Genomic profiling of Anaplastic Large Cell Lymphoma

Boi, Michela

UNIVERSITÉ DE GENÈVE **FACULTÉ DES SCIENCES**

Section des Sciences Pharmaceutiques ***Professeur Leonardo Scapozza***

INSTITUTE OF ONCOLOGY RESEARCH (IOR) ***Genomic and Lymphoma Group***

Francesco Bertoni, MD

Genomic Profiling of Anaplastic Large Cell Lymphoma

THÈSE

presentée à la Faculté des Sciences de l'Université de Genève pour obtenir le grade de Docteur ès Science, mention interdisciplinaire

par

MICHELA BOI

de
Grugliasco
TORINO (Italie)

Thèse n° 4586

TORINO
Centro Copie "I Duplicanti"
2013

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention interdisciplinaire

Thèse de *Madame Michela BOI*

intitulée :

" Genomic Profiling of Anaplastic Large 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, Bellinzona, Switzerland), G. HOPFGARTNER, professeur ordinaire (Section des sciences pharmaceutiques), G. INGHIRAMI, professeur (Department of Molecular Biotechnology and Health Science, Department of Pathology, Center for Experimental Research and Medical Studies, University of Turin, Italy), D. ROSSI, docteur (Department of Clinical and Experimental Medicine « Amedeo Avogadro », University of Eastern Piedmont, Novara, Italy), et Madame THOME-MIAZZA, professeur (Département de biochimie, Université de Lausanne, Suisse) autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 14 août 2013

Thèse - 4586 -

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".

"If we knew what we are doing, it wouldn't be research..."
Anonymous

Summary

In the Western Countries, Anaplastic Large Cell Lymphoma (ALCL) is one of the most common form of T-cell Non-Hodgkin Lymphoma (NHL). In the last 2008 World Health Organization (WHO) classification, ALCL has been separated in two distinct entities based on the presence or absence of translocations affecting the Anaplastic Lymphoma Kinase (*ALK*) gene mapping on chromosome 2. Nevertheless, ~15–40% of ALCL do not express ALK or other recurrent translocations. Although the clinical presentations, translocations and genetic events vary between ALK+ and ALK- ALCL and the relationship between these two subtypes remains unclear. In this regard, the WHO Classification classifies ALK- ALCL only as a provisional entity.

With the aim of identifying the genetic events underlying the pathogenesis of ALCL, and to demonstrate that ALK+ and ALK- subsets can definitively be considered different entities, we studied a series of 69 cases of ALCL (31 ALK-, 33 ALK+, 5 cutaneous ALCL) with high-density genome wide SNP-based arrays using GeneChip Human Mapping SNP 6.0 array (Affymetrix, Santa Clara, CA, USA). Due to the scarcity of cALCL samples, we focused our attention on the analysis of the ALK+ and ALK- ALCL subsets. The detection of aberrations and the comparison of lesions between the two main ALCL groups permitted the identification of those genes that are likely to be involved in ALCL pathogenesis.

Genomic profiles of the ALK+ group show very few aberrations, while ALK- ALCL samples have a more perturbated and complex profile. Among all the ALCLs, the most common losses affected 17p13.3-p12 (25%), in which *TP53* is located, and 6q21 (19%), region that encompass *PRDM1* and *ATG5*. Other common losses were identified at level of 13q32.3-q33.3 and 16q23.2 (16%). Gains of different regions of the long arm of chromosome 1 (1q, 1q25.1,1q32.2) were observed in more than 20% of patients, while chromosome 7 and 8q24.22 were gained in 14% and 17% of cases, respectively. The analysis of the minimal common regions (MCRs) in the ALK- patients revealed that the most common aberrations were the same identified in the whole ALCL series, but at a higher frequency. Losses of 17p13.3-p12 and 6q21 were observed in 42% and 35% of the cases respectively, and losses of distal regions of chromosome 13, 12q and 10 were identified in more than 20% of the cases. Gains at level of 1q were observed in almost

30% of ALK- cases. The ALK+ subset presented less lesions or at a lower frequency than what seen in ALK- cases: only gains of 1q, 1q32.2, 7q32.3 and 7p22.3-p21.3 were observed in more than 15% of samples.

The deletions at 6q21 targeted the gene *PRDM1*, coding for BLIMP1. The whole coding sequence of *PRDM1* has been sequenced in order to understand if other mechanisms, different from deletions, could be responsible for the inactivation of the gene. Two mutations were detected in ALK- patients, in which the nucleotide change determined the formation of a stop codon. One of the patients showed a copy neutral loss of heterozygosity (cnLOH), while the other displayed the loss of the wild type allele, suggesting a complete inactivation of the protein. Another case was observed with complete gene loss: one ALK+ case with homozygous deletion. In summary, *PRDM*1 was inactivated, by deletion or by somatic mutations, in 12/31 (39%) and in 1/33 (3%) of clinical specimens derived from ALK- and ALK+ ALCL respectively, in 4/7 (57%) ALCL cell lines, and in 1/5 (20%) cALCL. The detection of the same *PRDM1* loss observed in patients in ALCL cell lines, allowed us to perform gain-of-function experiments in order to unravel the function of BLIMP1 loss in ALCL pathogenesis. After the re-expression of *PRDM1*, ALCL cells underwent proliferation arrest, with a concomitant moderate increase in the percentage of apoptotic cells and an arrest in cell cycle, suggesting a tumor suppressor role for BLIMP1 in our *in vitro* ALCL model. All the results obtained with gain-of-function experiments *in vitro* have been confirmed by *in vivo* experiments.

In order to have a global overview of what are the transcriptional changes happening after BLIMP1 re-expression, we performed a Gene Expression Profiling (GEP) analysis. From the comparison among GEP obtained from cells infected for BLIMP1 re-expression and GEP obtained from cells infected with an empty vector or not infected at all, we could discriminate genes that were affected by BLIMP1 re-expression. Combining GEP data with previous BLIMP1 functional studies, it seems that BLIMP1 might exerts a pro-apoptotic effects if expressed in ALCL cells, acting on a series of genes that are involved in pathways regulating cell proliferation and survival.

As conclusion, the analysis of genomic profiles revealed that ALK- ALCL presents unique features when compared with ALK+ ALCL, despite the observation of some common lesions. This suggests that different pathogenetic events are required for the development of both subtypes. Moreover, the frequent inactivation of BLIMP1 in ALCL clinical specimens suggested a tumor suppressor role for this gene in both *in vivo* and *in vitro*

models for ALCL since its re-expression causes reduction in cell growth, increased cell death and cell cycle arrest.

Résumé

Le Lymphome Anaplastique à Grandes Cellules (Anaplastic Large Cell Lymphoma, ALCL) est l'un des plus commun NHL des cellules de type T présent en Occident. La classification de 2008 de la World Health Organization (WHO), a séparé ce lymphome en deux formes distinctes, basées sur la présence ou l'absence de translocations touchant le gène Kinase Lymphome Anaplasique (Anaplasic Lymphoma Kinase, ALK) présent sur le chromosome 2. Cependant, 15-40% du lymphome ALCL n'exprime ni le gène ALK ni d'autres translocations fréquentes. Malgré les représentations cliniques, les translocations et les événements génétiques varient entre ALK+ et ALK- . La relation entre ces deux sous-types n'est pas encore claire. À ce propos, la classification de la WHO indique ALK- ALCL comme un sous-groupe provisoire. En ayant pour but l'identification d'événements génétiques qui sont à la base de la pathogénie du lymphome ALCL, et pour démontrer que les sous-groupes ALK+ et ALK- peuvent être considérés comme deux entités distinctes, nous avons étudié une série de 69 cas de lymphome ALCL (31 ALK-, 33 ALK+, 5 cALCL) par moyen de array high-density genome wide SNP-based en usant l'array GeneChip Human Mapping SNP 6.0 (Affymetrix, Santa Clara, CA, USA). La détection d'aberrations et la comparaison des lésions entre les deux groupes de lymphome ALCL, ont permis l'identification de gènes qui sont potentiellement impliqués dans sa pathogénie.

Les profiles génomiques du groupe ALK+ comprennent très peu d'aberrations; au contraire, les échantillons ALK- ALCL présentent plus de lésions et ont donc un profile plus complexe et moins défini. Parmi les lymphomes ALCL, les pertes les plus communes sont à charge de 17p13.3-p12 (25%), où se trouve *TP53*, et de 6q21 (19%), région qui comprend *PRDM1* et *ATG5*. D'autres pertes communes ont été identifiées au niveau de 13q32.3-q33.3 et 16q23.2 (16%). Des acquisitions de différentes régions du bras long du chromosome 1 (1q, 1q25.1, 1q32.2) ont étés observées dans plus de 20% des patients, tandis que les chromosomes 7 et 8q24.22 ont étés trouvés respectivement dans le 14 et le 17% des cas. L'analyse des Minimal Common Regions (MCRs) des patients ALK- a mis en évidence que les aberrations les plus communes sont identiques à celles identifiées pour le lymphome ALCL, même si elles sont

présentes à une fréquence plus élevée. Des pertes de 17p13.3-p12 et 6q21 ont été observées respectivement dans le 42% et le 35% des cas et, en plus, pour un pourcentage de cas supérieur à 20, d'autres pertes ont été identifiées pour la région distale du chromosome 13, 12q et 10. Pour ce qui concerne les cas de lymphome ALK-, il y a des acquisitions de 1q dans au moins 30% des cas. Le sous-groupe ALK+ présente moins de lésions, ou des lésions à une fréquence plus basse que celles identifiées pour les cas ALK- : des acquisitions de 1q, 1q32.2, 7q32.3 et 7p22.3-p21.3 ont été observées dans plus de 15% des cas.

Les délétions au niveau de 6q21 comprennent le gène PRDM1, qui code pour BLIMP1. La région codante de PRDM1 a été séquencéé pour comprendre si d'autres mécanismes, délétions à part, peuvent être à l'origine de l'inactivation du gène. Deux mutations somatiques, donnant lieu à un codon de stop, ont été identifiées. L'une comporte un copy neutral loss of heterozygosity (cnLOH) et couvre le locus *PRDM1*; la deuxième implique la perte de l'autre allèle, ce qui suggère la présence d'une protéine tronquée. La perte complète du gène a été observée dans un cas ALK+ comportant une délétion homozygote. En résumé, *PRDM1* est inactivé à cause de délétions ou mutations somatiques, en 12 cas sur 31 (ce qui correspond à 39%) et en 1 cas sur 33 (3%) des échantillons cliniques provenant respectivement, de patients touchés par le lymphome ALCL ALK- et ALK+ ; en 4 cas sur 7 (57%) pour ce qui concerne les lignées cellulaires de lymphome ALCL et en 1 cas sur 5 (20%) pour cALCL. L'identification, dans les lignées cellulaires, des mêmes pertes à charge de *PRDM1* observées pour les patients, nous a permis de conduire des expériences de gain-of-function pour comprendre comment la perte de BLIMP1 est impliquée dans la pathogénie du lymphome ALCL. Suite à la reconstitution du gène *PRDM1*, le cellules du lymphome ALCL subissent un arrêt de la prolifération accompagné d'une légère augmentation du pourcentage de cellules qui sont en apoptose, de même qu'un arrêt du cycle cellulaire. Cela suggère un rôle de suppresseur de tumeur, dans notre model *in vitro* de lymphome ALCL, pour ce qui concerne BLIMP1. Tous les résultats obtenus *in vitro* ont étés confirmés par ceux effectués *in vivo*.

Pour comprendre l'ensemble général des changements au niveau de la transcription, qui ont lieu suite à la reconstitution de BLIMP1, nous avons conduit une analyse de GEP. En comparant les résultats obtenus à partir de cellules infectées pour qu'elles réexpriment BLIMP1 et ceux obtenus à partir de cellules infectées avec le vecteur seul,

ou pas infectées du tout, nous avons pu mettre en évidence quels gènes sont touchés par la réexpression de BLIMP1. De résultats obtenus précédemment par moyen d'études fonctionnelles sur BLIMP1, combinés avec ceux de GEP, semblent démontrer que BLIMP1 est capable d'exercer un effet pro-apoptotique en agissant sur une série de gènes faisant partie de pathway de la régulation de la prolifération et de la survie cellulaire si BLIMP1 est exprimé par les cellules du lymphome ALCL.

En conclusion, malgré l'observation de quelques lésions communes, l'analyse des profiles génomiques démontre que le lymphome ALCL ALK- présente des traits uniques si comparé avec le lymphome ALCL ALK+. Cela suggère que de différents événements pathogéniques peuvent être nécessaires pour le développement des deux sous-types de ce lymphome. De plus, la fréquente inactivation de BLIMP1 retrouvée dans les échantillons cliniques du lymphome ALCL, suggère un rôle de suppresseur de tumeur pour ce gène tant *in vivo* que dans le model *in vitro*. Cela est démontré par la réexpression de BLIMP1, qui comporte croissance cellulaire, augmentation de la mort cellulaire et un arrêt du cycle cellulaire.

Table of Contents

Abbreviation List

aa	amino acids
ABC	Activated B-Cell like
ACA	Abnormal Chromosomal Arm
ACVBP	doxorubicin, cyclophosphamide, vindesine, bleomycin, prenisolone
AITL	AngioImmunoblastic T-cell Lymphoma
ALCL	Anaplastic Large Cell Lymphoma
ALK	Anaplastic lymphoma Kinase
ASC	Antibody Secreting Cell
AWC	Abnormal Whole Chromosome
B-NHL	B-cell NHL
BAC	Bacterial Artificial Chromosome
BL	Burkitt Lymphoma
BLIMP1	B Lymphocyte Induced Maturation Protein 1
C-CGH	Conventional-CGH
cALCL	Cutaneous ALCL
CGH	Comparative Genomic Hybridization
CHOP	cyclophosphamide, doxorubicin, vincristine, prednisone
CLL	Chronic Lymphocytic Leukemia
CML	Chronic Myelogenous Leukemia
CN	Copy Number
CNS	Central Nervous System
CNV	Copy Number Variation
DAG	Diacylglycerol
DD	Differential Display
DLBCL	Diffuse Large B-Cell Lymphoma
EATL	Enteropathy-associated T-cell Lymphoma
EBV	Epstein Barr Virus
EMA	Epithelial Membrane Antigen
ENKTL	Extranodal NK/T-cell Lymphoma

FDA	Food and Drug Administration
FFS	Failure Free Survival
FISH	Fluorescent In Situ Hybridization
FL	Follicular Lymphoma
GC	Germinal Center
GCB	Germinal Center B-cell like
GEP	Gene Expression Profiling
GFP	Green Fluorescent Protein
GISTIC	Genomic Identification of Significant Targets in Cancer
GSEA	Gene Set Enrichment Analyis
HARP	Heparin Affinity Regulatory Peptide
HBNF	Heparin-Binding Neurotrophic Factor
HD/HL	Hodgkin Disease/Hodgkin Lymphoma
HSTL	HepatoSplenic T-cell Lymphoma
IgH	ImmunoGlobulin Heavy chain
IgHV	ImmunoGlobulin Heavy chain Variable region
IL	Interleukin
IMT	Inflammatory Myofibroblastic Tumor
IP3	Inositol triphosphate
IRS-1	Insulin Receptor Substrate-1
JNK	JUN N-terminal kinase
KD	Kinase Domain
kDa	Kilo Dalton
KO	Knock-Out
LOH	Loss Of Heterozygosity
LRA	Long Recurrent Abnormality
M-FISH	Molecular-FISH
MACOP-B	methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisone, bleomycin
MALT	Mucosa-Associated Lymphoid Tissue
MCL	Mantle Cell Lymphoma
MCR	Minimal Common Region
MK	MidKine

MMAE	MonoMethyl Auristatin E
mTOR	mammalian Target Of Rapamycin
MZL	Marginal Zone Lymphoma
NGS	Next Generation Sequencing
NHL	Non-Hodgkin Lymphoma
NKCL	Natural Killer Cell Lymphoma
NPM	NucleoPhosMin
NSCLC	Non Small Cell Lung Cancer
OS	Overall Survival
OSF-1	Osteoblast-Specific Factor-1
PKC	Protein Kinase C
PLCγ	PhoshoLipase C-γ
PRDI	Positive Regulatory Domain I
PTCL	Peripheral T-Cell Lymphoma
PTCL, NOS	PTCL, Not Otherwise Specified
PTN	PleioTroPhin
REAL	Revised European and American Lymphoma
RTK	Receptor Tyrosine Kinase
SHC	Src Homology 2 domain-containing
SKY-FISH	Spectral KarYotyping-FISH
SNP	Single Nucleotide Polymorphism
SPTL	Subcutaneous Panniculitis-like T-cell Lymphoma
SRA	Short Recurrent Abnormality
STAT3	Signal Transducer and Activator of Transcription 3
T-NHL	T-cell NHL
TCR	T-Cell Receptor
TNF	Tumor Necrosis Factor
UPD	UniParental Disomy
WGSA	Whole Genome Sampling Analysis
WHO	World Health Organization

Figures Index

Number *Page*

Introduction, Materials and Methods

Results

I. Genomic profiling of Anaplastic Large Cell lymphoma

II. Massive RNAseq defines novel tumorigenic gene fusions of Anaplastic Large Cell Lympoma

Appendix

Tables Index

Number *Page*

Introduction, Materials and Methods

Results

I. Genomic profiling of Anaplastic Large Cell lymphoma

Appendix

Introduction

1. CANCER GENOME

Cancer is one of the major cause of death worldwide, accounting for around 13% of all deaths in 2008 [1] and deaths for this cause worldwide are projected to continue rising, with an estimated 13,1 million deaths in 2030. Cancer includes more than 100 distinct diseases caused in part by diverse risk factors (chemicals and environmental factors, diet, hormones, smoke, viruses, radiations...) and with various epidemiology. It could rise from most of the cell types and organs of the human body and is characterized by relatively uncontrolled proliferation of neoplastic cells that can invade beyond normal tissue boundaries and metastatize to distant organs. Human cancers usually evolve through a multistage process that can extend over a period of decades, during which cancer cells acquire different biological capabilities. This process is driven by the progressive accumulation of genomic instability, mutations and epigenetic changes that lead, in the end, to an altered expression or functionality of genes involved in different cellular processes. In fact cells need to acquire the traits, also known as "hallmarks of cancer" [2] (Fig 1), that enable them to become tumorigenic and ultimately malignant. Hanahan and Weinberg recently reviewed the concept about the six characteristics of neoplastic cells that constitute an organizing principle for rationalizing the complexities of tumoral cells. The "hallmarks of cancer" include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis and activating invasion and metastasis (Fig 1). Importantly, genome instability generates the genetic diversity that allows the acquisition of these capabilities.

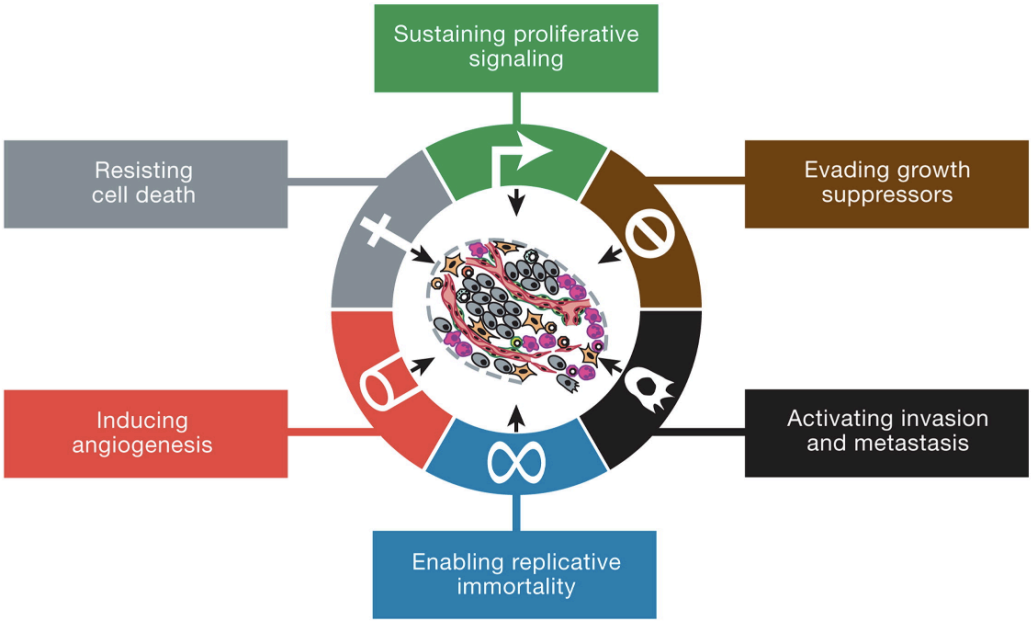


Figure 1. The six hallmarks of cancer: *distinctive and complementary capabilities that enable tumor growth and metastatic dissemination. Figure taken from Hanhan D et al, Cell 2011 [2].*

Cancer is fundamentally a disease of failure of regulation of tissue growth: genes responsible of the regulation of cell growth, proliferation and differentiation become altered. The affected genes are divided mostly into two broad categories: oncogenes and tumor suppressor genes. Oncogenes are genes that give an advantage to cancer cells when activated by mutations or over-expressed by gain/amplification or translocations. Activated oncogenes can bring cells that ought to die, such as damaged or neoplastic cells, to survive and proliferate and even a single mutation can exhert the oncogenic effect, in this case being the mutation dominant. Examples of oncogenic aberrations include recurrent mutations affecting *KRAS*, *BRAF*, *EGFR*, and *PIK3CA*, amplifications in *HER2* and *MYC*, and fusion genes involving *MYC*, *ABL*, ETS family members and *ALK* [3]. Conversely, tumor suppressor genes are genes that protect cells from neoplastic transformation, having, the proteins for which they code, a dampening or repressive effects on the regulation of cell cycle or in the promotion of apoptosis. When one of these genes is mutated, it can cause a loss or reduction of its function and the cell can progress

to cancer, but usually in the presence of additional genetic events. In fact, especially for tumor suppressor genes, both alleles that code for a particular protein must usually be affected before an effect is manifested ("two-hit hypotesis", also known as Knudson's hypotesis) [4]. *RB1*, *TP53*, *NF1*, *APC* and *CDKN2A* are classical examples of tumor suppressor genes inactivated by genomic aberrations [3]. The identification of key-genes involved in tumorigenesis can be very helpful for the choice of best molecular targeted agent and the appropriate combination therapy for a specific tumor. The products of oncogenes represent attractive therapeutic targets, because inhibition of these aberrant or abundant proteins is likely to reverse the oncogenic effect of the affected gene or to lead to cell death of cancer cells addicted to these aberrations [3]. Generally, it is more difficult to develop therapeutic strategies for cancers associated with mutations in tumor suppressor genes since this often required the restoration of the wild type gene function, an undoubtedly more difficult task.

1.1 Genetic aberrations in cancer

Cancer is a disease of the genome since cancer cells are characterized by cytogenetic abnormalities that disrupt the normal balance among cell proliferation, survival, and differentiation: many of these alterations are mediated by genetic changes associated with chromosomal abnormalities (Fig 2). Balanced chromosomal rearrangements and chromosomal imbalances represent the two main classes of chromosomal abnormalities found in human cancer, as reported in Figure 2. In haematological cancers and certain type of solid tumors, some of these lesions have emerged as prognostic and predictive markers. The development of new approaches for the molecular characterization of cytogenetic abnormalities has provided new insights into the mechanisms of tumorigenesis and has, in few instances, led to treatment that targets a specific genetic abnormality [5].

Lesions can determine the deregulated expression of genes with activation of oncogenes and inactivation of tumor suppressor genes and different evidences suggest that these aberrations are early or even initiating events in the tumorigenic process. But unfortunately the cause of chromosomal abnormalities remains poorly understood. Different studies have shown that certain environmental and occupational exposures and therapy with cytotoxic drugs can induce chromosomal aberrations, even if, for most cancer-associated chromosomal abnormalities, no specific initiating factors have been identified [3].

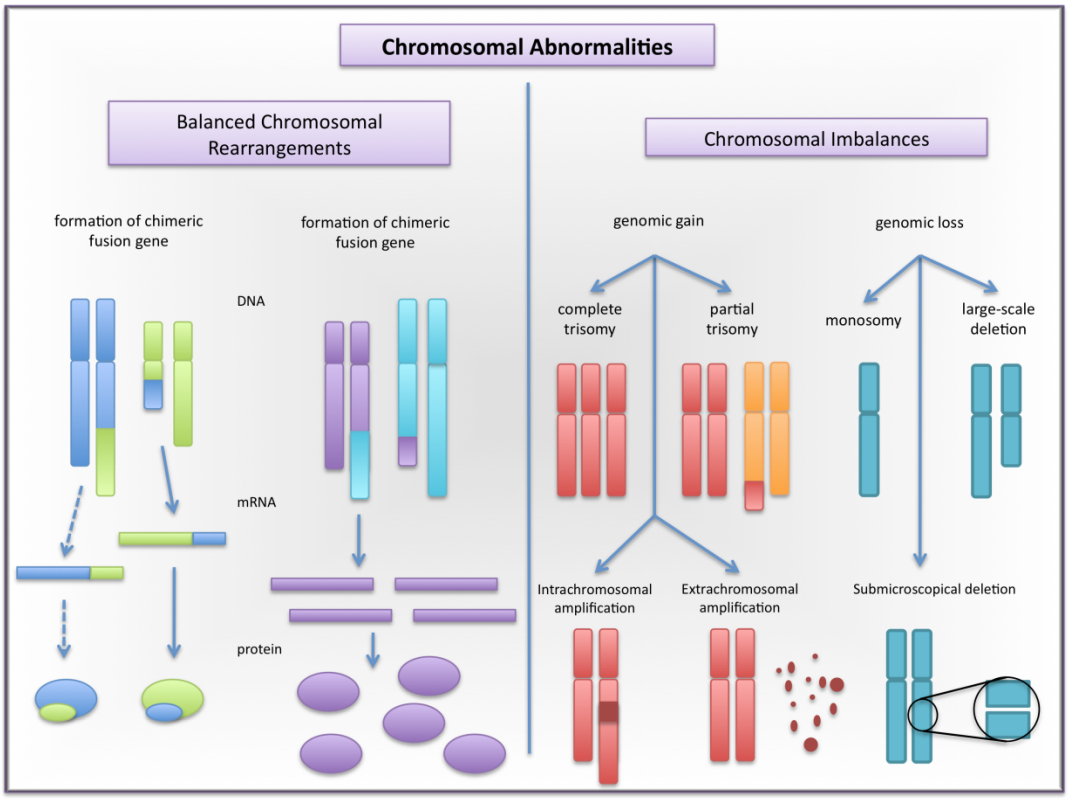


Figure 2. Chromosomal abnormalities in human cancer. *Balanced chromosomal rearrangements, represented by translocations, comprise lesions that lead to the formation of a chimeric fusion gene and lesions that lead to the aberrant juxtaposition of gene regulatory elements to the coding sequence of a structurally intact gene. The formation of a chimeric fusion gene results on the expression of a chimeric protein with new or altered activity. In the majority of cases only one of the two fusion genes, generated and not the reciprocal counterpart (indicated by the dashed arrow), contributes to cancer pathogenesis. The deregulated expression of a structurally normal gene results in deregulated expression of a normal protein. Chromosomal imbalances can be categorized into genomic gains and genomic losses, situations in which there is a change in the copy number of one or more genes, which is two for human beings.*

1.1.1 Translocations

Among genetic alterations, chromosomal translocations are considered the primary cause for many cancers, including lymphoma, leukemia and some solid tumors. In the late nineteenth and early twentieth centuries, translocations were first detected citogenetically as novel chromosomes that appeared prominently in tumor cells. The great German cytogenetist Theodor Boveri (1862-1915), considered as the first cancer geneticist, based on his observations of tumors, postulated that tumor cells possessed "growth-stimulatory chromosomes" that play a role in malignancies [6]. One hundred years later, we can affirm that Boveri's insights were correct. Translocations are indeed common in cancer cells, and in some cases they are able to produce oncogenes that are responsible for neoplastic transformation and uncontrolled growth of tumoral cells.

