHV4-34 subset #16, little is known about this less commonly encountered subset. Thus, the aim of Paper

It was to perform a detailed genetic characterization of clinically divergent stereotyped CLL subsets using high-resolution SNP arrays, to identify genetic events that may give clues to their pathogenesis and diverse clinical outcomes. In particular this study focused on the comparison of poor prognostic IGHV3-21 subset #2 and favourable prognostic IGHV4-34 subset #4 patients. In an attempt to identify stereotyped specific events, the genetic profiles of stereotyped subsets were compared to their non-stereotyped counterparts and a general CLL cohort.

SNP analysis revealed a higher number of poor prognostic IGHV3-21 samples to carry alterations, i.e. 97% of subset #2 and 92% of non-subset #2 samples. In contrast, the proportion of subset #4 samples bearing alterations was considerably less (76%). Similarly, 75% of subset #16 samples carried aberrations, whereas 88% of the more clinically heterogeneous non-subset 4/16 cases displayed alterations. Besides the marked differences in the overall frequencies of samples affected with CNAs, differences in the genomic complexity regarding the number, type and size of alterations encountered were noted between the study groups. Overall, subset #2 and non-subset #2 displayed a higher average number of aberrations compared to subset #4. More specifically, 48% of subset #2 and 39% of non-subset #2 samples carried  $\geq 3$  alterations compared to 18% of subset #4 and 13% of subset #16 (Figure 4.2 A). In fact, the majority (53%) of subset #4 patients carried only one alteration. The increased genome complexity observed for IGHV3-21 patients, particularly for subset #2, can be partially accounted for by the high frequency of the known recurrent alterations observed. Interestingly, known recurrent alterations were detected in 90% of subset #2 cases compared to, 54% of non-subset #2 patients. More noticeably, a low frequency of the known recurrent alterations was seen among the favourable

prognostic subset #4 patients (35%). Similarly, a low frequency was detected in subset #16 (25%). However, IGHV4-34 non-subset 4/16 patients displayed a similar frequency of known recurrent alterations compared to IGHV3-21 patients (59%).



**Figure 4.2 A.** Level of genomic complexity (CNA number and size).

Notably, poor prognostic deletions of 11q were more frequent in subset #2 and non-subset #2 (31% and 23%) relative to subset #4 and IGHV4-34 non-subset #4/16 patients (0% and 12%). Furthermore, deletions of 13q were particularly high in subset #2 cases and concurrent alterations at 11q and 13q were also found to be common in these patients. In contrast, deletion of the good prognostic marker del(13q) was the

only known recurrent aberration detected in subset #4 (35%) (Table 4.2 A). On considering the larger number of alterations and the high frequency of poor genomic markers carried by subset #2 in comparison to subset #4 patients, these findings suggest that subset #2 are more likely to acquire genomic events relative to subset #4. These events may represent important genetic events during the pathogenesis of IGHV3-21 CLL particularly when considering the frequent finding of concurrent 11q and 13q deletions, especially in subset #2 cases (Table 4.2 A). Hypothetically, these aberrations may be acquired during the phase of active stimulation by (unknown) antigens, which trigger the IGHV3-21 precursor cell to undergo rapid cell division and hence, under such proliferative pressure acquire such clonal alterations. In addition, it is interesting to speculate that these observations may account for the adverse clinical outcome reported for IGHV3-21 CLL, particularly for subset #2 patients.

**Table 4.2 A.** Known genomic alterations in IGHV3-21 and IGHV4-34 subsets compared to a general CLL cohort.

| <b>Subset/<br/>non-subset</b> | <b>del(11q)</b><br>(poor<br>prognosis) | <b>del(13q)</b><br>(good<br>prognosis) | <b>trisomy 12</b><br>(intermediate<br>prognosis) | <b>del(17p)</b><br>(poor<br>prognosis) |
|-------------------------------|--|--|--|--|
| Subset #2                     | <b>31%</b>                             | 79%                                    | 0  | 0                                      |
| Non-subset #2                 | <b>23%</b>                             | 31%                                    | 23%  | 7.7%                                   |
| Subset #4                     | <b>0</b>                               | <b>35%</b>                             | <b>0</b>   | <b>0</b>                               |
| Subset #16                    | 13%                                    | 0                                      | 13%  | 0                                      |
| Non-subset #4/16              | 12%                                    | 44%                                    | 6%   | 3%                                     |
| General CLL cohort            | 13%                                    | 54%                                    | 11%  | 4%                                     |



Recent reports suggest that not only the type of alteration but the size of the alteration and the nature of their breakpoints influence clinical outcome. In this study, alteration size accounted for another layer of genomic complexity. Size distribution analysis showed subset #2 and non-subset #2 patients to have a higher number of alterations greater than 1Mb in size compared to the other subgroups, particularly subset #4. This can be accounted for by the high frequency of large known recurrent alterations seen in IGHV3-21 patients, particularly in subset #2. However, further investigations revealed that both subset #2 and non-subset #2 have similar average number of CNAs/sample comparable to the level observed in our recent array-based study on Swedish CLL.<sup>183</sup> Hence, subset #2 and non-subset #2 cases do not appear to be more complex when accounting for all genomic alterations compared to other CLL patients in general. Their crux appears to come from their susceptibility to acquire known recurrent alterations and larger sized alterations, which might be reflective of their more active proliferative status.

On the other hand, subset #4, #16 and non-subset #4/16 showed a higher prevalence of smaller alterations <1Mbp, owing to the lack of large recurrent alterations such as del(11q), trisomy 12 and del(17p) in these groups. Additionally, 13q deletions in subset #4 were shown to be of a smaller average size compared to those encountered in the IGHV3-21 groups. Overall, on considering the lower number and size of alterations in subset #4 and the lack of poor prognostic alterations compared to subset #2, these features suggest that subset #4 patients are less likely to accumulate genomic alterations. This is suspected to be due to a low proliferative drive stemming from the IGHV4-34 BCR. Altogether, the lack of such genomic aberrations may in

part be responsible for the indolent disease course experienced by subset #4 CLL patients.

Novel events were identified on chromosomes 2q, 7q and 14q in IGHV4-34 patients; however these were not attributed to any particular subset. Recurrent gains at 2q and losses at 3p were demonstrated in two subset #2 cases. Recurrent CNN-LOH events were mainly detected on chromosome 13q, although independently of stereotypy status.

This study concludes that genomic aberrations were more common in subset #2 and non-subset #2 compared to subset #4. The particularly high frequency of del(11q) in subset #2 may be linked to the adverse outcome reported for these patients. In contrast, the lower prevalence and size of CNAs and the absence of poor-prognostic aberrations in subset #4 may reflect the indolent disease course seen in these patients.

### **4.3: Paper III: Array-based methylation profiling in prognostic subsets of CLL**

Aberrant DNA methylation is a key component in augmenting tumorigenesis, where genome-wide hypomethylation and regional hypermethylation of tumor suppressor gene (TSG) promoters are signature characteristics of many cancers.<sup>286</sup> As mentioned, in CLL, the candidate gene approach has revealed a significant association linking promoter methylation to transcriptional silencing for a small number of individual gene promoters, namely *DAPK1*,<sup>309</sup> *ZAP70*<sup>308</sup> and *TWIST-2*.<sup>307</sup> Since the global methylation profiles of CLL were yet to be discovered, Paper III

sought to investigate genome wide DNA methylation patterns in three divergent CLL prognostic subgroups; IGHV mutated, unmutated and IGHV3-21 patients. Considering the clinical and biological diversity of these subgroups, Paper III set out to perform a detailed genome wide, albeit somewhat promoter biased, epigenetic characterization using the Illumina 27K high-resolution DNA methylation microarray, as an important means to identify differentially and aberrantly methylated genes that may play a role in their pathogenesis and diverse clinical outcome. Moreover, as methylation changes are reversible with the use of de-methylating agents, whole genome scanning may help identify new therapeutic targets in CLL. Hence, Paper III investigated the significance of DNA methylation in regulating gene promoters by re-inducing the expression of methylated TSGs using the methyl-inhibitor 5-aza-2'-deoxycytidine and the histone deacetylase inhibitor TSA.

Corroborating other studies, Paper III demonstrated that CLL is hypomethylated compared to normal healthy individuals on the global level.<sup>304-306</sup> Furthermore, by applying stringent bioinformatic approaches we identified for the first time distinct differential methylation profiles in divergent CLL prognostic subgroups. A number of significantly differentially methylated genes occupying key cellular pathways and functions were revealed between the studied subgroups. For instance, 7 TSGs were preferentially methylated and potentially silenced in the poor prognostic IGHV unmutated subgroup (e.g. *VHL*, *ABI3*) (Table 4.3 A).<sup>363,364</sup> In contrast, these genes remained unmethylated and hence potentially expressed in the favourable prognostic IGHV mutated CLL. That said, the IGHV mutated subgroup did display methylation of 2 TSGs, *PPP1R3A*<sup>365</sup> and *WISP3*<sup>366</sup> that were unmethylated in IGHV unmutated patients (Table 4.3 A). Similarly, most methylated TSGs found in the IGHV

unmutated subgroup were found to be unmethylated in IGHV3-21 patients, one exception was the preferential methylation of the *SLC22A18* gene.<sup>367</sup> These observations suggest that the specific methylation patterns of certain TSGs may play a role in the pathogenic pathways that lead to these clinically divergent subgroups. That notwithstanding, many of the TSGs identified in this study have been implicated in several other cancers, however, their role in CLL has yet to be determined (Table 4.3 A).

**Table 4.3 A.** A brief summary of the differentially methylated genes identified between IGHV mutated and IGHV unmutated CLL

| Gene function              | IGHV Mutated-CLL | IGHV Unmutated-CLL | Examples  |
|----------------------------|------------------|--------------------|---|
| Aiding tumourigenesis      | Methylated       | Unmethylated       | <i>ADORA3</i> <sup>368</sup><br><i>ANGPT2</i> <sup>369</sup>                          |
| Suppressing tumourigenesis | Unmethylated     | Methylated         | <i>VHL</i> <sup>363</sup> , <i>ABI3</i> <sup>364</sup><br><i>LCOD1</i> <sup>370</sup> |
| Anti-apoptotic             | Methylated       | Unmethylated       | <i>BCL2</i> <sup>87</sup>   |
| Pro-apoptotic genes        | Unmethylated     | Methylated         | <i>BCL10</i> <sup>371</sup>   |

Certain cellular pathways, such as NFκB, PI3K/Akt and MEK/ERK, are dysregulated in CLL B cells leading to activation of anti-apoptotic pathways.<sup>116,117,120,372,373</sup> Our data strengthens this idea and further indicates that certain genes involved in such ‘proliferative’ pathways may be regulated by DNA methylation in CLL. For example, genes activating the NFκB pathway, such as *ADORA3*<sup>368</sup> and *CARD15*,<sup>374</sup> were specifically unmethylated in the unmutated CLL subgroup. Similarly, the *LOC340061*<sup>375</sup> and *PRF1*<sup>376</sup> genes, which activate the NFκB and MAPKinase pathways, respectively, were unmethylated and therefore potentially expressed in

poor-prognostic IGHV3-21 CLL compared to IGHV mutated CLL. In line with these findings, genes repressing the NFκB pathway (*LCOD1* gene)<sup>370</sup> and MAPKinase (*ZNF540* gene)<sup>377</sup> pathways were generally methylated in both IGHV unmutated and IGHV3-21 samples in contrast to IGHV mutated CLL. Again most genes have not been associated with CLL, while many are known to be involved in the tumourigenesis of several other cancers and leukaemias.

Genes known to facilitate tumourigenesis within other cancer types, such as *IFNBI*<sup>378</sup> and *IL17RC*,<sup>379</sup> were noted to be preferentially unmethylated in the IGHV unmutated and IGHV3-21 CLL subgroups in comparison to IGHV mutated CLL. Similarly the *LMO2* oncogene<sup>380</sup> and *RASGRP3*<sup>381</sup> which regulate B cell proliferation by facilitating B-cell receptor-Ras signaling, were unmethylated in IGHV3-21 samples. Thus, the above data indicates that poor-prognostic CLL subsets favor expression of genes aiding tumor proliferation.

