



Thèse

2010

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Array-comparative genomic hybridization study in diffuse large B-cell lymphoma

Scandurra, Marta

How to cite

SCANDURRA, Marta. Array-comparative genomic hybridization study in diffuse large B-cell lymphoma. Doctoral Thesis, 2010. doi: 10.13097/archive-ouverte/unige:12828

This publication URL: <https://archive-ouverte.unige.ch/unige:12828>

Publication DOI: [10.13097/archive-ouverte/unige:12828](https://doi.org/10.13097/archive-ouverte/unige:12828)

UNIVERSITÉ DE GENÈVE
Section des Sciences Pharmaceutiques

FACULTÉ DES SCIENCES
Professeur Leonardo Scapozza

ONCOLOGY INSTITUTE OF SOUTHERN SWITZERLAND (IOSI)
Laboratory of Experimental Oncology

Francesco Bertoni, MD

ARRAY-COMPARATIVE GENOMIC HYBRIDIZATION STUDY IN DIFFUSE LARGE B-CELL LYMPHOMA

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention interdisciplinaire

par

MARTA SCANDURRA

de

ROME (ITALY)

Thèse N°

4251

BELLINZONA

Atelier Laboratory of experimental Oncology
2010



**UNIVERSITÉ
DE GENÈVE**

FACULTÉ DES SCIENCES

**Doctorat ès sciences
Mention interdisciplinaire**

Thèse de *Madame Marta SCANDURRA*

intitulée :

**"Array-Comparative Genomic Hybridization Study
in Diffuse Large B-Cell Lymphoma"**

La Faculté des sciences, sur le préavis de Messieurs L. SCAPOZZA, professeur ordinaire et directeur de thèse (Section des sciences pharmaceutiques), F. BERTONI, docteur et codirecteur de thèse (Oncology Institute of Southern Switzerland – Laboratory of Experimental Oncology - Functional Genomics and Lymphoma Group – Bellinzona, Switzerland), Madame M. CUENDET, professeure associée (Section des sciences pharmaceutiques), Messieurs A. NERI, professeur associé (Università degli Studi di Milano - Facoltà di Medicina e chirurgia – Dipartimento di scienze mediche – Istituto di Ricovero e Cura a Carattere Scientifico - Milano, Italia), et S. DIRNHOFER, professeur (Universität Basel – Institut für Pathologie – Basel, Schweiz), et autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 27 septembre 2010

Thèse - 4251 -


Le Doyen, Jean-Marc TRISCONE

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Dedicato alla memoria di Nonno Pietro
(To the memory of my grandfather Pietro)

“Il caso aiuta le menti preparate”
(“Chance favors the prepared minds”)

Pasteur L, XIX century

CONTENTS

ABBREVIATIONS	9
ABSTRACT	11
RÉSUMÉ	13
GENERAL AIM	16
1. INTRODUCTION	17
1.1 GENETIC ABERRATIONS IN CANCER	17
1.1.1 GENOMIC GAINS	20
1.1.2 GENOMIC AMPLIFICATION	20
1.1.3 GENOMIC LOSSES	20
1.1.4 LOH AND UNIPARENTAL DISOMY (UPD)	21
1.2 GENOMIC PROFILING AND CANCER	23
1.2.1 ARRAY COMPARATIVE GENOMIC HYBRIDIZATION (aCGH)	24
1.3 B-CELL LYMPHOID NEOPLASMS	29
1.3.1 BIOLOGY OF B-CELLS	29
1.3.2 B-CELLS DIFFERENTIATION AND VDJ-REARRANGEMENT	30
1.3.3 GENETIC CHANGES IN B-CELLS WITHIN GERMINAL CENTERS	33
1.3.4 REGULATION OF B-CELL DEVELOPMENT IN THE HUMAN GC	38
1.4 DIFFUSE LARGE B-CELL LYMPHOMA	41
1.4.1 MORPHOLOGY	41
1.4.2 GENOMIC ABERRATIONS	44
1.4.3 BIOLOGY AND MOLECULAR PATHOGENESIS	46
1.4.4 THERAPY	52
2. MATERIALS AND METHODS	53
3. IDENTIFICATION OF RECURRENT GENETIC LESIONS IN DLBCL	67
3.1 ABSTRACT	67
3.2 AIM OF THE CHAPTER	68
3.3 RESULTS	69
3.3.1 DISTINCT RECURRENT ALTERATIONS IN DLBCL	69
3.3.2 DLBCL AND CELL OF ORIGIN	75
3.5 DISCUSSION	77

4. GENOMIC LESIONS ASSOCIATED WITH A DIFFERENT CLINICAL OUTCOME IN DLBCL TREATED WITH R-CHOP-21	81
4.1 ABSTRACT	81
4.2 AIM OF THE CHAPTER	82
4.3 RESULTS	83
4.3.1 GENOMIC REGIONS AND OUTCOME IN PATIENTS TREATED WITH R-CHOP	83
4.3.2 THE IMPACT OF 8p23.1 LOSS	89
4.3.4 NMF-CLUSTERING	90
4.4 DISCUSSION.....	95
5. MOLECULAR AND FUNCTIONAL CHARACTERIZATION OF A RECURRENT DNA GAIN ON 11q24.3.....	101
5.1 ABSTRACT	101
5.2 AIM OF THE CHAPTER	102
5.3 RESULTS	103
5.3.1 DISTINCT RECURRENT ALTERATIONS IN DLBCL AND IDENTIFICATION OF THE 11q24.3 GAIN.....	103
5.3.2 IDENTIFICATION OF PUTATIVE TARGETS WITHIN 11q24.3 GAIN	104
5.3.3 VALIDATION OF <i>ETS1</i> AND <i>FLI1</i> EXPRESSION IN DLBCL SAMPLES.....	108
5.3.4 LOSS OF FUNCTION STUDIES	108
5.4 DISCUSSION	121
6. CONCLUSION.....	129
7. BIBLIOGRAPHY.....	133
8. ACKNOWLEDGEMENTS	151
9. SUPPLEMENTARY MATERIAL	153
SUPPLEMENTARY TABLE 1	153
SUPERVISED ANALYSIS OF THE GENE EXPRESSION PROFILES OF CASES WITH OR WITHOUT 8p LOSS (8/54, 15%) IDENTIFIED 81 PROBES ALL UPREGULATED.....	153
SUPPLEMENTARY TABLE 2.....	183
10. LIST OF PUBLICATIONS.....	189

ABBREVIATIONS

Ab	antibody
ABC	Activated B-cell
ACA	Abnormal chromosome arm
(a)CGH	(array)comparative genomic hybridization
Ag	antigen
APC	Antigen presenting cells
AWC	Abnormal whole chromosome
BCL-2/6	B-cell lymphoma-2/6
BCR	B-cell receptor
BLIMP-1	B-lymphocyte induced maturation protein 1
BM	bone marrow
CIN	Chromosomal instability
CN	Copy number
CSR	Class switch recombination
DFS	Disease Free Survival
DLBCL	Diffuse large b-cell lymphoma
ds	double stranded
DSB	double strand break
ETS	E-twenty six
FDCs	Follicular dendritic cells
FISH	Fluorescence in situ hybridization
FLI1	Friend leukemia virus integration 1
GC	Germinal center
GCB	Germinal center B-cell
GCOS	GeneChip operating software
GEP	Gene Expression Profile
HMM	Hidden Markov Model

HSC	Hematopoietic stem cell
Ig	Immunoglobulin
Igh	Immunoglobulin heavy chain
IRF	Interferon regulatory factor
LOH	Loss of heterozygosity
LRA	Long recurrent abnormality
MALT	Mucosa-associated lymphoid tissues
mBPCR	Modified bayesian pice wise costant regression
MCR	Minimal common region
miRNA	MicroRNA
mmix	Master mix
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHL	Non-Hodgkin's lymphoma
NMF	Non-negative matrix factorization
NTC	No template control (negative reaction)
OS	Overall Survival
PC	Plasma cell
PFS	Progression Free Survival
R-CHOP	Rituximab- Cyclophosphamide, Hydroxydoxorubicin, Oncovin, Prednisone
SHM	Somatic hypermutation
shRNA	Short hairpin RNA
SNP	Single nucleotide polymorphism
SRA	Short recurrent abnormality
ssDNA	Single stranded DNA
TNF	Tumor necrosis factor
TNFAIP3	Tumor necrosis factor alpha-induced protein 3
UPD	Uniparental disomy
XBP-1	X-box binding protein 1

ABSTRACT

BACKGROUND: Diffuse large B-cell lymphoma (DLBCL) is the most common type of non-Hodgkin's lymphoma (NHL), accounting for almost 40% of the cases. DLBCL can be cured using anthracycline-based chemotherapy regimens in only 40%–50% of patients, suggesting that DLBCL is a heterogeneous diagnostic category and a large part of the patients cannot be cured¹. No information is available on the prognostic impact of genomic aberrations detected using high resolution single nucleotide polymorphism (SNP) array in DLBCL patients treated with R-CHOP. To better characterize the molecular mechanisms occurring in DLBCL and to identify new candidate cancer genes, we performed the genome-wide DNA profiles of 166 DLBCL samples. Moreover we aimed to evaluate the impact of recurrent aberrations on the clinical outcome of patients uniformly treated with R-CHOP. Furthermore we performed unsupervised clustering to identify biological subgroups according to the genetic profile.

RESULTS: Fifty eight recurrent regions of gains, 47 losses and 54 LOH were identified in 166 DLBCL samples. Also, five recurrent amplifications and seven homozygous deletions were found. Twenty lesions showed a statistical significant impact on Overall Survival (OS). Only 8p23.1 loss maintained its significance after multiple test correction (MTC) and was associated with additional aberrations, such as 17p- and 15q-. Unsupervised clustering using the non-negative matrix factorization (NMF) algorithm identified five clusters with distinct genetic profiles, clinical characteristics and outcome: 1 (61/166, 37%), 2 (22/166, 13%), 3 (5/166, 4%), 4 (29/166, 17%), 5 (7/166, 6%). Cluster 1 was characterized by a heterogeneous genomic profile, lacking the most recurrent lesions. Gene expression profiles (GEP) and morphology data suggested that cluster 1 could represent DLBCL with a high content of infiltrating T-cells, partially explaining the low rate of aberrations. This cluster has characteristics similar to the "host response" cluster previously identified by Monti et al., 2005².

A still uncharacterized recurrent gain at 11q24.3, observed in almost 30% of DLBCL cases, contained the genes coding for the transcription factors *ETS1* and *FLI1*, which are involved in lymphocyte development. Real Time PCR data confirmed high expression of *ETS1* and *FLI1* in cases with the recurrent 11q24.3 gain. Furthermore genomic profiling of DLBCL cell lines allowed us to identify the OCI-LY7 as a candidate cell model for further functional characterization of both factors, since it showed the same aberration identified in patients and expressed the two genes at high levels. To evaluate the functional role of the two genes, silencing was performed using the shRNA technology. Interestingly RT-PCR and immunoblot analysis showed that *ETS1* down-regulation by ShRNA in OCI-LY7 leads to the reactivation of the transcription factor *BLIMP1*, considered as one of the master regulator of plasma cell development. Furthermore, the down-regulation of *PAX-5*, a known target of *BLIMP1* and a known negative regulator of the plasma cells (PCs) differentiation, was also observed. Silencing of *ETS1* caused also a block in the proliferation rate of OCI-LY7 cell line. Finally, *FLI1* down-regulation in OCI-LY7, using the above mentioned ShRNA strategy, was also followed by the down-regulation of *PAX5* gene, apparently independently of *ETS1* expression, and by an apparent block of the proliferation rate of the OCI-LY7, leading us to the conclusion that *FLI1* could be an activator of *PAX5*.

CONCLUSIONS: Genomic profiling of a large series of DLBCL samples allowed the identification of genetic features and clusters, associated with a different outcome in patients treated with R-CHOP. We also identified a new amplicons containing two transcription factors, *ETS1* and *FLI1*, known to be involved in B-cell development. We then provided experimental data indicating a role of the two genes in DLBCL pathogenesis.

RÉSUMÉ

CONTEXTE: Les lymphomes diffus à larges cellules B (DLBCL) sont les lymphomes non-Hodgkiniens les plus répandus, représentant presque 40% des cas. Seulement 40 à 50% des patients souffrant des DLBCL peuvent être traités avec une chimiothérapie à base d'anthracycline, suggérant que les DLBCL représentent une catégorie diagnostique hétérogène avec la majorité des cas ne pouvant être correctement traités¹.

Aucune information n'est disponible sur l'impact pronostique des aberrations génomiques détectées en utilisant les «high resolution single nucleotide polymorphism (SNP) array» chez les patients traités avec R-CHOP.

Afin de mieux caractériser les mécanismes moléculaires survenant dans les DLBCL et d'identifier de nouveaux gènes potentiellement cancérigènes, nous avons i) étudié les profils d'ADN génomiques complets de 166 patients DLBCL ; ii) essayé d'évaluer l'impact des aberrations récurrentes sur l'issue clinique des patients traités uniformément avec R-CHOP ; iii) réalisé un clustering non supervisé pour identifier des sous-groupes biologiques en fonction du profil génétique.

RESULTATS: 58 régions récurrentes de gains, 47 de pertes et 54 pertes d'hétérozygotie (Loss of heterozygosity (LOH)) ont été identifiées dans les 166 échantillons de DLBCL. De plus, cinq amplifications récurrentes et sept délétions homozygotes ont été mises en évidence.

20 lésions ont montré un impact statistiquement significatif sur la survie générale (Overall Survival (OS)). Après traitement des données avec la méthode de la correction multiple de texte (Multiple Text Correction » (MTP)), seule la perte 8p23.1 associée avec des aberrations additionnelles telles que 17p- et 15q- a maintenu sa signification statistique. Le regroupement non-supervisé utilisant l'algorithme appelé « non-negative matrix factorization (NMF) algorithm » a mis en evidence cinq clusters aux profils génétiques caractéristiques et des

issues cliniques distincts : 1 (61/166, 37%), 2 (22/166, 13%), 3 (5/166, 4%), 4 (29/166, 17%), 5 (7/166, 6%). Le cluster 1 était caractérisé par un profil génomique hétérogène qui ne présentait pas les lésions les plus récurrentes. Les profils d'expression génétique ainsi que les études morphologiques ont suggéré que le cluster 1 pourrait représenter les DLBCL avec un nombre élevé de cellules T infiltrant la tumeur, apportant ainsi une explication partielle au faible taux d'aberrations. Ce cluster présente des caractéristiques similaires au « host response » cluster identifié par Monti et al., 2005².

Un gain récurrent encore non caractérisé à 11q24.3, observé dans presque 30% des cas de DLBCL, contient les gènes codant pour les facteurs de transcription *ETS1* et *FLI1*, tous deux impliqués dans le développement lymphocytaire. Les résultats de PCR en temps réel ont confirmé une expression élevée de *ETS1* et *FLI1* dans les cas de gain récurrent 11q24.3. En outre, l'étude des profils génomiques de lignées cellulaires DLBCL nous a permis d'identifier OCI-LY7 comme modèle cellulaire potentiel pour l'étude fonctionnelle de ces deux facteurs de transcription, cette lignée montrant les mêmes aberrations que celles identifiées chez les patients et exprimant les 2 gènes à des niveaux élevés.

Pour évaluer leur rôle fonctionnel, ces deux gènes ont été mis sous silence grâce à la technologie du shRNA. Il est intéressant de noter que la mise sous silence de *ETS1* a conduit à la réactivation du facteur de transcription *BLIMP-1* considéré comme le principal régulateur du développement des cellules plasmatisques. De plus, cette réactivation était accompagnée de la répression de *PAX-5*, cible connue de *BLIMP-1* et régulateur négatif de la différenciation des plasmocytes.

La mise sous silence de *ETS1* a également conduit à un blocage de la fréquence de prolifération de la lignée cellulaire OCI-LY7.

Finalement, la même stratégie de shRNA appliquée à *FLI1* avec la lignée OCI-LY7 a montré une répression de l'expression du gène *PAX5*,

apparemment indépendamment de l'expression de *ETS1*, avec un blocage apparent du taux de prolifération de OCI-LY7, nous conduisant à la conclusion que *FLI1* pourrait être un activateur de *PAX5*.

CONCLUSIONS : Le profilage génomique d'une grande série d'échantillons de DLBCL a permis l'identification de caractéristiques génétiques et de clusters associés à des issues cliniques différentes pour les patients traités avec R-CHOP. Nous avons également identifié un nouvel amplicon contenant deux facteurs de transcription, *ETS1* et *FLI1*, connus pour être impliqués dans le développement des cellules B. Nous avons ensuite généré des données expérimentales indiquant un rôle pour ces deux gènes dans la pathogénèse des DLBCL.

GENERAL AIM

The aim of the present work was to detect recurrent aberrations, through the application of the genome wide SNP array 250K, in a subset of DLBCL patients in order to evaluate two main aspects of the disease: the impact of causative lesions in DLBCL genome on the clinical outcome and the identification of those with a biological effect on genes involved in the pathogenesis.

To achieve the described overall goals, specific aims were defined:

i) Characterization of the molecular mechanisms occurring in DLBCL and identification of new candidate cancer genes after performing Array-comparative genomic hybridization (a-CGH) of 166 DLBCL samples in order to obtain the genome-wide DNA profiles. ii) The evaluation of the impact of recurrent aberrations on the clinical outcome of patients uniformly treated with R-CHOP. iii) Unsupervised clustering analysis was performed to identify biological subgroups according to the genetic profile. iv) Functional characterization of genes affected by an identified recurrent gain on 11q24.3.

1. INTRODUCTION

1.1 GENETIC ABERRATIONS IN CANCER

Most tumors are characterized by changes in chromosomal structure and number. Chromosomal aberrations are generally more numerous in malignant than in benign tumors and the karyotypic complexity and cellular heterogeneity observed is often associated with poor prognosis³. Genomic changes observed in cancer can be numerical or structural and can often coexist in complex karyotypes. Numerical aberrations comprise aneuploidy, DNA copy number (CN) changes, and “unbalanced” translocations, whereas structural aberrations include “balanced” translocation and inversions (Figure 1.1). Structural changes are defined balanced if they involve equal exchange of material between two chromosomal regions, whereas the term unbalanced refers to a non reciprocal exchange leading to gain or loss of genome portions. Translocations are given by exchange of chromosomal segments between non-homologous chromosomes, and can be balanced or unbalanced. Spatial proximity and sequence homology are factors probably contributing in increasing the propensity of translocated chromosome partners to rearrange³.

DNA copy number changes comprise deletions, gains or amplifications of genomic material (discussed in details in the next sections). CN aberrations can be also referred to as gene dosage alterations, i.e. alterations in the number of copies of a given sequence found in a cell. It was demonstrated that changes in DNA copy number are a mechanism that lead to altered gene expression in cancer cells⁴.

Further genomic anomalies involved in cancer are somatic point mutations, loss of heterozygosity (LOH) (discussed in section 1.5) and epigenetic modifications. The result of somatic point mutations, in which one or a few nucleotides are altered by sequence error and or deletions or insertion, can be nonsense

mutations, missense mutations or frameshift mutations. In nonsense mutations the new codon causes the protein to prematurely terminate, leading to a shorter and usually non functional product. Missense mutations can determine an incorrect amino acid into the protein sequence. The effect on protein function depends on the site of mutation and the nature of the amino acid replacement. In general, mutations that do not affect the protein sequence or function are called silence or synonymous mutations. Frameshift mutations cause the affected codon to be misread and subsequently also all the following codons, leading to a much different and often non functional product⁵.

Differential DNA methylation patterns exist in normal and cancer cells. DNA methylation involves the addition of a methyl group to the cytosine ring (carbon 5 position) by DNA methyltransferases. In the human genome, epigenetic modification to DNA by methylation is primarily found in repetitive DNA elements where it is thought to play a role in protecting against recombination events and to silence potentially destabilizing transposable elements ⁶. On the contrary, hypermethylation of gene promoter regions causes their transcriptional repression. In tumors, hypomethylation of the genome results in an increased mutation rate, including genome deletions and chromosomal copy number changes^{7, 8}. Methylation studies in various types of tumors have found it to be tightly connected to cancer development through local hypermethylation of tumor-suppressor genes causing transcriptional repression and/or global genomic hypomethylation resulting in the expression of oncogenes.

Another relevant epigenetic modification involved in chromatin remodeling is histone acetylation, which consist of the addition of acetyl groups to lysines from amino-terminal tail of core histones. Lysine acetylation and deacetylation reactions are catalyzed by histone acetyltransferase (HAT) and histone deacetylase (HDAC) respectively. Acetylated histone enhances chromatic decondensation and DNA accessibility, therefore the acetylation states correlates with transcriptional activation. The disruption of

acetylation/deacetylation balance due to the alterations of HATs and HDACs has been detected in many cancers, including lymphoma⁹.

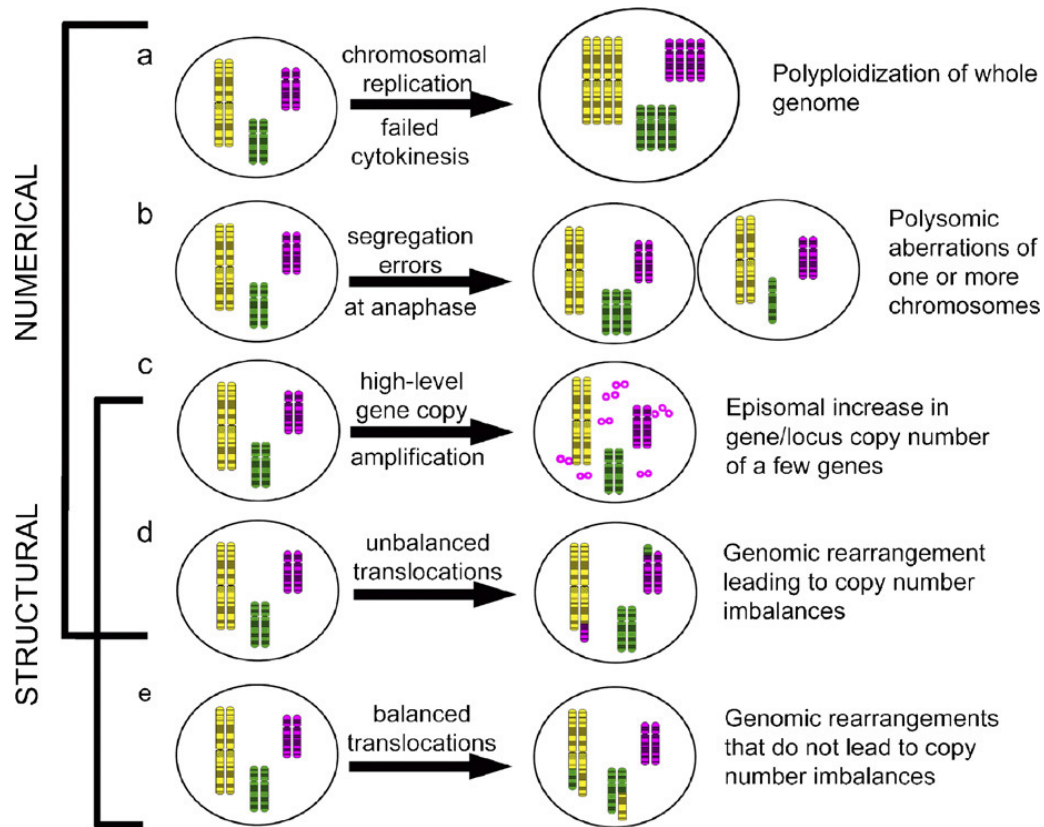


Figure 1.1 Acquisition of numerical and/or structural chromosome aberrations in cancer cells.
(modified by Bayani J et al 2007)³

Importantly, malignant upregulation of histone deacetylation process is being targeted by therapeutic interventions with HDAC inhibitors.

In general, cancer genes are characterized by a redundancy of deregulation mechanisms. Oncogene can be activated by gene dosage alterations, point mutations, translocations, promoter hypomethylation, constitutive active transcription factors, or absence of control by a tumor suppressor gene, if the latter is inactivated. Tumor suppressor genes can be inactivated by gene dosage alterations, LOH events, point mutations or epigenetic silencing by promoter hypermethylation. The identification of genetic alterations can lead to

the discovery of genes, transcripts and pathways that could play a relevant role in cancer.

1.1.1 GENOMIC GAINS

Most recurrent genomic gains probably contribute to tumorigenesis by enhancing the activity of genes in the affected chromosomal regions. Genomic gains commonly arise from chromosomal non-disjunction or unbalanced translocations, which cause complete or partial chromosomal trisomies, or from amplification events affecting DNA segments of different size. Numerous examples of large-scale genomic gains are associated with specific types of cancer. Since such aberrations involve multiple genes, the identification of their functionally relevant targets has proved to be difficult ¹⁰.

1.1.2 GENOMIC AMPLIFICATION

Gene amplification, defined as a high copy number increase of a restricted region of a chromosome arm, can also occur. Amplification is a reiterative process, in which multiple copies of a genome region are accumulated. Amplification requires a DNA double strand break and progression through the cell cycle with this damaged DNA ¹¹. A role for genome context in promoting amplification has also been suggested. Particular genome sequences prone to breakage have also been shown to set the boundaries of amplicons, further suggesting that genome position influences the propensity to amplify ¹¹.

1.1.3 GENOMIC LOSSES

The spectrum of genomic losses ranges from cytogenetically visible alterations, such as complete or partial chromosomal monosomies, to single gene or intragenic deletions that are detectable only by techniques that provide high spatial resolution. Most recurrent genomic losses probably contribute to malignant transformation by reducing the function of specific genes in the affected chromosomal regions ¹⁰. Extensive genomic deletions affecting multiple genes are frequent in tumors, making it difficult to identify which lost

gene contributes to the development of the cancer. The classic approach to identify a tumor-suppressor gene compares multiple tumors with a specific chromosomal deletion to determine the minimal genomic region that is lost in all cases ^{12, 13}. Candidate genes from this region are then screened for deletions, mutations, or epigenetic modifications that inactivate the remaining allele. Cancer-associated chromosomal losses may also act through inactivation of genes that do not encode proteins. For example, several genomic regions that are recurrently deleted in a variety of tumors contain microRNA genes ¹⁴. These genes encode small RNAs involved in post-transcriptional regulation of gene expression, and there is emerging evidence that the loss of specific microRNAs with tumor-suppressive activity may contribute to tumorigenesis ¹⁴.

1.1.4 LOH AND UNIPARENTAL DISOMY (UPD)

Classic LOH is defined as the loss of one allele of a heterozygous locus. Genetic mechanisms leading to LOH are highly variegated. LOH can be caused by deletion, which can involve whole chromosome or smaller region, by mitotic non disjunction, or by mitotic recombination between two homologous chromosomes, including break-induced replication and gene conversion (interstitial mitotic recombination event). LOH results in homozygosity or hemizyosity for a particular mutation. In addition, one allele can be inactivated through mutation and the other through promoter methylation. Therefore, LOH can be due to both DNA CN changes (one copy lost) and copy-number-neutral changes ¹⁵.

In many cases, tumor-suppressor genes can be inactivated both when a germline mutation occurs in one allele and somatic inactivation occurs in the other allele at a later stage or when sequential somatic inactivation of both alleles occurs. In both cases, somatic inactivation is often associated with LOH at the tumor-suppressor gene ¹⁵. UPD refers to the situation in which both copies of a chromosome pair have originated from one parent, either as isodisomy, in which two identical segments from one parent homologue are present, or as heterodisomy, where sequences from both homologues of the transmitting

parent are present ¹⁶ (Figure 1.2). This can occur during the transmission of a chromosome from parent to gamete and during early somatic cell divisions in the zygote (germ line recombination resulting in constitutive UPD), but it is also likely to occur simply as a consequence of somatic recombination during mitotic cell divisions (mitotic or somatic recombination).

Two mechanisms are supposed to be involved: in UPD of whole chromosomes, mitotic nondisjunction is followed by duplication of the remaining homologue in the monosomic cell or loss of one chromosome in the trisomic cell; partial uniparental disomy (pUPD) involves mitotic recombination events between chromatids, such as reciprocal exchange of chromosome material between two homologues ^{15, 17}.

In general, UPD can be an important step in cancer development as well as a contributing factor to other late onset diseases. In fact this phenomenon might be a common unrecognized mechanism in cancer development, because in UPD, the karyotype appears 'normal' when examined by conventional cytogenetic analysis, fluorescence in situ hybridization (FISH) or aCGH ¹⁶. However, analysis by high-resolution single nucleotide polymorphism (SNP) oligonucleotide genomic microarrays allows the identification of copy-number and copy-number-neutral changes. The major difference between a UPD and LOH might be a difference in the gene dosage; copy number loss leading to LOH mostly follows with mutation in the remaining allele (haploinsufficiency); however, loss of one allele and duplication of the remaining allele leads to UPD (homozygosity of the mutated or methylated allele) ¹⁵. Differences in gene dosage might have different pathologic results.

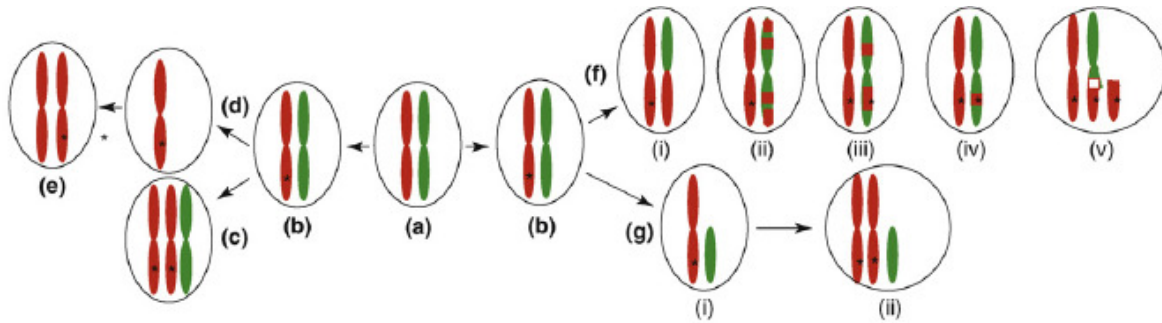


Figure 1.2 Model of the mechanism of UPD in cancer. (a) Both maternal and paternal alleles are present in the normal cell. (b) Mutated cell. During mitosis, incomplete segregation of the chromosome can occur as either (c) a trisomic cell (by non-disjunction) or (d) a monosomic cell (by anaphase lag), and (e) the remaining allele is duplicated by mitotic non-disjunction. Alternatively, (f) (i–iv) one or more mitotic recombination events can occur during mitosis, resulting in segmental UPD. (g) (i) Loss of a chromosome arm and (ii) reduplication by non-disjunction might follow. The white segment indicates deletion ¹⁵.

1.2 GENOMIC PROFILING AND CANCER

The main approaches for identifying genetic events that are involved in the human cancer have been facilitated through the development and application of a diverse series of high-resolution, high-throughput microarray platforms ¹⁸. Basically there are two types of array: those that carry PCR products or DNA fragments from cloned nucleic acids (e.g. cDNA, BACs, cosmids) and those that use oligonucleotides. The latter is mostly represented by Array-based comparative genome hybridization (aCGH) and provides the most robust methods to run genome-wide scans to find novel CN changes ¹⁹. Oligonucleotide arrays are used to interrogate SNPs in order to quantify copy number abnormalities and LOH events. Genomic profiling is important for the identification of diagnostic and prognostic factors, new therapeutic targets and for the development of a tumor classification system that could reflect the underlying biology ²⁰. Once candidate genes involved in these alterations are identified, other independent methodology must be performed in order to

confirm preliminary results obtained with this technology and to study pathological mechanism associated with the observed alterations.

1.2.1 ARRAY COMPARATIVE GENOMIC HYBRIDIZATION (aCGH)

Genome-wide DNA profiling, sometimes simply referred to genomic profiling in the context of DNA characterization studies, is a powerful approach for cancer research. It allows to simultaneously identify multiple alterations at genomic level from various specimens in an unbiased manner and short time.

Array CGH has proven to be a specific, sensitive, and fast technique, with considerable advantages compared to other methods used for the analysis of DNA copy number changes and many other applications (Figure 1.3). Array CGH enables analysis of the whole genome in a single experiment and allows the global profiling of copy number imbalances in tumors and an accurate determination of the breakpoints of regions that are gained and/or lost ^{20, 21}.

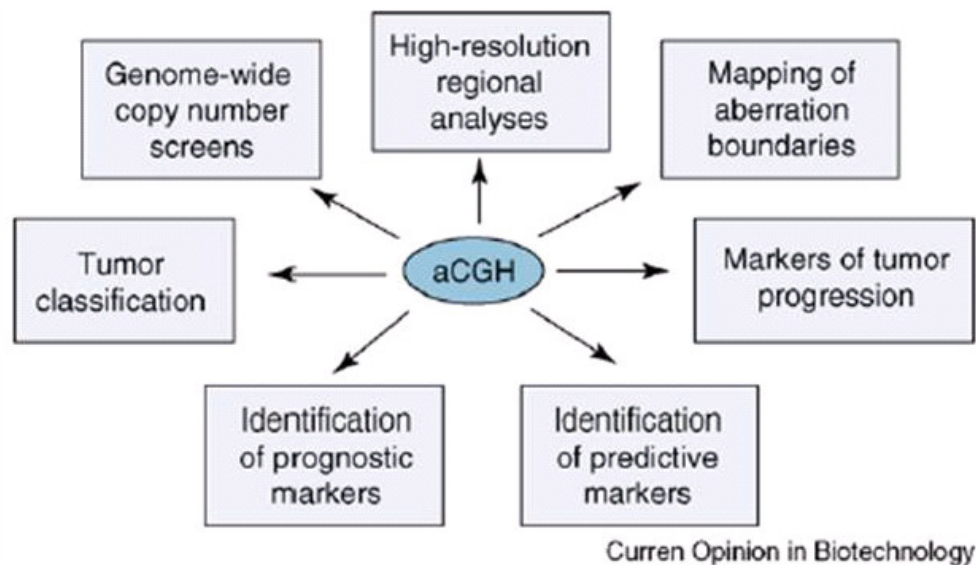


Figure 1.3 Application of aCGH in Cancer (taken from Kallioniemi, A) ²²

Conventional CGH is a molecular cytogenetic technique that detects and maps alterations in DNA copy number of DNA sequences²³. In this approach, the genomic DNAs of test (patient) and reference (normal) samples are differentially labeled. The DNAs are then hybridized to normal human metaphase chromosomes where DNA sequences from both sources compete for their targets. Subsequently the ratio of the hybridization signal intensity of the test and reference sample at any interrogated locus is detected. The assumption is that the relative amount of test and reference DNA bound to a given chromosomal locus is determined by relative abundance of that sequence in the two DNA samples ²¹ (figure 1.4). After measurement of the ratio of the intensities of the two different dyes along the target chromosome, region of impaired content can be detected.

The resolution of conventional CGH is limited by the length of the metaphase chromosomes and is approximately 10 megabases and could contain hundreds of genes. Microarray based CGH has been developed, which combines microarray technology with the CGH approach ^{22, 24}.

In the present work we focused our project on the use of oligoarray CGH platforms. Different oligoarrays combine different labeling and hybridization techniques, and all have yielded high-resolution copy number measurements. Affymetrix sells a commercial oligoarray CGH platform that contains short 25 mer oligonucleotides photo lithographically synthesized on the arrays. These are single-channel arrays, which means that only test DNA needs to be labeled and hybridized. The labeling of the test sample involves a restriction enzyme-based complexity reduction, which precludes the use of suboptimal DNA quality samples (Figure 1.5).

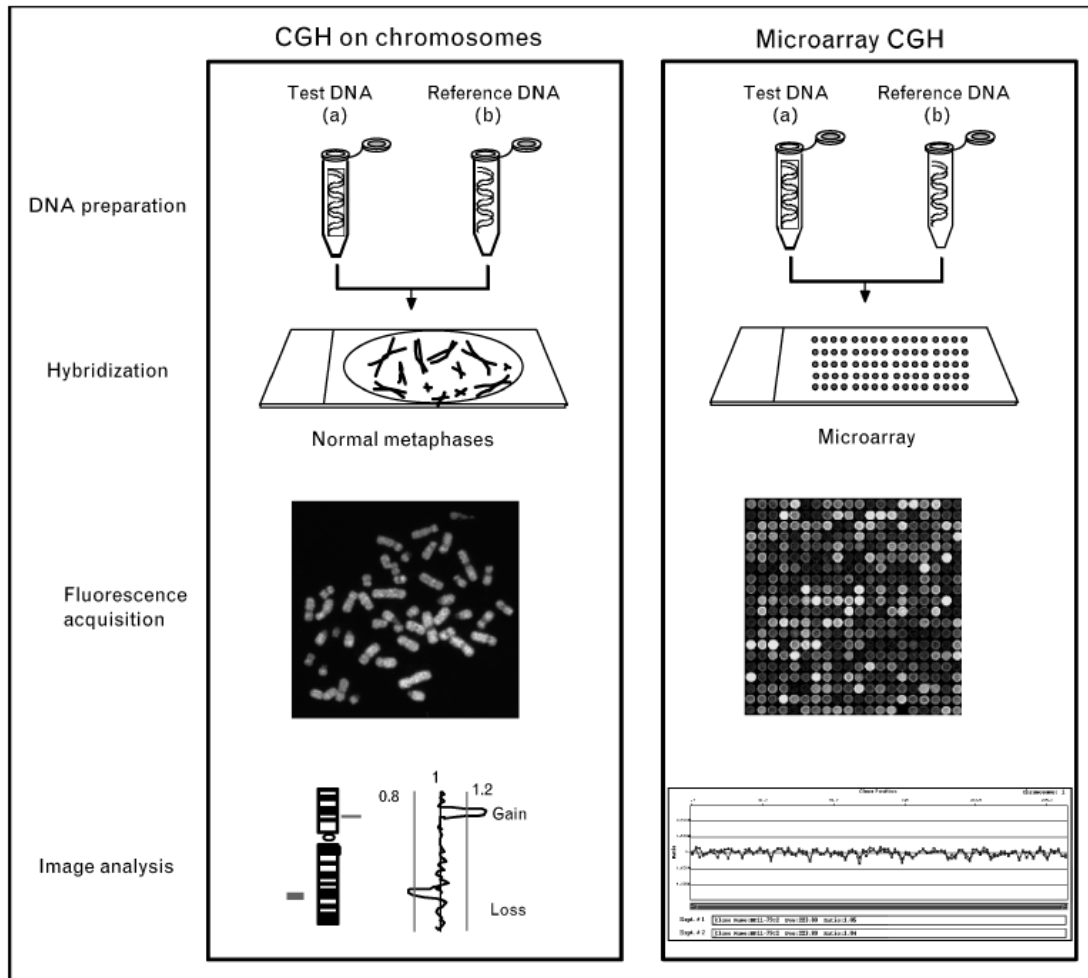


Figure 1.4 CGH methodologies ²⁵. (See text for detailed description)

The hybridization signal intensity (Phycoerythrin fluorescence) is detected by a scanner after having washed away the hybridization buffer from the array. The staining process is designed to amplify the signal of an annealed probe and comprises a first staining with streptavidin, followed by incubation with biotinylated anti-streptavidin antibody and finally with streptavidin-phycoerythrin conjugates.

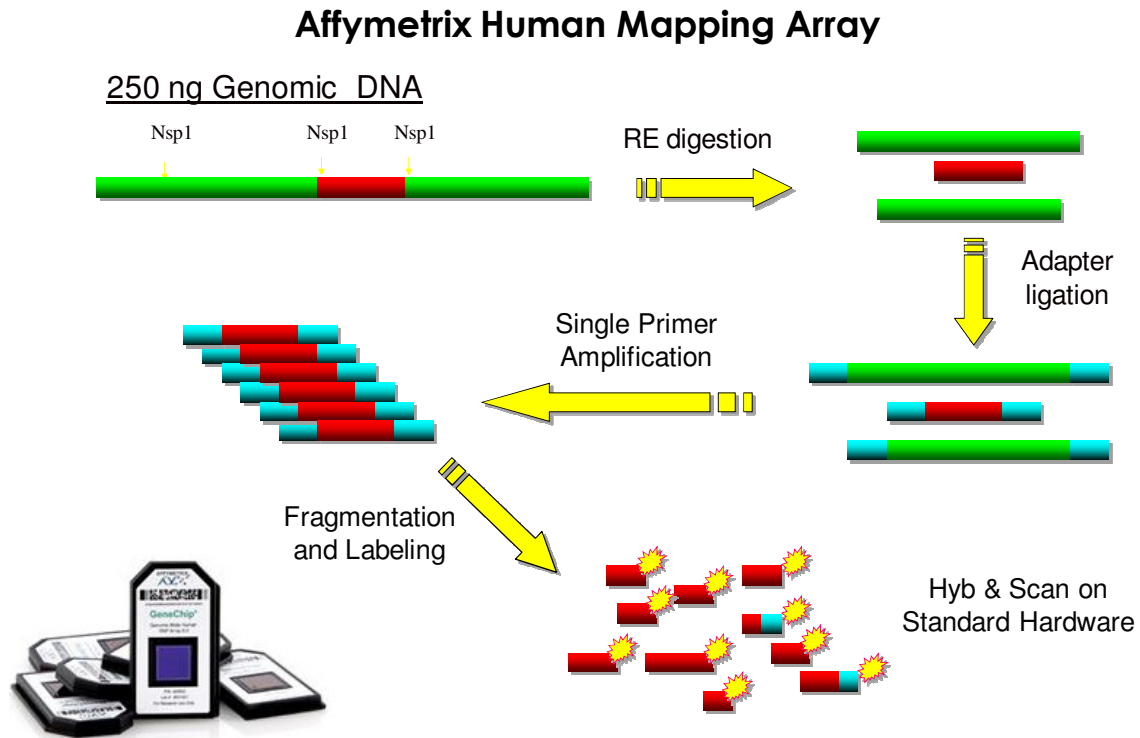


Figure 1.5 Gene Chip mapping assay overview (Affymetrix).

The results are obtained through software applications that compare the acquired hybridization signal intensities corresponding to the analyzed genome with built in average values from normal reference samples.

The technical noise per element on the array is relatively high, which is compensated by the large amount of elements on the array, currently over 2,000,000. The Gene Chip Mapping 500k array set consists of two independent arrays (Gene Chip Human Mapping 250k Nsp and Sty arrays, also used in the present work), that enable the genotyping for more than 500,000 loci with a single primer. Every 250k Array is made of 6.5 million of 5 μm x 5 μm features, each feature consisting of more than one million copies of a 25-mer oligonucleotide probe of defined sequence. For each SNP, redundancy of probes is assured by 24 or 40 different oligonucleotides tiled around the SNP.

Furthermore aCGH is not dependent on the tumor source since tumor DNA can be obtained either from cell lines or from fresh frozen tissues. An important contribution of aCGH to cancer research has thus been in determining putative

locations of cancer genes, especially at chromosomal sites undergoing DNA amplification ²⁵. A large number of sub regional chromosomal gains and DNA amplifications have been discovered by aCGH in some cancers. For example, CGH studies have implicated amplification of oncogenes whose activation in cancer was previously known to occur by chromosomal translocation only, such as amplification of the REL proto-oncogene in non-Hodgkin lymphomas ²⁵. The discovery of novel small regions of loss by aCGH could also help the search for new tumor suppressor genes ^{4, 26-30}. In this way, aCGH is extremely useful in the analysis of the biological basis of tumor progression.

In addition to tumor progression, the high number of CN aberrations has also been linked with poor prognosis, and the identification of specific genetic aberrations associated with patient outcome has been the goal of a number of aCGH studies ³¹⁻³³. Several different aberrations have been implicated in different tumor types and both increased and decreased copy number are found to have prognostic significance. For instance, 17q23.2-qter gains and 17p13.1-p13.3 losses were indicative of poor patient outcome in medulloblastomas ³⁴, whereas gain of 1q or loss of chromosome 13, were associated with poor prognosis in multiple myeloma ^{22, 32}.

Moreover, specific genetic aberrations discovered by aCGH have also been linked to different response to various cancer therapies ³⁵⁻³⁷. Increased number of genetic changes was shown to be associated with therapy resistance in ovarian cancer in which twice as many aberrations were observed in the treatment of the resistant tumor group as compared with the sensitive tumor group ³⁵. Furthermore, data from this study also showed that losses of 1p36.33 and gains of 17q11.2 were found more often in tumors that were resistant to treatment, whereas losses of 13q12-q13 were seen in sensitive cases ²².

1.3 B-CELL LYMPHOID NEOPLASMS

1.3.1 BIOLOGY OF B-CELLS

B cells are small lymphocytes (6–10 μm) and have a dense nucleus and little cytoplasm³⁸.

Their development involves the successive acquisition of properties that are essential for the function of the mature cells along with the loss of properties that are more characteristic of the immature cell.

In human adults, a primary B-cell repertoire, derived from pluripotent hematopoietic stem cells (HSC), is produced in the bone marrow through Immunoglobulin (Ig) gene rearrangements³⁹.

In the bone marrow, there are important external signals produced by the network of specialized non-lymphoid connective-tissue stromal cells that interact intimately with the developing lymphocytes. The contribution of the stromal cells is double⁴⁰. First, they form specific adhesive contacts with the developing lymphocytes by interactions between cell-adhesion molecules and their ligands. Second, they provide soluble and membrane-bound cytokines and chemokines that control lymphocyte differentiation and proliferation, such as stem-cell factor (SCF), a membrane-bound cytokine present on stromal cells that stimulates the growth of hematopoietic stem cells and earliest B-lineage progenitors. Moreover, the chemokine *CXCL12* (*stromal cell-derived factor 1*, *SDF-1*) is also essential for the early stages of B-cell development. It is produced constitutively by the stromal cells, and one of its roles may be to retain developing B-cell precursors in the marrow microenvironment⁴¹.

The B cells leaving the bone marrow are not yet exposed to external antigens and are called 'naïve' or 'virgin' B cells. The lymphocytes circulate through blood stream and are delivered to secondary or peripheral lymphoid organs. The contact between antigen-presenting cells (APC) and B and T cells is ensured in the latter, which normally include spleen, lymph nodes and mucosa-associated lymphoid tissues (MALT)⁴². Antigen binding is required for the final

stages of development. Following antigen recognition, B cells become activated either by themselves or with T cell help. Once activated, B cells proliferate and become enlarged, especially with regard to the cytoplasm. Subsequent cell division leads to the expansion of specific clonal lineages. Moreover, following selection, some of the cells differentiate into antibody-secreting plasma cells (PCs) or later into memory cells. Two types of genetic changes are crucial during differentiation ³⁸. First, somatic hypermutations (SHM) change the antigen binding properties so that B cells can bind antigens more avidly. Second, class switching recombination (CSR), which results from a change in the heavy chain constant region within the Ig, alters the way that Igs on B cells become recognized by effector cells ⁴³.

1.3.2 B-CELLS DIFFERENTIATION AND VDJ-REARRANGEMENT

The earliest B cells in the bone marrow are the progenitor B cells or pro-B cells, which differ from HSC by the expression of B-cell antigens (e.g. *CD19* and *CD79a*) and by their ability to rearrange Ig genes. This phenomenon, called VDJ rearrangement, implies a highly regulated series of genetic events resulting in the membranous expression of a functional BCR ⁴⁴.

The B cell receptor (BCR) is composed of two identical heavy-chain and two identical light-chain Ig polypeptides that are covalently linked by disulphide bridges. Other components of the BCR are the *CD79A* and *CD79B* molecules, which contain cytoplasmic immune receptor tyrosine-based activation motifs.

The variable regions of the Ig heavy chain are assembled from 123 variable (V), 27 diversity (D) and 6 junctional (J) genes present on the Ig heavy-chain gene locus on chromosome 14; the variable regions of the Ig light chains (which can be of κ or λ type) are assembled from V and J elements located at either the *Igk* or *Igl* gene locus on chromosome 2 and 22, respectively ⁴².