Translocations, in which a segment of one chromosome is transferred to a nonhomologous chromosome or to a new site on the same chromosome, can be distinguished into two main categories: balanced translocations and Robertsonian translocations. Balanced translocations comprise *nonreciprocal translocations*, one-way translocations in which a chromosomal segment is transferred to a nonhomologous chromosome, and *reciprocal translocations*, which involve the exchange of segments from two nonhomologous chromosomes [7]. No genetic material is lost during the exchange in balanced translocations. Conversely, the joining of the long q arms of two acentric chromosomes at a single centromere is the characteristic of Robertsonian translocations. In this case, chromosome p arms are lost, but because their shortness and repetitiveness, there are no phenotypic consequences after the loss of genetic material [7].

Translocations can broadly have two consequences depending on the chromosome breakpoint [5] (Fig 2):

1. It can result in the formation of a unique fusion or chimeric gene, with a potent oncogene function. An example of historic interest is the Philadelphia chromosome, the product of a reciprocal translocation involving small segments at the ends of q arms of chromosomes 9 and 22, in patients with Chronic Myelogenous Leukemia (CML). The translocation fused the coding sequence of *BCR* gene on chromosome 22 with the coding sequence of the *ABL* gene on chromosome 9. The BCR-ABL fusion protein encoded by the chimeric gene is a protein tyrosine kinase that constitutively activates signalling pathways involved in

cell proliferation and survival. An additional example of well-studied translocation of this type is the t(2;5)(p23;q35) which generates the chimeric fusion protein NPM-ALK, detected mostly in Anaplastic Large Cell Lymphoma, but also in other different type of cancer.

2. It can lead to the juxtaposition of the coding region of a gene in proximity of the transcriptionally active promoter/enhancer region of another gene, hence causing the over-expression of the former gene. The first translocation of this kind was described in patients of Burkitt Lymphoma (BL), in which *MYC* proto-oncogene was placed from chromosome 8 under the control of the powerful immunoglobulin heavy chain gene (IGH) promoter on chromosome 14. The result of the translocation is MYC over-expression in lymphoid cells.

Since malignancy is a multistep process, probably translocations are necessary but not sufficient to induce a fully malignant phenotype. Nevertheless, the identification of translocations could be crucial in determining the most appropriate therapy. In this regard, the Philadelphia story is a milestone in the hematologic field: the initial observation of a chromosomal abnormality in CML patients and the identification of this aberration as a translocation led to the molecular analysis and functional characterization of the genes involved in the t(9;22) translocation with the final result of development of a drug which specifically targets the defective gene product [8]. From the Philadelphia story, the take-home message is that the understanding of the full genetic and molecular effects of chromosome rearrangements is undoubtedly one of the best way for the development of cancer-specific drugs.

1.1.2 Homozygous Deletions, Losses, Gains and Amplifications

Copy number (CN) changes are defined as genomic regions with a number of DNA copies different from the normal CN (which is two for human beings). We can divide these aberrations in four categories: homozygous deletion (CN=0), loss (CN=1), gain (CN=3,4) and amplification (CN>4).

The spectrum of genomic *losses* ranges from cytogenetically visible alterations, such as complete or partial chromosomal monomies, to single gene or intragenic deletions, detectable only with the most sensitive techniques (Fig 2). Most recurrent genomic losses probably contribute to malignant transformation by reducing the function of specific genes in the affected chromosomal regions. In tumors, extensive genomic deletions are frequent and in this case the best approach to identify genes important for

tumorigenesis is the determination of minimal common region that is lost in all inter- and intra-tumoral cases analyzed. Candidate genes from this region are then screened for deletions, mutations, or epigenetic modifications that inactivate the remaining allele. This strategy has identified important tumor suppressor genes. According to the two-hit hypotesis, the first event is tipically viewed as a point mutation that inactivate one copy of a tumor suppressor gene; the second hit involves a deletion within the other chromosome that eliminates the remaining wild-type allele [4]. An important variant of the two-hit hypotesis is the *homozygous deletion* of a relevant gene, resulting from a focal deletion of a tumor suppressor locus followed by loss of the remaining allele. There are genes, also called haplo-insufficient tumor suppressor genes, that contribute to tumorigenesis by deletion of a single allele [9].

Most recurrent genomic *gains* probably contribute to tumorigenesis by enhancing the activity of specific genes in the affected chromosomal regions. Genomic gains commonly arise from chromosomal non-disjunction or unbalanced translocation, which cause complete or partial chromosomal trisomies, or from amplification events affecting DNA segments of different size [5]. Gains affecting small genomic regions or even single genes have been described less frequently than large gains.

Gene *amplification*, defined as a high copy number increase of a restricted region of a chromosome arm, can also occur. Regions of focal high level copy number change, which require a DNA double strand break and progression through the cell cycle with this damaged DNA, are likely to represent aberrations continuously under selection during tumor growth since amplified DNA is unstable [10]. They often harbor oncogenes and thus are useful for identifying genes or pathways that foster tumor development. 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 [10]. Nevertheless, change in DNA copy number is not the only way expression and function of a gene can be altered: other mechanisms for deregulation may be post-transcriptional, post-translational or involve alteration in expression of upstream genes.

1.1.3 Loss of Heterozygosity and UniParental Disomy

Loss of Heterozygosity (LOH) is the demonstration of the loss of one allele at a heterozygous locus. LOH is tipically the effect of a DNA deletion affecting one allele. However, other mechanisms can give origin to LOH, in the presence of a normal CN (the

so-called "copy neutral LOH" or "Uniparental Disomy" or UPD) (Fig 3): somatic recombination between two homologous chromosomes, including break-induced replication and gene conversion (interstitial mitotic recombination event), mitotic non-disjunction, and non-disjunction followed by duplication of the remaining chromosome may result in LOH.

LOH, or copy neutral LOH, can underlie the inactivation of tumor suppressor genes. The classical example of LOH inactivation of a gene with tumor suppressor activity is the case of retinoblastoma. Retinoblastoma, a rapidly developing cancer in the cells of retina, occurs sporadically in most cases, but in some families it displays autosomal dominant inheritance. With the classic model of the double hit, Knudson proposed, for the inherited cases of the diseases, that the first hit in the *Rb* gene is present in the germline, result of the inheritance from one parent. However the presence of the lesion is insufficient for tumor formation. The second hit in the gene, mostly consisting of a somatic mutation, seems to be necessary for promoting tumor formation [11]. Moreover, it is now clear that copy neutral LOH can also represent the duplication of a genomic region containing a mutated oncogene, which, thus, becomes homozygous (as for example the case of *JAK2* gene in chronic myeloproliferative disorders [12]). Importantly, because of the lack of change in copy number, UPD remains undetected by metaphase cytogenetics. Copy neutral LOH can also not represent a somatic event, but it might be due to the inheritance of the same alleles from both parents. Inherited UPD or autozygosity may constitute an independent predisposition factor for the development of malignancy [13].

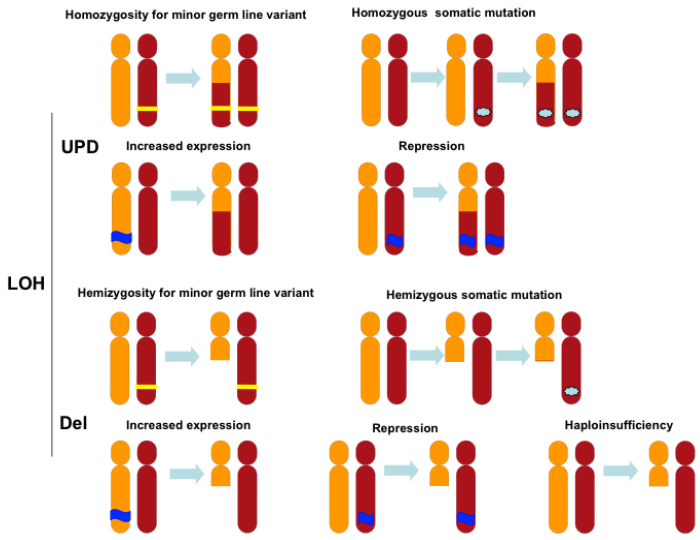


Figure 3. Pathogenetic actions of LOH, both UPD and deletion. *UPD can lead to the duplication of a disease-linked minor germline variant (top line, left) or an acquired mutation (top line, right). It can also lead to increased gene expression by the loss of a negative epigenetic mark (second line, left) or decreased gene expression by the duplication of a repressive epigenetic mark (second line, right). Deletion of chromosomal material can lead to the detection of a minor germline variant (third line, left) of acquired mutation (third line, right), resulting in hemizygosity. Furthermore, deletion can affect gene expression: it can lead to increased expression through loss of an imprinted allele, repression by loss of the expressed allele, or haploinsufficiency. Figure taken from Tuna M et al, Trends Mol Med 2009[14].*

There are similarities and important differences between the consequences of LOH due to deletion or copy neutral LOH. The major difference between them might be a difference in the gene dosage; CN loss leading to LOH is mostly present with a mutation in the remaining allele, but it can also be the only genetic event affecting the gene (haploinsufficiency). Conversely, loss of one allele and duplication of the remaining allele leads to copy neutral LOH (homozygosity of the mutated or methylated allele). Differences in gene dosage might have different pathologic results [13].

1.1.4 Chromothripsis

Chromothripsis, (Greek; chromos for chromosome, thripsis for shattered into pieces), is a very recently discovered mechanism for genetic instability in cancer cells. It has been identified thanks to the now available high-resolution high-throughput genomic tools. Stephens et al [15] described a type of cataclysmic event in which chromosomes are broken into many pieces and then stitched back together (Fig 4). Chromothriptic events are more likely to occur in a single catastrophic event than a series of subsequent and random alterations. Interestingly, these alterations are primarily limited to a single chromosome or, in some cases, a few chromosomes that appear to be coordinately altered.

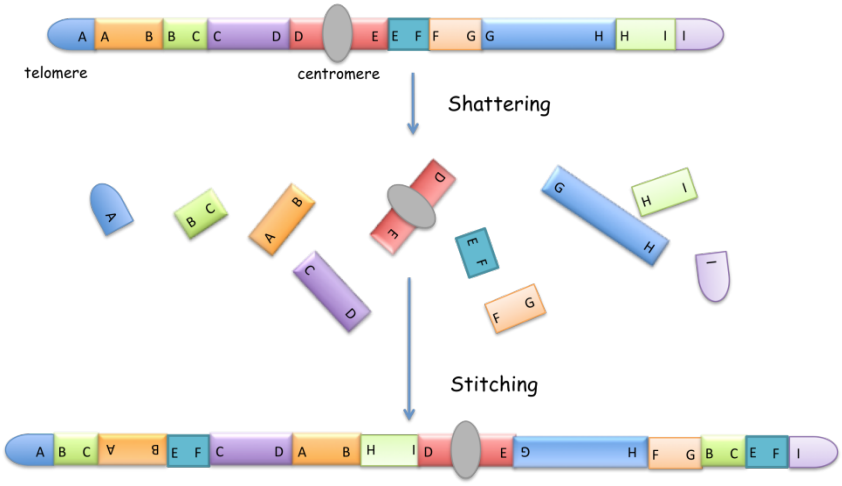


Figure 4. Stitching together shattered chromosomes by chromothripsis. *Chromothripsis occurs through a single catastrophic shattering event followed by the stitching of genomic fragments into derivative chromosomes.*

A combination of genome re-sequencing and analysis of SNPs array in cell lines and tumors suggests that this event occurs in 2-3% of cancers, spanning a wide variety of tumor types [16].

The origins of chromothriptic events during cancer genome evolution remain obscure, but at least two possible mechanisms seem plausible. First, chromothripsis may arise from the failure to restart stalled DNA replication forks in the S or G2 phases of the cell cycle before

entry into mitosis. Another possible explanation is the reintegration of broken chromosomal DNA from micronuclei formed during mitotic cell division [17]. The consequences of this event may lead to the generation of amplifications of one or more oncogenes or to the deletion of one or more tumor suppressor genes.

1.1.5 Somatic Mutations

The mutations found in a cancer cell genome have accumulated over the lifetime of the cancer patient. Somatic mutations occur in the genomes of all dividing cells, both normal and neoplastic: they can be considered as errors that occur during meiosis or DNA replication. The organism itself, by cellular processes as the hypermutation, can also induce them. Most of the occurred damage is repaired, however, a small fraction may be converted into fixed mutations. Mutation can result in several different types of change in sequences, which can have no effect, alter the product of a gene, or prevent the gene from functioning properly or completely. Very often, exchange of a single nucleotide with another one is observed, and this type of mutation is known as point mutation. Point mutations that occur in the protein-coding region of a gene may be classified in three types, depending upon the codification of the erroneous codon:

- *Silent Mutation* or Synonymous, if the mutated codon codes for the same amino acid;
- *Missense Mutation* or Non-Synonymous, if the mutated codon codes for a different amino acid;
- *Nonsense Mutation*, if the mutated codon codes for a stop codon and can truncate the protein.

Mutations can be classified also by their effect on protein function:

- *Loss-of-function Mutation*, if the mutation leads to a gene product having less or no function, typical in the case of tumor suppressor genes;
- *Gain-of-function Mutation*, if the mutation leads to a gene product having a new or abnormal function, as in the case of oncogenes;
- *Dominant Negative Mutation*, if the mutation leads to an altered gene product that acts antagonistically to the wild-type allele;

Cancer genomes carry two biological classes of somatic mutation classified on the basis of their consequences for cancer development: *driver* and *passenger* mutations. Driver mutations confer growth advantage on the cell in which they occur, are casually implicated in cancer development and have therefore been positively selected. By definition, these

mutations are in "cancer genes" implicated in tumorigenesis. Conversely, passenger mutations have not been subject to selection. They were present in the cell that was the progenitor of the final clonal expansion of the cancer, and/or are accumulated during cancer progression and they are biologically neutral and do not confer growth advantage. A challenge to any mutation screens is, therefore, to distinguish driver from passenger mutations [18].

1.1.6 Epigenetic Lesions

Epigenetic refers to heritable changes in gene expression without modifications in DNA sequence. There are two primary and interconnected epigenetic mechanisms: DNA methylation and covalent modifications of histones. In addition, it has also becoming apparent that RNA is intimately involved in the formation of a repressive chromatin state. DNA methylation occurs almost exclusively at CpG islands, CpG-rich regions preferentially located at the 5' end of genes that occupy almost 60% of human gene promoters, and contributes in the regulation of gene expression and the silencing of repeat elements in the genome. Histone modifications (including acetylation, methylation and phosphorylation) are important in transcriptional regulation and are often associated with DNA methylation [19]. Epigenetic mechanisms are essential for normal development and maintenance of tissue-specific gene expression patterns and the disruption of these processes can lead to altered gene function and malignant cellular transformation. An explosion of data in the field of epigenetics in the last decade has resulted in a better understanding of the epigenetic processes and has established the notion that cancer is at least in part an epigenetic disease. It is now accepted the concept that genetic and epigenetic alterations work in concert for the initiation and maintenance of many human tumor types [20, 21].
Pathologic epigenetic changes include global DNA hypomethylation in tumors, which contributes to cancer development by inducing generation of chromosomal instability and reactivation of transposable elements. Hypermethylation of the CpG islands in the promoter region of genes and chromatin alterations work together with histone modifications to silence well-characterized tumor suppressor genes [22].

2. Study of Genomic Aberrations in Lymphomas

Part of data reported in paragraph 2.1.2 have been published in: *Boi M et al. Genetic alterations in systemic nodal and extranodal non-cutaneous lymphomas derived from mature T cells and natural killer cells. Cancer Sci. 2012*

Decades before the availability of sequence technology, the first differences observed in our genetic material were mainly rare changes in quantity and structure of chromosomes. These included aneuploidies, rearrangements, heteromorphisms and fragile sites, all of which were large enough to be identified with the use of a microscope. The traditional technique for the detection of such abnormalities is cytogenetic that, through the chromosome banding, allows the reconstruction of a map of all the chromosomes and the determination of any aberrations.The first applications of this method dates back to the late '50s, a time when medical genetics has undergone significant change through the identification of methods capable of separating the metaphase chromosomes, thus identifying the correct number of human chromosomes and allowing the identification of new syndromes related to chromosomal deletions or trisomies (i.e., Down's Syndrome 1959) [23, 24]. In the '60s the techniques for analysis of chromosomes have improved becoming simpler and faster. Specific staining techniques have allowed pointing out each chromosome in a typical alternation of bands. Although the first staining protocols for chromosome banding date back to 1967, this technique is still commonly used. Among the various techniques for chromosome banding, the most widely used is the G-banding: the result is a clear division of chromosomes into bands, which identify chromosomal regions rich in GC or in euchromatin, and dark bands, areas rich in AT or heterochromatin (Fig 5).

In recent decades, the development of new technologies increasingly standardized and sophisticated has favored a more precise definition of the chromosomes allowing the appreciation of microdeletions or microduplications that can escape to a classical cytogenetic analysis.These aberrations can be detected by the most modern techniques of molecular cytogenetic involving the use of DNA probes. The Fluorescent *In Situ* Hybridization (FISH) represents the first molecular cytogenetic method that provides the use of a probe labeled with a fluorescent dye allowing the identification of specific sequences of DNA or RNA (Fig 5). The chromosomal targets that may be employed are: metaphase chromosomes, interphase nuclei, fresh biopsy tissues or paraffin sections [25, 26]. DNA is fixed and denatured "in situ" on the slides to make the two strands of DNA

accessible to hybridization of the labeled probe. The probe binds to the complementary gene region and the analysis is performed under a fluorescence microscope. FISH can be considered a valid support to conventional cytogenetics particularly in hematology where it can identify the genetic alterations with greater sensitivity and specificity than the standard cytogenetic. FISH can identify numerical chromosomal abnormalities such as trisomies, monosomies, tetraploidies and structural abnormalities such as inversions, translocations and interstitial deletions.

A further development of the standard FISH has seen the establishment of methods known as Spectral Karyotyping (SKY) or Multiplex (M)-FISH [27] (Fig 5) which are based on a combinatorial labeling strategy allowing in this way the analysis of the entire genome simultaneously, overcoming one of the limitations of FISH represented by a partial observation of the genome. This technique permits a more clear division of the chromosomes in metaphase thanks to the use of different combinations of fluorophores (24 colors) with a specific staining for each chromosome. With this technique both simple and complex chromosomal abnormalities can be identified rapidly and unambiguously in metaphase chromosomes of normal or tumoral cells.

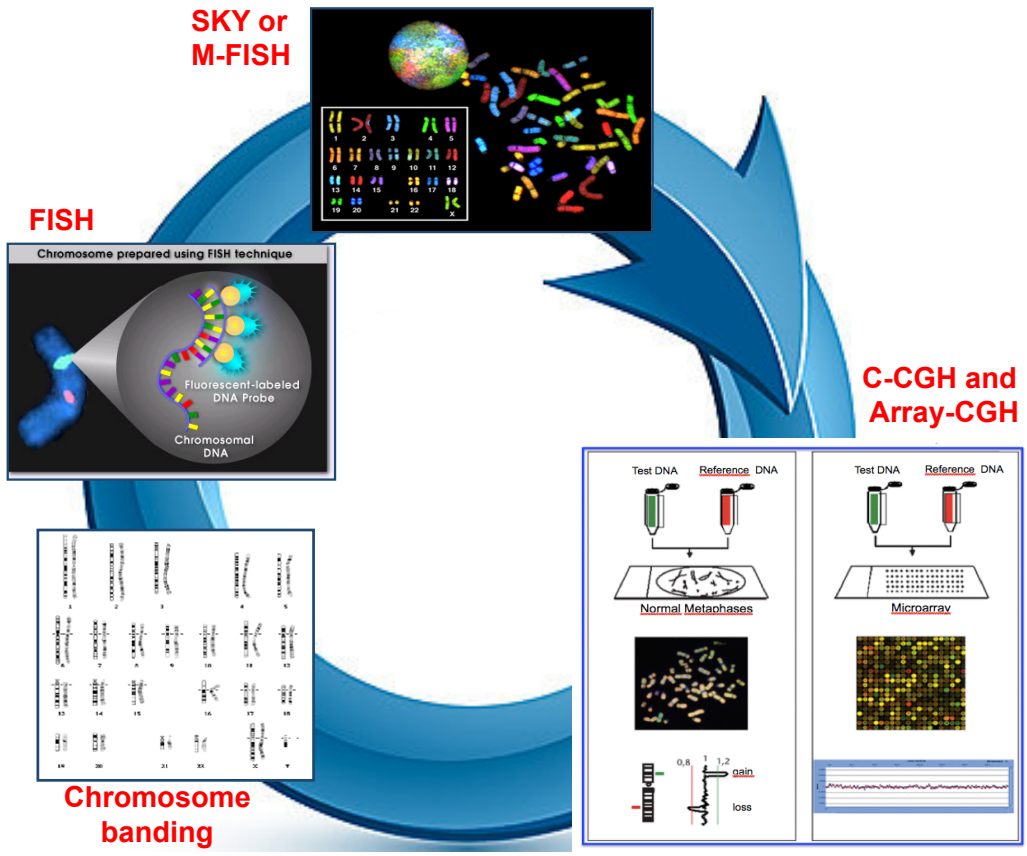


Figure 5. Evolution of techniques used in the study of genomic aberrations in lymphoma. *Starting from '60s with chromosome banding (in the bottom left part), cytogenetic studies have seen the progress of FISH ('80s) and the subsequent development of SKY or M-FISH. With the advent of molecular biology in the '90s, C-CGH and Array-CGH have improved the study of genomic aberrations.*

2.1 Array-CGH and Gene Expression Profiling

The physiological development of molecular cytogenetics has led to the definition of new methods capable of analyzing the entire genome with a higher resolution, looking for new cryptic abnormalities in regions that escape conventional cytogenetics. Genome-wide DNA profiling is a powerful approach for cancer research and in particular for the characterization of aberrations in lymphomas (Fig 5).

Conventional Comparative Genomic Hybridization (C-CGH) is a technique that analyzes the entire genome in a single experiment, and identifies genetic abnormalities such as

amplifications and/or deletions between two samples of DNA by competitively hybridizing the differentially labelled DNA to metaphase chromosomes [28, 29] (Fig 5). Through competition between test (i.e. tumor) and control diploid DNA, imbalances due to CN differences result in a shift in the fluorescence spectra (Fig 5). This technology later evolved so that the DNA targets are hybridized to microarrays containing cDNA fragments or bacterial artificial chromosomes (BACs). This technique has been a valuable tool in the analysis of lymphomas, as, knowledge of the genomic alterations in this heterogeneous group of diseases is important for the classification, prediction of outcome and for the improvement of understanding of the biology [30, 31]. Nevertheless, the resolution of C-CGH is limited by the length of the metaphase chromosomes and is approximately 10 megabases and only hundreds of genes could be identified into the chromosomal aberrations recognized.

More recently, microarray-based format (array-CGH) using large insert genomic clones, cDNAs or oligonucleotides have replaced metaphase chromosomes and it can be used to create a virtual karyotype (Fig 5). This advance technique provides the advantage over metaphases of a higher solution, and the ability to directly map the CN changes to the genome sequence. 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 CN changes and many other applications. It enables analysis of the whole genome in a single experiment and allows the global profiling of DNA CN changes in tumors and accurate determination of the breakpoint of regions that are gained and/or lost [32]. In the lymphomas field, the improved resolution of array-CGH formats has increased the number of genomic aberrations identified, and a picture is beginning to emerge of their clinical association. Importantly, a proportion of smaller amplifications or deletions detected by this technique could not be detected by metaphase CGH.

An evolution of this technique was the use of arrays of photolithographically synthesized short oligonucleotides, originally designed for detecting SNPs, DNA variation occurring when a single nucleotide in the genome differs between members of the same biological species or paired chromosomal in an individual, in copy number assessment in CGH experiments [33]. In oligonucleotide arrays, the probes are short sequences designed to match part of the sequence of known or predicted genes. In the majority of cases oligonucleotide arrays are produced through a specific technique of manifacturing: they are produced by printing short oligonucleotide sequences designed to represent a single gene or family of splice-variants by synthesizing this sequence directly onto the array

surface, a technique known as photolithography. Sequences may be longer (60-mer probes as the Agilent design), which are more specific to individual target gene, or shorter (25-mer probes produced by Affymetrix), which may be spotted in higher density across the array.

The platform used in this study is the AffymetrixGenome-Wide Human SNP 6.0. Over the years many efforts have been done to improve the resolution of the array platforms used for the detection of genomic imbalances. At the beginning, Affymetrix developed an assay termed whole-genome sampling analysis (WGSA) for highly multiplexed SNP genotyping of complex DNA. This method reproducibly amplifies a subset of the human genome through a single primer amplification reaction using restriction enzyme digested, adapter-ligated human genomic DNA. This assay was first developed for simultaneous genotyping of over 10,000 SNPs on a single array (GeneChipHuman Mapping 10K Array Xba 142). Then, the WGSA assay was extended to allow highly accurate SNP genotyping of over 100,000 SNPs using the two arrays GeneChipMapping 100K Set. The WGSA assay was again extended and improved some years later with the product known as the GeneChipMapping 500K Assay in which 500,000 SNPs are queried using a two-array set. The sixth-generation product, the AffymetrixGenome-Wide Human SNP Array 6.0, also uses the WGSA assay that has been the hallmark characteristic of all previous mapping arrays. This single array interrogates 906,600 SNPs by combining the Nsp I and Sty I PCR fractions prior to the DNA purification step and through a reduction in the absolute number of features associated with each individual SNP on the array. This array also contains 945,826 copy number probes designed to interrogate copy number variations (CNVs) in the genome. The Genome-Wide Human SNP Array 6.0 thus provides a robust, flexible, cost-effective approach for scoring SNP genotypes in large numbers of samples.

Genomic information alone, although useful in predicting the existence of specific genes and providing some informations about the nature of the proteins produced by them, cannot predict all of the post-transcriptional and post-translational modifications that RNA transcripts and proteins undergo to give rise to multiple gene transcripts (mRNAs) and gene products (proteins). Thus, it has been important to develop and apply a variety of techniques that can accurately measure gene structure and activity at the level of mRNA as well. Over the past forty years, several methods have been developed to allow and improve comparative analysis among normal and tumoral cells in terms of gene expression [34]. Starting from simple approaches that used 2D gel electophoresis [35] to

compare protein expression, methods that focus on mRNA have evolved and become more sophisticated over time, thanks to the advent of recombinant DNA technology. The earliest approach in this field was differential hybridization [36], for which mRNA samples to be compared (for example, normal and tumor cells) were radioactively labelled as cDNA probes by reverse transcription. The resulting complex cDNA probes are then used to hybridize to the cDNA phage plaques lifted onto nylon membranes. Comparison in the hybridization pattern, allowed the identification of differentially expressed genes in the two pair samples, even if the sensitivity of this technique was very low, since it would not be able to detect most of the genes expressed at low levels. At the same time, the subtractive hybridization [37] was developed. The new ingenious approach was used to remove most of the cDNAs that represented the genes commonly expressed in both cells being compared and enriched for cDNA probes that represent mRNA uniquely expressed in one cells but not the other. With the aim to overcome previous methods limitations, differential display (DD) [38] was proposed in the early 1990s by integrated two of the most simple and powerful molecular biology methods, namely PCR and DNA sequencing by gel electrophoresis. This approach worked with the amplification of the 3' termini of eukaryotic mRNA by reverse transcription-PCR using one of the three anchored oligo-dT primers (that is the run of Ts ending with a C, G or A), in combination with a set of short primers of arbitrary sequences. By changing primer sequences, different subsets of mRNA can be analyzed and displayed by denaturing polyacrylamide gel electrophoresis. Differences in gene expression are indicated by comparison of complementary DNA pattern between or among relevant RNA samples. Differentially expressed cDNA bands can be purified and sequenced for further molecular characterization. Initially, DD suffered from a high rate of false positive, but technical improvements have greatly reduced this number. Nevertheless, this approach allowed the identification of only few genes at a time, giving DD a low-tech, low-precision and low-throughput image. In the half of 1990s, cDNA microarrays [39] and GeneChip arrays [40] were developed: these techniques are, in essence, based on the differential-hybridization strategy, in which cDNA plaques are replaced with spotted cDNA or oligos, and radioactive labels are replaced with fluorescent ones. Oligonucleotide-based microarrays were the most used microarrays for genome-wide gene expression studies. Oligonucleotides arrays are built up through two major approaches: in the first, microarrays composed of short oligonucleotides (25 bases) are synthesized directly onto a solid matrix using photolitographic technology (Affymetrix). Alternatively, microarrays composed of longer

oligonucleotides (55-70 bases) can either be deposited by an ink-jet printing process (Agilent) or spotted by a robotic printing process onto glass slides (CodeLink). For the Affymetrix microarray system, only one sample is hybridized per chip exploiting a single-colour detection scheme, while Agilent technology uses a two-color scheme, whereby the same array is hybridized with two different samples. The chip is then scanned for the presence and strength of the fluorescent labels at each spot representing probe target hybrids. The level of fluorescence at a particular spot provides quantitative information about the expression of the particular gene corresponding to the spotted cDNA sequence [41]. Microarrays approach allowed the simultaneous analysis of the expression of mRNA from ten of thousands of genes and one of the great advantages over other methods is that each spot on the microarray contains a known sequence so, once a signal is detected, the nature of the gene is known.