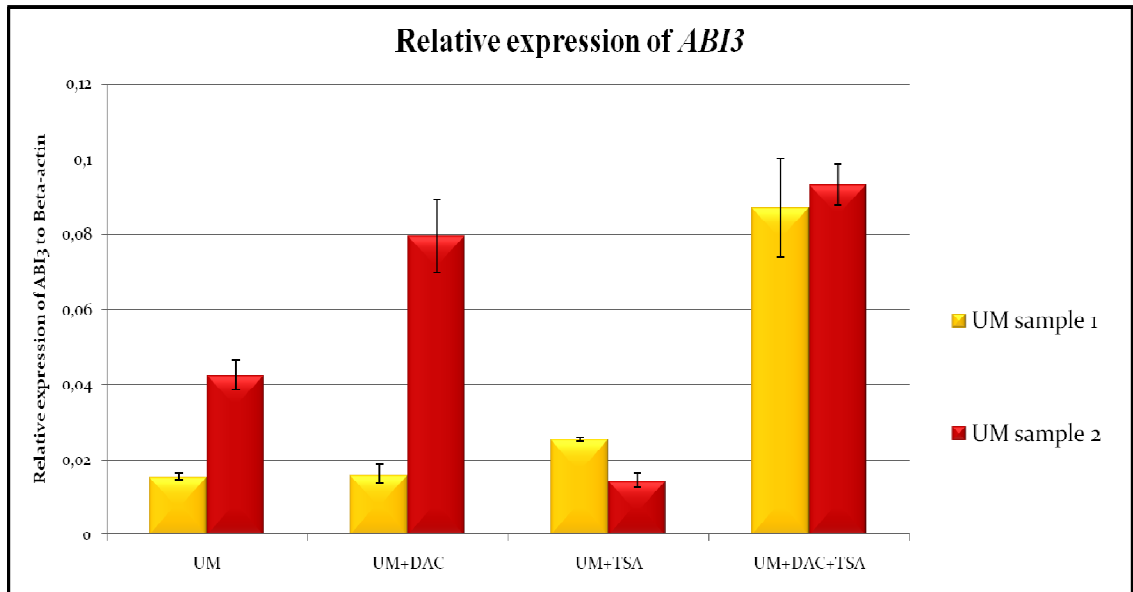
Interestingly, array analysis provided insights into the deregulation of anti- and pro-apoptotic genes in IGHV mutated CLL. For instance, *PLDI*, a known anti-apoptotic gene,<sup>382</sup> remained unmethylated in IGHV mutated CLL samples. On the other hand, pro-apoptotic genes such as *BCL10*<sup>371</sup> and *TP53I3*<sup>383</sup>, were methylated in IGHV mutated samples unlike IGHV unmutated CLL.

In short, genes related to prognosis, for instance the *ANGPT2*<sup>369</sup> gene, which has been implicated in progressive CLL disease states were shown to be preferentially methylated in the favourable prognostic mutated CLL cases while being unmethylated in unmutated CLL (Table 4.3 A). In addition, the *NGFR* gene which has been associated with a favorable prognosis in ALL,<sup>384</sup> was shown to be

unmethylated in IGHV mutated CLL and showed increased expression on the RNA level within this subset.

Independent verification of array results was performed for a select number of genes of different functionalities using an independent CLL sample set. MSP-PCR semi-quantitatively verified the methylation status of four methylated genes (*PRF1*, *ADORA3*, *IGSF4* and *BCL10*), RQ-PCR demonstrated the correlation between gene expression and methylation status for 4 genes (eg. *VHL*, *ABI3* and *NGFR*) in an independent CLL cohort and finally, bi-sulfite sequencing further quantified the degree of methylation for two genes (*ABI3* and *VHL*).

Finally, through the inhibition of DNA methylation using methyl- and HDAC inhibitors, this study evidenced the induced expression of methylated TSGs (e.g. *IGSF4*, *ABI3* and *VHL*) in unmutated CLL primary samples. In general, although a low induction of gene expression was observed with methyl inhibitor treatment only, a more pronounced effect was seen when the methyl inhibitor and histone acetyl inhibitor drugs were applied in combination. That said, the low induction can be partially explained by the fact that CLL primary cells do not proliferate at a high rate under *in vitro* conditions. These data indeed strengthen the role of DNA methylation in regulating promoter expression and open up the possibility of targeted therapy in CLL (Figure 4.3 A).



**Figure 4.3 A.** Re-expression of DNA methylated tumour suppressor gene *ABI3* using methyl inhibitor DAC (5-aza-2' deoxycytidine) and histone acetyl inhibitor Tricostatin (TSA).

In conclusion, Paper III shows distinct methylation profiles encompassing a number of candidate genes involved in cellular pathways regulating proliferation and apoptosis in IGHV unmutated/mutated and IGHV3-21 CLL and thus underlines the critical role of epigenetic changes during leukemogenesis.

#### **4.4: Paper IV: DNA methylation screening of CLL samples over time and within different CLL compartments**

On identifying differential methylation profiles in IGHV mutated and unmutated CLL in Paper III, we set out to further investigate their methylation profiles using a higher resolution 450K global DNA methylation microarrays. More specifically, we aimed to determine whether the DNA methylation profiles of these subgroups changed over time during the course of disease and whether DNA methylation of CLL cells differed according to the anatomical site from which the cells resided. The first part of this study, in accordance with Paper III, again identified IGHV mutated and

unmutated CLL to have differential DNA methylation profiles, albeit on a larger scale than previously reported. Specifically we found 2239 differentially methylated sites between these subgroups compared to a mere 99 sites identified previously. We found a relatively even distribution of differentially methylated sites across the whole genome. However, we noted that the largest proportion of differentially methylated sites was located outside of the CpG islands and within gene bodies. Only recently with the advent of whole genome analysis has the relevance of this latter finding come to light. Traditionally DNA methylation analysis has solely focused on CpG islands, particularly within gene promoters, since it is here that gene regulation is primarily regulated. In Paper III, we identified DNA methylation of *ADORA3*, positioned approximately 1kb away from the promoter, to have a regulatory function on gene expression. On a much larger scale Irizarry *et al*<sup>295</sup> further corroborates this finding, showing the majority of altered DNA methylation sites in colon cancer to reside outside of CpG islands in regions up to 2kb distant from the CpG island in regions called CpG island shores. Moreover, these sites were shown to correlate to gene expression.

Gene ontology analysis identified a number of enriched physiological and molecular functionalities, most of which have not been implicated in CLL before. Many genes were known as members of established signaling pathways, whereas others were shown to be involved in other cancers. For instance, CpG sites within DNA repair genes, DNA methyltransferase genes *DNMT3B*,<sup>264</sup> members of the polycomb complex *EED*,<sup>385</sup> *JAIRD2*<sup>386</sup> and polycomb target genes were detected as differentially methylated. Specifically 230 CpG sites covering 158 polycomb target genes previously characterized in embryonic fibroblasts<sup>387</sup> presented as differentially



methyated between IGHV mutated and unmutated CLL. Approximately half of the sites presented with a higher methylation level in unmutated CLL, whereas the other half were found to have a higher level within the mutated subgroup. Preferential methylation of developmental polycomb target genes has been frequently recognised in other forms of cancer.<sup>388,389</sup> That said, the exact biological relevance of this finding is largely unexplored in cancer. What is known is that tumor-specific targeting of *de novo* methylation at these targets is pre-programmed by an established epigenetic system namely H3K27me3 that normally has a role in marking embryonic genes for repression. Unlike normal cells, the EZH2 polycomb complex maintaining H3K27me3 recruits DNA methyltransferase to these sites in cancer cells.<sup>389</sup>

A focused search for genes formerly implicated in CLL, found a number of survival related (*TCF3*,<sup>390</sup> *LEF1*<sup>391</sup>), proliferation (*CD80*,<sup>392</sup> *CD86*<sup>393</sup>) and apoptosis (*FAS*)<sup>394</sup> and BCR signaling related genes (*IBTK*).<sup>395</sup> Moreover, 12 differentially methylated genes identified in Paper III were found to overlap to this present study and were shown to be largely concordant in terms of methylation status. For example, the preferential methylation of tumour suppressor genes *ABI3*<sup>364</sup> in IGHV unmutated and *WISP3*<sup>366</sup> in IGHV mutated subgroups was re-confirmed. In support of Corcoran *et al*,<sup>308</sup> we found the CLL prognostic gene *ZAP70*,<sup>166</sup> to be preferentially methylated in IGHV mutated patients relative to unmutated cases. Furthermore, Corcoran *et al*,<sup>308</sup> have shown that methylation of *ZAP70* correlates well to *ZAP70* expression. Recently, methylation analysis of *ZAP70* intron 1 by pyrosequencing has been shown to be a robust method correlating to *ZAP70* and CD38 expression, IGHV mutation status, overall survival and time to treatment. Given the lack of sufficient standardisation with current flow cytometry quantitation methods, quantification by

pyrosequencing can provide an alternative mode of analysis.<sup>396</sup> For the first time, CLL prognostic genes *LPL*<sup>170,173</sup> and *CLLUI*<sup>177,178,394</sup> were noted to be differentially methylated between IGHV mutated and unmutated CLL. More specifically, higher DNA methylation within the gene body of *CLLUI* and the transcription start site TSS1500 and gene body region of *LPL* in IGHV mutated CLL relative to IGHV unmutated patients was demonstrated. In recent times, expression of both *LPL* and *CLLUI* has been noted to be higher in IGHV unmutated patients compared to IGHV mutated cases.<sup>177,178,267</sup> In CLL, high expression of these genes is associated with poor clinical outcome.<sup>170,173</sup> In view of these observations, it is interesting to speculate that perhaps DNA methylation of these genes may contribute to their differential expression seen in CLL.

One of the enriched canonical pathways highlighted through IPA analysis was the molecular mechanisms of cancer pathway. Within this IPA derived pathway, we identified pathways previously implicated in CLL namely the NF- $\kappa$ B/TNFR1 and TGF- $\beta$  pathways. Of particular interest is the TNF/NF- $\kappa$ B pathway. Here ten members of this pathway were found to be differentially methylated. For instance, a CLL autocrine growth factor *TNF*,<sup>397,398</sup> *TNFRSF8* (CD30), a receptor for thioredoxin (Trx)<sup>399</sup> a survival factor for CLL cells<sup>400,401</sup> and two TNF receptor genes, *TNFSF1A* and *TNFRSF1B*, recently shown to be redox-controlled by protein-disulfide isomerase (PDI) in complex with thioredoxin-1 (Trx1), were found to be differentially methylated in mutated and unmutated CLL. Although these findings are interesting, relating the role of DNA methylation to the functionalities of such a pathway in CLL is difficult. The variable incidence of multiple or single CpG sites across genes at variable locations along different gene regions (gene body, TSS1500

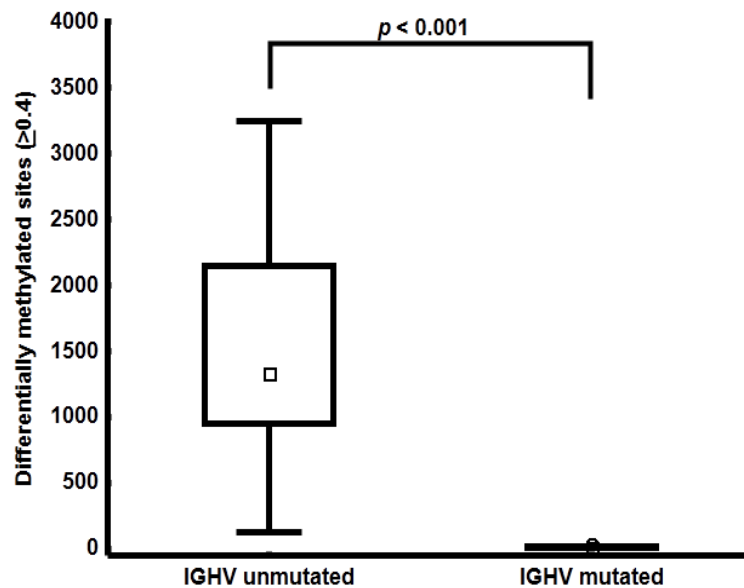
etc) makes interpretation challenging. Sometimes multiple CpG sites across the same gene at different regions will have variable methylation values. These challenges are further complicated by the fact that the biologically relevant level of DNA methylation needed to effect gene expression is currently unknown.