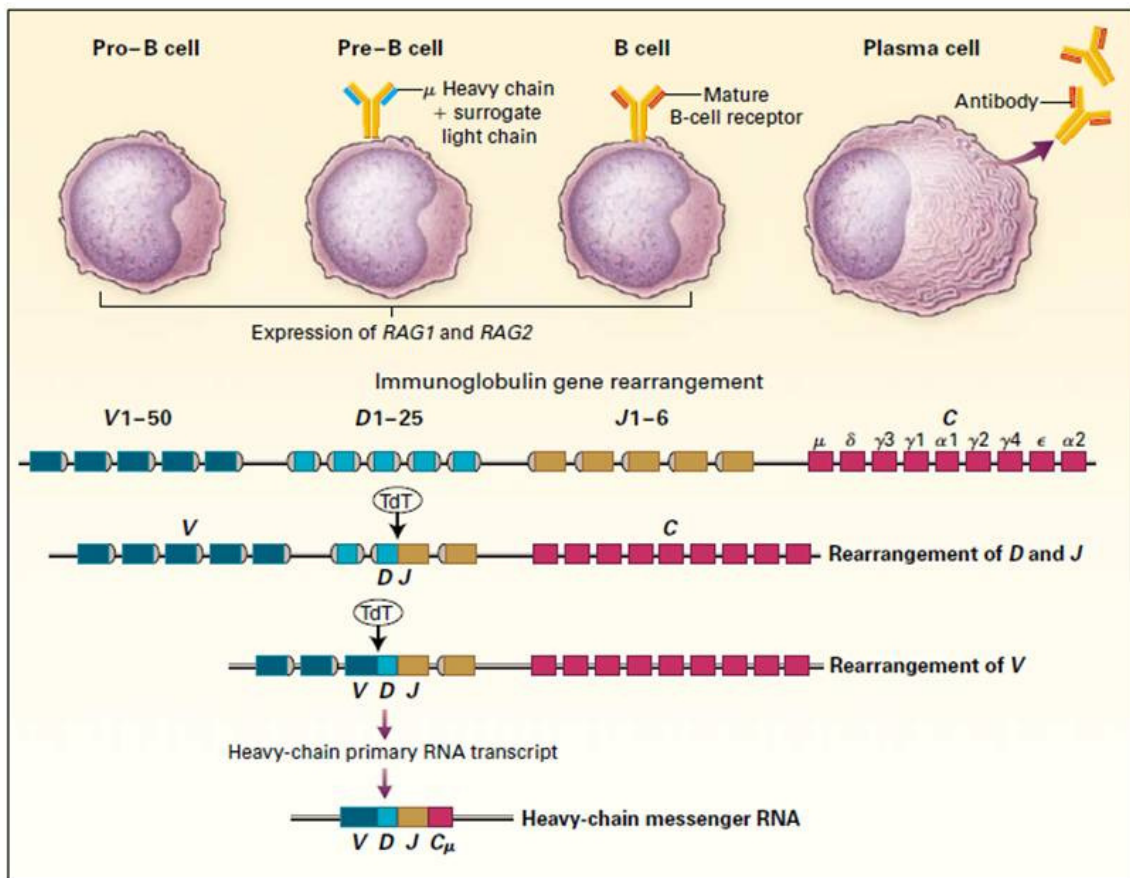


Fig. 1.6. Immunoglobulin heavy chain genes rearrangement during the B-cell ontogenesis
(Modified from Daves, 2000)⁴⁵.

As there are many different V, D and J segments in the germline, each B cell will generate a unique antibody. Ig gene rearrangement is dependent upon the activation of the recombination activating genes *RAG1* and *RAG2* in the early stages of B-cell maturation in the bone marrow ⁴⁶.

A pro-B-cell is transformed into an early pre-B cell once the D and J gene segments are recombined in the Ig heavy-chain gene locus (with the genes for light chains still remaining in germline configuration). Further rearrangements attach one of the V gene segments to the D-J segment and give rise to the late pre-B cell, which expresses a rearranged VDJ-C_μ heavy chain on the cell surface (figure 1.6).

The Stages of B cell Development

	Stem cell	Early pro-B cell	Late pro-B cell	Large pre-B cell	Small pre-B cell	Immature B cell	Mature B cell
							
H-chain genes	Germline	D-J rearranging	V-DJ rearranging	VDJ rearranged	VDJ rearranged	VDJ rearranged	VDJ rearranged
L-chain genes	Germline	Germline	Germline	Germline	V-J rearranging	VJ rearranged	VJ rearranged
Surface Ig	Absent	Absent	Absent	μ chain transiently at surface as part of pre-B-cell receptor. Mainly intracellular	Intracellular μ chain	IgM expressed on cell surface	IgD and IgM made from alternatively spliced H-chain transcripts

Figure 1.7. Stages of B-cells Development (see text for details) (Modified from *Immunobiology, Janeway, 7th edition*) ⁴⁰

Subsequent rearrangement of the Ig light-chain gene locus leads to the surface expression of a complete IgM molecule and to a new B-cell stage designating the immature B cell. The latter subsequently gives rise to the mature B cell that expresses both IgM and IgD at the cell surface as a result of alternative transcriptional termination and splicing of Ig heavy-chain mRNA rather than by DNA recombination ⁴² (figure 1.7). In contrast to immature B cells, mature B cells expressing a functional BCR, may leave the bone marrow and have the capacity of responding to the binding of an antigen by proliferating and differentiating into either PCs or memory B cells (figure 1.7 and 1.8).

As B cells develop from pro-B cells to mature B cells, each step of differentiation is also characterized by the expression of proteins other than Ig. One of the first identifiable proteins expressed on the surface of B cell is *CD45R*. This is a B-cell-specific form of the *CD45* protein; T-cells, monocytes, and neutrophils express other variants of this protein ⁴⁰. *CD45R* is a tyrosine phosphatase required for receptor signaling in B cells. Another protein is *CD19*, which also participates in BCR signaling, and *CD43* that functions both as an adhesion molecule that may guide cell-cell interactions, for example those of B cell precursors with stromal

cells, and also as a signaling molecule, although not as part of the B-cell receptor complex. Kit, the receptor for bone marrow stromal cells, is highly expressed on stem cells and in the late pro-B cell stage and Kit signaling promotes the proliferation of pro- and pre-B cells.

1.3.3 GENETIC CHANGES IN B-CELLS WITHIN GERMINAL CENTERS

During T-cell-dependent antibody responses to exogenous antigen, germinal centers (GCs) are formed by proliferating B cells in the follicles of peripheral lymphoid tissues, including the spleen, lymph nodes, Peyer's patches and tonsils. Naïve B cells first migrate to the T-cell-rich area (known as the T-cell zone) of lymphoid tissues due to their upregulation of specific chemokine receptors in response to their encounter with antigen presented by dendritic cells. In the T-cell zones, B cells become fully activated as a result of their interaction with CD4⁺ T cells and antigen presenting cells ⁴⁷. The initiation of the GC response requires the interaction of co-stimulatory B-cell-surface receptors with ligands expressed by T cells and/or antigen-presenting cells, of which the most important is that between the tumor-necrosis factor (*TNF*)-receptor family member *CD40*, which is expressed by all B cells, and its ligand *CD154* expressed by helper T cells. Activated B cells can then either develop directly into antibody-secreting cells in specialized extra follicular sites of plasmablast growth and differentiation, such as the medullary cords of lymph nodes, or mature into GC-precursor B cells and move to the primary follicle, a structure made of recirculating IgM⁺IgD⁺ B cells within a network of follicular dendritic cells (FDCs). Here, B cells start to proliferate rapidly and push the IgM⁺IgD⁺ B cells aside to form the mantle zone around the GC, yielding a structure known as the secondary follicle. After a few days of vigorous proliferation, the characteristic structure of the GC becomes apparent with a dark zone consisting almost exclusively of densely packed proliferating B cells known as centroblasts, and a light zone comprised of smaller, non-dividing centrocytes situated within a mesh

of FDCs, T cells and macrophages. Centroblasts diversify their IgV genes by somatic hyper mutation (SHM), and those cells that express newly generated modified antibodies are selected for improved antigen binding in the light zone. Some centrocytes eventually differentiate into memory B cells or plasma cells (figure 1.8).

The main characteristic of centroblasts is the high proliferation rate that is required for the generation of large number of modified antibodies, from which the few B cells that display antibodies with improved antigen-binding are selected. Accordingly, gene expression profile analysis has shown that the differentiation of an antigen-activated B cell into a centroblast is accompanied by a dramatic up-regulation of genes associated with cell proliferation and the down-regulation of genes encoding negative regulators of clonal expansion ⁴⁸. Moreover, centroblasts express several pro-apoptotic molecules ^{49, 50}, which allow the rapid execution of cell death by default or in response to exogenous signals. The major benefit of this process is the rapid elimination of GC B cells with newly generated immunoglobulin mutations that produce a non-functional or non-binding antibody. This feature and the spatial concentration of proliferation, mutation and selection of antigen-activated B cells may have the driving force for the evolution of the highly specialized GC structure ⁴⁷.

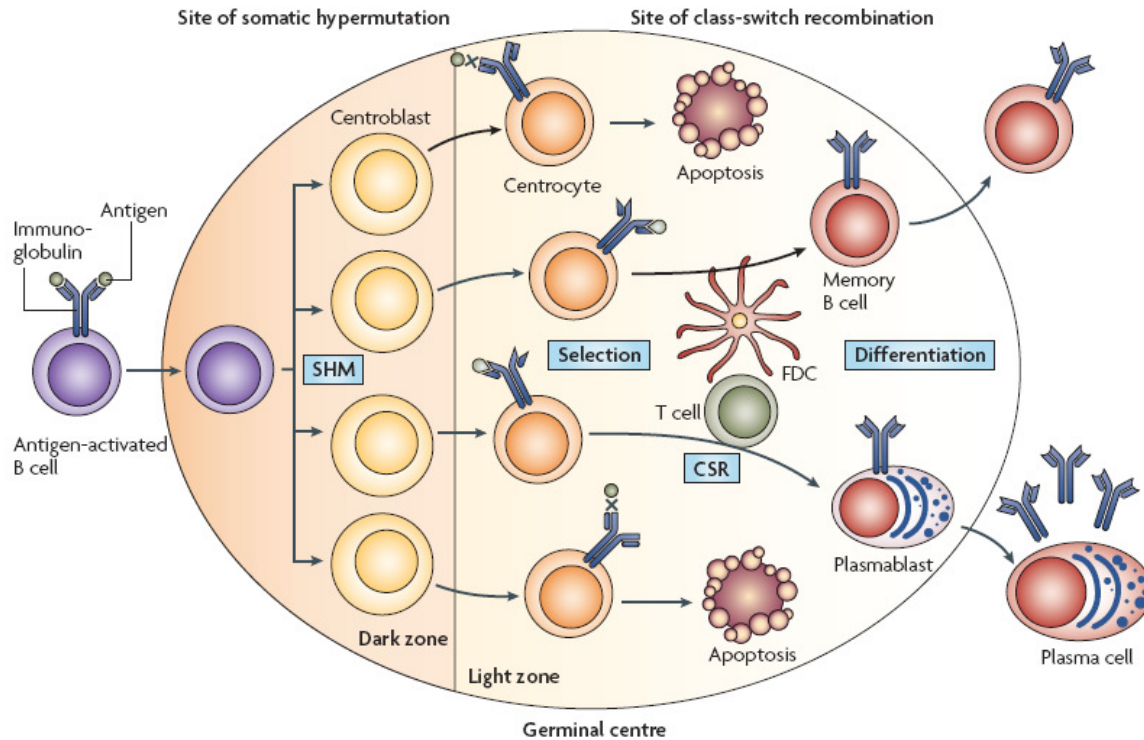


Figure 1.8. Biology of B-cell differentiation (Modified from Klein, U, 2008 ⁴⁷)

The low level of *NF-κB* activation may be due to the fact that the dark zone is largely avoided of T cells that express CD154 ⁵⁰. These observations suggest that centroblasts are not subject to active CD40-signalling. However, the established GCs can be dissolved by inhibiting the CD40-CD154 interaction ⁵¹, suggesting that CD40 stimulation may specifically occur in the light zone, leading to the delivery of survival signals necessary for cell exit from the GC or their recirculation within GC. Very little is known about how the centroblast-specific transcriptional programme is established but some candidate factors are known, include several transcription factors like B-cell lymphoma 6 (*BCL6*), the ETS family member *SPI-B* and interferon-regulatory factor 8 (*IRF8*) ⁴⁷.

1.3.3.1 SOMATIC HYPERMUTATION

The GC, dark zone in particular, is the main structure where SHM occur. SHM is a process that modifies the Ig variable region (IgV) of the rearranged antibody genes in B cells during an immune response (figure 1.9). SHM is associated with DNA strands breaks and introduces mostly single nucleotide exchanges, but also small deletion and duplications (about 6% of all mutation events) ⁵² into the rearranged IgV region. The base changes are distributed throughout V region, but not completely randomly: there are some 'hotspots' of mutation that indicate a preference for characteristic short motifs of four or five nucleotides, and perhaps also for secondary structural features. Some of the mutant immunoglobulin molecules bind antigen better than the original B-cell receptors, and they are preferentially selected to mature into antibody-secreting cells ⁵³.

The SHM requires the activity of activation-induced cytidine deaminase (AID), which is an enzyme that catalyses the initiation event by deaminating cytidines directly on DNA ⁵⁴. Then, error-prone repair that involve different DNA repair factors leads to the introduction of somatic mutations ⁵⁵.

SHM, beside acting on the IgH, κ and λ loci, also affects some non-immunoglobulin genes, most notably the *BCL6* proto-oncogene, the master transcriptional regulator of centroblasts. *BCL6* is a nuclear phosphoprotein that belongs to a large BTB/POZ family of nuclear factors that contain zinc-finger motif and acts as transcriptional repressor ⁵⁶. Centroblasts have the unique ability to tolerate this state of proliferation and genomic instability due in part to *BCL6*-mediated repression of the *ATR* (ataxia telangiectasia and Rad3 related), *TP53* and *p21* genes ⁵⁷. Thus, *BCL6* may allow GC B cells to sustain the physiological genotoxic stress that is associated with high proliferation, and sustain the DNA breaks that are induced by SHM without eliciting the p53-dependent and p53-independent growth arrest ⁴⁷.

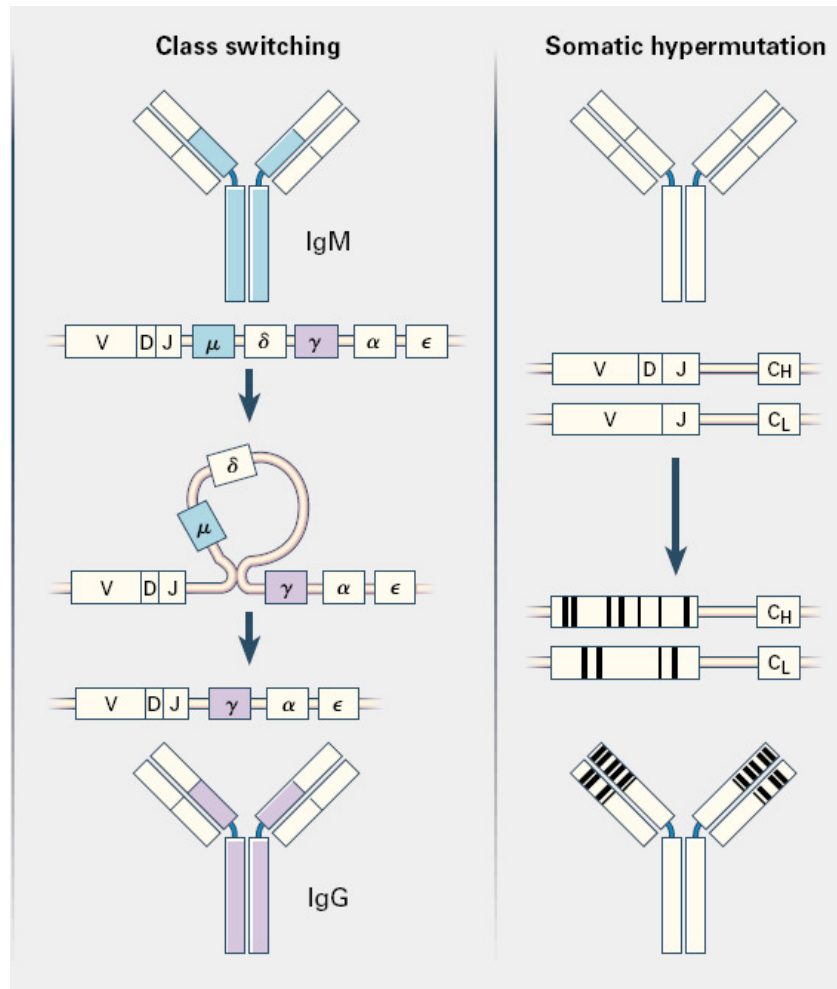


Figure 1.9. Molecular Processes Modifying the Genes Encoding Antibody Molecule (Modified from Kuppers R, 2003⁵³).

1.3.3.2 CLASS SWITCH RECOMBINATION

The light zone of the GC is the site for three main B-cell development processes: the selection of centroblast that produced high affinity antibodies, the process of class switch recombination (CSR) and the initiation of centrocytes into plasma cells or memory B cells (figure 1.9).

Compared with centroblasts, centrocytes are very heterogeneous, because they can undergo diverse developmental stages, including the differentiation back into centroblasts and into post-GC cells with the acquisition of a high-affinity B-cell receptor⁴².

CSR is an irreversible somatic recombination mechanism by which B cells can switch their immunoglobulin class expression from IgM and IgD to either IgG, IgA and IgE. This process alters the effector functions of an antibody that can be specifically required to clear a particular antigen by leaving the antibody specificity defined by the rearranged V(D)J region of the heavy-chain unchanged, but allowing this same region to associate with different constant regions. CSR occurs by DNA recombination involving non-homologous end-joining processes between specific repetitive regions of several hundred base pairs (known as switch regions) that precede the immunoglobulin constant region genes ⁴⁷. This process needs expression of activation-induced cytidine deaminase (AID) and can occur in both T-dependent and T-independent manner, and it is the combination of specific cytokines and co-stimulatory signals that determine the nature of resulting immunoglobulin class. For CSR, *CD40-CD154* interaction is critical, and recent studies demonstrated an important role for IRF4 that seem to regulate *AID* expression through intermediate molecules rather than through a direct mechanism. IRF4 may be up-regulated by the *CD40-NF- κ B* signalling pathway that is activated in centrocytes⁴⁷.

Thus, SHM and CSR are not temporally linked. Also, the processes are spatially separated, with SHM occurring in the dark zone and CSR in the light zone of GCs (figure 1.7).

1.3.4 REGULATION OF B-CELL DEVELOPMENT IN THE HUMAN GC

The GC reaction and the differentiation of B cells into PCs and memory B cells are regulated by a network of transcription factors (Figure 1.10). PC differentiation is most likely initiated by the downregulation of *PAX5*, the 'identity' gene of B cells. This factor seems to be essential for maintaining the identity of mature B cells, including naive, GC and memory B cells ⁴⁷. In addition, an important role of *CD40* stimulation in the differentiation of GC B cells into

memory B cells has been suggested a decade ago based on *in vitro* experiments of purified human GC B cells ⁵⁸. It remains to be determined whether continued stimulation through CD40 directs the centrocytes towards the memory B-cell pathway, whereas the release of CD40 signaling is a prerequisite for their differentiation into PCs *in vivo* ⁴⁷.

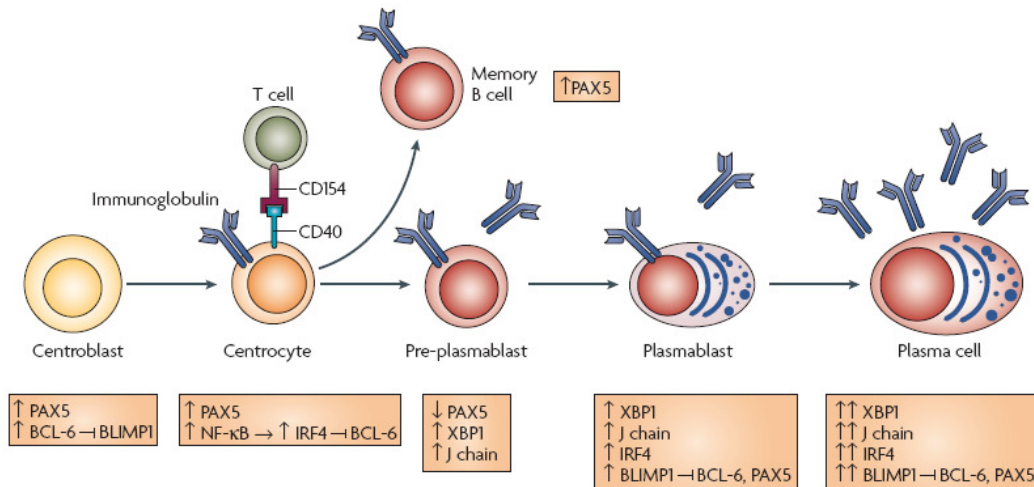


Figure 1.10. Termination of the GC transcriptional programme and post-GC B-cell development (Modified from Klein U, 2008 ⁴⁷).

BCL6, which is strongly up regulated in GC B cells, is the master regulator of the GC reaction by suppressing apoptosis and promoting proliferation.

BCL6 is thought to terminate the GC transcriptional program and allow GC B cells to differentiate into memory B cells or plasma cells. Expression of B lymphocyte induced maturation protein (*BLIMP*)-1, a zinc finger-containing transcriptional repressor, is induced in subsets of GC B cells and in PCs. *BLIMP1* promotes PC development by suppressing genes associated with the GC program (*BCL6*, *PAX5*, *SPI-B*) and induces the expression of those essential for PC development. The X-box binding protein 1 (*XBP*-1), which is up regulated as part of the endoplasmic reticulum (ER) stress response, is expressed at a high level in PCs and is essential for inducing the secretory phenotype of the PCs. More recently, interferon regulatory factor (*IRF*)-4 was identified as an additional master gene in PC differentiation and as an essential regulator of plasma cell differentiation ⁵⁹. *IRF4* has been suggested to act upstream of or in parallel to

BLIMP1 for the generation of PCs. *IRF4* also represses *BCL6*, thereby terminating the GC transcriptional programme. Other works suggested that the expression of *BLIMP1*, *IRF4* and *XBPI* is independently regulated, but that the three transcription factors are all required for the establishment of the terminally differentiated PCs ⁶⁰.

1.4 DIFFUSE LARGE B-CELL LYMPHOMA

1.4.1 MORPHOLOGY

Diffuse Large B-cell Lymphoma (DLBCL) is the most common type of non-Hodgkin's lymphoma (NHL), accounting for 35-40% of all cases.

DLBCL is heterogeneous disease in terms of histological and clinical features, suggesting the existence of biologically distinct subclasses within this entity ⁶¹.

Characteristically, DLBCL express CD45 with the pan-B-cell markers CD20, CD79 and surface Ig, more often IgM or compared to IgA. DLBCL cells variable express one or more B-cell associated antigens: CD5, BCL-2, PAX5, CD138, CD10, BCL6, MUM1/IRF4 ⁶². The last three markers have been proposed to distinguish Germinal Centers (GC)-DLBCL (CD10⁺ or CD10⁻, BCL6⁺ and MUM1 from non-GC DLBCL that are characterized by a different overall survival ^{63, 64}. The presence of CD5 seems to be related with a worse overall survival (OS) ^{65, 66}.

Table 1. DLBCL variants and subtypes

COMMON MORPHOLOGIC VARIANTS	IMMUNOHISTOCHEMICAL SUBGROUPS	DLBCL SUBTYPES	TRANSFORMED DLBCL
CENTROBLASTIC (80%)	CD5-POSITIVE DLBCL	T-CELL/HISTIOCYTE-RICH LARGE B-CELL LYMPHOMA	SMALL LYMPHOCYTIC LYMPHOMA/ RICHTER'S SYNDROME
IMMUNOBLASTIC (10%)	GCB	PRIMARY DLBCL OF THE CNS	LYMPHOPLASMACYTIC LYMPHOMA
ANAPLASTIC	NON-GCB	PRIMARY CUTANEOUS DLBCL, LEG TYPE	EXTRANODAL OR SPLENIC MARGINAL ZONE LYMPHOMA
PLASMABLASTIC LYMPHOMA		EBV-POSITIVE DLBCL OF THE ELDERLY	FOLLICULAR LYMPHOMA

Table 1. DLBCL variants and subtypes

		ALK-POSITIVE LARGE B-CELL LYMPHOMA	
		PRIMARY MEDIASTINAL (THYMIC) LARGE B-CELL LYMPHOMA	
		INTRAVASCULAR LARGE B-CELL LYMPHOMA	
		DLBCL ASSOCIATED WITH CHRONIC INFLAMMATION	
		LARGE B-CELL LYMPHOMA ARISING IN HUMAN HERPESVIRUS	
		PRIMARY MEDIASTINAL (THYMIC) LARGE B-CELL LYMPHOMA	

The expression of *BCL2* protein is seen in 30-50% of the cases and is probably linked with a worse prognosis ^{67, 68}. Expression of CD138 may be seen in cases showing plasmablastic morphology but is rarely seen in other cases ⁶². DLBCL may be primary, or may arise secondarily from indolent B-cell lymphomas (Transformed DLBCL), such as small chronic lymphocytic leukemia (Richter's syndrome), lymphoplasmacytic lymphoma, extra nodal or splenic marginal zone lymphoma, lymphocyte predominant Hodgkin lymphoma, or follicular lymphoma (FL) (figure1.11). DLBCLs can arise in nodal or extra nodal sites ⁶⁹. DLBCL is a lymphoma characterized by the presence of large neoplastic cells, with a diffuse growth pattern. Their morphology resembles the proliferating cells of the GC (centroblastic variant) or immunoblasts. Centroblastic DLBCL is the most represented form, observed in more than 80% of cases, while

immunoblastic DLBCL is represented only in the 10% of cases. The rest of the cases are represented by other variants (T-cell/histiocyte-rich-B-cell lymphoma, Anaplastic B-cell lymphoma, Plasmablastic, etc) ⁷⁰(Table 1).

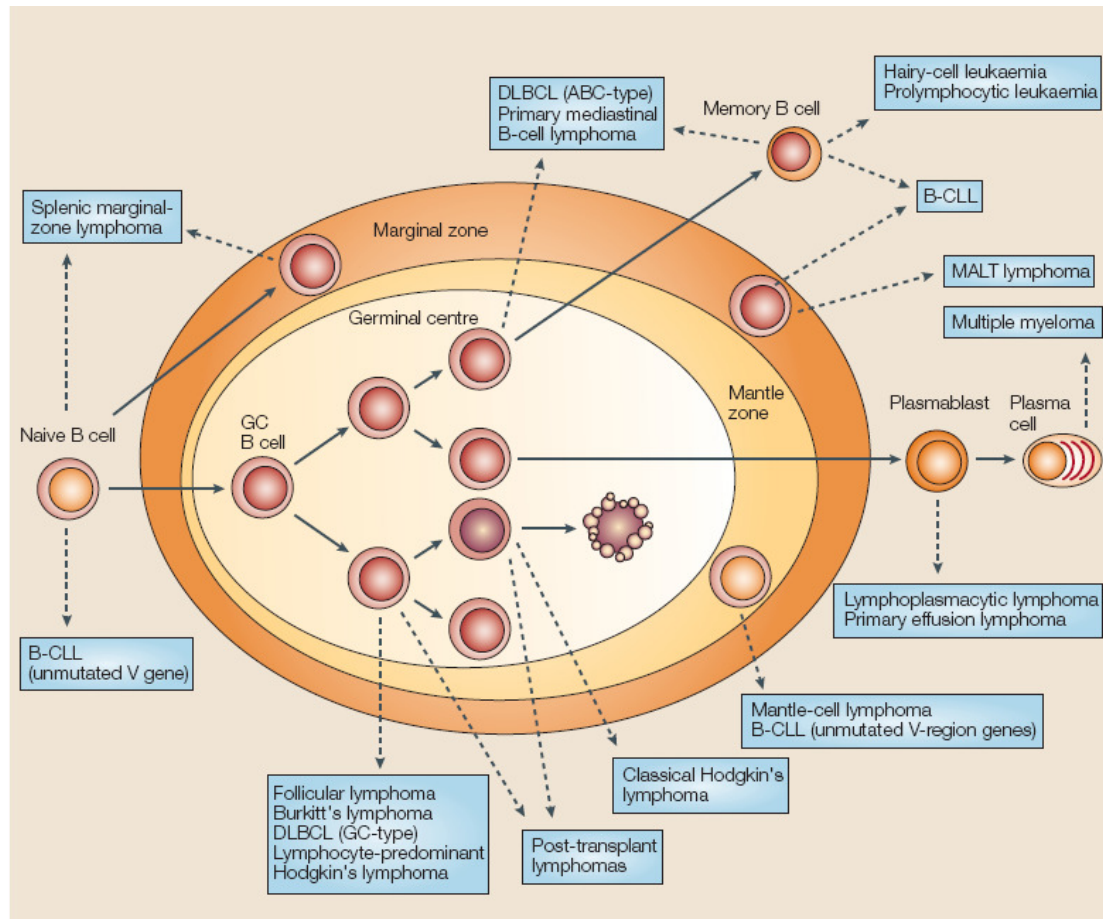


Figure 1.11. Cellular Origin of Human B cell Lymphomas (taken from Kuppers R, 2005 ⁴⁴).

1.4.2 GENOMIC ABERRATIONS

The most frequently detected genetic alteration in DLBCL involves a rearrangement of the *BCL6* gene in chromosomal band 3q27 with various translocation partners, resulting in promoter substitution and *BCL6* deregulation, which can be detected in 30% to 40% of the cases ⁷¹. Alternatively, in some other cases *BCL6* expression can be deregulated by mutations in the 5' regulatory sequences of the gene. *BCL6* targets a variety of genes, including genes involved in the cell cycle control and maintenance of genomic stability, such as *CDKN1B/p27*, *CCND2/cyclin D2*, and *TP53*, allowing B-cells to sustain the physiologic occurrence of DNA breaks during the GC reaction ⁷². Furthermore, *BCL6* represses expression of genes required for terminal differentiation, such as *PRDM-1* and its gene product *BLIMP1* ⁷³. Deregulation of *BCL6* expression might therefore contribute to lymphomagenesis by blocking terminal differentiation and providing resistance to apoptotic signals in response to DNA damage.

Array-CGH-based analysis of a large series of DLBCL revealed as common alteration losses of two regions in chromosome 6q (6q21-q23.3 and 6q16) comprising *PRDM-1* and *TNFAIP3* genes and gains in 18q, gains and amplifications in 2p, and gains in chromosomes 3q and 6p ⁴.

A translocation of the *BCL-2* gene in chromosomal band 18q21 to the Ig locus in 14q32, the hallmark translocation of FL, is found in approximately 20% of DLBCL, and the typical translocation t(8;14) involving the *MYC* gene locus in 8q24 is present in approximately 5% of DLBCL. Amplification of the *BCL-2*, *MYC*, and *REL* gene loci are found in around 20% of DLBCL ⁷⁴. Inactivation by somatic mutations of *TP53*, a multifunctional transcription factor that plays a major role in the DNA damage response by mediating cell cycle arrest, DNA repair, or apoptosis, can be detected in approximately 20% of DLBCL and is associated with the clinical course of GCB subtype ⁷⁵(Figure 1.12).

In addition to the occurrence of chromosomal imbalances, there is increasing evidence that epigenetic changes, such as DNA methylation and histone modification, may play an important role in DLBCL pathogenesis. Few studies

exist in malignant lymphomas and most published reports focus on the transcriptional silencing of putative tumor suppressor genes through CpG island hypermethylation. In DLBCL, inactivation of the *HIC1* gene by deletion and promoter hypermethylation was shown to be associated with an inferior clinical outcome ⁷⁶ and a very recent study investigated DNA hypermethylation levels of more than 500 CpG islands in DLBCL and uncovered potential target genes that may be transcriptionally regulated by CpG hypermethylation ^{77,78}.

Another important mechanism in the transformation process of DLBCL might be the aberrant targeting of a variety of proto-oncogenes by the process of somatic hyper mutation, which is not detectable in normal GC B-cells or in other GC-derived lymphomas. Aberrant somatic mutation of the proto-oncogenes *PIM1*, *MYC*, *RHOH/TTF*, and *PAX5* can be detected in approximately 50% of DLBCL, resulting in nucleotide changes preferentially in the 5' sequences of these genes with potentially oncogenic effects ⁷¹.

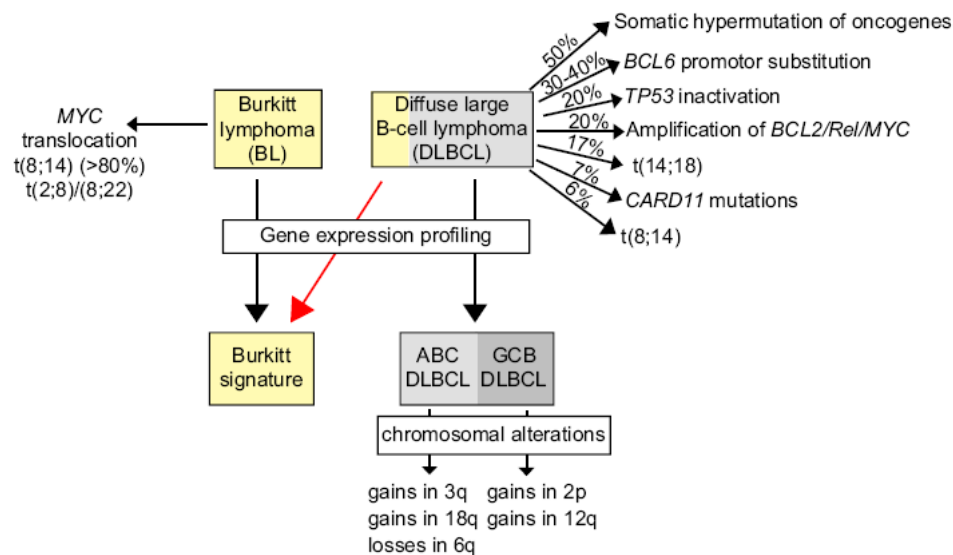


Figure 1.12. Molecular Features of DLBCL (Takken from Hartman, E M) ⁶¹.

1.4.3 BIOLOGY AND MOLECULAR PATHOGENESIS

DLBCL arise from normal lymphocytes at different stages of B-cell differentiation (as shown in figure 1.11).

At distinct stages of development, B-cells differentiation are characterized by the particular structure of the BCR and expression patterns of differentiation markers. Analysis of these features was used to determine the origin of various human B-cell lymphomas. The rationale for such a classification of B-cell lymphomas is based on the observation that malignant B cells seem to be 'frozen' at a particular differentiation stage, which reflects their origin. One of the main concepts emerging from these studies is that most types of B-cell lymphoma are derived from GC or post-GC B cells ⁴⁴.

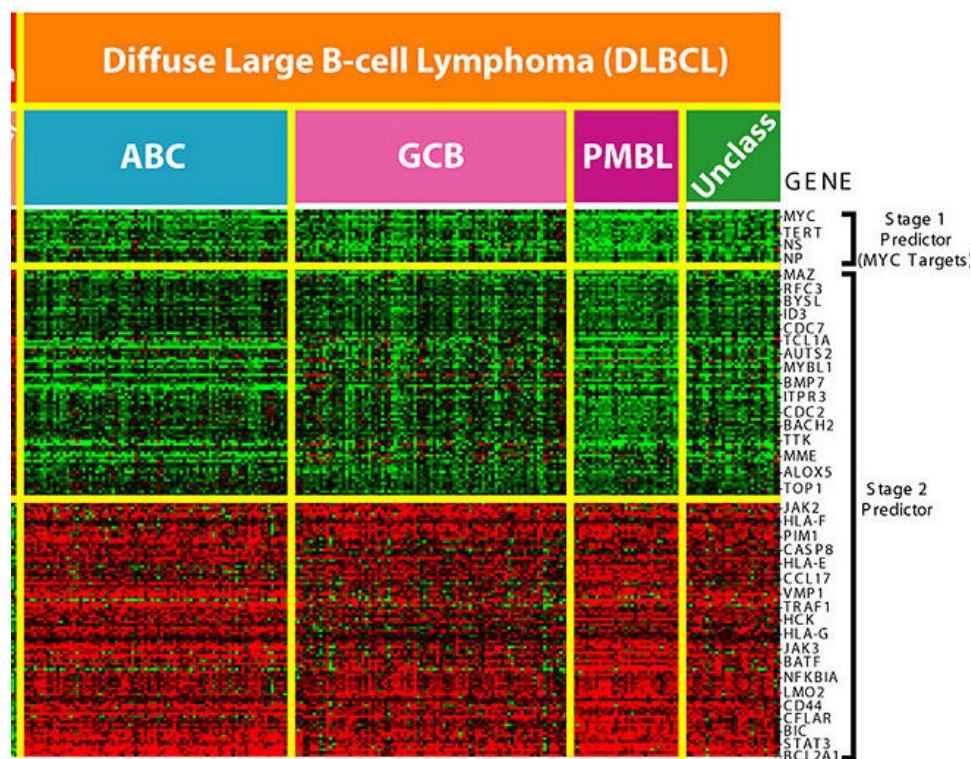


Figure 1.13. Molecular subtypes of DLBCL: ABC, GCB and PMBL. Green color represents downregulated genes while red color represents upregulated genes (modified from www.wrongdiagnosis.com/l/lymphoma/wiki.htm).

As DLBCL is a very heterogeneous disease, Gene-expression profiling studies allowed the identification of two main DLBCL subtypes, one with a profile resembling GC B cells (GCB-type), and the other resembling in-vitro-activated B cells (ABC-type)⁷⁹⁻⁸¹. A third type of DLBCL was identified as primary mediastinal B-cell lymphoma (PMBCLs) and is believed to derive from post-GC B cells of the thymus^{44, 82} (figure 1.13). Importantly, patients with these DLBCL subtypes have different survival rates⁸³, indicating that the subgroups represent distinct biological disease entities⁴.

Table 2. Identification of DLBCL subtype-specific genomic aberrations by aCGH (modified from Lenz G, 2008⁴)

ABC-DLBCL	GCB-DLBCL	PMBL
TRISOMY CHR 3/GAIN 3p (FOXP-1)/short 3q AMPL (NFKBIZ)	1p LOSS (TP73) TRANSLOCATION 3q27.3 (BCL-6)	6q LOSS
6q ARM LOSS (TNFAIP3, BLIMP-1)	2p AMPL (REL)	9p GAIN (JAK2)
9p LOSS/HOMOZYGOUS DEL (CDKN2A, CDKN2B, INK4a/ARF)	10q LOSS (PTEN)	20p GAIN
18q ARM GAIN (BCL-2, NFATC1, MALT1)	12q GAIN (MOM2)	
19q ARM GAIN (SPI-B)	13q AMPL (MIHG1, miR 17-92)/13q LOSS (ING1)	

GC-type DLBCL harbors the expression pattern of genes characteristic of normal GC B-cells (GC signature genes), such as *CD10*, and signaling molecules and transcription factors such as *BCL6*. Typical alterations of GCB DLBCL are the translocation t(14;18)(q32;q21) and the amplification of the *REL* locus in 2p.11

Conventional and array-based CGH analysis ^{4, 63, 81} revealed more frequent gains of a chromosomal region in 12q, which might cause upregulation of MDM2 gene in GCB-DLBCL, encoding a negative regulator of the tumor suppressor p53. Moreover amplification of MIHG1 (mir-17-92) microRNA on chromosome 13, deletion of PTEN tumor suppressor gene on chromosome 10, have been described as characteristic features of the GCB DLBCL ^{4, 63, 81} (table 2; figure 1.14).

It has been very recently reported a recurrent somatic mutations affecting the polycomb-group oncogene EZH2, which encodes a histone methyltransferase responsible for trimethylating Lys27 of histone H3 (H3K27). The mutations, which result in the replacement of a single tyrosine in the SET domain of the EZH2 protein (Tyr641), occur in almost 20% of GCB-DLBCLs and 7% of FLs and are absent from ABC-DLBCLs ⁸⁴.

The ABC-type tumors express genes characteristic of in vitro activated peripheral blood B-cells (ABC signature genes) as well as some genes normally expressed by PCs and mainly, a large number of NF- κ B target genes, thus suggesting their post-GC origin ⁸². NF- κ B is usually retained in an inactive form in the cytoplasm, by binding to a member of the I κ B family proteins, I κ B- α . In response to signaling, due to different pathways, I κ B- α is phosphorylated by I κ B kinase complex (IKK) and degraded by the ubiquitin-proteasome pathway. This leads to the activation of NF- κ B. Only ABC-type exhibits constitutive activation of IKK complex and cell lines of this DLBCL subtype can be killed by blocking NF- κ B, suggesting that this pathway may represent a new therapeutic target for this prognostic unfavorable subgroup ⁸⁵ (figure 1.14).

ABC-DLBCLs carry more frequent gains in chromosomes 3q and 18q, as well as losses in 6q. Interestingly, the deleted region in 6q includes the *PRDM-1* gene locus (6q21- q22.1), whose gene product *BLIMP1* is a target of *BCL6*. Indeed, mutational analysis of *PRDM-1/BLIMP1* revealed its inactivation by mutations in a considerable proportion of ABC-DLBCL, suggesting that *BLIMP1* inactivation may play a pathogenetic role in this subset of DLBCL ⁷³.

To investigate whether constitutive NF- κ B activation in ABC-DLBCL represents a primary pathogenetic event or reflects the intrinsic program of the tumor cell of origin, Compagno M et al performed a screening of the mutational status of the complete coding sequence of 31 NF- κ B-pathway genes in a subset of DLBCL samples. Genes found mutated after filtering for known polymorphisms and synonymous mutations were further analyzed in a validation panel composed of 87 DLBCL (23 ABC, 44 GCB and 20 unclassified/non-GC). This strategy identified a total of 48 sequence changes distributed in 6 different genes, including the NF- κ B negative regulator A20 (*TNFAIP3*) and the positive regulators *CARD11*, *TNFRSF11A* (*RANK*), *TRAF2*, *TRAF5* and *MAP3K7* (*TAK1*)²⁷.

The frequent 6q23.3 loss involving *TNFAIP3* locus (containing the A20 gene) was predominantly associated with ABC-type²⁷. The A20 gene, which encodes an ubiquitin-modifying enzyme involved in termination of NF- κ B responses, is most commonly affected, with 30% of patients displaying biallelic inactivation by mutations and/or deletions. When reintroduced in cell lines carrying biallelic inactivation of the gene, A20 induced apoptosis and cell growth arrest suppressing NF- κ B activation and indicating a tumor suppressor role²⁷.

Functional studies of GC and ABC-DLBCL cell lines revealed additional molecular differences. For example, phosphodiesterase 4B (*PDE4B*), an inactivator of cyclic AMP (cAMP) which mediates apoptosis via *AKT* inactivation, is highly expressed in ABC-type. The activation of the same pathway in GC-type cell lines would lead to an increase of cell death, whereas there's no effect on ABC-type cell lines. Also the response to IL-4 seems to be different between the two DLBCL subtypes⁴.

Another characteristic of the ABC-DLBCL lesion is an amplicon on chromosome 3 that includes *NFKBIZ*, which encodes an *I κ B*-like protein that binds to NF- κ B heterodimers and enhances transactivation of some NF- κ B targets, such as IL-6. *NFKBIZ* is a particularly intriguing target gene, given the already mentioned constitutive activation of the NF- κ B pathway in ABC-DLBCL and the important role of IL-6 signaling through *STAT3* in a subset of ABC-DLBCLs. Gains of chromosome arm 3p matched with elevated mRNA levels of the oncogene

FOXP1, are observed in ABC-DLBCL and never observed in GCB-DLBCL ⁴. Similarly, gain of 18q21-q22 occurs in ABC-DLBCL cases, but less frequently in GCB-DLBCL (10%) and PMBCL (16%) cases. The two most up regulated genes in this locus are *BCL-2* and *NFATC1*.

	GCB DLBCL	ABC DLBCL
Cell of origin	Germinal center B-cell	Postgerminal center B-cell
Oncogenic mechanisms	BCL2 translocation Amplification of the CREL locus	Constitutive NFkB activation
Ongoing Ig mutation	Yes	No
Immunohistochemistry	Expression of CD10, Bcl6	Expression of MUM1/IRF4
Clinical outcome	Favorable	Poor

Figure 1.14. Molecular characteristics of GCB and ABC (taken from Rosenwald A, 2008 ⁶¹).

Activating mutations of *CARD11* have been identified in an ABC-DLBCL subset (9.6% of cases reported by Lenz G, 2008 ⁴) that likely contribute to the constitutive activation of the NF- κ B pathway in these cases. *CARD11* is a protein that participates in the constitutive NF- κ B activation in ABC-DLBCL and its downregulation by small hairpin RNAs, was toxic to ABC-DLBCL, but not to GCB-DLBCL cell lines. Therefore, *CARD11* might represent a novel, promising molecular target for a subset of DLBCL ²⁸.

The *SPI-B* locus has been shown to be translocated and inserted in the proximity of the Ig 3'a-enhancer in an ABC-DLBCL cell line, resulting in relatively high *SPI-B* levels. The same study showed that *SPI-B* levels were higher in ABC-DLBCL patients than in the GCB-DLBCL ones. RNA interference of *SPI-B* in an ABC-

DLBCL cell line identified the requirement of *SPI-B* for proliferation and cell survival, reinforcing the notion that *SPI-B* has oncogenic potential ⁴ (figure 1.14). In addition, knockdown of proximal BCR subunits (IgM, Igk, CD79A and CD79B) killed ABC-DLBCLs with wild type CARD11 but not other lymphomas. The BCRs in these ABC-DLBCLs formed prominent clusters in the plasma membrane with low diffusion, similarly to BCRs in antigen-stimulated normal B cells. Somatic mutations affecting the immunoreceptor tyrosine based activation motif (ITAM) signaling modules of CD79B and CD79A were detected frequently in ABC-DLBCL biopsy samples but rarely in other DLBCLs. In 18% of ABC-DLBCLs, one functionally critical residue of CD79B, the first ITAM tyrosine, was mutated ⁸⁶. These mutations increased surface BCR expression and attenuated Lyn kinase, a feedback inhibitor of BCR signaling. These findings establish chronic active BCR signaling as a new pathogenetic mechanism in ABC-DLBCL, suggesting several therapeutic strategies. Taken together, these observations demonstrate that NF- κ B activation in DLBCL is caused by genetic lesions affecting multiple genes, the loss or activation of which may promote lymphomagenesis.

Regarding PMBL, frequent gains, including amplifications of chromosomal region 2p14-p16, have been found in up to 47% of cases ^{63, 87, 88}(table 2). This region of the genome may however also be amplified in other B-cell diseases, including GCB-DLBCL and 50% of classical Hodgkin's lymphoma (cHL) ⁸⁹. It is interesting to note that the latter is closely related to PMBL in terms of gene expression profile analysis ^{88, 90}. Moreover a characteristic aberration of PMBL type seems to be the amplification on chromosome 9p, possibly affecting *JAK2* gene and interestingly similar lesions have also been detected in Hodgkin lymphoma.

1.4.4 THERAPY

Patients are treated with the polychemotherapy regimen cyclophosphamide doxorubicin vincristin, prednisone combined with the monoclonal antibody anti CD-20 Rituximab (R-CHOP)⁹¹. GC-DLBCL type are characterized by a more favorable prognosis in terms of overall-survival (60% of 5-years survival) compared to ABC-type (only 35% 5-years survival) after R-CHOP treatment (see figure 1.14). It was also recently reported by G. Lenz and colleagues that ABC-DLBCL patients carrying trisomy of chromosome 3 or deletion of the INK α /ARF (chromosome 9p) or both aberrations, had a worse clinical outcome compared to non aberrant ABC-DLBCL when treated with R-CHOP⁴. Moreover, the same group, performed a gene expression profile study developing a multivariate gene-expression-based survival-predictor model on 233 patients with DLBCL who received R-CHOP⁹².

The multivariate model created from three gene-expression signatures — termed “germinal-center B-cell,” “stromal-1,” and “stromal-2” — predicted survival in patients who received R-CHOP, showing a higher overall and progression-free survival rates in GC-DLBCL cases than those with ABC-DLBCL⁹². The study also indicated the neoangiogenesis as a possible new therapeutic target for DLBCL.

Although the reported response rate to this chemotherapy regimen is very high, 30-40% of the cases do not still respond to it. The emergence of new strategies, including dose dense chemotherapy approaches and the identification of new rational therapeutic targets by gene expression profiling, has resulted in improvements in outcome for patients with this disease in recent years⁹³.

2. MATERIALS AND METHODS

CLINICAL SAMPLES

Frozen tumoral biopsies of 166 DLBCL patients, taken at diagnosis, have been collected. Cases were selected only based upon the availability of frozen material and for having a fraction of malignant cells in the pathologic specimen representing >70% of overall cellularity as determined by morphologic and immunophenotypic studies. Cases of DLBCL transformed from low grade lymphomas, of PMBL, HIV-related DLBCL and post-transplant DLBCL were excluded.

DNA SAMPLES

Genomic DNA was isolated using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany), following the manufacturers instructions. DNA samples were quantified by spectrophometric measurements using a Nanodrop (Nanodrop technologies, Wilmington, DE, USA). DNA concentration was adjusted to 50 ng/μl with reduced EDTA TE Buffer. DNA integrity was verified by electrophoresis of at least 200 ng total genomic DNA on a 1% agarose gel prepared and run in TBE buffer. After electrophoresis, DNA was stained in aqueous solution of ethidium bromide (Applichem GmbH, Darmstadt, DE) (0.5 μg/ml) and visualized using the Alphaimager 3400 (Alphainnotech Corporation, Fremont, CA, USA). Genomic DNA degradation was detected as a smear given by DNA fragments of variable lengths. Degraded genomic DNA samples were discarded. In contrast, high quality genomic DNA supported as a major band at approximately 10-20 kb.

GENE CHIP MAPPING 500K ASSAY PROTOCOL

250 ng of genomic DNA (50 ng/μl) were digested with a mix prepared on ice containing 1x NE buffer 2 (New England, Biolabs, MA, USA), 1x BSA (New England, Biolabs), 0.5 U/ μl NspI (New England, Biolabs) in a final volume of 20 μl. This digestion mix was incubated at 37°C for 120' and 65°C for 20' in a MJR Thermal Cycler 200 (MJ Research, Cambridge, MA, USA).