This type of analysis can currently provide a correlation of certain pathological phenotypes with specific gene signatures, leading to a more precise and reproducible diagnosis. In addition to the histological analysis, GEP data will probably contribute to treatment strategies as a result of more precise prognosis definition and by individualizing drug susceptibility-specific patterns.

2.1.1 Study of B-cell Lymphomas

Recurrent chromosomal abnormalities are important prognostic indicators for various hematological diseases. The identification of genes involved in neoplastic hematologic disorders has contributed significantly to the elucidation of the biology of hematopoiesis and to the identification of functional consequences of genetic changes, leading to a new era of drug therapy [42]. Non-Hodgkin B-cell Lymphomas (B-NHL) is a set of B-lymphocyte malignancies with a great variety in terms of clinic and pathologic characteristics. In most cases, they are characterized by recurrent chromosomal translocations involving the regulatory regions of immunoglobulins genes. These translocations are considered primary events in the pathogenesis of B-NHL and the consequence is an aberrant expression of translocated oncogenes resulting in dysregulation of cell cycle inhibition or physiological apoptotic mechanisms. However, it has become clear that lymphomagenesis is a multistep process, involving the stepwise accumulation of molecular alterations as well as macro- and micro-environmental factors.

Diffuse Large B-Cell Lymphoma (DLBCL) is the most common B-NHL in the Western Countries accounting for 25-30% of all NHL. DLBCL is very heterogeneous in term of

histological and clinic features, suggesting the existence of biologically distinct subclasses within the same entity [43]. GEP studies in fact have contributed substantially to the understanding of the heterogeneity among DLBCL cases (Table 1). Alizadeh et al. [44] identified at least two major groups among DLBCL specimens: the germinal center B-cell like (GCB) DLBCL subgroups shows strong expression of genes that normally are highly expressed in germinal center B-cells, suggesting a GC B-cell as the cell of origin in this subtype. Conversely, the activated B-cell like (ABC) DLBCL subtype is characterized by a gene-expression signature with features reminiscent of *in vitro* activated B-cells, pointing to a post-GC B-cells as the cell of origin of these cases. These latter specimens are characterized also by the constitutive activation of NF-κB pathway. Interestingly, the heterogeneity between the two DLBCL subtypes is also recognizable in terms of genomic aberrations. The translocation t(14;18)(q32;q21), the amplification of the *REL* locus in 2p [45] and frequent gains at 12q level are considered typical alterations of the GCB subtype [46, 47]. On the contrary, ABC DLBCL carries more frequent gains at 3q and 18q levels, as well as losses in 6q. Interestingly, the deleted region in 6q includes the *PRDM1* gene locus (6q21) [48], whose gene product BLIMP1 is a target of *BCL6*, gene located at 3q27 locus and very often alterated in DLBCL through translocations and mutations. Indeed, a considerable proportion of ABC DLBCL reveal inactivation of *PRDM1*/BLIMP1 by somatic mutations, suggesting that the inactivation of this gene may play a pathogenetic role in this subset of DLBCL [49, 50]. Moreover, Lenz and colleagues [51] found activating mutations of *CARD11* in a DLBCL subset that likely contribute to the constitutive activation of the NF-kB pathway in these cases. Besides translocations involving *BCL6* gene [52], others are frequently identified in DLBCL: a translocation of the *BCL2* gene in chromosomal band 18q21 to the immunoglobulin-heavy chain (*IGHV*) locus in 14q32 is found in approximately 20% of cases [53]and the t(8;14), typical of Burkitt lymphoma, involving the *MYC* gene locus in 8q24 can be detected in approximately 6% of specimens [54].

Burkitt Lymphoma (BL), a highly aggressive B-NHL, is a rapidly proliferating lymphoma that often presents in extranodal sites or as an acute leukemia. Three clinical variants can be recognized: the endemic, the sporadic and the immunodeficiency-associated cases, each manifesting differences in clinical presentation, morphology and biology (Table 1). The cytogenetic hallmark of all variants of BL is the translocation involving *MYC* gene in chromosome 8q24 and the *IGHV* locus on chromosome 14q32 in more than 80% of cases, or the light chain loci [55]. The consequence of this translocation is the

deregulated expression of *MYC*, which is a key gene in the regulation of a wide variety of cellular processes such as proliferation, differentiation and apoptosis. BL usually carries relatively few secondary aberrations, most frequently gains in 12q, 20q, 22q, Xq, and losses of 13q [56]. Gains and amplifications of 1q and gains in 7q were also described in a substantial subset of cases, and were reported to be associated with a poor clinical outcome [57]. Concurrent translocations of *MYC* and *BCL2*, involved in t(14;18) were reported in a small percentage of BL patients, and these dual translocations were associated with a particular poor prognosis. GEP analysis revealed, as expected, an overactivation of the *MYC* signature, due to the overexpression of *MYC* and its target genes (Table 1). Moreover, another gene cluster was found to be highly expressed in BL: it consisted of a distinctive subset of genes expressed in normal GC B-cells [58]. However, the differential diagnosis between atypical BL and other aggressive lymphomas, especially DLBCL, can be difficult in the diagnostic setting, because morphologic and immunophenotypical features may overlap. Furtermore, on the genetic level, translocations involving the *MYC* gene, although a typical cytogenetic feature of BL, are not specific of this entity, but can also be identified in 5% to 10% of DLBCL [59] in which determines a poor outcome [60]. However, the precise discrimination between BL and DLBCL is of significant clinical importance because of the different therapeutic approaches used in the treatment of these patients.

Among the more common B-NHL, **Follicular Lymphoma** (FL) accounts for about 30% of all adult NHL. Generally, FL patients present an indolent clinical course, however, they display a wide variability with a possible progression of the disease or transformation to an aggressive lymphoma [61] (Table 1). The cytogenetic peculiarity of FL is the translocation t(14;18)(q32;q21), which can be detected in approximately 85% of the cases. This translocation juxtaposes the *BCL2* gene on chromosome 18q21 to *IgH* enhancer on chromosome 14, leading to a constitutive expression of *BCL2*. Array-CGH data have highlighted the presence of secondary chromosomal alterations in FL, confirming what has been suggested by different studies: the translocation alone is not sufficient to produce a fully malignant phenotype [62, 63]. The most frequently reported aberrations in FL include gains in chromosomes 1q, 2p, 7, 8q, 12q, 18q and X, as well as losses in 1p, 6q, 10q, 13q and 17p [64]. Among these alterations, losses of chromosome 1p as well as of the long or short arm of chromosome 17, and gains of 1q, 12, and X were identified as late events and negative prognostic factors [65] (Table 1). Transformation of FL into a more aggressive lymphoma, most commonly DLBCL, is reported in 10% to 60%

of cases and this event confers a poor clinical outcome. Gains of chromosome 7, 12q, and X and losses of 4q, 13q, and 17p are recurrent cytogenetic lesions associated with transformation, as well as inactivation of *TP53* in 17p13.1, the *CDKN2A* gene in 9p21.3, and a deregulated expression of *MYC*. GEP studies demonstrate the importance of the microenvironment influence in the pathogenesis, evolution and progression [66] (Table 1). The translocation t(11;14)(q13;q32), that leads to the over-expression of cyclin D1 gene, due to the juxtaposition of the *CCND1* gene on chromosome 11q13 to the *IgH* enhancer, is the cytogenetic hallmark of **Mantle Cell Lymphoma** (MCL) [67]. MCL is an aggressive B-NHL that comprises about 3-10% of NHLs and presents a considerable clinical variability with a generally poor clinical outcome [68]. Also in this case the t(11;14) alone is not sufficient to promote tumorigenesis and, in support of this hypotesis, high number of molecular secondary alterations have been reported in MCL [30, 31, 69] (Table 1). In CGH-based studies, the most frequently detected alterations include gains in 3q, 7p, 8q, 12q, and 18q and losses in 1p, 6q, 8p, 9p, 9q, 11q, 13q, and 17p [69] and many of these regions have been confirmed by higher resolution array-CGH studies [30, 31]. Trisomy 12 has been reported in about 20% of cases. Deletions of the *CDKN2A* locus on 9p21, amplification of *BMI1* on 10p11 and amplifications of *CDK4* on 12q14 were identified as single-gene alterations involved in the pathogenesis of MCL. Rare MCL cases are negative for cyclin D1 and the t(11;14), but present an expression profile and other features indistinguishable from conventional MCL [70]: these cases show high expression of cyclin D2 or cyclin D3. In some of these cases it is possible to identify a t(2;12)(p12;p13) fusing cyclin D2 gene to the kappa light chain gene locus [71].

Genome-wide DNA profiling has helped the elucidation of molecular pathogenesis of **Marginal Zone Lymphomas** (MZLs) [72] (Tale 1). MZLs are divided into three distinct subtypes: extranodal MZLs of Mucosa-Associated Lymphoid Tissue (MALT), nodal MZLs and splenic MZLs. Nevertheless, the genetic relationship between the three MZL categories is still unresolved. Similarities in histologic features as well as some genetic lesions (gains of chromosome 3 and 18 occurring at higher frequency than other types of lymphomas) suggest that the subtypes differ from other lymphomas, but there are no clear criteria able to differentiate among them. An analysis of 218 MZL patients revealed that MALT lymphoma presented significantly more frequently gains at 3p, 6p, 18p, and loss of *TNFAIP3/A20* gene located in 6q23, whereas splenic MZL was associated with losses of 7q31 and 8p. Nodal MZL did not show statistically significant differences compared with MALT lymphoma while lacking the splenic MZL-related 7q losses.

Conversely to specific genomic lesions, gains at 3q and 18q have been confirmed to be common to all three subtypes [72].

Pioneering studies of Dohner et al [73] with FISH were the first ones to identify an association between the genetic alterations (such as trisomy 12, deletion of 11q22-23 (*ATM*), 17p13 (*TP53*) and 13q14), and survival in **Chronic Lymphocytic Leukemia** (B-CLL) (Table 1). It is the most heterogeneous haematological neoplasm, both from the molecular and the clinical point of view. Indeed, it can present distinct clinical patterns: an indolent type, occurring in patients who do not require immediate therapeutic intervention and can survive more than 10 years, and the other, much more aggressive, with survival of even a few months. It has been demonstrated that in 50%-80% of B-CLL cases, the cell of origin presents more than 2% of somatic mutations in the variable region of the immunoglobulin heavy chain (*IGHV*) gene (mutated B-CLL); the remaining part of B-CLL patients has a "germline" configuration in that region with less than 2% of somatic mutations (unmutated B-CLL). The presence of mutations beyond a threshold of 2% is associated with a better prognosis respect to germline *IGHV* B-CLL cases, which show a more aggressive clinical course [74]. Thus, it was suggested that B-CLL might include two disparate malignancies; one derived from an unmutated *IGHV*, pre-germinal center B cell, and the other from a mutated *IGHV* B cell that has passed through the germinal center. However, using genomic scale gene expression profiling, Rosenwald at al[75]show that a common gene expression signature, irrespective of *IGHV* mutational status, characterizes B-CLL suggesting that B-CLL cases share a common mechanism of transformation and/or cell of origin (Table 1). Nonetheless, the expression of hundreds of other genes correlated with the *IGHV* mutational status, including many genes that are modulated in expression during mitogenic B cell receptor signaling. At a genomic level, more than 80% of B-CLL cases show the presence of alterations: the most common include losses of 13q14.3 (*miR-15a*, *miR-16-1*; 55% of cases), trisomy 12 (20%) and, less commonly, deletions of 11q22-23 (*ATM*), 17p13 (*TP53*) and 6q21 [73]. The distribution of these abnormalities varies based on mutational status of *IGHV* gene and they are associated with different clinical outcome. The most common detected anomaly in patients with B-CLL is the deletion of chromosome region 13q14 which whenoccurs alone, generally has a favorable prognostic value, however it is possible to observe within this population subjects with worst prognosis. In about 20% of patients cytogenetic analysis does not reveal any abnormalities and such patients have an intermediate prognosis. Trisomy of chromosome 12, usually present in a limited portion of the leukemic

cells is, therefore, interpreted as a secondary event; however, it is associated with intermediate prognosis. Definitely unfavorable are deletions 11q22-q23 and 17p13. Very recently, new genomic regions associated with poor outcome have been identified [76]. Gains at 2p (*MYCN* and *REL* genes) and 8q (*MYC*) and losses affecting 8p (*TRAILR1*, *TRAILR2* and *MCPH1* tumor suppressor genes) were identified for their poor prognostic impact in a series of more than 300 B-CLL patients. Gains at 2p were also associated with an increase risk of histological transformation to Richter Syndrome, the transformation to an aggressive DLBCL.

In the last decade, significant progress in the understanding of the molecular biology and the definition of underlying genetic alterations in B-NHL, which in part, can explain the clinical heterogeneity of the different entities, has been made. Newly available technologies and their application in the analysis of B-NHL will help even more to reveal distinct molecular and genetic features of many entities, most likely leading to a more personalized targeted therapy.

Table 1. **Summary of genomic studies in B-cell malignancies.**

	study	type of study	year	platform	num of cases	ref
DLBCL						
	Alizadeh AA	GEP	2000	Lymphochip	96 (normal and malignant lymphocytes)	[44]
	Rosenwald A	GEP	2002	Lymphochip	240 DLBCL	[45]
	Tagawa H	array-CGH	2005	Agilent (oligonucleotide-array custom-made)	99 DLBCL	[46]
	Bea S	CGH	2005	Vysis	224 DLBCL	[48]
	Chen W	array-CGH	2006	Spectral Genomics	64 DLBCL	[47]
	Mandelbaum J	array-CGH	2010	Affymetrix SNP 6.0	158 (139 primary samples and 19 cell lines)	[50]
BL						
	Hummel M	GEP	2006	Affymetrix U133A GeneChips	220 mature aggressive B-cell lymphomas	[55]
	Dave SS	GEP	2006	oligonucleotide arrays custom-made + Affymetrix U133 plus 2.0 (only for a subset)	303 aggressive lymphomas	[58]
	Garcia JL	CGH	2003	metaphase chromosome from normal individuals	34 BL	[56]
FL						
	Dave SS	GEP	2004	Affymetrix U133A/B	191 FL	[66]
	Viardot A	CGH	2002	metaphase chromosome from normal individuals	124 FL	[65]
MCL						
	Fu K	GEP	2005	Affymetrix U133A/B	6 MCL	[70]
	Kohlhammer K	array-CGH	2004	BAC/PAC array custom-made	53 MCL	[31]
	Rubio-Moscardo F	array-CGH	2005	BAC/PAC array (UCSF Human Array 2.0)	113 (68 primary and 45 cell lines)	[30]
	Salaverria I	CGH	2007	Vysis	77 MCL	[69]
MZL						
	Rinaldi A	array-CGH	2011	Affymetrix SNP 250k	218 MZL	[72]
CLL						
	Rosenwald A	GEP	2001	Lymphochip	37 CLL and other B cell malignancies	[75]
	Rinaldi A	array-CGH	2011	Affymetrix SNP 6.0	323 CLL	[76]

2.1.2 Study of T-cell Lymphomas

The T-cell lymphomas are a heterogeneous group of rare diseases that are divided into two main groups: those arising from precursor cells correspond to the single entity called "lymphoblastic lymhoma", while the lymphomas derived from mature lymphocytes are referred to as "Peripheral T-Cell Lymphomas" (PTCLs). The PTCLs represent <15% of all non-Hodgkin lymphomas (NHLs) [77]. Their classification is based on the presentation of the lymphoma, i.e. disseminated, nodal or extranodal/cutaneous disease. In sharp contrast with B-cell lymphomas, which are assigned to a normal cellular counterpart, the cellular origin of PTCLs remains difficult to determine, although molecular profiling has recently allowed significant progress [78]. Some of the entities recognized by the WHO classification are very rare and very scarce molecular data are available, but the genetic events, underlying systemic nodal and extranodal non-cutaneous mature T-cell and NK-cell lymphomas are reported here [79] (Table 2).

Peripheral T-Cell Lymphomas, not otherwise specified (PTCL, NOS) are a heterogeneous category of mature T- NHL, which do not correspond to any of the other specifically defined entities in the current WHO classification. Although representing the majority of T-NHL, the genetics of PTCL, NOS is poorly characterized (Table 2). Studies evaluating genomic aberrations in PTCL, NOS by the use of classical cytogenetic analysis, comparative genomic hybridization (CGH), or microarray-based CGH (arrayCGH) have shown a complex karyotype, characterized by a high genomic instability. Most frequent gains occur on chromosomal regions 1q32-qter, 2p, 7q22-ter (*CDK6* gene), 8q24, 9q33-qter,11q13, and 17q (17cen-q21). Recurrent losses mainly affect 6q21, 9p21 (*CDKN2A* and *CDKN2B* tumor suppressor genes), 10 (10cen-p12, 10q23-q24), 13q21, 14q, 16q, 17p13 (*TP53*) [80, 81]. Chromosomal translocations involving T-cell receptor genes are detected in only 1% or less of all PTCL, NOS. A t(5;9)(q33;q22) translocation causing the fusion of the tyrosine kinase domain of *SYK* to the N-terminal pleckstrin homology domain and proline-rich region of *ITK* [82] has been reported in a fraction of PTCL, NOS. Importantly, the overexpression of *SYK*, involved in proliferation and pro-survival signaling, can be seen in the majority of PTCL, NOS, despite the absence of *SYK/ITK* translocations in most of the cases. This observation provides a rationale for the usage of SYK inhibitors, such the fostamatinib disodium, currently under investigation in B-cell tumors. At a gene expression level, when compared to normal T-cells, PTCL, NOS profiles are characterized by the deregulation of genes involved in important cell functions such as matrix deposition, cytoskeleton organization,

cell adhesion, apoptosis, proliferation, transcription and signal transduction [83]. In particular, the over-expression of *PDGFRα*and the deregulation of NF-κB provide the rational for some of the clinical trials currently running for patients with PTCL, NOS.

Little is known about the genetics of **Angioimmunoblastic T-cell Lymphoma** (AITL), an aggressive systemic disease, characterized by consistent association with Epstein Barr virus (EBV), which suggests a pathogenetic role for this virus, eliciting chronic antigen-driven T cell-mediated responses (Table 2). Besides high frequencies of trisomies of chromosomes 3 and 5, however not specific for AITL, the most frequent gains have been reported at 11q13, 19, 20q13 and 22q, whilst losses have been reported at 13q22-q32, 8p22, and 9p21 (*CDKN2A*) [81, 84]. Despite that, in tumor specimens, neoplastic cells are largely over-numbered by reactive cells making GEP studies difficult[85], AITL has been shown to clearly differ from other PTCL. Also, AITL appeared to derive from follicular helper T-cells as normal counterparts [85], from which the lymphoma diverges, for example, for a lower expression of *CD57* and higher *CD10* expression [86]. Moreover, GEP analysis highlighted that a set of genes involved in vascular biology was upregulated in AITL, notably the vascular endothelial growth factor (*VEGFA*), suggesting a valuable target for therapeutic intervention given that several VEGF inhibitors are under investigation in lymphomas [87].

Extranodal NK/T-cell Lymphoma (ENKTL) is a clinically aggressive entity which represents about 1-2% of all NHL. A variety of cytogenetic aberrations have been reported, but no specific chromosomal translocations have been identified (Table 2). The commonest cytogenetic abnormality is del(6q21-q25) or i(6)(p10), but it is currently unclear whether this is a primary or progression-associated event [88]. Only few reports using array-CGH have been published and many of them did not distinguish between aggressive NK-cell leukemias and extranodal NK/T-cell lymphomas. They have documented chromosome losses of 1p, 4q, 5, 6q, 7q, 11q, 12q, 15q, 17q and gain of 2q, 10q, and 13q [89, 90]. Several investigators have focused on loss of 6q, a region that includes many tumor suppressor genes. Very recent data indicate *PRDM1*, coding for BLIMP1, possibly together with *FOXO3* as the genes inactivated after the 6q21 deletions [91, 92]. In order to better characterize the genomic alterations of extranodal NK/T-cell lymphoma array-CGH technique has been used: gain of 1q was found in both primary cutaneous ENKTL and non-cutaneous ENKTL. Deletions at 9p, region including *CDKN2A* and *CDKN2B*, 12p, and 12q are present only in the ENKTL. A proportion of cases exhibit partial deletion of *FAS* or mutations of *TP53*, *b-catenin*, *KRAS* or *KIT*. Different gene

expression studies [93, 94], evaluating both RNA and microRNAs (miRNAs) and comparing ENKTL and normal NK-cells, have identified a series of deregulated genes and pathways, many of which could represent the rational for targeted therapies: PDGFRα, Aurora-A, Survivin, AKT, JAK-STAT, MYC, TP53 and NF-κB pathways.

Enteropathy-associated T-cell Lymphoma (EATL) is defined as an intestinal lymphoma of intraepithelial T lymphocytes and it has been recently subdivided in two types: EATL Type 1 and EATL Type 2, with the first one strongly associated with celiac sprue and the HLA-DQ2 haplotype. The two types of EATL are associated with similar, but not identical, genetic alterations [95-97] (Table 2): gains of 9q33-q34 and 16q12.1 deletions are frequent in both types and they seem to form a common genetic link between the two subtypes. EATL Type 1 frequently displays 1q22-q44 and 5q gains, while EATL Type 2 is more often characterized by 8q24 (*MYC*) gains. *NOTCH1* or *NEK6* have been proposed as genes possibly affected by the 9q gains[97].

Most **Hepatosplenic T-cell Lymphoma** (HSTL) cases, rare entities with peculiar clinical, morphological and phenotypic features, are characterized by the presence of an isochromosome 7q, which can be the sole karyotypic abnormality, suggesting its primary role in the pathogenesis [98, 99] (Table 2). Trisomy 8 and loss of chromosome Y seem to be associated with progressive disease. A very recent work has reported GEP data obtained from nine hepatosplenic lymphomas highlighting the activation of many, potentially targetable pathways, such as VEGF, MAP kinase, JAK-STAT, mTOR, Notch signaling, cytokine-cytokine receptor interaction, cell adhesion, and also of SYK [99].

A rare, difficult-to-diagnose, and poorly characterized subtype of extranodal cytotoxic T-cell lymphoma is represented by **Subcutaneous Panniculitis-like T-cell Lymphoma** (SPTL). CGH analysis revealed large numbers of DNA copy number changes in this disease [100] (Table 2): the most common alterations were losses of chromosomes 1pter, 2pter, 10qter, 11qter, 12qter, 16, 19, 20, and 22 and gains of chromosomes 2q and 4q. DNA copy number aberrations in SPTL, such as loss of 10q, 17q and chromosome 19, partially overlap with those seen in other cutaneous T-cell lymphomas, such as mycosis fungoides or Sezary syndrome whereas 5q and 13q gains characterize SPTL [100]. Allelic *NAV3* aberrations, previously found in Mycosis Fungoides and Sezary Syndrome, were identified in 44% of SPTL patients.

In summary, systemic and extranodal T-NHL are a very heterogeneous group of disorders with few characteristic genomic lesions. However, a series of possible therapeutic targets are being currently evaluated, also supported by data derived from

gene expression and array-CGH studies (Table 2). Therefore, the use of modern molecular approaches are strongly needed to improve the knowledge of the underlying pathogenesis of the individual T-cell lymphoma subtypes, providing rational bases for identifying further new therapeutic targets and improving the usage of currently available anti- tumor compounds.