With the prospect of aberrant DNA methylation marks as candidate prognosticators, it is essential that such marks be stable in terms of DNA methylation level. To date, a limited number of studies exist documenting the stability of DNA methylation over time. Of the few studies available variable results ensue. This is perhaps in part reflective of the different diseases and nature of the cohort studied in these reports. However, more pressing inconsistencies stem from the alternative techniques employed and differences in the processing and interpretation of DNA methylation data. For example, a report examining bladder cancer metachronous tumours from two individuals, has shown DNA methylation to be relatively stable over time both within the individual and between patients.<sup>329</sup> In contrast, a study by Bjornsson *et al*<sup>328</sup> has identified inter-individual age variation and differences occurring over time within individuals, that would otherwise be over-looked by group averaging of the data. These changes were shown to be clustered within family members highlighting the possibility that genetic background may effect DNA methylation changes.<sup>324,328</sup> Nevertheless, a study of DNA methylation in 40 monozygotic twins found 35% of twin couples to have a discordant level of total DNA methylation, particularly those over the age of 28 years.<sup>330</sup> In this instance, different environmental influences and exposures were thought to play a role.

In light of the above findings we dedicated the second part of the study to examine DNA methylation over time in paired CLL diagnostic and follow-up patients. Based on average DNA methylation values across the groups, this study for the first time established that global DNA methylation is relatively stable over time. Both IGHV mutated and unmutated patients incurred only a few, small recurrent changes over time. Interestingly however, on analyzing the intra-individual changes a larger difference in the number of non recurrent differentially methylated sites were noted over time between IGHV mutated and unmutated cases, that would otherwise be missed on group averaging of the data. More specifically, intra-individual analysis of IGHV unmutated patients showed a large difference whereas intra-individual investigation of IGHV mutated patients showed no significant difference in the number of non-recurrent differentially methylated sites over time (Figure 4.4 A). That said, in global terms, these non-reccurent changes in IGHV-umutated cases equated to a very small overall global change (<1%), when taking into account the 485,000 sites available on the array. Nevertheless, it is interesting to speculate that perhaps these findings are reflective of the more indolent and aggressive disease course associated with IGHV mutated and unmutated patients respectively. Notably, global DNA methylation within IGHV unmutated patients over time was relatively stable despite the implementation of treatment between diagnosis and follow-up. Thus it appears that DNA methylation of CLL cells is not largely effected by therapy, at least not in respect to overall DNA methylation.

It is difficult to ascertain whether or not some of these non-recurrent DNA methylation changes act as driver epi-mutations causing the expansion of the CLL clone. A more likely explanation, given the lack of recurrent changes over time,

indicates that the changes encountered merely act as passenger events arising from the higher proliferative nature in the preexisting unmutated clone. Nevertheless, it is interesting to speculate that perhaps these changes may contribute to CLL leukemogenesis by contributing to overall epigenomic instability.



**Figure 4.4.** Box plot visualizing the difference in number of non recurrent differentially methylated sites overtime within and between patient matched diagnostic and follow-up IGHV mutated and IGHV unmutated CLL patients.

Resting PB CLL cells are known to communicate with a pro-survival microenvironment in proliferation centers of the LN resulting in proliferation (~1%/day) of the CLL clone.<sup>89</sup> Recently, a higher expression of genes involved in proliferation and active BCR signaling were shown in the LN relative to the blood of patient-matched CLL samples.<sup>118</sup> In light of the anti-correlation of DNA methylation to gene expression, we hypothesized that DNA methylation may partake in the mechanism of differential gene expression seen within the different compartments of

CLL. Furthermore, in mind of the possibility of aberrant DNA methylation marks as candidate prognostic markers, it is important to establish whether such marks are stable over different sampling sites. Therefore, in the third part of the study, we examined the DNA methylation profiles of patient-matched CLL cells derived from the PB and LN, to ascertain the influence of the microenvironment on CLL patient DNA methylation profiles. This study for the first time found that global DNA methylation is rather similar in patient matched PB and LN samples. Only few, small recurrent changes were identified. Moreover, the number of non-recurrent changes although relatively low in global terms (<1%), were found to vary in number across the intra-individual PB and LN pairs. Therefore it seems that unlike gene expression, DNA methylation in CLL cells is not heavily influenced by the microenvironment.

In sum, on a larger scale to Paper III, this study reconfirms IGHV mutated and IGHV unmutated CLL to have differential methylation profiles. More specifically, CLL prognostic genes, polycomb target genes, and key members of TGF- $\beta$  and NF- $\kappa$ B/TNFR1 pathways were noted to be differentially methylated. Finally, global DNA methylation was shown to be relatively stable over time and to be remarkably similar in CLL cells derived from different anatomical CLL compartments.

## **5: CONCLUSIONS**

This thesis applied IGHV gene sequencing and high resolution SNP and DNA methylation microarray techniques in order to interrogate genetic and methylation events that may contribute to CLL leukemogenesis in diverse prognostic subgroups and subsets of CLL.

### **5.1. Paper I**

- Within the Swedish population-based cohort studied, the geographic bias in IGHV3-21 gene frequency previously identified in Sweden was validated, albeit at a lower frequency than formerly reported in our hospital- based cohort study.
- IGHV3-21 patients were shown to have an inferior outcome independently of stereotypy. Hence, no prognostic value could be attributed to IGHV3-21 stereotypy status.

### **5.2: Paper II**

- Poor prognostic IGHV3-21 subset #2 patients commonly carried more genomic aberrations and larger sized alterations compared to favourable prognostic IGHV4-34 subset #4 patients. Thus, it appears that subset #2 patients are more genomically complex relative to subset #4.

- IGHV4-34 subset #4 patients exclusively carried the favourable prognostic del(13q) marker, whereas poor prognostic subset #2 patients carried a high frequency of the poor prognostic del(11q) marker. Hence, these findings may in part be reflective of the favourable prognosis seen in subset #4 and the adverse outcome seen in subset #2 patients.

### 5.3. Paper III

- Distinct differential methylation profiles were detected in three divergent prognostic subsets of CLL, IGHV mutated, unmutated and IGHV3-21. Differentially methylated genes, such as TSGs, were preferentially methylated in the poor prognostic IGHV unmutated subgroup compared to the favourable prognostic IGHV mutated CLL. Moreover, genes involved in cell proliferation were specifically unmethylated in the poor prognostic IGHV unmutated and IGHV3-21 subgroups. Hence, these findings suggest that the methylation status of certain key cellular genes may contribute to CLL pathogenesis and the divergent clinical outcome seen in certain CLL subsets.
- Through the use of methyl inhibitor and histone deacetylation drugs, the re-expression of methylated TSGs (*ABI3*, *VHL*) was demonstrated. Hence the reversible nature of DNA methylation was functionally established.



## 5.4: Paper IV

- On a larger scale to Paper III, this study reconfirmed that IGHV mutated and IGHV unmutated CLL are characterized by differential methylation profiles. Differentially methylated sites between these subgroups primarily presented outside of CpG islands. Preferential DNA methylation of CpG sites within known CLL prognostic genes *CLLUI* and *LPL* was notably higher in IGHV mutated cases relative to IGHV unmutated patients. Key members of TGF- $\beta$  and NF- $\kappa$ B/TNFR1 pathways and large numbers of polycomb targets were also shown to be differentially methylated between IGHV mutated and unmutated CLL.
- Global DNA methylation in paired diagnostic and follow-up IGHV-UM (treated) CLL and IGHV-M (untreated) CLL patients was deemed to be relatively stable over time. Hence, this finding supports the notion that aberrant DNA methylation events are early events in CLL pathogenesis.
- DNA methylation changes in patient-matched peripheral blood and lymph node samples were found to be relatively similar. Hence, it seems the differences in the PB and LN micro-environment does not heavily influence the DNA methylation profiles of CLL cells.

Taken together, these studies provide a greater insight into the genetic alterations and DNA methylation events occurring in divergent CLL prognostic subgroups.

## 6: REFERENCES

1. Rai KR PD. Chronic lymphocytic leukemia. In Hematology: Basic Principles and Practice (ed 2nd ed). New York: Churchill Livingstone; 1995.
2. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999;94:1840-1847.
3. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999;94:1848-1854.
4. Tobin G, Thunberg U, Karlsson K, et al. Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. *Blood*. 2004;104:2879-2885.
5. Stamatopoulos K, Belessi C, Moreno C, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood*. 2007;109:259-270.
6. Murray F, Darzentas N, Hadzidimitriou A, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*. 2008;111:1524-1533.
7. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000;343:1910-1916.
8. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin*. 2009;59:225-249.
9. Redaelli A, Laskin BL, Stephens JM, Botteman MF, Pashos CL. The clinical and epidemiological burden of chronic lymphocytic leukaemia. *Eur J Cancer Care (Engl)*. 2004;13:279-287.
10. Dighiero G, Hamblin TJ. Chronic lymphocytic leukaemia. *Lancet*. 2008;371:1017-1029.
11. Swedish lymphoma registry, report 2000-2006  
URL:<http://www.ocsyd.se/VP-verksamhet/Kvalitetsreg/svenskalymfomregistretrapport2000-2006.pdf>.
12. Houlston RS, Sellick G, Yuille M, Matutes E, Catovsky D. Causation of chronic lymphocytic leukemia--insights from familial disease. *Leuk Res*. 2003;27:871-876.
13. Goldin LR, Pfeiffer RM, Li X, Hemminki K. Familial risk of lymphoproliferative tumors in families of patients with chronic lymphocytic leukemia: results from the Swedish Family-Cancer Database. *Blood*. 2004;104:1850-1854.
14. Crowther-Swanepoel D, Houlston RS. The molecular basis of familial chronic lymphocytic leukemia. *Haematologica*. 2009;94:606-609.
15. Goldin LR, Slager SL, Caporaso NE. Familial chronic lymphocytic leukemia. *Curr Opin Hematol*. 2010;17:350-355.

16. Sellick GS, Goldin LR, Wild RW, et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. *Blood*. 2007;110:3326-3333.
17. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet*. 2008;40:1204-1210.
18. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nat Genet*. 2010;42:132-136.
19. Rawstron AC. Prevalence and characteristics of monoclonal B-cell lymphocytosis (MBL) in healthy individuals and the relationship with clinical disease. *J Biol Regul Homeost Agents*. 2004;18:155-160.
20. Rawstron AC, Bennett FL, O'Connor SJ, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med*. 2008;359:575-583.
21. Rawstron AC, Green MJ, Kuzmicki A, et al. Monoclonal B lymphocytes with the characteristics of "indolent" chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts. *Blood*. 2002;100:635-639.
22. Rawstron AC, Yuille MR, Fuller J, et al. Inherited predisposition to CLL is detectable as subclinical monoclonal B-lymphocyte expansion. *Blood*. 2002;100:2289-2290.
23. Blair A, Purdue MP, Weisenburger DD, Baris D. Chemical exposures and risk of chronic lymphocytic leukaemia. *Br J Haematol*. 2007;139:753-761.
24. Khalade A, Jaakkola MS, Pukkala E, Jaakkola JJ. Exposure to benzene at work and the risk of leukemia: a systematic review and meta-analysis. *Environ Health*. 2010;9:31.
25. Chiorazzi N, Ferrarini M. B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. *Annu Rev Immunol*. 2003;21:841-894.
26. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med*. 2005;352:804-815.
27. Chiorazzi N, Ferrarini M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood*. 2011;117:1781-1791.
28. Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood*. 2008;111:5446-5456.
29. Nowakowski GS, Hoyer JD, Shanafelt TD, et al. Using smudge cells on routine blood smears to predict clinical outcome in chronic lymphocytic leukemia: a universally available prognostic test. *Mayo Clin Proc*. 2007;82:449-453.
30. Johansson P, Eisele L, Klein-Hitpass L, et al. Percentage of smudge cells determined on routine blood smears is a novel prognostic factor in chronic lymphocytic leukemia. *Leuk Res*. 2010;34:892-898.
31. Matutes E, Owusu-Ankomah K, Morilla R, et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia*. 1994;8:1640-1645.
32. Zwiebel JA, Cheson BD. Chronic lymphocytic leukemia: staging and prognostic factors. *Semin Oncol*. 1998;25:42-59.