Ligation mix contained 1.5 μM Adaptor Nsp I, 1x T4 DNA Ligase buffer (New England, Biolabs) and 32 U/ μl T4 DNA Ligase (New England, Biolabs), and in the ligation program was 16°C for 180' and 70°C for 20'. PCR amplification of ligated DNA fragments was performed in three replicates for each sample (four replicates in case of low yield). The PCR master mix was prepared on ice with the Titanium DNA amplification Kit (Clontech Laboratories, Inc./Takara Bio Company, CA, USA) and contained 1x Clontech Titanium Taq PCR Buffer, 1M GC-Melt, 350 μM each dNTP, 4.5 μM PCR Primer 002, and 1x Clontech Titanium Taq DNA Polymerase in a final volume of 100 μl with 10 μl of diluted ligated DNA for each sample. The PCR program, run on a MJR thermal cycler 200 with heated lid, was : denaturation at 94°C for 3', followed by 30 cycles of 90°C for 30", 60°C for 30", and 68°C for 15", then finally 68°C for 7'. PCR products were purified using the DNA Amplification Clean-up Kit (Clontech, Takara Bio Europe) using a QIAvac Multiwell unit (Qiagen). Briefly, a Clean-up plate was connected to a vacuum source and 8 μl of 0.1 M EDTA were added to each PCR reaction and all PCR reactions for each sample were consolidated into one well of the plate. With application of a vacuum the wells were dried and washed three times with AQ-gene water. The vacuum was released and the Purified PCR products were eluted adding 45 μl of RB buffer to each well and shaking the plate at room temperature for 10'. The purified PCR products were recovered by pipetting the eluate out of each well and quantified with the Nanodrop (Nanodrop technologies). 90 μg of purified PCR product in a volume of 45 μl (RB Buffer) were used for the next fragmentation step. Fragmentation was performed in a final volume of 55 μl containing 1x Fragmentation buffer and 0.25 U/ μl Fragmentation reagent. Fragmentation was performed at 37°C for 35', followed by DNase inactivation step at 95 °C for 15'. DNA fragments were then labeled with 0.857 mM GeneChip DNA labeling Reagent, 1x TdT Buffer and 1.5 U/ μl Terminal deoxynucleotidyl Transferase (TdT). The mix was prepared on ice and incubated at 37°C for 240', and 95°C for 15'. The hybridization cocktail was prepared on ice mixing 70 μl of fragmentated and labeled DNA target with 56 mM MES, 5% DMSO (Sigma Aldrich Chemie GmbH), 5.77 mM EDTA (Ambion

Euopr Ltd., Huntingdon, UK), 0.115 mg/ml Herring Sperm DNA (Invitrogen), 1x Oligonucleotide Control Reagent, 11.5 mg/ml Human Cot-1 (Invitrogen), 0.0115% Tween 20 (Sigma Aldrich Chemie GmbH), 26.9 M TMACl (Tetramethyl Ammonium Chloride) (Sigma Aldrich Chemie GmbH), to a final volume of 260 µl. The hybridization mix was denatured at 99°C for 10' in a Heat block, cooled down on ice for 10" and centrifugated at 2000 rpm 1'. After an incubation at 49°C for 1', 200 µl of denatured Hybridization mix were injected into the GeneChip Human Mapping 250k Array Nsp I, that was then placed in a GeneChip Hybridization Oven 640 (Affymetrix), at 60 rpm and incubated for 16-18 hours.

Buffers and solutions for array hybridization, washing and staining were prepared according to Chapter 5 of the GeneChip Mapping 500K Assay Manual and, when recommended, were filtered with the stericup vacuum (Millipore) driven disposable filtration system (Millipore AG). To wash, stain and scan a probe array, we first registered the experiment in GCOS. Briefly, each experiment was assigned with a name, the probe array type used, the name and other specifications of the sample, the name of the project.

Washing and staining were performed with Genechip Fluidics Station 450. The scripts for the fluidics station were downloaded from the Affymetrix web site at http://www.affymetrix.com/support/technical/fluidics_scripts.affx.

We followed the fluidics protocol "Mapping500k1v2_450". The arrays were finally scanned using the GeneChip scanner 3000 7G. Library files "Mapping500k_Nsp", required to scan and analyze the 250k Nspl arrays, were downloaded from the Affymetrix web site. GCOS automatically generated cell intensity data (CEL files) from each array.

DATA ANALYSIS

Data acquisition was performed using the GCOS v1.4 and Genotyping Analysis Software (GTYPE) v4.1. CHP files with raw perfect match and mismatch intensities and genotype calls of each interrogated SNP on each array were generated applying the GTYPE "Batch Analysis" tool to the raw intensity data

(CEL files) of all the processed arrays. The GTYPE Batch Analysis calculation is based on Bayesian Linear Model with Mahalanobis distance classifier (BRLMM). CHP files constitute the input to the CN analysis pipeline by the GTYPE Chromosome Copy number Analysis Tool (CNAT) v4.01. The modified Bayesian Piecewise Regression (mBPCR) method⁹⁴ was used to estimate the profile of the log2-ratio of CNs, starting from their raw values obtained with CNAT v4.01. mBPCR is a regression procedure which assumes that the data are noisy observations of a piecewise constant function. The latter represents the true log2-ratio profile, since we can divide the genome in regions of constant CN. Thus, the algorithm estimates, in a Bayesian way, the number of segments in the profiles, their endpoints (also called breakpoints) and the value of the log2-ratio in each interval. After normalization of each profile to a median log2-ratio of zero, thresholds for loss and gain defined as six times the median absolute deviation symmetrically around zero with an associated p lower than <0.001 after Bonferroni multiple test correction. LOH profiles were obtained applying the method with haplotype correction for tumor-only LOH inference available in the dChip software. This algorithm is a hidden Markov model⁹⁵, where for each SNP the probability to belong to a LOH region depends on both the probability of homozygosity in a normal population and the probability to be influenced by the status of the previous SNP in the sequence. If two SNPs are close, it is supposed that they are more likely to have the same LOH status.

To identify the minimal recurrent alterations that are supposed to contain loci relevant for the tumor, we applied the algorithm developed by Lenz et al⁶. Four distinct types of minimal common regions (MCR) were defined: short recurrent abnormality (SRA), long recurrent abnormality (LRA), abnormal chromosome arm (ACA) and abnormal whole chromosome (AWC). These types were derived as follows. A sample was defined as having an abnormal chromosome arm of a particular class (e.g., deletion or gain) if abnormal segments of that class covered more than 60% of the arm in that sample. An arm was defined as wild type if wild-type segments covered more than 95% of the arm. A sample was defined as having an abnormal whole chromosome if it was found to be

abnormal in both arms. A sample was defined as wild type for an entire chromosome if it was wild type for both chromosome arms.

For MCR occurring in at least 15% of cases, differences in frequencies between subgroups were evaluated using a Fisher's exact test followed by multiple test correction (MTC). The commonly affected regions were compared with the Database of Genomic Variants (<http://projects.tcag.ca/variation/>): regions showing an overlap above 80% between probes and known copy number variations (CNV) were considered bona fide CNV.

The Database for Annotation, Visualization and Integrated Discovery (DAVID) 2008 web-accessible program was used to map genes contained within MCR, retrieved using the web-accessible Galaxy (<http://galaxy.psu.edu/>), to KEGG and Biocarta Pathways, and statistically examine the enrichment of gene members for each pathway.

DAVID bioinformatics resources consist of an integrated biological knowledge base and analytic tools aimed at systematically extracting biological meaning from large gene/protein lists. The procedure first requires uploading a gene list containing any number of common gene identifiers followed by analysis using one or more text and pathway-mining tools such as gene functional classification, functional annotation chart or clustering and functional annotation table ⁹⁶.

NON-NEGATIVE MATRIX FACTORIZATION (NMF)

In order to identify biological subgroups according to the genetic profile, unsupervised clustering was performed using the non-negative matrix factorization (NMF) method ^{97, 98}. NMF aims to reduce the dimensionality of a data matrix while keeping the important information that it encodes. Standard factorization of a data matrix uses decompositions that allow negative entries, such as principal component analysis (PCA) and singular value decomposition (SVD). However, for many data sets such as copy number intensities, the original data matrices are non-negative. In this case, a standard factorization has difficulty for interpretation. In contrast to cancellations due to negative entries in

matrix factors, the non-negativity in NMF ensures factors contain coherent parts of the original data.

Given an original data matrix X of dimensions n by m , the idea is to obtain matrices W and H such that $\|X - WH\|$ is minimized, where $\|\cdot\|$ represents a given norm. In this manner, the matrix X is compactly represented by two lower rank matrices W and H . The dimensions of W and H depend on the rank that is chosen by the user. The main idea is that the matrices W and H encode the same information of X , but using lesser parameters. It has been shown that, under some simple circumstances, NMF with rank K resembles a relaxed form of the well-known K-means clustering: matrix factor W contains cluster centroids and H contains cluster membership indicators, that is, W gives us the weights in which each probe is used to separate the samples into clusters, while H can be analyzed by columns, where each column of H describes the weight in which each sample could be assigned to the different clusters. This justifies the use of NMF for data clustering. Although implementations of NMF are available in many languages, efficiency is not taken as the main concern in most cases. As we need to deal with large matrices, we used an in-house solution that has been optimized for our problem.

ANALYSIS OF CLINICAL DATA

The impact of the MCR on the clinical course was determined with the log-rank test ⁹⁹ according to Overall survival (OS), Progression free survival (PFS) and Disease free survival (DFS).

OS is defined as the time from entry onto the clinical trial until death as a result of any cause. PFS is defined as the time from entry onto a study until lymphoma progression or death as a result of any cause. DFS is the time of occurrence of disease-free state or attainment of a complete remission (CR) to disease recurrence or death as a result of lymphoma or acute toxicity of treatment.

Response to R-CHOP, OS, PFS and DFS were defined according to the criteria of Cheson et al. ¹⁰⁰, defining the time of diagnosis as start of the study. To control false positive rate in log-rank test of the MCR according to OS, PFS e DFS, MTC

was performed with the Benjamini-Hochberg-method (q_{value}). To identify MCR with an independent impact on survival according to the gain, loss, amplification ($CN > 4$), homozygous deletion or LOH we used the Cox proportional-hazard model¹⁰¹. A p-value of < 0.05 was considered as statistically significant. Chi-square test was used to compare differences in clinical parameters among the different NMF subgroups previously identified. Statistical analyses on clinical data were performed with SPSS 17.0.1 (SPSS, Chicago, IL, USA) and R statistical package (<http://www.r-project.org/>).

GENE EXPRESSION

Previously published raw gene expression profiles (GEP) data, obtained using the Affymetrix GeneChip U133 plus 2.0 (Lenz G, 2008), were available for 54 of the 166 IC-DLBCL: 30 GCB-like, 19 ABC, and five unclassified. Array data were kindly provided by Dr W.C. Chan, University of Nebraska, Omaha, USA.

Data were analyzed using Partek Genomics Suite 6.4 (Partek, St. Louis, MO). Signal intensities were normalized by Partek RMA. Statistical differences were calculated by ANOVA analysis and the genes with a p-value < 0.005 were analyzed for enrichment of KEGG and Biocarta pathways using DAVID 2008 web-accessible program⁹⁶.

CELLS CULTURE

OCI-LY7 cell line were maintained in RPMI 1640 (GIBCO, Invitrogen) supplemented with 10% fetal bovine serum (FBS), 1% of sodium pyruvate (GIBCO, Invitrogen) 0.1% MEM Non-Essential-amino acid (NEAA) (GIBCO, Invitrogen), 0.1% 2-mercaptoethanol (GIBCO, Invitrogen), 1.1% of penicillin and of streptomycin (pen/strep) (GIBCO, Invitrogen) in a 5% CO₂ humidified atmosphere incubator at 37°C. The same growing medium was used for VAL, DOHH2, LY10, LY19, SUDHL6 DLBCL cell lines. For K422 and U2932 cell lines RPMI 1640 (GIBCO, Invitrogen) medium supplemented with 10% bovine growth serum was used. HEK-293T packaging cells (ATCC, Manassas, VA) were cultured in D-MEM (GIBCO) medium complemented with 10% of FBS, 100x Pen/Strep (GIBCO, Invitrogen), NEAA (GIBCO, Invitrogen) and 100x Glutamine (GIBCO, Invitrogen).

RETROTRANSCRIPTION AND RT-PCR ANALYSIS

Total RNA (500 ng up to 1 µg) from each sample was reverse-transcribed using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen). Briefly, 500ng up to 1 µg total RNA (RNA concentration ≥ 50 ng/µl) was mixed with 50 ng random hexamers, and 1 mM dNTP mix, and the final volume of 10 µl was reached with DEPC water. The first mix was incubated at 65°C for 5' in a MJR Thermal Cycler 200, then cool on ice for at least 1'. In the meantime, the second mix was prepared, added to the first mix and the whole was further incubated at 25°C for 2' in the thermal cycler. At that point, the enzyme SuperScript II RT was added to the final mix, which consisted of the first mix with 1x RT buffer, 5nM MgCl₂, 10 mM DTT, 40 U RNase OUT and 50 U SuperScript II RT. The following program was run on the thermal cycler: 25°C for 10', 42°C for 50', 70°C for 15'. The reaction was then chilled on ice and 2 U RNase-H were added in each tube, with subsequent incubation at 37°C for 20'. The results were diluted 1:5 with DEPC water.

cDNA solution (2.5 µl) was amplified by PCR in a 25 µl solution containing PCR Buffer 10X (20 nM Tris-HCl, 50 nM KCl), 1.5 nM MgCl₂, 5 nM dNTPs, 2.5 U Taq DNA polymerase (Roche) and 30 pmol of each primer. The following program was run on the thermal cycler: 95°C for 10' and then 95° for 15", 55-58°C for 30", 72°C for 30" for 30 cycles. The oligonucleotides used for PCR were designed for *FLI-1* (primer F 5'-GTGCACAGGGGAGTGAGG-3'; primer R 5'-TCACTGGCTGATTGATCCAC-3'), *PAX-5* (primer F 5'-GGAGGAGTGAATCAGCTTGG-3'; primer R 5'-GGCTTGATGCTTCCTGTCTC-3'), *ETS-1* (primer F 5'-GGAGCAGCCAGTCATCTTTC-3'; primer R 5'-TTGAATCCCAGCCATCTC-3'), *BLIMP1* (primer F 5'-GAAACATGACCGGCTACAAGA-3'; primer R 5'-CGAGCTGAGTAAAGCCCTTG-3') and *GAPDH* (primer F 5'-GGCTGTGGGCAAGGTCATCCCTGA-3'; primer R 5'-TCCACCACCCTGTGCTGTA-3').

METHYLATION

The methylation status of CDKN2A was investigated by methylation-specific polymerase chain reaction (MSPCR), as previously reported ¹⁰². DNA used as a template was chemically modified by bisulfite treatment (EpiTect Bisulfite Kit, 4 QIAGEN Milan, Italy). MSPCR was performed with primers specific for methylated and unmethylated sequences ¹⁰² on 50 ng of bisulfite-treated gDNA in a Verity thermal cycler (Applied Biosystems, Milan, Italy). All MSPCR reactions were performed with positive (SUDHL6 for methylation status and LNCAP for unmethylation status) and negative (normal gDNA) controls for both unmethylated and methylated alleles. gDNA solution of patients and controls (2.5 µl) was amplified by PCR in a 25 µl solution containing PCR Buffer 10X (20 nM Tris-HCl, 50nM KCl), 1.5 nM MgCl₂, 5 nM dNTPs, 2.5 U Taq DNA polymerase (Roche) and 30 pmol of each primer. PCR products were separated by agarose gel electrophoresis and visualized under UV illumination.

REAL-TIME-PCR

Total RNA (1 µg) was reverse-transcribed with random hexamers using SuperScript First-Strand Synthesis System for PCR (Invitrogen).

Real-time PCR was performed with the ABI PRISM 7000 Sequence Detection System (Applied Biosystems) and the Absolute QPCR ROX Mix (AB-1139) (ABgene, Epsom, UK). We used the TaqMan GeneExpression Assay (20X mix of unlabelled PCR primers and FAM-labeled TaqMan MGB probes) for *ETS-1* (*Hs00901425_m1*), *FLI-1* (*Hs00231107_m1*) and β -2-microglobulin (*Hs9999907_m1*). Polymerase chain reactions were performed using a 96-wells plate and optical adhesive covers (Applied Biosystems) in a final volume of 25 µl with 1X Absolute QPCR ROX Mix (ABgene), 1X mix of primers and probes, 2 µl diluted cDNA, and DEPC-water to reach the final volume. Cycling parameters were: 95°C for 15', followed by 40 cycles of 95°C for 15'' and 60°C for 1'. All samples, were analyzed in triplicate. A NTC was added to each plate and for each assay, and contained all PCR reagents without any template DNA. The comparative CT method was used for relative quantification. Gene expression levels were

normalized to the endogenous control β 2-microglobuline and divided by the calibrator (CD19 positive B cells).

WESTERN BLOT (WB) ANALYSIS

For Western Blot (WB) analysis, cells were lysed with 25 μ l of lysis buffer containing 20 μ l of HEPES pH 7,9 20 mM, 50 μ l of NaCl 50 mM, 10 μ l of EDTA 10 mM, 10 μ l of EGTA 2 mM, 25 μ l of NP-40 0.50%, 0.5 μ l of DTT 1 mM, 5 μ l of Na₃Vo₄ 1 mM, 25 μ l of NaF 50 mM, 0.5 of PMSF 0.2 mM, 50 μ l of protein Kinase cocktail 1X (Complete Mini, Roche), and the final volume of 500 μ l was reached with DEPC water.

Protein quantification was done by BCA protein Assay kit (Pierce, IL, USA) according to the manufacturer's instructions. This assay enables a colorimetric detection and quantification of total protein based on bicinchoninic acid (BCA). The method combined the reduction of Cu²⁺ to Cu⁺ by proteins in alkaline medium, known as biuret reaction, with the colorimetric detection of Cu⁺ with a reagent containing BCA. The chelation of two molecules of BCA with one cuprous cation (Cu⁺) give a purple coloration and the complex strongly absorbs at 570 nm. Absorbance can be detected by 96-well plate reader (Beckman Coulter AD340, Beckman Coulter GmbH), and it is directly proportional to the protein concentration.

Prior to gel electrophoresis, protein were denatured at 100°C for 4'. 20 μ g of total protein lysates were loaded on 8% or 10% sodium dodecylsulfate-polyacrilamide gel electrophoresis (SDS-PAGE). Electrophoresis was performed in 1X running buffer (TGE). Then, proteins were transferred onto nitrocellulose membranes using 1X transfer buffer. Filters were first blocked (0.2% I-BLOCK in PBS with 0.1% Tween 20) 1 h at room temperature (RT), and then incubated over-night at 4°C with primary antibodies, prepared in 0.1% TBS-T according to optimized dilutions: anti*FLI1*, anti *ETS-1*, anti *PAX5*, and anti- α -TUBULIN (Santa Cruz Biotechnology) and anti *BLIMP1* (BD) 1:1000 were used. After 5 washes of 5 minutes each using 0.1% TBS-T, filters were incubated for 1 h at RT with secondary anti-mouse or anti-rabbit antibodies using 1:2000 dilution for the

detection of all primary antibody (Amersham, Piscataway, NJ). After 5 washes with 0.1% TBS-T, the detection of immunocomplexes was performed with an enhanced chemiluminescence system (ECL; Pierce).

LENTIVIRUS INFECTION

Lentivirus supernatants (SNs), containing ShRNA against *FLI1* (2 different clones of pLKO transfer vector: 61C and 61E), *ETS-1* (3 different clones of pLKO transfer vector: 60A, 60C, 60D) or GFP (scramble control), were produced by co-transfecting HEK-293T packaging cells with pLKO transfer vector (2.5 µg/µl), pCMV-dR8.74 vector (1.875 µg/µl) and pMD2.VSVG vector (0.625 µg/µl) (Clontech) using Lipofectamine (Invitrogen) in six-well plates. After 24h, 2ml of fresh medium were added to each well and lentivirus SN from each condition was harvested after 48 and 72 h, filtrated (0.22-µm pore), and used directly, to infect OCI-LY7 (1 ml of SN every 1×10^5 /ml cells). The infection was performed using FACS Tubes and a dose of 6 µg/ml of polybrene was added to each infection tube, to allow the interaction between virus and target cells. OCI-LY7 cells were left for 6 hours at 37°C, to allow the infection, and then the cells were washed with PBS 1X and centrifuged at 1200 rpm for 10'. After centrifugation SN was removed and infected cells were re-suspended in fresh complete medium and plated into 24 well plates. Each stable clone was generated in quadruplicates (or more).

Twenty-four hours after the infection, 500 µl of fresh medium containing puromycin (20 µg/ml final concentration) could be added and clones were selected for 10-12 days with the antibiotic. RNA and protein were extracted always after 10-12 days.

INTRACELLULAR STAINING

OCI-LY7 infected with ShRNAs targeting *FLI-1* (61C, 61E) were selected for 10 days with puromycin and then harvested and washed in PBS 1x plus 1% FBS and collected by centrifugation at 1500 rpm for 5 minutes at 4°C. Supernatant (SN) was removed and cells were fixed and permeabilized by re-suspending cell pellet in 200 µl per sample using BD Cytfix/Cytoperm Buffer and 40 µl per well

were aliquoted in a 96 well plate, in order to have 4 replicates per condition. Cells were incubated for 20-30 minutes on ice and then washed by adding 200 μ l of 1x Perm Wash Buffer per replicate. Cells were then collected by centrifugation at 1300 rpm for 2 minutes at 4°C and SN was then removed.

Intracellular staining was performed by re-suspending each cell pellet in 50 μ l of the appropriate antibody dilution in 1x Perm Wash Buffer. Incubate 20-30 minutes on ice, light protected. Antibodies (Abs) used were anti-ETS-1, anti-FLI-1 (Polyclonal, goat-anti rabbit, Santa Cruz Biotechnology), and anti-PAX5 (monoclonal, goat-anti-mouse, Santa Cruz), at a dilution rate of 1:250 μ l. As negative control for polyclonal antibodies, we used an anti- rabbit serum. Cells were then washed in 200 μ l 1xPerm Wash Buffer per sample and collected by centrifugation at 1300 rpm for 2 minutes at 4°C and SN was then removed.

Because anti-ETS-1 and anti-FLI-1 were polyclonal Ab raised in rabbit, we incubated replicates for those Abs and negative controls, with a secondary goat-anti-rabbit/biotynilated Ab at a concentration of 2 μ g/ml for 20 minutes in ice. Anti-PAX5 was a monoclonal Ab raised in mouse and so replicates for this Ab was incubated for 20 minutes in ice with a secondary goat-anti-mouse/biotinilated Ab, at the same concentration. Monoclonal Abs do not need a negative control as they are proved to be highly specific.

Cells were then washed in 200 μ l 1xPerm Wash Buffer per sample and collected by centrifugation at 1300 rpm for 2 minutes at 4°C and SN was then removed.

Because all secondary antibodies were biotynilated, we performed a third step of incubation using the streptavidine at a concentration of 2 μ g/ml, in order to induce a fluorescent reaction due to the presence of biotin in secondary Abs. All samples were then incubated for 20 minutes with this solution and cells were finally washed in 200 μ l 1x Perm Wash Buffer per sample and collected by centrifugation at 1300 rpm for 2 minutes at 4°C and SN was then removed.

All replicates were resuspended in 150 μ l of PBS 1X and then analyzed by FACS-canto.

CFSE LABELLING ASSAY

5-10*10³ OCI-LY7 cells were spun down in a 5 ml FACS tube, washed with 3 ml of PBS 1X and centrifuged at 1500 rpm for 5 minutes. After centrifugation, pellet was re-suspended in 500 µl of PBS 1X with the addition of 500 µl of CFSE solution (final concentration was 0.5 µM). CFSE solution was prepared fresh with a dilution of CFSE 1:10000 in PBS. Cells were then incubated under agitation for 8 minutes in the dark at room temperature. Then, 1 ml of complete media for OCI-LY7 (see section CELLS CULTURE for details) was slowly added and centrifugation at 1500 rpm for 5 minutes was repeated. OCI-LY7 cells were then re-suspended in 150 µl of complete medium and cultured in a 5% CO₂ humidified atmosphere incubator at 37°C. CFSE data acquisition was performed using FACS analysis after 3, 6 and 9 days. Cells were washed again in PBS 1X and re-suspended in 150 µl of PBS 1X directly for the analysis with FACS-Canto.

3. IDENTIFICATION OF RECURRENT GENETIC LESIONS IN DLBCL

3.1 ABSTRACT

In this chapter we present the molecular characterization of DLBCL by genomic profiling approach, already introduced in Chapter 1. DLBCL is the commonest type of non-Hodgkin's lymphoma (NHL), occurring in approximately 40% of patients. DLBCL can be cured using anthracycline-based chemotherapy regimens in only 40%–50% of patients, suggesting that DLBCL is a heterogeneous diagnostic category and an important percentage of patients cannot yet be cured.

To improve molecular characterization of DLBCL and to identify new candidate cancer genes, we obtained the genome wide DNA profiles of 166 DLBCL samples. Using an algorithm previously described by Lenz G. et al in 2008 ⁴, we have been able to identify 171 minimal common regions (MCR), among which we found 47 recurrent regions of gains, 58 losses and 54 LOH ¹⁰³. Moreover, we detected also five recurrent amplifications and seven homozygous deletions. The most frequent aberrations observed were gains on 1q, 2p, 6p, 7q, 11q and 18q, whereas deletions were found on 1p, 6q and 15q. Part of these aberrations have already been reported in literature ^{4, 63, 81}, indicating that our data are consistent with what has been described so far, but some others (such as 11q gain) have never been characterized.

Furthermore, we tried to understand the power of classification between GCB and non-GCB DLBCL, comparing Immunohistochemistry (IHC) and GEP data for a subset of 54 patients of the present work. The two methods showed significant differences, indicating that new tools are needed.

In conclusion our data support the evidence of the genomic complexity of DLBCL and reinforce the role of integrative genomic approaches for improving the understanding of molecular pathogenesis of DLBCL.

3.2 AIM OF THE CHAPTER

Our aim was to detect causative lesions in DLBCL genome as the one being recurrent in our data set and affecting the expression of genes, which might be involved in pathways that participate in the lymphoma pathogenesis.

Finally we compared recurrent aberrations occurring in a subset of patients classified into the main DLBCL biological entities (GCB and ABC DLBCL) and for which, GEP data were available.

3.3 RESULTS

3.3.1 DISTINCT RECURRENT ALTERATIONS IN DLBCL

Genomic analysis was performed with the Affymetrix 250k SNP arrays Figure 3.1 shows the frequencies of gains, losses and of LOH in 166 DLBCL samples.

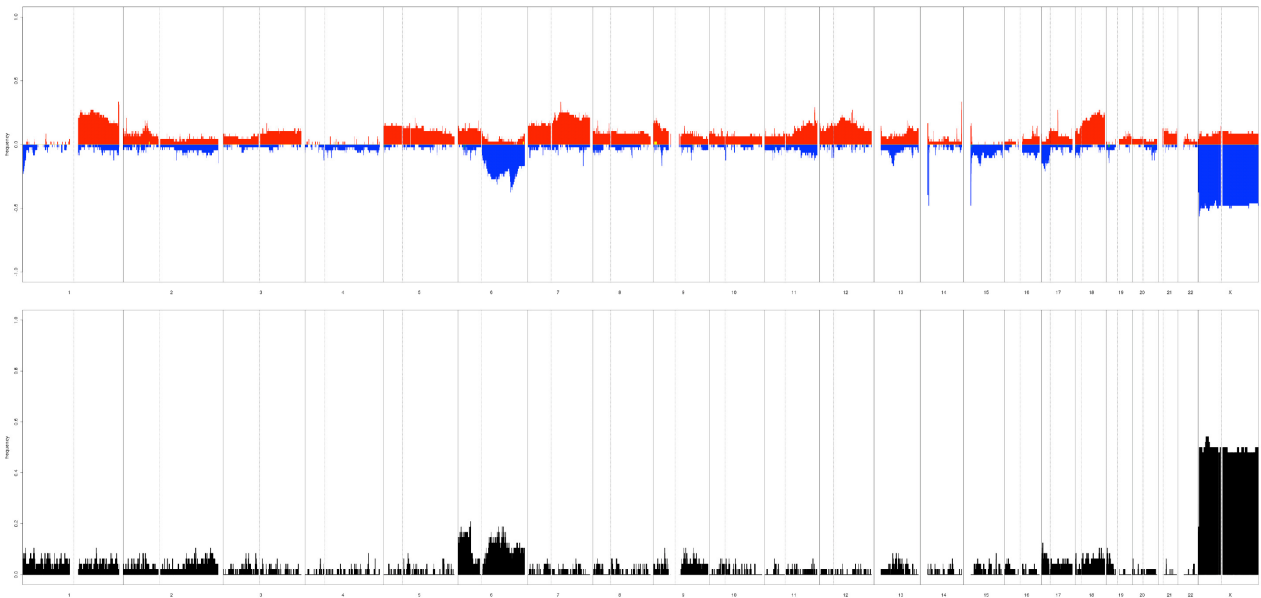


Figure 3.1. Frequency of DNA gains (up) and losses (down) (upper panel) and LOH (lower panel) observed in 166 DLBCL samples. X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

In order to reduce the large amount of genomic profiling array data, recurrent minimal common regions (MCRs) were defined using the algorithm by Lenz et al ⁴ (figure 3.2; refer also to chapter 2). Four distinct types of MCR were defined: short recurrent abnormality (SRA), long recurrent abnormality (LRA), abnormal chromosome arm (ACA) and abnormal whole chromosome (AWC). A total of 171 MCR were identified. Recurrent regions of gains were 58 (21 SRA, 9 LRA, 6 AWC, 22 ACA), losses 47 (33 SRA, 7 LRA, 7 ACA), LOH 54 (47 SRA, 2 LRA and 5 ACA). Five recurrent amplifications and seven homozygous deletions were identified. Table 3.1 shows the 41 MCR occurring in more than 15% of the cases.

MINIMAL COMMON REGIONS (MCR)

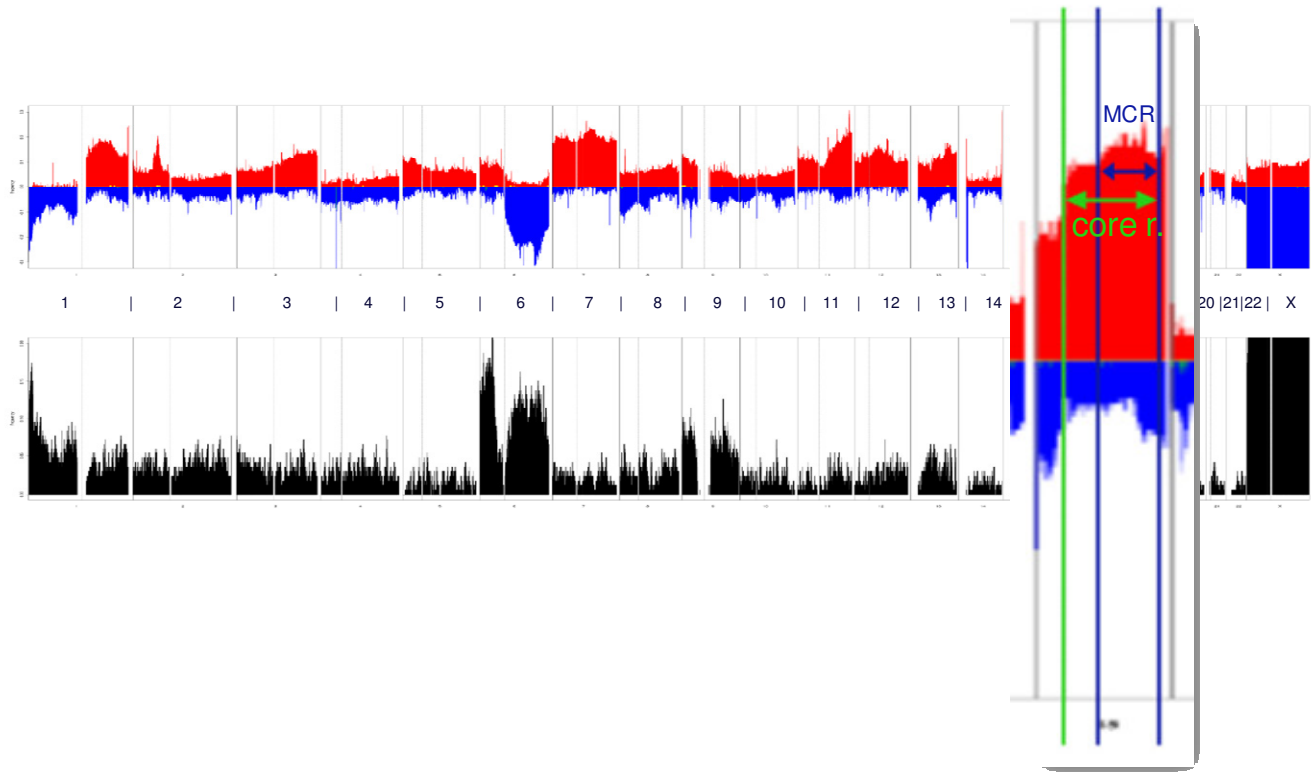


Figure 3.2. Example of a minimal common region (MCR). Core r.=core region. The plot represents also the frequency of DNA gains (up) and losses (down) (upper panel) and LOH (lower panel) observed in 166 DLBCL samples. X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

Among gains, the most recurrent lesions were 1q44 (40/166, 25%), 2p16.1-15 (34/166, 20%), 7q21.11 (44/166, 27%), 11q24.3 (49/166, 30%), 14q11.2 (32/166, 19%), 17q21.31 (42/166, 25%), 18q23 (36/166, 23%). The 1q44 and the 2p16.1-15 regions, contained a large number of genes. The other gained regions included a smaller variety of genes, such as *SEMA3D* in 7q21.11 and a subset of 7 genes (*C11orf45*, *ETS-1*, *FLI-1*, *KCNJ1*, *KCNJ5*, *P53AIP1* and *RICS*) in 11q24.3. *ETS-1* and *FLI-1* are known to be members of a family of transcription factors, often

involved in cancer, including lymphomas ¹⁰⁴. The 14q11.2 region included 8 genes (*A26C2*, *OR11H12*, *OR4K1*, *OR4K2*, *OR4K5*, *OR4M1*, *OR4N2*, *OR4Q3*), while 17q21.31 contained 6 genes (*ARL17*, *ARL17P1*, *KIAA1267*, *LRRC37A*, *LRRC37A2*, *NSF*) and finally 18q23 included only *ADNP2*, *PARD6G*.

Recurrent losses mainly affected 1p36.33-p36.31 (43/166, 26%), 6q16.3 (47/166, 28%), 6q21.1 (52/166, 31%), 6q23.3-24.1 (53/166, 32%), 14q.11 (50/166, 30%), 15q 11.2 (66/166, 40%), 17p13.3-p11.2 (28/166, 17%) and 18p11.32 (30/166, 18%). As shown in the table 3.1 many other regions on chromosomes 14 and 15 seem to be highly aberrated, but they might be considered bona fide CNV since they had a total overlap with known CNV, although the first lesion is located within the T-cell receptor alpha (*TCRA*) locus, and could also be due to infiltrating normal T-cells or by a T-cell rearrangement occurred in the lymphoma B-cell ¹⁰⁵. The 15q deletion targeted five genes, among which *OR4N4* and *OR4M2* are G-protein-coupled receptors, members of the olfactory receptor genes family. The deletion of the 19p13 region has not been previously reported and comprised a series of genes, including *C3* (complement precursor 3) and *TNFSF14* (tumor necrosis factor ligand superfamily 14). Also in this case deletion occurring on chromosome 1 and 17 contained a large number of genes and for the same reasons, previously explained, they were difficult to analyze. Although 6q23.3-24.1 contained 23 genes we have been able to characterize *TNFAIP3*, in collaboration with the group of Laura Pasqualucci in New York, as recurrently deleted or otherwise mutated in ABC-DLBCL patients ²⁷(*BCLAF1*, *C6orf115*, *C6orf91*, *CCDC28A*, *CITED2*, *FAM54A*, *HEBP2*, *HECA*, *IFNGR1*, *IL20RA*, *IL22RA2*, *KIAA1244*, *MAP3K5*, *MAP7*, *OLIG3*, *PBOV1*, *PDE7B*, *PERP*, *PEX7*, *REPS1*, *SLC35D3*, *TNFAIP3*, *TXLNB*).

The 6q16.3 and the 6q21.1 included respectively three (*CCNC*, *MCHR2*, *PRDM13*) and two genes (*C6orf203*, *KIAA1553*). Moreover in the 14q11.2 eight genes were mapping (*A26C2*, *OR11H12*, *OR4K1*, *OR4K2*, *OR4K5*, *OR4M1*, *OR4N2*, *OR4Q3*), while the 18p11.32 contained ten genes (*ADCYAP1*, *C18orf56*, *CETN1*, *CLUL1*, *COLEC12*, *ENOSF1*, *THOC1*, *TYMS*, *USP14*, *YES1*).

Homozygous deletions occurred at 3p14.2 (2/166, 1%), 6q23.3-q24.1 (2/166, 1%), 9p21.3 (5/166, 3%), 10q23.31 (2/166, 1%), 14q11.2 (3/166, 2%), 15q11.2 (2/166, 1%), 19p13.3-p13.2 (2/166, 1%). Known tumor suppressor genes were included in 4/7 of these MCR: *FHIT* (3p14.2), *TNFAIP3/A20* (6q23.3), *CDKN2A* (9p21.3) and *FAS* (10q23.31). Of interest, we have shown that *FHIT* is a very common target of deletion in HIV-related DLBCL ²⁶

Amplifications occurred at 2p16.1 (7/166, 4%), 2p15 (3/166, 2%), 9p24.3-p21.3 (3/166, 2%), 13q32.1 (2/166, 1%) and 19q13.3 (2/166, 1%). Known oncogenes and lymphoma genes were present in 3/5 lesions: *BCL11A* and *REL* (2p16), *JAK2* (9p), *SPIB* (19q). The 2p15 included *LOC51057*, *MDH1*, *UGP2*. The 13q32.1 region contained *DNAJC3*, *HS6ST3*, *OXGR1*, *UGCGL2* and mapped approximately 4 Mb telomeric to the transforming MIR17-92 cluster.

Table 3.1. SRA and LRA occurring in more than 15% of the cases. (*more than 80% of the probes of the region overlap with known copy number variations; ° this SRA is included in a LRA; p-values, bold if <0.05)

Status	Region	Frequency	Start	End	Size	Type	Notes
gain	7q	16%				WCA	
	11p15.5-p15.4	15%	2.27E+06	2.99E+06	7.25E+05	SRA	
	11q22.1-q25	19%	9.86E+07	1.34E+08	3.58E+07	LRA	
	11q24.3-q24.3	30%	1.28E+08	1.29E+08	9.39E+05	SRA	
	12q21.31-q21.31	19%	8.07E+07	8.07E+07	4.73E+04	SRA	
	12q21.31-q21.31	19%	8.31E+07	8.32E+07	3.35E+04	SRA	
	13q31.3-q31.3	17%	8.93E+07	9.13E+07	1.99E+06	SRA	
	14q11.2-q11.2	19%	1.93E+07	1.95E+07	1.53E+05	SRA	*
	14q32.33-q32.33	29%	1.06E+08	1.06E+08	5.70E+05	SRA	*

Table 3.1. SRA and LRA occurring in more than 15% of the cases.

	15q11.2-q11.2	19%	1.84E+07	2.01E+07	1.66E+06	SRA	*
	17q21.31-q21.31	25%	4.15E+07	4.17E+07	2.01E+05	SRA	*
	18q	20%				ACA	
	18q11.1-q23	20%	1.69E+07	7.61E+07	5.93E+07	LRA	
	18q23-q23	22%	7.60E+07	7.61E+07	1.38E+05	SRA	*
	1q21.1-q42.3	17%	1.44E+08	2.34E+08	9.02E+07	LRA	
	1q23.3-q23.3	20%	1.60E+08	1.61E+08	2.15E+05	SRA	
	1q44-q44	25%	2.46E+08	2.47E+08	1.29E+06	SRA	*
	2p16.1-p15	17%	6.06E+07	6.22E+07	1.55E+06	LRA	
	2p16.1-p15	20%	5.86E+07	6.37E+07	5.18E+06	SRA	*
	7p22.3-q36.3	23%	1.41E+05	1.59E+08	1.59E+08	LRA	
	7q21.11-q21.11	27%	8.41E+07	8.46E+07	4.58E+05	SRA	
LOH	1p36.32-p36.23	17%	5.28E+06	9.05E+06	3.77E+06	SRA	
	6p21.33-p21.32	22%	3.07E+07	3.30E+07	2.24E+06	SRA	
	6p22.3-p22.3	18%	2.11E+07	2.12E+07	9.86E+04	SRA	
	6p24.3-p24.3	19%	7.45E+06	9.68E+06	2.22E+06	SRA	
	6p25.3-p21.2	15%	2.94E+05	3.90E+07	3.87E+07	LRA	
	6q16.2-q16.3	16%	9.97E+07	1.01E+08	9.41E+05	SRA	°
loss	14q11.2-q11.2	30%	1.93E+07	1.95E+07	1.53E+05	SRA	*
	14q11.2-q11.2	40%	2.15E+07	2.20E+07	5.66E+05	SRA	*

Table 3.1. SRA and LRA occurring in more than 15% of the cases.

	15q11.2-q11.2	40%	1.84E+07	1.99E+07	1.52E+06	SRA	*
	15q12-q21.2	16%	2.40E+07	5.01E+07	2.62E+07	LRA	
	15q15.1-q21.1	19%	3.90E+07	4.38E+07	4.85E+06	SRA	
	17p13.3-p11.2	17%	1.89E+04	1.60E+07	1.60E+07	SRA	
	18p11.32-p11.32	18%	2.10E+05	8.85E+05	6.75E+05	SRA	
	1p36.33-p36.31	26%	7.76E+05	5.40E+06	4.62E+06	SRA	°
	6q	17%				ACA	
	6q13-q27	30%	7.56E+07	1.69E+08	9.32E+07	LRA	
	6q16.3-q16.3	28%	1.00E+08	1.01E+08	4.20E+05	SRA	°
	6q21-q21	31%	1.07E+08	1.07E+08	1.26E+05	SRA	°
	6q23.3-q24.1	32%	1.36E+08	1.40E+08	3.94E+06	SRA	°
	9p21.3-p21.3	19%	2.18E+07	2.22E+07	3.57E+05	SRA	

Table 3.2. Differences in genomic lesions between GCB and ABC as evaluated by applying Fisher's exact test (p-value)

Genomic lesion	ABC-GEP	%	ABC-IHC	%	p-value
del 6q16.3	2/19	10%	15/35	42%	0.016
del 6q23.3-q24.1	3/19	15%	16/35	46%	0.037
del 6q13-q27	2/19	10%	16/35	46%	0.014
12q21.31 gain	0/19	0%	8/35	23%	0.039
del 15q11.2	5/19	26%	22/35	63%	0.021
del 17p13.3-p11.2	0/19	0%	7/35	20%	0.04
	GCB-GEP	%	GCB-IHC	%	p-value
del 1p	6/30	20%	0/25	0%	0.02
del 6q16.3	7/30	23%	8/25	32%	0.031
18q11.1-q23 gain	2/30	6%	7/25	28%	0.031
del 18q23	8/30	27%	0/25	0%	0.05
18q23 gain	1/30	3%	7/25	28%	0.017

3.3.2 DLBCL AND CELL OF ORIGIN

Gene expression data were available for 54 patients coming from University of Nebraska Medical Center, Omaha. Genomic profiles of patients classified by GEP and Immunohistochemistry (IHC) showed many differences. Interestingly, considering IHC data, 18q11-q23 gain (7/25; $p_{\text{value}} = 0.031$) and losses of 6q16.3 (13/25; $p_{\text{value}} = 0.047$) were detected in the GCB-DLBCL group (Table 3.2; Figure 3.3, Panel B). Gain of 12q21.3 (8/35; $p_{\text{value}} = 0.039$) and losses of 6q13-q27 (16/35; $p_{\text{value}} = 0.014$), 15q11.2 (22/35; $p_{\text{value}} = 0.021$) and 17p13.3-11.2 (7/35; $p_{\text{value}} = 0.043$) were observed among non-GCB patients (table 3.2; Figure 3.3, Panel D).

According to GEP data, GCB-DLBCL were mainly characterized by losses on 1p (6/30; $p_{\text{value}} = 0.026$), on 6q16.3 (7/30; $p_{\text{value}} = 0.047$) and on 18q.23 (8/30; $p_{\text{value}} = 0.05$) (table 3.2; figure 3.3, Panel A). Recurrent aberrations among ABC-DLBCL

patients were losses of 6q23.3-q24.3 (3/19; $p_{\text{value}} = 0.037$) and of 15q11.2 (5/19; $p_{\text{value}} = 0.021$) (Table 3.2; Figure 3.3, Panel A and C). Statistical analysis of the comparison between GCB and ABC-DLBCL failed to reveal any significant difference among subgroups.

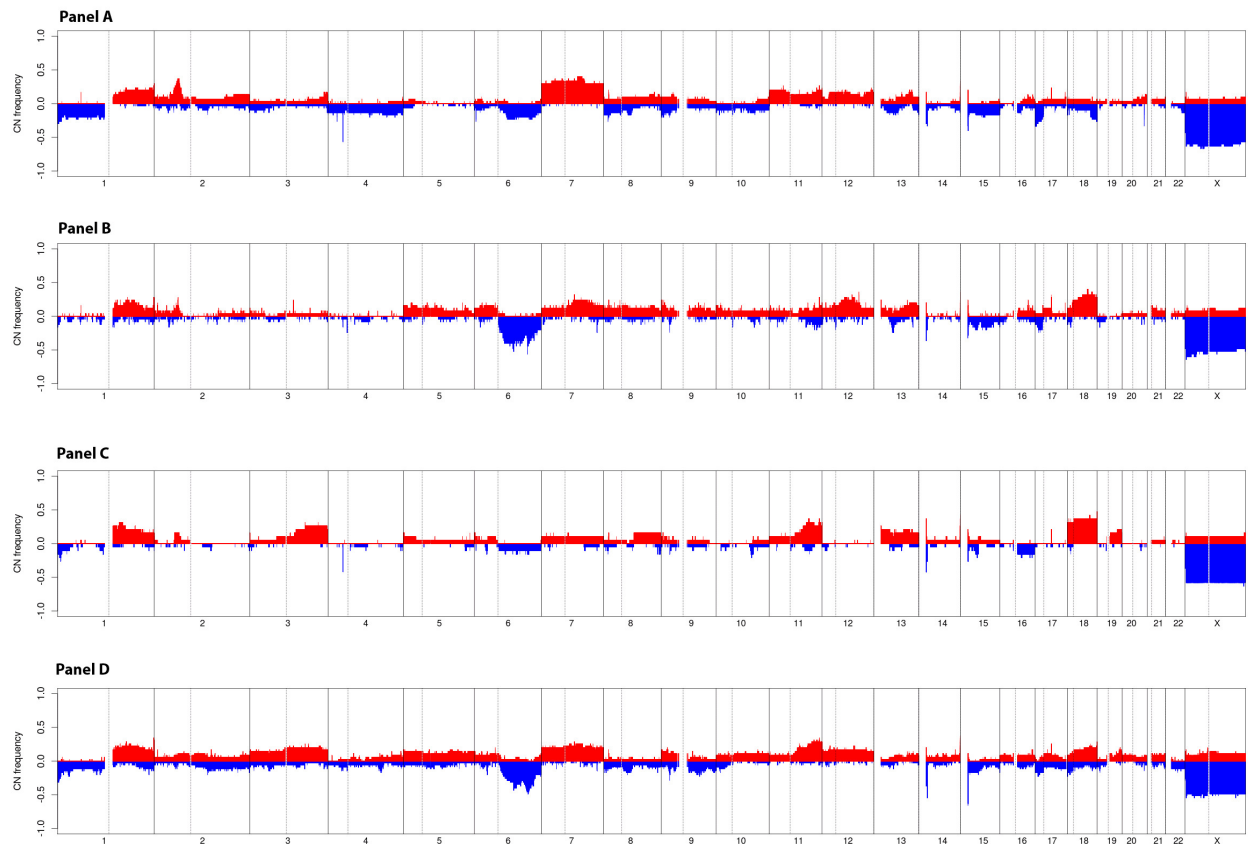


Figure 3.3. CN changes in 30 samples classified as GCB-DLBCL by GEP (A), in 25 samples classified as GCB-DLBCL cases by IHC (B), in 19 samples classified as post GC-DLBCL by GEP (C), and in 35 samples classified as post GC-DLBCL by IHC, (D). X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

3.5 DISCUSSION

A series of 171 recurrent lesions, MCR, were identified, partially reflecting what already reported in DLBCL ^{4, 27, 63, 81} and confirming the efficacy of the adopted method. Our data further support the evidence of the genomic complexity of DLBCL and they fortify the role of aCGH in the understanding of the molecular pathogenesis of the disease.

Importantly, the high resolution of the arrays allowed the detection of a discrete number of recurrent aberrations. The relevance of the analysis of these regions is that presumably they represent aberrations involving important genomic loci that could contribute to lymphomagenesis with different mechanisms of action, underlining the complexity and the variety that characterized DLBCL disease.

Gains occurring in chromosome 1q44, 7q (whole arm) and 2p16.1-15 have been already reported in DLBCL ⁸¹. The proto-oncogene *BCL11A* (B-cell lymphoma/leukemia 11A) at 2p16 is regarded as candidate target gene of these imbalances, as it was often co-amplified with the *REL* proto-oncogene in non-Hodgkin's lymphoma (NHL) as well as in classical Hodgkin's lymphoma ⁸⁸. *BCL11A* encodes a Kruppel-like zinc-finger transcription factor, which represses transcription. Initially, the *BCL11A* gene was identified as recurrent translocation partner with the immunoglobulin heavy chain locus in B-cell NHLs.