Table 2. **Summary of genomic studies in T-cell malignancies.**

	study	type of study	year	platform	num of cases	ref
PTCL, NOS						
	Piccaluga PP	GEP	2007	Affymetrix HG-U133 2.0 plus	28 PTCL, NOS+ 6 AITL, 6 ALCL and 20 normal T cells	[83]
	Zettl A	CGH	2004	Vysis	42 PTCL, NOS and 37 ALCL	[80]
	Fujiwara SI	array-CGH	2008	Affymetrix Mapping 50k Hind 240 arrays	33 PTCL, NOS and 40 AITL	[81]
AITL						
	Piccaluga PP	GEP	2007	Affymetrix HG-U133 2.0 plus	28 PTCL, NOS+ 6 AITL, 6 ALCL and 20 normal T cells	[85]
	Iqbal J	GEP	2010	Affymetrix HG-U133 2.0 plus	36 ITL + 28 ALCL + 12 ATLL + 14 T/NKCL + 44 PTCL, NOS + 10 other rare PTCL entities + 16 cell lines	[86]
	Fujiwara SI	array-CGH	2008	Affymetrix Mapping 50k Hind 240 arrays	33 PTCL, NOS and 40 AITL	[81]
ENKTL						
	Huang Y	GEP	2010	Affymetrix HG-U133 2.0 plus	7 primary and 2 cell lines	[93]
	Karube K	GEP and array-CGH	2011	Agilent Human Genome 4x44k Oligonucleotide array (GEP); Agilent Human Genome CGH 44A microarray (array-CGH)	35 ENKTL	[91]
	Nakashima Y	array-CGH	2005	BAC/PAC array custom-made	27 ENKTL	[89]
	Iqbal J	GEP and array-CGH	2009	Affymetrix HG-U133 2.0 plus (GEP); BAC array custom-made (array-CGH)	8 cell lines and 7 primary ENKTL	[90]
EATL						
	Deleeuw RJ	array-CGH	2007	BAC array custom-made	30 EATL	[95]
	Ko YH	array-CGH	2010	BAC/PAC array custom-made	5 EATL	[96]
HSTL						
	Travert M	GEP and array-CGH	2012	Affymetrix HG-U133 2.0 plus (GEP); Agilent Sureprint CGH array 4x180k (Array-CGH)	9 HSTL and cell lines	[99]
SPTL						
	Hahtola S	array-CGH	2008	metaphase chromosome from normal individuals	9 SPTL	[100]

2.2 Next Generation Sequencing

Next generation sequencing (NGS, also known as massive parallel sequencing or as second-generation sequencing) technologies are revolutionising the ability to characterize cancers at a genomic, transcriptomic and epigenetic levels. For the past 15 years, Sanger sequencing (capillary-based DNA sequencing or first-generation sequencing, first developed in 1975 [101]) and fluorescence-based electrophoresis technologies have been extensively used in somatic and germline genetic studies. Modern Sanger sequencing used automated instruments that detect fluorescently labeled nucleotide sequences. Capillary electrophoresis resolved DNA strands of increasing length during DNA elongation where single-stranded amplified DNA template is randomly terminated by incorporating fluorescent dideoxynucleotides. Nevertheless, Sanger sequencing presented a relatively restricted sequencing capacity, limiting the use of this approach in case of single gene or hot-spot sequence analysis. Improvements in instrumentation coupled with the development of high performance computing and bioinformatics have allowed the progress of massive parallel sequencing, overcoming the limited scalability of traditional Sanger sequencing. Many different high-throughput technologies for massive parallel DNA sequencing emerged in late 1996, and became commercially available since 2005. Different engineering configurations and sequencing chemistry characterized different NGS platforms. However, they share the technical paradigm of massive parallel sequencing via spatially separated, clonally amplified DNA templates. All approaches provide three critical steps: DNA sample preparation, immobilization, and sequencing. Generally, the first part of the technique implies the preparation of a "sequencing library", which involves the fragmentation of DNA and the addition of defined sequences, known as “adapters,” to the ends of randomly fragmented DNA. The addition of adapters is required for two main reasons: first, to anchor the DNA fragments of the sequencing library to a solid surface and then, to define the site in which the sequencing reactions begin. The amplification of the sequencing library DNA is required only for some high-throughput sequencing systems in order to form spatially distinct and detectable sequencing features. Amplification can be performed *in situ*, in emulsion or in solution to generate clusters of clonal DNA copies. Sequencing is performed using either DNA polymerase synthesis for fluorescent nucleotides or the ligation of fluorescent oligonucleotides. In the physical sequencing phase, the individual bases in each fragment are identified one after the other in the right order. Perhaps more important than the

sequencing throughput provided by this technology and its relative low cost compared with traditional sequencing methods is the type of data it generates. The result of a sequenced segment is called a "read". Each nucleotide of the regions studied will be included in many reads, and, therefore, repeatedly analyzed. In the re-assembly phase, bioinformatics tools are used to align overlapping reads, which allows the original genome to be assembled into contiguous sequences. The longer the read length, the easier it is to reassemble the genome. The million of analyzed reads are then compared with a reference human genome. Instead of long reads generated from a PCR-amplified sample, massively parallel sequencing methods provide much shorter reads (~21 to ~400 base pairs), but millions of them. The short reads generated in the sequencing of each DNA molecule can be counted and quantified, allowing the identification of mutations, the quantification of its frequency and accurate copy number assessment of each genomic region. In addition, with the introduction of approaches that allow for the sequences of both ends of a DNA molecule (that is, paired end massively parallel sequencing or mate pair sequencing), it has become possible to detect balanced and unbalanced somatic rearrangements, such as fusion genes, in a genome-wide fashion in a single assay [102]. Nevertheless, PCR amplification of DNA to be analyzed represents a limit of second generation sequencing techniques. In fact the possible introduction of base sequence errors or the promotion of the use of certain sequences over others, may lead to the identification of mutations not tumor-related and to the change of the frequency and abundance of various DNA fragments that existed before amplification [103]. The sequencing from a single DNA molecule (without the need for PCR amplification and its potential for distortion of abundance levels), together with the ultimate miniaturization and the minimal use of biochemicals can overcome this limit. This method of sequencing from a single DNA molecule is called "third-generation sequencing" [104]. Numerous are the advantages of these very recent approaches over second NGS techniques: higher throughput, faster turnaround time, longer read lengths, higher consensus accuracy to enable rare variant detection, small amount of starting material and low cost [103]. Many of the third generation sequencing platforms integrate DNA sequencing with other more general utility, such as comprehensive characterization of transcriptomes and translation and identification of patterns of methylation. High-throughput sequencing technologies can be used also to sequence cDNA, in an approach called "RNA-seq", in order to get informations about differential expression of genes, including gene alleles and differently spliced transcripts, non-coding RNAs, post-transcriptional mutations or editing and gene

fusions. Besides RNA-seq, high troughput sequencing techniques have a wide range of applications, such as: chromatin immunoprecipitation coupled to DNA microarray (ChIP-chip) or sequencing (ChIP-seq), whole genome genotyping, genome wide structural variation and sequencing of mitochondrial genome.

Advances in the field of genomics, and especially in sequencing approaches, have broadened the horizons and expectations towards a new era of personalized therapy.

3. ANAPLASTIC LARGE CELL LYMPHOMA

In 1985, Harald Stein and collegues [105] identified a subset of NHLs characterized by anaplastic cytology, an unusual sinus growth pattern and strong expression of the antigen Ki-1 recognized by an antibody developed in Kiel, West Germany [106]. Subsequently, Ki-1 was identified as an activation antigen, now designated CD30 or TNFRSF8. The authors recognized, at the time, that Ki-1 lymphoma as described in this first publication was not a homogeneous entity.

In the late 1980s and early 1990s, a recurrent t(2;5) was observed in Ki-1 lymphomas and described [107-109], and some years later, in 1994, the translocation was cloned by Steve Morris and others [110] and the genes involved, *ALK* on 2p23 and *NPM* on 5q35, were identified.

Anaplastic Large Cell Lymphoma (ALCL) was first included as an entity in the Revised European and American Lymphoma (REAL) classification in 1994, and, subsequently, in the World Health Organization (WHO) classification of lymphoid neoplasms in 2001. The new edition of WHO classification (2008) [1] recognizes, within the spectrum of mature T-cell neoplasms, two types of systemic ALCL according to ALK protein expression in tumor samples:

- a distinct entity, named ALK+ ALCL, which is characterized by *ALK* gene rearrangements and consequent ALK protein expression;
- a provisional entity, the so-called ALK- ALCL, which cannot be distinguished morphologically from ALK+ ALCL but differs from this entity because of the lack of ALK protein.

3.1 Clinical Features

Clinically, ALCL can be distinguished in:

1. primary or *de novo* ALCL,
2. secondary ALCL, if the lymphoma arises from a preceding lymphoproliferative disorders such as Lymphomatoid Papulosis, Mycosis Fungoides or Hodgkin Disease, or it can be identified in the setting of HIV infection.

Primary ALCL can be further divided into two distinct forms, namely, primary systemic ALCL and primary cutaneous ALCL (cALCL) [111].

Primary systemic ALCL affects patients with a median age lower than that of the most other forms of NHL with a typical bimodal age distribution, with a larger peak in the

second and third decades of life (median age at diagnosis 10.2 to 11 years), and a smaller peak in the sixth and seventh decades. ALK status seems to follow the bimodal age distribution, with a strong correlation between the presence of one of the ALK translocations and the younger age of patients; conversely, older patients present the ALK- subtype of the lymphoma. ALCL represents approximately 10% to 15% of pediatric/adolescent NHLs and only 2% of adult NHLs [112]. Male predominance is typically present.

Approximately 60% to 70% of patients have advanced stage III/IV disease at diagnosis due to peripheral and abdominal lymphoadenopathy. Extranodal disease is frequent (almost one third of cases) with skin, bone, lung, liver and soft tissue being the most common sites. Bone marrow involvement is uncommon as well as involvement of central nervous system and the gastrointestinal tract. Recently, in adults, breast involvement has been described, with most cases being ALK- ALCL and frequently in association to breast implants [113]. Patients with peripheral blood involvement, though its rarity, have a poor prognosis irrespective the ALK status.

Consistent differences in terms of 5-year survival between ALK+ and ALK- cases have been reported in numerous studies, with ALK+ ALCL showing a significantly improved overall survival. The 5-year overall survival rate in ALK+ ALCL is approximately 80%, in contrast to only 48% in ALK- ALCL as shown in Fig 6 [114, 115].

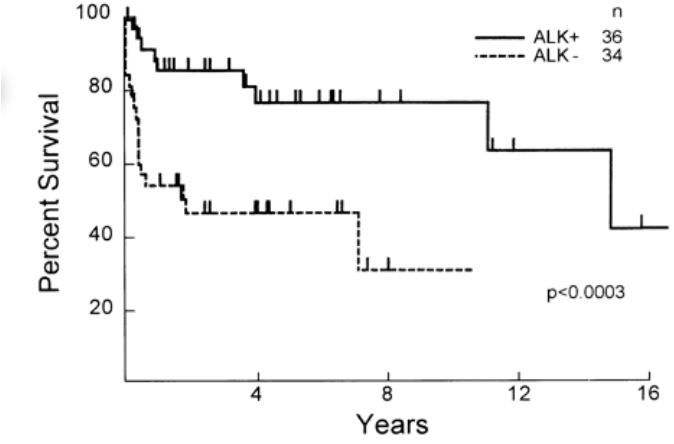


Figure 6. Overall Survival curve of 36 ALK+ ALCL and 34 ALK- ALCL patients [115].

The favorable prognosis of ALK+ ALCL may be somewhat related to the younger age of patients or to the well characterized ALK biology, as there is no significant outcome difference, both in terms of free failure survival (FFS) and overall survival (OS) (Fig 7), among ALK+ and ALK- patients if stratified according to the median age at diagnosis [116].

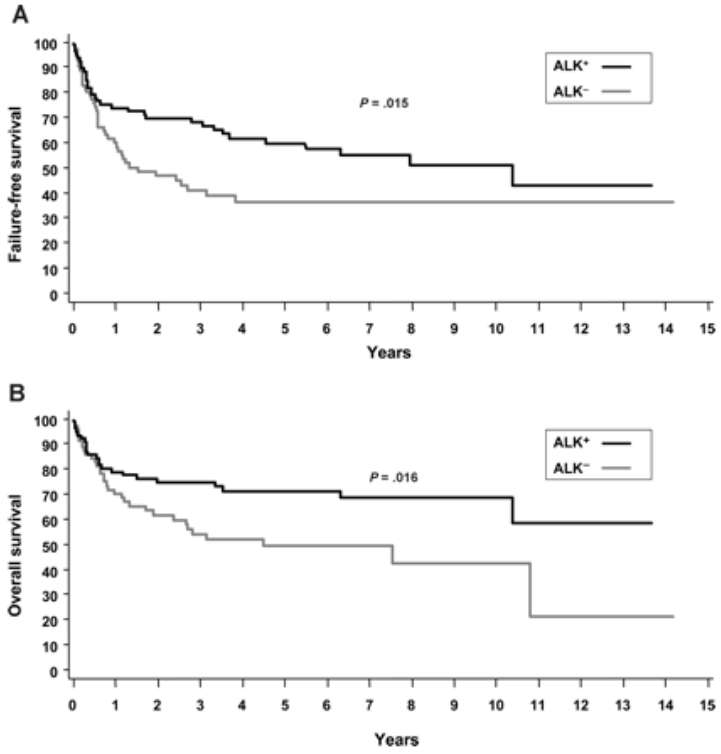


Figure 7. **FFS (A) and OS (B) of 87 ALK+ ALCL and 72 ALK- ALCL [116].**

Patients affected by primary cutaneous ALCL, presenting localized skin lesions, have distinct clinical features: they are older, with a median age at diagnosis in the seventh decade, male predominace, and the presence of red solitary nodules often accompained with ulceration. Approximately 20% of patients experience spontaneous regression of their primary lesion. However, half of the patients show recurrences, most often limited to the skin in the same anatomic region as the primary lesion. About 20% of patients go on to develop systemic lymphoma. Overall, patients with cALCL have an improved survival compared to those with systemic disease [111].

3.2 Histologic Features

According to the current WHO classification, ALCL is not sustained by a unique hystotype but actually includes five morphologic variants: common, giant cell-rich, small-cell type, lympho-histiocytic and Hodgkin-like (Fig 8) [117]. Indeed, common ALCL variant (Fig 8, panel A) corresponds to the description of the tumor given by Stein and coworkers in 1985 [105]. The "hallmark cells", with large, eccentric nuclei with a characteristic horseshoe, kidney shape, are present to a certain degree in all the histologic variants of ALCL.

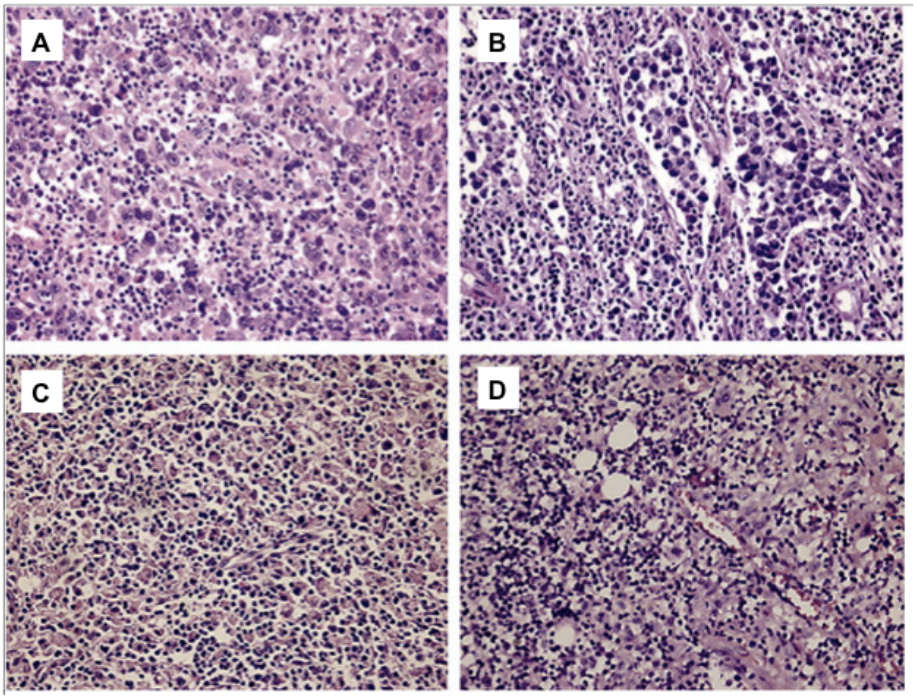


Figure 8. Morphology of primary systemic ALCL. *(A) a common histologic variant showing a diffuse and cohesive pattern (H&E, 400X); (B) a common histologic variant, showing sinusoidal involvement (ALCL-HD like) (H&E, 400X); (C) a lymphohistiocytic variant (H&E, 400X); (D) a small cell variant (H&E, 400X). Figure taken from Huang W et al, Exp and Mol Pathol 2009, [118].*

In the classical variant, which represents approximately 70% of all ALCL cases, the neoplastic cells grow cohesively and preferentially involve lymph node sinuses, particularly in lymph node not involved extensively. With greater involvement, ALCL replaces the paracortical regions or may diffusely replace lymph node architecture. Cells are anaplastic, large, bizzarre and irregularly shaped and often have a polylobated nuclei; the cytoplasm is abundant and often basophilic. In the giant cell-rich type, a large number of the tumor cells contain more than one nucleus, while the small cell variant (Fig 8, panel D) is characterized by a mixture of small, medium-sized, and large lymphoid cells, with nuclei being often irregular. This variant may, however, contain areas with the morphology typical of the ALCL common type, such as sheets of CD30+ blasts, and can transform into the common type and vice versa. Few neoplastic cells associated with numerous reactive lymphocytes and histiocytes compose the lympho-histiocytic variant (Fig 8, panel C), the second most common morphologic variant, identified in 5-10% of ALCL cases [119]. Finally, the so-called ALCL of the Hodgkin-like type deserves special attention. The morphologic and immunophenotipic features of ALCL and Hodgkin disease (HD) are very similar, making often difficult the diagnosis between the two different entities. Few decades ago, investigators created a category under the term of ALCL-HD related in which borderline cases, or gray-zone cases could be collected. The tumors falling in this category show features of both ALCL and HD: these cases contain relatively dense nodules or sheets of tumor cells with characteristic similar to those present in HD but usually present in sinuses as in ALCL (Fig 8, panel B). The change from the term ALCL-HD related to ALCL-HD like reflect the tendency to believe that these grey-zone lymphomas represents ALCL mimicking HD [120].

3.3 Phenotypic Features

Irrespective to the histiotype, neoplastic cells of ALCL carry a distinctive phenotypic profile. They commonly show a strong CD30 expression on the cell membrane and in the Golgi region [121]. CD30, which represents a member of the receptor superfamily of tumor necrosis factor (TNF), is a glycoprotein of 120 kDa carried by lymphoid elements following activation and composed by three different domains, intracytoplasmic, transmembranic and external. The external domain of CD30 is cleaved by a metalloproteinase so that it can be detected and measured in the serum. Notably, CD30 overexpression has been reported to induce NF-κB activation [122]. Nevertheless, the

positivity for CD30 is not specific for this type of lymphoma, being expressed also by Reed-Sternberg cells of HD and DLBCL cells. Positivity for TIA-1, granzyme B and perforin is recorded in about 85% of ALCL cases, while the epithelial membrane antigen (EMA) is carried by 60%-70% of cases. Approximately two-thirds of ALCL cases show a T-cell phenotype, 10%-20% show a B-cell phenotype, while the remainder express neither T- or B-lineage marker (null-cell phenotype): the current WHO classification only recognize T/null-cell ALCL as a clinicopathological entity, and include B-cell ALCL with all other DLBCL cases. In fact, CD3 expression is appreciated in only half of the cases and it usually occurs at the cytoplasmic level as expected in activated cells. Others T-cell associated antigens can be explored, as CD2, CD5 and CD7; moreover, the expression of CD4 and CD8 is variable.

In order to help the differentiation between ALCL and HD or DLBCL, the search of BSAP, the *PAX5* gene product, has proved to be a very useful tool, being present in HD and DLBCL cases and not in ALCL ones. Nevertheless, a recent study reported the observation of the positivity for PAX5 staining in some few cases of ALCL [123]. The development of antibodies against ALK has further refined the immunoistochemical analysis of ALCL: it could be useful in the differential diagnosis between ALCL and HD, since ALK is not detected neither in normal lymphocytes, nor in Hodgkin lymphoma [117].

3.4 Molecular Genetic Features

Independently from the expression of ALK, ALCL share a cluster of transcripts which allow their stratification and distinction from other T-cell lymphomas, suggesting a common signature for both ALK+ and ALK- cases and, possibly, an unique origin for both subtypes [86, 124]. However, although ALK+ ALCL and ALK- ALCL are largely overlapping in terms of their gene expression profiles, different groups have reported a series of genes with a differential expression between the two lymphoma subtypes [79, 86, 124-126]. Excluding ALK, BCL6, PTPN12, CEBPB, SERPINA1 and GAS1 have been reported as the most discriminating genes between the two subsets [124, 126]. Moreover, the transcripts overexpressed in ALK+ tumors comprised CCND3 and genes involved in immune response, NF-κB pathway, leucocyte transendothelial migration, focal adhesion and signal transduction (SYK, LYN, CDC37) [125]. Finally, in another study [86], among the top-ranked genes were ALK, TNFRSF8 (CD30), MUC1, TH17-cell-associated molecules *(IL-17A, IL-17F, GATA3, RORC*), and a small group of

immunoregulatory cytokines/receptors regulating STAT3 (IL-26, IL- 31RA) or JAK3 (IL-9) activation. ALK+ ALCL also showed a lower expression of transcripts related to TCR components and TCR signaling or activation, alongside a higher expression of the cytotoxic molecules *GZMB* and *PRF1* [86]. Conversely, ALK- ALCL showed greater expression of a set of cytokine/receptors (CCL1, CCL22, CCR8, CCR4, IL-13RA2, CXCL14, TGFBR1), of the transcription factor genes HOXC6 and HOXA3 and several antiapoptotic factors (BCL2, BIRC6, BIC) but low expression of certain proapoptotic genes (BAX, BCL2L1, BNIP3) [125, 126].

Chromosomal abnormalities have been identified in both ALCL subtypes, with the most common aberration being translocations involving *ALK* gene in the ALK+ cases (Fig 9). Some lesions are common among ALK+ ALCL and ALK- ALCL, but others are so different to suggest a possible different pathogenesis. However, chromosomal abnormalities in ALCL will be widely described in the following paragraphes.

The distinction between ALCL subtypes can be improved through the analysis of differentially espressed miRNAs, used as molecular classifiers. Recently, microRNAs have emerged as tiny but potent molecules able to regulate cell differentiation and proliferation: they are noncoding RNA (18-22 nucleotides long) that regulate the expression of their target proteins through binding at the 3'-UTR regions by translational inhibition. Recently, Merkel and collegues [127] identified a distinct profile of miRNAs that characterize ALCL; furthermore this profile distinguishes ALK+ cases from ALK- ones. Despite the high concordance in miRNAs expression between ALK+ and ALK- ALCL suggests some similarities between the two disease entities, some of them have been recognized as differentially espressed: miR-155 is overexpressed in ALK- specimens, while different members of miR-17-92 cluster (miR-886-3p, miR-20b, miR-106a, miR-17 and miR-20a) are overexpressed in ALK+ cases. Interestingly, miR-101, which directly targets the mTOR transcript, was commonly deregulated in both subsets. Its forced re-expression attenuates cell proliferation in ALK+ cellular models, suggesting that NPM-ALK-driven cell proliferation is at least partially dependent on miR-101 target proteins.

Taken together, molecular data hint that, despite the presence of a gene expression profiling and miRNAs expression in part overlapping among the two subsets, they are different entities because they show different chromosomal aberrations. It is plausible to think that both subsets share a common origin, but different molecular events happen

over time and ALCL cells could encounter a different pathogenetic process, resulting in the end in the manifestation of ALK+ ALCL or ALK- ALCL.

3.4.1 ALK gene and Translocations

A number of different receptor tyrosine kinases (RTKs) are implicated in oncogenesis, besides their vital roles in the transmission of extracellular cues that control diverse cellular functions, including proliferation, survival and differentiation. Anaplastic Lymphoma Kinase (*ALK)*, also known as CD246, belongs to the insuline receptor superfamily together with IGF-1 and MET, with a role in both neural development and oncogenesis [128]. *ALK* was originally identified as a result of the cloning of the fusion gene NPM-ALK observed in ALCL [110]. An association between ALCL and the t(2;5)(p23 ;q35) chromosomal rearrangement was reported in the late 1980s [108, 109], and the genes reported in this translocation were identified in 1994 as those encoding nucleolar protein nucleophosmin (*NPM*) at 5q35 and the novel *ALK* RTK at 2p23 [110]. *ALK* gene is composed by 26 exons; the 6226 bp cDNA encodes for a 177 kDa polypeptide that, after post-translational modifications such as *N*-glycosilation, becomes a mature ALK RTK of approximately 200-220 kDa. As with other RTKs, ALK possesses an extracellular ligand-binding region, a transmembrane-spanning domain and a cytoplasmic kinase catalytic region. The extracellular region includes several motifs, including a 26-aa N-terminal signal peptide sequence, as well as the binding site (located at residues 391-401) for the endogenous ligands reported to activate ALK, pleiotrophin (PTN) and midkine (MK) [129, 130]. The 28-aa transmembrane domain is followed by a 64-aa juxtamembrane domain which contains a binding site of insulin receptor substrate-1 (IRS-1), facilitating a tyrosine phosphorylation-dependent interaction. The Kinase Domain (KD or TK) in the intracellular part includes a three-tyrosine-containing motif (Y1278, Y1282 and Y1283) within its activation loop, representing the major self-phosphorylation site of the insulin receptor superfamily. The 244-aa ALK C-terminus contains a phosphotyrosine-dependent binding site for the substrate protein Src homology 2 domain-containing (SHC), as well as an interaction site for phosphotyrosine-dependent binding of phospholipase C-γ (Fig 9) [128, 131].

In normal tissue, spatiotemporal expression of ALK is restricted, peaking during the prenatal period and decreasing rapidly after birth. In human beings, ALK mRNA is predominantly transcribed in the brain and intestinal nervous tissue, scattering in the testis, placenta and fetal liver. The expression of the putative ligand PTN is similar to that

of ALK, with high levels in the nervous system during fetal development and decreased expression after birth. The other putative mammalian ALK ligand, MK, was originally described as a retinoic-acid inducible, developmentally regulated, heparin-binding neurotrophic factor that shows 45% identity to PTN [128]. Functions attributed to MK and PTN are similar, including the direction of neurite connections, effects upon neuronal migration and a possible role in angiogenesis. MK is expressed in the brain as well as in other organs, with higher levels generally at midgestation and, similar to PTN, downregulation at birth. PTN, MK and ALK appear to play a role in the regulation of cell survival: some studies [132] have reported that ALK, at least in part, could regulate the already known activity of PTN and MK in the ragulation of apoptosis.
Other hypothetical ligands have been considered: OSF-1 (osteoblast-specific factor-1), HARP (heparin affinity regulatory peptide) and HBNF (heparin-binding neurotrophic factor). Nevertheless the role of these molecules remains controversial. ALK has also been proposed to be a dependence receptor [133] : dependence receptors induce apoptosis in their non-ligated state, but suppress apoptotic signalling in response to ligand binding.

The full length ALK TRK is known to be expressed by several tumor types, both in cell lines and in primary specimens, such as neuroblastomas, neuroectodermal tumors, glioblastomas and melanoma. In addition, anti-ALK immunoreactivity has been observed in tissue immunostaining studies of other malignancies as breast carcinoma, malignant peripheral nerve-sheath tumors and lipogenic tumors [128]: unfortunately, the pathogenic significance of the presence of the full length ALK receptor in the aforementioned tumors remains uncertain. It has been postulated that the growth of tumors could be driven by an autocrine and/or paracrine loops involving PTN and/or MK, but, besides some preliminary studies on glioblastoma [134], this has not yet been unequivocally proven for any tumor type. The full length ALK protein has been reported to be expressed also in a rare subtype of poor-prognosis B-cell NHL that probably comprises less than 0.5-1% of all B-cell lymphomas [135]. However the same group, some years later described the presence of some ALK fusions, coexpressed with full length ALK in these NHLs.
As other RTKs, even ALK has been implicated in oncogenesis due to genetic abnormalities ranging from point mutations, gene amplification, or fusion to heterologous genes; mechanisms that results in the activation of the kinase catalytic domain. This kinase activation induces growth factors-independent proliferation, cellular transformation

and anti-apoptotic signaling. However, the most common mechanism of constitutive ALK activation currently known to occur in cancer involves chromosomal translocations that interrupt *ALK* gene at 2p23 and fuse it with another gene (Fig 9), resulting in the creation of oncogenic *ALK* fusion genes and proteins [128, 136]. The most common ALK fusion protein found in ALCL, and the first one identified in this lymphoma, is NPM-ALK resulting from the translocation t(2 ;5)(p23 ;q35) (Fig 9) [110]. This type of chimera, as well as several of the so-called "variants" are reported in Fig 9.

All ALK fusion proteins share some critical features:

- the same region of ALK (aa 1058-1620 of the full length normal receptor) that comprises the complete intracellular segment, including the TK catalytic domain (most breaks occur in the intron between exons 16 and 17);
- an N-terminal partner protein that is normally wide expressed in normal cells, the gene promoter of which controls the aberrant expression of the ALK chimeric mRNA and the expression of its encoded fusion protein;
- the presence of an oligomerization domain of some type in the sequence of the ALK fusion partner, which mediates constitutive self association of the ALK fusion and, in turn, causes constant ALK KD activation.