33. Kalil N, Cheson BD. Chronic lymphocytic leukemia. *Oncologist*. 1999;4:352-369.
34. Wadhwa PD, Morrison VA. Infectious complications of chronic lymphocytic leukemia. *Semin Oncol*. 2006;33:240-249.
35. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood*. 1975;46:219-234.
36. Binet JL, Lepoprier M, Dighiero G, et al. A clinical staging system for chronic lymphocytic leukemia: prognostic significance. *Cancer*. 1977;40:855-864.
37. Shanafelt TD, Byrd JC, Call TG, Zent CS, Kay NE. Narrative review: initial management of newly diagnosed, early-stage chronic lymphocytic leukemia. *Ann Intern Med*. 2006;145:435-447.
38. Ansell SM, Li CY, Lloyd RV, Phyllyk RL. Epstein-Barr virus infection in Richter's transformation. *Am J Hematol*. 1999;60:99-104.
39. Tsimberidou AM, Keating MJ. Richter syndrome: biology, incidence, and therapeutic strategies. *Cancer*. 2005;103:216-228.
40. Tsimberidou AM, Keating MJ. Richter's transformation in chronic lymphocytic leukemia. *Semin Oncol*. 2006;33:250-256.
41. Eichhorst B, Hallek M, Dreyling M. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21 Suppl 5:v162-164.
42. Hussein MA, Gundacker H, Head DR, et al. Cyclophosphamide followed by fludarabine for untreated chronic lymphocytic leukemia: a phase II SWOG TRIAL 9706. *Leukemia*. 2005;19:1880-1886.
43. Onkologiskt centrum, Södra sjukhusregionen. 2008-2010. National guidelines for investigation and treatment of chronic lymphocytic leukaemia [http://www.oc.gu.se/pdf/KLL\\_Nationella\\_Riktlinjer\\_20060601.pdf](http://www.oc.gu.se/pdf/KLL_Nationella_Riktlinjer_20060601.pdf).
44. <http://www.cancer.gov/drugdictionary/?expand=All>.
45. Tam CS, Keating MJ. Chemoimmunotherapy of chronic lymphocytic leukemia. *Best Pract Res Clin Haematol*. 2007;20:479-498.
46. Tsimberidou AM, Tam C, Abruzzo LV, et al. Chemoimmunotherapy may overcome the adverse prognostic significance of 11q deletion in previously untreated patients with chronic lymphocytic leukemia. *Cancer*. 2009;115:373-380.
47. Osterborg A, Foa R, Bezares RF, et al. Management guidelines for the use of alemtuzumab in chronic lymphocytic leukemia. *Leukemia*. 2009;23:1980-1988.
48. Pettitt AR, Matutes E, Oscier D. Alemtuzumab in combination with high-dose methylprednisolone is a logical, feasible and highly active therapeutic regimen in chronic lymphocytic leukaemia patients with p53 defects. *Leukemia*. 2006;20:1441-1445.
49. Dreger P, Corradini P, Kimby E, et al. Indications for allogeneic stem cell transplantation in chronic lymphocytic leukemia: the EBMT transplant consensus. *Leukemia*. 2007;21:12-17.
50. Goldsby RA KT, Osborne BA, Kuby J. *Immunology* (ed Fifth). New York, NY: W.H Freeman and Company; 2003.
51. Blom B, Spits H. Development of human lymphoid cells. *Annu Rev Immunol*. 2006;24:287-320.
52. Kang J, Der SD. Cytokine functions in the formative stages of a lymphocyte's life. *Curr Opin Immunol*. 2004;16:180-190.

53. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. *Blood*. 2008;112:1570-1580.
54. Lefranc MPL G. *The Immunoglobulin Facts Book*. London, UK: Academic Press; 2001.
55. Lefranc MP, Giudicelli V, Ginestoux C, et al. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res*. 1999;27:209-212.
56. Tonegawa S. Somatic generation of antibody diversity. *Nature*. 1983;302:575-581.
57. Kurosawa Y, Tonegawa S. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J Exp Med*. 1982;155:201-218.
58. Lefranc MP, Giudicelli V, Kaas Q, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res*. 2005;33:D593-597.
59. Market E, Papavasiliou FN. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biol*. 2003;1:E16.
60. McBlane JF, van Gent DC, Ramsden DA, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell*. 1995;83:387-395.
61. van Gent DC, McBlane JF, Ramsden DA, Sadofsky MJ, Hesse JE, Gellert M. Initiation of V(D)J recombination in a cell-free system. *Cell*. 1995;81:925-934.
62. Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem*. 2002;71:101-132.
63. Akira S, Okazaki K, Sakano H. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science*. 1987;238:1134-1138.
64. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. *Cell*. 2002;109 Suppl:S45-55.
65. MacLennan IC. Germinal centers. *Annu Rev Immunol*. 1994;12:117-139.
66. Chien NC, Pollock RR, Desaynard C, Scharff MD. Point mutations cause the somatic diversification of IgM and IgG2a antiphosphorylcholine antibodies. *J Exp Med*. 1988;167:954-973.
67. Kleinstein SH, Louzoun Y, Shlomchik MJ. Estimating hypermutation rates from clonal tree data. *J Immunol*. 2003;171:4639-4649.
68. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M. Variability in the lambda light chain sequences of mouse antibody. *Nature*. 1970;228:1045-1047.
69. Clarke SH, Huppi K, Ruezinsky D, Staudt L, Gerhard W, Weigert M. Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. *J Exp Med*. 1985;161:687-704.
70. Hartley SB, Cooke MP, Fulcher DA, et al. Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. *Cell*. 1993;72:325-335.
71. Manser T. Textbook germinal centers? *J Immunol*. 2004;172:3369-3375.
72. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. *Immunity*. 2007;27:190-202.
73. de Vinuesa CG, Cook MC, Ball J, et al. Germinal centers without T cells. *J Exp Med*. 2000;191:485-494.
74. Dono M, Zupo S, Colombo M, et al. The human marginal zone B cell. *Ann N Y Acad Sci*. 2003;987:117-124.

75. Weill JC, Weller S, Reynaud CA. Human marginal zone B cells. *Annu Rev Immunol.* 2009;27:267-285.
76. Weller S, Faili A, Garcia C, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc Natl Acad Sci U S A.* 2001;98:1166-1170.
77. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem.* 2007;76:1-22.
78. Milstein C, Neuberger MS, Staden R. Both DNA strands of antibody genes are hypermutation targets. *Proc Natl Acad Sci U S A.* 1998;95:8791-8794.
79. Rogozin IB, Diaz M. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *J Immunol.* 2004;172:3382-3384.
80. Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta.* 1992;1171:11-18.
81. Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. *Nat Immunol.* 2001;2:530-536.
82. Shapiro GS, Wysocki LJ. DNA target motifs of somatic mutagenesis in antibody genes. *Crit Rev Immunol.* 2002;22:183-200.
83. Alfarano A, Indraccolo S, Circosta P, et al. An alternatively spliced form of CD79b gene may account for altered B-cell receptor expression in B-chronic lymphocytic leukemia. *Blood.* 1999;93:2327-2335.
84. Payelle-Brogard B, Magnac C, Alcover A, Roux P, Dighiero G. Defective assembly of the B-cell receptor chains accounts for its low expression in B-chronic lymphocytic leukaemia. *Br J Haematol.* 2002;118:976-985.
85. Vuillier F, Dumas G, Magnac C, et al. Lower levels of surface B-cell-receptor expression in chronic lymphocytic leukemia are associated with glycosylation and folding defects of the mu and CD79a chains. *Blood.* 2005;105:2933-2940.
86. Thompson AA, Talley JA, Do HN, et al. Aberrations of the B-cell receptor B29 (CD79b) gene in chronic lymphocytic leukemia. *Blood.* 1997;90:1387-1394.
87. Gottardi D, Alfarano A, De Leo AM, et al. In leukaemic CD5+ B cells the expression of BCL-2 gene family is shifted toward protection from apoptosis. *Br J Haematol.* 1996;94:612-618.
88. Vrhovac R, Delmer A, Tang R, Marie JP, Zittoun R, Ajchenbaum-Cymbalista F. Prognostic significance of the cell cycle inhibitor p27Kip1 in chronic B-cell lymphocytic leukemia. *Blood.* 1998;91:4694-4700.
89. Messmer BT, Messmer D, Allen SL, et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. *J Clin Invest.* 2005;115:755-764.
90. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer.* 2006;6:107-116.
91. Belinsky SA, Nikula KJ, Palmisano WA, et al. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc Natl Acad Sci U S A.* 1998;95:11891-11896.
92. Feinberg AP, Tycko B. The history of cancer epigenetics. *Nat Rev Cancer.* 2004;4:143-153.

93. Yamada Y, Jackson-Grusby L, Linhart H, et al. Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. *Proc Natl Acad Sci U S A*. 2005;102:13580-13585.
94. Chiorazzi N, Hatzi K, Albesiano E. B-cell chronic lymphocytic leukemia, a clonal disease of B lymphocytes with receptors that vary in specificity for (auto)antigens. *Ann N Y Acad Sci*. 2005;1062:1-12.
95. Klein U, Tu Y, Stolovitzky GA, et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med*. 2001;194:1625-1638.
96. Alizadeh AA, Majeti R. Surprise! HSC are aberrant in chronic lymphocytic leukemia. *Cancer Cell*. 2011;20:135-136.
97. Kikushige Y, Ishikawa F, Miyamoto T, et al. Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell*. 2010;20:246-259.
98. Bomstein Y, Yuklea M, Radnay J, et al. The antiapoptotic effects of blood constituents in patients with chronic lymphocytic leukemia. *Eur J Haematol*. 2003;70:290-295.
99. Burger JA, Tsukada N, Burger M, Zvaifler NJ, Dell'Aquila M, Kipps TJ. Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1. *Blood*. 2000;96:2655-2663.
100. Lagneaux L, Delforge A, Bron D, De Bruyn C, Stryckmans P. Chronic lymphocytic leukemic B cells but not normal B cells are rescued from apoptosis by contact with normal bone marrow stromal cells. *Blood*. 1998;91:2387-2396.
101. Panayiotidis P, Jones D, Ganeshaguru K, Foroni L, Hoffbrand AV. Human bone marrow stromal cells prevent apoptosis and support the survival of chronic lymphocytic leukaemia cells in vitro. *Br J Haematol*. 1996;92:97-103.
102. Pedersen IM, Kitada S, Leoni LM, et al. Protection of CLL B cells by a follicular dendritic cell line is dependent on induction of Mcl-1. *Blood*. 2002;100:1795-1801.
103. Burger JA, Ghia P, Rosenwald A, Caligaris-Cappio F. The microenvironment in mature B-cell malignancies: a target for new treatment strategies. *Blood*. 2009;114:3367-3375.
104. Burger JA, Kipps TJ. Chemokine receptors and stromal cells in the homing and homeostasis of chronic lymphocytic leukemia B cells. *Leuk Lymphoma*. 2002;43:461-466.
105. Ghia P, Chiorazzi N, Stamatopoulos K. Microenvironmental influences in chronic lymphocytic leukaemia: the role of antigen stimulation. *J Intern Med*. 2008;264:549-562.
106. Kumar CC. Signaling by integrin receptors. *Oncogene*. 1998;17:1365-1373.
107. Monroe JG, Dorshkind K. Fate decisions regulating bone marrow and peripheral B lymphocyte development. *Adv Immunol*. 2007;95:1-50.
108. Redondo-Munoz J, Ugarte-Berzal E, Garcia-Marco JA, et al. Alpha4beta1 integrin and 190-kDa CD44v constitute a cell surface docking complex for gelatinase B/MMP-9 in chronic leukemic but not in normal B cells. *Blood*. 2008;112:169-178.
109. Caligaris-Cappio F. Role of the microenvironment in chronic lymphocytic leukaemia. *Br J Haematol*. 2003;123:380-388.