Aberrations occurring at chromosome 7 are very frequently represented in DLBCL. Gain of 7q is a common genetic lesion among DLBCL, possibly more common in germinal center B-cell-like (GCB) than in activated B-cell-like (ABC) types ^{4, 106}. Our group conducted a study on genomic lesions and Bone Marrow (BM) infiltration in DLBCL (Chigrinova E, submitted). Gains of chromosome 7 were highly significant inversely correlated with BM infiltration. Gains of 7q seem to occur at similar percentage among DLBCL arising from both immune-privileged and immune-competent anatomical sites ¹⁰⁷, and they could be inversely associated with a T-cell infiltrate ⁶³.

Gain on 11q contains two interesting genes, *ETS1* and *FLI1*, whose role in DLBCL was not known and will be elucidated in the present work in chapter 5.

Consistently with other published data, we found losses occurring at chromosome 1p, 15q, 17p, 18p and 6q and all of them have been already described in DLBCL ^{4, 27, 45, 81}. Interestingly 18p loss has been found to be recurrent in PMBCL patients, which represent a third subgroups of DLBCL.

As reported in the present work, deletions in the short arm of chromosome 17 (17p), probably involving the tumor suppressor *TP53* occur in up to 20% DLBCLs. Although inactivation of both alleles of a tumor suppressor gene is usually required for tumor development, the overlap between *TP53* deletions and mutations is poorly understood in DLBCLs, suggesting the possible existence of additional tumor suppressor genes in 17p ⁷⁶. Telomeric to the *TP53*, Stocklein et al ⁷⁶, identified the presence of the tumor suppressor hypermethylated in Cancer 1 (*HIC1*). Methylation-specific PCR demonstrated hypermethylation of *HIC1* exon 1a in a substantial subset of DLBCLs, which is accompanied by simultaneous *HIC1* deletion of the second allele in 90% of cases ⁷⁶. DLBCL patients with complete inactivation of both *HIC1* and *TP53* may be characterized by an even inferior clinical course than patients with inactivation of *TP53* alone, suggesting a functional cooperation between these two proteins. These findings strongly imply *HIC1* as a novel tumor suppressor in a subset of DLBCLs.

Losses on chromosome 6q are known to be associated with ABC-DLBCL and believed to target two tumor suppressor genes: *PRDM1* and *A20 (TNFAIP3)*. The *TNFAIP3* gene, is now known to be an important tumor suppressor gene in DLBCL²⁷. Its deletion contributes to the constitutive activation of *NF-kB* pathway, which is one of the main features of ABC-DLBCL-subtypes. Notably, Bertoni's group reported that the same gene is frequently deleted in marginal zone lymphoma (MZL) ²⁹ and inactivation of *A20* by the typical two-hit mechanism, has been detected in Hodgkin lymphoma and PMBL, all frequently associated with constitutive activation of the *NF-kB* pathway. Thus, *A20* inactivation may represent a common mechanism for constitutive *NF-kB* activation, which may contribute to lymphomagenesis by stimulating cell proliferation and survival. The

identification of genetic lesions in ABC-DLBCL and MZLs that deregulate NF- κ B activity offers the possibility of targeted therapeutic intervention.

Interestingly we have observed that an important difference distinguishing Richter's Syndrome (RS) from de novo DLBCL is the absence of deletions affecting chromosome 6q in the form transformed from chronic lymphocytic leukemia (CLL) ¹⁰⁸. Chromosome 6q, as already mentioned, contains two tumor-suppressor genes, namely *BLIMP1* and *TNFAIP3/A20*, which are frequently inactivated by deletions or mutations in ABC-DLBCL. Differently from RS, chromosome 6q deletion is also involved in another model of lymphoma transformation, namely follicular lymphoma (FL) transformation into DLBCL. The absence of 6q deletions in RS indicates that inactivation of genes mapped on 6q could not be relevant for transformation of CLL into DLBCL and suggests that different types of transformation from indolent lymphoproliferations to DLBCL might be driven by distinct genetic lesions.

Among homozygous deletions, some targeted well known tumor suppressor genes already discussed, such as *CDKN2A* (9p21) or *FAS* (10q23.31), and *TNFAIP3/A20* itself ²⁷. A homozygous deletion affected 19p13, possibly overlapping to what reported in mantle cell lymphoma ¹⁰⁹. However, the lesion was still relatively large in size with several genes as possible targets, including *C3* (complement precursor 3), *TNFSF9* and *TNFSF14* (tumor necrosis factor ligand superfamily 9 and 14).

For the current work a part of the cases has been classified both by GEP or IHC. Comparison between the two methodologies highlighted some differences and the most important one was the presence of 18q gain in GCB-DLBCL cases identified by IHC. This result is consistent with the previous observation by Haarer et al ¹¹⁰, where IHC data are comparable with GEP data only in the 70% of cases. For this reason it will be interesting to evaluate this difference once more after the introduction of the new algorithm proposed by Choi et al ¹¹¹, which added two new markers (*FOXP1* and *GCET1*) to the well known Hans's algorithm and this issue is currently investigated in our lab.

In conclusion our findings are consistent with the genomic complexity and widespread molecular changes associated with DLBCL.

Further studies are needed to elucidate whether our selected regions and genes included, can be consider relevant candidates in the pathogenesis of the disease. The next chapter will show data regarding the characterization of one of the identified lesions.

4. GENOMIC LESIONS ASSOCIATED WITH A DIFFERENT CLINICAL OUTCOME IN DLBCL TREATED WITH R-CHOP-21

4.1 ABSTRACT

As previously discussed, DLBCL is defined as a neoplasm of large B-lymphoid cells with nuclear size more than twice the size of a normal lymphocyte, which has a diffuse growth pattern. DLBCL, despite being such an heterogeneous type of NHL, represent a potentially curable disease when treated with RCHOP-21 (rituximab, cyclophosphamide, doxorubicine, vincristine and prednisone repeated every 21 days) which is considered the current standard chemotherapy. Nevertheless, the progression free survival (PFS) in elderly patients is still below 55% and in younger patients does not exceed 85%.

Gene expression profiling (GEP) patterns and genomic aberrations have been identified with an impact on the clinical course of DLBCL patients treated with CHOP ^{1, 80, 112}. Differently from GEP data, no information is available on the prognostic impact of genomic aberrations detected using high resolution single nucleotide polymorphism (SNP) array in DLBCL patients treated with R-CHOP.

To evaluate the impact on the clinical outcome of genomic aberrations, detected using a high-density genome wide-SNP-based array in a population of DLBCL patients treated with R-CHOP-21. 166 DNA samples were analyzed using the GeneChip Human Mapping 250K Nspl (Affymetrix, Santa Clara, CA, USA). Genomic anomalies were analyzed regarding their impact on the clinical course of 124 patients treated with R-CHOP-21. Unsupervised clustering was performed to identify genetic related subgroups of patients with a different clinical outcome¹⁰³.

Results showed that twenty recurrent genetic lesions had an impact on the clinical course. Loss of genomic material at 8p23.1 showed the strongest statistical significance and was associated with additional aberrations, such as 17p loss and 15q loss. Unsupervised clustering identified five DLBCL clusters with distinct genetic profiles, clinical characteristics and outcome ¹⁰³.

ArrayCGH allowed us to identify new genomic aberrations prevailing in poor-risk patient. Moreover, we identified different genomic related subgroups with distinct clinical characteristics and courses ¹⁰³.

4.2 AIM OF THE CHAPTER

In this part of the study, we analyzed the impact of recurrent aberrations revealed by the high-density genome wide SNP-based array on clinical outcome in a subset of DLBCL patients treated with RCHOP-21 and presented in chapter 3. To identify those aberration influencing the Overall survival (OS), the PFS and the disease free survival (DSF) we integrated genomic profiling data and clinical parameters obtained for 124 patients. Moreover we performed unsupervised clustering using the non-negative matrix factorization (NMF) ⁹⁷algorithm to identify biological subgroups according to the genetic profile.

4.3 RESULTS

4.3.1 GENOMIC REGIONS AND OUTCOME IN PATIENTS

TREATED WITH R-CHOP

Table 4.1 shows the clinical features of the 124 cases treated with R-CHOP-21 and for which follow-up data were available. The clinical characteristics were comparable with published series of DLBCL patients ¹, ¹¹³. The age was advanced and there was a prevalence of female.

The complete remission (CR) rate was 81% (100/124), partial remission (PR) 12% (15/124), stable (SD) or progressive disease (PD) 7% (9/124).

Table 2.1 Clinical characteristics of 124 DLBCL patients treated with R-CHOP. PS= Presentation Sites; IPI= International Prognostic Index; GCB= Germinal Center B-Cells; BM= Bone Marrow; LDH= Lactate dehydrogenase; HCV= Hepatitis C virus.

Clinical characteristics	n	valid	percent
Median age, years (range)		64 (18-86)	
Sex (m:f)	58:63	121	48:52
PS >1	24	119	20
LDH > normal	62	110	57
Extranodal sites >1	24	115	24
Stage III/IV	80	122	66
Bulky disease >10cm	24	109	22
IPI score >1	77	103	75
B-Symptoms	40	119	34
Beta2-MG > normal	24	43	56
GCB	48	78	55
HCV infection	10	76	13
Only extranodal disease	15	104	14
BM involvement	21	111	20

Among the complete responders, 27 (22%) eventually relapsed. After a median follow-up of 23 months (range, 1-123 months), 3-year estimated of OS was 88%, PFS 65% and DFS 71%, respectively. Disease caused 9/19 deaths (47%).

Univariate analysis identified elevated serum levels of LDH, B-symptoms at time of diagnosis, bone marrow involvement and the achievement of a CR with R-CHOP as strongest factors influencing the clinical course of the patients (data not shown).

Figure 4.1 shows the frequencies of CN changes for the 100 patients who obtained a CR after R-CHOP and for the 24 who did not (panel A). The following lesions were more frequent in the non-CR patients: loss of 8p (5/24, 21%, vs 5/100, 5%; $p=0.023$; $q_{\text{value}}=0.51$) and 8p21.2-p23.1 (5/24, 21%, vs 6/100, 6%; $p=0.036$; $q_{\text{value}}=0.58$), homozygous deletion at 15q11.2 (2/24, 8%, vs 0/100, 0%; $p=0.036$; $q_{\text{value}}=\text{n.d.}$) and LOH of 8p23.1 (6/24, 25%, vs 4/100, 4%; $p=0.003$; $q_{\text{value}}=0.24$), 9p (3/24, 12.5%, vs 2/100, 2%; $p=0.049$; $q_{\text{value}}=\text{n.d.}$) and 15q21.3 (5/23, 22%, vs 5/100, 5%; $p=0.019$; $q_{\text{value}}=0.51$).

LOH analysis of the 124 patients (figure 4.1, panel B) showed a recurrent 9p21.3 UPD (including the CDKN2A gene) in 13/100 of the CR cases and 6/24 of the non-CR cases. As described in chapter 1, UPD is often due to epigenetic inactivation of alleles. To investigate whether mechanisms of methylation might be involved in the silencing of CDKN2A and if such observation might be associated with the detection of a UPD, we decided to perform PCR analysis using primers on methylated and unmethylated sequences of CDKN2A. Methylation was detected in 3/13 (07-213; 07-485; 07-506) of CR cases and in 2/6 (07-362; 08-262) of the non responders, presumably in both alleles (figure 4.2), suggesting that UPD can be associated with the silencing of the gene CDKN2A.

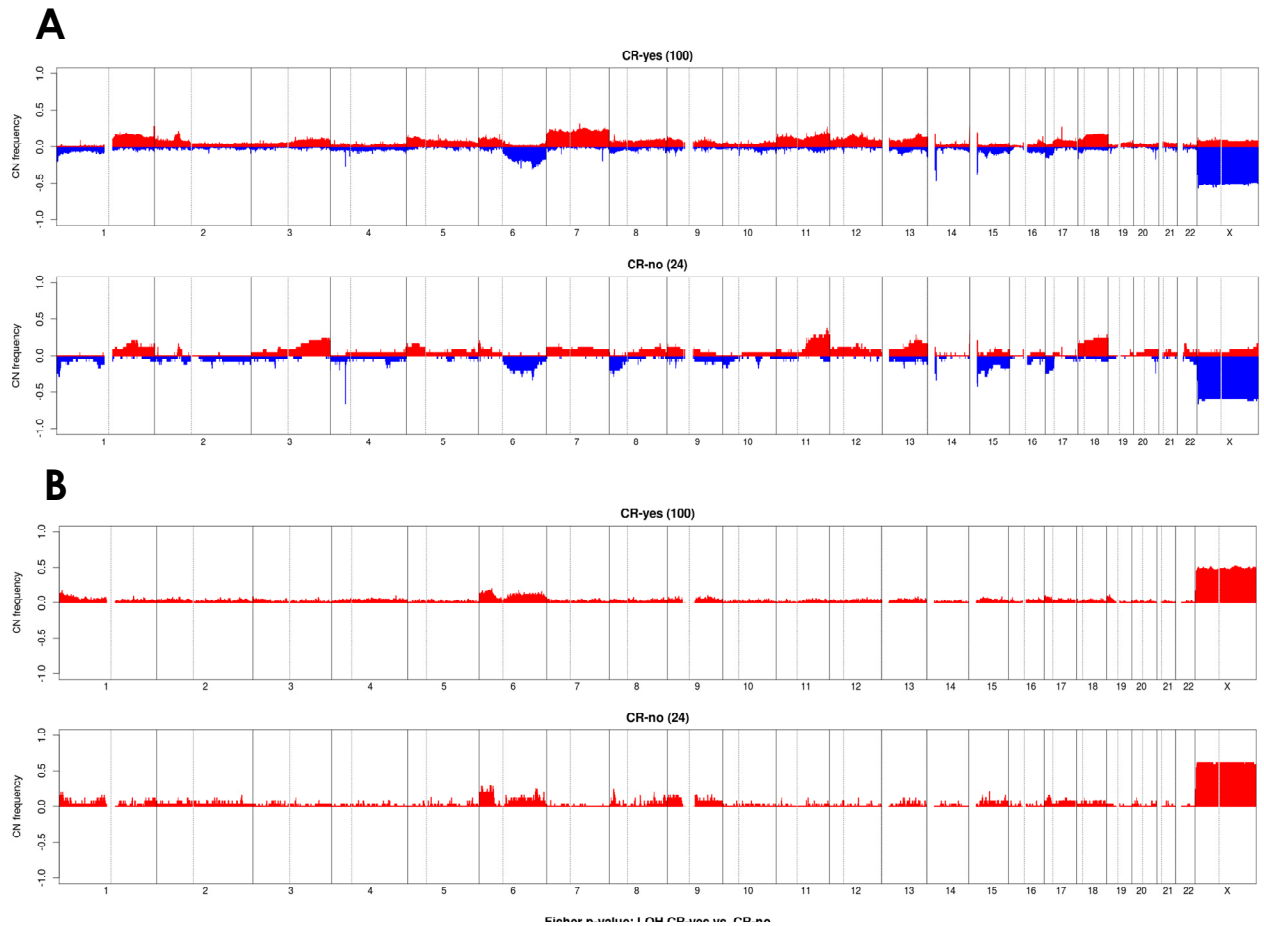


Figure 4.1. Panel A: Frequency of DNA gains (red) and losses (blue) observed in 124 patients according to complete remission after treatment. Upper panel shows frequency plot of CR patients, lower panel shows frequency plot of non-CR patients. Panel B: Frequency of DNA LOH (red) observed in 124 patients according to complete remission after treatment. Upper panel shows frequency plot of CR patients, lower panel shows frequency plot of non-CR patients. X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

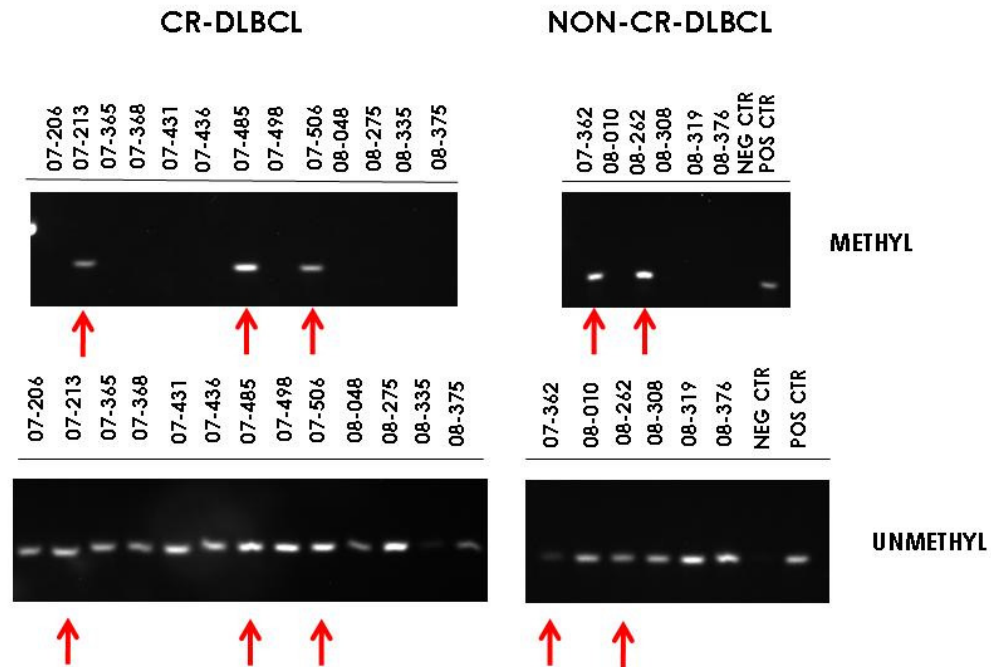


Figure 4.2. RT-PCR analysis showing methylation analysis of CDKN2A in a panel of CR and NON-CR DLBCL patients. Upper panel shows methylation rate of CDKN2A in CR and NON-CR samples showing UPD after LOH analysis; lower panel shows expression of unmethylated CDKN2A in CR and NON-CR samples showing UPD after LOH analysis. Red arrows indicate patients with methylation of CDKN2A.

MCR were then evaluated for their impact on the patients' clinical course. Twenty regions showed a statistical significant impact on OS, PFS and/or DFS (Table 4.2).

Table 4.2. Twenty regions with a statistical significant impact on OS, PFS and/or DFS. * ≤ 5 patients with the genomic alteration. hdel, homozygous deletion; ampl, amplification.

Region	survival time: p-value / q-value
12q42.2 LOH*	PFS: <0.001 / <0.001; OS: 0.007 / n.s; DFS: 0.034 / n.s.
6q24.1-q24.2 LOH	PFS: 0.021 / n.s.
9p21.1 LOH	PFS: 0.047 / n.s.
10q23.31 hdel*	PFS: < 0.001 / < 0.001; OS: <0.001 / <0.001
15q11.2 hdel*	PFS: < 0.001 / 0.001; OS: <0.001 / <0.001
18p11.32 loss	PFS: 0.027 / n.s.
8p23.1 LOH	OS: 0.015 / n.s.
6q23.3-q24.1 hdel*	OS: 0.035 / n.s.
8p23.1 loss	OS: 0.002 / 0.1
8p23.1-p21.2 loss	OS: 0.012 / n.s.
8p loss	OS: 0.01 / n.s.
9p24.2 loss	OS: 0.047 / n.s.
9q31.1 LOH	OS: 0.04 / n.s.
15q12-q21.2 loss	OS: 0.034 / n.s.
4p12-p11 loss	DFS: 0.028 / n.s.
4q32.2 LOH*	DFS: 0.045 / n.s.
6q22.32-q22.33 LOH	DFS: 0.003 / n.s.
6q25.1 LOH*	DFS: 0.007 / n.s.
9p24.3-p21.3 ampl*	DFS: <0.001 / 0.001
14q11.2 hdel*	DFS: 0.001 / 0.077

In order to identify possible pathways affected by these genomic regions, we mapped the genes contained within MCR to known biologic pathways. Seven pathways appeared as statistically significantly enriched: natural killer cell mediated cytotoxicity (21 genes, $p=6.8 \times 10^{-12}$, $q_{\text{value}}=1.4 \times 10^{-9}$), regulation of autophagy (12 genes, $p=1.7 \times 10^{-11}$, $q_{\text{value}}=1.7 \times 10^{-9}$), antigen processing and presentation (15 genes, $p=2.6 \times 10^{-9}$, $q_{\text{value}}=1.7 \times 10^{-7}$), Jak-STAT signaling pathway

(17 genes, $p=3.2 \times 10^{-7}$, $q_{\text{value}}=1.6 \times 10^{-5}$), Toll-like receptor signaling pathway (13 genes, $p=3.4 \times 10^{-6}$, $q_{\text{value}}=1.4 \times 10^{-4}$), cytokine-cytokine receptor interaction ($p=5.2 \times 10^{-6}$, $q_{\text{value}}=1.7 \times 10^{-4}$) and apoptosis (7 genes, $p=0.001$, $q_{\text{value}}=0.032$) (Table 4.3).

Table 4.3. Significantly enriched pathways among the MCR with an impact on OS, PFS and DFS.

Term	Count	p-value	Genes	Benjamini
Apoptosis	7	0.011	CHP, TNFRSF10C, FAS, PPP3CC, TNFRSF10B, TNFRSF10A, TNFRSF10D,	0.273
Natural killer cell mediated cytotoxicity	21	2.98×10^{-12}	CHP, IFNB1, IFNA6, IFNA8, SHC4, IFNA21, PPP3CC, TNFRSF10B, IFNA7, IFNA17, IFNA2, IFNA16, IFNA5, TNFRSF10C, IFNA4, FAS, IFNA1, IFNA10, TNFRSF10A, IFNA14, TNFRSF10D,	6×10^{-10}
Regulation of autophagy	12	1.05×10^{-11}	IFNA6, IFNA4, IFNA8, IFNA21, IFNA1, IFNA10, IFNA7, IFNA17, IFNA14, IFNA5, IFNA16, IFNA2	1.06×10^{-9}
Jak-STAT signaling pathway	16	1.04×10^{-6}	IFNA6, IFNB1, IFNA8, IFNA21, JAK2, IFNA17, IFNA7, IFNA2, IFNW1, IFNA16, IFNA5, SPRED1, IFNA4, IFNA1, IFNA10, IFNA14	5.24×10^{-5}
Toll-like receptor signaling pathway	13	2.17×10^{-6}	IFNA6, IFNB1, IFNA8, IFNA21, IFNA17, IFNA7, IFNA2, IFNA16, IFNA5, IFNA4, IFNA1, IFNA10, IFNA14	8.73×10^{-5}
Antigen processing and presentation	15	1.48×10^{-9}	IFNA6, IFNA8, IFNA21, PDIA3, CTSB, IFNA17, IFNA7, IFNA2, IFNA16, IFNA5, IFNA4, IFNA1, IFNA10, B2M, IFNA14	9.91×10^{-8}
Cytokine-cytokine receptor interaction	19	1.12×10^{-5}	IFNB1, IFNA6, IFNA8, IFNA21, TNFRSF10B, IFNA7, IFNA17, IFNA2, IFNW1, IFNA16, IFNA5, TNFRSF10C, IFNA4, FAS, IFNA1, IFNA10, TNFRSF10A, IFNA14, TNFRSF10D	3.76×10^{-4}

Among the four homozygous deletions possibly associated with a poor clinical outcome, two targeted known tumor suppressor genes: *TNFAIP3/A20* at 6q23.3, and *FAS* at 10q23.31.

The remaining two homozygous deletions, comprising gene coding the *TCRA* at 14q11.2 and *OR4N4* and *OR4M2* at 15q11.2, overlapped completely with known CNV.

Although statistically significant, there were only a very low number of cases bearing the four recurrent homozygous deletions, making impossible to ascertain their real impact on the outcome.

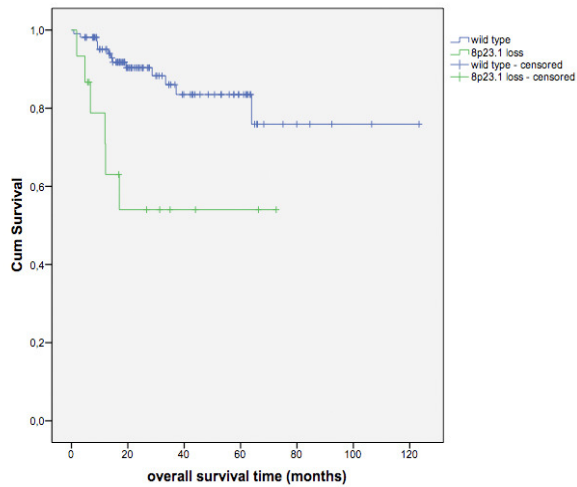
None of the genomic aberrations maintained their role in predicting the outcome in a multivariate analysis including IPI-factors (age >60, poor performance status, elevated LDH, advanced stage and ≥ 2 extranodal sites).

4.3.2 THE IMPACT OF 8p23.1 LOSS

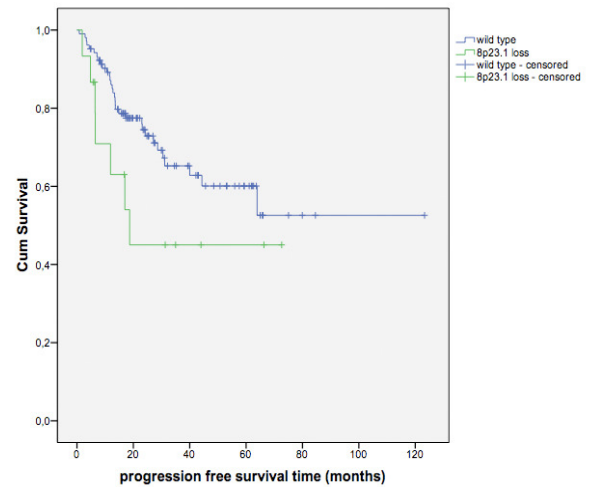
The lesion showing the strongest association with a worse outcome was 8p23.1 loss (Figure 4.3). To elucidate its possible biologic significance, we compared the pattern of genomic aberrations in patients with and without the aberration. Cases with 8p23.1 deletion presented a series of associated lesions. The highest statistical significance was observed for losses occurring on the long arm of chromosome 15 (15q12-q21.2, 11/22, 50%, vs 15/144, 10%; $p = 0.00004$; $q_{\text{value}} = 0.001$), on 17p (9/22, 41%, vs 14/144, 10%; $p\text{-value} = 0.0006$; $q_{\text{value}} = 0.01$), on 10q23.33-q24.1 (7/22, 32%, vs 10/144, 7%; $p = 0.002$; $q_{\text{value}} = 0.03$) and for 11q gains (9/22, 41%, vs 15/144, 11%; $p = 0.0009$; $q_{\text{value}} = 0.02$).

Despite using low stringent criteria ($p\text{-value} < 0.005$; no MTC), a supervised analysis of the gene expression profiles of cases with or without 8p (8/54, 15%) identified only 81 probes, all up-regulated, as differentially expressed, without any enrichment for specific biologic pathways (Supplementary Table 1 in Annex).

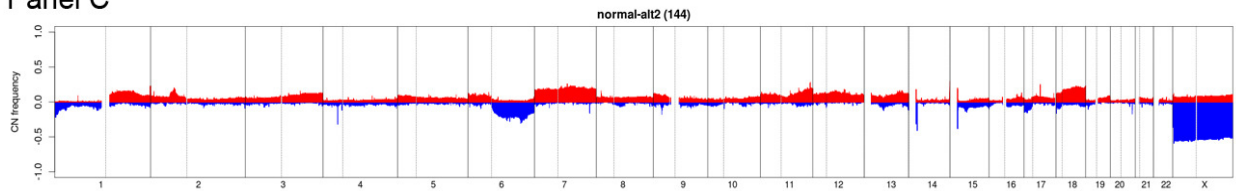
Panel A



Panel B



Panel C



Panel D

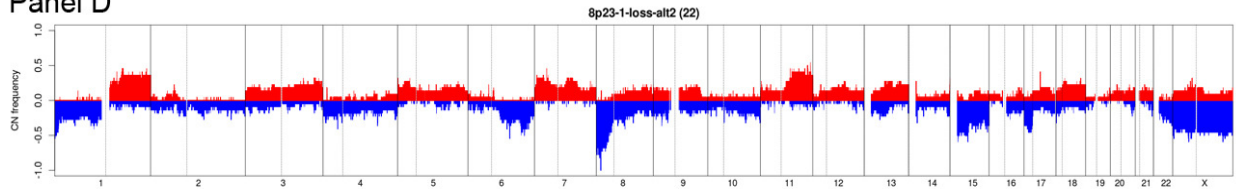


Figure 4.3. Kaplan-Meier graphs: 8p23.1 loss according to OS ($p=0.002$; Panel A) and PFS ($p=0.119$; Panel B) and Frequency of DNA gains (red) and losses (blue) observed in 124 patients according to 8p23.1 loss. Panel C: Frequency plot of wild-type patients, Panel D: Frequency Plot of 8p23.1 loss patient. X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

4.3.4 NMF-CLUSTERING

In order to identify subgroups of patients with different genetic backgrounds and with possible distinct clinical outcome, we performed unsupervised clustering using the NMF algorithm. Five clusters were identified with a distinct genetic profile (Figure 4.4).

Cluster 1 consisted of 77/166 patients (37%). Their genetic profile was very heterogeneous and only very few recurrent aberrations could be observed. The second group (cluster 2) was represented by 28/166 patients (13%) and characterized by gain of 1q and 3q, loss of 1p, 6q and 17p. The third group included 8/166 patients (4%), with the gain of 2p, 11p and 11q and losses of 1p, 6q, 9p, 15q, 17p and 18p. The fourth group consisted of 38/166 patients (17%). Their genetic profile showed gain of 1q, 2p, 6p, 7, 12q and losses of 6q, 8p, 15q and 17p. Finally, the fifth group was represented by 15/166 patients (6%). It was characterized by the gains of 3q, 11q, 13q, 18q and losses of 1p, 6q, 9p and 15q.

Clinical data were available for 61/66 patients of cluster 1, 22/28 of cluster 2, 5/8 of cluster 3, 29/38 of cluster 4 and 7/15 of cluster 5. Due to the small number of patients, we did not further analyze clusters 3 and 5. Due to the distinct genomic profiles we compared cluster 1 versus clusters 2 and 4 and cluster 2 versus cluster 4 (Table 4.5). Cluster 1 showed a lower percentage of older patients ($p=0.068$), significantly higher rate of BM involvement (16/57, 28% vs 1/19, 5% and 2/24, 8%; $p=0.008$) and HCV infection (9/38, 24% vs 1/12, 8% and 0/20; $p=0.014$).

Clusters 2 and 4 showed differences in the number of patients with a high IPI-score (9/16, 56% vs. 7/26, 27%; $p=0.057$) and in single IPI factors, CR-rate (17/22, 77% vs. 25/29, 86%; $p=n.s.$) and relapse rate (5/17, 29% vs. 2/25, 8%; $p=0.068$). Cluster 1 was not only characterized by a distinct genomic profile and different clinical features, such as high HCV infection and BM infiltration, but, similar to cluster 2, the outcome was relatively poor when compared to cluster 4: high relapse rate (26%), 3-year estimated OS was 83%, PFS 61% and DFS 66%, respectively. Moreover, unlike cluster 2 and 4 it did not reach a plateau in PFS (figure 4.5).

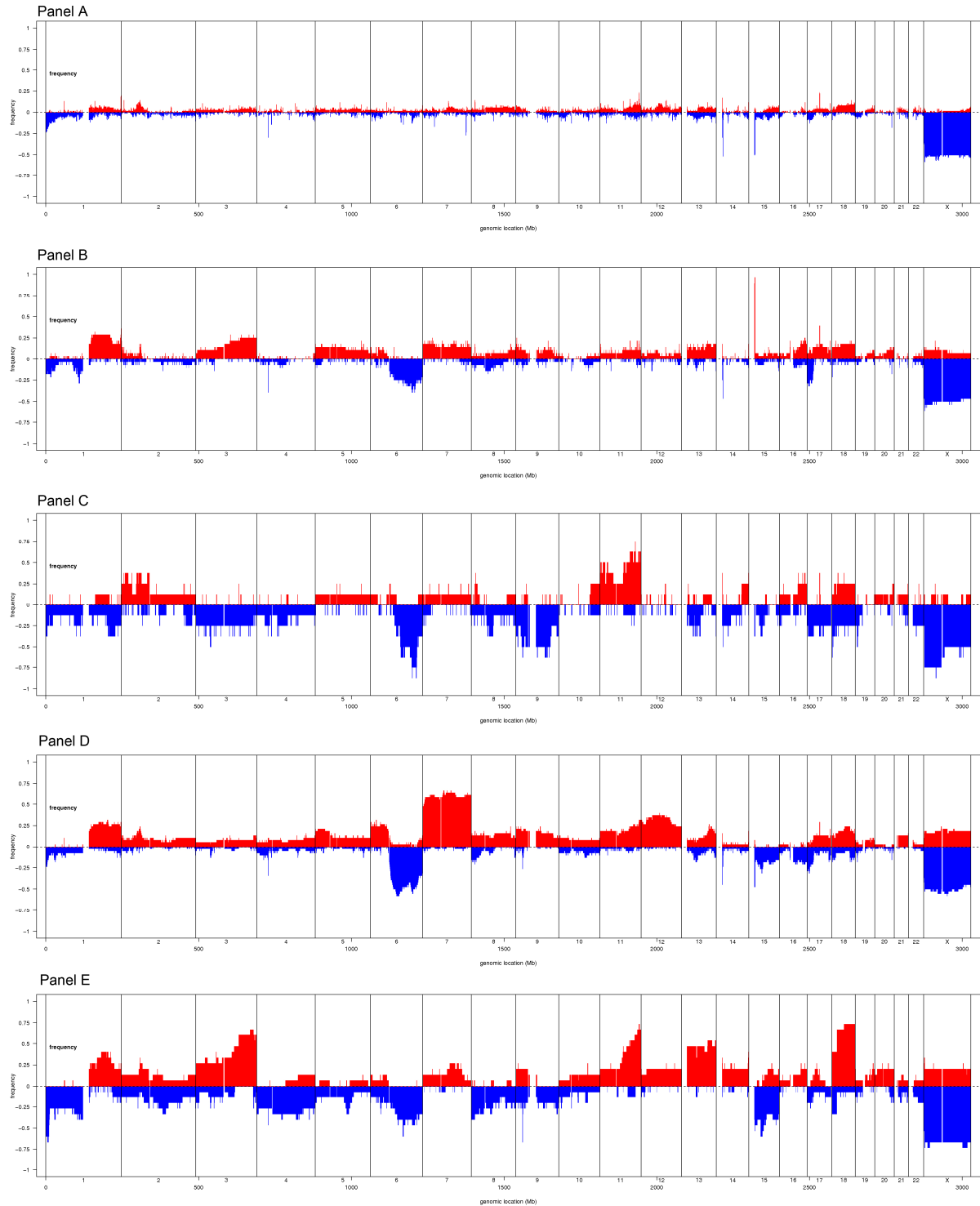


Figure 4.4. Frequency plot of 166 patients showing gains (red) and losses (blue) according to the different NMF clusters. Panel A: Cluster 1, panel B: Cluster 2, panel C: Cluster 3, panel D: cluster 4, panel E: cluster 5. X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases showing the aberrations.

Table 4.4. Patient characteristics according to the three main NMF clusters.

Parameter	Cluster 1 (%) (n=61)	Cluster 2 (%) (n=22)	Cluster 4 (%) (n=27)
Median age (range), years	63 (18-82)	68 (24-83)	66 (42-86)
Age > 60	56	64	79
Elevated LDH	55	65	48
Advanced stage	67	60	71
ECOG PS >2	19	33	18
>2 Extranodal sites	28	15	19
IPI ≥ 2	67	60	71
Bone marrow involvement	28	5	8
B-Symptoms	33	45	26
Bulky disease	17	29	22
HCV-infection	24	8	0
GCB	25	10	15
Median follow up (range), months	26 (1-106)	20 (9-92)	23 (2-123)
CR-rate	82	77	86
Relapse rate	26	29	8

A supervised analysis of gene expression profiles of cases classified as cluster 1 versus cases belonging to clusters 2 and 4 identified 546 probes (p-value < 0.005; no MTC) (Supplementary table 2). The up-regulated transcripts were 323, while the down-regulated were 223. Genes of the T-cell receptor signaling pathway were the only significantly over-represented group of genes (p-value = 2.E-06; Benjamini-corrected p-value = 4.E-04). All 8/166 DLBCL cases described as T-cell rich were classified as cluster 1.

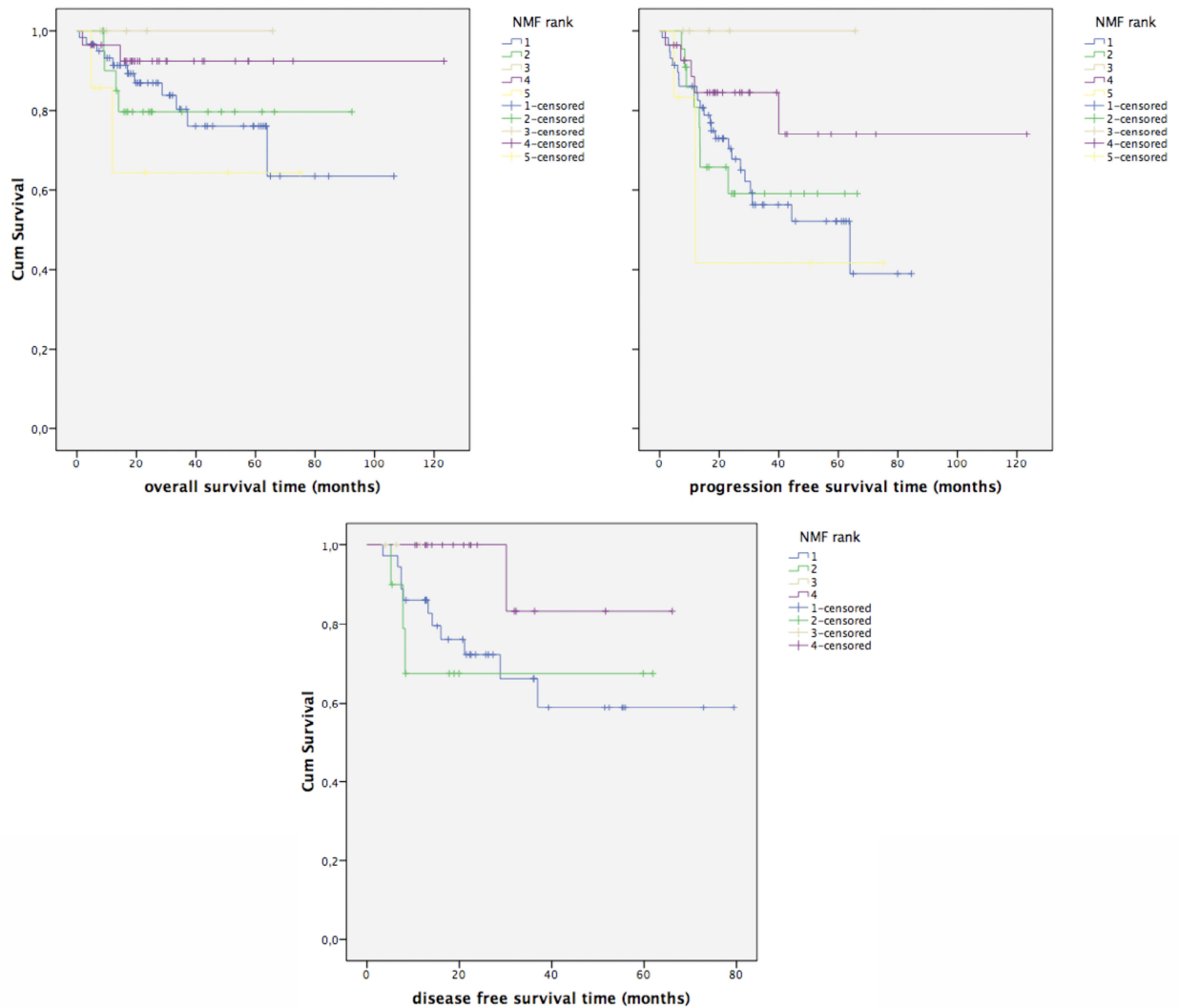


Figure 4.5. Kaplan-Meier curves of OS (Panel A), PFS (Panel B) and DFS (Panel C) according to the different NMF clusters. Blue line refers to cluster 1; green line refers to cluster 2; beige line refers to cluster 3; purple line refers to cluster 4; yellow line refers to cluster 5.

4.4 DISCUSSION

DLBCL is an aggressive disease characterized by heterogeneous biological features likely leading to distinct clinical presentations and courses ¹¹⁴. Several arrayCGH-studies of CHOP treated DLBCL patients have been performed ^{4, 114}. However, to our knowledge, there is only few data about the impact of aberrations identified with high resolution SNP-array on the clinical course of DLBCL cases after CHOP in combination with the monoclonal anti-CD20 antibody rituximab (R-CHOP), which currently represents the state of the art treatment ⁹¹. Furthermore, clustering analysis is widely used in GEP to identify disease subgroups with a distinct clinical course, whereas up to now there are no data regarding clustering analysis of high resolution genomic profiling arrays in DLBCL combined with survival data. Therefore, we conducted a retrospective multi-center study to evaluate the impact of unbalanced genomic aberrations as detected using a high-density genome wide-SNP-based array on OS, PFS and DFS in a population of 124 DLBCL patients treated with R-CHOP. We also performed unsupervised clustering to identify possible subgroups of patients sharing a common biology and, possibly, clinical outcome.

Despite our series could have been biased due to the requirement of available frozen material, the observed clinical parameters were comparable to previously published series of patients treated with R-CHOP outside clinical trials ¹, thus reflecting the general population of DLBCL patients. To evaluate the impact of genomic aberration on clinical outcome, we first compared the profile of patients achieving versus patients not achieving CR. Losses of 8p, a bona fide CNV at 15q11.2, and LOH at 15q21.3 and at 9p (*CDKN2A*) were associated with treatment failure with R-CHOP, but none were of high statistical significance.

PCR analysis of the methylated and unmethylated status of *CDKN2A* in a subset of CR and non-CR patients apparently showing UPD of the 9p21.3 locus after LOH analysis, allowed us to demonstrate the hypothesis that epigenetic

mechanisms inactivating genes, such as CDKN2A, could be associated with UPD events. In fact, 3/13 cases among CR and 2/6 of non-CR patients showed methylation of CDKN2A at least in one allele. Interestingly, the two non-CR cases seemed to show methylation of both alleles of CDKN2A, suggesting that there might be an association of UPD not only with epigenetic silencing of target genes, but also with the lack of response to the standard chemotherapy in DLBCL patients. A study on a bigger number of patients could clarify if such result happened by chance or if it could be interesting to perform further characterization.

Then, we investigated the impact of single MCRs on OS, PFS and DFS: Twenty regions showed a significant correlation on clinical outcome (p-value < 0.05). Loss of genomic material at 8p23.1 showed the most statistically significant impact on OS. Loss of 8p has been previously shown associated with poor characteristics and worse outcome in lymphoid neoplasms ¹¹⁵⁻¹¹⁷. Since the pattern of deletions and LOH in this regions was very heterogeneous, no clear candidate genes could be identified: different, only partially overlapping, deletions affecting the short arm of chromosome 8, including the loss of the whole 8p, showed a strong or less strong association with a worse outcome. It was recently reported by Bosh et al ¹¹⁸, that the 8p23.1 region contains a 3.8–4.5 Mb segment which can be found in different orientations (defined as genomic inversion) among individuals. Moreover, an impact of 8p23.1 inversion on gene expression levels cannot be excluded, since four genes (*NEIL2*, *MSRA*, *CTSB* and *BLK*) from this region have statistically significant different expression levels depending on the inversion status ¹¹⁸. To which extent these gene expression differences can influence the function of the respective proteins and if they are directly related to the inversion of the region remains to be proven. Taken together these observations have indicated that further characterization of the 8p23.1 loss in DLBCL might be worth. Indeed, validation of these findings is on going on a very large series of cases in collaboration with a US consortium. As also reported for 17p deletion, which could affect more than a single gene ⁷⁶, 8p loss could deregulate different genes, including TRAIL-R1 and TRAIL-R2 ¹¹⁷.

Losses of 17p (*TP53*, *HIC1*) and 15q (*TP53BP1*) were observed at a statistically higher prevalence in association with the loss of 8p, suggesting that the possible poor outcome after R-CHOP could be due to the concomitant deregulation of a series of genes mapped on different chromosomes. Indeed, genes involved in DNA repair and apoptosis are mapped in these regions and could be all inactivated in patients bearing these deletions: *TRAIL-R1* and *TRAIL-R2* (8p), *TP53* and *HIC1* (17p) and *TP53BP1* (15q) ^{75, 117, 119}. Interestingly, in DLBCL, loss of *TP53* correlates with low levels of *TRAIL-R2* ⁷⁵, and, loss of 8p is a frequent phenomenon in patients with chronic lymphocytic leukemia with 17p loss, possibly identifying a subgroup of patients with an extremely poor outcome ¹¹⁵. Regarding *TP53*, in our series, the loss of 17p13.1 did not show any impact on the clinical outcome, similarly to what previously reported by Young et al ⁷⁵ who showed a negative influence of *TP53* mutations but not of 17p13.1 loss on the clinical course in their series of DLBCL patients.

Functional annotation analysis and pathway mapping identified seven significantly enriched pathways, among the genes of the regions with an impact on survival. The most important genes mapped were *TRAIL-R*, *FAS* (natural killer mediated cytotoxicity and apoptosis) and *JAK* (*JAK-STAT* pathway). Both, *TRAIL-R* and *FAS* are genes coding for proteins with a crucial role in regulation of cell death and they are known to be involved in mechanisms for tumor escape ¹²⁰. Therefore, the loss we observed strengthens the notion that both may play an important role in this disease. On the contrary, *JAK2*, a protein able to promote cell transformation by constitutive activation of the *JAK-STAT* pathway ^{121, 122}, was amplified. The specific role of this aberration in DLBCL is yet unclear, whereas it has been shown that *JAK2* genomic gains are frequent in PMBL and Hodgkin's disease ⁸⁷.

Unsupervised NMF-clustering identified five clusters with a distinct genetic profile, and some with clear different clinical characteristics and distinct courses, without a significant bias on geographic provenance. Cluster 1 patients were characterized by a heterogeneous genomic profile, lacking of most recurrent lesions and probably the worst long term outcome among the

different clusters. These patients had a high rate of HCV infection and BM involvement and a slightly younger age. GEP and morphology data suggested that this cluster could represent DLBCL cases having a high content of infiltrating T-cells, which would also explain the low rate of genomic aberrations due to a lower content of neoplastic cells. These data could delineate a DLBCL subtype with distinct clinical features and a high rate of inflammatory cells. Similar characteristics can be found in a DLBCL subgroup, termed host response (HR) cluster, identified applying unsupervised clustering on gene expression data ².

Similarly to our cluster 1, HR cluster reveals increased expression of T-cell molecules and, when compared with other DLBCL cases, the patients are significantly younger, have a higher incidence of bone marrow involvement and a lack of the recurrent chromosomal translocations involving *BCL2* or *BCL6* ². Since we observed a high prevalence of HCV infection and since viruses can induce epigenetic transforming changes in human cells ¹²³, at least part of cluster 1 DLBCL, as well of the HR cluster, might have developed mainly through epigenetic changes more than genomic aberrations. Alternatively, a transforming process predominantly driven by somatic mutations could be hypothesized. Similar to cluster 1, the outcome of the cluster 2 patients was relatively poor. This subgroup was characterized by 1q gain, 3q gain, 6q loss and 17p loss. On the converse, cluster 4 patients, mainly characterized by gains of chromosome 7 and 12, presented a good outcome. Based upon their genomic profiles, cluster 2 was reminiscent of post-GCB-DLBCL while cluster 4 of the GCB-type. Among the remaining two clusters, cluster 3 was characterized by losses of 6q, 9, and 11q gains and had a trend for a good outcome. Instead cluster 5 (gains of 3q, 11q and 18q; losses of 6q, 8p, 15q and 17p) showed a poor outcome. However, the number of patients classified in the last two clusters was too small to draw any conclusion.

In conclusion, the application of a genome-wide SNP-based arrayCGH allowed us to identify genetic features possibly associated with a different outcome. When confirmed in other series, loss of 8p might identify a group of patients with worse outcome when treated with R-CHOP, and characterized by a series of

concomitant additional aberrations such as 17p and 15q losses. By applying unsupervised clustering, a distinct subgroup of patients characterized by a low frequency of recurrent aberrations, high prevalence of HCV infection and bone marrow involvement and poor outcome with R-CHOP could be identified; this subgroup deserves to be better characterized in order to design *ad hoc* therapeutic strategies.

5. MOLECULAR AND FUNCTIONAL CHARACTERIZATION OF A RECURRENT DNA GAIN ON 11q24.3

5.1 ABSTRACT

Analyzing the genome-wide DNA profiles of 166 DLBCL cases collected at time of diagnosis and analyzed with the Affymetrix 250k array (as described in Chapter 3 and 4), we detected a recurrent DNA gain on chromosome 11q24.3 occurring in 30% of the patients. Genomic gains are one of the mechanisms leading to oncogenes activation. Genes affected by such aberrations deserve further characterization in terms of expression and functionality and can be exploited as possible prognostic markers and /or therapeutic targets.