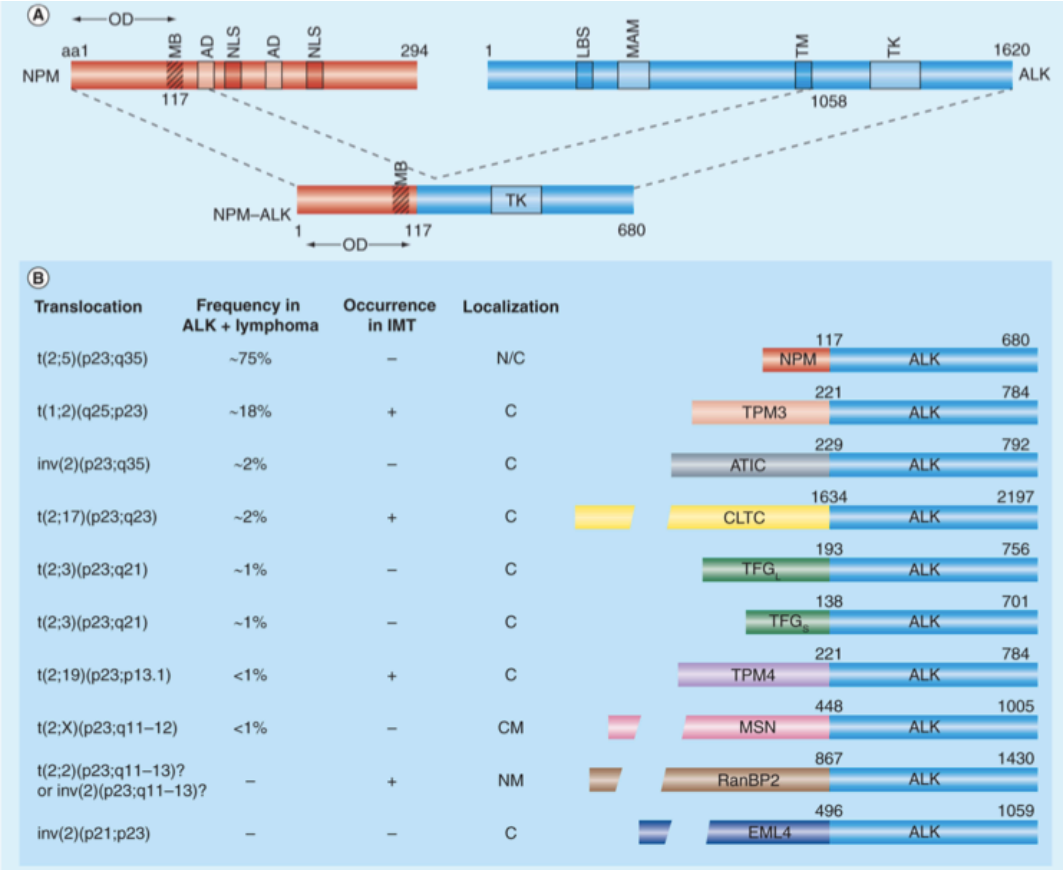


Figure 9. ALK fusion proteins. *(A) NPM-ALK fusion protein deriving from t(2;5)(p23;q35). Fusion of the chromosome 5 gene encoding NPM to the chromosome 2 gene encoding ALK generates the chimeric tyrosine kinase, NPM–ALK. NPM contains an OD (residues 1–117) a putative MB (residues 104–115), two ADs (Asp/Glu-rich acidic domain; residues 120–132 and 161–188) that function as acceptor regions for nucleolar targeting signals and two NLS (residues 152–157 and 191–197). ALK contains a single MAM domain, a region of approximately 170 aa homologous to the extracellular portions of a number of functionally diverse proteins that may have an adhesive function (residues 480–635). The ligand-binding site (LBS) for pleiotrophin and midkine (ALK residues 391–401) is indicated. OD: oligomerization domain, MB: metal-binding domain, AD: acidic amino acid domain, NLS*: nuclear localization signal, *TK: tyrosine kinase domain. (B) "Variants" translocations involving ALK gene: chromosomal rearrangement, frequency in ALK+ lymphomas, occurrence in Inflammatory Myofibroblastic Tumor (IMT) and localization of the chimera (C: cytoplasmic; N: nuclear; CM: cell membrane; NM: nuclear membrane) are reported [128].*

Oligomerization of ALK fusion proteins via these domains mimics the ligand-mediated aggregation and activation of the full length wild type ALK receptor, constitutively activating the chimera and continuosly transmitting growth-promoting cellular signals. ALK fusion proteins not only self-associate, they also hetero-associate with the normal counterpart of their partner proteins; consequently, the localization of ALK fusions in tumor cells is determined largely by the normal subcellular distribution of their partner proteins. This activation in the fusion proteins leads to the phosphorylation and aberrant activation of multiple downstream substrates and signaling cascades involved in the neoplastic transformation of cells.

As already said, the most common ALK fusion protein in NHL is NPM-ALK, being found in approximately 75-80% of ALK+ ALCL (Fig 9) [137]. NPM1 encodes a 23 kDa multifunctional protein that is involved in the transport of pre-ribosomal particles and ribosome biogenesis, regulation of cell division, DNA repair, transcription and genomic stability. The NPM protein includes a nucleolar localization signal favouring the nuclear and nucleolar cellular localization of the chimera.

In addition to NPM-ALK, other fusions involving ALK have been described, as reported in Fig 9. Among them, other chromosomal translocations identified involve *TPM [138]*, *TFG [139]*, *CLTC* [140], *ATIC* [141], *TSPYL2 [140]*, *MSN [142]*, *RanBP2 [143]*, *EML4 [144]*, *KIAA1618 [145]*, *MYH9 [146], KIFB5 [147], CARS [148], SEC31L1 [149]* as *ALK* partner. The presence of the aforementioned chimeras has been observed in ALCL as well as in other lymphoid and non-lymphoid neoplasms, such as a rare subtype of DLBCL (named ALK+ DLBCL), Inflammatory Myofibroblastic Tumors (IMT), Non-Small-Cell Lung Cancers (NSCLCs), colon cancers, breast cancer, renal cell carcinomas, esophageal squamous cell carcinoma and extramedullary plasmacytomas. The discovery that EML4-ALK fusion is observed in 6% of NSCLC has opened a new scenario for ALK targeted therapy, one of the most attractive and promising story of the last decade: this has lead to the development of different ALK inhibitors, that in some cases, are currently under investigation, with the Crizotinib as the first approved drug for the treatment of these well-defined subsets of NSCLC patients. However, some of the ALK translocations, such as NPM-ALK and ATIC-ALK transcripts, have been detected in cells of healthy individuals with relatively high frequency, suggesting that this aberration could be a common event, necessary but not sufficient for the neoplastic transformation of cells [150]: other molecular events might be required for lymphoid cells to be fully transformed and lead to ALCL.

3.4.2 ALK signalling in ALCL

Fusion proteins deriving from translocations involving *ALK* gene, have a clear oncogenic potential: the aberrant constitutively activation of the tyrosine kinase domain enhances cell proliferation and survival leading also to cytoskeletal rearrangements and changes in cell shape. ALK activates different well-characterized pathways through the interaction with downstream molecules that trigger intracellular signalling cascades. In particular, the Ras-ERK (Fig 10) pathway is essential mostly for ALCL proliferation, whereas the JAK3-STAT3 pathway (Fig 11) and the PI3K-Akt (Fig 12) pathway have been shown to be vital primarily for cell survival and phenotypic changes [136]. Most of the knowledge we have about ALK signalling came from studies on NPM-ALK fusion protein, but it is conceivable to think that all the other variants could act on the same downstream pathways and with the same mode of action assuming that dimerization occurs.

The increased growth of ALCL is mainly attributed to the activation of Ras-ERK pathway. In ALCL cells the chimera functions as a docking molecule for several downstream adaptors or scaffolding molecules such as IRS-1, SHC and SRC, which bind to the fusion protein at level of specific tyrosine residues of ALK (Fig 10). These mediators are able to activate Ras, which in turn phosphorylates ERK1 and ERK2. The SHP2/GRB2 complex interacts with ALK and SHC to enhance phosphorylation of ERK family members through SRC. Activation of phosholipase C-γ (PLCγ) is also thought to contribute to the fusion protein-mediated transformation. PLC-γ is directly bound and activated by ALK. This association leads to the formation of diacylglycerol (DAG) and inositol triphosphate (IP3), which, in turn, mobilize calcium stores from the endoplasmic reticulum and activate protein kinase C (PKC). Activated ERK signalling pathway and phosphorilated JUN N-terminal kinase (JNK) with its downstream targets (JUN and AP1), result in uncontrolled cell-cycle progression and cell growth due to the downregulation of p21 and concomitant upregulation of cyclin D3 and cyclin A. As reported in Fig 10, ERK signalling pathway is also able in turn to activate mammalian target of rapamycin (mTOR) which leads to the phosphorylation of the mTOR targets with the final end to promote cell-cycle progression [136].

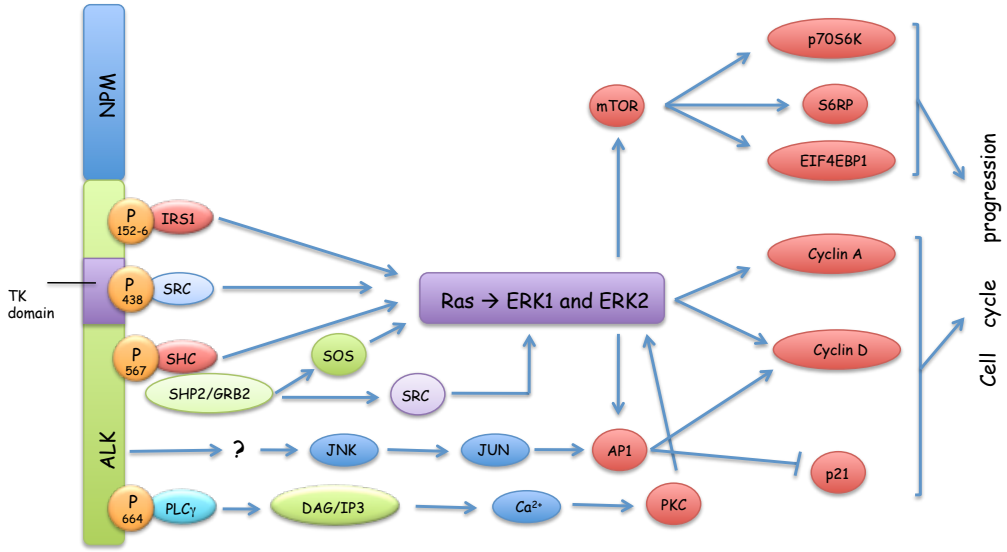


Figure 10. ALK and Ras-ERK pathway. *NPM-ALK is depicted as example for an oncogenic fusion protein involving ALK.*

Survival in ALCL is promoted by the activity of the key player Signal Transducer and Activator of Transcription 3 (STAT3) (Fig 11). STAT3 can be directly phosphorylated and so activated by the fusion protein or, the fusion transcript can activate JAK3, which in turn can contribute to STAT3 activation as shown in Fig 11. SHP1 is reported to be a negative regulator by dephosphorilating ALK, but in ALCL it is lost owing to gene methylation, resulting in the enhancement of STAT3 activation. In these cells, STAT3 allows the enhanced transcription of both anti-apoptotic factors and cell cycle regulators such as BCL2, BCL-XL, C/EBPβ, Survivin, MCL1 and different cyclins. The transforming property of the chimeras seems to be mediated also by STAT5B, although the mechanism of activation in this case requires JAK2. On the contrary, STAT5A has been reported to act as a tumor suppressor in ALCL, where it is epigenetically silenced and, if re-expressed, in turn an inhibit fusion proteins expression [136].

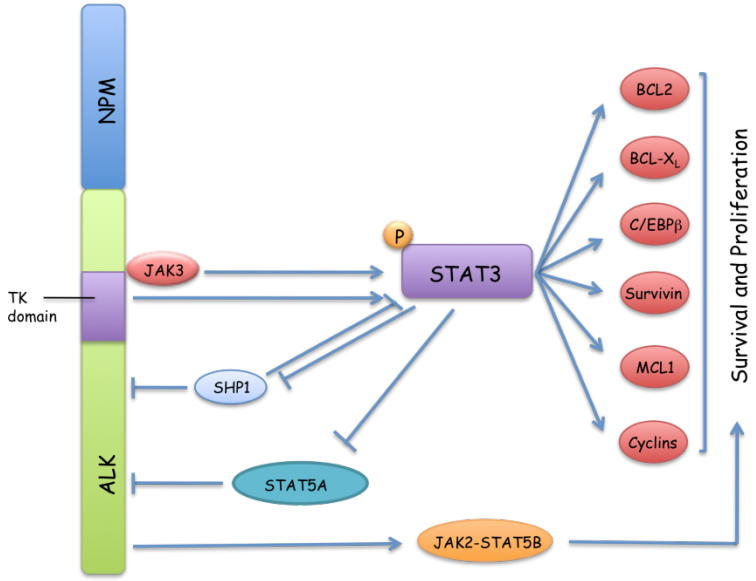


Figure 11. ALK and JAK3-STAT3 pathway. *NPM-ALK is depicted as example for an oncogenic fusion protein involving ALK.*

NPM-ALK or another variant translocation binds to and activates PI3K through the regulatory p85 subunit of PI3K in ALCL cells, leading to the phosphorylation of its downstream effectors AKT1 and AKT2, which in turn enhance the survival of cells by blocking the function of pro-apoptotic proteins, such as BAD (Fig 12). AKT1 and AKT2 are able to phosphorylate FOXO3A, thus blocking FOXO3A-mediated transcription of target genes that promote apoptosis, cell-cycle arrest (p27 and Cyclin D2) and metabolic processes (BIM), forcing in this way G1 phase cell-cycle block. mTOR can sinergistically act promoting the survival pathway, leading to the transcription of its target genes such as p70S6K, S6RP and EIF4EBP1 (Fig 12) [136].

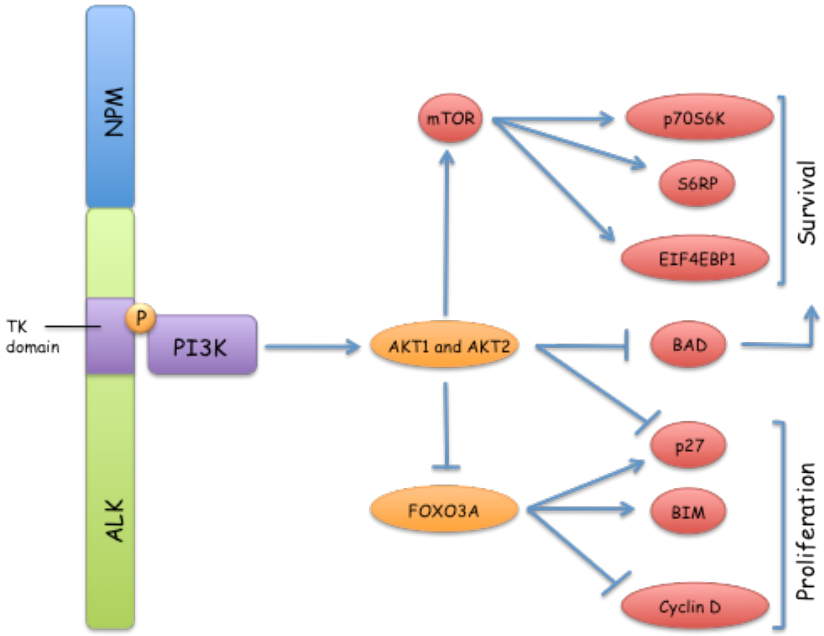


Figure 12. ALK and PI3K-Akt pathway. *NPM-ALK is depicted as example for an oncogenic fusion protein involving ALK.*

It has been reported that NPM-ALK could also activate c-Src kinase, a tyrosine kinase receptor that plays a relevant role in cell migration, as well as in cell proliferation and growth. The chimera following its association with a specific tyrosine residue in position 418 activates c-Src. Src family kinases may also contribute to the sustained activation state of Cdc42 in ALCL cells, responsible of the regulation of the shape and migration of ALCL cells [151].

Many of the members involved in the activated pathways could be potential therapeutic targets: treatment of ALCL cells with specific inhibitors (Mek, ERK, Ras, mTOR, STAT3, JAK3, Akt) or RNA interference leads to a cell-cycle arrest, increase of apoptotic cells rate and proliferation arrest. It is plausible to think the clinical use of a combination of these drugs together with specific anti-ALK small molecules in order to obtain a synergistic effect or to use them in case of acquired resistance to ALK inhibitors.

3.4.3 Secondary Genomic Aberrations

Besides the well-defined primary chromosomal translocations, ALK+ ALCL carries frequent secondary chromosomal imbalances including losses of chromosome 4 (4q13-q28), 6q (6q13-q22), 11q (11q14-q23) and 13q (13q21-q31; 13q32-q33), and gains at 7p11-pter and chr17 [79, 80, 152-154]. Interestingly, the 17p13 locus, encompassing *TP53,* has been reported as a target of both losses and gains. This represents a unique example among lymphomas and tumors in general and, in particular, gains of 17p, together with losses at chromosome 4, are uncommon among other types of T-cell lymphomas, including ALK- ALCL, making them a characteristic of ALK+ ALCL [80, 155]. So far, no genes have been clearly identified as altered by these lesions.

3.4.4 ALK- ALCL Subset

In the current WHO classification, ALK- ALCL is included as a provisional entity, because of the lack of clear distinct features. It is defined as a CD30+ peripheral T-cell neoplasm that is not reproducibly distinguishable on morphological grounds from ALK+ ALCL, but lacks the ALK protein expression. It must be distinguished from primary cALCL, other subtypes of CD30+ T- or B-cell lymphoma with anaplastic features and classical HL. Among all systemic ALCL, ALK- ALCL constitute 15-50% of cases [156]. In the previous edition of the WHO classification ALK- ALCL was considered as a unique entity together with ALK+ ALCL cases, under the name of “ALCL”. However it has been recognized as a different entity from ALK+ ALCL mainly because of the older median age of patients and the more aggressive clinical course. The principal differential diagnosis of ALK- ALCL is with PTCL, NOS and HL, but with complete immunophenotypic and molecular studies it can be distinguished in the majority of cases. In this regard, staining for PAX5 might be useful for differential diagnosis with HL. By contrast, the distinction between ALK- ALCL and PTCL, NOS is not always clear-cut, but loss of T-cell marker in ALK- cases occur with greater frequency than tipically seen in PTCL, NOS.

On the contrary, ALK- ALCL by definition is characterized by the absence of chromosomal translocations involving *ALK* gene. However, two translocations have been recently reported: the t(6;7)(p25.3;q32.3), involving the *DUSP22* gene and the FRA7H fragile site [157], reported in both systemic and cutaneous ALK- ALCL cases, and the translocation affecting the gene coding for *IRF4*, which is located less than 50

Kb telomeric to the *DUSP22* gene [158, 159]. The first lesion seems to cause a down-regulation of *DUSP22* expression level and an up-regulation of *MIR29A*, mapped within FRA7H. Their biologic and clinical significances are to be determined, since *MIR29A* is believed to act as tumor suppressor gene in ALK+ ALCL, in which it appears to be epigenetically silenced [160]. In terms of losses and gains, ALK- ALCL tends to differ from both ALK+ ALCL and PTCL, NOS, although overlapping features can be found [79, 80, 152]. The most frequent genetic alterations are gains of 1q41-qter, 5q, 6p, 7p 8q, 12q and 17q (17q12-q21), and losses at 6q (6q21-q22) and 13q (13q21-q22). Importantly, the majority of these chromosomal imbalances differ from those identified in ALK+ ALCL, supporting the concept that they are different biological entities.
The main differences in clinical presentation between ALK- ALCL and ALK+ ALCL are in terms of age at diagnosis, and involvement of extranodal sites. ALK- patients are usually older, with a median age of 55-60 years, and present a less common extranodal spread in only 20% of cases, involving mainly skin, liver and the gastro-intestinal tract.

3.5 Therapy

ALCL is currently treated with combinational chemotherapy and radiotherapy, in case of localized disease. Autologous and allogenic stem cell transplant could be considered in case of relapsed disease. However, many new molecular therapies have been developed on the basis of results obtained trough molecular studies. Conventional and targeted treatment will be widely described in the following sections.

3.5.1 Conventional Treatment

Although standard chemotherapy, such as CHOP (cyclophosphamide, doxorubicin, vincristine, prednisone), is given to newly diagnosed systemic ALCL patients, long-term disease outcome varies depending on the subtype. The ALK+ ALCL type responds extremely well to CHOP, rendering over a 70% long-term disease-free survival, whereas the ALK- type has a poorer prognosis with less than 45% to 50% achieving long-term disease-free survival. Higher doses of chemotherapy followed by a stem cell transplant may be prescribed for relapsed patients or patients with a low chance of being cured by CHOP chemotherapy [161]. German investigators have reported an improved failure-free survival when patients receive CHOP plus etoposide [162]. The doxorubicin, cyclophosphamide, vindesine, bleomycin, prenisolone (ACVBP) regimen,

which includes a high dose consolidation phase after the initial induction, is highly active in DLBCL and has also been utilized in patients with ALCL [163]. The five-year survival rate was 63% utilizing ACVBP: results that did not appear greatly different from that achieved utilizing CHOP. In a pediatric series, vinblastine was shown to be a very active drug as a single agent in patients with relapsed ALCL [164]. However, when vinblastine was added to an intensive, leukemia-like chemotherapy regimen, it did not reduce the frequency of relapse [165]. These patients generally receive CHOP-like or MACOP-B (methotrexate, doxorubicin, cyclophosphamide, vincristine, prednisone, bleomycin) regimens. Patients are generally responsive to doxorubicin-containing chemotherapy, but relapses are frequent, with a generally 5-year survival poorer than ALK+ cases.

3.5.2 Targeted Therapy and New Discoveries

The discovery of deregulated ALK in common cancers as NSCLC and neuroblastoma, as well as in ALCL, have increased the industry interest in the development of ALK inhibitors. In case of translocation, overexpression or activating point mutations, ALK inhibitors could represent an effective therapeutic strategy; however, currently, only one ALK inhibitor, Crizotinib, has been approved from FDA (Food and Drug Administration) for the treatment of NSCLC patients bearing the EML4-ALK translocation [166]. The only compound approved from FDA is the Pfizer product PF-02341066 (Crizotinib or Xalkori), which was initially developed as an inhibitor of c-Met (Fig 13, panel A) [167], but active in a similar way for ALK. In August 2011, it was approved for patients with locally advanced or metastatic NSCLC who have an ALK rearrangement, despite the low frequency of this chromosomal aberration (4-5%). Crizotinib belongs to the TK inhibitor family, the most prominent member of which is Imatinib (Gleevec; Novartis). Based on experience with this drug, it is known that resistance mutations can eventually occurred after prolonged treatment. This underscores the need for a wide range of chemically different ALK inhibitors [168]. The approval of Crizotinib is one more important landmark in the development of “molecular oncology”, as cancer care becomes defined by specific genetic alterations that can increasingly be treated by targeted therapies. Obviously, this type of treatment can be considered only for ALK+ ALCL; in contrast, for ALK- ALCL patients other treatment options would have to be taken into account. Several additional different classes of small molecules, identified as ALK inhibitors, have marked antitumor activity in preclinical models.

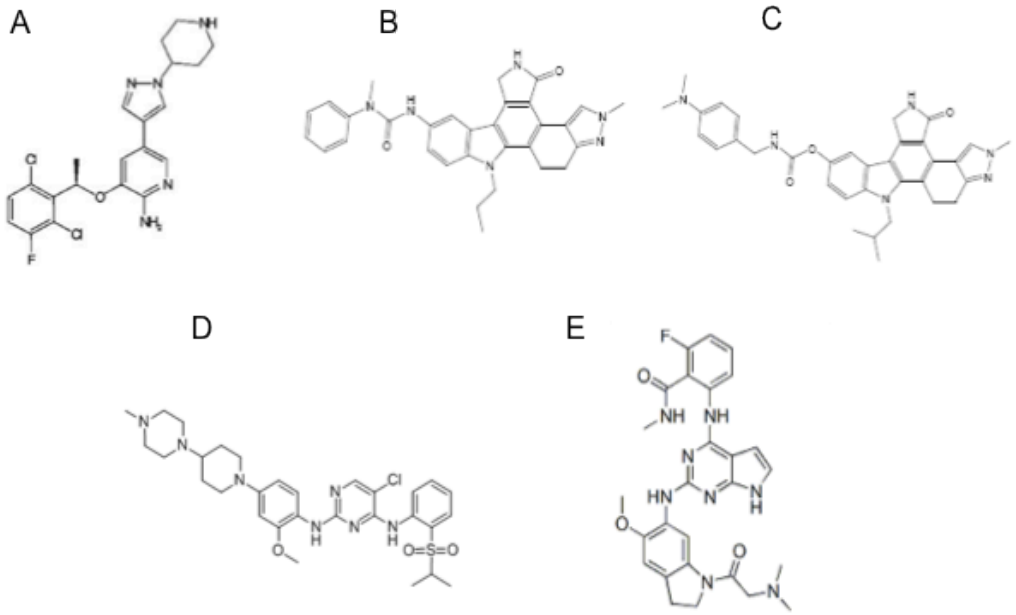


Figure 13. Structures of ALK inhibitors*. (A) structure of PF-02341066 (Crizotinib or Xalkori) (Pfizer); (B) structure of CEP-14083 (Cephalon); (C) structure of CEP-14513 (Cephalon); (D) structure of NVP-TAE684 (Novartis); (E) structure of GSK1838705 (GlaxoSmithKline).*

No naturally selective inhibitors of ALK exist, nevertheless, two structurally related natural product compounds have been reported to inhibit ALK as well as other kinases: staurosporine and 7-hydroxystaurosporine (UCN-01), isolated from *Streptomyces staurospreus* [168]. CEP-14083 and CEP-14513 (Cephalon, Frazer, PA) (Fig 13, panel B and C) are derivatives of staurosporine, whose clinical use is hampered by unfavorable physicochemical properties such as low solubility and stability. Cephalon produced also compound 18, a second-generation inhibitor, with improved bioavailability in *in vivo* models [169]. Three additional natural occurring or derivative compounds that affect ALK activity are geldanamycin, 17-allylamino-17-demethoxygeldanamycin (17-AAG) and herbimycin A. These modulators act trough the binding to heat shock protein 90 (HSP-90) leading to an indirect ALK inhibition, thereby enhancing the proteasome-mediated degradation of ALK protein (Fig 14) [170]. Other inhibitors classes include pyridine,

thiazole, aminopyridine, diaminopyridine, and fused ring system. After a cellular screen of a kinase-directed small molecule library to search for compounds that were selectively cytotoxic to cells oncogenically transformed by NPM-ALK, the compound NVP-TAE684 (Novartis) was discovered (Fig 13, panel D). It exherts its activity hindering NPM-ALK phosphorylation and, since the activity of others insulin receptors is not impaired, its activity is highly specific for NPM-ALK fusion protein [171]. Because of its good pharmacokinetic properties and excellent oral bioavailability, GSK1838705 (GloxoSmithKline), an ATP-competitive inhibitor of ALK, is very interesting (Fig 13, panel E) [172]. It is clear that only ALCL forms that bear the ALK translocations would be amenable to this type of treatment, but this option could be effective also in other ALK-expressing cancer if the growth of solid tumor is dependent on ALK expression.

Moreover, ALK+ ALCL cases could benefit of other type of treatments and immunotherapy would represent one of these choices. The capacity of immune system both to maintain cancer in an equilibrium state and to eliminate tumor cells could be exploited in order to increase or modulate the quality of this immune response against transformed cells. Some specific characteristics make the ALK portion of the fusion protein a potentially relevant tumor antigen for vaccination [173]. First, wild type ALK is expressed only in very few and specific regions of the central nervous system (CNS), being undetectable in normal somatic cells. This minimizes the risk of autoimmune reactions following ALK vaccination. Second, the fusion protein involving ALK is required for the transformation and maintenance of neoplastic lymphoid cells and thus the immune escape of tumor cells lacking ALK is unlikely. Finally, translocated ALK is observed in more than half of ALCL patients, especially in children and young adults, who will benefit of a long-term therapy for protection from relapses. Vaccinations with both HLA-specific peptides [174] and DNA vaccination [173] approaches have been reported. The great advantage of the latter strategy is that it can be done independently from HLA haplotype. After ALK vaccination, a total protection from tumor was observed in mice immunized, and a chemotherapeutic treatment in combination with ALK vaccination, mimicking the scenario of ALCL patients, reveals an improvement of the cure rate [173].