110. Burger JA, Gandhi V. The lymphatic tissue microenvironments in chronic lymphocytic leukemia: in vitro models and the significance of CD40-CD154 interactions. *Blood*. 2009;114:2560-2561; author reply 2561-2562.
111. de la Fuente MT, Casanova B, Garcia-Gila M, Silva A, Garcia-Pardo A. Fibronectin interaction with alpha4beta1 integrin prevents apoptosis in B cell chronic lymphocytic leukemia: correlation with Bcl-2 and Bax. *Leukemia*. 1999;13:266-274.
112. Plate JM, Long BW, Kelkar SB. Role of beta2 integrins in the prevention of apoptosis induction in chronic lymphocytic leukemia B cells. *Leukemia*. 2000;14:34-39.
113. Mohle R, Failenschmid C, Bautz F, Kanz L. Overexpression of the chemokine receptor CXCR4 in B cell chronic lymphocytic leukemia is associated with increased functional response to stromal cell-derived factor-1 (SDF-1). *Leukemia*. 1999;13:1954-1959.
114. Burger JA, Burger M, Kipps TJ. Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells. *Blood*. 1999;94:3658-3667.
115. Bertilaccio MT, Scielzo C, Muzio M, Caligaris-Cappio F. An overview of chronic lymphocytic leukaemia biology. *Best Pract Res Clin Haematol*. 2010;23:21-32.
116. Longo PG, Laurenti L, Gobessi S, Sica S, Leone G, Efremov DG. The Akt/Mcl-1 pathway plays a prominent role in mediating antiapoptotic signals downstream of the B-cell receptor in chronic lymphocytic leukemia B cells. *Blood*. 2008;111:846-855.
117. Muzio M, Apollonio B, Scielzo C, et al. Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood*. 2008;112:188-195.
118. Herishanu Y, Perez-Galan P, Liu D, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood*. 2011;117:563-574.
119. Herreros B, Rodriguez-Pinilla SM, Pajares R, et al. Proliferation centers in chronic lymphocytic leukemia: the niche where NF-kappaB activation takes place. *Leukemia*. 2010;24:872-876.
120. Furman RR, Asgary Z, Mascarenhas JO, Liou HC, Schattner EJ. Modulation of NF-kappa B activity and apoptosis in chronic lymphocytic leukemia B cells. *J Immunol*. 2000;164:2200-2206.
121. Hewamana S, Lin TT, Rowntree C, et al. Rel a is an independent biomarker of clinical outcome in chronic lymphocytic leukemia. *J Clin Oncol*. 2009;27:763-769.
122. Ghia P, Strola G, Granziero L, et al. Chronic lymphocytic leukemia B cells are endowed with the capacity to attract CD4+, CD40L+ T cells by producing CCL22. *Eur J Immunol*. 2002;32:1403-1413.
123. Granziero L, Ghia P, Circosta P, et al. Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. *Blood*. 2001;97:2777-2783.
124. Ghia P, Guida G, Stella S, et al. The pattern of CD38 expression defines a distinct subset of chronic lymphocytic leukemia (CLL) patients at risk of disease progression. *Blood*. 2003;101:1262-1269.
125. Jaksic O, Paro MM, Kardum Skelin I, Kusec R, Pejisa V, Jaksic B. CD38 on B-cell chronic lymphocytic leukemia cells has higher expression in lymph nodes than in peripheral blood or bone marrow. *Blood*. 2004;103:1968-1969.



126. Swerdlow SH, Murray LJ, Habeshaw JA, Stansfeld AG. Lymphocytic lymphoma/B-chronic lymphocytic leukaemia--an immunohistopathological study of peripheral B lymphocyte neoplasia. *Br J Cancer*. 1984;50:587-599.
127. Deaglio S, Morra M, Mallone R, et al. Human CD38 (ADP-ribosyl cyclase) is a counter-receptor of CD31, an Ig superfamily member. *J Immunol*. 1998;160:395-402.
128. Deaglio S, Vaisitti T, Aydin S, Ferrero E, Malavasi F. In-tandem insight from basic science combined with clinical research: CD38 as both marker and key component of the pathogenetic network underlying chronic lymphocytic leukemia. *Blood*. 2006;108:1135-1144.
129. Calissano C, Damle RN, Hayes G, et al. In vivo intraclonal and interclonal kinetic heterogeneity in B-cell chronic lymphocytic leukemia. *Blood*. 2009;114:4832-4842.
130. Deaglio S, Malavasi F. Chronic lymphocytic leukemia microenvironment: shifting the balance from apoptosis to proliferation. *Haematologica*. 2009;94:752-756.
131. Patten PE, Buggins AG, Richards J, et al. CD38 expression in chronic lymphocytic leukemia is regulated by the tumor microenvironment. *Blood*. 2008;111:5173-5181.
132. Lund FE, Yu N, Kim KM, Reth M, Howard MC. Signaling through CD38 augments B cell antigen receptor (BCR) responses and is dependent on BCR expression. *J Immunol*. 1996;157:1455-1467.
133. Zupo S, Isnardi L, Megna M, et al. CD38 expression distinguishes two groups of B-cell chronic lymphocytic leukemias with different responses to anti-IgM antibodies and propensity to apoptosis. *Blood*. 1996;88:1365-1374.
134. Fais F, Ghiotto F, Hashimoto S, et al. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest*. 1998;102:1515-1525.
135. Tobin G, Thunberg U, Johnson A, et al. Somatically mutated Ig V(H)3-21 genes characterize a new subset of chronic lymphocytic leukemia. *Blood*. 2002;99:2262-2264.
136. Messmer BT, Albesiano E, Efremov DG, et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med*. 2004;200:519-525.
137. Tobin G, Thunberg U, Johnson A, et al. Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. *Blood*. 2003;101:4952-4957.
138. Widhopf GF, 2nd, Rassenti LZ, Toy TL, Gribben JG, Wierda WG, Kipps TJ. Chronic lymphocytic leukemia B cells of more than 1% of patients express virtually identical immunoglobulins. *Blood*. 2004;104:2499-2504.
139. Herve M, Xu K, Ng YS, et al. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J Clin Invest*. 2005;115:1636-1643.
140. Borsche L, Lim A, Binet JL, Dighiero G. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood*. 1990;76:562-569.
141. Catera R, Silverman GJ, Hatzi K, et al. Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol Med*. 2008;14:665-674.

142. Lanemo Myhrinder A, Hellqvist E, Sidorova E, et al. A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood*. 2008;111:3838-3848.
143. Chu CC, CATERA R, Hatzi K, et al. Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood*. 2008;112:5122-5129.
144. Rosen A, Murray F, Evaldsson C, Rosenquist R. Antigens in chronic lymphocytic leukemia--Implications for cell origin and leukemogenesis. *Semin Cancer Biol*.2010; 20:400-9.
145. Sasso EH, Silverman GJ, Mannik M. Human IgM molecules that bind staphylococcal protein A contain VHIII H chains. *J Immunol*. 1989;142:2778-2783.
146. Seppala I, Kaartinen M, Ibrahim S, Makela O. Mouse Ig coded by VH families S107 or J606 bind to protein A. *J Immunol*. 1990;145:2989-2993.
147. Bomben R, Dal Bo M, Capello D, et al. Comprehensive characterization of IGHV3-21-expressing B-cell chronic lymphocytic leukemia: an Italian multicenter study. *Blood*. 2007;109:2989-2998.
148. Tobin G, Soderberg O, Thunberg U, Rosenquist R. V(H)3-21 gene usage in chronic lymphocytic leukemia--characterization of a new subgroup with distinct molecular features and poor survival. *Leuk Lymphoma*. 2004;45:221-228.
149. Ghia P, Stamatopoulos K, Belessi C, et al. Geographic patterns and pathogenetic implications of IGHV gene usage in chronic lymphocytic leukemia: the lesson of the IGHV3-21 gene. *Blood*. 2005;105:1678-1685.
150. Thorselius M, Krober A, Murray F, et al. Strikingly homologous immunoglobulin gene rearrangements and poor outcome in VH3-21-using chronic lymphocytic leukemia patients independent of geographic origin and mutational status. *Blood*. 2006;107:2889-2894.
151. Bomben R, Dal Bo M, Capello D, et al. Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicentre study. *Br J Haematol*. 2009;144:492-506.
152. Ghia EM, Jain S, Widhopf GF, 2nd, et al. Use of IGHV3-21 in chronic lymphocytic leukemia is associated with high-risk disease and reflects antigen-driven, post-germinal center leukemogenic selection. *Blood*. 2008;111:5101-5108.
153. Abramenko I, Bilous N, Kryachok I, et al. IGHV3-21 gene expression in patients with B-cell chronic lymphocytic leukemia in Ukraine. *Exp Oncol*. 2007;29:226-230.
154. Lin K, Manocha S, Harris RJ, Matrai Z, Sherrington PD, Pettitt AR. High frequency of p53 dysfunction and low level of VH mutation in chronic lymphocytic leukemia patients using the VH3-21 gene segment. *Blood*. 2003;102:1145-1146.
155. Matthews C, Catherwood M, Morris TC, Alexander HD. Routine analysis of IgVH mutational status in CLL patients using BIOMED-2 standardized primers and protocols. *Leuk Lymphoma*. 2004;45:1899-1904.
156. Kostareli E, Hadzidimitriou A, Stavroyianni N, et al. Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Leukemia*. 2009;23:919-924.
157. Sutton LA, Kostareli E, Hadzidimitriou A, et al. Extensive intracлонаl diversification in a subgroup of chronic lymphocytic leukemia patients with

- stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. *Blood*. 2009;114:4460-4468.
158. Pugh-Bernard AE, Silverman GJ, Cappione AJ, et al. Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. *J Clin Invest*. 2001;108:1061-1070.
  159. Krober A, Seiler T, Benner A, et al. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood*. 2002;100:1410-1416.
  160. Tobin G, Thunberg U, Laurell A, et al. Patients with chronic lymphocytic leukemia with mutated VH genes presenting with Binet stage B or C form a subgroup with a poor outcome. *Haematologica*. 2005;90:465-469.
  161. Ibrahim S, Keating M, Do KA, et al. CD38 expression as an important prognostic factor in B-cell chronic lymphocytic leukemia. *Blood*. 2001;98:181-186.
  162. Hamblin TJ, Orchard JA, Ibbotson RE, et al. CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. *Blood*. 2002;99:1023-1029.
  163. Thunberg U, Johnson A, Roos G, et al. CD38 expression is a poor predictor for VH gene mutational status and prognosis in chronic lymphocytic leukemia. *Blood*. 2001;97:1892-1894.
  164. Boonstra JG, van Lom K, Langerak AW, et al. CD38 as a prognostic factor in B cell chronic lymphocytic leukaemia (B-CLL): comparison of three approaches to analyze its expression. *Cytometry B Clin Cytom*. 2006;70:136-141.
  165. Rassenti LZ, Jain S, Keating MJ, et al. Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood*. 2008;112:1923-1930.
  166. Crespo M, Bosch F, Villamor N, et al. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N Engl J Med*. 2003;348:1764-1775.
  167. Rassenti LZ, Huynh L, Toy TL, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med*. 2004;351:893-901.
  168. Krober A, Bloehdorn J, Hafner S, et al. Additional genetic high-risk features such as 11q deletion, 17p deletion, and V3-21 usage characterize discordance of ZAP-70 and VH mutation status in chronic lymphocytic leukemia. *J Clin Oncol*. 2006;24:969-975.
  169. Catherwood MA, Matthews C, Niblock R, Dobbin E, Morris TC, Alexander HD. ZAP-70 mRNA quantification in B-cell chronic lymphocytic leukaemia. *Eur J Haematol*. 2006;76:294-298.
  170. Mansouri M, Sevov M, Fahlgren E, et al. Lipoprotein lipase is differentially expressed in prognostic subsets of chronic lymphocytic leukemia but displays invariably low catalytical activity. *Leuk Res*. 2010;34:301-306.
  171. Rosenwald A, Alizadeh AA, Widhopf G, et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med*. 2001;194:1639-1647.
  172. Nuckel H, Huttmann A, Klein-Hitpass L, et al. Lipoprotein lipase expression is a novel prognostic factor in B-cell chronic lymphocytic leukemia. *Leuk Lymphoma*. 2006;47:1053-1061.