To identify genes included in the gain, we analyzed the expression level of the genes mapped in the 11q24.3 MCR, in a subset of DLBCL patients and in a series of DLBCL cell lines by Real-Time-PCR. Seven genes mapped in the locus 11q24.3 (*ETS1*, *FLI1*, *C11ORF45*, *KCNJ11*, *KCNJ5*, *P53AIP1*, *RICS*), but a first analysis on a pool of CD19+ B-cells and on a pool of cells deriving from thymus and lymphonode, showed that only two of them, *ETS1* and *FLI1*, were found to be expressed. Second analysis by Real-Time-PCR confirmed their high expression in 3/17 DLBCL patients carrying the 11q24.3 gain. Further genomic analysis of DLBCL cell lines allowed us to identify the gain of 11q24.3 in the OCI-LY7 DLBCL cell line. Real-time-PCR and Western Blot (WB) analysis confirmed a higher expression of the two proteins in this cell line. Thus we selected OCI-LY7 as a candidate model to perform loss of function studies for further characterization of *ETS1* and *FLI1* role in DLBCL.

Our data show a possible involvement of these two genes in an alternative mechanism of lymphomagenesis. In fact down-regulation of *ETS1* in OCI-LY7, using shRNA strategy, leads to the reactivation of *BLIMP1*, the master regulator

of B-cells differentiation into plasma cells ^{124, 125} (PCs) and to the consequent down-regulation of one of its target, *PAX5*, involved as a negative regulator in the same pathway ⁴⁷. The *ETS1* knock-down caused also a block in the proliferation rate of OCI-LY7 cell line. *FLI1* down-regulation, also using shRNA strategy, was followed by the down-regulation of *PAX5* gene, independently of *ETS1*, and by an apparent block of the proliferation rate of the OCI-LY7, leading us to the conclusion that *FLI1* could be an activator of *PAX5*.

Taken together these observation allowed us to propose a possible new role for *ETS 1* and *FLI1* in the pathogenesis of DLBCL by contributing to a possible alternative molecular mechanism due to the presence of the 11q24.3 gain, in which *ETS1* up-regulation seems to play a key role in maintaining B-cells in a Germinal Center-status (GC-status) by blocking the activity of *BLIMP1*. On the other side, *FLI1* takes part to the same scenario by activating *PAX5* and giving the advantage to malignant cells to not further differentiate into normal PCs.

5.2 AIM OF THE CHAPTER

Our aim was to detect causative lesions in DLBCL genome as the one being recurrent in our data set and with a biological effect on genes, which might be involved in pathways that participate in lymphomagenesis. In Chapter 3, we have identified the 11q24.3 gain in 30% of DLBCL patients. Here we have tried to identify candidate genes targeted by the recurrent aberration and to understand their biological role in DLBCL.

5.3 RESULTS

5.3.1 DISTINCT RECURRENT ALTERATIONS IN DLBCL AND IDENTIFICATION OF THE 11q24.3 GAIN

Among the 58 recurrent gains, identified in 166 DLBCL samples, the 11q24.3 gain occurred in 30% of patients (section 3.3.1). As shown in figure 5.1 it was a relatively small gain (1Mb) (Upper panel).

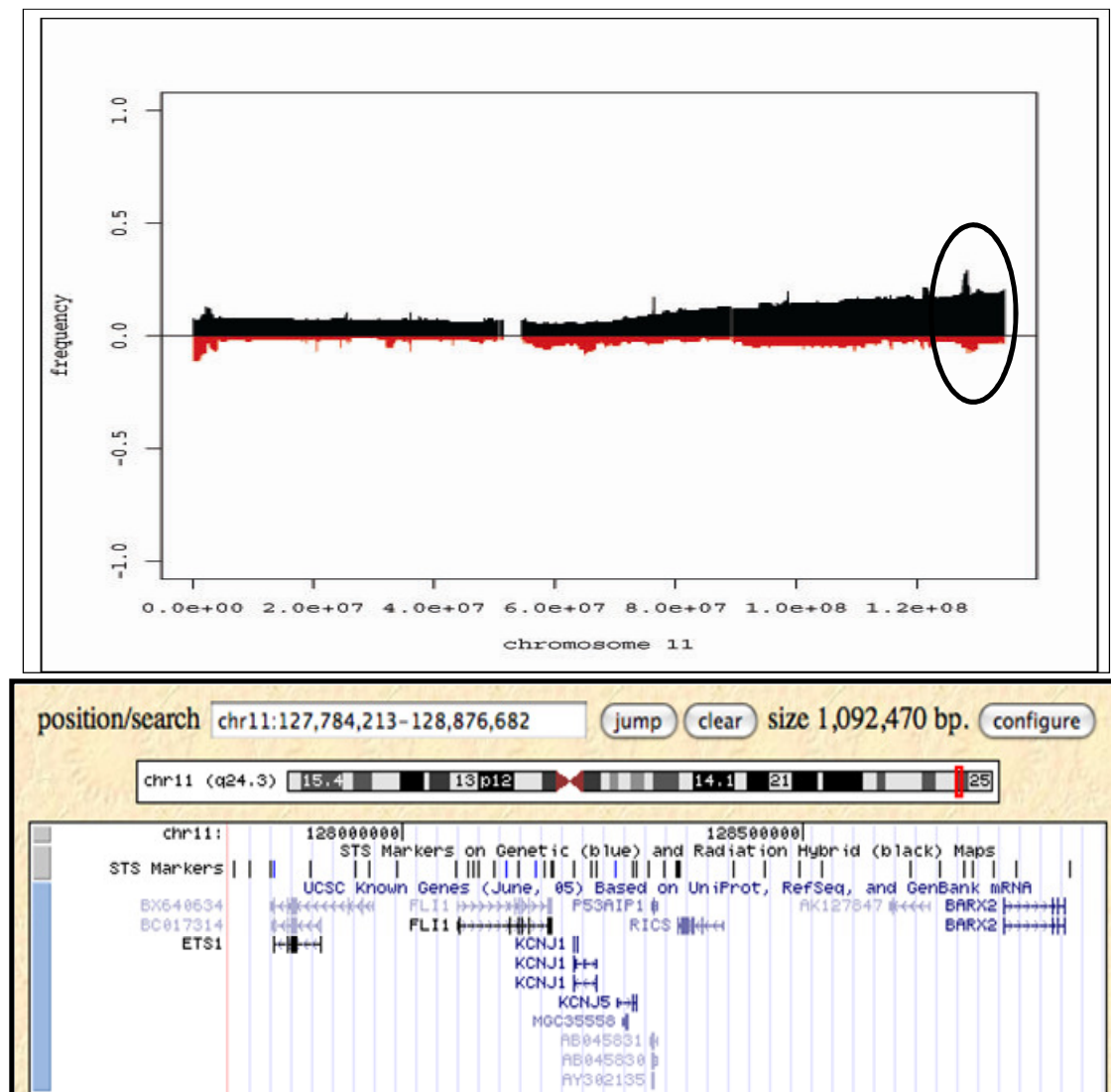


Figure 5.1. The 11q24.3 occurring in 30% of patients is a very specific peak of gain (upper panel); genes localized in 11q.24.3 locus. UCSC data base identified a set of seven genes.

5.3.2 IDENTIFICATION OF PUTATIVE TARGETS WITHIN 11q24.3 GAIN

We were interested in the putative targets of the 11q24.3 gain. Figure 5.1 (bottom panel) represents the minimal gained region and the corresponding annotation with respect to Refseq genes visualized on the genome browser of the database UCSC (<http://genome.ucsc.edu/>). Seven genes mapped in the 11q24.3 (*ETS1*, *FLI1*, *C11ORF45*, *KCNJ1*, *KCNJ5*, *P53AIP1*, *RICS*) and their expression was evaluated using RT-PCR analysis on a pool of CD19 positive B-cells obtained from a healthy donor and on a pool of cells deriving from thymus and lymphonode. Only *ETS1* and *FLI1*, members of ETS family transcription factors, were expressed. Then we decided to perform a real-time-PCR in all the cell lines, to quantify the expression level of these two factors. The level of *FLI1* expression was higher in SUDHL6, DOHH2, OCI-LY7 and OCI-LY19 DLBCL cell lines, while *ETS1* was higher in SUDHL6, DOHH2 and OCI-LY7 (Figure 5.3). To understand if possible genomic aberrations could be associated with the observed higher expression, we also performed genomic analysis of the entire panel of DLBCL cell lines previously mentioned. A gain of 11q24.3 was detected in OCI-LY7, which could explain the higher rate of mRNA expression of *ETS1* and *FLI1* in this cell line (figure 5.2 and 5.3). A loss of the region 11q23.1-qter was demonstrated in KARPAS422, elucidating the lack of expression of both genes in this cell line. SUDHL-6 showed a deletion of the 11q24.3, but the high expression of the two genes *ETS1* and *FLI1* in this cell line, might be explained by the presence of a possible translocation. Indeed a small interstitial deletion was observed centromeric to *ETS1*.

OCI-LY7 GENOMIC PROFILE

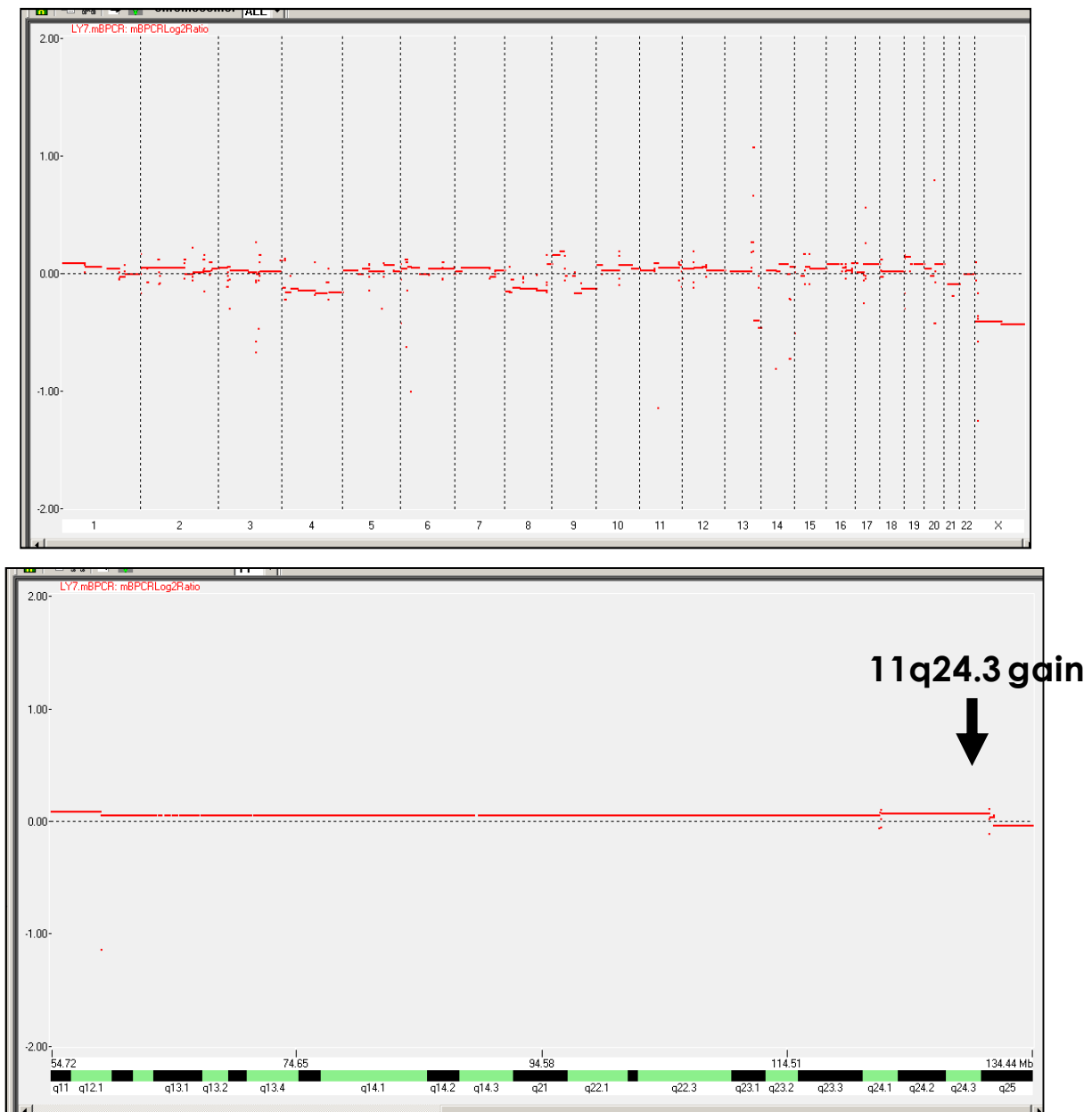


Figure 5.2. Genomic profile of OCI-LY7 (upper panel); figure shows also a detail of the 11q24.3 gain (lower panel)

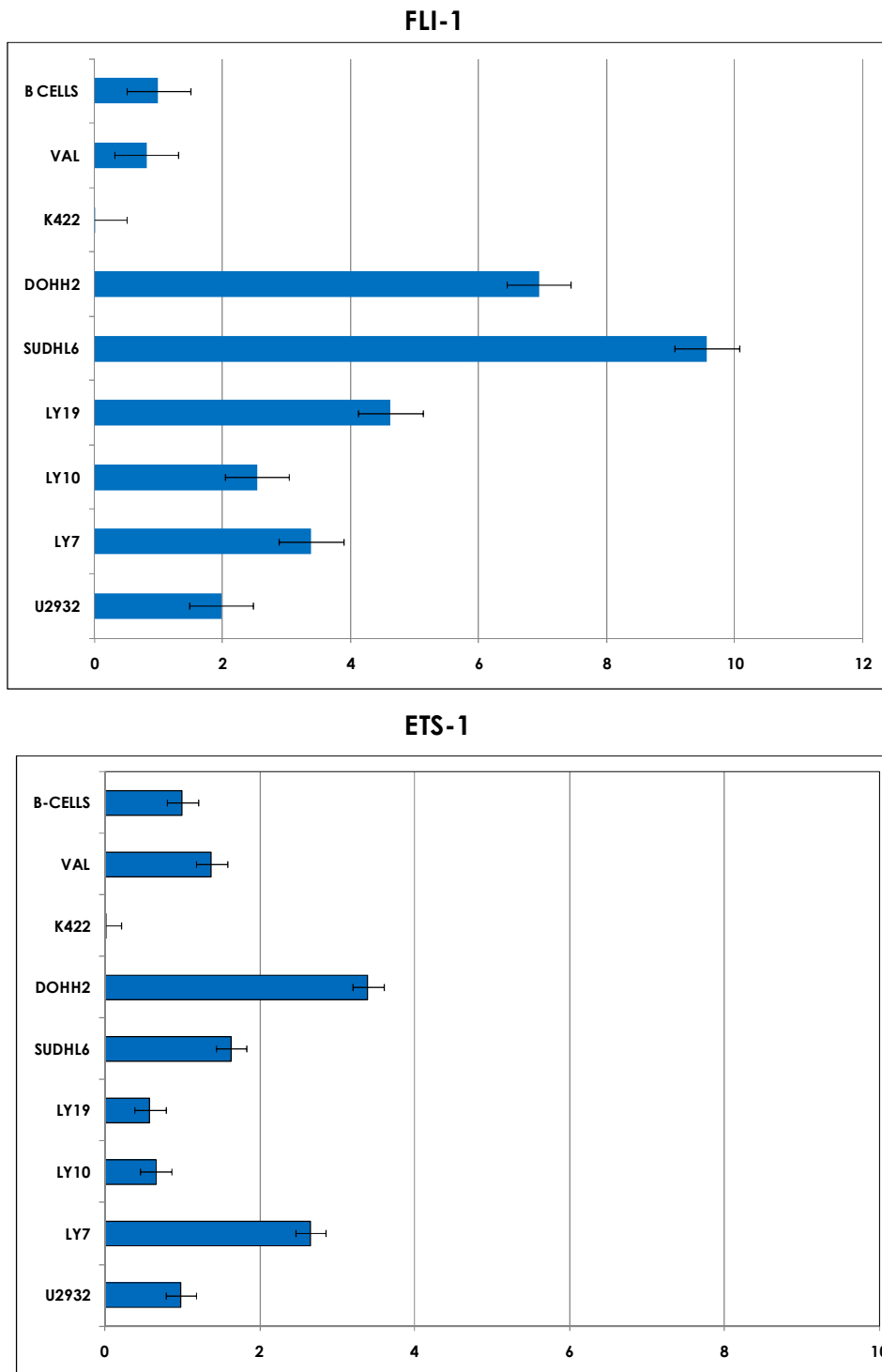


Figure 5.3. Validation of the expression of ETS1 and FLI1 in a panel of DLBCL cell line by RealTime-PCR. Upper panel FLI1, lower panel ETS1. On the x-axis the fold induction of mRNA expression quantified with relative quantification method ($\Delta\Delta^{ct}$), normalized to endogenous control (beta-2microglobulin) and compared to the calibrator (B-cells); on the y-axis are listed DLBCL samples analyzed.

Furthermore, we performed WB analysis on the same panel of DLBCL cell lines and we could confirm the expression at protein level of both genes in all samples, except again in KARPAS422. We could interestingly note that *ETS1* protein is more highly expressed in OCI-LY7, compared to *FLI1* in the same cell line (figure 5.4).

OCI-LY7 carried the 11q24.3 gain, found to be 30% recurrent in the DLBCL subset of patients analyzed in the present work and expressed both *FLI1* and *ETS1* at high level. Thus, we chose OCI-LY7 as a model to perform next loss of function studies for both candidate genes, since also none of the other genes mapping in this region were found to be expressed also in this cell line.

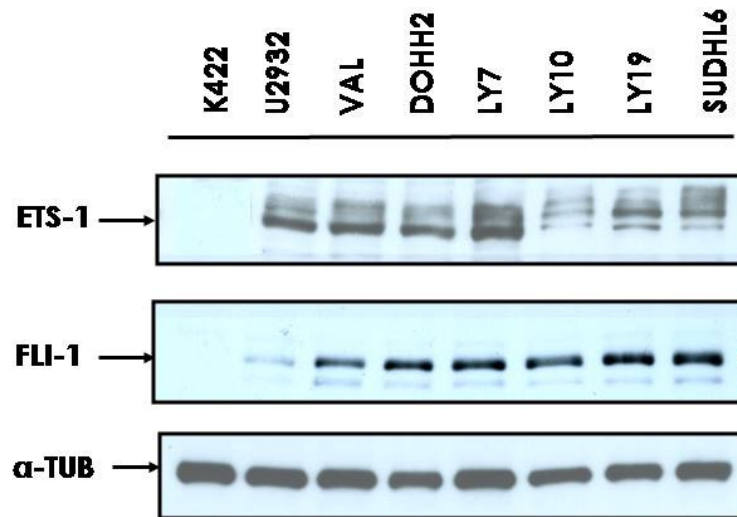


Figure 5.4: Western Blot analysis shows the protein level of *ETS1* and *FLI1* in a panel of DLBCL cell line. Tubulin level was measured as loading control.

5.3.3 VALIDATION OF *ETS1* AND *FLI1* EXPRESSION IN DLBCL SAMPLES

We performed Real-Time-PCR analysis in a subset of 17 DLBCL samples, already analyzed with the 250k Gene-Mapping array, in order to evaluate the mRNA expression of *ETS1* and *FLI1* in DLBCL patients possibly bearing the aberration on 11q. As shown in figure 5.5, 3/17 cases (07-205, 07-212, 07-215), carrying the gain on 11q24.3, seems to highly express both candidate genes (indicated by the arrow).

5.3.4 LOSS OF FUNCTION STUDIES

Down-regulation experiments were performed to silence the genes *ETS1* and *FLI1* by shRNA strategy (using lentivirus SN), in order to investigate the possible functional consequences of 11q24.3 gain.

Cells were infected with lentiviral SN for 6 hours (estimated time for the transduction) and the cationic polymer polybrene (hexadimethrine bromide) was added to the infection mixture, at a concentration of 6 µg/ml. This molecule neutralizes the charge repulsion between viruses and sialic acid on the cell surface allowing the interaction between lentivirus and the target cells. We also tested different concentration rate of the polybrene and we chose 6 µg/ml, because at this dilution no toxic effect was displayed by the cells. OCI-LY7 needs to be washed after the estimated time of 6 hours, because a prolonged time of contact with the virus could be lethal. Once the cells internalize the genetic material released by the virus, not only they become able to actively produce the shRNA against the gene of interest, but they also acquire the resistance to the puromycin. In fact, vectors carrying the shRNA were designed in order to put the shRNA and an antibiotic selection marker for puromycin under the same promoter.

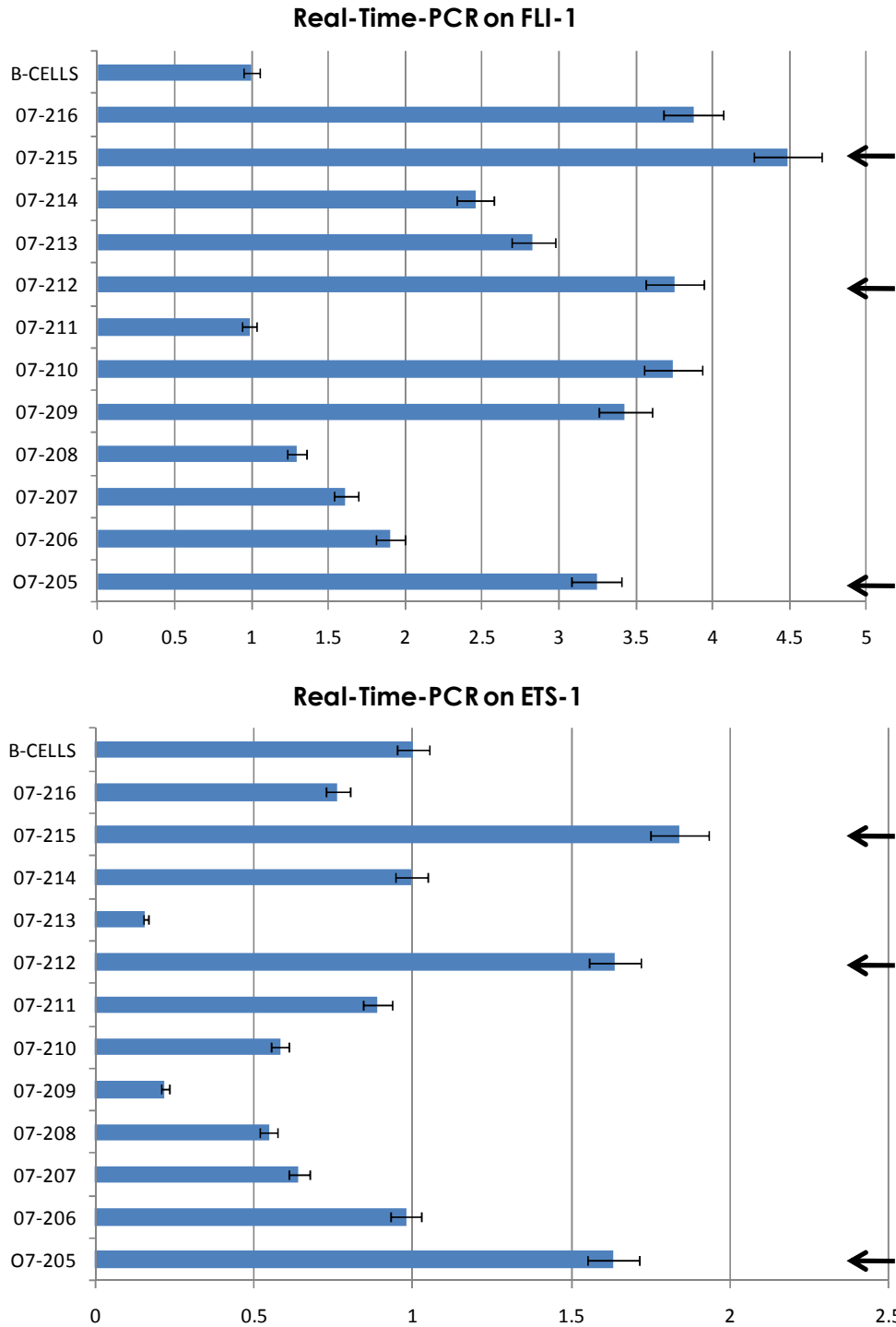


Figure 5.5. Validation of the expression of ETS1 and FLI1 in a subset of 17 DLBCL samples RealTime-PCR. Upper panel FLI1, lower panel ETS1. On the x-axis the fold induction of mRNA expression quantified with relative quantification method ($\Delta\Delta^{ct}$), normalized to endogenous control (β -2microglobulin) and compared to the calibrator (B-cells); on the y-axis are listed DLBCL samples analyzed. Arrows indicate patients carrying the 11q24.3 gain.

Thus, correctly transduced cells could be easily selected by the addition of the right dose of puromycin to the cells in culture, 24h after the infection. Also in this case we tested different concentration of the puromycin and we chose a selection dose of 20 µg/ml, because at this concentration wild type cells were all dead within 36 hours and only the infected cells were able to survive.

We tested the transduction efficiency of OCI-LY7 using a different type of vector carrying a shRNA against *ERK-2* and GFP labeled, instead of carrying Puromycin as a selective marker. The results of cellular infection using a GFP-labeled shRNA in the OCI-LY7 cell line accessed by FACS measurement revealed a low transduction efficiency for this cell line, with only 20% of positive cells. Despite the obvious low transduction efficiency of the OCI-LY7, we decided to proceed anyway with silencing experiments, because OCI-LY7 represents a good model as they carry the aberration of interest in the present work and efficiency in lymphoma cells is expected to be relatively low.

5.3.3.1 *ETS1* DOWNREGULATION LEADS TO *BLIMP1* REACTIVATION AND TO *PAX5* DOWNREGULATION.

To evaluate the role of *ETS1* in OCI-LY7, , stable knock down was performed in the cell line carrying the 11q24.3 gain, using three different shRNAs targeting *ETS1* (60A, 60C, 60D). Cells were selected for 10-12 days using 20 µg/ml of puromycin. As the rate of surviving cells after 10-12 days of antibiotic selection was very low, we proceeded by making 4-5 replicates of infection for each shRNA during every experiments.

The results of lentivirus transduction in OCI-LY7 using shRNAs abrogating the expression of the gene *ETS1* are shown in figure 5.6. A strong reduction both at mRNA and protein level was appreciated with all three shRNAs used, 60A, 60C, 60D, compared to the scramble control (figure 5.7).

In 2008, Jhon S.A. et al ¹²⁶, reported that *ETS1* is involved in B-cells differentiation pathway and especially that it must be down-regulated during late phases of PCs development. They showed that *ETS1* might interfere with the activity of

BLIMP1, the master regulator of PCs differentiation. Starting from these evidences we hypothesized that the up-regulation of *ETS1*, due to the 11q24.3 gain, could be responsible for *BLIMP1* down-regulation and for the maintenance of B-cells into a GC status, blocking their further differentiation into PCs. To test this hypothesis we evaluate *BLIMP1* level after *ETS1* down-regulation in OCI-LY7 transduced with the three different shRNAs (60A, 60C, 60D) and as shown in figure 5.7, we could detect a re-activation of *BLIMP1* at the protein level.

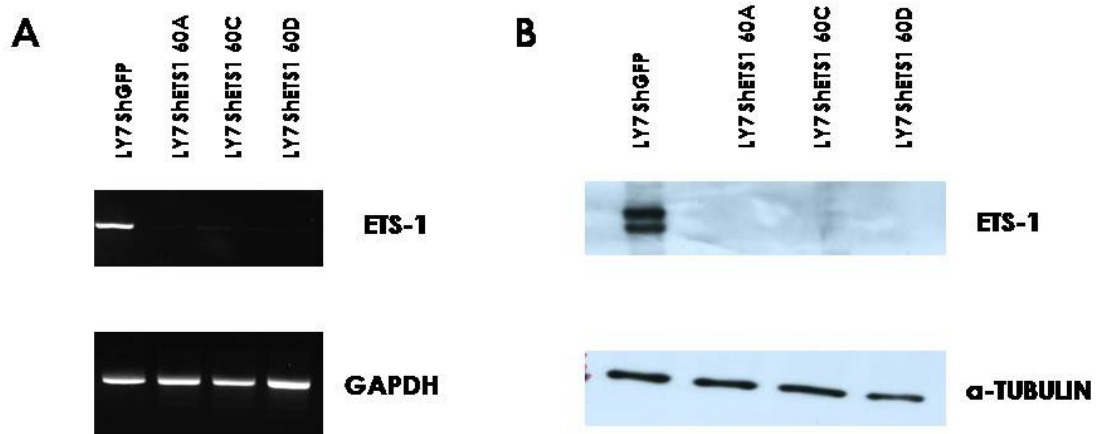


Figure 5.6: RT-PCR (A) and WB (B) analysis measuring the down regulation of *ETS1* in LY7 treated with three different shRNAs. Cells were infected with lentiviral SN containing three different shRNAs targeting *ETS1* and one shRNA against GFP (used as scramble control). Panel A shows RT-PCR analysis, where a strong downregulation of *ETS1*, was detected compared to scramble control. GAPDH level was measured as normalizing control. Panel B shows a strong reduction at protein level of *ETS1*, compared to scramble control. Tubulin level was measured as loading control.

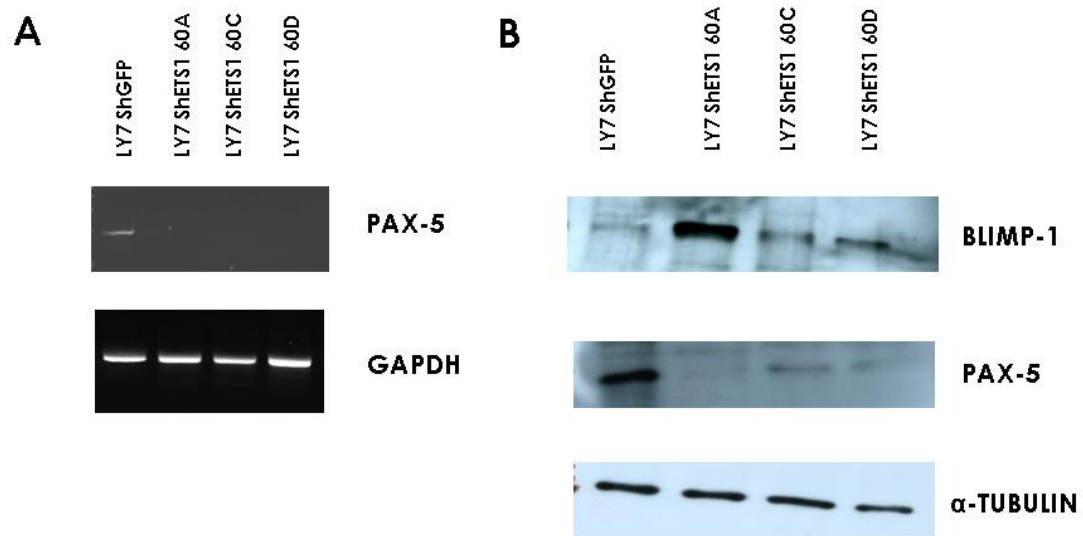


Figure 5.7: RT-PCR (A) and WB (B) analysis measuring the down regulation of PAX5 and the reactivation of BLIMP1 in LY7 treated with three different shRNAs targeting ETS1. Cells were infected with lentiviral SN containing three different shRNAs targeting ETS1 and one shRNA against GFP (used as scramble control). Panel A shows RT-PCR analysis, where a strong downregulation of PAX5, was detected compared to scramble control. GAPDH level was measured as normalizing control. Panel B shows an increase of BLIMP1 at protein level and the downregulation of PAX5 protein after ETS1 knock-down, compared to scramble control. Tubulin level was measured as loading control.

To confirm the real re-activation of *BLIMP1* we investigated the possible down-regulation of one of its target genes involved in the same pathway, *PAX5*. The latter is a known transcription factor expressed during early phases of B-cells development, but it must be down-regulated by *BLIMP1* during PCs differentiation, because it is implied in the identity of memory B-cells ⁴⁷. Figure 5.7 shows also that together with *BLIMP1* reactivation (panel B), we could observe *PAX5* down-regulation both at mRNA and protein (panel A and B) level, indirectly confirming our initial hypothesis.

Taken together, these observations allowed us to propose an alternative mechanism occurring in DLBCL. In fact the up-regulation of *ETS1* due to the

presence of the 11q24.3 gain, might compromise *BLIMP1* activity and might cause a block of B-cells differentiation into PCs.

5.3.3.2 *ETS1* DOWNREGULATION AFFECTS OCI-LY7 CELL PROLIFERATION ACTIVITY.

To test the effect of *ETS1* down-regulation on cells proliferation, we performed Carboxyfluorescein (CFSE) assay. It normally refers to a fluorescent dye, with an excitation and emission of 492 nm. CFSE is incorporated into the cells and it is useful to monitor the rate of cells growth, because the lower is the fluorescence emitted the higher is the dilution of the CFSE, due to the capability of the cells to proliferate.

OCI-LY7 transduced with shRNA (60A, 60C, 60D) against *ETS1* and with the scramble control, were maintained under puromycin selection for 10 days and then labeled with CFSE. To evaluate the impact of *ETS1* down-regulation on cells growth, FACS analysis was performed after 3, 6 and 9 days.

After 3 days from the labeling, the proliferation rate of OCI-LY7 transduced with lentiviruses containing the shRNAs targeting *ETS1* (60A, 60C, 60D) was decreased. In fact, the absence of a second peak (meaning a cell population normally proliferating) on the left panel for all the three clones, means that OCI-LY7 were alive, but not under active proliferation and so CFSE could not be diluted. The observed block in cell division at day 3, is maintained until day 6 (Fig. 5.8). OCI-LY7 treated with the scramble shRNA proliferate normally both at day 3 and 6, as shown in figure 5.8. FACS analysis was performed also at day 9 and even at this time the inhibition of the proliferation was maintained in the samples stably knocking-down *ETS1*, although a small number of cells survived after 9 days from the CFSE labeling. The last observation could mean that the abrogation of *ETS1* expression might have an impact not only on cells proliferation, but also on their survival.

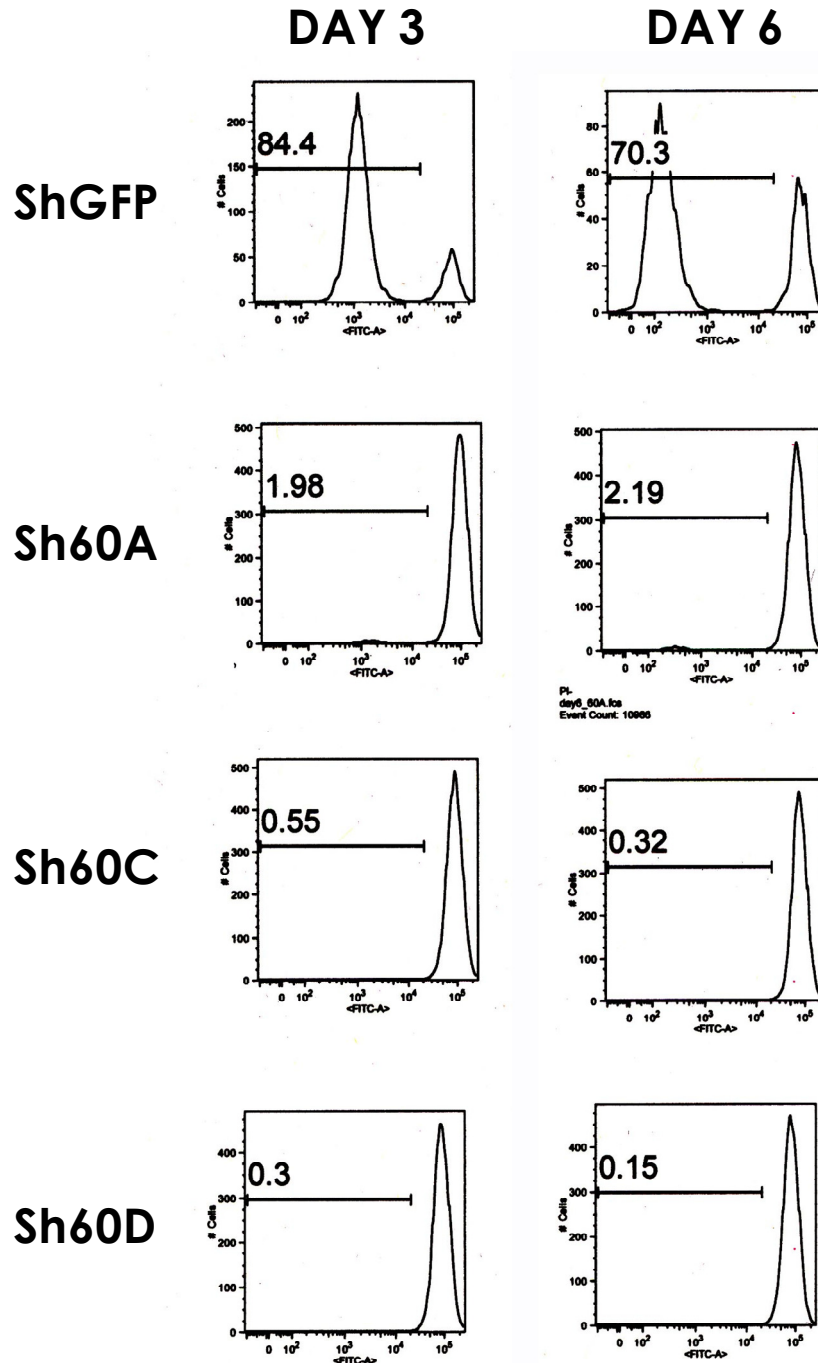


Figure 5.8. CFSE labelling of LY7 cell line transduced with shRNAs targeting *ETS1* analyzed by FACS. OCI-LY7 cells were treated with CFSE after lentivirus transduction. FACS acquisition was performed after 3 and 6 days. X axis represents concentration of CFSE; Y axis represents cells count. GFP is the scramble control profile; sh60A, sh60C and sh60D represent proliferation profile of LY7 cells treated with three different shRNAs against *ETS1*.

5.3.3.3 *FLI1* IS A CANDIDATE ACTIVATOR OF *PAX5*

To evaluate the role of *FLI1* in the OCI-LY7, we performed a stable knock down of this gene using two different shRNAs (61C, 61E). After the infection, cells were selected for 10-12 days using 20 µg/ml of puromycin. As previously discussed, the rate of surviving cells after antibiotic selection was very low and so we proceeded by making 4-5 replicates of infection for each shRNA during every experiments.

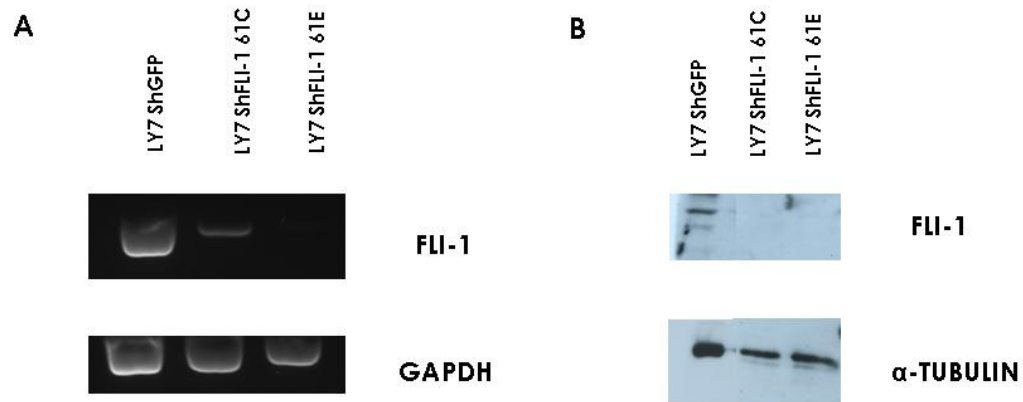


Figure 5.9. PCR and WB analysis of samples extracted from OCI-LY7 treated with shRNA against *FLI1*. Cells were transduced with lentiviral SN containing two different shRNAs targeting *FLI1* and one shRNA against GFP (used as scramble control). Panel A shows RT-PCR analysis, where a strong downregulation of *FLI1*, was detected compared to scramble control. GAPDH level was measured as normalizing control. Panel B shows a reduction of *FLI-1* protein level in the same samples, compared to scramble control, by WB. Tubulin level was measured as loading control.

Figure 5.9 shows that we obtained the down-regulation of *FLI1* at mRNA (panel A) and protein level (panel B). To further confirm these results, we also performed an intracellular staining (IS) in the OCI-LY7 transduced with the two shRNAs against *FLI1*, after maintaining the cells for 10 days under antibiotic selection. Using this technique, cells are first fixed and permeabilized in order to introduce inside the cells an antibody recognizing our target gene. The

secondary antibody, which is in our case biotinylated, is then internalized and after the addition of the streptavidine solution, the amount of the protein of interest inside the cells can be measured by FACS analysis, thanks to the fluorescent reaction generated by the biotin-streptavidine system.

As shown in figure 5.10 we could detect a decrease of the protein level of *FLI1*. With this experiment, we have been able also to show that when *FLI1* is down-regulated, *ETS1* is still expressed in the same clone. The last observation is important, because in literature it is known that *FLI1* together with *PAX5*, *ETS1* and *GABPa*, is co-recruited at *mb-1* promoter binding site during B-cell development in order to activate it ¹²⁷. The ability of *PAX5* to promote B cell development is due to its role in activating the expression of B lineage specific genes, including *mb-1* which encodes the Ig-a protein. Moreover, as previously discussed, *PAX5* must be down-regulated during PCs differentiation in particular, after the GC-development stage and *BLIMP1* is responsible for its inactivation.

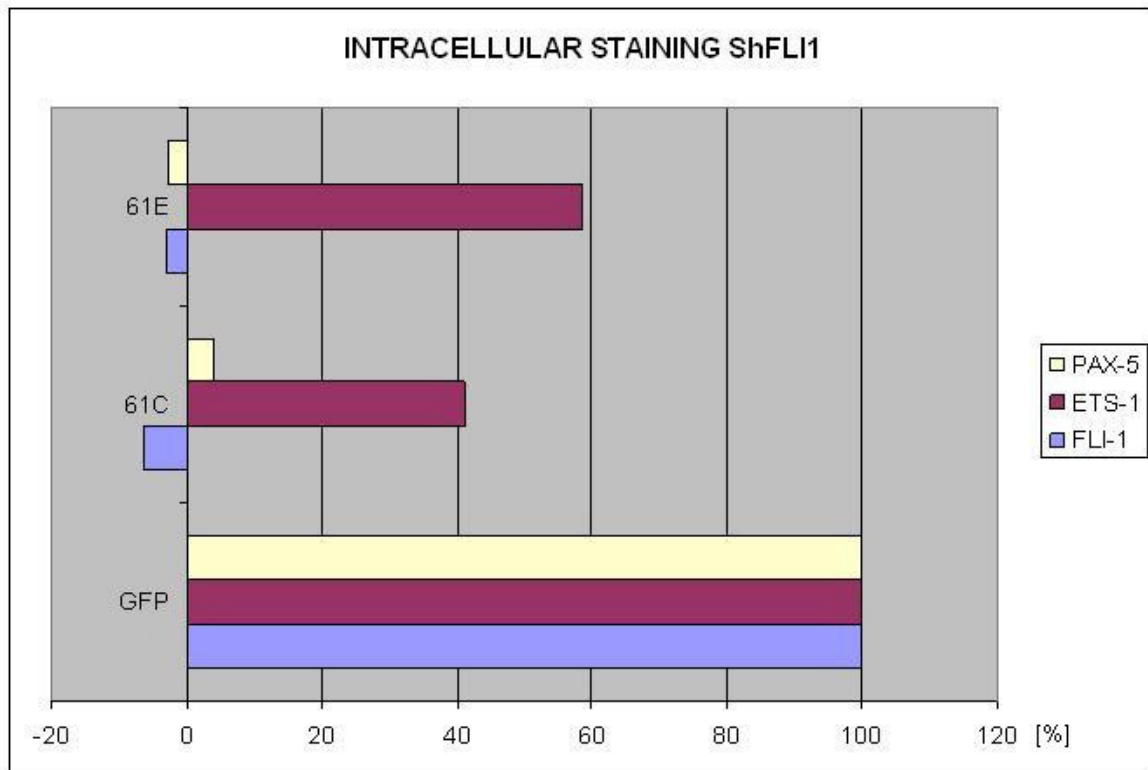


Figure 5.10. Intracellular Staining (IS) showing a reduction of FLI1 protein level in OCI-LY7 samples treated with shRNAs against FLI1 (61C, 61E) compared to scramble control (GFP). The graph shows also the maintenance of ETS1 protein level and the down-regulation of PAX5 protein. Single experiment representative for the same transduction experiment performed for figure 5.9 and 5.11.

This overview allowed us to make another hypothesis about *FLI1*: because we could observe only a minor effect on the regulation of *ETS1* operated by *FLI1* in our system, we wanted to investigate if *FLI1* might have an impact on factors, such as *PAX5*, already involved in the alternative mechanism that we proposed in section 5.4.3.

Up to now, as shown in figure 5.11, we have been able to demonstrate by RT-PCR (Panel A) and WB (Panel B), that the abrogation of the *FLI1* transcript by shRNAs (61C, 61E) leads to the down-regulation of *PAX5* both at mRNA and protein level.

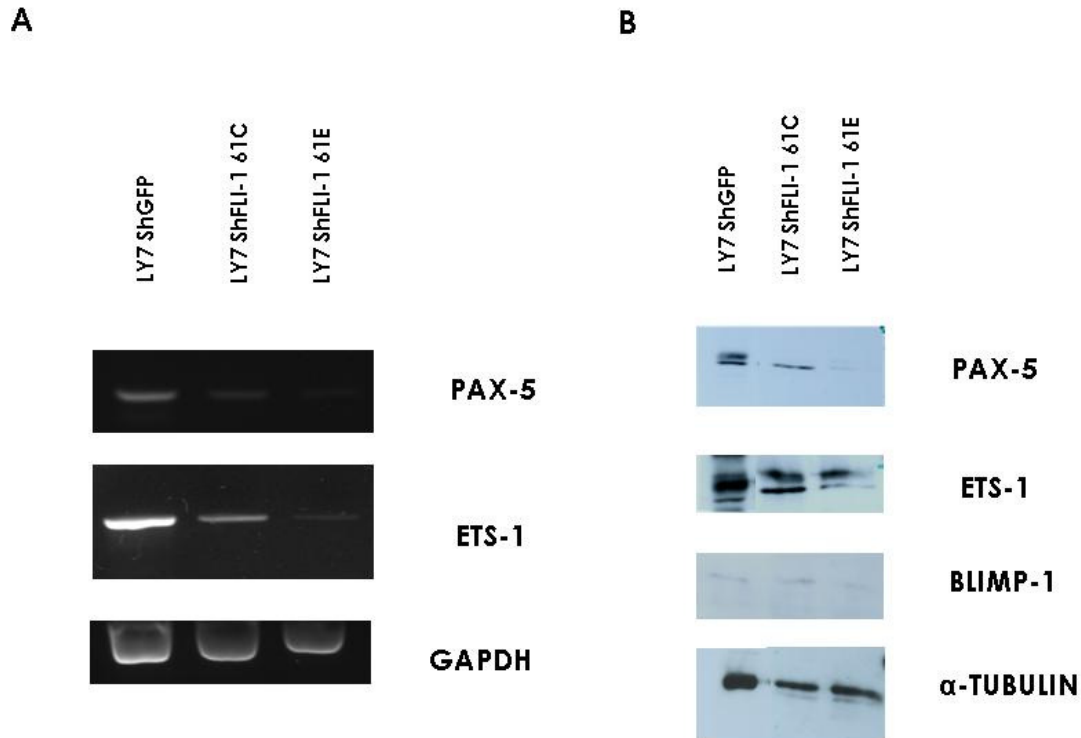


Figure 5.11. PCR and WB analysis of samples extracted from OCI-LY7 treated with shRNA against *FLI1*. Cells were infected with lentiviral SN containing two different shRNAs targeting *FLI1* and one shRNA against *GFP* (used as scramble control). Panel A shows RT-PCR analysis, where a downregulation of *PAX5* was detected following down-regulation of *FLI1* (already shown in figure 5.10) compared to scramble control. *ETS1* level seems to be marginally affected by *FLI1* downregulation. *GAPDH* level was measured as normalizing control. Panel B shows a reduction of *PAX5* protein level in the same samples by WB, no reactivation of *BLIMP1* and the maintenance of *ETS1* expression, after *FLI1* downregulation (all compared to scramble control). *Tubulin* level was measured as loading control.

Because IS experiments (figure 5.10), RT-PCR and WB analysis (Figure 5.11) showed that *ETS1* is not completely affected after *FLI1* silencing, we supposed that *BLIMP1* was not re-activated and that *FLI1* could have been a good candidate as an activator of *PAX5*. The inactivation of the latter, might be mainly caused by *FLI1* knock-down. As shown in figure 5.11 also *BLIMP1* is not reactivated after *FLI1* abrogation, thus reinforcing our hypothesis regarding the second ETS factor.

5.3.3.4 *FLI1* DOWNREGULATION AFFECTS OCI-LY7 CELL PROLIFERATION ACTIVITY

Similarly to what was done for *ETS1*, to test the effect of *FLI1* down-regulation on cells proliferation, we performed CFSE assay..

OCI-LY7 transduced with shRNA (61C, 61E) against *FLI1* and with the scramble control, were maintained under puromycin selection for 10 days and then labeled with CFSE. To evaluate the impact of *FLI1* down-regulation on cells growth, FACS analysis was performed after 3, 6 and 9 days.