For ALK+ ALCL patients, inhibitors of chimeras-activated downstream pathways can be taken into account as shown in Fig 14. JAK3, for example, a signal transducer from ALK to STAT3, can be a possible target and PF-956980 (Pfizer) is one of the JAK3 small molecule inhibitor described [175]. SHP1, a phosphatase tumor suppressor, is constitutively deactivated through promoter methylation in 50% of ALK+ ALCL, which in

turn deactivates the JAK3/STAT3 pathway. Its reactivation after treatment with the methyl-transferase inhibitor 5-aza-deoxycytidine [176], leads to reduced levels of JAK3, phospho-JAK3 and phospho-STAT3 with the final goal to sensitize cells to doxorubicin-induced apoptosis. STAT3 can also be directly deactivated by the small molecule S3I-201 (National Cancer Institute) which acts by inhibiting DNA binding and STAT3 complex formation [177]. Other compounds for the inhibition of STAT3 downstream targets, such as Bcl2 and Bcl-XL, are described: ABT-737 and its derivative ABT-236 (Abbott laboratories) [178, 179]. The RAS/ERK pathway, which is mainly responsible for cell proliferation, may be inhibited by farnesyltransferase or geranylgeranyltransferase inhibitors [180]. Moreover, rapamycin and its orally available analog temsirolimus, may be used for the selectively suppression of mTOR protein, while the overactivation of AKT can be suppressed with classical ATP competitive inhibitors, phosphatidyl analogs that block the essential binding of AKT to phosphatidylinositol (3,4,5)-triphosphate, pseudosubstrate inhibitors and allosteric inhibitors [181].

Immunological strategies, such as anti-CD30 antibodies, are currently considered as therapeutic option for both forms of ALCL. First studies on the use of CD30 antibodies for ALCL and HL patients back to more than 15 years ago. In one of the first promising studies, the anti-CD30 monoclonal antibody was bound to saporin, a protein with ribosome-inactivating activity. A rapid reduction of the tumor was observed, but the clinical response were transient [182]. Despite high expectations, also treatment with a fully human anti-CD30 immunoglobulin G1 monoclonal antibody (MDX-060), although well tolerated, showed low rates of clinical response [183]. Specific response only in ALCL patients was observed after treatment with the anti-CD30 antibody SGN-30 (Seattle Genetics), but the clinical activity was very limited with side effects rather mild [184]. An improvement to these anti-CD30 antibody-based treatment strategies is SGN-35 (Brentuximab Vedotin or Adcetris; Seattle Genetics). In this case the SGN-30 antibody is coniugated to the antitubulin agent monomethyl auristatin E (MMAE) with an enzyme-cleavable dipeptide linker: the antibody portion of the drug attaches to CD30 on the surface of malignant cells, delivering MMAE which is responsible for the anti-tumour activity. This improvement has been very succesfull in clinical trials, with a tumor regression observed in 86% of ALCL patients [185]. On the basis of these studies, SGN-35 was approved from FDA in 2011 for the treatment of ALCL and HL. One appealing

feature of the CD30-antibody strategy is that, in contrast to ALK inhibitors, it can be used for all ALCL patients, ALK-, ALK+ and cALCL ones.

Novel molecular targets for future ALCL therapies could be represented by miRNAs. Reintroduction or downregulation of differentially expressed miRNAs could be considered as a novel strategy for ALCL treatment (Fig 14). In ALCL, miR-17-92 cluster has been reported to be overexpressed in ALK+ ALCL cases; conversely, miR-155 is overexpressed in ALK- patients. Also miR-29b, with a known tumor suppressor activity, and miR-101, both with a low expression in both forms of ALCL, could be used for ALCL therapy [127]. Fig 13 represents a visual resume of all targeted and future treatments, widely described in this paragraph, for ALCL patients.
It is clear that, despite the several treatment options current available for ALCL, ALK- patients benefit of less strategies, and their treatment remains a big challenge for the future. But the recent increase in the therapeutic armamentarium for lymphomas should give us confidence that also for ALK- ALCL, specific molecular targets, instrumental for development of new drugs, will be identified [168].

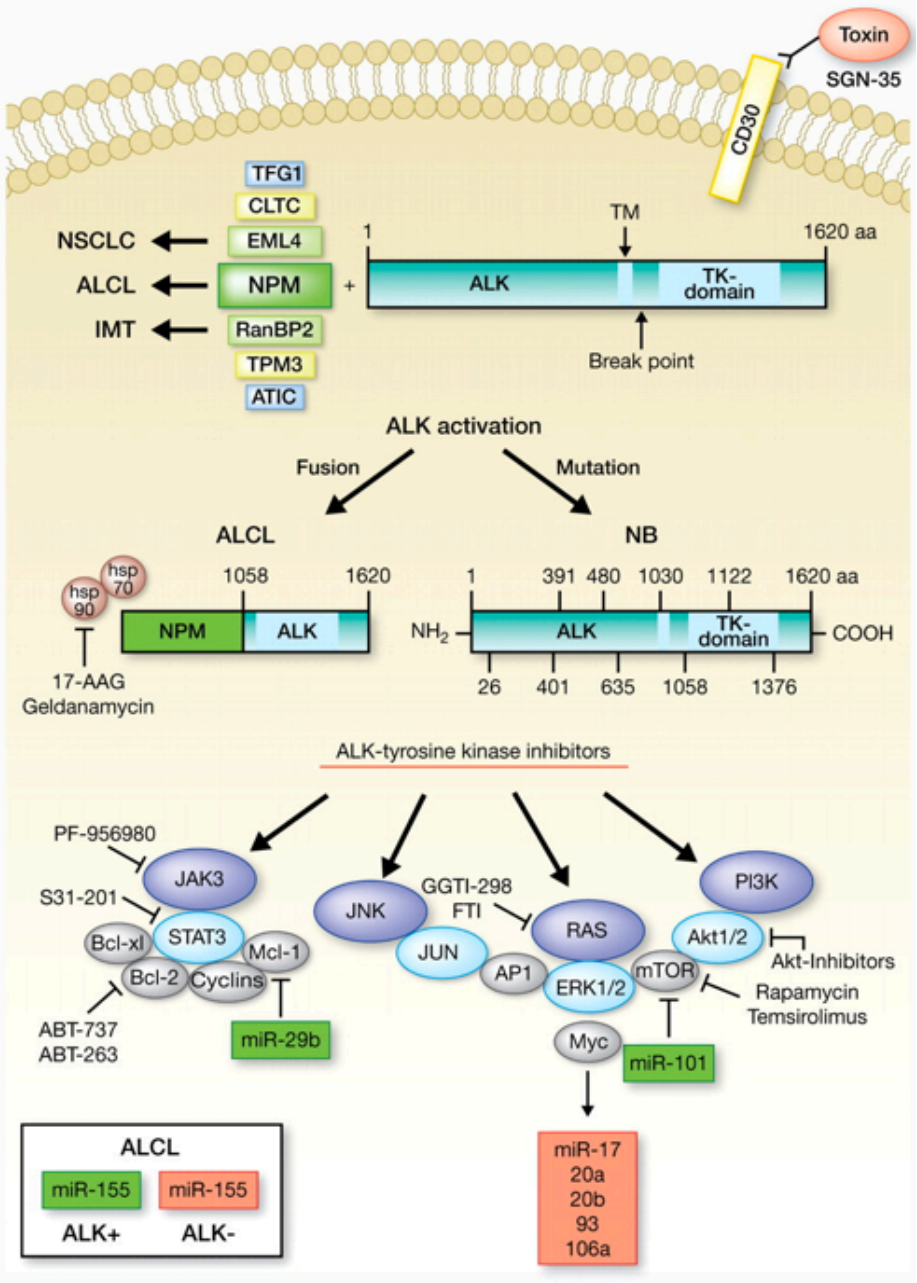


Figure 14. Sites of action of targeted ALCL therapeutics. *ALK can be activated through translocations, full length overexpression and activating point mutations. All these cases are amenable to ALK-TK inhibition. Another approach is to inhibit one of the many ALK-activated downstream pathways such as PI3K/Akt1/mTOR, JAK/Stat3 and RAS/ERK. Other options which can be considered also for ALK- ALCL patients are the immunotherapy anti-CD30 or strategies considering novel molecular targets such as miRNAs, which may be considered for future ALCL therapies [168].*

4. PRDM1/BLIMP1

In 1991, Maniatis and colleagues [186] described for the first time a novel transcriptional repressor of the IFNβ gene. They used a screening of a cDNA expression library for proteins which could interact with the positive regulatory domain I (PRDI) sites in the IFNβ promoter and with this type of research they found the transcriptional repressor that they named PRDI-binding factor I (PRDI-BF1). Three years later, the group of Davis isolated a murine cDNA that was induced following the cytokine-induced differentiation of Bcl-1 cell line and named this factor B lymphocyte maturation-induced protein 1 (Blimp1) [187]. Beside the discovery of the murine homolog of the previous discovered PRDI-BF1, this study is also the demonstration that the ectopic expression, and consequent overexpression, of Blimp1 is sufficient to drive mature B cells into antibody-secreting cells (ASC). Additional studies showed the homology between PRDI-BF1 and murine Blimp1, although the mouse protein presents a slight different N terminus from the humans [188], the human and the mouse proteins are highly homologous and are interchangeable in functional assays [189]. For simplicity BLIMP1 has been adopted also to name the human protein.

4.1 Gene and protein structure

The gene encoding BLIMP1, *PRDM1*, is located on chromosome 6q21 in human and on chromosome 10 in mice [190]. Human *PRDM1* encodes two major isoforms designated BLIMP1α and BLIMP1β, which arise from alternative promoters as shown in Fig 15. The TATA-less, GC-rich promoter, used for the transcription of the full-length form, is immediately upstream of exon 1; in contrast, BLIMP1β is transcribed from a novel promoter and a new exon, 1β, located upstream of exon 4 of the full-length gene, lacking the first 101 amino acids of BLIMP1α [191]. While the full-length BLIMP1α is responsible for plasma cells differentiation [192], no clear functions have been attributed to BLIMP1β; but because it contains the DNA-binding domain but bears a disrupted regulatory domain, it has been suggested that it might inhibits BLIMP1α functions [191]. A schematic representation of gene and mRNA of BLIMP1α and BLIMP1β are reported in Fig 15.

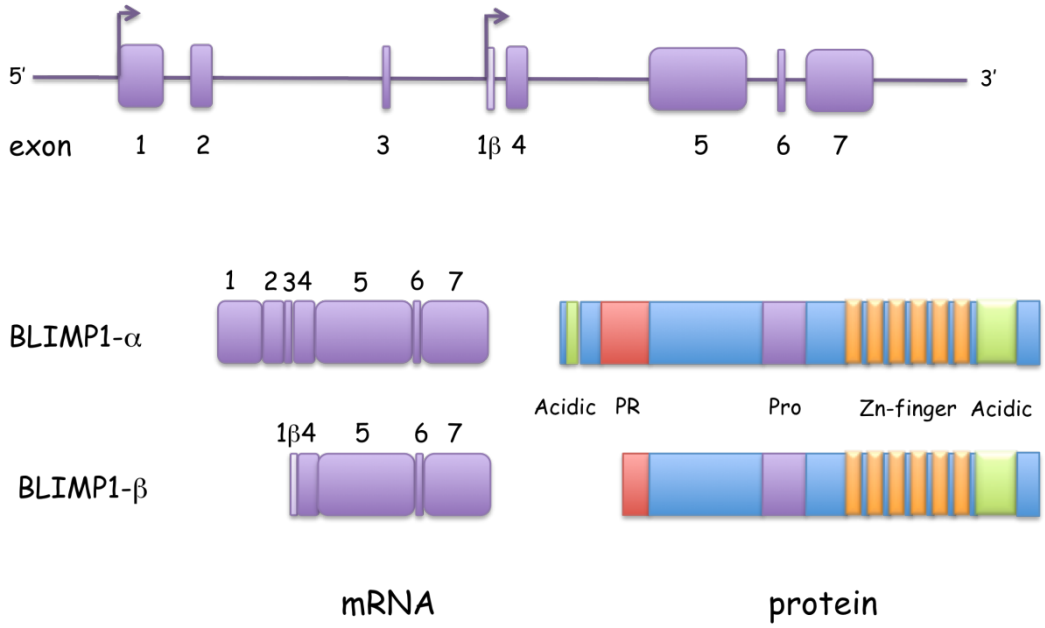


Figure 15. Schematic representation the genomic structure and the expressed mRNAs of BLIMP1α and BLIMP1β. *BLIMP1β lacks exons 1–3 and the amino-terminal 101 aa of BLIMP1α and has new exon 1β between exons 3 and 4. BLIMP1β transcript is transcribed from alternative promoter between exons 3 and 4. Numbered open boxes represent the exons. PR: PRDI-BF1-RIZ1 homologous; Pro: proline rich; Zn: zinc.*

Human BLIMP1 protein contains approximately 789 amino acids and a predicted molecular weight of 88 kDa. The protein contains different domains (Fig 15):

- five C_2H_2 zinc finger motifs located in the C-terminus;
- PR domain;
- the consensus binding site (PRDI site);
- proline-rich region.

The five Krüppel-type zinc fingers are clearly implicated in the DNA-binding motif, the first two of which are sufficient for the recognition of PRDI region in the *IFNβ* promoter [193]. The zinc finger domain has also been described to confer the repressor activity by recruiting the histone methyl transferase G9a, which methylates lysine 9 on histone 3 (H3K9), leading to a repressive histone modification. The PR domain, so called after its identification in the PRDI-BF1 and RIZ proteins, has similarity to SET domains found in histone methyl transferases (HMT). In fact, even if BLIMP1 has no demonstrable

HMT activity, it is able to recruit the G9a HMT. The PRDI site is an 11-bp sequence (A/C)AG(T/C)**GAAAG**(T/C)(G/T), similar to the binding sites for IFN regulatory factor (IRF)1 and IRF2: in fact, BLIMP1 and IRF1/2 compete for binding to the site in the *IFNβ* promoter [194]. The proline-rich region, located N-terminal to the zinc fingers, is required for transcriptional repressor activity along with them, and it has been shown to mediate the association with transcriptional corepressors such as Groucho family members [195] and histone deacetylases 1 and 2 [196].
Different mechanisms are used by BLIMP1 to silence its target genes in a context-dependent manner. Individual domains of the BLIMP1 protein are able to recruit specific co-repressor complexes or cromatine-modifying enzymes to mediate transcriptional repression. Thus, BLIMP1 serves as scaffold to recruit proteins or co-repressor complexes which modify histones (by deacetylation, H3K9 methylation and arginine methylation) and, in doing so, assembles silent chromatin over the target loci. However, the mechanism used by BLIMP1 to modify histones and repress target genes has to be exactly determined. Moreover, it is unclear if chromatin modification is the only mechanism by which BLIMP1 repress transcription.

4.2 BLIMP1 functions

In order to understand the function of BLIMP1 protein, numerous efforts to create BLIMP1 knock-out (KO) mice have been made. Unfortunately, attempts to create viable BLIMP1 KO were unsuccesful, but these studies revealed a surprising role for the protein in fetal development: embryos with deletions of either exons 1-3 or 6-8 died around embryonic day (E) 10.5 from loss of mesenchymal bronchial arches, placental defects and severe hemorrage [197], while mice with removal of only exons 7-8 showed a prolonged embryonic lifespan up to around E15 [198]. BLIMP1 can be detected in embryos already at E6.5 with a dynamic expression pattern: it is induced in the foregut pocket, primordial germ cells, anterior mesendoderm, limb buds, and myotomes at early gestation, and in chondroblasts, mesenchymal tissues, hair follicles, and sensory vibrissae in mid to late gestation [199]. To bypass embryonic lethality, mice reconstituted with Blimp1-deficient fetal liver stem cells or conditional Blimp1-KO mice have been utilized in studies on the immune functions of BLIMP1.

The development of several BLIMP1 reporter mouse strains have allowed the analysis of B-cell specific BLIMP1 expression *in vivo*: BLIMP1 is present in all ASCs independent

from their locations and differentiation history, but not in naïve and memory B cells. In this regard, mice with B cells specific deletion display normal early B cell development and unperturbed or even increased GC formation following immunization [200]. Moreover, proliferating B cells, lacking BLIMP1, arrest their differentiation to plasma cells, resulting in a severe reduction of Ig production of all isotypes. On the contrary, ectopic expression of BLIMP1 alone is sufficient to induce an ASCs phenotype in B cells, establishing, in this way, BLIMP1 as the master regulator for plasma cells differentiation [187, 201].
Importantly, BLIMP1 is expressed in other cells of the immune system: it is also expressed in human peripheral blood monocytes and granulocytes, and its overexpression is able to induce differentiation into cells with a macrophage phenotype in a monocytic cell line [202]. BLIMP1 is also involved in osteoclastogenesis [203] and in the modulation of the function of dendritic cells [204].

Mice lacking BLIMP1 in their T cells have provided numerous and important informations about its role in the T lineage. Two different strategies have been used to generate mice with BLIMP1-deficient T cells: the first approach leaded to mice in which BLIMP1 deletion was specifically restricted to T lymphocytes, while the latter one generated chimeric mice in which cells from both the myeloid and the lymphoid lineages were BLIMP1 deficient [198, 205]. Despite the differences in the two strategies, both studies revealed that lack of BLIMP1 results in critical alterations of T cells functions and homeostasis, culminating with the spontaneous development of inflammatory disease [189] in colon [198], as well as in lungs and liver [205]. Despite the molecular mechanisms associated with the development of colitis have not been clearly characterized, it has been suggested that in mice with BLIMP1-deficient T cells there is an increased production of IFN-γ with a concomitant decreased levels of IL10, indicating deregulated T cells activation and/or defective development and function of Treg and Th1 cells.

4.3 BLIMP1 and B-cell development

BLIMP1 is a key gene whose function places it central to the transcriptional network that controls B cell terminal differentiation (Fig 16). B cells can differentiate into either short-lived plasma cells or germinal center B cells after an appropriate activation and costimulation. Short-lived plasma cells stop proliferating and acquire the capability to produce antibodies to quickly combat infection, thanks to the elaboration a large secretory apparatus [206]. Conversely, germinal center B cells, specialized for

proliferation and affinity maturation, have a minimal secretory apparatus. They are responsible for the production of higher affinity, isotype-switched antibodies [207].

The commitment to differentiation was determined by the outcome of a competition between two sets of transcription factors [206]: on one side there are PAX5 (paired box gene 5) and BCL6 (B cell lymphoma 6), and in the other BLIMP1 and XBP1 (X-box-binding protein1) (Fig 16). PAX5 is a critical factor for B cell development and the maintenance of B cell identity, since its loss from mature B cells leads to the loss of several proteins critical for B cells functions, such as CD19. It functions also as repressor of XBP1, whose function induces formation of the secretory apparatus necessary for the production of large amounts of antibody [208]. BCL6, another transcriptional repressor, is required for the formation of GC in B cells. It is expressed in the earlier stage of B cell development, where it might act to promote cell division and maintain viability reducing the expression of factors such as p21 and p53. It might function also to maintain an activated but undifferentiated state through the repression of STAT3-mediated B-cell differentiation [209]. A regulatory loop for B-cell differentiation came out from the observation that BCL6 repress BLIMP1 expression, and BLIMP1, in turn, is able to repress both BCL6 and PAX5 [208]. Thus, in this model, plasma cell differentiation was triggered by a reduction in BCL6 expression with a concomitant increase in BLIMP1 (Fig 16). The increased levels of BLIMP1 induced the plasma cell gene-expression program, with the subsequent loss of PAX5 and BCL6 turning off the B-cell gene-expression program. Finally, the presence of IRF4 as well is required for plasma cells differentiation [210, 211]: it has been suggested that it may act upstream or in parallel to BLIMP1, repressing BCL6, thereby terminating the GC transcriptional programme. Expression of BLIMP1, IRF4 and XBP1 seems to be independently regulated, but all of them are jointly required for the establishment of the terminally differentiated plasma cells [207], as reflected by the dramatic effects of their individual deficiencies on plasma cells development.

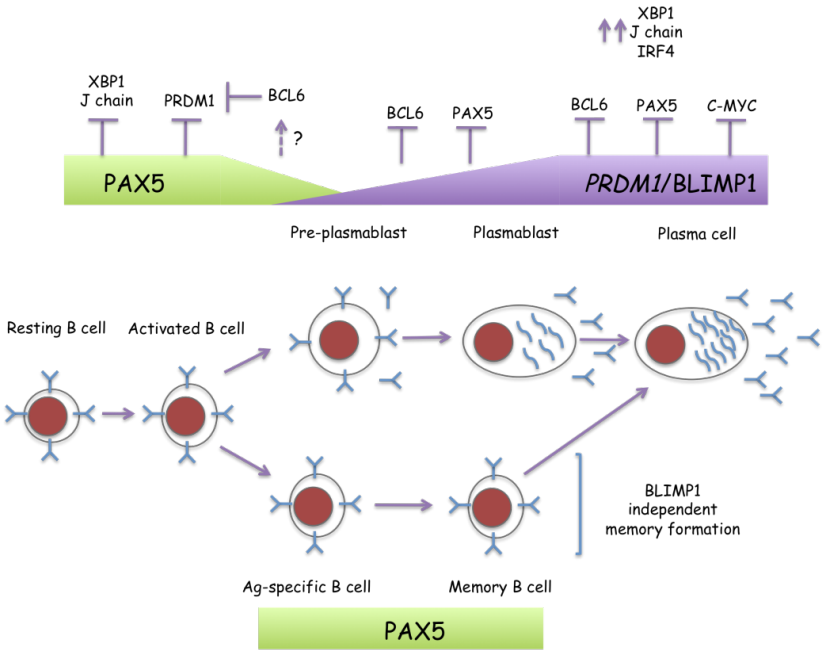


Figure 16. BLIMP1 controls plasma cells differentiation. *The stages of cellular differentiation from a resting B cell to an antibody-secreting plasma cell are reported, together with the relative levels for PAX5, BLIMP1 and key target genes. The pre-plasmablast stage is characterized by low BLIMP1-independent antibody secretion. Plasmablasts have intermediate BLIMP1 expression, proliferate rapidly and secrete antibodies. Increasing BLIMP1 levels are associated with long-lived plasma cells that secrete large quantities of antibody and exit the cell cycle. Memory B cells, by contrast, are BLIMP1-independent.*

Tight control over BLIMP1 expression is essential, since premature expression of BLIMP1 in primary B cells leads to cell death, suggesting that only cells "ready" to initiate the transcriptional program driven by BLIMP1 are able to survive and differentiate [212]. BLIMP1 has long been recognized as the master regulator of plasma cells development, since exogenous expression is sufficient to induce plasmablast differentiation [187], while, on the contrary, mice in which the expression of BLIMP1 is conditionally depleted are no longer able to generate plasma cells [192]. It has been subsequently shown that XPB1, a transcription factor required for the secretory phenotype of plasma cells, is a downstream target of BLIMP1 [207].

Several stimuli that promote B cell differentiation are able to induce BLIMP1 expression. LPS, which activates TLR4, together with several cytokines, including IL-2, IL-5, IL-10, and IL-21, which activate STAT3, a direct activator of *PRDM1* transcription, are some examples. *PRDM1* transcription is also induced by the transcription factors IRF4 and AP-1. Besides the activation, also the repression of BLIMP1 is important in the process of the B cell differentiation: BCL6, PAX5 and BACH2 are the main regulators of this program.

4.4 BLIMP1 and T-cell development

For many years it was assumed that BLIMP1 function in immune response was B-cell specific, however, it is now accepted that BLIMP1 has a critical role also in maintaining T cell homeostasis [198, 205]. The idea that BLIMP1 is more important for T cell function than development is reinforced by the observation of the pattern of BLIMP1 mRNA expression in T cells: low in thymocytes and naive T cells and high in antigen-experienced T cells. Studies in which BLIMP1 was ectopically expressed or, on the contrary, deleted, reinforced the concept that BLIMP1 is involved in the regulation of responsiveness and homeostasis of peripheral T cells by attenuating both proliferation and survival [189]. One key mechanism by which BLIMP1 modulates T cell responsiveness is through repression of IL-2: IL-2 production, induced by T cell activation, is inversely correlated with BLIMP1 expression levels. This cytokine acts in both autocrine and paracrine manners, regulating the initial expansion of naive T cells upon TCR stimulation *in vitro* and *in vivo*. Thus, T cell activation induces *il2* transcription, IL-2 signaling, in turn, induces *PRDM1* transcription, and BLIMP1 feeds back to repress *il2* transcription [189]. The negative feedback loop involving *PRDM1*/BLIMP1 and IL-2 is reported in Fig 17.

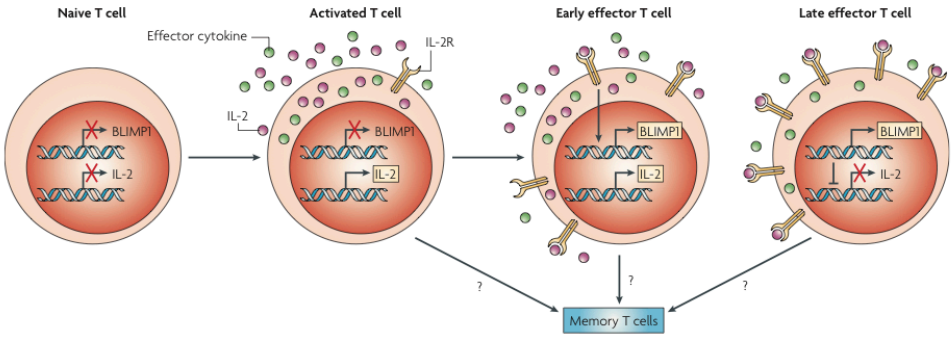


Figure 17. Model of interaction of BLIMP1 and IL-2 in T-cell differentiation*. Naive T cells express neither BLIMP1 nor IL-2. Following encounter with antigen, activated T cells secrete high quantities of IL-2 and upregulate the IL-2 receptor (IL-2R). Engagement of IL-2R by IL-2 results in PRDM1/BLIMP1 transcription in early effector cells. BLIMP1 expression increases in late effector cells and repress IL-2 expression. Figure taken from Nutt SI et al, Nat Rev Immunol 2007 [213].*

T cell homeostasis as well is crucial for a functional immune system, as the accumulation of T cells resulting from lack of regulatory T cells or an inability to shut down immune responses can lead to inflammation and autoimmune pathology. BLIMP1 is an essential regulator of T cell homeostasis: in peripheral T cells, it is induced upon activation and is required for T cell quiescence and to attenuate TCR signaling in response to stimulation.

T cells can be divided in several subsets, each with a distinct function. The main distinction is among T helper and T cytotoxic cells, also known as CD4+ and CD8+ T cells for the presence of CD4 and CD8 antigens respectively on their surface. T helper cells assist other cells of the immune system in different processes such as the maturation of B cells and activation of cytotoxic T cells and macrophages. They can also differentiate in one of several subtypes: Th1, Th2, TFh, Treg and Th17, each of them secreting different cytokines to facilitate different types of immune response. The observation of high levels of BLIMP1 in Treg cells and the inflammatory disease associated with BLIMP1 ablation in T cells, suggest a role for this transcription factor in Treg cell differentiation and/or function, where it is also required for IL-10 production, an immune-suppressive cytokine important for intestinal mucosal tolerance. Both functions are impaired in case of development of spontaneous inflammatory disease or colitis in Blimp1-deficient mice [214]. Different evidences suggest that BLIMP1 is also involved in

Th2 commitment: BLIMP1 is able to repress Th1 genes such as *Bcl6* and *tbet* and the observation of an increased production of IFN-γ by BLIMP1-deficient cells, confirmed the fact that BLIMP1 may attenuate Th1 differentiation [198]. Moreover, BLIMP1 transcripts and protein are more abundant in Th2 than in Th1 cells differentiated *in vitro*. BLIMP1-deficient CD4+ T cells are deflected from terminal Th1 or Th2 differentiation and are more diverted toward TFh differentiation showing greater proliferative potential. So, BLIMP1 appears to be a fundamental regulator in Th differentiation, but another important player in this process is BCL6, showing the opposite behaviour. BLIMP1-deficient T cells have increased BCL6 mRNA levels and BCL6 protein can repress Th2 differentiation by interfering with the function of GATA3, a master regulator of Th2 differentiation [215].
Moreover, BLIMP1 expression has recently been reported in cytotoxic T cells, responsible for the destruction of virally infected cells and tumoral cells. Both effector and memory CD8+ cells showed the expression of this transcription factor, and it has been reported that the absence of BLIMP1 can lead to an excessive proliferation and increased numbers of memory CD8+ T cells [198].
In conclusion, BLIMP1 is a critical factor for most terminal effector cell differentiation in both CD4+ and CD8+ T cells. Terminal differentiation in memory T cells from both subsets, is characterized by low proliferative potential and high effector molecule secretion. These cells, also known as CD4+ and CD8+ memory T cells, are antigen-specific T cells that persist long-term after infection has resolved. After the re-exposure with the cognate antigen, they quickly expand, providing the immune system with "memory" against past infections. They can be further subdivided in central memory T cells (T_{CM}), which reside in lymphoid organs, and effector memory T cells (T_{EM}), dominant in tissue at sites of infection.