173. Kaderi MA, Kanduri M, Buhl AM, et al. LPL is the strongest prognostic factor in a comparative analysis of RNA-based markers in early chronic lymphocytic leukemia. *Haematologica*.2011;96:1153-1160.
174. Heintel D, Kienle D, Shehata M, et al. High expression of lipoprotein lipase in poor risk B-cell chronic lymphocytic leukemia. *Leukemia*. 2005;19:1216-1223.
175. Nikitin EA, Malakho SG, Biderman BV, et al. Expression level of lipoprotein lipase and dystrophin genes predict survival in B-cell chronic lymphocytic leukemia. *Leuk Lymphoma*. 2007;48:912-922.
176. van't Veer MB, Brooijmans AM, Langerak AW, et al. The predictive value of lipoprotein lipase for survival in chronic lymphocytic leukemia. *Haematologica*. 2006;91:56-63.
177. Buhl AM, Jurlander J, Jorgensen FS, et al. Identification of a gene on chromosome 12q22 uniquely overexpressed in chronic lymphocytic leukemia. *Blood*. 2006;107:2904-2911.
178. Buhl AM, Jurlander J, Geisler CH, et al. CLLU1 expression levels predict time to initiation of therapy and overall survival in chronic lymphocytic leukemia. *Eur J Haematol*. 2006;76:455-464.
179. Josefsson P, Geisler CH, Leffers H, et al. CLLU1 expression analysis adds prognostic information to risk prediction in chronic lymphocytic leukemia. *Blood*. 2007;109:4973-4979.
180. Buhl AM, Novotny GW, Josefsson P, et al. The CLLU1 expression level is a stable and inherent feature of the chronic lymphocytic leukemia clone. *Leukemia*. 2009;23:1182-1186.
181. Buhl AM, James DF, Neuberg D, Jain S, Rassenti LZ, Kipps TJ. Analysis of CLLU1 expression levels before and after therapy in patients with chronic lymphocytic leukemia. *Eur J Haematol*. 2011;86:405-411.
182. Van Dyke DL, Shanafelt TD, Call TG, et al. A comprehensive evaluation of the prognostic significance of 13q deletions in patients with B-chronic lymphocytic leukaemia. *Br J Haematol*.2010;148:544-550.
183. Gunnarsson R, Isaksson A, Mansouri M, et al. Large but not small copy-number alterations correlate to high-risk genomic aberrations and survival in chronic lymphocytic leukemia: a high-resolution genomic screening of newly diagnosed patients. *Leukemia*.2010;24:211-215.
184. Chena C, Avalos JS, Bezares RF, et al. Biallelic deletion 13q14.3 in patients with chronic lymphocytic leukemia: cytogenetic, FISH and clinical studies. *Eur J Haematol*. 2008;81:94-99.
185. Hernandez JA, Rodriguez AE, Gonzalez M, et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica*. 2009;94:364-371.
186. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A*. 2002;99:15524-15529.
187. Liu Y, Corcoran M, Rasool O, et al. Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia. *Oncogene*. 1997;15:2463-2473.
188. Mertens D, Philippen A, Ruppel M, et al. Chronic lymphocytic leukemia and 13q14: miRs and more. *Leuk Lymphoma*. 2009;50:502-505.

189. Ouillette P, Erba H, Kujawski L, Kaminski M, Shedden K, Malek SN. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14. *Cancer Res.* 2008;68:1012-1021.
190. Klein U, Lia M, Crespo M, et al. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell.*2010;17:28-40.
191. Palamarchuk A, Efanov A, Nazaryan N, et al. 13q14 deletions in CLL involve cooperating tumor suppressors. *Blood.*2010;115:3916-3922.
192. Dohner H, Stilgenbauer S, James MR, et al. 11q deletions identify a new subset of B-cell chronic lymphocytic leukemia characterized by extensive nodal involvement and inferior prognosis. *Blood.* 1997;89:2516-2522.
193. Rotman G, Shiloh Y. ATM: from gene to function. *Hum Mol Genet.* 1998;7:1555-1563.
194. Stilgenbauer S, Liebisch P, James MR, et al. Molecular cytogenetic delineation of a novel critical genomic region in chromosome bands 11q22.3-923.1 in lymphoproliferative disorders. *Proc Natl Acad Sci U S A.* 1996;93:11837-11841.
195. Austen B, Powell JE, Alvi A, et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood.* 2005;106:3175-3182.
196. Bullrich F, Rasio D, Kitada S, et al. ATM mutations in B-cell chronic lymphocytic leukemia. *Cancer Res.* 1999;59:24-27.
197. Gunn SR, Hibbard MK, Ismail SH, et al. Atypical 11q deletions identified by array CGH may be missed by FISH panels for prognostic markers in chronic lymphocytic leukemia. *Leukemia.* 2009;23:1011-1017.
198. Gunnarsson R, Mansouri L, Isaksson A, et al. Array-based genomic screening at diagnosis and during follow-up in chronic lymphocytic leukemia. *Haematologica.*2011;96:1161-1169.
199. Gardiner A, Parker H, Glide S, et al. A new minimal deleted region at 11q22.3 reveals the importance of interpretation of diminished FISH signals and the choice of probe for ATM deletion screening in chronic lymphocytic leukemia. *Leuk Res.*2012;36:307-310.
200. Cardinaud B, Moreilhon C, Marcet B, et al. miR-34b/miR-34c: a regulator of TCL1 expression in 11q- chronic lymphocytic leukaemia? *Leukemia.* 2009;23:2174-2177.
201. Herold T, Jurinovic V, Mulaw M, et al. Expression analysis of genes located in the minimally deleted regions of 13q14 and 11q22-23 in chronic lymphocytic leukemia-unexpected expression pattern of the RHO GTPase activator ARHGAP20. *Genes Chromosomes Cancer.*2011;50:546-558.
202. Auer RL, Riaz S, Cotter FE. The 13q and 11q B-cell chronic lymphocytic leukaemia-associated regions derive from a common ancestral region in the zebrafish. *Br J Haematol.* 2007;137:443-453.
203. Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T. Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. *Leukemia.* 2007;21:2442-2451.
204. Dierlamm J, Wlodarska I, Michaux L, et al. FISH identifies different types of duplications with 12q13-15 as the commonly involved segment in B-cell lymphoproliferative malignancies characterized by partial trisomy 12. *Genes Chromosomes Cancer.* 1997;20:155-166.

205. Haidar MA, El-Hajj H, Bueso-Ramos CE, et al. Expression profile of MDM-2 proteins in chronic lymphocytic leukemia and their clinical relevance. *Am J Hematol.* 1997;54:189-195.
206. Merup M, Juliusson G, Wu X, et al. Amplification of multiple regions of chromosome 12, including 12q13-15, in chronic lymphocytic leukaemia. *Eur J Haematol.* 1997;58:174-180.
207. Wu X, Bayle JH, Olson D, Levine AJ. The p53-mdm-2 autoregulatory feedback loop. *Genes Dev.* 1993;7:1126-1132.
208. Porpaczy E, Bilban M, Heinze G, et al. Gene expression signature of chronic lymphocytic leukaemia with Trisomy 12. *Eur J Clin Invest.* 2009;39:568-575.
209. Stilgenbauer S, Sander S, Bullinger L, et al. Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. *Haematologica.* 2007;92:1242-1245.
210. Zenz T, Eichhorst B, Busch R, et al. TP53 Mutation and Survival in Chronic Lymphocytic Leukemia. *J Clin Oncol.* 2010;28:4473-4479.
211. Rossi D, Cerri M, Deambrogi C, et al. The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clin Cancer Res.* 2009;15:995-1004.
212. Zenz T, Krober A, Scherer K, et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood.* 2008;112:3322-3329.
213. Dicker F, Herholz H, Schnittger S, et al. The detection of TP53 mutations in chronic lymphocytic leukemia independently predicts rapid disease progression and is highly correlated with a complex aberrant karyotype. *Leukemia.* 2009;23:117-124.
214. Merup M, Moreno TC, Heyman M, et al. 6q deletions in acute lymphoblastic leukemia and non-Hodgkin's lymphomas. *Blood.* 1998;91:3397-3400.
215. Stilgenbauer S, Bullinger L, Benner A, et al. Incidence and clinical significance of 6q deletions in B cell chronic lymphocytic leukemia. *Leukemia.* 1999;13:1331-1334.
216. Finn WG, Kay NE, Kroft SH, Church S, Peterson LC. Secondary abnormalities of chromosome 6q in B-cell chronic lymphocytic leukemia: a sequential study of karyotypic instability in 51 patients. *Am J Hematol.* 1998;59:223-229.
217. Amiel A, Mulchanov I, Elis A, et al. Deletion of 6q27 in chronic lymphocytic leukemia and multiple myeloma detected by fluorescence in situ hybridization. *Cancer Genet Cytogenet.* 1999;112:53-56.
218. Sherratt T, Morelli C, Boyle JM, Harrison CJ. Analysis of chromosome 6 deletions in lymphoid malignancies provides evidence for a region of minimal deletion within a 2-megabase segment of 6q21. *Chromosome Res.* 1997;5:118-124.
219. Cuneo A, Rigolin GM, Bigoni R, et al. Chronic lymphocytic leukemia with 6q- shows distinct hematological features and intermediate prognosis. *Leukemia.* 2004;18:476-483.
220. Oscier DG, Stevens J, Hamblin TJ, Pickering RM, Lambert R, Fitchett M. Correlation of chromosome abnormalities with laboratory features and clinical course in B-cell chronic lymphocytic leukaemia. *Br J Haematol.* 1990;76:352-358.

221. Mayr C, Speicher MR, Kofler DM, et al. Chromosomal translocations are associated with poor prognosis in chronic lymphocytic leukemia. *Blood*. 2006;107:742-751.
222. Mitelman F, Mertens F, Johansson B. Prevalence estimates of recurrent balanced cytogenetic aberrations and gene fusions in unselected patients with neoplastic disorders. *Genes Chromosomes Cancer*. 2005;43:350-366.
223. Martin-Subero JI, Ibbotson R, Klapper W, et al. A comprehensive genetic and histopathologic analysis identifies two subgroups of B-cell malignancies carrying a t(14;19)(q32;q13) or variant BCL3-translocation. *Leukemia*. 2007;21:1532-1544.
224. Schlette E, Rassidakis GZ, Canoz O, Medeiros LJ. Expression of bcl-3 in chronic lymphocytic leukemia correlates with trisomy 12 and abnormalities of chromosome 19. *Am J Clin Pathol*. 2005;123:465-471.
225. Kujawski L, Ouillette P, Erba H, et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. *Blood*. 2008;112:1993-2003.
226. Reindl L, Bacher U, Dicker F, et al. Biological and clinical characterization of recurrent 14q deletions in CLL and other mature B-cell neoplasms. *Br J Haematol*. 2010;151:25-36.
227. Grubor V, Krasnitz A, Troge JE, et al. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). *Blood*. 2009;113:1294-1303.
228. Gunn SR, Bolla AR, Barron LL, et al. Array CGH analysis of chronic lymphocytic leukemia reveals frequent cryptic monoallelic and biallelic deletions of chromosome 22q11 that include the PRAME gene. *Leuk Res*. 2009;33:1276-1281.
229. Pfeifer D, Pantic M, Skatulla I, et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. *Blood*. 2007;109:1202-1210.
230. Chapiro E, Leporrier N, Radford-Weiss I, et al. Gain of the short arm of chromosome 2 (2p) is a frequent recurring chromosome aberration in untreated chronic lymphocytic leukemia (CLL) at advanced stages. *Leuk Res*. 2010;34:63-68.
231. Woyach JA, Heerema NA, Zhao J, et al. Dic(17;18)(p11.2;p11.2) is a recurring abnormality in chronic lymphocytic leukaemia associated with aggressive disease. *Br J Haematol*. 2010;148:754-759.
232. Juliusson G, Merup M. Cytogenetics in chronic lymphocytic leukemia. *Semin Oncol*. 1998;25:19-26.
233. Schwaenen C, Nessling M, Wessendorf S, et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. *Proc Natl Acad Sci U S A*. 2004;101:1039-1044.
234. Sellmann L, Gesk S, Walter C, et al. Trisomy 19 is associated with trisomy 12 and mutated IGHV genes in B-chronic lymphocytic leukaemia. *Br J Haematol*. 2007;138:217-220.
235. Ibbotson R, Athanasiadou A, Sutton LA, et al. Coexistence of trisomies of chromosomes 12 and 19 in chronic lymphocytic leukemia occurs exclusively in the rare IgG-positive variant. *Leukemia*. 2012;26:170-172.
236. Athanasiadou A, Stamatopoulos K, Gaitatzi M, Stavroyianni N, Fassas A, Anagnostopoulos A. Recurrent cytogenetic findings in subsets of patients with chronic lymphocytic leukemia expressing IgG-switched stereotyped immunoglobulins. *Haematologica*. 2008;93:473-474.