FLI1 is known to be an important factor for the maintenance of cell proliferation¹²⁸. After 3 days from the labeling, the proliferation rate of OCI-LY7 infected with lentiviruses containing the shRNAs targeting *FLI1* (61C, 61E) was lower. Once again, the absence of a second peak (meaning a cell population normally proliferating) on the left panel for both clones, means that OCI-LY7 transduced were alive, but not under active proliferation and so CFSE could not be diluted. The observed block in cells division at day 3, is still present at day 6 (Fig. 5.12). OCI-LY7 treated with the scramble shRNA proliferated normally both at day 3 and 6, as shown in figure 3.13. FACS analysis was performed also at day 9 and even at this time the inhibition of the proliferation was maintained in the samples stably knocking-down *FLI1*, although a small number of cells survived after 9 days from the CFSE labeling. The last observation suggests that also the abrogation of *FLI1* expression has an important impact not only on cell proliferation as expected, but also on their survival.

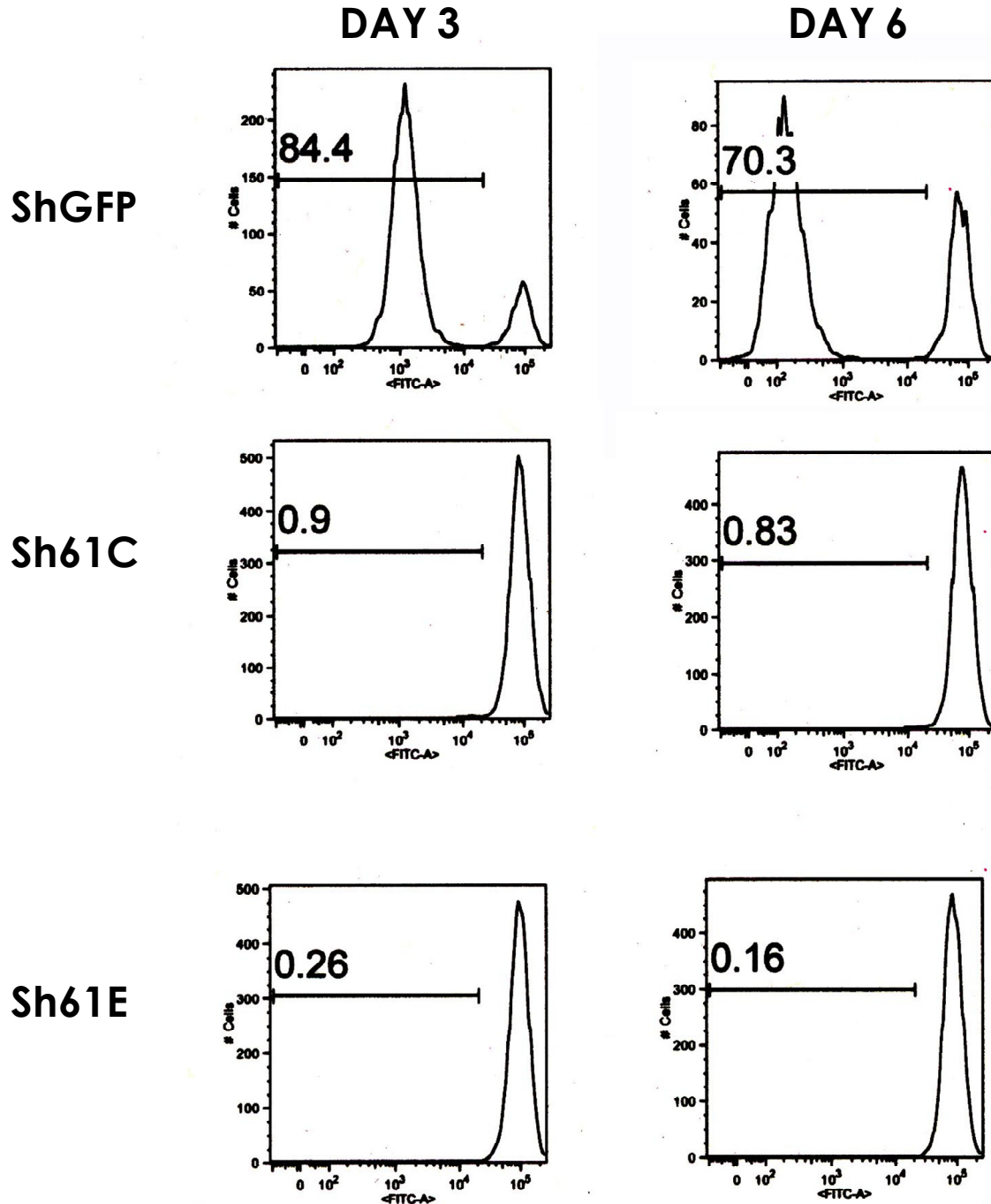
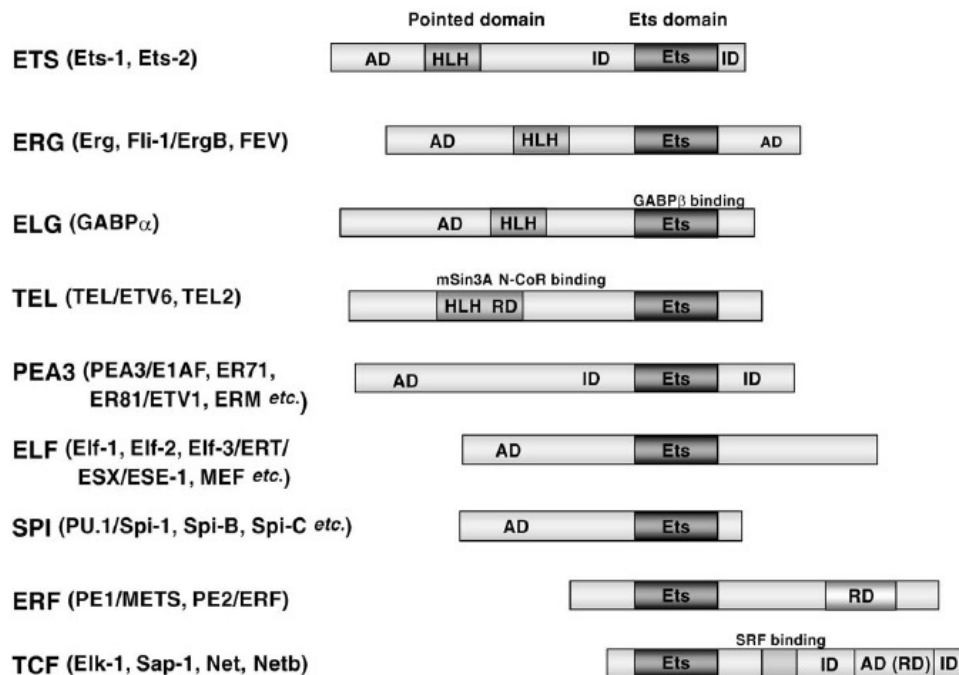


Figure 5.12. CFSE labeling in OCI-LY7 cell line infected with shRNA targeting FLI1 analyzed by FACS. OCI-LY7 cells were treated with CFSE after lentivirus transduction. FACS acquisition was performed after 3 and 6 days. X axis represents CFSE concentration; Y axis represents cell count. GFP is the scramble control profile; sh61C and sh61E represent proliferation profile of OCI-LY7 cells treated with two different ShRNAs against FLI1.

5.4 DISCUSSION

In the first part of our project, applying aCGH to study 166 DLBCL samples with the purpose of identifying relevant CN changes linked to malignancy, we detected a recurrent gain at 11q24.3 in 30% of cases. Due to the hypothesis that highly increased DNA CN might determine the over expression of genes possibly playing a role in cancer development and progression, we focused our interest on the genes mapped in 11q24.3 region. Among the seven genes mapping in the 11q24.3 gain, only *ETS1* and *FLI1* were expressed with a high rate in DLBCL patients bearing the lesion, both at RNA and protein level. These two candidate genes are members of the ETS family of transcription factors, which is mainly characterized by an evolutionarily-conserved DNA-binding domain that regulates expression of a variety of viral and cellular genes by binding to a purine-rich GGAA/T core sequence in cooperation with other transcriptional factors and cofactors (Figure 5.13). Most ETS family proteins are nuclear targets for activation of Ras-MAP kinase signaling pathway and some of them affect proliferation of cells by regulating the immediate early response genes and other growth-related genes ¹⁰⁴. Several ETS family proteins are preferentially expressed in specific cell lineages and are involved in their development and differentiation by increasing the enhancer or promoter activities of the genes encoding growth factor receptors and integrin families specific for the cell lineages. Many ETS family proteins also modulate gene expression through protein-protein interactions with other cellular partners. Deregulated expression or formation of chimeric fusion proteins of ETS family due to proviral insertion or chromosome translocation is associated with leukemias and specific types of solid tumors. Several ETS family proteins also participate in malignancy of tumor cells including invasion and metastasis by activating the transcription of several protease genes and angiogenesis-related genes ¹²⁹.

Subfamily (member)**Figure 5.13. Overview of ETS family members (Oikawa T, 2002¹²⁹)**

A role for ETS factors in oncogenesis has been appreciated and as presented in Chapter 1, it was recently observed that the *SPI-B* can be constitutively activated in ABC-DLBCL⁴. Enforced expression of *SPI-B* in human B cells directly represses *PRDM-1* (which encodes *BLIMP1*) and *XBP-1* (encoding X-box binding protein-1) gene expression, which are required for PC formation and Ab production. *SPI-B*-deficient mice do not show increased numbers of PCs or elevated serum Ab titers, which is probably explained by the impaired GC formation also seen in these mice, therefore suggesting additional contributions of *SPI-B* during B-cell differentiation¹³⁰.

ETS1, mapped in the 11q24.3 region, was recently described in murine B-cell differentiation. Evidence also suggests that the transcription factor *ETS1* may play a key role in regulating *BLIMP1* activity, since mice lacking *ETS1* exhibit a strikingly increased number of IgM-secreting plasma cells in their lymphoid organs^{126, 131}. Moreover, B cells lacking *ETS1* undergo enhanced differentiation into IgM-secreting PCs when cultured in the presence of a synthetic TLR9 ligand

(CpG oligodeoxynucleotide (ODN))¹³¹. The mechanism by which *ETS1* regulates the development of IgM-secreting plasma cells is unknown but likely involves modulating the expression or activity of genes that govern the PC fate. Interestingly, several ETS proteins can direct lineage-commitment decisions via their ability to form protein-protein complexes with key lineage-specific transcription factors, leading to inactivation of the function of those transcription factors. In line with the observation that *ETS1* and *SPI-B* are differentially expressed in naive and GC B cells, it is tempting to speculate that *SPI-B* prevents premature differentiation of activated GC B cells, whereas *ETS1* contributes to the prevention of spontaneous differentiation and Ab secretion by naïve B cells¹²⁵.

As already mentioned, ETS proteins regulate the expression of genes that are critical for the control of cellular proliferation, differentiation, and programmed cell death and previous studies indicate that *FLI1*, the other gene mapping in 11q24.3, plays an important role in the regulation of megakaryocyte development¹³². Studies regarding the role of *FLI1* in immune function or dysfunction are limited. *FLI1* can bind specific ETS-binding sites and transcriptionally activate a number of genes including those encoding Egr-1, multiple megakaryocytic genes, *bcl-2*, *mb-1*, and hTERT^{127, 133, 134}. In fact, together with *FLI1*, *ETS1* and *PAX5* are co-recruited at mouse *mb-1* promoter, as transcription activators¹²⁷. The latter encodes the mouse-homologous of the essential transmembrane signaling protein CD79A. *FLI1* appears to be also a negative regulator of Marginal Zone (MZ) B cell development and a positive regulator of follicular B cell development, possibly via regulating surface expression of Ig-A. In addition, altered expression of E2A and Id mRNAs may also contribute to the observed increase in MZ B to follicular B cell ratio. These findings highlight a novel *in vivo* function for *FLI1* in the regulation of B cell development and function¹³², which must be further investigated.

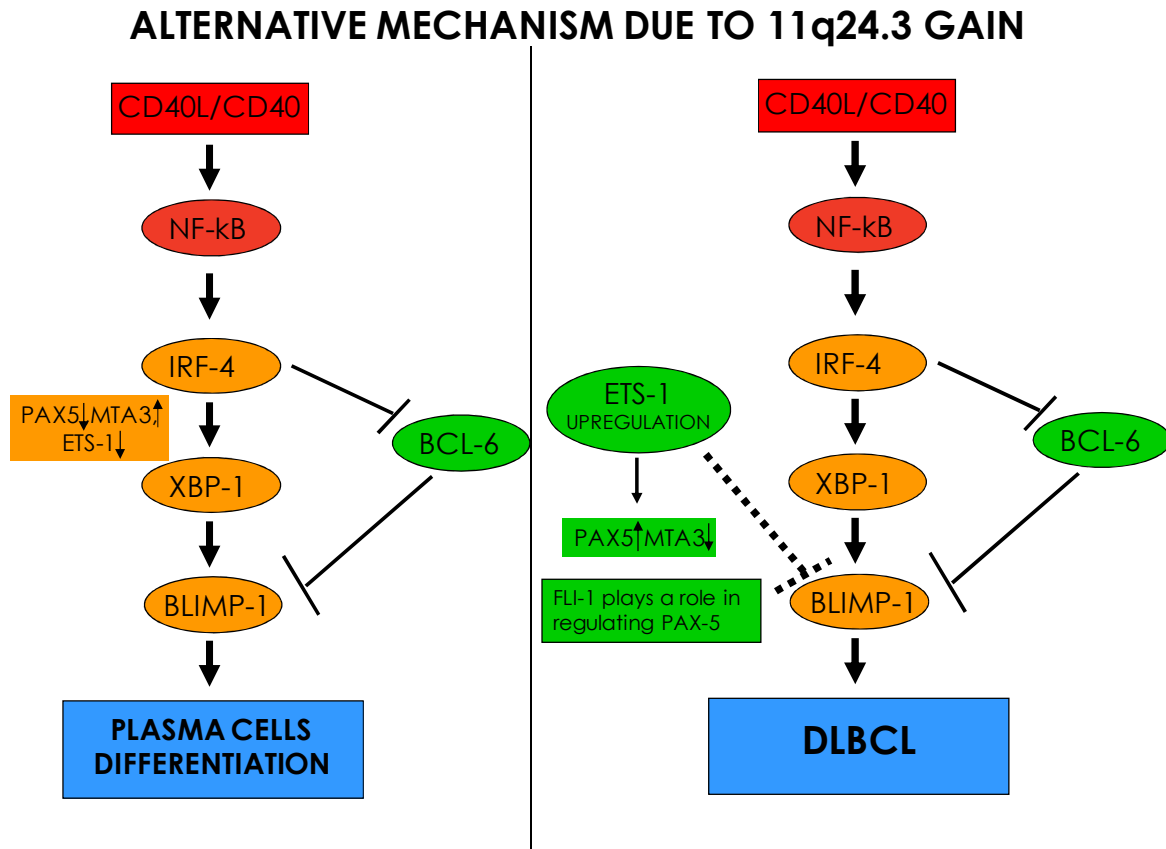


Figure 5.14. B-cell differentiation pathway in normal conditions and during DLBCL pathogenesis.

Normally (left panel), after CD40-CD40L stimulation, NF-κB pathway is activated. This results in the transcriptional silencing of BCL6 through IRF-4 activation and in the activation of XBP-1 and BLIMP1. Together with the down-regulation of BCL6 also PAX-5 and ETS-1 must be down-regulated. All these mechanisms are likely to be crucial for the transition of GC B cells from centroblast to centrocytes committed to PC differentiation. During DLBCL pathogenesis (right panel), BLIMP1 activity could be blocked by the up-regulation of ETS1. This alteration might be responsible for the blocking of the entrance of B-cells into PC differentiation. During this alternative pathway also FLI1 up-regulation could play a role in regulating PAX-5, contributing to the impairment of PCs differentiation.

Since we detected an over expression of both genes in patients, first we tried to find a cell model to evaluate their function in DLBCL. Genomic profiling of a subset of DLBCL cell lines, allowed us to identify OCI-LY7 as a candidate model. This cell carried the same 11q24.3 gain. Moreover real-Time-PCR and WB analysis confirmed the expression of both ETS factors in OCI-LY7.

Our hypothesis started from the evidence that John SA et al in 2008 ¹²⁶ demonstrated a putative role for *ETS1* in interfering with *BLIMP1* activity. In fact *ETS1* must be down-regulated during the late phases of PCs differentiation, exactly when *BLIMP1* is acting in order to allow B-cells development ¹²⁶. So we wondered whether the up-regulation of *ETS1* due to the presence of 11q24.3 gain might be responsible for the down-regulation of *BLIMP1* as an alternative or additional mechanism leading to the block of the PCs development (figure 5.14). Our data have been able to validate this hypothesis, as we observed a re-activation of *BLIMP1* protein, after down-regulation of *ETS1* using three different ShRNAs targeting this factor (60A, 60C, 60D) (figure 5.6). An indirect support that the proposed mechanism was not an artefact, was the observed down-regulation of *PAX5*, which, as previously discussed, is a known target of *BLIMP1*. Taken together these data indicate a possible involvement of *ETS1* in blocking PCs development in a GC status, by abrogating the expression of *BLIMP1*. Thus, the mechanism that we tried to elucidate might be really active as alternative way, leading to DLBCL (Figure 5.14), with *ETS1* expression still maintained during the late phases of such differentiation pathway, especially in cases bearing the 11q24.3 gain.

Data regarding *FLI1* deserve a better clarification, although there are good evidences that *FLI1* might be a new activator candidate of *PAX5*. In fact, independently of *ETS1* expression and *BLIMP1* reactivation, *PAX5* is down-regulated in OCI-LY7 both at mRNA and protein level, also when *FLI1* is inactivated by shRNA strategy, with no or minimal effect on *ETS1* gene (figure 5.15). We do not know yet whether *FLI1* acts directly or indirectly on *PAX5* promoter, but using different bioinformatic tools, such as PROMOSER, GENOMATICS, TFSEARCH and ALGGEN-PROMO, we found a series of ETS Binding Sites (EBs) into *PAX5* promoter, which might be recognized by *FLI1*. This could reinforce the observation that *FLI1* is a direct activator of *PAX5* expression.

Moreover after *ETS1* or *FLI1* down-regulation we could notice a block in the cell proliferation, as showed by FACS analysis of the CFSE assay performed in OCI-LY7 stably infected with shRNAs targeting both factors, compared to the normal

control. This evidence suggests that the presence of the two *ETS* factors give a prominent advantage to OCI-LY7 cell in their capability of proliferating. A characterization of the antigens expressed by the OCI-LY7 before and after the abrogation of *ETS1*, could clarify if the cells are slowly changing phenotype and differentiating in to PCs (as we hypothesized) or if their destiny is to undergo cell death. On the contrary, *FLI1* is known to have an important role in cell proliferation, so the observation that knock-down of this gene could lead to a block of cell growth, is more consistent with previous published data ¹³².

In conclusion, the aim of this study, the identification of recurrent lesions and the characterization of the function of the affected genes, was reached. Indeed, we could demonstrate that both genes, mapping in the 11q24.3 gained region, might play an important role in the mechanism of B-cells differentiation into PCs. *ETS1* is important during the first phases of B-cells development ¹²⁵ and generates a possible arrest of the cells in a GC status if its expression is maintained until the late stages of the pathway, mainly because it might cause the inactivation of *BLIMP1* activity. *FLI1* is probably important at the same time for the activation of *PAX5* during early phases of the same pathway. A prolonged expression of the latter (*PAX5*) would impair PCs commitment and would probably allow B-cells to differentiate in memory B-cells ⁴⁷, so it could be of great interest to understand the role of *FLI1* even in memory B-cells commitment. Moreover, both genes have also a direct effect on cell growth and survival.

Although further experiments are needed in order to understand if all these interactions take place at transcriptional level or at protein level, the present work set solid bases and demonstrate the great usefulness of the combination of genomic approaches and functional studies in cancer.

FLI1 DOWNREGULATION, LEADS TO PAX5 DOWNREGULATION INDEPENDENTLY OF ETS1 AND BLIMP1 ACTIVITY

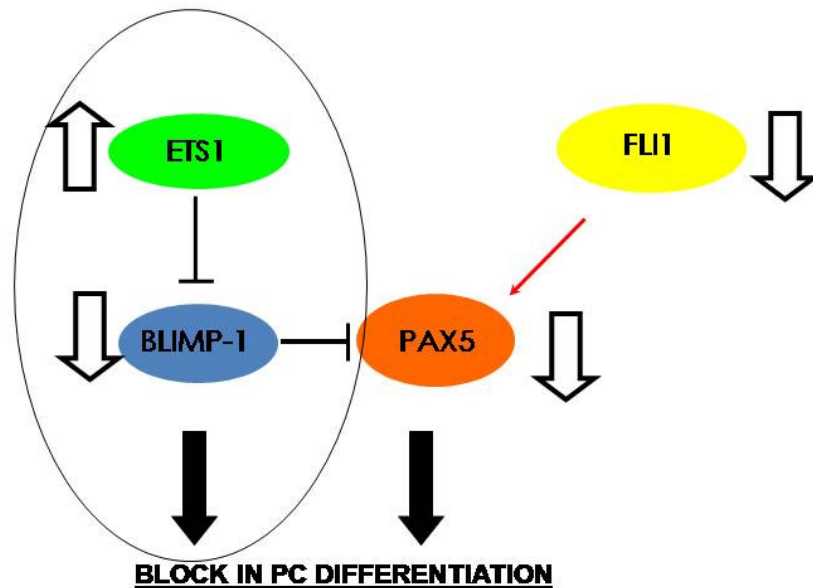


Figure 5.15. Possible role of FLI1 in DLBCL pathogenesis. ETS1 upregulation causes BLIMP1 down-regulation that should justify the presence of PAX-5 in OCI-LY7. During FLI1 knock down in OCI-LY7 PAX-5 expression was abrogated. Thus, if actually ETS1 is maintained up-regulated and BLIMP1 is not consequently reactivated, this could be a first confirmation that FLI1 acts directly on PAX5 promoter.

6. CONCLUSION

The main purpose of this study was to demonstrate the importance of the genomic profiling in combination with clinical analysis and functional studies.

Since DLBCL is a very heterogeneous type of lymphoma, many arrayCGH-studies aimed at elucidating its biological features, which are often associated with distinct clinical presentations and courses ^{81, 106, 114, 135}.

In Chapters 2 and 3, we have been able to show how such genomic approach is useful in order to characterize recurrent aberrations occurring in DLBCL samples. Defining MCRs as previously reported ⁴, we identified a variety of lesions, first in order to confirm the heterogeneity of this disease and then to understand their influence on the clinical course of our cohort of patients. Losses of 8p, a bona fide CNV at 15q11.2, and LOH at 15q21.3 and at 9p (*CDKN2A*) were associated with treatment failure with R-CHOP, but none were of high statistical significance. A significant impact on the outcome was displayed only by 8p23.1 loss, but the role of this lesion must be further investigated, although it has been previously associated with a worse clinical course in other B-cell malignancies ¹¹⁵⁻¹¹⁸. Since this region contains a large variety of genes, due to the size, no clear candidate genes has been identified so far. Moreover Bosh N et al ¹¹⁸ recently reported that the segment included in 8p23.1 can be found in different orientations, suggesting that mutations, such as inversions, might have an important effect on the alteration of genes mapping in this locus ¹¹⁸.

Furthermore, we also aimed to investigate the presence of biological DLBCL subgroups according to the genomic profiles of our subset of patients. Indeed, through the application of NMF algorithm, we have been able to identify five different clusters, some with clear distinct clinical characteristics and distinct courses. Cluster 1 patients were characterized by a heterogeneous genomic profile, by the worst long term outcome among the different clusters and by the high rate of inflammatory cells. Based upon their genomic profiles, cluster 2 was reminiscent of post-GC DLBCL while cluster 4 of the GCB-type. Analysis of their

outcome confirmed such observation, since cluster 2 was characterized by a poorer clinical course compared to cluster 4, resembling what has been already shown for these DLBCL subtypes⁸⁰.

In chapter 5 we presented a functional characterization of the 11q24.3 gain, found to be recurrent in 30% of cases, with the purpose of understanding its biological role in DLBCL. Indeed, we succeeded in identifying two possible candidate genes in this region, *ETS1* and *FLI1*, which are both members of the ETS family transcription factors¹²⁸. We also found a possible DLBCL cell line model, OCI-LY7, bearing the same gain on 11q and identified among our subset of patients. Since there are evidences in literature of the possible involvement of both ETS factors in B-cells differentiation^{126, 127}, we hypothesized an alternative mechanism leading to DLBCL, in which *ETS1* and *FLI1* might play a role (figure 5.14 and 5.15). Thus, we decided to perform loss of function studies using shRNA strategy on OCI-LY7, with the aim of knocking down both genes. After *ETS1* downregulation we could detect a reactivation of *BLIMP1*, the master regulator of PCs development^{124, 136, 137}. This observation was confirmed by the consequent downregulation of *PAX5*, a known target of *BLIMP1*⁴⁷, which must be repressed during the late phases of this pathway¹³⁷. We have also been able to show that *FLI1* is a good candidate as an activator of *PAX5*. After *FLI1* knocking-down we observed a downregulation of *PAX5*, independently of *BLIMP1* and *ETS1* activity, suggesting an important role of this ETS factor in regulating the expression of a gene involved in PCs development. Finally, we demonstrated that the abrogation of both *ETS1* or *FLI1* had an important impact on OCI-LY7 proliferation, leading to a block in their division.

Taken together these experiments showed that the mechanism we tried to elucidate might be a possible alternative way contributing to DLBCL, where *ETS1*, upregulated by the presence of the 11q24.3 gain, could be responsible for the impairment of *BLIMP1* activity, leading to a block of PCs differentiation. At the same time, high level of *FLI1*, due to the same gain, are responsible for the activation of *PAX5* and for the consequent maintenance of the cells in a GC status.

In conclusion all the data described in the present work, reinforce the importance of combining genomic approach with clinical evaluation and functional studies in order to better characterize human cancers.

7. BIBLIOGRAPHY

1. Sehn, L.H. *et al.* Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol* **23**, 5027-5033 (2005).
2. Monti, S. *et al.* Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood* **105**, 1851-1861 (2005).
3. Bayani, J. *et al.* Genomic mechanisms and measurement of structural and numerical instability in cancer cells. *Semin Cancer Biol* **17**, 5-18 (2007).
4. Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A* **105**, 13520-13525 (2008).
5. Griffiths, A.J.F., Gelbart, W.M., Lewontin, R.C. & Miller, J.H. *Modern Genetic Analysis*, Edn. 2nd Edition. (Freeman, W.H., 2002).
6. Feinberg, A.P. Genomic imprinting and gene activation in cancer. *Nat Genet* **4**, 110-113 (1993).
7. Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L. & Jaenisch, R. DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89-93 (1998).
8. Das, P.M. & Singal, R. DNA methylation and cancer. *J Clin Oncol* **22**, 4632-4642 (2004).

9. Lafon-Hughes, L., Di Tomaso, M.V., Mendez-Acuna, L. & Martinez-Lopez, W. Chromatin-remodelling mechanisms in cancer. *Mutat Res* **658**, 191-214 (2008).
10. Frohling, S. & Dohner, H. Chromosomal abnormalities in cancer. *N Engl J Med* **359**, 722-734 (2008).
11. Gajduskova, P. *et al.* Genome position and gene amplification. *Genome Biol* **8**, R120 (2007).
12. Hinds, P.W. & Weinberg, R.A. Tumor suppressor genes. *Curr Opin Genet Dev* **4**, 135-141 (1994).
13. Knudson, A.G. Two genetic hits (more or less) to cancer. *Nat Rev Cancer* **1**, 157-162 (2001).
14. Zhang, L. *et al.* microRNAs exhibit high frequency genomic alterations in human cancer. *Proc Natl Acad Sci U S A* **103**, 9136-9141 (2006).
15. Tuna, M., Knuutila, S. & Mills, G.B. Uniparental disomy in cancer. *Trends Mol Med* **15**, 120-128 (2009).
16. Robinson, W.P. Mechanisms leading to uniparental disomy and their clinical consequences. *Bioessays* **22**, 452-459 (2000).
17. Murthy, S.K., DiFrancesco, L.M., Ogilvie, R.T. & Demetrick, D.J. Loss of heterozygosity associated with uniparental disomy in breast carcinoma. *Mod Pathol* **15**, 1241-1250 (2002).
18. Cowell, J.K. & Hawthorn, L. The application of microarray technology to the analysis of the cancer genome. *Curr Mol Med* **7**, 103-120 (2007).

19. Feuk, L., Carson, A.R. & Scherer, S.W. Structural variation in the human genome. *Nat Rev Genet* **7**, 85-97 (2006).
20. Costa, J.L., Meijer, G., Ylstra, B. & Caldas, C. Array comparative genomic hybridization copy number profiling: a new tool for translational research in solid malignancies. *Semin Radiat Oncol* **18**, 98-104 (2008).
21. Oostlander, A.E., Meijer, G.A. & Ylstra, B. Microarray-based comparative genomic hybridization and its applications in human genetics. *Clin Genet* **66**, 488-495 (2004).
22. Kallioniemi, A. CGH microarrays and cancer. *Curr Opin Biotechnol* **19**, 36-40 (2008).
23. Pinkel, D. *et al.* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* **20**, 207-211 (1998).
24. Chen, X., Jorgenson, E. & Cheung, S.T. New tools for functional genomic analysis. *Drug Discov Today* **14**, 754-760 (2009).
25. Lapierre, J.M. & Tachdjian, G. Detection of chromosomal abnormalities by comparative genomic hybridization. *Curr Opin Obstet Gynecol* **17**, 171-177 (2005).
26. Capello, D. *et al.* Genome wide DNA-profiling of HIV-related B-cell lymphomas. *Br J Haematol* **148**, 245-255 (2010).
27. Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-721 (2009).

28. Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676-1679 (2008).
29. Novak, U. *et al.* The NF- κ B negative regulator TNFAIP3 (A20) is inactivated by somatic mutations and genomic deletions in marginal zone lymphomas. *Blood* **113**, 4918-4921 (2009).
30. Rinaldi, A. *et al.* Single nucleotide polymorphism-arrays provide new insights in the pathogenesis of post-transplant diffuse large B-cell lymphoma. *Br J Haematol* **149**(4):569-77 (2010).
31. Bergamaschi, A. *et al.* Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* **45**, 1033-1040 (2006).
32. Carrasco, D.R. *et al.* High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. *Cancer Cell* **9**, 313-325 (2006).
33. Kim, M.Y. *et al.* Recurrent genomic alterations with impact on survival in colorectal cancer identified by genome-wide array comparative genomic hybridization. *Gastroenterology* **131**, 1913-1924 (2006).
34. Mendrzyk, F. *et al.* Genomic and protein expression profiling identifies CDK6 as novel independent prognostic marker in medulloblastoma. *J Clin Oncol* **23**, 8853-8862 (2005).
35. Bernardini, M. *et al.* High-resolution mapping of genomic imbalance and identification of gene expression profiles associated with differential chemotherapy response in serous epithelial ovarian cancer. *Neoplasia* **7**, 603-613 (2005).

36. Han, W. *et al.* Genomic alterations identified by array comparative genomic hybridization as prognostic markers in tamoxifen-treated estrogen receptor-positive breast cancer. *BMC Cancer* **6**, 92 (2006).
37. Kim, S.W. *et al.* Analysis of chromosomal changes in serous ovarian carcinoma using high-resolution array comparative genomic hybridization: Potential predictive markers of chemoresistant disease. *Genes Chromosomes Cancer* **46**, 1-9 (2007).
38. Ollila, J. & Vihinen, M. B cells. *Int J Biochem Cell Biol* **37**, 518-523 (2005).
39. Monroe, J.G. & Dorshkind, K. Fate decisions regulating bone marrow and peripheral B lymphocyte development. *Adv Immunol* **95**, 1-50 (2007).
40. Murphy K.M., T.P., Walport M. *Janeway's Immunobiology*, Edn. 7th edition
41. Wilson, A. & Trumpp, A. Bone-marrow haematopoietic-stem-cell niches. *Nat Rev Immunol* **6**, 93-106 (2006).
42. Sagaert, X., Sprangers, B. & De Wolf-Peeters, C. The dynamics of the B follicle: understanding the normal counterpart of B-cell-derived malignancies. *Leukemia* **21**, 1378-1386 (2007).
43. Goodman, M.F., Scharff, M.D. & Romesberg, F.E. AID-initiated purposeful mutations in immunoglobulin genes. *Adv Immunol* **94**, 127-155 (2007).

44. Kupperts, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* **5**, 251-262 (2005).
45. Dave, B.J. *et al.* Cytogenetic characterization of diffuse large cell lymphoma using multi-color fluorescence in situ hybridization. *Cancer Genet Cytogenet* **132**, 125-132 (2002).
46. Gellert, M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* **71**, 101-132 (2002).
47. Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol* **8**, 22-33 (2008).
48. Klein, U. *et al.* Gene expression dynamics during germinal center transit in B cells. *Ann N Y Acad Sci* **987**, 166-172 (2003).
49. MacLennan, I.C. Somatic mutation. From the dark zone to the light. *Curr Biol* **4**, 70-72 (1994).
50. MacLennan, I.C. Germinal centers. *Annu Rev Immunol* **12**, 117-139 (1994).
51. Schonbeck, U. & Libby, P. The CD40/CD154 receptor/ligand dyad. *Cell Mol Life Sci* **58**, 4-43 (2001).
52. Goossens, T., Klein, U. & Kupperts, R. Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. *Proc Natl Acad Sci U S A* **95**, 2463-2468 (1998).
53. Kupperts, R. Somatic hypermutation and B cell receptor selection in normal and transformed human B cells. *Ann N Y Acad Sci* **987**, 173-179 (2003).

54. Larijani, M. *et al.* Methylation protects cytidines from AID-mediated deamination. *Mol Immunol* **42**, 599-604 (2005).
55. Liu, M. & Schatz, D.G. Balancing AID and DNA repair during somatic hypermutation. *Trends Immunol* **30**, 173-181 (2009).
56. Chang, C.C., Ye, B.H., Chaganti, R.S. & Dalla-Favera, R. BCL-6, a POZ/zinc-finger protein, is a sequence-specific transcriptional repressor. *Proc Natl Acad Sci U S A* **93**, 6947-6952 (1996).
57. Shaffer, A.L., Rosenwald, A. & Staudt, L.M. Lymphoid malignancies: the dark side of B-cell differentiation. *Nat Rev Immunol* **2**, 920-932 (2002).
58. Arpin, C. *et al.* Generation of memory B cells and plasma cells in vitro. *Science* **268**, 720-722 (1995).
59. Klein, U. & Dalla-Favera, R. Unexpected steps in plasma-cell differentiation. *Immunity* **26**, 543-544 (2007).
60. Sciammas, R. *et al.* Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* **25**, 225-236 (2006).
61. Hartmann, E.M., Ott, G. & Rosenwald, A. Molecular biology and genetics of lymphomas. *Hematol Oncol Clin North Am* **22**, 807-823, (2008).
62. Colomo, L. *et al.* Diffuse large B-cell lymphomas with plasmablastic differentiation represent a heterogeneous group of disease entities. *Am J Surg Pathol* **28**, 736-747 (2004).

63. Bea, S. *et al.* Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood* **106**, 3183-3190 (2005).
64. Lopez-Guillermo, A. *et al.* Diffuse large B-cell lymphoma: clinical and biological characterization and outcome according to the nodal or extranodal primary origin. *J Clin Oncol* **23**, 2797-2804 (2005).
65. Ennishi, D. *et al.* Does rituximab really induce hepatitis C virus reactivation? *J Clin Oncol* **26**, 4695-4696; author reply 4696 (2008).
66. Niitsu, N. *et al.* Clinicopathologic characteristics and treatment outcome of the addition of rituximab to chemotherapy for CD5-positive in comparison with CD5-negative diffuse large B-cell lymphoma. *Ann Oncol*.
67. Dunleavy, K., Staudt, L.M. & Wilson, W.H. The BCL-2 biomarker in the era of molecular diagnosis of diffuse large B-cell lymphoma. *Leuk Lymphoma* **48**, 1061-1063 (2007).
68. Obermann, E.C., Csato, M., Dirnhofer, S. & Tzankov, A. BCL2 gene aberration as an IPI-independent marker for poor outcome in non-germinal-centre diffuse large B cell lymphoma. *J Clin Pathol* **62**, 903-907 (2009).
69. Armitage, J.O., Bertrand Coiffier, M.D., Ph.D. , Riccardo Dalla-Favera, M.D., Nancy Lee Harris, M.D. & Peter M. Mauch, M.D. *Non-Hodgkin Lymphomas* Edn. 2nd edition (2009).
70. Pileri, S.A. *et al.* Diffuse large B-cell lymphoma: one or more entities? Present controversies and possible tools for its subclassification. *Histopathology* **41**, 482-509 (2002).

71. Lossos, I.S. & Levy, R. Diffuse large B-cell lymphoma: insights gained from gene expression profiling. *Int J Hematol* **77**, 321-329 (2003).
72. Keller, C.E. *et al.* Intrachromosomal rearrangement of chromosome 3q27: an under recognized mechanism of BCL6 translocation in B-cell non-Hodgkin lymphoma. *Hum Pathol* **37**, 1093-1099 (2006).
73. Pasqualucci, L. *et al.* Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J Exp Med* **203**, 311-317 (2006).
74. Lin, P. & Medeiros, L.J. High-grade B-cell lymphoma/leukemia associated with t(14;18) and 8q24/MYC rearrangement: a neoplasm of germinal center immunophenotype with poor prognosis. *Haematologica* **92**, 1297-1301 (2007).
75. Young, K.H. *et al.* Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAILreceptor-2, predict for poor survival in diffuse large B-cell lymphoma. *Blood* **110**, 4396-4405 (2007).
76. Stocklein, H. *et al.* Detailed mapping of chromosome 17p deletions reveals HIC1 as a novel tumor suppressor gene candidate telomeric to TP53 in diffuse large B-cell lymphoma. *Oncogene* **27**, 2613-2625 (2008).
77. Martin-Subero, J.I. *et al.* New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling. *Blood* **113**, 2488-2497 (2009).
78. Wang, X.M. *et al.* Identification and functional relevance of de novo DNA methylation in cancerous B-cell populations. *J Cell Biochem* **109**, 818-827.

79. Lossos, I.S. Diffuse large B cell lymphoma: from gene expression profiling to prediction of outcome. *Biol Blood Marrow Transplant* **14**, 108-111 (2008).
80. Rosenwald, A. & Staudt, L.M. Gene expression profiling of diffuse large B-cell lymphoma. *Leuk Lymphoma* **44 Suppl 3**, S41-47 (2003).
81. Tagawa, H. *et al.* Comparison of genome profiles for identification of distinct subgroups of diffuse large B-cell lymphoma. *Blood* **106**, 1770-1777 (2005).
82. Lossos, I.S. Molecular pathogenesis of diffuse large B-cell lymphoma. *J Clin Oncol* **23**, 6351-6357 (2005).
83. Alizadeh, A.A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-511 (2000).
84. Morin, R.D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* **42**, 181-185 (2010).
85. Stoffel, A. The NF-kappaB signalling pathway: a therapeutic target in lymphoid malignancies? *Expert Opin Ther Targets* **9**, 1045-1061 (2005).
86. Davis, R.E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88-92 (2010).
87. Bentz, M. *et al.* Gain of chromosome arm 9p is characteristic of primary mediastinal B-cell lymphoma (MBL): comprehensive molecular cytogenetic analysis and presentation of a novel MBL cell line. *Genes Chromosomes Cancer* **30**, 393-401 (2001).

88. Weniger, M.A. *et al.* Gains of REL in primary mediastinal B-cell lymphoma coincide with nuclear accumulation of REL protein. *Genes Chromosomes Cancer* **46**, 406-415 (2007).
89. Barth, T.F. *et al.* Gains of 2p involving the REL locus correlate with nuclear c-Rel protein accumulation in neoplastic cells of classical Hodgkin lymphoma. *Blood* **101**, 3681-3686 (2003).
90. Rosenwald, A. *et al.* Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. *J Exp Med* **198**, 851-862 (2003).
91. Coiffier, B. State-of-the-art therapeutics: diffuse large B-cell lymphoma. *J Clin Oncol* **23**, 6387-6393 (2005).
92. Lenz, G. *et al.* Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* **359**, 2313-2323 (2008).
93. Wu, G. & Keating, A. Biomarkers of potential prognostic significance in diffuse large B-cell lymphoma. *Cancer* **106**, 247-257 (2006).
94. Rancoita, P.M., Hutter, M., Bertoni, F. & Kwee, I. Bayesian DNA copy number analysis. *BMC Bioinformatics* **10**, 10 (2009).
95. Beroukhi, R. *et al.* Inferring loss-of-heterozygosity from unpaired tumors using high-density oligonucleotide SNP arrays. *PLoS Comput Biol* **2**, e41 (2006).
96. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).

97. Lee, D.D. & Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788-791 (1999).
98. Zheng, C.H., Huang, D.S., Zhang, L. & Kong, X.Z. Tumor clustering using nonnegative matrix factorization with gene selection. *IEEE Trans Inf Technol Biomed* **13**, 599-607 (2009).
99. Bland, J.M. & Altman, D.G. The logrank test. *BMJ* **328**, 1073 (2004).
100. Cheson, B.D. *et al.* Revised response criteria for malignant lymphoma. *J Clin Oncol* **25**, 579-586 (2007).
101. Cox, D.R. Regression models and life tables. *J R Stat Soc* **34**, 187-220 (1972).
102. Chu, L.C., Eberhart, C.G., Grossman, S.A. & Herman, J.G. Epigenetic silencing of multiple genes in primary CNS lymphoma. *Int J Cancer* **119**, 2487-2491 (2006).
103. Scandurra M, M.M., Greiner TC, Rancoita PMV, De Campos CP, Chan WC, Vose JM, Inghirami G, Chiappella A, Baldini L, Ponzoni M, Ferreri AJM, Franceschetti S, Gaidano G, Montes-Moreno S, Piris MA, Facchetti F, Tucci A, Nomdedeu J Fr, Lazure T, Lambotte O, Uccella S, Pinotti G, Pruneri G, Martinelli G, Tibiletti M G, Chigrinova E, Rinaldi A, Zucca E, Kwee I, Bertoni F. Genomic lesions associated with a different clinical outcome in diffuse large B-cell lymphoma treated with R-CHOP-21. *British Journal of Haematology* (2010).
104. Oikawa, T. ETS transcription factors: possible targets for cancer therapy. *Cancer Sci* **95**, 626-633 (2004).

105. Fuschiotti, P. *et al.* Analysis of the TCR alpha-chain rearrangement profile in human T lymphocytes. *Mol Immunol* **44**, 3380-3388 (2007).
106. Takeuchi, I. *et al.* The potential of copy number gains and losses, detected by array-based comparative genomic hybridization, for computational differential diagnosis of B-cell lymphomas and genetic regions involved in lymphomagenesis. *Haematologica* **94**, 61-69 (2009).
107. Booman, M. *et al.* Genomic alterations and gene expression in primary diffuse large B-cell lymphomas of immune-privileged sites: the importance of apoptosis and immunomodulatory pathways. *J Pathol* **216**, 209-217 (2008).
108. Scandurra, M. *et al.* Genomic profiling of Richter's syndrome: recurrent lesions and differences with de novo diffuse large B-cell lymphomas. *Hematol Oncol* (2009).
109. Tagawa, H. *et al.* Genome-wide array-based CGH for mantle cell lymphoma: identification of homozygous deletions of the proapoptotic gene BIM. *Oncogene* **24**, 1348-1358 (2005).
110. Haarer, C.F., Roberts, R.A., Frutiger, Y.M., Grogan, T.M. & Rimsza, L.M. Immunohistochemical classification of de novo, transformed, and relapsed diffuse large B-cell lymphoma into germinal center B-cell and nongerminal center B-cell subtypes correlates with gene expression profile and patient survival. *Arch Pathol Lab Med* **130**, 1819-1824 (2006).
111. Choi, W.W. *et al.* A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clin Cancer Res* **15**, 5494-5502 (2009).

112. Feugier, P. *et al.* Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. *J Clin Oncol* **23**, 4117-4126 (2005).
113. Armitage, J.O. & Weisenburger, D.D. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *J Clin Oncol* **16**, 2780-2795 (1998).
114. Stokke, T. *et al.* Loss of chromosome 11q21-23.1 and 17p and gain of chromosome 6p are independent prognostic indicators in B-cell non-Hodgkin's lymphoma. *Br J Cancer* **85**, 1900-1913 (2001).
115. Forconi, F. *et al.* Genome-wide DNA analysis identifies recurrent imbalances predicting outcome in chronic lymphocytic leukaemia with 17p deletion. *Br J Haematol* **143**, 532-536 (2008).
116. Martinez-Climent, J.A. *et al.* Loss of a novel tumor suppressor gene locus at chromosome 8p is associated with leukemic mantle cell lymphoma. *Blood* **98**, 3479-3482 (2001).
117. Rubio-Moscardo, F. *et al.* Characterization of 8p21.3 chromosomal deletions in B-cell lymphoma: TRAIL-R1 and TRAIL-R2 as candidate dosage-dependent tumor suppressor genes. *Blood* **106**, 3214-3222 (2005).
118. Bosch, N. *et al.* Nucleotide, cytogenetic and expression impact of the human chromosome 8p23.1 inversion polymorphism. *PLoS One* **4**, e8269 (2009).

119. Takeyama, K. *et al.* Integrative analysis reveals 53BP1 copy loss and decreased expression in a subset of human diffuse large B-cell lymphomas. *Oncogene* **27**, 318-322 (2008).
120. Khong, H.T. & Restifo, N.P. Natural selection of tumor variants in the generation of "tumor escape" phenotypes. *Nat Immunol* **3**, 999-1005 (2002).
121. Knoop, L., Hermans, C., Ferrant, A. & Constantinescu, S.N. Clinical implications of JAK2 mutations in myeloproliferative disorders. *Acta Clin Belg* **63**, 93-98 (2008).
122. Knoop, L., Hornakova, T., Royer, Y., Constantinescu, S.N. & Renauld, J.C. JAK kinases overexpression promotes in vitro cell transformation. *Oncogene* **27**, 1511-1519 (2008).
123. Ferrari, R., Berk, A.J. & Kurdiani, S.K. Viral manipulation of the host epigenome for oncogenic transformation. *Nat Rev Genet* **10**, 290-294 (2009).
124. John, S.A. & Garrett-Sinha, L.A. Blimp1: a conserved transcriptional repressor critical for differentiation of many tissues. *Exp Cell Res* **315**, 1077-1084 (2009).
125. Schmidlin, H., Diehl, S.A. & Blom, B. New insights into the regulation of human B-cell differentiation. *Trends Immunol* **30**, 277-285 (2009).
126. John, S.A., Clements, J.L., Russell, L.M. & Garrett-Sinha, L.A. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J Biol Chem* **283**, 951-962 (2008).

127. Maier, H. *et al.* Requirements for selective recruitment of Ets proteins and activation of mb-1/Ig-alpha gene transcription by Pax-5 (BSAP). *Nucleic Acids Res* **31**, 5483-5489 (2003).
128. Gallant, S. & Gilkeson, G. ETS transcription factors and regulation of immunity. *Arch Immunol Ther Exp (Warsz)* **54**, 149-163 (2006).
129. Oikawa, T. & Yamada, T. Molecular biology of the Ets family of transcription factors. *Gene* **303**, 11-34 (2003).
130. Schmidlin, H. *et al.* Spi-B inhibits human plasma cell differentiation by repressing BLIMP1 and XBP-1 expression. *Blood* **112**, 1804-1812 (2008).
131. Wang, D. *et al.* Ets-1 deficiency leads to altered B cell differentiation, hyperresponsiveness to TLR9 and autoimmune disease. *Int Immunol* **17**, 1179-1191 (2005).
132. Zhang, X.K. *et al.* The transcription factor Fli-1 modulates marginal zone and follicular B cell development in mice. *J Immunol* **181**, 1644-1654 (2008).
133. Jackers, P., Szalai, G., Moussa, O. & Watson, D.K. Ets-dependent regulation of target gene expression during megakaryopoiesis. *J Biol Chem* **279**, 52183-52190 (2004).
134. Lesault, I., Quang, C.T., Frampton, J. & Ghysdael, J. Direct regulation of BCL-2 by FLI-1 is involved in the survival of FLI-1-transformed erythroblasts. *EMBO J* **21**, 694-703 (2002).
135. Tibiletti, M.G. *et al.* BCL2, BCL6, MYC, MALT 1, and BCL10 rearrangements in nodal diffuse large B-cell lymphomas: a multicenter evaluation of a new set of fluorescent in situ hybridization

- probes and correlation with clinical outcome. *Hum Pathol* **40**, 645-652 (2009).
136. Shaffer, A.L. *et al.* Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* **17**, 51-62 (2002).
137. Shapiro-Shelef, M. & Calame, K. Regulation of plasma-cell development. *Nat Rev Immunol* **5**, 230-242 (2005).

8. ACKNOWLEDGEMENTS

There are many persons I would like to thank and since three years is a long long period of someone's life, I hope I won't forget anyone.