4.5 BLIMP1 and Lymphomas

Different evidences, among which the frequent loss of the genomic region 6q21 encompassing *PRDM1* gene, suggest that an abnormal regulation or an abnormal activity of BLIMP1 may be one causal event in some B cells and NK cells malignancies. If this transcription factor acts as a tumor suppressor, lack of its function could be important in development of some forms of lymphomas. However, loss of BLIMP1 alone is not sufficient for B cells lymphomagenesis, as inferred from mice with conditional deletion of *PRDM1* in the B cell lineage which do not develop B cell tumors spontaneously [189]. Nevertheless, more and more studies recently reported the frequent inactivation of

BLIMP1, through deletion or mutation, in different lymphoma subtypes, providing evidence that it may indeed functions as a tumor suppressor gene.

In 2006, two different groups [49, 216], after the observation of the frequent loss at genomic locus 6q21, identified BLIMP1 inactivation by structural alterations in approximately 24% of DLBCL cases analyzed. Interestingly, all of these cases were restricted to the ABC-DLBCL subtype, suggesting a role for this gene in this particular B cell lymphoma subtype. Nor GCB-DLBCL cases or unclassified cases have been reported to show inactivation of BLIMP1. The alterations observed in these studies included gene truncations, nonsense mutations, frameshift deletions and splice site mutations that leaded to the generation of aberrant transcripts encoding truncated BLIMP1 proteins. Moreover, in all cases, both BLIMP1 alleles were inactivated by deletion or mutations. These results suggest that BLIMP1 inactivation follows the typical two-hit mechanisms in different ABC-DLBCL cases, strongly suggesting that BLIMP1 may act as a tumor suppressor gene, whose loss of function may be critical for the pathogenesis of ABC-DLBCL. More recently the group of Laura Pasqualucci [50] reported further data on the role of BLIMP1 in DLBCL: converging evidences from human genetics, functional studies and mouse models indicated that BLIMP1 must be considered a *bona fide* tumor suppressor gene for ABC-DLBCL, in which the gene inactivation contributes to lymphomagenesis by blocking plasma cells differentiation. In this report, BLIMP1 was inactivated in 53% of ABC-DLBCL cases analyzed by multiple mechanisms, such as deletions, homozygous deletions, truncating or missense mutations and transcriptional repression by BCL6, which was observed to be translocated and so constitutively active in many cases. BCL6 aberrations and BLIMP1 inactivation appear mutually exclusive. Analysis of 6q21 deletions revealed, also in this cohort of samples, that the majority of the mutated cases had lost expression of the wild type allele due to deletion, epigenetic silencing or UPD, demonstrating biallelic inactivation of the gene. As in the previous study of the group, also in this analysis, a moderate percentage of ABC-DLBCL cases showed no alteration at BLIMP1 or BCL6 levels, but lacked BLIMP1 protein, despite the presence of IRF4, a transcriptional repressor that is known to be invariably coexpressed with BLIMP1 in normal GC B cells and in all plasma cells. This phenomenon suggests that other mechanisms, such as for example promoter hypermethylation, could be responsible for the absence of BLIMP1 protein. Functional studies performed with the re-introduction of BLIMP1 into a DLBCL cell line, suggest the central role of this gene in the block of terminal differentiation and in the modulation of

expression of genes involved in cell cycle progression. Other studies in mouse models, confirmed that BLIMP1 disruption promotes the development of lymphomas often resembling human ABC-DLBCL, but, since the low penetrance, long latency and clonality of DLBCL in these mice, it has been suggested that additional oncogenic events are required for the pathogenesis of this disease [217]. Constitutive activation of the canonical NF-κB pathway, which is also a typical feature of human ABC-DLBCL, strongly synergized with BLIMP1 loss in lymphomagenesis, suggesting that both NF-κB signaling and BLIMP1 loss play critical roles in the pathogenesis of ABC-DLBCL [217].

Two years ago, *PRDM1* has been reported to be a gene with a tumor suppressor function in Natural Killer Cell Lymphoma (NKCL) as well [91, 92]. In both cases, array-CGH analysis highlighted 6q21 loss as one of the most frequent aberration identified in clinical specimens and cell lines. In the study of Karube et al [91], the two narrow regions inside the 6q21 loss include different genes, among which *PRDM1* and *FOXO3*. Forced re-expression of BLIMP1 and FOXO3 leaded to the suppression of cell proliferation in both cases with a concomitant induction of apoptosis and cell-cycle arrest. The genomic analysis for the presence of mutations inside these two genes revealed that only one or two cases (one clinical specimen and one cell line) for FOXO3 and BLIMP1 respectively, presented a genomic mutation accompained by the genomic loss of the gene, considering the idea of a biallelic inactivation. For the other cases, in which only heterozygous genomic loss is observed, haploinsufficiency has been suggested as the mechanism responsible for gene inactivation. In both cases, functional studies suggest a tumor suppressor role for these genes. Moreover, *PRDM1* gene has been reported to be inactivated by a combination of deletion and promoter hypermethylation [92] in NKCL. In support of its role as a tumor suppressor gene, the reconstitution of BLIMP1 in NKCL cell lines led to G2/M cell-cycle arrest, increased apoptosis and a strong negative selection pressure with progressive elimination of BLIMP1-expressing cells. In conclusion, recent and numerous evidences suggest that *PRDM1*/BLIMP1 could act as a tumor suppressor gene in different types of lymphoma, deriving from B and NK cells, and have a role in the pathogenesis of these neoplastic diseases. For this reason, the study of this gene in other malignancies, especially in other lymphoma subtypes, could improve the understanding of the role of this gene in the pathogenesis. If BLIMP1 exherts a role in lymphomagenesis in different type of lymphomas, deriving also from different cells of origin, could be of great interest to try to develop different thrapeutic strategies which allow the reintroduction of *BLIMP1* gene in tumor cells. This putative target therapy could be helpful

in association with the conventional ones which very often are not sufficient for the healing of patients.

Aim of the Study

Anaplastic Large Cell Lymphoma (ALCL) is a malignancy of mature T cells. It accounts for about 5% of all NHL with a typical bimodal distribution, presenting two major peak of incidence, one in children and young adults (10-30 years of age) and the other one in the elderly (60-70 years of age). ALCL can be further subdived into two main subtypes on the basis of the presence or absence of a typical chromosomal translocation involving the Anaplastic Lymphoma Kinase (*ALK*) gene. ALK translocation leads to the constitutive activation of the tyrosine kinase ALK, which, in turn, is able to activate downstream pathways promoting cell cycle progression, cell proliferation and survival of neoplastic cells. While the pathogenesis of the ALK+ ALCL subset is now well known, so far the molecular events responsible for the ALK- ALCL subset are still largely unknown. Even if recent evidence suggests that ALK+ and ALK- subtypes, despite sharing some features, represent two distinct and different entities, ALK- ALCL represents a distinct subset of ALCL in the current WHO classification although only provisionally. For this reason, the aim of this project is to demonstrate that ALK+ and ALK- subsets can definitively be considered different entities.

In order to achieve this aim, our first goal is to unravel the characteristic genomic aberrations affecting each ALCL subtype. The use of Affymetrix SNP 6.0 platforms allows us to detect genomic CN changes at high-resolution level along the whole genome of patients presenting ALCL. The array-CGH approach identifies gains and losses of genomic material and the comparison of the aberrations between the two ALCL groups permits us to identify those genes that are likely to be involved in the lymphoma pathogenesis concerned. We focused our attention in particular on ALK- ALCL CN changes, since less information is available regarding its pathogenesis. The identification of genomic locus preferentially gained or lost in one subtype over the other are analyzed in detail in order to unravel genes whose function might be critical for malignant transformation, and hence might represent possible therapeutic targets. The second part of the project provides the functional characterization of one of the genes identified, via gain-of-function experiments. Thanks to the collaboration of the University of Turin, *in vivo* studies have been performed in order to confirm the data obtained in *in vitro* experiments and to assess definitively a role of the gene identified and studied.

Moreover, we collaborated in a study with the University in Turin in which the use of RNA-seq allows us to discover the presence of new putative pathogenetic translocations in the ALK- subset and their consequences were studied *in vivo* with a PTD model.

Materials and Methods

1. GENOMIC ANALYSIS

Tumor Panel

All clinical specimens were derived from involved sites and obtained in the course of routine diagnostic procedures before therapy initiation. Cases were selected based on the availability of frozen material with a fraction of neoplastic cells in the specimen representing more than 70% of overall cellularity as determined by morphologic and/or immunophenotypic studies. Informed consent was obtained in accordance to the Declaration of Helsinki following the procedures approved by the local ethical commitees and institutional review boards of each participating institution.

DNA samples

Genomic DNA was isolated from frozen material using DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany), following the manufacturers instructions. DNA samples were quantified by spectrophotometric 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 100 ng total genomic DNA on a 1% agarose gel prepared and run in TBE buffer. In the gel was added 1 μl of GelRed Nucleic Acid Stain (Brunschwig, Basel, Switzerland) for staining DNA in agarose gel. After electrophoresis, DNA was visualized using the Alphimager 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.

Array-CGH

Genomic profiles were obtained using Affymetrix Human Mapping GeneChip 6.0 arrays (Affymetrix, Santa Clara, CA, USA) following the workflow presented in Fig 18. 500 ng of genomic DNA were used for this analysis: two replicates of each sample are required for this protocol, one for Nsp and one for Sty digestion. One positive control and one negative control were included with every set of samples run. Sty Digestion Master Mix, composed by 1X NE buffer 2, 1X BSA and 1U/ml Sty, was prepared on ice in a final

volume of 14,75 μl and added to 250 ng DNA. This digestion mix was incubated at 37°C for 120' and 65°C for 20' in a Tetrad 2 Peltier Thermal cycler (BIO-RAD, Bio-Rad Laboratories, Marnes-la-Coquette, France). Ligation mix contained 0.75 μl Adaptor Sty I (50 mM), 2.5 μl T4 Ligase Buffer (10X) and 2 ml T4 Ligase (400U/μl). This mix was addedd to previously Sty digested samples and they were loaded onto the thermal cycler. The ligation program was 16°C for 180' and 70°C for 20' and once it was finished, the final volume of 100 μl for each sample was reached with molecular biology-grade water. PCR amplification of ligated DNA fragments was performed in three replicates for each sample. 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 mM each dNTP, 4.5 mM 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 Tetrad 2 Peltier Thermal cycler (BIO-RAD, Hercules, CA, USA) with heathed lid, was: denaturation at 94°C for 3', followed by 30 cycles of 94°C for 30", 60°C for 30" and 68°C for 15", then followed by one cycle at 68°C for 7'. After the amplification, the correct PCR distribution, between 200 bp and 1100 bp, was controlled on a 2% agarose TBE gel. The same workflow was followed for the digestion, ligation, amplification with NSP restriction enzyme: the only difference was in the preparation of the reaction in quadruplicates. After having pooled all the PCR products reaching the final volume of 700 μl, PCR products were purified using magnetic beads and connecting the plate to a vacuum source. With application of a vacuum the wells were dried and washed two times with 75% Et-OH. 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, Wilmington, DE, USA). 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 0.5 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 GeneChip DNA labeling Reagent, 1X TdT Buffer and 30U/μl Terminal deoxynucleotidyl Transferase (TdT). The mix was prepared on ice and incubated at 37°C for 4 hours and 95°C for 15'. The hybridization cocktail was prepared on ice mixing 70 μl of fragmented and labeled DNA target with MES, Denhardt's Solution, EDTA, 10mg/ml Herring Sperm DNA,

Oligonucleotide Control Reagent, 1mg/ml Human Cot-1 DNA, Tween-20, DMSO and TMACL (Tetramethyl Ammonium Cloride), to a final volume of 263 μl. After that the hybridization mix was denatured at 95°C for 10' and 49°C for 1' in a thermal cycler, 200 μl of denatured hybridization mix were injected into the Genome-Wide Human SNP Array 6.0, that was then placed in a GeneChip Hybridization Oven 640 (Affymetrix, Santa Clara, CA, USA), rotating at 60 rpm at 50°C and incubated for 16-18 hours.

Buffers and solutions for array hybridization, washing and staining were prepared according to AffymetrixGenome-Wide Human SNP Nsp/Sty 6.0 User Guide and, when recommended, were filtered with the stericup vacuum (Millipore, Billerica, MA, USA) driven disposable filtration system (Millipore, Billerica, MA, USA). To wash, stain and scan a probe array, we first assigned a name to each sample and then we registered it in AGCC software.

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

The arrays were finally scanned using the Gene Chip scanner 3000 7G. AGCC automatically generated cell intensity data (CEL files) from each array.

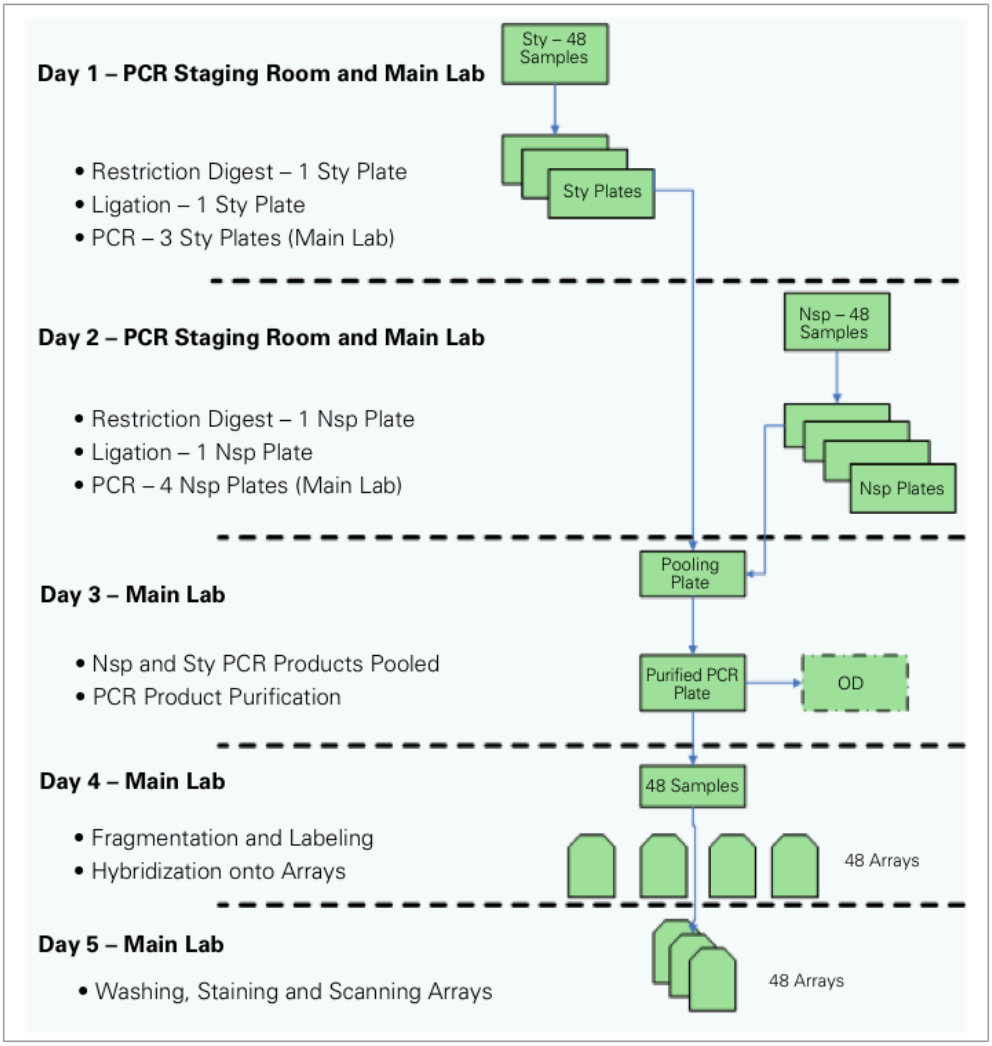


Figure 18. Workflow recommended for processing 48 samples. *Figure taken from*AffymetrixGenome-Wide Human SNP Nsp/Sty 6.0 User Guide.

Data Analysis

Data analysis has been performed following the workflow reported in Fig 19. Genomic profiles were obtained using Affymetrix Human Mapping GeneChip6.0 arrays (Affymetrix, Santa Clara, CA, USA). Initial array intensity data was obtained using Affymetrix GeneChip Command Console Software (AGCC) with 270 Caucasian HapMap samples as a reference dataset for normalization. Raw copy number (CN) and allelic imbalance (AI) were extracted from CEL files using the Affymetrix Power Tools software package

with standard parameters, (Affymetrix,Santa Clara, CA). We applied 'probe weeding' by removing probes (about 20%) with an interprobe distance of less than 180 bp. Without probe weeding we experienced that hot spots of high probe density were prone to artifacts of tiny amplifications or deletions. A minimum probe distance of 180 bp ensures that probes are more or less independent from the PCR fragment size. Allelic imbalance (AI) was corrected for copy number. The copy number for the X chromosome was centered to the median value.

Because of the sensitivity to genomics waves and tendency of hypersegmentation of existing algorithms, we devised a newsegmentation algorithm 'Fast first-derivative segmentation' (FFSEG; Kwee et al, submitted) based on edge detection using first-derivatives of the raw copy number signal. FFSEG is considerably more robust against genomics waves and hypersegmentation, is extremely fast and compares well with the established algorithms such as mBPCR [218] and circular binary segmentation [219]. For details on mBPCR and CBS, see references.

FFSEG uses non-segmented, raw copy number data that has not necessarily been corrected for genomic waves. For each chromosome, FFSEG locally computes first derivatives by convolving the raw signal with the SGED kernel

$$F(x) = -\text{sign}(x) * \exp(-(x/n)^{**}2)$$

where x is the position index and n is a scale parameter. This choice of kernel was preferred rather than the more commonly used derivative of a Gaussian because it showed better localization of the edge positions due to the sharp boundary at x=0. Candidate edges were defined as statistically significant outliers, $p<0.05$ after Bonferroni multiple test correction. The detection of edges was repeated at multiple scales, n = (8, 16, 32, 64, 128, 256, 512, 1024, 2048). We found that this set of scales detected most edges in our test cases. The set of edges were then combined and duplicate positions were removed. This final set served as candidate positions of true chromosomal breakpoints. A subsequent pruning step removed breakpoint candidates by testing the mean signal of its flanking regions (t-test, $p<0.05$). Finally, the copy number level for each region is estimated by its mean value. For larger scales, we firstly bin the raw signal onto a coarse resolution, detect the approximate positions of edge candidates and refine the edges by re-estimating their positions with the SGED kernel using the full resolution of probes. This sped up the computation considerably without loss of accuracy.

We applied a multiple filtering strategy to minimize false positives. Segments containing less than 10 probes or with a size smaller than 2000 bp were discarded. Segments were filtered for copy number variation if they overlapped at least 50% with a known CNV according to the Database of Genomic Variants (http://projects.tcag.ca/variation/). Segments spanning the centromere were filtered out.

After CNV filtering, we employed 'gap filling'. Small gaps of less than 500 kb and flanked by aberrated regions larger than the gap size itself were closed. Larger gaps between 500 kb and 10 Mb were filled if they were flanked by regions with at least 1.5 times their size. Finally, neighbouring segments were iteratively merged if their difference was smaller than 0.15 (log2ratio).

As final quality control, filtered and merged segmented profiles, together with their raw profiles, were visually inspected. Profiles were considered as poor quality and discarded from further analyses if they showed severe over-segmentation or no aberrations at all as evaluated by two independent investigators.

Copy number segments were discretized as heterozygous deleted (HDEL), loss (LOSS), normal (NORM), gain (GAIN), or amplified (AMPL) using thethresholds -0.60, -0.10, 0.10 and 0.60 (log2ratio). AII segments were classified as LOH if the absolute AI was >0.25.

Clustered heatmaps were generated using an Euclidean distance and Ward linkage on the discretized CN/LOH values after resampling segments uniformly to about 10000 probes. The X/Y chromosomes were omitted for computing the clustering.

Partial frequencies were computed for LOH and each CN aberration class. The overall frequencies of LOSS and GAIN, correspond to the cumulative frequencies of (LOSS+HDEL) and (GAIN+AMPL), respectively.

Minimal common regions (MCR) were calculated according to the algorithm by Lenz et al [220]. Briefly, the MCR algorithm iteratively searches for high frequency peaks and determines the smallest region of maximum overlap as its corresponding MCR. 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 maximum segment size of 25 Mb and gap size of 500 kb for SRA analysis, and a minimal segment size of 15 Mb and gap size of 10 Mb for LRA analysis were applied [220]. MCR frequencies were calculated as the fraction of samples bearing the same aberration contained in the MCR without regard to its MCR type. MCRs containing genes encoding the immunoglobulin heavy chain (IGHV) genes and the kappa and lambda light chains were discarded since CN changes in these regions probably represent

physiological rearrangements occurring in B-cells. Focal abberations were addtionally identified using GISTIC 2.0 [221]on the filtered and merged segments.

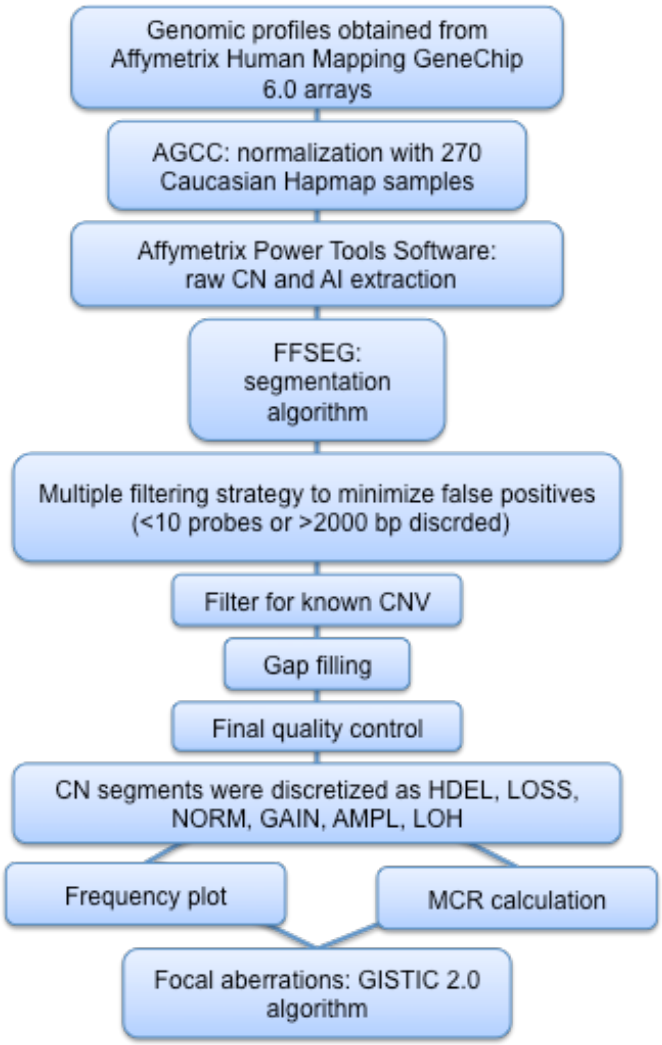


Figure 19. Workflow followed for data analysis.

Mutational Analysis

Whole Genome Amplification (WGA) of genomic DNA was performed with REPLI-g Mini/Midi Kit (QIAGEN, Hilden, Germany). Briefly, 10 ng of genomic DNA was denatured by adding denaturation buffer. After denaturation has been stopped by addition of neutralization buffer, a master mix containing buffer and DNA polymerase is added. The isothermal amplification reaction proceeded overnight at 30°C.

One μl of WGA product 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, Basel, Switzerland) and 10 pmol of each primer. Polymerase Chain Reaction (PCR) reactions were composed by: one activation cycle of 95°C for 10', 35 cycles of a denaturing step of 95°C for 30", a specific annealing step for each couple of primers, and an elongation step of 72°C for 30", followed by a final elongation step of 72°C for 7'. Primers and specific annealing conditions for PCR amplification of all *PRDM1* coding exons and *TP53* (exons from 4 to 8) are described in Table 3. The PCR products were purified using QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany). Purified amplicons were sequenced directly from both strands (Mycrosynth, Switzerland) and compared with corresponding germ-line sequences (NM_001198.3, BLIMP1; NM_000546.5, TP53) with BLAST software. All mutations were confirmed on independent PCR products, and germ-line polymorphisms, including changes listed in the NCBI SNP database or present in available matched normal DNA, were excluded. Identity of matched normal DNA was verified by analyzing known polymorphisms.

Table 3. Primers for PRDM1 and TP53 sequencing.

GENE	Primer FW	Primer REV	annealing
PRDM1-ex1	5'-TGACGCCAAACACATGTTAAA-3'	5'-GTTCCAGCTCACACTCGTCA-3'	60°C 30 sec
PRDM1-ex2	5'-TATACGGCTTCTTGGCTCTT-3'	5'-AGGAACAGTTGAAGGCTGG-3'	58°C 1 min
PRDM1-ex3	5'-AGATGGTCTCCCCCTATGGT-3'	5'-AAGCAAGCAACAAACTGTTTC-3'	58°C 1 min
PRDM1-ex4	5'-GCCCTGATTTCTGCTGATTC-3'	5'-GTCCCTAGCTTAAGCCACCT-3'	60°C 45 sec
PRDM1-ex1b	5'-TAGATGTTCATCCCGTTCTGA-3'	5'-ACTTGAGAATGACCAAAATG-3'	56°C 1 min 30 sec
PRDM1-ex5	5'-TTGAGTGAGTGGCCCAGAG-3'	5'-AGGGAAGTCACTTGTCCAAA-3'	60°C 1 min 30 sec
PRDM1-ex6	5'-AAACTCCCTGCTAGCCTGTG-3'	5'-GCCATCTCAAGTCATCAGCA-3'	60°C 45 sec
PRDM1-ex7	5'-CACAAGGAGGCTTCTCACCT-3'	5'-GATTTCAGTAACTTTGGAGTT-3'	58°C 1 min
TP53-ex4	5'-TCCTCTGACTGCTCTTTTCAC-3'	5'-TGAAGTCTCATGGAAGCCAG-3'	60°C 30 sec
TP53-ex5	5'- GTTTCTTTGCTGCCGTCTTC-3'	5'-AGCAATCAGTGAGGAATCAG-3'	58°C 30 sec
TP53-ex6	5'-TCTGATTCCTCACTGATTGCTC-3'	5'-CCACTGACAACCACCCTTAAC-3'	61°C 30 sec
TP53-ex7	5'-TCATCTTGGGCCTGTGTTATC-3'	5'-AGTGTGCAGGGTGGCAAG-3'	60°C 30 sec
TP53-ex8	5'-AGGACCTGATTTCCTTACTGCC-3'	5'-ATAACTGCACCCTTGGTCTCC-3'	60°C 30 sec

Methylation

The methylation status of BLIMP1 was investigated by bisulfite sequencing PCR method. 500 ng of DNA used as template was chemically modified by bisulfite treatment (EZ DNA Methylation Kit, ZYMO Research, Irvine, CA, USA). This technique involves treating methylated DNA with bisulfite, which converts unmethylated cytosines into uracil. Methylated cytosins remain unchanged during the treatment. Once converted, the methylation profile of DNA can be determined by PCR amplification followed by DNA sequencing. PCR amplification was performed with primers specific for converted bisulfite converted DNA, designed with Methprimer tool. 2 μl of bisulfite converted DNA 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, Basel, Switzerland) and 30 pmol of each primer. The thermal profile consisted in: one activation cycle of 95°C for 10', 35 cycles of a denaturing step of 95°C for 30", an annealing step of 58°C for 30", and an elongation step of 72°C for 30", followed by a final elongation step of 72°C for 7'. PCR products were separated by agarose gel electrophoresis and visualized with Alphimager 3400 (Alphainnotech Corporation, Fremont, CA, USA). Negative control (DNA not treated) was used in each reaction. The PCR products were purified using

QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany). Purified amplicons were sequenced directly from both strands (Mycrosynth, Switzerland) and analyzed for the methylation detection.