237. Chapiro E, Radford-Weiss I, Bastard C, et al. The most frequent t(14;19)(q32;q13)-positive B-cell malignancy corresponds to an aggressive subgroup of atypical chronic lymphocytic leukemia. *Leukemia*. 2008;22:2123-2127.
238. Balatti V, Bottoni A, Palamarchuk A, et al. NOTCH1 mutations in CLL associated with trisomy 12. *Blood*. 2012;119:329-331.
239. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475:101-105.
240. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med*. 2011;208:1389-1401.
241. Wang L, Lawrence MS, Wan Y, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med*. 2011;365:2497-2506.
242. Bickmore W ed *Karyotype Analysis and Chromosome Banding*. . Chichester.: John Wiley & Sons Ltd; (Apr 2001).
243. Buhmann R, Kurzeder C, Rehklaue J, et al. CD40L stimulation enhances the ability of conventional metaphase cytogenetics to detect chromosome aberrations in B-cell chronic lymphocytic leukaemia cells. *Br J Haematol*. 2002;118:968-975.
244. Decker T, Schneller F, Kronschnabl M, et al. Immunostimulatory CpG-oligonucleotides induce functional high affinity IL-2 receptors on B-CLL cells: costimulation with IL-2 results in a highly immunogenic phenotype. *Exp Hematol*. 2000;28:558-568.
245. Pinkel D, Landegent J, Collins C, et al. Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc Natl Acad Sci U S A*. 1988;85:9138-9142.
246. Roylance R. Methods of molecular analysis: assessing losses and gains in tumours. *Mol Pathol*. 2002;55:25-28.
247. Speicher MR, Gwyn Ballard S, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet*. 1996;12:368-375.
248. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. *Science*. 1996;273:494-497.
249. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*. 1992;258:818-821.
250. Pinkel D, Seagraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet*. 1998;20:207-211.
251. Ylstra B, van den Ijssel P, Carvalho B, Brakenhoff RH, Meijer GA. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res*. 2006;34:445-450.
252. Wicker N, Carles A, Mills IG, et al. A new look towards BAC-based array CGH through a comprehensive comparison with oligo-based array CGH. *BMC Genomics*. 2007;8:84.
253. Gunnarsson R, Staaf J, Jansson M, et al. Screening for copy-number alterations and loss of heterozygosity in chronic lymphocytic leukemia--a comparative study of four differently designed, high resolution microarray platforms. *Genes Chromosomes Cancer*. 2008;47:697-711.



254. Hagenkord JM, Monzon FA, Kash SF, Lilleberg S, Xie Q, Kant JA. Array-based karyotyping for prognostic assessment in chronic lymphocytic leukemia: performance comparison of Affymetrix 10K2.0, 250K Nsp, and SNP6.0 arrays. *J Mol Diagn.*2010;12:184-196.
255. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet.* 2009;10:295-304.
256. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16:6-21.
257. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature.* 1986;321:209-213.
258. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature.* 1993;366:362-365.
259. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis.*2010;31:27-36.
260. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* 1997;13:335-340.
261. He XJ, Chen T, Zhu JK. Regulation and function of DNA methylation in plants and animals. *Cell Res.*2011;21:442-465.
262. Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet.* 2002;3:662-673.
263. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science.* 2001;293:1089-1093.
264. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* 1999;99:247-257.
265. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science.* 1975;187:226-232.
266. Popp C, Dean W, Feng S, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature.*2010;463:1101-1105.
267. Jost JP. Nuclear extracts of chicken embryos promote an active demethylation of DNA by excision repair of 5-methyldeoxycytidine. *Proc Natl Acad Sci U S A.* 1993;90:4684-4688.
268. Zhu B, Zheng Y, Angliker H, et al. 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. *Nucleic Acids Res.* 2000;28:4157-4165.
269. Metivier R, Gallais R, Tiffoche C, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature.* 2008;452:45-50.
270. Hirasawa R, Chiba H, Kaneda M, et al. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* 2008;22:1607-1616.
271. Nakamura T, Arai Y, Umehara H, et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol.* 2007;9:64-71.
272. Ooi SK, Qiu C, Bernstein E, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature.* 2007;448:714-717.
273. Hashimshony T, Zhang J, Keshet I, Bustin M, Cedar H. The role of DNA methylation in setting up chromatin structure during development. *Nat Genet.* 2003;34:187-192.
274. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature.* 2008;454:766-770.

275. Mohn F, Weber M, Rebhan M, et al. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell*. 2008;30:755-766.
276. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*. 2007;130:77-88.
277. Macleod D, Charlton J, Mullins J, Bird AP. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes Dev*. 1994;8:2282-2292.
278. Fuhrmann G, Chung AC, Jackson KJ, et al. Mouse germline restriction of Oct4 expression by germ cell nuclear factor. *Dev Cell*. 2001;1:377-387.
279. Epsztejn-Litman S, Feldman N, Abu-Remaileh M, et al. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol*. 2008;15:1176-1183.
280. Feldman N, Gerson A, Fang J, et al. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol*. 2006;8:188-194.
281. Lee TI, Jenner RG, Boyer LA, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*. 2006;125:301-313.
282. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer*. 2006;6:846-856.
283. Bernstein BE, Mikkelsen TS, Xie X, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006;125:315-326.
284. Horard B, Tatout C, Poux S, Pirrotta V. Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol Cell Biol*. 2000;20:3187-3197.
285. Vire E, Brenner C, Deplus R, et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*. 2006;439:871-874.
286. Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res*. 1998;72:141-196.
287. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med*. 2003;349:2042-2054.
288. Kalari S, Pfeifer GP. Identification of driver and passenger DNA methylation in cancer by epigenomic analysis. *Adv Genet*. 2010;70:277-308.
289. Baylin SB. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol*. 2005;2 Suppl 1:S4-11.
290. Song F, Smith JF, Kimura MT, et al. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A*. 2005;102:3336-3341.
291. Ehrlich M. DNA methylation in cancer: too much, but also too little. *Oncogene*. 2002;21:5400-5413.
292. Herman JG. Hypermethylation of tumor suppressor genes in cancer. *Semin Cancer Biol*. 1999;9:359-367.
293. Robertson KD. DNA methylation and human disease. *Nat Rev Genet*. 2005;6:597-610.
294. Doi A, Park IH, Wen B, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet*. 2009;41:1350-1353.

295. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet.* 2009;41:178-186.
296. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science.* 2003;300:455.
297. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet.* 2007;8:286-298.
298. Weber M, Davies JJ, Wittig D, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet.* 2005;37:853-862.
299. Martin-Subero JI, Ammerpohl O, Bibikova M, et al. A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One.* 2009;4:e6986.
300. Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature.* 1978;274:775-780.
301. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet.* 1988;78:151-155.
302. Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.* 1994;54:4855-4878.
303. Magewu AN, Jones PA. Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. *Mol Cell Biol.* 1994;14:4225-4232.
304. Lyko F, Stach D, Brenner A, et al. Quantitative analysis of DNA methylation in chronic lymphocytic leukemia patients. *Electrophoresis.* 2004;25:1530-1535.
305. Stach D, Schmitz OJ, Stilgenbauer S, et al. Capillary electrophoretic analysis of genomic DNA methylation levels. *Nucleic Acids Res.* 2003;31:E2.
306. Wahlfors J, Hiltunen H, Heinonen K, Hamalainen E, Alhonen L, Janne J. Genomic hypomethylation in human chronic lymphocytic leukemia. *Blood.* 1992;80:2074-2080.
307. Raval A, Lucas DM, Matkovic JJ, et al. TWIST2 demonstrates differential methylation in immunoglobulin variable heavy chain mutated and unmutated chronic lymphocytic leukemia. *J Clin Oncol.* 2005;23:3877-3885.
308. Corcoran M, Parker A, Orchard J, et al. ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. *Haematologica.* 2005;90:1078-1088.
309. Raval A, Tanner SM, Byrd JC, et al. Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell.* 2007;129:879-890.
310. Rahmatpanah FB, Carstens S, Hooshmand SI, et al. Large-scale analysis of DNA methylation in chronic lymphocytic leukemia. *Epigenomics.* 2009;1:39-61.
311. Rush LJ, Raval A, Funchain P, et al. Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer Res.* 2004;64:2424-2433.
312. Yu MK, Bergonia H, Szabo A, Phillips JD. Progressive disease in chronic lymphocytic leukemia is correlated with the DNA methylation index. *Leuk Res.* 2007;31:773-777.

313. Fabris S, Bollati V, Agnelli L, et al. Biological and clinical relevance of quantitative global methylation of repetitive DNA sequences in chronic lymphocytic leukemia. *Epigenetics*.2011;6:188-194.
314. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363:2424-2433.
315. Yan XJ, Xu J, Gu ZH, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*.2011;43:309-315.
316. Wigle TJ, Knutson SK, Jin L, et al. The Y641C mutation of EZH2 alters substrate specificity for histone H3 lysine 27 methylation states. *FEBS Lett*. 2011;585:3011-3014.
317. Ernst T, Chase AJ, Score J, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet*.2010;42:722-726.
318. Chen S, Bohrer LR, Rai AN, et al. Cyclin-dependent kinases regulate epigenetic gene silencing through phosphorylation of EZH2. *Nat Cell Biol*.2010;12:1108-1114.
319. Varambally S, Cao Q, Mani RS, et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science*. 2008;322:1695-1699.
320. Sander S, Bullinger L, Klapproth K, et al. MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. *Blood*. 2008;112:4202-4212.
321. Martin-Subero JI, Kreuz M, Bibikova M, et al. New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling. *Blood*. 2009;113:2488-2497.
322. Pike BL, Greiner TC, Wang X, et al. DNA methylation profiles in diffuse large B-cell lymphoma and their relationship to gene expression status. *Leukemia*. 2008;22:1035-1043.
323. Tong WG, Wierda WG, Lin E, et al. Genome-wide DNA methylation profiling of chronic lymphocytic leukemia allows identification of epigenetically repressed molecular pathways with clinical impact. *Epigenetics*.2010;5:499-508.
324. Yang HH, Hu N, Wang C, et al. Influence of genetic background and tissue types on global DNA methylation patterns. *PLoS One*.2010;.5:e9355.
325. Boks MP, Derks EM, Weisenberger DJ, et al. The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One*. 2009;4:e6767.
326. Mishra DK, Chen Z, Wu Y, Sarkissyan M, Koeffler HP, Vadgama JV. Global methylation pattern of genes in androgen-sensitive and androgen-independent prostate cancer cells. *Mol Cancer Ther*.2010;9:33-45.
327. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. *J Nutr*. 2002;132:2333S-2335S.
328. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *Jama*. 2008;299:2877-2883.
329. Markl ID, Cheng J, Liang G, Shibata D, Laird PW, Jones PA. Global and gene-specific epigenetic patterns in human bladder cancer genomes are relatively stable in vivo and in vitro over time. *Cancer Res*. 2001;61:5875-5884.

330. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A*. 2005;102:10604-10609.
331. Armstrong KM, Bermingham EN, Bassett SA, Treloar BP, Roy NC, Barnett MP. Global DNA methylation measurement by HPLC using low amounts of DNA. *Biotechnol J*. 2011;6:113-117.
332. Kuo KC, McCune RA, Gehrke CW, Midgett R, Ehrlich M. Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res*. 1980;8:4763-4776.
333. Khulan B, Thompson RF, Ye K, et al. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res*. 2006;16:1046-1055.
334. Irizarry RA, Ladd-Acosta C, Carvalho B, et al. Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res*. 2008;18:780-790.
335. Rauch T, Li H, Wu X, Pfeifer GP. MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. *Cancer Res*. 2006;66:7939-7947.
336. Clark SJ, Statham A, Stirzaker C, Molloy PL, Frommer M. DNA methylation: bisulphite modification and analysis. *Nat Protoc*. 2006;1:2353-2364.
337. Fouse SD, Nagarajan RO, Costello JF. Genome-scale DNA methylation analysis. *Epigenomics*. 2010;2:105-117.
338. Costello JF, Hong C, Plass C, Smiraglia DJ. Restriction landmark genomic scanning: analysis of CpG islands in genomes by 2D gel electrophoresis. *Methods Mol Biol*. 2009;507:131-148.
339. Hatada I, Hayashizaki Y, Hirotsune S, Komatsubara H, Mukai T. A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc Natl Acad Sci U S A*. 1991;88:9523-9527.
340. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462:315-322.
341. [http://www.illumina.com/products/infinium\\_humanmethylation\\_27beadchip\\_kits.ilmn](http://www.illumina.com/products/infinium_humanmethylation_27beadchip_kits.ilmn).
342. Fouse SD, Shen Y, Pellegrini M, et al. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell*. 2008;2:160-169.
343. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*. 2007;39:457-466.
344. Zhang X, Yazaki J, Sundaresan A, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell*. 2006;126:1189-1201.
345. Brunner AL, Johnson DS, Kim SW, et al. Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res*. 2009;19:1044-1056.
346. Down TA, Rakyen VK, Turner DJ, et al. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol*. 2008;26:779-785.

347. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A*. 1996;93:9821-9826.
348. Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. Bisulfite sequencing of DNA. *Curr Protoc Mol Biol*;Chapter 7:Unit 7 9 1-17.
349. Mikeska T, Felsberg J, Hewitt CA, Dobrovic A. Analysing DNA methylation using bisulphite pyrosequencing. *Methods Mol Biol*.2011;791:33-53.
350. Tost J, Gut IG. Analysis of gene-specific DNA methylation patterns by pyrosequencing technology. *Methods Mol Biol*. 2007;373:89-102.
351. Mack GS. Epigenetic cancer therapy makes headway. *J Natl Cancer Inst*. 2006;98:1443-1444.
352. Smedby KE, Hjalgrim H, Melbye M, et al. Ultraviolet radiation exposure and risk of malignant lymphomas. *J Natl Cancer Inst*. 2005;97:199-209.
353. Ghia P, Stamatopoulos K, Belessi C, et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia*. 2007;21:1-3.
354. Lefranc MP, Giudicelli V, Ginestoux C, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res*. 2009;37:D1006-1012.
355. <http://www.affymetrix.com>.
356. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics*. 2006;22:7-12.
357. Goransson H, Edlund K, Rydaker M, et al. Quantification of normal cell fraction and copy number neutral LOH in clinical lung cancer samples using SNP array data. *PLoS One*. 2009;4:e6057.
358. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001;125:279-284.
359. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics*. 2002;18:1427-1431.
360. Stevenson FK, Caligaris-Cappio F. Chronic lymphocytic leukemia: revelations from the B-cell receptor. *Blood*. 2004;103:4389-4395.
361. Matthews C, Catherwood MA, Morris TC, Alexander HD. V(H)3-48 and V(H)3-53, as well as V(H)3-21, gene rearrangements define unique subgroups in CLL and are associated with biased lambda light chain restriction, homologous LCDR3 sequences and poor prognosis. *Leuk Res*. 2007;31:231-234.
362. Hojjat-Farsangi M, Jeddi-Tehrani M, Razavi SM, et al. Immunoglobulin heavy chain variable region gene usage and mutational status of the leukemic B cells in Iranian patients with chronic lymphocytic leukemia. *Cancer Sci*. 2009;100:2346-2353.
363. Latif F, Tory K, Gnarra J, et al. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science*. 1993;260:1317-1320.
364. Matsuda S, Yokozaki S, Yoshida H, Kitagishi Y, Shirafuji N, Okumura N. Insulin receptor substrate protein 53 (IRSp53) as a binding partner of antimetastasis molecule NESH, a member of Abelson interactor protein family. *Ann Oncol*. 2008;19:1356-1357.
365. Hayashida Y, Goi T, Hirono Y, Katayama K, Urano T, Yamaguchi A. PPP1R3 gene (protein phosphatase 1) alterations in colorectal cancer and its relationship to metastasis. *Oncol Rep*. 2005;13:1223-1227.

366. Kleer CG, Zhang Y, Pan Q, et al. WISP3 is a novel tumor suppressor gene of inflammatory breast cancer. *Oncogene*. 2002;21:3172-3180.
367. Yamada HY, Gorbsky GJ. Tumor suppressor candidate TSSC5 is regulated by UbcH6 and a novel ubiquitin ligase RING105. *Oncogene*. 2006;25:1330-1339.
368. Bar-Yehuda S, Stemmer SM, Madi L, et al. The A3 adenosine receptor agonist CF102 induces apoptosis of hepatocellular carcinoma via de-regulation of the Wnt and NF-kappaB signal transduction pathways. *Int J Oncol*. 2008;33:287-295.
369. Maffei R, Marasca R, Martinelli S, et al. Angiopoietin-2 expression in B-cell chronic lymphocytic leukemia: association with clinical outcome and immunoglobulin heavy-chain mutational status. *Leukemia*. 2007;21:1312-1315.
370. Nagasaki K, Schem C, von Kaisenberg C, et al. Leucine-zipper protein, LDOC1, inhibits NF-kappaB activation and sensitizes pancreatic cancer cells to apoptosis. *Int J Cancer*. 2003;105:454-458.
371. Grimwade D, Du MQ, Langabeer S, Rogers J, Solomon E. Screening for mutations of Bcl10 in leukaemia. *Br J Haematol*. 2000;109:611-615.
372. Cuni S, Perez-Aciego P, Perez-Chacon G, et al. A sustained activation of PI3K/NF-kappaB pathway is critical for the survival of chronic lymphocytic leukemia B cells. *Leukemia*. 2004;18:1391-1400.
373. Hofbauer SW, Pinon JD, Brachtl G, et al. Modifying akt signaling in B-cell chronic lymphocytic leukemia cells. *Cancer Res*. 2010;70:7336-7344.
374. Kanazawa N, Okafuji I, Kambe N, et al. Early-onset sarcoidosis and CARD15 mutations with constitutive nuclear factor-kappaB activation: common genetic etiology with Blau syndrome. *Blood*. 2005;105:1195-1197.
375. Ishikawa H, Barber GN. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature*. 2008;455:674-678.
376. Zarnack K, Eichhorn H, Kahmann R, Feldbrugge M. Pheromone-regulated target genes respond differentially to MAPK phosphorylation of transcription factor Prf1. *Mol Microbiol*. 2008;69:1041-1053.
377. Xiang Z, Yuan W, Luo N, et al. A novel human zinc finger protein ZNF540 interacts with MVP and inhibits transcriptional activities of the ERK signal pathway. *Biochem Biophys Res Commun*. 2006;347:288-296.
378. Francois DT, Katona IM, June CH, Wahl LM, Mond JJ. Examination of the inhibitory and stimulatory effects of IFN-alpha, -beta, and -gamma on human B-cell proliferation induced by various B-cell mitogens. *Clin Immunol Immunopathol*. 1988;48:297-306.
379. Haudenschild DR, Curtiss SB, Moseley TA, Reddi AH. Generation of interleukin-17 receptor-like protein (IL-17RL) in prostate by alternative splicing of RNA. *Prostate*. 2006;66:1268-1274.
380. Natkunam Y, Zhao S, Mason DY, et al. The oncoprotein LMO2 is expressed in normal germinal-center B cells and in human B-cell lymphomas. *Blood*. 2007;109:1636-1642.
381. Coughlin JJ, Stang SL, Dower NA, Stone JC. RasGRP1 and RasGRP3 regulate B cell proliferation by facilitating B cell receptor-Ras signaling. *J Immunol*. 2005;175:7179-7184.
382. Jang YH, Namkoong S, Kim YM, Lee SJ, Park BJ, Min DS. Cleavage of phospholipase D1 by caspase promotes apoptosis via modulation of the p53-dependent cell death pathway. *Cell Death Differ*. 2008;15:1782-1793.

383. Contente A, Dittmer A, Koch MC, Roth J, Dobbelstein M. A polymorphic microsatellite that mediates induction of PIG3 by p53. *Nat Genet.* 2002;30:315-320.
384. Troeger A, Gudowius S, Escherich G, et al. High nerve growth factor receptor (p75NTR) expression is a favourable prognostic factor in paediatric B cell precursor-acute lymphoblastic leukaemia. *Br J Haematol.* 2007;139:450-457.
385. Margueron R, Justin N, Ohno K, et al. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature.* 2009;461:762-767.
386. Pasini D, Cloos PA, Walfridsson J, et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature.*2010;.464:306-310.
387. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* 2006;20:1123-1136.
388. Hahn MA, Hahn T, Lee DH, et al. Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Res.* 2008;68:10280-10289.
389. Schlesinger Y, Straussman R, Keshet I, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet.* 2007;39:232-236.
390. Saborit-Villarroya I, Vaisitti T, Rossi D, et al. E2A is a transcriptional regulator of CD38 expression in chronic lymphocytic leukemia. *Leukemia.*2011;25:479-488.
391. Gutierrez A, Jr., Tschumper RC, Wu X, et al. LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood.*2010;116:2975-2983.
392. Lenschow DJ, Sperling AI, Cooke MP, et al. Differential up-regulation of the B7-1 and B7-2 costimulatory molecules after Ig receptor engagement by antigen. *J Immunol.* 1994;153:1990-1997.
393. Chaperot L, Plumas J, Jacob MC, et al. Functional expression of CD80 and CD86 allows immunogenicity of malignant B cells from non-Hodgkin's lymphomas. *Exp Hematol.* 1999;27:479-488.
394. Osorio LM, Aguilar-Santelises M. Apoptosis in B-chronic lymphocytic leukaemia. *Med Oncol.* 1998;15:234-240.
395. Liu W, Quinto I, Chen X, et al. Direct inhibition of Bruton's tyrosine kinase by IBtk, a Btk-binding protein. *Nat Immunol.* 2001;2:939-946.
396. Chantepie SP, Vaur D, Grunau C, et al. ZAP-70 intron1 DNA methylation status: determination by pyrosequencing in B chronic lymphocytic leukemia. *Leuk Res.*2010;34:800-808.
397. Cordingley FT, Bianchi A, Hoffbrand AV, et al. Tumour necrosis factor as an autocrine tumour growth factor for chronic B-cell malignancies. *Lancet.* 1988;1:969-971.
398. Ferrajoli A, Keating MJ, Manshouri T, et al. The clinical significance of tumor necrosis factor-alpha plasma level in patients having chronic lymphocytic leukemia. *Blood.* 2002;100:1215-1219.
399. Schwertassek U, Balmer Y, Gutscher M, et al. Selective redox regulation of cytokine receptor signaling by extracellular thioredoxin-1. *Embo J.* 2007;26:3086-3097.
400. Backman E, Bergh AC, Lagerdahl I, et al. Thioredoxin, produced by stromal cells retrieved from the lymph node microenvironment, rescues chronic



- lymphocytic leukemia cells from apoptosis in vitro. *Haematologica*. 2007;92:1495-1504.
401. Nilsson J, Soderberg O, Nilsson K, Rosen A. Thioredoxin prolongs survival of B-type chronic lymphocytic leukemia cells. *Blood*. 2000;95:1420-1426.