I would like to begin by thanking all the participating centers to this big project, for providing samples and clinical data:

- IOSI, Bellinzona
- Omaha, Nebraska, USA
- HSR, Milan, IT
- IEO, Milan, IT
- CNIO, Madrid, SP
- Hopitaux de Paris, Paris, FR
- Novara, IT
- Torino, IT
- Varese, IT
- Brescia, IT
- Policlinico, Milan, IT
- HSCSP, Barcelona, SP

The first personal thank of this chapter is for my boss, Dr. Francesco Bertoni, who gave me the possibility of participating to this amazing DLBCL project and to whom I owe all my professional improvements and my being successful in the last three years. A special thanks to Michael, not only for working hard on the clinical part and for sharing with me most of the enthusiasm that drove me till the end of my PhD, but also for curing me every time I was sick. I need to thank Paola, for helping me with all the bioinformatic tools of the present work and for sharing with me unforgettable moments. Moreover, I have to thank Andrea for all his precious suggestions and for listening to me any time I needed it, as a very good friend. A very special thank to all members of my group: Ivo for performing most of the analysis regarding arrays, Michela, Ekaterina, Monica, Cassio, Dr Emanule Zucca (for his great enthusiasm!) and last but not least Afua.

ShRNAs have been my nightmare for the last year, but in the end I succeeded in working with them. For this functional part I really need to thank Professor Roberto Piva and his group, for teaching me and helping me in realizing my experiments. A special thank goes to my friend Dora, for helping me with all the cell-biology experiments and another very special one is for Luisa for the great help with bacteria. I'd like to thank also my supervisor at University of Geneva, Professor Leonardo Scapozza, for coordinating my PhD in the last three years and for critical correction of the thesis.

In the end, I want to add some personal acknowledgements mainly dedicated to people sharing my life outside the lab. Thank you so much to Guido, for motivating me in never giving up, even when everything appeared so dark around me. Thank you because I would have never realized my dream, without you. You are the best part of it. Thank you to my brother Riccardo, to my uncle Cesare and to my sister in law Gloria for trusting in me every single day of my life. Thank you so much to Anna Bentivoglio for loving me and having faith on me as a real daughter.

And the last thank, which is the most important one, is for my parents, Anna Maria and Antonio, the only reason why I've been able to do every single step of the last 28 years. Thank you not only for understanding and motivating me, but mainly because I owe you both everything I've done so far and I'll be able to do for the rest of my life. In the end, thank you very much to all the people I was lucky enough to meet during the last three years and to all of them who trusted in me and in my enthusiasm.

This work was supported by: Oncosuisse grant OCS-1939-8-2006; Swiss National Science Foundation (grants 205321-112430, 205320-121886/1); Cantone Ticino ("Computational life science / Ticino in rete" program); Fondazione per la Ricerca e la Cura sui Linfomi (Lugano, Switzerland); PHS grant, UO1 CA 114778.

9. SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE 1

SUPERVISED ANALYSIS OF THE GENE EXPRESSION PROFILES OF CASES WITH OR WITHOUT 8p LOSS (8/54, 15%) IDENTIFIED 81 PROBES ALL UPREGULATED.

Probeset ID	Gene Title	Gene Symbol	p-value(1 vs. 2 and 4)	Ratio(1 vs. 2 and 4)	Fold-Change(1 vs. 2 and 4)	Fold-Change(1 vs. 2 and 4) (Description)
239699_s_at	postmeiotic segregation increased 2-like 1	PMS2L1	8,97E-06	0,679272	-1,47216	1 down vs 2 and 4
1563469_at	MRNA; cDNA DKFZp313M0417 (from clone DKFZp313M0417)	---	4,70E-05	1,44602	1,44602	1 up vs 2 and 4
202125_s_at	trafficking protein, kinesin binding 2	TRAK2	7,19E-05	1,60551	1,60551	1 up vs 2 and 4
219187_at	FK506 binding protein like	FKBPL	8,76E-05	0,746058	-1,34038	1 down vs 2 and 4
210113_s_at	NLR family, pyrin domain containing 1	NLRP1	9,48E-05	1,63322	1,63322	1 up vs 2 and 4
1552619_a_at	anillin, actin binding protein	ANLN	0,000108787	0,720666	-1,3876	1 down vs 2 and 4
218764_at	protein kinase C, eta	PRKCH	0,000124117	2,41495	2,41495	1 up vs 2 and 4
203232_s_at	ataxin 1	ATXN1	0,000143911	2,53198	2,53198	1 up vs 2 and 4
206106_at	mitogen-activated protein kinase 12 /// histone deacetylase 10	HDAC10 /// MAPK12	0,000145815	0,613011	-1,63129	1 down vs 2 and 4
210138_at	regulator of G-protein signaling 20	RGS20	0,000155474	0,751534	-1,33061	1 down vs 2 and 4
226538_at	Mannosidase, alpha, class 2A, member 1	MAN2A1	0,000155613	2,15058	2,15058	1 up vs 2 and 4
209846_s_at	butyrophilin, subfamily 3, member A2	BTN3A2	0,000190915	2,51129	2,51129	1 up vs 2 and 4
244252_at	hypothetical	LOC399884	0,00019591	0,79012	-1,26562	1 down vs 2

	LOC399884		1	7		and 4
210415_s_at	outer dense fiber of sperm tails 2	ODF2	0,00020428 8	0,70110 5	-1,42632	1 down vs 2 and 4
221262_s_at	solute carrier family 2 (facilitated glucose transporter), member 11	SLC2A11	0,00020770 7	0,78651 4	-1,27143	1 down vs 2 and 4
212614_at	AT rich interactive domain 5B (MRF1-like)	ARID5B	0,00021455 9	2,37364	2,37364	1 up vs 2 and 4
64064_at	GTPase, IMAP family member 5	GIMAP5	0,00021884 7	2,13167	2,13167	1 up vs 2 and 4
227150_at	metal-regulatory transcription factor 1	MTF1	0,00022505 5	1,85971	1,85971	1 up vs 2 and 4
220892_s_at	phosphoserine aminotransferase 1	PSAT1	0,00023895 1	0,61492 2	-1,62622	1 down vs 2 and 4
221903_s_at	cylindromatosis (turban tumor syndrome)	CYLD	0,00024898 5	1,84295	1,84295	1 up vs 2 and 4
227701_at	chromosome 10 open reading frame 118	C10orf118	0,00025874 2	1,48214	1,48214	1 up vs 2 and 4
204820_s_at	butyrophilin, subfamily 3, member A3 /// butyrophilin, subfamily 3, member A2	BTN3A2 /// BTN3A3	0,00027736 8	2,12311	2,12311	1 up vs 2 and 4
220050_at	chromosome 9 open reading frame 9	C9orf9	0,00027851 3	0,74292 4	-1,34603	1 down vs 2 and 4
222761_at	basic, immunoglobulin-like variable motif containing	BIVM	0,00028027 9	1,50329	1,50329	1 up vs 2 and 4
201855_s_at	ATM/ATR-Substrate Chk2-Interacting Zn2+-finger protein	ASCIZ	0,00028383 8	1,70577	1,70577	1 up vs 2 and 4
228933_at	Nance-Horan syndrome (congenital cataracts and dental anomalies)	NHS	0,00028733 6	1,50034	1,50034	1 up vs 2 and 4
231241_at	Similar to PCAF associated factor 65 beta isoform b	LOC645744	0,00029876 9	1,36281	1,36281	1 up vs 2 and 4
229202_at	Transcribed locus	---	0,00030449 9	1,8317	1,8317	1 up vs 2 and 4
203915_at	chemokine (C-X-C motif) ligand 9	CXCL9	0,00035884	3,4867	3,4867	1 up vs 2 and 4
217147_s_at	T cell receptor	TRAT1	0,00035990	3,10953	3,10953	1 up vs 2 and 4

	associated transmembrane adaptor 1		5			
209972_s_at	JTV1 gene	JTV1	0,00037027 1	0,70722 9	-1,41397	1 down vs 2 and 4
229495_at	Aminoacylase 1-like 2	ACY1L2	0,00037440 7	0,85048 7	-1,1758	1 down vs 2 and 4
1559114_a_at	Chemokine (C-X-C motif) receptor 7	CXCR7	0,00037598 3	0,80949 2	-1,23534	1 down vs 2 and 4
218805_at	GTPase, IMAP family member 5	GIMAP5	0,00038891 1	2,5946	2,5946	1 up vs 2 and 4
204595_s_at	stanniocalcin 1	STC1	0,00039538 5	0,53969 1	-1,85291	1 down vs 2 and 4
212929_s_at	family with sequence similarity 21, member B /// family with sequence similarity	FAM21A /// FAM21B /// FAM21C /// FAM21D	0,00040309 9	1,63725	1,63725	1 up vs 2 and 4
202136_at	zinc finger, MYND domain containing 11	ZMYND11	0,00043418 3	1,74488	1,74488	1 up vs 2 and 4
223653_x_a†	bruno-like 4, RNA binding protein (Drosophila)	BRUNOL4	0,00045080 1	0,80448	-1,24304	1 down vs 2 and 4
213295_at	cylindromatosis (turban tumor syndrome)	CYLD	0,00045711	2,12781	2,12781	1 up vs 2 and 4
206804_at	CD3g molecule, gamma (CD3-TCR complex)	CD3G	0,00046452 9	2,7684	2,7684	1 up vs 2 and 4
205689_at	pecanex-like 2 (Drosophila)	PCNXL2	0,00047119 5	1,24825	1,24825	1 up vs 2 and 4
240843_at	Transcribed locus	---	0,00047945 3	1,51045	1,51045	1 up vs 2 and 4
203645_s_at	CD163 molecule	CD163	0,00049639 4	5,4842	5,4842	1 up vs 2 and 4
213135_at	T-cell lymphoma invasion and metastasis 1	TIAM1	0,00049908 1	2,31252	2,31252	1 up vs 2 and 4
236697_at	Transcribed locus	---	0,00050201 3	0,73644 3	-1,35788	1 down vs 2 and 4
230449_x_a†	Transcribed locus, weakly similar to XP_519878.1 similar to ubiquitin-conjugati	---	0,00050437 9	0,62572	-1,59816	1 down vs 2 and 4
206545_at	CD28 molecule	CD28	0,00051456 2	2,7572	2,7572	1 up vs 2 and 4

237593_at	Transcribed locus	---	0,00053288 5	0,83395	-1,19911	1 down vs 2 and 4
209770_at	butyrophilin, subfamily 3, member A1	BTN3A1	0,00053676 9	1,57772	1,57772	1 up vs 2 and 4
242540_at	chromosome 11 open reading frame 47	C11orf47	0,00056796 1	0,81245 9	-1,23083	1 down vs 2 and 4
207268_x_at	abl interactor 2	ABI2	0,00057015 2	0,77766 4	-1,2859	1 down vs 2 and 4
211722_s_at	histone deacetylase 6	HDAC6	0,00057245 5	0,82476 3	-1,21247	1 down vs 2 and 4
222062_at	interleukin 27 receptor, alpha	IL27RA	0,00058143 1	2,0362	2,0362	1 up vs 2 and 4
241396_at	neural precursor cell expressed, developmentally down-regulated 4-like	NEDD4L	0,00058398 8	0,57087 1	-1,75171	1 down vs 2 and 4
212993_at	MRNA; cDNA DKFZp667B1718 (from clone DKFZp667B1718)	---	0,00059750 2	2,16493	2,16493	1 up vs 2 and 4
215049_x_at	CD163 molecule	CD163	0,00060001	5,16851	5,16851	1 up vs 2 and 4
204070_at	retinoic acid receptor responder (tazarotene induced) 3	RARRES3	0,00061489	2,77286	2,77286	1 up vs 2 and 4
217921_at	mannosidase, alpha, class 1A, member 2	MAN1A2	0,00061794 7	1,59201	1,59201	1 up vs 2 and 4
205831_at	CD2 molecule	CD2	0,00062410 3	3,07973	3,07973	1 up vs 2 and 4
211005_at	linker for activation of T cells	LAT	0,00062686 1	2,17492	2,17492	1 up vs 2 and 4
215659_at	Gasdermin-like	GSDML	0,00063051 5	1,4158	1,4158	1 up vs 2 and 4
1560557_at	CDNA clone IMAGE:4823654	---	0,00063712 2	0,85554 9	-1,16884	1 down vs 2 and 4
211793_s_at	abl interactor 2	ABI2	0,00064148 9	0,73908 9	-1,35302	1 down vs 2 and 4
218259_at	MKL/myocardin-like 2	MKL2	0,00064167 4	1,60678	1,60678	1 up vs 2 and 4
1559590_at	choline dehydrogenase	CHDH	0,00064230 2	0,71401	-1,40054	1 down vs 2 and 4
239462_at	zinc finger protein 284	ZNF284	0,00067576 5	1,36861	1,36861	1 up vs 2 and 4
212478_at	required for meiotic nuclear division 5 homolog A (S.	RMND5A	0,00067815	0,76541 5	-1,30648	1 down vs 2 and 4

	cerevisiae)					
229390_at	family with sequence similarity 26, member F	FAM26F	0,00073682 4	2,40713	2,40713	1 up vs 2 and 4
201854_s_at	ATM/ATR-Substrate Chk2-Interacting Zn2+-finger protein	ASCIZ	0,00075086	1,44874	1,44874	1 up vs 2 and 4
1568847_at	Homo sapiens, clone IMAGE:3453596, mRNA	---	0,00076030 1	0,82464 3	-1,21265	1 down vs 2 and 4
239035_at	5,10-methylenetetrahydrofolate reductase (NADPH)	MTHFR	0,00076125 1	1,41454	1,41454	1 up vs 2 and 4
236692_at	similar to DTW domain containing 2	LOC729839	0,00077302 8	1,40844	1,40844	1 up vs 2 and 4
226977_at	similar to bovine IgA regulatory protein	LOC492311	0,00077646 9	1,36513	1,36513	1 up vs 2 and 4
236404_at	Transcribed locus	---	0,00079006	1,57276	1,57276	1 up vs 2 and 4
204597_x_at†	stannocalcin 1	STC1	0,00079126 8	0,59108	-1,69182	1 down vs 2 and 4
232848_at	hCG1818123	hCG_1795283	0,00081653 7	0,80358 4	-1,24443	1 down vs 2 and 4
205167_s_at	cell division cycle 25 homolog C (S. pombe)	CDC25C	0,00082815 1	0,63990 8	-1,56273	1 down vs 2 and 4
235458_at	hepatitis A virus cellular receptor 2	HAVCR2	0,00085752 3	2,25804	2,25804	1 up vs 2 and 4
205692_s_at	CD38 molecule	CD38	0,00088321 6	2,44133	2,44133	1 up vs 2 and 4
225353_s_at	complement component 1, q subcomponent, C chain	C1QC	0,00090681 6	3,21697	3,21697	1 up vs 2 and 4
225417_at	enhancer of polycomb homolog 1 (Drosophila)	EPC1	0,00090935 4	1,47312	1,47312	1 up vs 2 and 4
208451_s_at	complement component 4A (Rodgers blood group) /// complement component 4B (Child	C4A /// C4B	0,00091819 7	2,38396	2,38396	1 up vs 2 and 4
232617_at	cathepsin S	CTSS	0,00092724 8	1,90372	1,90372	1 up vs 2 and 4
229035_s_at	kelch domain containing 4	KLHDC4	0,00092975 5	1,44427	1,44427	1 up vs 2 and 4
213958_at	CD6 molecule	CD6	0,00093199	1,80191	1,80191	1 up vs 2 and 4

			5			
241060_x_a †	---	---	0,00093479 2	1,4549	1,4549	1 up vs 2 and 4
1558957_s_ at	Transcribed locus	---	0,00093706 8	1,18256	1,18256	1 up vs 2 and 4
221894_at	aaF domain containing kinase 2	ADCK2	0,00094273 7	0,69889 9	-1,43082	1 down vs 2 and 4
232645_at	hypothetical protein LOC153684	LOC153684	0,00095143 2	1,41009	1,41009	1 up vs 2 and 4
202788_at	mitogen-activated protein kinase- activated protein kinase 3	MAPKAPK3	0,00095583 6	1,64127	1,64127	1 up vs 2 and 4
1566924_at	CDNA: FLJ20971 fis, clone ADSU01565	---	0,00096754 1	0,79445 2	-1,25873	1 down vs 2 and 4
209881_s_at	linker for activation of T cells	LAT	0,00097095 8	1,7568	1,7568	1 up vs 2 and 4
219885_at	schlafen family member 12	SLFN12	0,00100507	1,53594	1,53594	1 up vs 2 and 4
243400_x_a †	lymphotoxin beta receptor (TNFR superfamily, member 3)	LTBR	0,00101201	0,81059 2	-1,23367	1 down vs 2 and 4
209062_x_a †	nuclear receptor coactivator 3	NCOA3	0,00101741	0,68088 1	-1,46868	1 down vs 2 and 4
210031_at	CD247 molecule	CD247	0,00101862	2,44702	2,44702	1 up vs 2 and 4
206656_s_at	chromosome 20 open reading frame 3	C20orf3	0,00102149	1,63266	1,63266	1 up vs 2 and 4
237430_at	Transcribed locus	---	0,0010355	0,84401 3	-1,18482	1 down vs 2 and 4
213539_at	CD3d molecule, delta (CD3-TCR complex)	CD3D	0,0010466	2,33993	2,33993	1 up vs 2 and 4
203518_at	lysosomal trafficking regulator	LYST	0,00106086	2,0583	2,0583	1 up vs 2 and 4
224551_s_at	spectrin, beta, non- erythrocytic 4	SPTBN4	0,00107224	0,79700 5	-1,2547	1 down vs 2 and 4
226771_at	ATPase, Class I, type 8B, member 2	ATP8B2	0,00107435	1,82799	1,82799	1 up vs 2 and 4
240070_at	V-set and transmembrane domain containing 3	VSTM3	0,00108152	2,21455	2,21455	1 up vs 2 and 4
228362_s_at	Family with sequence similarity 26, member F	FAM26F	0,00108274	2,14702	2,14702	1 up vs 2 and 4
204787_at	V-set and immunoglobulin	VSIG4	0,00108534	4,89993	4,89993	1 up vs 2 and 4

	domain containing 4					
210817_s_at	calcium binding and coiled-coil domain 2	CALCOCO2	0,00108706	1,6901	1,6901	1 up vs 2 and 4
225773_at	ring finger and SPRY domain containing 1	RSPRY1	0,00108902	1,48412	1,48412	1 up vs 2 and 4
226528_at	metaxin 3	MTX3	0,00110702	1,4737	1,4737	1 up vs 2 and 4
1563032_at	CDNA clone IMAGE:5311737	---	0,00110904	0,91886 8	-1,0883	1 down vs 2 and 4
230015_at	progressive rod-cone degeneration	PRCD	0,0011158	0,61213 6	-1,63362	1 down vs 2 and 4
205945_at	interleukin 6 receptor	IL6R	0,00113768	2,1649	2,1649	1 up vs 2 and 4
206363_at	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)	MAF	0,00114145	2,22206	2,22206	1 up vs 2 and 4
244255_at	hypothetical protein LOC286114	LOC286114	0,00114237	0,85817 3	-1,16527	1 down vs 2 and 4
236008_at	Chromosome 9 open reading frame 73	C9orf73	0,00115489	1,42268	1,42268	1 up vs 2 and 4
215837_x_a †	PRO1621 protein	PRO1621	0,00116899	0,86088 8	-1,16159	1 down vs 2 and 4
212907_at	solute carrier family 30 (zinc transporter), member 1	SLC30A1	0,00117836	2,08606	2,08606	1 up vs 2 and 4
230989_s_at	testis-specific serine kinase 6	TSSK6	0,00118064	0,79542 3	-1,25719	1 down vs 2 and 4
206118_at	signal transducer and activator of transcription 4	STAT4	0,00118255	1,81899	1,81899	1 up vs 2 and 4
240537_s_at	hypothetical gene supported by AK131031; BC070380	LOC440356	0,00119148	0,81561 8	-1,22606	1 down vs 2 and 4
202269_x_a †	guanylate binding protein 1, interferon- inducible, 67kDa	GBP1	0,001216	2,78732	2,78732	1 up vs 2 and 4
209816_at	patched homolog 1 (Drosophila)	PTCH1	0,00122448	0,87098 6	-1,14812	1 down vs 2 and 4
213193_x_a †	T cell receptor beta variable 19 /// T cell receptor beta constant 1	TRBC1 /// TRBV19	0,00123019	2,41744	2,41744	1 up vs 2 and 4
229629_at	Transcribed locus	---	0,00124082	2,16326	2,16326	1 up vs 2 and 4
226352_at	junction-mediating and regulatory protein	JMY	0,00125007	1,63945	1,63945	1 up vs 2 and 4

233784_at	CDNA: FLJ22733 fis, clone HSI15907	---	0,00125476	0,69328 6	-1,44241	1 down vs 2 and 4
205544_s_at	complement component (3d/Epstein Barr virus) receptor 2	CR2	0,00125559	0,30379 9	-3,29166	1 down vs 2 and 4
209602_s_at	GATA binding protein 3	GATA3	0,00126657	1,46562	1,46562	1 up vs 2 and 4
221375_at	olfactory receptor, family 1, subfamily G, member 1	OR1G1	0,00127424	0,81118 1	-1,23277	1 down vs 2 and 4
1554677_s_at	CKLF-like MARVEL transmembrane domain containing 4	CMTM4	0,0012816	0,78509 8	-1,27373	1 down vs 2 and 4
213604_at	transcription elongation factor B (SIII), polypeptide 3 (110kDa, elongin A)	TCEB3	0,00128957	1,58335	1,58335	1 up vs 2 and 4
240750_at	---	---	0,00130788	0,86710 5	-1,15326	1 down vs 2 and 4
243972_at	Transcribed locus	---	0,00131944	0,87788 1	-1,13911	1 down vs 2 and 4
212548_s_at	FRY-like	FRYL	0,00132114	1,4119	1,4119	1 up vs 2 and 4
211178_s_at	proline-serine- threonine phosphatase interacting protein 1	PSTPIP1	0,00132385	1,89208	1,89208	1 up vs 2 and 4
1569069_s_at	tudor domain containing 3	TDRD3	0,00133427	1,57058	1,57058	1 up vs 2 and 4
229391_s_at	family with sequence similarity 26, member F	FAM26F	0,00134147	2,75795	2,75795	1 up vs 2 and 4
234841_x_a †	odorant binding protein 2B /// odorant binding protein 2A	OBP2A /// OBP2B	0,00134466	0,80820 1	-1,23732	1 down vs 2 and 4
227732_at	ataxin 7-like 1	ATXN7L1	0,00134765	1,5801	1,5801	1 up vs 2 and 4
1559268_at	ADAM metallopeptidase domain 23	ADAM23	0,00135313	0,86215	-1,15989	1 down vs 2 and 4
1556769_a_at	Homo sapiens, clone IMAGE:3883659, mRNA	---	0,00136964	0,84618 2	-1,18178	1 down vs 2 and 4
212990_at	synaptotagmin 1	SYNJ1	0,00137774	1,77514	1,77514	1 up vs 2 and 4
212613_at	butyrophilin, subfamily 3, member A2	BTN3A2	0,00138491	1,77373	1,77373	1 up vs 2 and 4
210915_x_a †	T cell receptor beta variable 19 /// T cell	TRBC1 /// TRBV19	0,00139077	2,55443	2,55443	1 up vs 2 and 4

	receptor beta constant 1					
213527_s_at	zinc finger protein 688	ZNF688	0,00139623	1,4399	1,4399	1 up vs 2 and 4
218939_at	leucine zipper-EF-hand containing transmembrane protein 1	LETM1	0,00140829	0,80416	-1,24353	1 down vs 2 and 4
205924_at	RAB3B, member RAS oncogene family	RAB3B	0,00141514	0,82858 5	-1,20688	1 down vs 2 and 4
215286_s_at	putative homeodomain transcription factor 2	PHTF2	0,00142009	0,77353 2	-1,29277	1 down vs 2 and 4
226520_at	Primary neuroblastoma cDNA, clone:Nbla11485	---	0,00142744	1,47267	1,47267	1 up vs 2 and 4
226741_at	solute carrier family 12 (potassium/chloride transporters), member 6	SLC12A6	0,00144264	1,40621	1,40621	1 up vs 2 and 4
220997_s_at	diaphanous homolog 3 (Drosophila)	DIAPH3	0,00144717	0,87427 3	-1,14381	1 down vs 2 and 4
242462_at	CDNA clone IMAGE:5301514	---	0,00145331	0,71480 3	-1,39899	1 down vs 2 and 4
212062_at	ATPase, Class II, type 9A	ATP9A	0,00145443	2,14943	2,14943	1 up vs 2 and 4
211339_s_at	IL2-inducible T-cell kinase	ITK	0,00146233	2,61192	2,61192	1 up vs 2 and 4
204084_s_at	ceroid-lipofuscinosis, neuronal 5	CLN5	0,0014729	2,02891	2,02891	1 up vs 2 and 4
202124_s_at	trafficking protein, kinesin binding 2	TRAK2	0,00148601	1,48329	1,48329	1 up vs 2 and 4
1553533_at	junctophilin 1	JPH1	0,00149142	0,75971 1	-1,31629	1 down vs 2 and 4
226074_at	protein phosphatase 1M (PP2C domain containing)	PPM1M	0,00150109	1,57055	1,57055	1 up vs 2 and 4
234820_at	MAS1 oncogene-like	MAS1L	0,00152773	0,84602 2	-1,182	1 down vs 2 and 4
218019_s_at	pyridoxal (pyridoxine, vitamin B6) kinase	PDXK	0,0015399	1,52524	1,52524	1 up vs 2 and 4
221584_s_at	potassium large conductance calcium-activated channel, subfamily M, alpha member	KCNMA1	0,00154933	1,96902	1,96902	1 up vs 2 and 4

212876_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 4	B4GALT4	0,00156068	0,76534 3	-1,3066	1 down vs 2 and 4
210118_s_at	interleukin 1, alpha	IL1A	0,00156646	0,71002 4	-1,4084	1 down vs 2 and 4
1557047_at	YEATS domain containing 2	YEATS2	0,00156815	0,76092 1	-1,3142	1 down vs 2 and 4
203513_at	spastic paraplegia 11 (autosomal recessive)	SPG11	0,00157819	1,67335	1,67335	1 up vs 2 and 4
211824_x_at	NLR family, pyrin domain containing 1	NLRP1	0,00158101	1,43393	1,43393	1 up vs 2 and 4
226155_at	KIAA1600	KIAA1600	0,00159434	1,72261	1,72261	1 up vs 2 and 4
1569225_at	sex comb on midleg- like 4 (Drosophila)	SCML4	0,00159692	1,95485	1,95485	1 up vs 2 and 4
1553957_at	zinc finger protein 564	ZNF564	0,00159722	1,26483	1,26483	1 up vs 2 and 4
206765_at	potassium inwardly- rectifying channel, subfamily J, member 2	KCNJ2	0,00161951	2,66605	2,66605	1 up vs 2 and 4
236658_at	Transcribed locus	---	0,00162038	0,79166 1	-1,26317	1 down vs 2 and 4
235432_at	nephronophthisis 3 (adolescent)	NPHP3	0,00162169	1,54849	1,54849	1 up vs 2 and 4
236248_x_at	transcriptional adaptor 2 (ADA2 homolog, yeast)-beta	MGC21874	0,00163041	1,63403	1,63403	1 up vs 2 and 4
202524_s_at	sparc/osteonectin, cwcw and kazal-like domains proteoglycan (testican) 2	SPOCK2	0,0016339	1,98503	1,98503	1 up vs 2 and 4
212370_x_at	family with sequence similarity 21, member B /// family with sequence similarity	FAM21A /// FAM21B	0,00165679	1,58417	1,58417	1 up vs 2 and 4
216562_at	---	---	0,00166849	0,81172 9	-1,23194	1 down vs 2 and 4
215417_at	exocyst complex component 6B	EXOC6B	0,00166963	0,86801 1	-1,15206	1 down vs 2 and 4
220577_at	GTPase, very large interferon inducible 1	GVIN1	0,00169227	1,68141	1,68141	1 up vs 2 and 4
240962_at	Transcribed locus	---	0,0016946	0,81729 1	-1,22355	1 down vs 2 and 4
212332_at	retinoblastoma-like 2 (p130)	RBL2	0,00169794	1,66156	1,66156	1 up vs 2 and 4
225929_s_at	ring finger protein 213	RNF213	0,00170595	1,92427	1,92427	1 up vs 2 and 4

209988_s_at	achaete-scute complex homolog 1 (Drosophila)	ASCL1	0,00172663	3,17412	3,17412	1 up vs 2 and 4
1557369_at	hypothetical protein LOC285401	LOC285401	0,00172678	0,76466 1	-1,30777	1 down vs 2 and 4
1559425_at	MRNA; cDNA DKFZp667O095 (from clone DKFZp667O095)	---	0,00173811	2,04119	2,04119	1 up vs 2 and 4
232484_at	MRNA; cDNA DKFZp434O1311 (from clone DKFZp434O1311)	---	0,00174134	0,81431 8	-1,22802	1 down vs 2 and 4
205926_at	interleukin 27 receptor, alpha	IL27RA	0,00174471	1,5765	1,5765	1 up vs 2 and 4
1561439_at	CDNA clone IMAGE:5301065	---	0,00174591	0,79814 6	-1,2529	1 down vs 2 and 4
201160_s_at	cold shock domain protein A	CSDA	0,00175145	1,69961	1,69961	1 up vs 2 and 4
240433_x_at	CDNA clone IMAGE:4811412	---	0,00175572	0,84162 3	-1,18818	1 down vs 2 and 4
215183_at	Clone HQ0072	---	0,00175706	0,82542 1	-1,2115	1 down vs 2 and 4
211055_s_at	inversin	INVS	0,00176182	0,80163 8	-1,24745	1 down vs 2 and 4
223192_at	solute carrier family 25, member 28	SLC25A28	0,00177928	1,39759	1,39759	1 up vs 2 and 4
232043_at	guanine nucleotide binding protein (G protein), gamma 7	GNG7	0,00178146	0,77936 9	-1,28309	1 down vs 2 and 4
231577_s_at	guanylate binding protein 1, interferon-inducible, 67kDa	GBP1	0,00179241	2,7618	2,7618	1 up vs 2 and 4
209348_s_at	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)	MAF	0,00180257	2,11602	2,11602	1 up vs 2 and 4
1557890_at	hypothetical protein LOC729178	LOC729178	0,00180477	0,88420 2	-1,13096	1 down vs 2 and 4
239626_x_at	Transmembrane 9 superfamily member 3	TM9SF3	0,0018167	0,78628 4	-1,2718	1 down vs 2 and 4
202761_s_at	spectrin repeat containing, nuclear envelope 2	SYNE2	0,00182006	1,93133	1,93133	1 up vs 2 and 4
242268_at	CUG triplet repeat, RNA binding protein 2	CUGBP2	0,00183338	2,11317	2,11317	1 up vs 2 and 4

220400_at	vacuolar protein sorting 13 homolog B (yeast)	VPS13B	0,00183364	0,873697	-1,14456	1 down vs 2 and 4
234633_at	keratin associated protein 4-14 /// hypothetical LOC653240	KRTAP4-14 /// LOC653240	0,00183668	0,855376	-1,16908	1 down vs 2 and 4
1557581_x_at	CDNA clone IMAGE:3923254	---	0,00184261	1,12703	1,12703	1 up vs 2 and 4
236460_at	Transcribed locus	---	0,00185642	0,772793	-1,29401	1 down vs 2 and 4
229584_at	leucine-rich repeat kinase 2	LRRK2	0,00187496	2,65582	2,65582	1 up vs 2 and 4
241825_at	chromosome 2 open reading frame 60	C2orf60	0,00187579	1,26975	1,26975	1 up vs 2 and 4
233411_at	CDNA FLJ12123 fis, clone MAMMA1000133	---	0,00188067	1,38466	1,38466	1 up vs 2 and 4
235291_s_at	hypothetical protein LOC643977	FLJ32255	0,00188373	1,93246	1,93246	1 up vs 2 and 4
222786_at	carbohydrate (chondroitin 4) sulfotransferase 12	CHST12	0,00188435	1,383	1,383	1 up vs 2 and 4
239089_at	Transcribed locus	---	0,00188986	1,15964	1,15964	1 up vs 2 and 4
241435_at	---	---	0,00189474	1,67865	1,67865	1 up vs 2 and 4
205459_s_at	neuronal PAS domain protein 2	NPAS2	0,00192166	0,829343	-1,20577	1 down vs 2 and 4
224162_s_at	F-box protein 31	FBXO31	0,00193862	1,34105	1,34105	1 up vs 2 and 4
235252_at	Kinase suppressor of ras 1	KSR1	0,00196958	1,69124	1,69124	1 up vs 2 and 4
211464_x_at	caspase 6, apoptosis-related cysteine peptidase	CASP6	0,00198203	1,2571	1,2571	1 up vs 2 and 4
235211_at	Transcribed locus	---	0,00198559	0,815826	-1,22575	1 down vs 2 and 4
210363_s_at	sodium channel, voltage-gated, type II, beta	SCN2B	0,00200031	0,855415	-1,16902	1 down vs 2 and 4
1560407_at	MAP/microtubule affinity-regulating kinase 1	MARK1	0,00200751	0,753131	-1,32779	1 down vs 2 and 4
213732_at	Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	TCF3	0,00200867	0,842276	-1,18726	1 down vs 2 and 4

218373_at	AKT interacting protein	AKTIP	0,00202054	1,37285	1,37285	1 up vs 2 and 4
219839_x_at	T-cell leukemia/lymphoma 6	TCL6	0,00202253	0,841448	-1,18843	1 down vs 2 and 4
218380_at	NLR family, pyrin domain containing 1	NLRP1	0,00203309	1,64815	1,64815	1 up vs 2 and 4
203217_s_at	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	ST3GAL5	0,00204238	1,66446	1,66446	1 up vs 2 and 4
227756_at	family with sequence similarity 81, member A	FAM81A	0,00204336	0,686135	-1,45744	1 down vs 2 and 4
222782_s_at	GEM interacting protein	GMIP	0,00204359	0,859328	-1,1637	1 down vs 2 and 4
244080_at	Full-length cDNA clone CS0DI011YD16 of Placenta Cot 25-normalized of Homo sapien	---	0,00205235	0,82741	-1,20859	1 down vs 2 and 4
220585_at	hexokinase domain containing 1	HKDC1	0,00205633	0,86856	-1,15133	1 down vs 2 and 4
1555613_at	zeta-chain (TCR) associated protein kinase 70kDa	ZAP70	0,00206747	2,16671	2,16671	1 up vs 2 and 4
1558976_x_at	family with sequence similarity 100, member A	FAM100A	0,0020783	0,828204	-1,20743	1 down vs 2 and 4
238027_at	hypothetical protein LOC202051	LOC202051	0,00209353	0,759932	-1,31591	1 down vs 2 and 4
214946_x_at	family with sequence similarity 21, member B /// family with sequence similarity	FAM21A /// FAM21B /// FAM21C /// FAM21D	0,00209439	1,44293	1,44293	1 up vs 2 and 4
229367_s_at	GTPase, IMAP family member 6	GIMAP6	0,00210183	2,483	2,483	1 up vs 2 and 4
206666_at	granzyme K (granzyme 3; tryptase II)	GZMK	0,00211136	3,1646	3,1646	1 up vs 2 and 4
225785_at	receptor accessory protein 3	REEP3	0,00211997	1,78052	1,78052	1 up vs 2 and 4
226312_at	rapamycin-insensitive companion of mTOR	RICTOR	0,00216906	1,60587	1,60587	1 up vs 2 and 4
209680_s_at	kinesin family member C1	KIFC1	0,00217306	0,710961	-1,40655	1 down vs 2 and 4
218559_s_at	v-maf musculoaponeurotic fibrosarcoma	MAFB	0,00217438	2,37709	2,37709	1 up vs 2 and 4

	oncogene homolog B (avian)					
240665_at	---	---	0,00217646	1,92887	1,92887	1 up vs 2 and 4
232506_s_at	chromosome 15 open reading frame 41	C15orf41	0,00218332	0,888707	-1,12523	1 down vs 2 and 4
213520_at	RecQ protein-like 4	RECQL4	0,00219189	0,714759	-1,39907	1 down vs 2 and 4
201448_at	TIA1 cytotoxic granule-associated RNA binding protein	TIA1	0,00219702	1,82195	1,82195	1 up vs 2 and 4
227627_at	serum/glucocorticoid regulated kinase family, member 3	SGK3	0,00220856	1,69211	1,69211	1 up vs 2 and 4
227717_at	FLJ41603 protein	FLJ41603	0,00221382	0,515558	-1,93965	1 down vs 2 and 4
221250_s_at	MAX dimerization protein 3	MXD3	0,00223033	0,751244	-1,33113	1 down vs 2 and 4
244326_at	Transcribed locus	---	0,00223794	0,769801	-1,29904	1 down vs 2 and 4
226804_at	family with sequence similarity 20, member A	FAM20A	0,00224865	1,97902	1,97902	1 up vs 2 and 4
231776_at	eomesodermin homolog (Xenopus laevis)	EOMES	0,00225437	2,74193	2,74193	1 up vs 2 and 4
220723_s_at	hypothetical protein FLJ21511	FLJ21511	0,002274	0,824028	-1,21355	1 down vs 2 and 4
214032_at	zeta-chain (TCR) associated protein kinase 70kDa	ZAP70	0,00228054	2,00138	2,00138	1 up vs 2 and 4
216663_s_at	zinc finger, MYND-type containing 10	ZMYND10	0,00228221	0,790492	-1,26503	1 down vs 2 and 4
203799_at	CD302 molecule	CD302	0,00229076	1,95906	1,95906	1 up vs 2 and 4
218204_s_at	FYVE and coiled-coil domain containing 1	FYCO1	0,00229407	1,5569	1,5569	1 up vs 2 and 4
226508_at	polyhomeotic homolog 3 (Drosophila)	PHC3	0,00230135	1,54016	1,54016	1 up vs 2 and 4
235996_at	Ras association (RalGDS/AF-6) domain family 8	RASSF8	0,00230481	0,794271	-1,25902	1 down vs 2 and 4
221905_at	cylindromatosis (turban tumor syndrome)	CYLD	0,00230604	1,77038	1,77038	1 up vs 2 and 4
219368_at	nucleosome assembly protein 1-like 2	NAP1L2	0,00230724	1,48484	1,48484	1 up vs 2 and 4

213373_s_at	caspase 8, apoptosis-related cysteine peptidase	CASP8	0,00231299	1,41737	1,41737	1 up vs 2 and 4
235791_x_at	chromodomain helicase DNA binding protein 1	CHD1	0,00231564	1,62038	1,62038	1 up vs 2 and 4
1569303_s_at	regulator of G-protein signaling 20	RGS20	0,0023231	0,85317	-1,1721	1 down vs 2 and 4
1553619_x_at	tripartite motif-containing 43 /// similar to bloodthirsty /// similar to tripar	LOC643445 /// LOC653192 /// TRIM43	0,00232367	0,771344	-1,29644	1 down vs 2 and 4
230450_at	Transcribed locus, moderately similar to XP_517655.1 similar to KIAA0825 protei	---	0,00233816	0,723827	-1,38155	1 down vs 2 and 4
209975_at	cytochrome P450, family 2, subfamily E, polypeptide 1	CYP2E1	0,00235326	0,71584	-1,39696	1 down vs 2 and 4
228348_at	lines homolog 1 (Drosophila)	LINS1	0,00236028	1,40141	1,40141	1 up vs 2 and 4
219777_at	GTPase, IMAP family member 6	GIMAP6	0,00236489	2,33611	2,33611	1 up vs 2 and 4
212204_at	transmembrane protein 87A	TMEM87A	0,00238669	1,59273	1,59273	1 up vs 2 and 4
1565535_x_at	---	---	0,00240493	0,859289	-1,16375	1 down vs 2 and 4
204861_s_at	NLR family, apoptosis inhibitory protein /// neuronal apoptosis inhibitory prote	NAIP /// NAIP1B	0,00246894	1,51568	1,51568	1 up vs 2 and 4
218278_at	WD repeat domain 74	WDR74	0,00247581	0,784155	-1,27526	1 down vs 2 and 4
229123_at	Zinc finger protein 225	ZNF225	0,00248143	1,47758	1,47758	1 up vs 2 and 4
214091_s_at	glutathione peroxidase 3 (plasma)	GPX3	0,0024823	2,56081	2,56081	1 up vs 2 and 4
209839_at	dynammin 3	DNM3	0,00248301	0,813825	-1,22876	1 down vs 2 and 4
239744_at	Transcribed locus	---	0,0024861	1,35685	1,35685	1 up vs 2 and 4
240283_at	---	---	0,0025065	0,84411	-1,18468	1 down vs 2 and 4
210360_s_at	metastasis suppressor 1	MTSS1	0,00250804	0,890106	-1,12346	1 down vs 2 and 4
218232_at	complement component 1, q	C1QA	0,00250815	3,41647	3,41647	1 up vs 2 and 4

	subcomponent, A chain					
229962_at	leucine rich repeat containing 37, member A3	LRRC37A3	0,00250977	0,814478	-1,22778	1 down vs 2 and 4
244251_at	Lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa	LCP2	0,00251086	1,61245	1,61245	1 up vs 2 and 4
213388_at	similar to phosphodiesterase 4D interacting protein isoform 2	LOC727942	0,00251888	1,65325	1,65325	1 up vs 2 and 4
1561604_at	CDNA clone IMAGE:4796102	---	0,00254155	0,823175	-1,21481	1 down vs 2 and 4
213000_at	MORC family CW-type zinc finger 3	MORC3	0,00254364	1,5143	1,5143	1 up vs 2 and 4
218552_at	enoyl Coenzyme A hydratase domain containing 2	ECHDC2	0,00256054	1,4647	1,4647	1 up vs 2 and 4
208634_s_at	microtubule-actin crosslinking factor 1	MACF1	0,00256871	1,70382	1,70382	1 up vs 2 and 4
220418_at	ubiquitin associated and SH3 domain containing, A	UBASH3A	0,00257055	2,07696	2,07696	1 up vs 2 and 4
205052_at	AU RNA binding protein/enoyl-Coenzyme A hydratase	AUH	0,00257435	1,67855	1,67855	1 up vs 2 and 4
202746_at	integral membrane protein 2A	ITM2A	0,00258123	2,6758	2,6758	1 up vs 2 and 4
229680_at	Coiled-coil domain containing 44	CCDC44	0,00259707	0,833834	-1,19928	1 down vs 2 and 4
218217_at	serine carboxypeptidase 1	SCPEP1	0,00259844	2,25639	2,25639	1 up vs 2 and 4
210943_s_at	lysosomal trafficking regulator	LYST	0,00259862	1,79924	1,79924	1 up vs 2 and 4
234151_at	CDNA: FLJ20976 fis, clone ADSU01764	---	0,00261782	1,58004	1,58004	1 up vs 2 and 4
202441_at	ER lipid raft associated 1	ERLIN1	0,00263816	1,57869	1,57869	1 up vs 2 and 4
230746_s_at	Stanniocalcin 1	STC1	0,00264199	0,592645	-1,68735	1 down vs 2 and 4
1552303_a_at	transmembrane protein 106A /// similar	LOC728772 /// TMEM106A	0,00265547	1,30144	1,30144	1 up vs 2 and 4

	to transmembrane protein 106A					
210572_at	protocadherin alpha 2	PCDHA2	0,00265703	0,86742 2	-1,15284	1 down vs 2 and 4
221245_s_at	frizzled homolog 5 (Drosophila)	FZD5	0,00266804	1,69199	1,69199	1 up vs 2 and 4
204417_at	galactosylceramidase	GALC	0,00267076	1,89816	1,89816	1 up vs 2 and 4
205453_at	homeobox B2	HOXB2	0,00268171	1,4709	1,4709	1 up vs 2 and 4
202270_at	guanylate binding protein 1, interferon-inducible, 67kDa	GBP1	0,00268289	2,58146	2,58146	1 up vs 2 and 4
229165_at	Mitochondrial ribosomal protein L12	MRPL12	0,00268962	0,78247 4	-1,278	1 down vs 2 and 4
216778_s_at	cylicin, basic protein of sperm head cytoskeleton 1	CYLC1	0,00269179	0,89123 8	-1,12203	1 down vs 2 and 4
205105_at	mannosidase, alpha, class 2A, member 1	MAN2A1	0,00270996	1,96976	1,96976	1 up vs 2 and 4
224556_s_at	LIM homeobox 6	LHX6	0,00271222	0,83489 5	-1,19776	1 down vs 2 and 4
1556601_a_at	Spermatogenesis associated 13	SPATA13	0,00271685	1,64119	1,64119	1 up vs 2 and 4
226109_at	chromosome 21 open reading frame 91	C21orf91	0,00273933	1,73406	1,73406	1 up vs 2 and 4
222670_s_at	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	MAFB	0,00275019	1,87098	1,87098	1 up vs 2 and 4
216592_at	melanoma antigen family C, 3	MAGEC3	0,00276	0,86318 4	-1,1585	1 down vs 2 and 4
214428_x_at	complement component 4A (Rodgers blood group) /// complement component 4B (Child	C4A /// C4B	0,00276126	1,65225	1,65225	1 up vs 2 and 4
215951_at	TBC1 domain family, member 2B	TBC1D2B	0,00277862	0,82210 3	-1,21639	1 down vs 2 and 4
200884_at	creatine kinase, brain	CKB	0,00279594	0,64128 1	-1,55938	1 down vs 2 and 4
232024_at	GTPase, IMAP family member 2	GIMAP2	0,00280598	2,69019	2,69019	1 up vs 2 and 4
236693_at	CDNA clone IMAGE:4797645	---	0,00282103	0,78873 3	-1,26786	1 down vs 2 and 4
221404_at	interleukin 1 family,	IL1F6	0,00282939	0,79837	-1,25255	1 down vs 2

	member 6 (epsilon)			1		and 4
220655_at	TNFAIP3 interacting protein 3	TNIP3	0,00284393	1,47094	1,47094	1 up vs 2 and 4
214753_at	Phosphonoformate immuno-associated protein 5	PFAAP5	0,00286811	1,79696	1,79696	1 up vs 2 and 4
204925_at	cystinosis, nephropathic	CTNS	0,00287336	1,40403	1,40403	1 up vs 2 and 4
235627_at	profilin family, member 4	PFN4	0,00287957	0,81845 2	-1,22182	1 down vs 2 and 4
2028_s_at	E2F transcription factor 1	E2F1	0,00289211	0,81046 1	-1,23387	1 down vs 2 and 4
235306_at	GTPase, IMAP family member 8	GIMAP8	0,00289908	2,32352	2,32352	1 up vs 2 and 4
222009_at	cementum protein 1	CEMP1	0,00293267	0,8011	-1,24828	1 down vs 2 and 4
221002_s_at	tetraspanin 14	TSPAN14	0,00294962	1,46994	1,46994	1 up vs 2 and 4
225914_s_at	calcium binding protein 39-like	CAB39L	0,00295493	0,66789 3	-1,49725	1 down vs 2 and 4
224285_at	G protein-coupled receptor 174	GPR174	0,00295793	1,40977	1,40977	1 up vs 2 and 4
201061_s_at	stomatin	STOM	0,00296228	2,09344	2,09344	1 up vs 2 and 4
223085_at	ring finger protein 19A	RNF19A	0,0029848	1,74231	1,74231	1 up vs 2 and 4
212486_s_at	FYN oncogene related to SRC, FGR, YES	FYN	0,00299152	1,82121	1,82121	1 up vs 2 and 4
219556_at	chromosome 16 open reading frame 59	C16orf59	0,00299404	0,73789 1	-1,35521	1 down vs 2 and 4
209687_at	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	CXCL12	0,00299877	3,13835	3,13835	1 up vs 2 and 4
232744_x_at	CDNA FLJ11681 fis, clone HEMBA1004865	---	0,00300843	1,43641	1,43641	1 up vs 2 and 4
204860_s_at	NLR family, apoptosis inhibitory protein /// neuronal apoptosis inhibitory prote	NAIP /// NAIP1B	0,00301405	1,70912	1,70912	1 up vs 2 and 4
228532_at	chromosome 1 open reading frame 162	C1orf162	0,00301923	2,24023	2,24023	1 up vs 2 and 4
202832_at	GRIP and coiled-coil domain containing 2	GCC2	0,00302353	1,59537	1,59537	1 up vs 2 and 4
221899_at	Phosphonoformate immuno-associated protein 5	PFAAP5	0,0030286	1,77799	1,77799	1 up vs 2 and 4
239158_at	Transcribed locus	---	0,00303358	0,87602	-1,14152	1 down vs 2

				7		and 4
39582_at	cylindromatosis (turban tumor syndrome)	CYLD	0,00303742	1,59189	1,59189	1 up vs 2 and 4
212188_at	potassium channel tetramerisation domain containing 12	KCTD12	0,00304681	1,95923	1,95923	1 up vs 2 and 4
219243_at	GTPase, IMAP family member 4	GIMAP4	0,00305001	2,30452	2,30452	1 up vs 2 and 4
218846_at	mediator complex subunit 23	MED23	0,00305082	1,48743	1,48743	1 up vs 2 and 4
214028_x_a †	tudor domain containing 3	TDRD3	0,00305714	1,35281	1,35281	1 up vs 2 and 4
201975_at	CAP-GLY domain containing linker protein 1	CLIP1	0,00307757	1,74718	1,74718	1 up vs 2 and 4
219233_s_at	gasdermin-like	GSDML	0,00308551	1,53819	1,53819	1 up vs 2 and 4
1569433_at	sterile alpha motif domain containing 5	SAMD5	0,00309177	0,89525 4	-1,117	1 down vs 2 and 4
203124_s_at	solute carrier family 11 (proton-coupled divalent metal ion transporters), membe	SLC11A2	0,00309407	1,60871	1,60871	1 up vs 2 and 4
209379_s_at	KIAA1128	KIAA1128	0,00309713	1,4974	1,4974	1 up vs 2 and 4
221081_s_at	DENN/MADD domain containing 2D	DENND2D	0,00311566	1,82786	1,82786	1 up vs 2 and 4
233075_at	hect domain and RLD 2 pseudogene 7	HERC2P7	0,00312394	0,70548 9	-1,41746	1 down vs 2 and 4
1566934_at	TYRO3P protein tyrosine kinase pseudogene	TYRO3P	0,0031317	0,82443 2	-1,21296	1 down vs 2 and 4
230350_at	Transcribed locus	---	0,00313233	1,43203	1,43203	1 up vs 2 and 4
242807_at	fibronectin type III and SPRY domain containing 1-like	FSD1L	0,00313759	0,70196 1	-1,42458	1 down vs 2 and 4
201535_at	ubiquitin-like 3	UBL3	0,0031377	1,61016	1,61016	1 up vs 2 and 4
204224_s_at	GTP cyclohydrolase 1 (dopa-responsive dystonia)	GCH1	0,00315208	2,24764	2,24764	1 up vs 2 and 4
214511_x_a †	Fc fragment of IgG, high affinity lb, receptor (CD64)	FCGR1B	0,00316651	3,59366	3,59366	1 up vs 2 and 4
213835_x_a †	GTP binding protein 3 (mitochondrial)	GTPBP3	0,00317256	0,69751 3	-1,43367	1 down vs 2 and 4
228034_x_a	alkB, alkylation repair	ALKBH5	0,0031748	0,82289	-1,21523	1 down vs 2

†	homolog 5 (E. coli)			3		and 4
230444_at	Transcribed locus	---	0,00317557	1,55361	1,55361	1 up vs 2 and 4
210038_at	protein kinase C, theta	PRKCQ	0,00317856	1,95135	1,95135	1 up vs 2 and 4
215711_s_at	WEE1 homolog (S. pombe)	WEE1	0,0031932	0,62813 3	-1,59202	1 down vs 2 and 4
1555056_at	cyclin G2	CCNG2	0,00319419	0,74397 9	-1,34412	1 down vs 2 and 4
228898_s_at	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subf	SMARCB1	0,00319614	0,79873 4	-1,25198	1 down vs 2 and 4
210105_s_at	FYN oncogene related to SRC, FGR, YES	FYN	0,00320013	1,79012	1,79012	1 up vs 2 and 4
243918_at	Transcribed locus	---	0,00320663	0,66629 8	-1,50083	1 down vs 2 and 4
218553_s_at	potassium channel tetramerisation domain containing 15	KCTD15	0,00321656	0,79018 2	-1,26553	1 down vs 2 and 4
222915_s_at	B-cell scaffold protein with ankyrin repeats 1	BANK1	0,00322024	0,45111 5	-2,21673	1 down vs 2 and 4
240602_at	HBS1-like (S. cerevisiae)	HBS1L	0,00322822	1,2221	1,2221	1 up vs 2 and 4
206320_s_at	SMAD family member 9	SMAD9	0,00323028	0,80965 3	-1,2351	1 down vs 2 and 4
1555271_a_at	telomerase reverse transcriptase	TERT	0,00324552	0,84023 4	-1,19015	1 down vs 2 and 4
218276_s_at	salvador homolog 1 (Drosophila)	SAV1	0,00324934	1,3838	1,3838	1 up vs 2 and 4
220877_at	---	---	0,0032522	0,76580 2	-1,30582	1 down vs 2 and 4
221528_s_at	engulfment and cell motility 2	ELMO2	0,00327053	1,55034	1,55034	1 up vs 2 and 4
203666_at	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	CXCL12	0,00327278	1,81831	1,81831	1 up vs 2 and 4
235812_at	transmembrane protein 188	TMEM188	0,00328068	1,69283	1,69283	1 up vs 2 and 4
206236_at	G protein-coupled receptor 4	GPR4	0,00329456	0,81780 6	-1,22278	1 down vs 2 and 4
242750_at	methylmalonic aciduria (cobalamin deficiency) cblA type	MMAA	0,0033029	1,48825	1,48825	1 up vs 2 and 4
212760_at	ubiquitin protein ligase E3 component n-	UBR2	0,00331652	1,48336	1,48336	1 up vs 2 and 4

	recognin 2					
240617_at	Transcribed locus, strongly similar to XP_001075280.1 similar to I-kappa-B-rela	---	0,00331896	0,81231 9	-1,23104	1 down vs 2 and 4
212687_at	LIM and senescent cell antigen-like domains 1	LIMS1	0,00332703	1,87935	1,87935	1 up vs 2 and 4
212607_at	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	AKT3	0,00333027	1,4938	1,4938	1 up vs 2 and 4
240265_at	TRAF3 interacting protein 3	TRAF3IP3	0,00333617	1,6342	1,6342	1 up vs 2 and 4
235229_at	Transcribed locus, weakly similar to XP_001117086.1 similar to Olfactory recept	---	0,00333919	2,11541	2,11541	1 up vs 2 and 4
227274_at	Transcribed locus, strongly similar to XP_524440.1 similar to Phospholipid hydr	---	0,00334113	1,56489	1,56489	1 up vs 2 and 4
1564002_a_at	chromosome 6 open reading frame 199	C6orf199	0,00335201	1,48675	1,48675	1 up vs 2 and 4
1559753_at	CDNA FLJ36163 fis, clone TESTI2026014	---	0,00336445	0,80248 4	-1,24613	1 down vs 2 and 4
227258_at	chromosome 10 open reading frame 46	C10orf46	0,00336869	0,79620 9	-1,25595	1 down vs 2 and 4
216950_s_at	Fc fragment of IgG, high affinity Ia, receptor (CD64)	FCGR1A	0,00338422	3,34675	3,34675	1 up vs 2 and 4
224675_at	mesoderm development candidate 2	MESDC2	0,0033895	1,50229	1,50229	1 up vs 2 and 4
236656_s_at	Full length insert cDNA Y137C01	---	0,00339474	1,31406	1,31406	1 up vs 2 and 4
218918_at	mannosidase, alpha, class 1C, member 1	MAN1C1	0,00340642	1,84906	1,84906	1 up vs 2 and 4
244164_at	Transcribed locus	---	0,00340659	0,85896 2	-1,1642	1 down vs 2 and 4
227654_at	chromosome 20 open reading frame 175	C20orf175	0,00342022	1,44801	1,44801	1 up vs 2 and 4
215291_at	megakaryoblastic leukemia (translocation) 1	MKL1	0,00342634	0,79764 8	-1,25369	1 down vs 2 and 4
228616_at	Polymerase (RNA)	POLRMT	0,00344134	0,85619	-1,16796	1 down vs 2

	mitochondrial (DNA directed)			5		and 4
242463_x_a †	zinc finger protein 600	ZNF600	0,00346225	1,582	1,582	1 up vs 2 and 4
1555042_at	transient receptor potential cation channel, subfamily V, member 5	TRPV5	0,0034641	0,82798 2	-1,20776	1 down vs 2 and 4
212119_at	ras homolog gene family, member Q	RHOQ	0,0034915	1,56196	1,56196	1 up vs 2 and 4
205456_at	CD3e molecule, epsilon (CD3-TCR complex)	CD3E	0,00349792	1,76237	1,76237	1 up vs 2 and 4
226412_at	splicing factor, arginine/serine-rich 18	SFRS18	0,00350796	1,60988	1,60988	1 up vs 2 and 4
229074_at	Transcribed locus	---	0,00350989	1,69606	1,69606	1 up vs 2 and 4
1553395_a_at	CD200 receptor 1	CD200R1	0,00351457	1,51025	1,51025	1 up vs 2 and 4
218501_at	Rho guanine nucleotide exchange factor (GEF) 3	ARHGEF3	0,00351555	2,09428	2,09428	1 up vs 2 and 4
209585_s_at	multiple inositol polyphosphate histidine phosphatase, 1	MINPP1	0,00352069	1,722	1,722	1 up vs 2 and 4
1558720_at	---	---	0,00355905	0,79688 9	-1,25488	1 down vs 2 and 4
233915_at	KIAA0825 protein	KIAA0825	0,00356893	0,83293 2	-1,20058	1 down vs 2 and 4
1554608_at	trans-golgi network protein 2	TGOLN2	0,00357206	1,52167	1,52167	1 up vs 2 and 4
208479_at	potassium voltage-gated channel, shaker-related subfamily, member 1 (episodic at	KCNA1	0,00358524	0,84871 8	-1,17825	1 down vs 2 and 4
1555809_at	cysteine-rich secretory protein LCCL domain containing 2	CRISPLD2	0,00360276	0,85645 4	-1,1676	1 down vs 2 and 4
202953_at	complement component 1, q subcomponent, B chain	C1QB	0,00360596	4,23875	4,23875	1 up vs 2 and 4
236401_at	Full-length cDNA clone CSODI016YI23 of	---	0,00361795	1,58151	1,58151	1 up vs 2 and 4

	Placenta Cot 25-normalized of Homo sapien					
210602_s_at	cadherin 6, type 2, K-cadherin (fetal kidney)	CDH6	0,0036375	0,850887	-1,17524	1 down vs 2 and 4
218591_s_at	nucleolar protein 10	NOL10	0,00366476	0,887471	-1,1268	1 down vs 2 and 4
1556661_at	Hypothetical protein LOC150384	LOC150384	0,0036667	0,868823	-1,15098	1 down vs 2 and 4
234090_at	CDNA FLJ14180 fis, clone NT2RP2003799	---	0,00367515	0,848652	-1,17834	1 down vs 2 and 4
232933_at	KIAA1656 protein	KIAA1656	0,00368013	0,854818	-1,16984	1 down vs 2 and 4
240583_at	RIO kinase 3 (yeast)	RIOK3	0,00368177	0,892861	-1,12	1 down vs 2 and 4
210354_at	interferon, gamma	IFNG	0,0036844	3,22022	3,22022	1 up vs 2 and 4
202747_s_at	integral membrane protein 2A	ITM2A	0,00369166	2,31485	2,31485	1 up vs 2 and 4
205804_s_at	TRAF3 interacting protein 3	TRAF3IP3	0,00369426	1,66987	1,66987	1 up vs 2 and 4
228361_at	E2F transcription factor 2	E2F2	0,00370584	0,690311	-1,44862	1 down vs 2 and 4
217922_at	mannosidase, alpha, class 1A, member 2	MAN1A2	0,00370587	1,41713	1,41713	1 up vs 2 and 4
1554095_at	RNA binding motif protein 33	RBM33	0,00370734	0,712651	-1,40321	1 down vs 2 and 4
206877_at	MAX dimerization protein 1	MXD1	0,00371249	0,856723	-1,16724	1 down vs 2 and 4
232001_at	hypothetical gene supported by AY007155	LOC439949	0,00374309	1,57474	1,57474	1 up vs 2 and 4
209234_at	kinesin family member 1B	KIF1B	0,00374358	1,45424	1,45424	1 up vs 2 and 4
221139_s_at	cysteine sulfinic acid decarboxylase	CSAD	0,00375763	1,42338	1,42338	1 up vs 2 and 4
244097_at	---	---	0,00376008	0,809787	-1,23489	1 down vs 2 and 4
224892_at	pallidin homolog (mouse)	PLDN	0,00378309	1,90436	1,90436	1 up vs 2 and 4
214559_at	dopamine receptor D3	DRD3	0,0037843	0,896375	-1,1156	1 down vs 2 and 4
212331_at	retinoblastoma-like 2 (p130)	RBL2	0,00379103	1,66159	1,66159	1 up vs 2 and 4
1558682_at	high mobility group AT-hook 2	HMGA2	0,00379445	0,886587	-1,12792	1 down vs 2 and 4

201941_at	carboxypeptidase D	CPD	0,00379669	1,64689	1,64689	1 up vs 2 and 4
211383_s_at	WD repeat domain 37	WDR37	0,00380807	1,56797	1,56797	1 up vs 2 and 4
202437_s_at	cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	0,00382063	2,57334	2,57334	1 up vs 2 and 4
209604_s_at	GATA binding protein 3	GATA3	0,0038224	2,18429	2,18429	1 up vs 2 and 4
223129_x_at	myosin regulatory light chain interacting protein	MYLIP	0,0038274	1,52727	1,52727	1 up vs 2 and 4
222536_s_at	zinc finger protein 395	ZNF395	0,00385664	0,730863	-1,36825	1 down vs 2 and 4
224374_s_at	elastin microfibril interfacer 2	EMILIN2	0,00387797	1,75066	1,75066	1 up vs 2 and 4
212914_at	chromobox homolog 7	CBX7	0,00388603	1,48561	1,48561	1 up vs 2 and 4
230655_at	Homo sapiens, clone IMAGE:5418468, mRNA	---	0,00388977	0,761764	-1,31274	1 down vs 2 and 4
1561626_at	hypothetical protein DJ1033E15.1	DJ1033E15.1	0,00390256	0,882825	-1,13273	1 down vs 2 and 4
212693_at	MDN1, midasin homolog (yeast)	MDN1	0,00391136	1,47703	1,47703	1 up vs 2 and 4
221816_s_at	PHD finger protein 11	PHF11	0,00392253	1,60238	1,60238	1 up vs 2 and 4
214532_x_at	POU class 5 homeobox 1 /// POU class 5 homeobox 1 pseudogene 1 /// POU class 5 h	POU5F1 /// POU5F1P1 /// POU5F1P3 /// POU5F1P4	0,00392373	0,792373	-1,26203	1 down vs 2 and 4
227483_s_at	Keratin associated protein 5-11	KRTAP5-11	0,00392809	0,832141	-1,20172	1 down vs 2 and 4
1552473_at	guanidinoacetate N-methyltransferase	GAMT	0,00392869	0,790709	-1,26469	1 down vs 2 and 4
203837_at	mitogen-activated protein kinase kinase 5 /// hypothetical protein LOC729	LOC729144 /// MAP3K5	0,00393141	1,7279	1,7279	1 up vs 2 and 4
235463_s_at	LAG1 homolog, ceramide synthase 6	LASS6	0,00394847	1,7222	1,7222	1 up vs 2 and 4
217776_at	retinol dehydrogenase 11 (all-trans/9-cis/11-cis)	RDH11	0,00395277	1,46809	1,46809	1 up vs 2 and 4
221758_at	armadillo repeat containing 6	ARMC6	0,00396948	0,794836	-1,25812	1 down vs 2 and 4
204646_at	dihydropyrimidine dehydrogenase	DPYD	0,00397657	2,08902	2,08902	1 up vs 2 and 4

228098_s_at	myosin regulatory light chain interacting protein	MYLIP	0,00399502	1,79505	1,79505	1 up vs 2 and 4
212111_at	syntaxin 12	STX12	0,00401184	1,56121	1,56121	1 up vs 2 and 4
1553115_at	naked cuticle homolog 1 (Drosophila)	NKD1	0,00402673	0,81618 2	-1,22522	1 down vs 2 and 4
227694_at	chromosome 1 open reading frame 201	C1orf201	0,0040274	0,78165 4	-1,27934	1 down vs 2 and 4
64418_at	CDNA FLJ34482 fis, clone HLUNG2004067	---	0,00403417	1,39891	1,39891	1 up vs 2 and 4
47069_at	proline rich 5 (renal)	PRR5	0,00403822	1,45059	1,45059	1 up vs 2 and 4
210746_s_at	erythrocyte membrane protein band 4.2	EPB42	0,00404887	0,85662 2	-1,16738	1 down vs 2 and 4
227010_at	CDNA clone IMAGE:4453251	---	0,0040509	0,75586 6	-1,32299	1 down vs 2 and 4
206974_at	chemokine (C-X-C motif) receptor 6	CXCR6	0,0040529	2,39314	2,39314	1 up vs 2 and 4
244163_at	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphor	SEMA3A	0,00405345	0,89836 9	-1,11313	1 down vs 2 and 4
236446_at	Transcribed locus	---	0,00406059	1,37429	1,37429	1 up vs 2 and 4
229006_at	CDNA FLJ36097 fis, clone TESTI2020956	---	0,00406127	0,77596 5	-1,28872	1 down vs 2 and 4
237228_at	Zinc finger, DHHC-type containing 1	ZDHHC1	0,00406154	0,82022	-1,21919	1 down vs 2 and 4
222603_at	endoplasmic reticulum metalloproteinase 1	ERMP1	0,00406513	1,71824	1,71824	1 up vs 2 and 4
1555832_s_at	Kruppel-like factor 6	KLF6	0,00406857	1,67312	1,67312	1 up vs 2 and 4
209285_s_at	chromosome 3 open reading frame 63	C3orf63	0,00407188	1,61075	1,61075	1 up vs 2 and 4
224240_s_at	chemokine (C-C motif) ligand 28	CCL28	0,00407397	0,73809 9	-1,35483	1 down vs 2 and 4
1554897_s_at	rhomboid, veinlet-like 2 (Drosophila)	RHBDL2	0,00408206	0,86161 7	-1,16061	1 down vs 2 and 4
234748_x_at	---	---	0,00408235	0,84187 2	-1,18783	1 down vs 2 and 4
203118_at	proprotein convertase subtilisin/kexin type 7	PCSK7	0,00412786	1,24972	1,24972	1 up vs 2 and 4
1555520_at	patched homolog 1 (Drosophila)	PTCH1	0,00414991	0,72607 7	-1,37726	1 down vs 2 and 4

239681_at	chromosome 17 open reading frame 56	C17orf56	0,00415252	0,83855 9	-1,19252	1 down vs 2 and 4
244082_at	Fission 1 (mitochondrial outer membrane) homolog (S. cerevisiae)	FIS1	0,00418878	0,84694 7	-1,18071	1 down vs 2 and 4
235360_at	CDNA clone IMAGE:30408657	---	0,00420838	1,41206	1,41206	1 up vs 2 and 4
1558922_at	Full length insert cDNA clone YF43G08	---	0,0042134	1,43638	1,43638	1 up vs 2 and 4
211175_at	G protein-coupled receptor 45	GPR45	0,00422825	0,81153 6	-1,23223	1 down vs 2 and 4
213136_at	protein tyrosine phosphatase, non-receptor type 2	PTPN2	0,00423353	1,44374	1,44374	1 up vs 2 and 4
1562342_at	MRNA; cDNA DKFZp667G1416 (from clone DKFZp667G1416)	---	0,0042367	0,90432 2	-1,1058	1 down vs 2 and 4
202939_at	zinc metalloproteinase (STE24 homolog, S. cerevisiae)	ZMPSTE24	0,00423895	1,58383	1,58383	1 up vs 2 and 4
216956_s_at	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	ITGA2B	0,00424079	0,78109 8	-1,28025	1 down vs 2 and 4
223583_at	tumor necrosis factor, alpha-induced protein 8-like 2	TNFAIP8L2	0,00424357	1,40783	1,40783	1 up vs 2 and 4
226876_at	family with sequence similarity 101, member B	FAM101B	0,00425357	1,72849	1,72849	1 up vs 2 and 4
1554922_at	zinc finger protein 678	ZNF678	0,00426794	0,86008 8	-1,16267	1 down vs 2 and 4
218735_s_at	zinc finger protein 544	ZNF544	0,00427583	1,52298	1,52298	1 up vs 2 and 4
222895_s_at	B-cell CLL/lymphoma 11B (zinc finger protein)	BCL11B	0,0042847	2,46033	2,46033	1 up vs 2 and 4
218341_at	phosphopantothienoyl cysteine synthetase	PPCS	0,00428515	1,46371	1,46371	1 up vs 2 and 4
1553179_at	ADAM metalloproteinase with thrombospondin type 1 motif, 19	ADAMTS19	0,00429233	0,82551 4	-1,21137	1 down vs 2 and 4
236089_at	Transcribed locus	---	0,00429628	0,82633 9	-1,21016	1 down vs 2 and 4

226181_at	tubulin, epsilon 1	TUBE1	0,00430082	1,42201	1,42201	1 up vs 2 and 4
218306_s_at	hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RCC1 (CHC1)-	HERC1	0,00430701	1,53668	1,53668	1 up vs 2 and 4
215005_at	EF-hand calcium binding protein 2	EFCBP2	0,00432222	0,82306 6	-1,21497	1 down vs 2 and 4
231607_at	Transcribed locus	---	0,00434079	0,79159 9	-1,26327	1 down vs 2 and 4
207681_at	chemokine (C-X-C motif) receptor 3	CXCR3	0,00434877	1,7821	1,7821	1 up vs 2 and 4
243938_x_at	dynein, axonemal, heavy chain 5	DNAH5	0,00438563	0,8841	-1,13109	1 down vs 2 and 4
1553690_at	shugoshin-like 1 (S. pombe)	SGOL1	0,00439608	0,73256 1	-1,36507	1 down vs 2 and 4
1555431_at	interleukin 31 receptor A	IL31RA	0,00439644	1,10673	1,10673	1 up vs 2 and 4
223682_s_at	hypothetical protein MGC11102	MGC11102	0,00441395	0,76773 9	-1,30253	1 down vs 2 and 4
236402_at	CDNA FLJ42263 fis, clone TKIDN2014570	---	0,0044204	0,70015 3	-1,42826	1 down vs 2 and 4
240746_s_at	---	---	0,00443051	0,86071 9	-1,16182	1 down vs 2 and 4
205258_at	inhibin, beta B	INHBB	0,00443726	0,76101 2	-1,31404	1 down vs 2 and 4
213849_s_at	protein phosphatase 2 (formerly 2A), regulatory subunit B, beta isoform	PPP2R2B	0,00444116	1,71635	1,71635	1 up vs 2 and 4
219320_at	myosin head domain containing 1	MYOHD1	0,00446143	0,79546 7	-1,25712	1 down vs 2 and 4
233356_at	BRCA2 region, mRNA sequence CG029	---	0,00446157	0,82464 7	-1,21264	1 down vs 2 and 4
1566285_at	PA4=candidate oncogene (3' region) [human, HEN-16, HEN-16T transformed endocervi	---	0,00447409	0,86169 2	-1,16051	1 down vs 2 and 4
210321_at	granzyme H (cathepsin G-like 2, protein h-CCPX)	GZMH	0,00448161	2,80469	2,80469	1 up vs 2 and 4
242967_at	---	---	0,00448788	0,89497 9	-1,11734	1 down vs 2 and 4
239177_at	CDNA FLJ38849 fis,	---	0,00450497	0,88258	-1,13304	1 down vs 2

	clone MESAN2008936			2		and 4
213344_s_at	H2A histone family, member X	H2AFX	0,00451013	0,73363 5	-1,36308	1 down vs 2 and 4
1552302_at	transmembrane protein 106A /// similar to transmembrane protein 106A	LOC728772 /// TMEM106A	0,00451154	1,34055	1,34055	1 up vs 2 and 4
213396_s_at	Full-length cDNA clone CS0DJ012YG05 of T cells (Jurkat cell line) Cot 10-normali	---	0,00451404	1,55128	1,55128	1 up vs 2 and 4
212112_s_at	syntaxin 12	STX12	0,00454697	1,6293	1,6293	1 up vs 2 and 4
202679_at	Niemann-Pick disease, type C1	NPC1	0,00454975	1,61018	1,61018	1 up vs 2 and 4
225509_at	hypothetical protein LOC56757	LOC56757	0,00456382	1,63243	1,63243	1 up vs 2 and 4
230875_s_at	ATPase, Class VI, type 11A	ATP11A	0,00459853	1,8241	1,8241	1 up vs 2 and 4
224986_s_at	3-phosphoinositide dependent protein kinase-1	PDPK1	0,00460243	1,40574	1,40574	1 up vs 2 and 4
211068_x_at	family with sequence similarity 21, member C /// family with sequence similarity	FAM21C /// FAM21D	0,00461615	1,57816	1,57816	1 up vs 2 and 4
212192_at	potassium channel tetramerisation domain containing 12	KCTD12	0,00462161	2,14144	2,14144	1 up vs 2 and 4
237286_at	Integrator complex subunit 9	RC74	0,00462237	0,78298 6	-1,27716	1 down vs 2 and 4
225492_at	---	---	0,00462672	1,54431	1,54431	1 up vs 2 and 4
240084_at	chromobox homolog 2 (Pc class homolog, Drosophila)	CBX2	0,00463961	0,82447 6	-1,21289	1 down vs 2 and 4
217546_at	metallothionein 1M	MT1M	0,00464455	3,61607	3,61607	1 up vs 2 and 4
211796_s_at	T cell receptor beta variable 19 /// T cell receptor beta variable 7-2 /// T cel	TRBC1 /// TRBV19 /// TRBV3-1 /// TRBV5-4 /// TRBV7-2	0,00470574	2,35893	2,35893	1 up vs 2 and 4
243179_at	CDNA FLJ33993 fis, clone DFNES2007757	---	0,0047205	1,32809	1,32809	1 up vs 2 and 4
221834_at	Lon peptidase 2, peroxisomal	LONP2	0,00472622	1,58043	1,58043	1 up vs 2 and 4
219996_at	ankyrin repeat and	ASB7	0,00472763	0,79403	-1,25939	1 down vs 2

	SOCS box-containing 7			7		and 4
221436_s_at	cell division cycle associated 3	CDCA3	0,00473772	0,68674 7	-1,45614	1 down vs 2 and 4
212616_at	chromodomain helicase DNA binding protein 9	CHD9	0,00473812	1,50848	1,50848	1 up vs 2 and 4
1554006_a_at	lethal giant larvae homolog 2 (Drosophila)	LLGL2	0,00477002	0,89217 1	-1,12086	1 down vs 2 and 4
1569274_at	Hypothetical protein LOC283551	LOC283551	0,00477546	0,83647 4	-1,19549	1 down vs 2 and 4
202193_at	LIM domain kinase 2	LIMK2	0,00478158	1,43537	1,43537	1 up vs 2 and 4
233612_at	CDNA FLJ12420 fis, clone MAMMA1003049	---	0,00480673	0,82035 4	-1,21899	1 down vs 2 and 4
241153_at	Transcribed locus, weakly similar to XP_001137328.1 hypothetical protein [Pan t	---	0,00482432	0,84392 5	-1,18494	1 down vs 2 and 4
210659_at	chemokine-like receptor 1	CMKLR1	0,00483767	1,57151	1,57151	1 up vs 2 and 4
1562103_at	Janus kinase 1 (a protein tyrosine kinase)	JAK1	0,00484467	0,79519 1	-1,25756	1 down vs 2 and 4
203741_s_at	adenylate cyclase 7	ADCY7	0,00484723	1,68534	1,68534	1 up vs 2 and 4
209607_x_at	sulfotransferase family, cytosolic, 1A, phenol- preferring, member 3 /// sulfotra	SULT1A3 /// SULT1A4	0,00486691	1,47725	1,47725	1 up vs 2 and 4
236901_at	Transcribed locus	---	0,0048723	2,00985	2,00985	1 up vs 2 and 4
235263_at	DKFZp434A0131 protein	DKFZP434A0131	0,00488956	0,85325 8	-1,17198	1 down vs 2 and 4
217580_x_at	Transcribed locus	---	0,00488966	0,72473 8	-1,37981	1 down vs 2 and 4
216233_at	CD163 molecule	CD163	0,00490806	1,59266	1,59266	1 up vs 2 and 4
212842_x_at	RANBP2-like and GRIP domain containing 5 /// RANBP2-like and GRIP domain contain	RGPD4 /// RGPD5 /// RGPD6 /// RGPD8	0,00491406	1,39474	1,39474	1 up vs 2 and 4
201446_s_at	TIA1 cytotoxic granule- associated RNA binding protein	TIA1	0,0049141	1,56517	1,56517	1 up vs 2 and 4
244514_at	Transcribed locus	---	0,0049207	0,83335 9	-1,19996	1 down vs 2 and 4

203836_s_at	mitogen-activated protein kinase kinase 5	MAP3K5	0,00492661	1,73606	1,73606	1 up vs 2 and 4
231656_x_at	oxysterol binding protein-like 10	OSBPL10	0,00493192	0,617539	-1,61933	1 down vs 2 and 4
1554697_at	ADAM metalloproteinase with thrombospondin type 1 motif, 9	ADAMTS9	0,00493877	0,845042	-1,18337	1 down vs 2 and 4
216934_at	---	---	0,00494934	0,853586	-1,17153	1 down vs 2 and 4
205563_at	KISS-1 metastasis-suppressor	KISS1	0,00495982	0,830535	-1,20404	1 down vs 2 and 4
209515_s_at	RAB27A, member RAS oncogene family	RAB27A	0,00498443	1,84326	1,84326	1 up vs 2 and 4
1569996_at	CDNA clone IMAGE:5265638	---	0,00498995	0,787255	-1,27024	1 down vs 2 and 4
202378_s_at	leptin receptor overlapping transcript	LEPROT	0,00499672	1,59991	1,59991	1 up vs 2 and 4
221260_s_at	family with sequence similarity 130, member A1	FAM130A1	0,00499929	1,40437	1,40437	1 up vs 2 and 4
Probeset ID	Gene Title	Gene Symbol	p-value(1 vs. 2 and 4)	Ratio(1 vs. 2 and 4)	Fold-Change (1 vs. 2 and 4)	Fold-Change(1 vs. 2 and 4) (Description)
239699_s_at	postmeiotic segregation increased 2-like 1	PMS2L1	8,97E-06	0,679272	-1,47216	1 down vs 2 and 4
1563469_at	MRNA; cDNA DKFZp313M0417 (from clone DKFZp313M0417)	---	4,70E-05	1,44602	1,44602	1 up vs 2 and 4

SUPPLEMENTARY TABLE 2

A SUPERVISED ANALYSIS OF GENE EXPRESSION PROFILES OF CASES CLASSIFIED AS CLUSTER 1 VERSUS CASES BELONGING TO CLUSTERS 2 AND 4 IDENTIFIED 546 PROBES. THE UP-REGULATED TRANSCRIPTS WERE 323, WHILE THE DOWN-REGULATED WERE 223.

Probeset ID	Gene Title	Gene Symbol	p-value(8p deletion)	p-value(8p loss vs. wt)	Ratio(8p loss vs. wt)	Fold-Change(8p loss vs. wt)
206779_s_at	acetylserotonin O-methyltransferase	ASMT	9,78E-06	9,78E-06	6,29437	6,29437
210551_s_at	acetylserotonin O-methyltransferase	ASMT	1,31E-05	1,31E-05	5,30017	5,30017
235428_at	Transcribed locus, strongly similar to XP_001095617.1 syntaxin binding protein	---	2,03E-05	2,03E-05	2,07827	2,07827
221261_x_at	melanoma antigen family D, 4B /// melanoma antigen family D, 4	MAGED4 /// MAGED4B	4,70E-05	4,70E-05	2,11725	2,11725
206574_s_at	protein tyrosine phosphatase type IVA, member 3	PTP4A3	5,97E-05	5,97E-05	2,24274	2,24274
230789_at	suppressor of hairy wing homolog 2 (Drosophila)	SUHW2	9,27E-05	9,27E-05	1,93841	1,93841
219740_at	vasohibin 2	VASH2	0,000132126	0,000132126	2,09433	2,09433
205139_s_at	uronyl-2-sulfotransferase	UST	0,000178207	0,000178207	1,92311	1,92311
224407_s_at	serine/threonine protein kinase MST4	RP6-213H19.1	0,000192566	0,000192566	2,30513	2,30513
214051_at	hypothetical protein MGC39900	MGC39900	0,000201561	0,000201561	3,00384	3,00384
219360_s_at	transient receptor potential cation channel, subfamily M, member 4	TRPM4	0,000201761	0,000201761	4,49634	4,49634

211080_s_at	NIMA (never in mitosis gene a)-related kinase 2	NEK2	0,000298917	0,000298917	1,92007	1,92007
203562_at	fasciculation and elongation protein zeta 1 (zygin I)	FEZ1	0,000331718	0,000331718	6,21268	6,21268
205197_s_at	ATPase, Cu ⁺⁺ transporting, alpha polypeptide (Menkes syndrome)	ATP7A	0,000418277	0,000418277	1,65303	1,65303
1557217_a_at	Fanconi anemia, complementation group B	FANCB	0,000425087	0,000425087	1,73069	1,73069
202261_at	vacuolar protein sorting 72 homolog (S. cerevisiae)	VPS72	0,000536022	0,000536022	1,87421	1,87421
206502_s_at	insulinoma-associated 1	INSM1	0,000571437	0,000571437	1,98663	1,98663
212570_at	endonuclease domain containing 1	ENDOD1	0,00060069	0,00060069	1,68072	1,68072
218905_at	integrator complex subunit 8	INTS8	0,00072932	0,00072932	1,62675	1,62675
209470_s_at	glycoprotein M6A	GPM6A	0,000731237	0,000731237	5,24851	5,24851
204603_at	exonuclease 1	EXO1	0,00075341	0,00075341	2,09081	2,09081
212541_at	FAD1 flavin adenine dinucleotide synthetase homolog (S. cerevisiae)	FLAD1	0,000762095	0,000762095	1,66104	1,66104
230758_at	gem (nuclear organelle) associated protein 8	GEMIN8	0,000767248	0,000767248	1,63643	1,63643
235343_at	---	---	0,000807773	0,000807773	2,37095	2,37095
238493_at	zinc finger protein 506	ZNF506	0,000814466	0,000814466	1,76003	1,76003
228565_at	mixed lineage kinase 4	KIAA1804	0,000889823	0,000889823	2,22687	2,22687
210313_at	leukocyte immunoglobulin-	LILRA4	0,000891328	0,000891328	3,53436	3,53436

	like receptor, subfamily A (with TM domain), member 4					
209695_at	protein tyrosine phosphatase type IVA, member 3	PTP4A3	0,000939893	0,000939893	2,68607	2,68607
1564691_at	CDNA: FLJ22859 fis, clone KAT01752	---	0,00100105	0,00100105	1,62894	1,62894
1557177_at	Homo sapiens, clone IMAGE:4340526, mRNA	---	0,00110608	0,00110608	1,99356	1,99356
232596_at	diaphanous homolog 3 (Drosophila)	DIAPH3	0,00122161	0,00122161	1,63581	1,63581
229360_at	suppressor of hairy wing homolog 2 (Drosophila)	SUHW2	0,0013299	0,0013299	1,73355	1,73355
228281_at	chromosome 11 open reading frame 82	C11orf82	0,00136162	0,00136162	1,9402	1,9402
210481_s_at	C-type lectin domain family 4, member M	CLEC4M	0,00141179	0,00141179	1,74176	1,74176
216997_x_at	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)	TLE4	0,00142457	0,00142457	1,68609	1,68609
218951_s_at	phosphatidylinositol-specific phospholipase C, X domain containing 1	PLCXD1	0,00147243	0,00147243	1,96918	1,96918
223985_at	fibronectin type III and SPRY domain containing 1-like	FSD1L	0,00148479	0,00148479	1,62146	1,62146
207409_at	leukocyte cell-derived chemotaxin 2	LECT2	0,00162425	0,00162425	1,64402	1,64402
211761_s_at	calcyclin binding protein	CACYBP	0,00163005	0,00163005	1,8142	1,8142
202323_s_at	acyl-Coenzyme A binding domain containing 3	ACBD3	0,00168734	0,00168734	2,17189	2,17189

236769_at	hypothetical protein LOC158402	LOC158402	0,0017369	0,0017369	1,97264	1,97264
37004_at	surfactant, pulmonary-associated protein B	SFTPB	0,00175405	0,00175405	3,07024	3,07024
1552695_at	solute carrier family 2 (facilitated glucose transporter), member 13	SLC2A13	0,00185842	0,00185842	2,04764	2,04764
225532_at	Cdk5 and Abl enzyme substrate 1	CABLES1	0,00191328	0,00191328	1,85186	1,85186
204899_s_at	Sin3A-associated protein, 30kDa	SAP30	0,00192232	0,00192232	1,56655	1,56655
209810_at	surfactant, pulmonary-associated protein B	SFTPB	0,00196128	0,00196128	3,11077	3,11077
228483_s_at	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa	TAF9B	0,00210114	0,00210114	1,66836	1,66836
226397_s_at	Transcribed locus	---	0,00221572	0,00221572	2,25588	2,25588
230893_at	DnaJ homology subfamily A member 5	DNAJA5	0,00221711	0,00221711	1,56245	1,56245
219510_at	polymerase (DNA directed), theta	POLQ	0,00227834	0,00227834	1,63084	1,63084
206498_at	oculocutaneous albinism II (pink-eye dilution homolog, mouse)	OCA2	0,00238068	0,00238068	1,64658	1,64658
218546_at	chromosome 1 open reading frame 115	C1orf115	0,00241741	0,00241741	2,37525	2,37525
223313_s_at	melanoma antigen family D, 4B /// melanoma antigen family D, 4	MAGED4 /// MAGED4B	0,00247836	0,00247836	1,54261	1,54261
235331_x_at	polycomb group ring finger 5	PCGF5	0,00249803	0,00249803	1,8546	1,8546
202322_s_at	geranylgeranyl	GGPS1	0,00250471	0,00250471	1,65136	1,65136

	diphosphate synthase 1			471		
231972_at	CDNA: FLJ21028 fis, clone CAE07155	---	0,00278625	0,00278625	1,6	1,6
220991_s_at	ring finger protein 32	RNF32	0,00291608	0,00291608	1,6235	1,6235
217640_x_at	chromosome 18 open reading frame 24	C18orf24	0,00305327	0,00305327	1,91711	1,91711
1555772_a_at	cell division cycle 25 homolog A (S. pombe)	CDC25A	0,00307714	0,00307714	1,97603	1,97603
1559067_a_at	Transcribed locus	---	0,00329519	0,00329519	2,24909	2,24909
203532_x_at	cullin 5	CUL5	0,00330657	0,00330657	1,62845	1,62845
227687_at	hydroletharus syndrome 1	HYLS1	0,00331712	0,00331712	2,0412	2,0412
239229_at	phosphate regulating endopeptidase homolog, X-linked (hypophosphatemia, vitamin	PHEX	0,00333513	0,00333513	2,3994	2,3994
228256_s_at	erythrocyte membrane protein band 4.1 like 4A	EPB41L4A	0,00349661	0,00349661	2,18769	2,18769
1552694_at	solute carrier family 2 (facilitated glucose transporter), member 13	SLC2A13	0,0040178	0,0040178	1,64508	1,64508
201926_s_at	CD55 molecule, decay accelerating factor for complement (Cromer blood group)	CD55	0,00404787	0,00404787	1,96567	1,96567
222746_s_at	B-box and SPRY domain containing	BSPRY	0,00406031	0,00406031	2,07249	2,07249
233750_s_at	chromosome 1 open reading frame 25	C1orf25	0,00409832	0,00409832	1,57398	1,57398
226147_s_at	polymeric immunoglobulin	PIGR	0,00412046	0,00412046	4,40004	4,40004

	receptor					
209469_at	glycoprotein M6A	GPM6A	0,00418498	0,00418498	3,39637	3,39637
1565951_s_at	choroideremia-like (Rab escort protein 2)	CHML	0,00420623	0,00420623	1,6873	1,6873
211502_s_at	PFTAIRE protein kinase 1	PFTK1	0,00420855	0,00420855	1,82215	1,82215
201644_at	tissue specific transplantation antigen P35B	TSTA3	0,00421329	0,00421329	1,5238	1,5238
243798_at	Transcribed locus	---	0,00428561	0,00428561	2,26873	2,26873
218741_at	centromere protein M	CENPM	0,00442278	0,00442278	1,77951	1,77951
220223_at	ATPase family, AAA domain containing 5	ATAD5	0,00446521	0,00446521	1,46953	1,46953
205987_at	CD1c molecule	CD1C	0,00446956	0,00446956	3,10145	3,10145
237563_s_at	hypothetical LOC440731	LOC440731	0,00462397	0,00462397	1,89062	1,89062
210674_s_at	protocadherin alpha 9 /// protocadherin alpha subfamily C, 2 /// protocadherin a	PCDHA1 /// PCDHA10 /// PCDHA11 /// PCDHA12 /// PCDHA13 /// PCDHA2 /// PCDHA3 ///	0,00488334	0,00488334	1,79313	1,79313
221706_s_at	uncharacterized hematopoietic stem/progenitor cells protein MDS032	MDS032	0,00489514	0,00489514	1,64152	1,64152
226615_at	xenotropic and polytropic retrovirus receptor	XPR1	0,00497165	0,00497165	1,59175	1,59175

10. LIST OF PUBLICATIONS

A) PUBLICATIONS

1. Chigrinova E, Mian M, **Scandurra M**, Greiner TC, Chan WC, Vose JM, Inghirami G, Chiappella A, Baldini S, Ponzoni M, Ferreris AJM, Franceschetti S, Gaidano G, Tuccis A, Facchetti S, Lazure T, Lambotte O, Montes-Moreno S, Piris MA, Nomdedeu JFr, Uccella S, Rancoita PMV, Kwee I, Zucca E, Bertoni F. **Diffuse Large B-cells lymphoma with concordant bone marrow involvement has peculiar genomic profile and poor clinical outcome.** *Hematological Oncology* 2010; in press
2. **Scandurra M**, Mian M, Greiner TC, Rancoita PMV, De Campos CP, Chan WC, Vose JM, Inghirami G, Chiappella A, Baldini L, Ponzoni M, Ferreri AJM, Franceschetti S, Gaidano G, Montes-Moreno S, Piris MA, Facchetti F, Tucci A, Nomdedeu J Fr, Lazure T, Lambotte O, Uccella S, Pinotti G, Pruneri G, Martinelli G, Tibiletti M G, Chigrinova E, Rinaldi A, Zucca E, Kwee I, Bertoni F. **Genomic lesions associated with a different clinical outcome in diffuse large B-cell lymphoma treated with R-CHOP-21** *British Journal of Haematology* 2010; in press.
3. Rinaldi A, Capello D, **Scandurra M**, Greiner TC, Chan WC, Bhagat G, Rossi D, Morra E, Paulli M, Rambaldi A, Rancoita PMV, Inghirami G, Ponzoni M, Moreno SM, Piris MA, Mian M, Chigrinova E, Zucca E, Dalla Favera R, Gaidano G, Kwee I, & Bertoni F. **SNP-arrays provide new insights in the pathogenesis of post-transplant diffuse large B-cell lymphoma.** *British Journal of Haematology* 2010; in press.
4. Capello D, **Scandurra M**, Poretti G, Rancoita PMV, Mian M, Gloghini A, Deambrogi C, Martini M, Rossi D, Greiner TC, Chan WC, Ponzoni M, Montes Moreno S, Piris MA, Canzonieri V, Spina M, Tirelli U, Inghirami G, Rinaldi A, Zucca E, Dalla Favera R, Cavalli F, Larocca LM, Kwee I, Carbone A, Gaidano G, & Bertoni F. **Genome wide DNA-profiling of HIV-related B-cell lymphomas.** *British Journal of Haematology* 2010; in press.
5. **Scandurra M**, Rossi D, Deambrogi C, Rancoita PMV, Chigrinova E, Mian M, Cerri M, Rasi S, Sozzi E, Forconi F, Ponzoni M, Montes-Moreno S, Piris MA, Inghirami G, Zucca E, Gattei V, Rinaldi A, Kwee I, Gaidano G, & Bertoni F. **Genomic profiling of Richter's syndrome: recurrent lesions and differences with de novo diffuse large B-cell lymphomas.** *Hematological Oncology* 2010.
6. Compagno M, Lim WK, Grunn A, Nandula SV, Bertoni F, Ponzoni M, **Scandurra M**, Califano A, Bhagat G, Chadburn A, Dalla Favera R, Pasqualucci L. **Mutations at multiple genes cause deregulation of the NF-kB pathway in diffuse large B-cell lymphoma.** *Nature* 2009;459(7247):717-21.
7. Tibiletti MG, Martin V, Bernasconi B, Del Curto B, Pecciarini L, Uccella S, Pruneri G, Ponzoni M, Mazzucchelli L, Martinelli G, Ferreri A, Pinotti G, Assannelli A, **Scandurra M**, Doglioni C, Zucca E, Capella C, & Bertoni F. **BCL2, BCL6, MYC, MALT1 and BCL10 rearrangements in nodal diffuse large B-cell lymphomas: a multicenter evaluation of a new set of FISH probes and correlation with clinical outcome.** *Human Pathology* 2009; 40 (5): 645-652.
8. Forconi F, Rinaldi A, Kwee I, Sozzi E, Raspadori D, Rancoita PMV, **Scandurra M**, Rossi D, Deambrogi C, Capello D, Zucca E, Marconi D, Bomben R, Gattei V, Lauria F, Gaidano G & Bertoni F. **Genome-wide DNA analysis identifies recurrent imbalances predicting outcome in chronic lymphocytic leukaemia with 17p deletion.** *British Journal Haematology* 2008; 143(4): 532-536.

B) ABSTRACTS

1. Kwee I, De Campos CP, Mian M, Rancoita PMV, Rinaldi A, Boi M, **Scandurra M**, Greiner TC, Chan WC, Rossi D, Gaidano G, Carbone A, Bhagat G, Arcaini L, Baldini L, Ponzoni M, Ighirami G, Piris MA, Dalla Favera R, Zucca E, Cavalli F, & Berton F. **High-resolution genomic profiling of 533 B-cell lymphomas defines distinct tumor signatures, genomic aberrations correlated with outcome and pathogenetic subgroups. Proceedings of the AACR conference "Advances in Cancer Research: From the laboratory to the clinic" 2010: 82 (abst. B59).**
2. Rinaldi A, Capello D, **Scandurra M**, Greiner TC, Chan WC, Rossi D, Bhagat G, Paulli M, Inghirami G, Ponzoni M, Montes Moreno S, Piris MA, Mian M, Rancoita PMV, Zucca E, Dalla Favera R, Gaidano G, Kwee I, & Berton F. **SNP-Arrays Provide New Insights Into the Pathogenesis of Post-Transplant Diffuse Large B-Cell Lymphoma (PT-DLBCL). Blood (ASH Annual Meeting Abstracts), 2009: 114 (22): 185 (oral session).**
3. **Scandurra M**, Mian M, De Campos CP, Rancoita PMV, Greiner TC, Chan WC, Vose JM, M.D. Inghirami G, Chiappella A, Baldini L, Ponzoni M, Ferreri AJM, Montes Moreno S, Piris MA, Franceschetti S, Gaidano G, Facchetti F, Tucci A, Lazure T, Lambotte O, Uccella S, Tibiletti MG, Pinotti G, Pruner G, Martinelli G, Nomdedeu JF, Chigrinova E, Rinaldi A, Zucca E, Kwee I, & Berton F. **Subgroups of Diffuse Large B-Cell Lymphoma (DLBCL) with Different Genomic Lesions and Clinical Course During Treatment with R-CHOP. Blood (ASH Annual Meeting Abstracts), 2009: 114 (22): 1521 (poster).**
4. **Scandurra M**, Rancoita PMV, Greiner TC, Chan AC, Vose JM, Inghirami G, Chiappella A, Baldini L, Ponzoni M, Ferreri AJM, Montes-Moreno S, Piris MA, Franceschetti S, Gaidano G, Facchetti F, Tucci A, Lazure T, Lambotte O, Uccella S, Pinotti G, Pruner G, Martinelli G, Nomdedeu JF, Tibiletti MG, Chigrinova E, De Campos C, Rinaldi A, Zucca E, Kwee I, & Berton F. **Array-CGH identifies regions, including the FOXP1 locus, associated with different clinical outcome in diffuse large B-cell lymphomas (DLBCL) treated with R-CHOP. Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 478 (oral session).**
5. Pasqualucci L, Compagno M, Lim WK, Grunn A, Nandula SV, **Scandurra M**, Berton F, Ponzoni M, Califano A, Bhagat G, Chadburn A, and Dalla-Favera R. **Mutations in Multiple Genes Cause Deregulation of the NFkB Pathway in Diffuse Large B-Cell Lymphoma. Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 801 (oral session).**
6. Berton F, **Scandurra M**, Cerri M, Deambrogi C, Rancoita PMV, Rasi S, Forconi F, Ponzoni M, Montes-Moreno S, Piris M, Inghirami G, Rinaldi A, Zucca E, Kwee I, Gaidano G, & Rossi D. **Genome wide-DNA profiling of Richter's syndrome-diffuse large B-cell lymphoma (RS-DLBCL): differences with de novo DLBCL and possible mechanisms of transformation from chronic lymphocytic leukemia (CLL). Blood (ASH Annual Meeting Abstracts), Nov 2008; 112: 2067 (poster session).**
7. **Scandurra M**, Kwee I, Ponzoni M, Rancoita PMV, Greiner TC, Chan AC, Vose JM, Inghirami G, Chiappella A, Montes-Moreno S, Piris M, Franceschetti S, Gaidano G, Ferreri AJM, Baldini L, Facchetti F, Rossi G, Canzonieri V, Spina M, Gloghini A, Carbone A, Pruner G, Martinelli G, Poretti G, Rinaldi A, Tibiletti MG, Carbone G, Zucca E, & Berton F. **Array-CGH in diffuse large B-cell lymphomas (DLBCL) treated with R-CHOP identifies recurrent 11q24.3 aberrations. Annals of Oncology 2008; 19: iv103 (184).**
8. Capello D, Poretti G, **Scandurra M**, Kwee I, Rancoita PMV, Deambrogi C, Gloghini A, Larocca LM, Rinaldi A, Rossi D, Ponzoni M, Montes Moreno S, Zucca E, Piris MA, Spina M, Tirelli U, Carbone A, Gaidano G, & Berton F. **Genome-wide array-based comparative genomic hybridization of HIV-related non-Hodgkin lymphoma: identification of recurrent genetic lesions specifically associated with the disease. Haematologica 2008; EHA 2008: in press (oral)(0896)**