2. HUMAN CELL CULTURE

Cell Lines

We used different established human cell lines derived from Anaplastic Large Cell Lymphoma (ALCL): SUDHL1, Supm2, TS, JB6, L82, Karpas 299 (ALK+ ALCL) and FE-PD, MAC-1 (cALCL, commonly used as ALK- ALCL model). All of them were cultured with RPMI 1640 media supplemented with fetal bovine serum (10%) (PAA, Pasching, Austria) and Penicillin-Streptomycin-Neomycin (~5,000 units penicillin, 5 mg streptomycin and 10 mg neomycin/mL, Sigma-Aldrich, St Louis, MO, USA).
HEK-293T packaging cells (ATCC) were cultured in DMEM (GIBCO) medium complemented with 10% of FBS (PAA) and Penicillin-Streptomycin-Neomycin (~5,000 units penicillin, 5 mg streptomycin and 10 mg neomycin/mL, Sigma-Aldrich, St Louis, MO, USA). All the cell lines have undergone karyotypes analysis (conventional or with the addition of FISH probes when needed, done by our collaborator M.G. Tibiletti) and search for Mycoplasma infection by biochemical assay (Lonza AG, Cologne, Germany). Controls have been repeated during the course of the project.

Transfection

Lipofectamine is a common trasfection reagent (Invitrogen, Life technologies, Zug, Switzerland) used to introduce siRNA or plasmid DNA into *in vitro* cell cultures by lipofection. Lipofectamine reagent contains lipid subunits that can form liposomes in an aquaeous environment, which capture the transfection materials, such as DNA plasmids. The DNA-containing liposomes can fuse with the plasma membrane of living cells, create pores along the membrane allowing nucleic acids to cross into the cytoplasm and contents to be available to the cell for replication or expression. For each condition, 0.4×10^6 or 0.8×10^6 cells were plated in a 24-wells plate in 500 μl of medium. Two different mixes were prepared: the first one containing 2 μl of lipofectamine in 50 μl of Optimem medium (GIBCO, Life technologies, Zug, Switzerland) and the second one containing 1 μg of DNA plasmids in 50 μl of Optimem medium. After having incubated both mixes for

5' at room temperature, we combined them and incubated the final mix for 20' at room temperature to allow the formation of complexes between DNA and lipofectamine. Then, we added the 100 μl DNA-lipofectamine complexes directly to each well containing cells and mixed gently by rocking the plate back and forth. Cells were incubated for 4 hours at 37°C in a CO_2 incubator. We tried different conditions of transfection, removing and changing the medium or not. At different time point, cells were controlled for the expression of GFP with a fluorescence microscope.

Electroporation

Transfection of ALCL cell lines was carried out also by electroporation using a Nucleofector device (Amaxa, Lonza AG, Cologne, Germany). For each condition, 5×10^6 cells were resuspended in 82 μl of Nucleofector solution V (Nucleofector kit) together with 18 μl of supplements and 2 μg of plasmid used for control (pWPI) or plasmid for the re-expression of BLIMP1 (pWPI-HA-BLIMP1) (kindly provided by Pasqualucci L.), both of them containing green fluorescent protein (GFP). Cells were then placed in an electroporation cuvette. Transfection with each plasmid was carried out with the application of different electroporation programs, G-016 and A-030. Immediately after electroporation, 500 μl of prewarmed RPMI medium with 10% FBS was added to the cuvette and cells were transferred into culture plates containing prewarmed medium. At different time point, cells were controlled for the expression of GFP with a fluorescence microscope.

Lentiviral Infection

Lentivirus supernatant, containing vector for BLIMP1 expression (pWPI-HA-Blimp1) or empty vector (pWPI) were produced by co-trasfecting HEK-293T packaging cells with pWPI transfer vector (4.5 mg) in combination with third generation helper plasmids (pCMV-dR8.74 vector: 3.30 mg and pMD2.VSVG: 1.2 mg) using jetPRIME reagent (Polyplus transfection, Illkirch, France) in six-well plates.

The viral supernatant was collected after 48 hours, filtered, ultracentrifugated and used directly to infect ALCL cell lines. The infection was performed using 24-well plates and a dose of 6 μg/ml of polybrene (Millipore, Billerica, MA, USA) was added to each well to allow the interaction between virus and target cells. One cycle of spin infection (1800 rpm; 45'; 30°C) was performed and cells were left for 8 hours at 37°C, 5% CO_2 to allow the

interaction between virus and target cells. Then the cells were washed with PBS 1X and re-suspended in fresh complete medium. Cells were collected at different days after infection for different assays, RNA and proteins extracted and then processed respectively for qRT-PCR and Western Blotting analysis.

Evaluation of Infection Efficiency

Cells infected with pWPI-HA-Blimp1 and control cells (cells infected with empty vector and cells not transfected) were harvested and washed once in PBS and analyzed using a FACScan flow cytometer (Becton Dickinson, NJ, USA). The analysis of the percentage of GFP positive cells was performed using the CellQuest Pro (Becton Dickinson, NJ, USA).

Cell proliferation

Cells were recovered and an equal number of cells were plated in triplicates. Cells were collected, diluted 1:1 with Trypan Blue and counted daily up to 10 days after the recovery.

Apoptosis Assay

Cells were harvested and washed once in PBS and then stained with propidium iodide (PI 1 μg/ml, Sigma-Aldrich, St Louis, MO, USA) in PBS and analyzed using a FACScan flow cytometer (Becton Dickinson, NJ, USA). The analysis of the percentage of cell death was performed using the CellQuest Pro (Becton Dickinson, NJ, USA).

Cell Cycle Analysis

Cells were harvested at different time points, washed once in PBS and then fixed in 70% ethanol at 4^{o}C for at least one hour. Cells were stained with propidium iodide (PI 50 μg/ml, Sigma-Aldrich, St Louis, MO, USA) or 7-Aminoactinomycin D (7AAD 50 μg/ml, BD-Pharmingen, NJ, USA) in PBS containing RNAse-A (75 kU/ml, Sigma-Aldrich, St Louis, MO, USA) and analyzed for DNA content using a FACScan flow cytometer (Becton Dickinson, NJ, USA). The analysis of cell cycle will be performed using the FlowJo software (TreeStar, Ashland, OR, USA).

3. IMMUNOBLOTTING AND IMMUNOISTOCHEMISTRY

Western Blotting

Cells were solubilized in hot SDS lysis buffer (2.5% SDS, Tris-Hcl 250 mM pH 7.5) and sonicated for 10". The protein content in the different samples was determined using the BCA protein assay (Pierce Chemical Co., Thermo Scientific, Rockford, IL, USA). This assay enables a colorimetric detection and quantification of total protein based on bicinchoninic acid (BCA). The method combined the reduction of Cu^{2+} 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 cupros cation (Cu+) gives a purple coloration and the complex strongly absorbs at 570 nm. Absorbance can be detected by 96-well plate reader (Beckman Coulter AD340, Brea, CA, USA), and it is directly proportional to the protein concentration. Lysates (25 μg) were fractionated by SDS-PAGE using 8%-10% polyacrylamide gels, based upon the expected molecular weight. The resolved proteins were blotted to a nitrocellulose membrane by electric transfer, and the membranes were blocked in 0.2% I-Block (Apllied Biosystem, Foster City, CA, USA) in PBS for one hour at room temperature. Membranes were incubated overnight at 4°C with primary antibodies prepared in TBS-T (20 mM Tris-HCl [pH 7.6], 137 mM NaCl, 0.5% Tween 20) according to optimized dilutions. Membranes were washed in TBS-T three times for 10' each and then incubated in TBS-T containing the appropriate horseradish peroxidase-conjugated anti-mouse, anti-rabbit or anti-rat secondary antibodies (Amersham Life Science, GE Healthcare, Glattbrugg, Switzerland) for one hour at room temperature. The membranes were washed three times for 10' each in TBS-T and then processed for enhanced chemiluminescence detection according to the manufacturers instructions (ECL; Amersham Life Science, GE Healthcare, Glattbrugg, Switzerland). Equal loading of samples was confirmed by probing for GAPDH or β-tubulin.

The following antibodies were used:

anti-BLIMP1 (6D3, SC-47732, Santa Cruz, CA, USA), 1:1000, rat
anti-α-GAPDH (MAB374, Millipore, Billerica, MA, USA), 1:10000, mouse
anti -β-tubulin (H-235, Santa Cruz, CA, USA), 1:1000, rabbit

Immunohistochemistry

Formalin-fixed, paraffin-embedded sections of ALCL cases (n=38) from the University of Turin, Department of Pathology were stained on a semi-automated immunostainer (Leika, Bond-Max) [124]. Briefly paraffin-embedded sections were first dewaxed. Epitope retrieval was performed in 0.01M citrate buffer, pH 6.0. Antibodies used were anti-PRDM1 (mouse monoclonal ROS195G, kindly provided by Dr. G. Roncador, 1:20). TMA samples were considered negative or positive for BLIMP1 depending on intensity of the staining: BLIMP1 staining was considered negative if positivity was detected in <10% of cells, moderately positive if detected in 10-50% of cells and positive if detected in >50% of cells. Slides were analyzed by light microscopy and reviewed by two independent operators (GI and AB).

4. QUANTITATIVE REALTIME-PCR ANALYSIS

RNA extraction and retrotranscription

Total RNA was extracted using the TRIZOL reagent (Invitrogen, Life technologies, Zug, Switzerland) and purified using the RNeasy total RNA Isolation Kit (QIAGEN, Hilden, Germany). The concentration of total RNA was determined spectrophotometrically at 260 nm using the NanoDrop (NanoDrop Technologies, Wilmington, DE, USA). One microgram of total RNA was reverse-transcribed using the Superscript First-Strand Synthesis System for RT-PCR kit (Invitrogen, Life technologies, Zug, Switzerland). Briefly, one microgram of total RNA was combined to 10 μl of 2X RT Reaction Mix and 2 μl RT Enzyme Mix and the final volume of 20 μl was reached with DEPC water. The mix was gently mixed and incubated at 25°C for 10', at 50°C for 30' and at 85°C for 5' in a Tetrad 2 Peltier Thermal Cycler (BIO-RAD, Hercules, CA, USA). The reaction was then chilled on ice and 1 μl RNase-H was added in each tube, with subsequent incubation at 37°C for 20'. The results were diluted 1:5 with DEPC water.

Quantitative Real-Time PCR

Quantitative Real Time PCR amplification was performed using Fast SYBR Green Master Mix on a StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA,

USA). cDNA solution (1 μl) was amplified in a 10 μl solution containing 2X Fast SYBR Green Master Mix, 0.3 pmol of each primer and DEPC water to reach the final volume. Primer sets (Table 4) were designed using Primer3 [222]. Polymerase chain reactions were performed using a 96-wells plate and optical adhesive covers (Applied Biosystem, Foster City, CA, USA). The following program was run on the thermal cycler: 95°C for 15', followed by 40 cycles of 95°C for 15" and 60°C for 1'. In order to ensure that the disired amplicon was detected, a dissociation curve was run following the Real Time PCR. All samples were analyzed in triplicates. A no-template control (NTC) was added to each plate and for each assay, and contained all PCR reagents without any template DNA. The relative quantity of the specific mRNA for each sample was calculated based on mean Ct values using the delta-delta Ct with a correction for experimental variations by normalization to the housekeeping gene GAPDH.

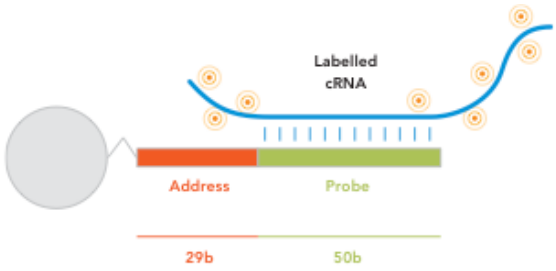
Table 4. Primer sets used for Quantitative Real Time PCR

GENE	Primer FW	Primer REV
PRDM1	5'-ACATGACCGGCTACAAGACC-3'	5'-GGCATTCATGTGGCTTTTCT-3'
GAPDH	5'-CGACCACTTTGTCAAGCTCA-3'	5'-CCCTGTTGCTGTAGCCAAAT-3'
PMAIP1	5'-GAGATGCCTGGGAAGAAGG-3'	5'-TTCTGCCGGAAGTTCAGTTT-3'
FGG	5'-TGCATTAAGAGTGGAACTGGAA-3'	5'-TGTTAGGCGGTACTTGTCAG-3'
SERPINA3	5'-GTTCAGAGAGATAGGTGAGC-3'	5'-CTGGTGAAGGCTTCCTCAAT-3'
SHIP1	5'-CCCTGCAAGAAATCACCAGT-3'	5'-ATCCGGTTCTCGTGCTCAG-3'

5. GENE EXPRESSION

Gene Expression Profiling (GEP)

Whole-genome gene expression was assessed using a direct hybridization approach and the general workflow is presented in Fig 20 and Fig 21. Illumina Whole-Genome Gene Expression BeadChip (Illumina, San Diego, CA, USA) consists of oligonucleotides immobilized to beads held in microwells on the surface of an array substrate (Scheme 1).



***Scheme 1**. Illustration of the binding of labelled cRNA to the oligonucleotides immobilized to beads on the array.*

500 ng of RNA starting material was amplified using a single-round in vitro transcription (IVT) amplification with the Illumina TotalPrep RNA Amplification Kit (Ambion, Life technologies, Zug, Switzerland). The Illumina RNA Amplification Kit is a complete system for generating biotinylated, amplified RNA for hybridization and was developed based on the principles of the popular Eberwine technique, using reverse transcription followed by a T7 RNA polymerase-based linear amplification step. Briefly, 11 μl of RNA appropriately diluted in water, were mixed with a Reverse Transcription Master Mix, containing T7 Oligo(dT) primer, 10X First Strand Buffer, dNTP Mix, RNase Inhibitor and Array Script reaching the final volume of 20 μl. The reaction was incubated at 42°C for two hours. A Second Strand Master Mix was prepared with Nuclease-free Water, 10X Second Strand Buffer, dNTP Mix, DNA Polymerase and RNase H. 80 μl of the second mix were transferred to each sample. Tubes were placed in a thermal cycler and were incubated at 16°C for two hours. Then the reactions were placed on ice and the cDNA obtained was purified (Fig 20). We added 250 μl of cDNA Binding Buffer to each sample and mixed the reaction which was then pipetted on the centre of cDNA Filter Cartridge. After having centrifuged them and washed the cDNA Filter Cartridge, the cDNA was eluted with 20 μl of 55°C pre-heated Nuclease-free Water. The cDNA sample was transferred in a PCR tube and 7.5 μl of IVT Master Mix (T7 10X Reaction Buffer, T7 Enzyme Mix and Biotin-NTP MIX) were added to cDNA. Once assembled, tubes were placed in a thermal cycler at 37°C for 14 hours. The reaction was stopped by adding 75 μl of Nuclease-free Water to each cRNA sample to bring the final volume to 100 μl. The cRNA purification was performed adding 350 μl of cRNA Binding Buffer to each cRNA sample. 250 μl of ACS reagent grade 100% ethanol were added to each sample and each mixture was pipetted

onto the center of the filter in the cRNA Filter Cartridge. After centrifugation and a washing of the filter, cRNA was eluted with 200 μl of 55°C pre-heated Nuclease-free Water.

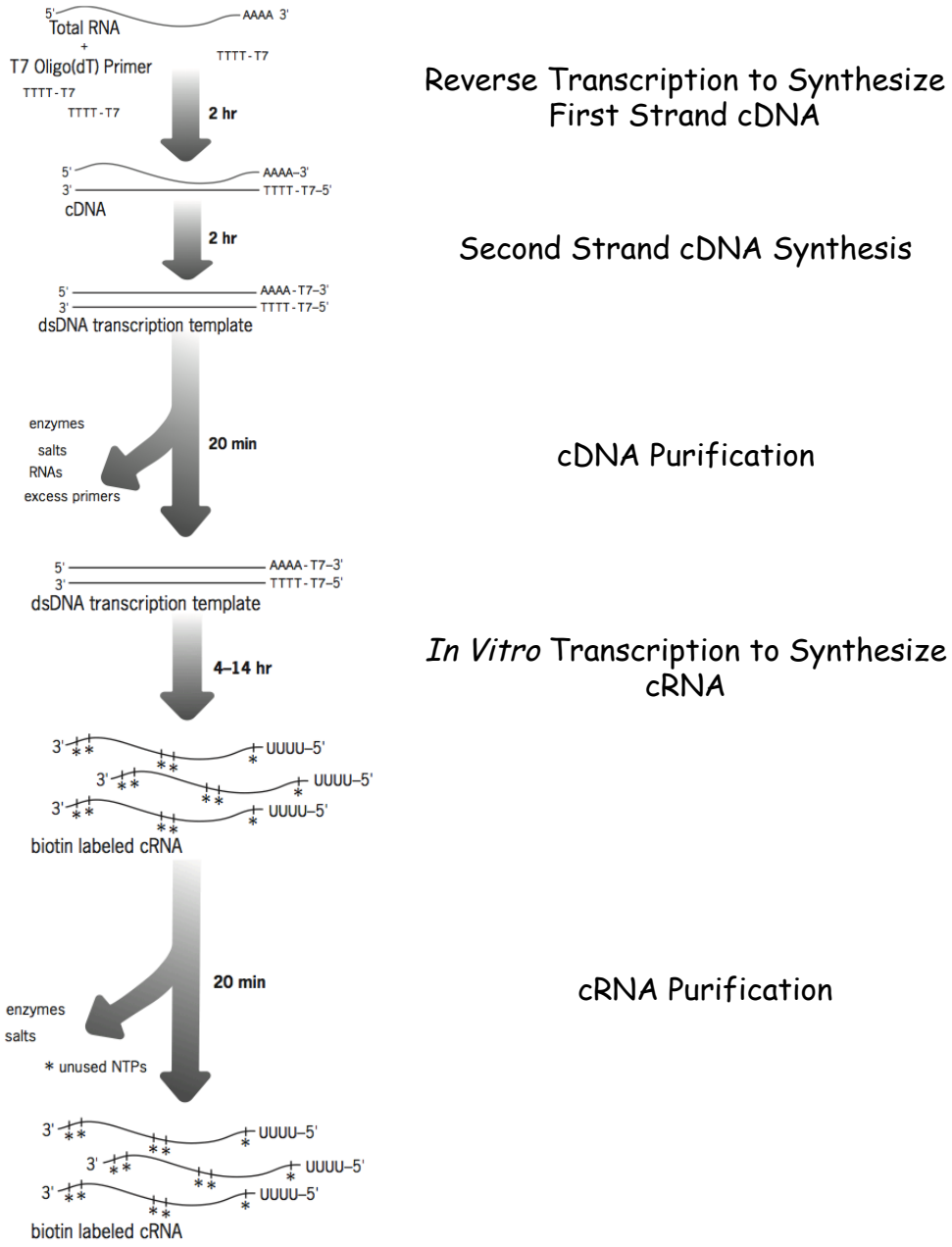


Figure 20. Illumina Total Prep RNA Amplification Procedure. *Figure adapted from Illumina TotalPrep RNA Amplification Kit, Protocol.*

The next step forecast the cRNA hybridization (Fig 21). The cRNA samples were preheated at 65°C for 5'. After they have been cooled, the appropriate volume of cRNA for each sample was added to Hyb Mix and RNase-free Water and the mixture was hybridized to the BeadChip.

The Human HT-12 v4 Expression BeadChip content provides genome-wide transcriptional coverage of well-characterized genes, gene candidates, and splice variants, delivering high-throughput processing of 2 samples per BeadChip. Each array on the Human HT-12 v4 Expression BeadChip targets more than 47,000 probes derived from the NCBI RefSeq Release 38 (November 7, 2009).

To allow the hybridization of cRNA to the BeadChip, the latter one was placed into the 58°C Illumina Hybridization Oven for 14 hours. Washings were performed in Hybex Waterbath as described in the Whole-Genome Gene Expression Direct Hybridization Assay Guide. In the end, BeadChips were placed in a Wash Tray where Block E1 buffer was added, and after 10', they were washed and dryed (Fig 21).

In the last step, Cy3-SA is introduced to bind to the analytical probes that have been hybridized to the BeadChip. This allows for differential detection of signals when the BeadChips are scanned. The iScan Reader uses a laser to excite the fluor of the single-base extension product on the beads of the BeadChip sections. Light emissions from these fluors are then recorded in high-resolution images of the BeadChip sections. Data from these images are analyzed using Illumina's GenomeStudio Gene Expression Module.

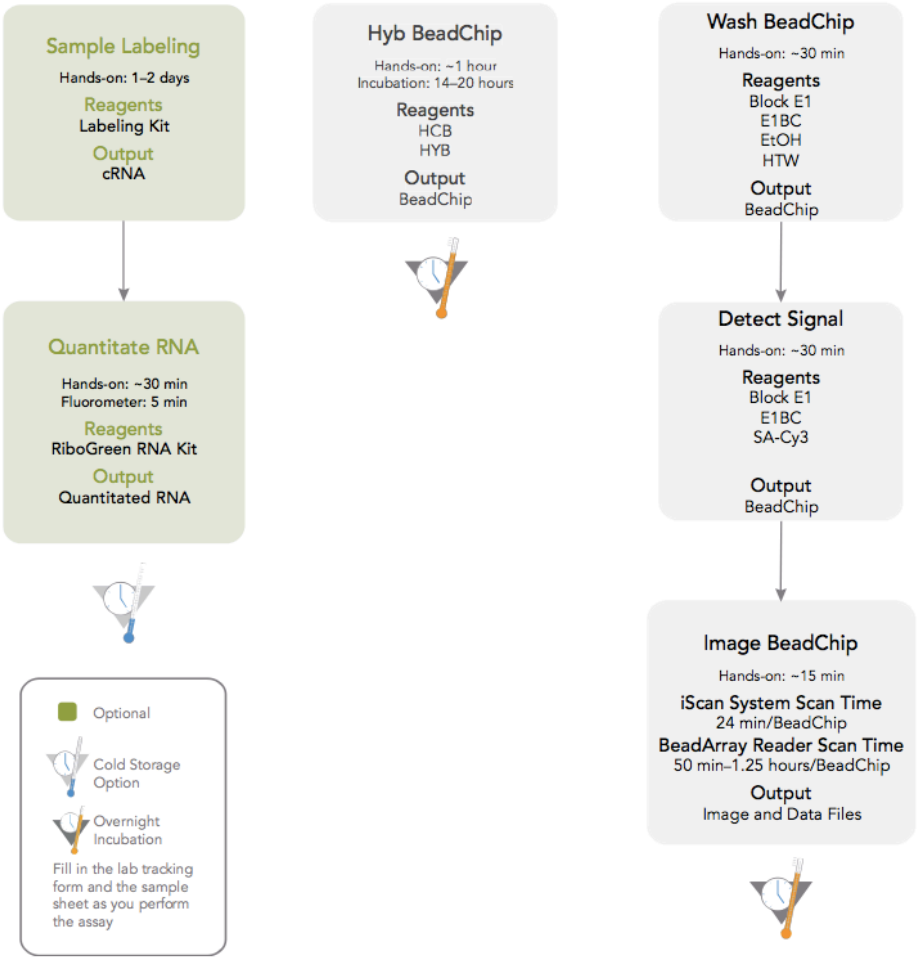


Figure 21. Direct Hybridization Assay Workflow. *Figure taken from Whole-Genome Gene Expression Direct Hybridization Assay Guide.*

Partek Analyis

GEP data were analyzed using Partek Genomic Suite (Partek, St Louis, Mi, USA). Signal intensities were normalized by Partek RMA. After quantile normalization and the removal of batch effect have been applied, the analysis of variance (ANOVA) was used to identify significantly expressed (both up and down regulated) genes, and a gene list was generated for subsequent bioinformatic analysis.

Gene Set Enrichment Analysis (GSEA)

Gene Set Enrichment Analysis (GSEA) is a computational method that determines whether an *a priori* defined set of genes shows statistically significant, concordant differences between two biological states (http://www.broadinstitute.org/gsea/index.jsp). The method derives its power by focusing on gene sets, that is, groups of genes that share common biological function, chromosomal location, or regulation. We compared the gene expression levels from 2 different phenotypes (TS cells after BLIMP1 re-expression versus TS cells not infected together with TS cells transduced with empty vector) and picked up the genes which had significant different expression for GSEA by using Molecular Signatures Database (V3.0). Gene set enrichment analysis was carried out by computing overlaps with canonical pathways (CP) and gene ontology (GO) obtained from the Broad Institute. Genes in Gene Set, Genes in Overlap, and P value were used to rank the pathways enriched in each phenotype.

Database for Annotation, Visualization and Integrated Discovery (DAVID) tool

Gene Ontology (GO) term enrichment was performed using DAVID bioinformatics resources (http://david.niaid.nih.gov) to determine whether particular gene ontology terms occur more frequently than expected by chance in a given set of genes. DAVID tool is able to provide a rapid means to reduce large lists of genes into functionally related groups of genes to help unravel the biological content captured by high throughput technologies. We used default settings for the category GOTERM_BP_ALL, and selected those terms with $P<0.05$ (for FDR no greater than 5%) representing central biochemical pathways/metabolic functions.

6. IN VIVO EXPERIMENTS

Supm2 cell line, a subclone of TS ALCL cell line, was infected with pWPI-HA-BLIMP1 lentivirus for the re-expression of BLIMP1 and with pWPI, empty vector for control. After the infection we checked the efficiency of transfection through FACS analysis of percentage of GFP positive cells. Male NSG (NOD/Shi-*scid*/IL-2Rγ^{null}) mice were injected subcutaneously with $1x10^6$ Supm2 cells implanted in both flanks. Three mice were implanted with cells infected with lentivirus for BLIMP1 and three mice were implanted

with cells infected with empty vector. At day 20, tumors were explanted and analyzed by flow cytometry for GFP expression. Tumor growth was measured over time. Mice were treated properly and ethically in accordance with European Community guidelines.

Results

I . GENOMIC PROFILING OF ANAPLASTIC LARGE CELL LYMPHOMA

Data reported in the following section have been submitted for publication as: *Boi M et al.* PRDM1/BLIMP1 Is Commonly Inactivated In Anaplastic Large T-Cell Lymphoma

1. Identification of genetic recurrent lesions in ALCL

To understand the pathogenetic events of ALCL, high-density genome-wide DNA profiling was performed on 69 ALCL samples, including 31 ALK- ALCL, 33 ALK+ ALCL and five cutaneous ALCL (cALCL). The same platform was also used for the analysis of seven ALCL cell lines: five derived from ALK+ and two from ALK- ALCL. Figure 22 shows the frequencies of gains and losses observed in ALCL cohort, while frequency plot of ALCL cell lines is reported in Figure 23. The frequency plots allow the estimation of frequency of all aberrations along the genome. For each chromosome short and long arm are recognizable. In the upper part of the graph, the frequency of gains and amplifications are reported (in red), while losses and homozygous deletions are represented in blue in the lower part of the frequency plot.

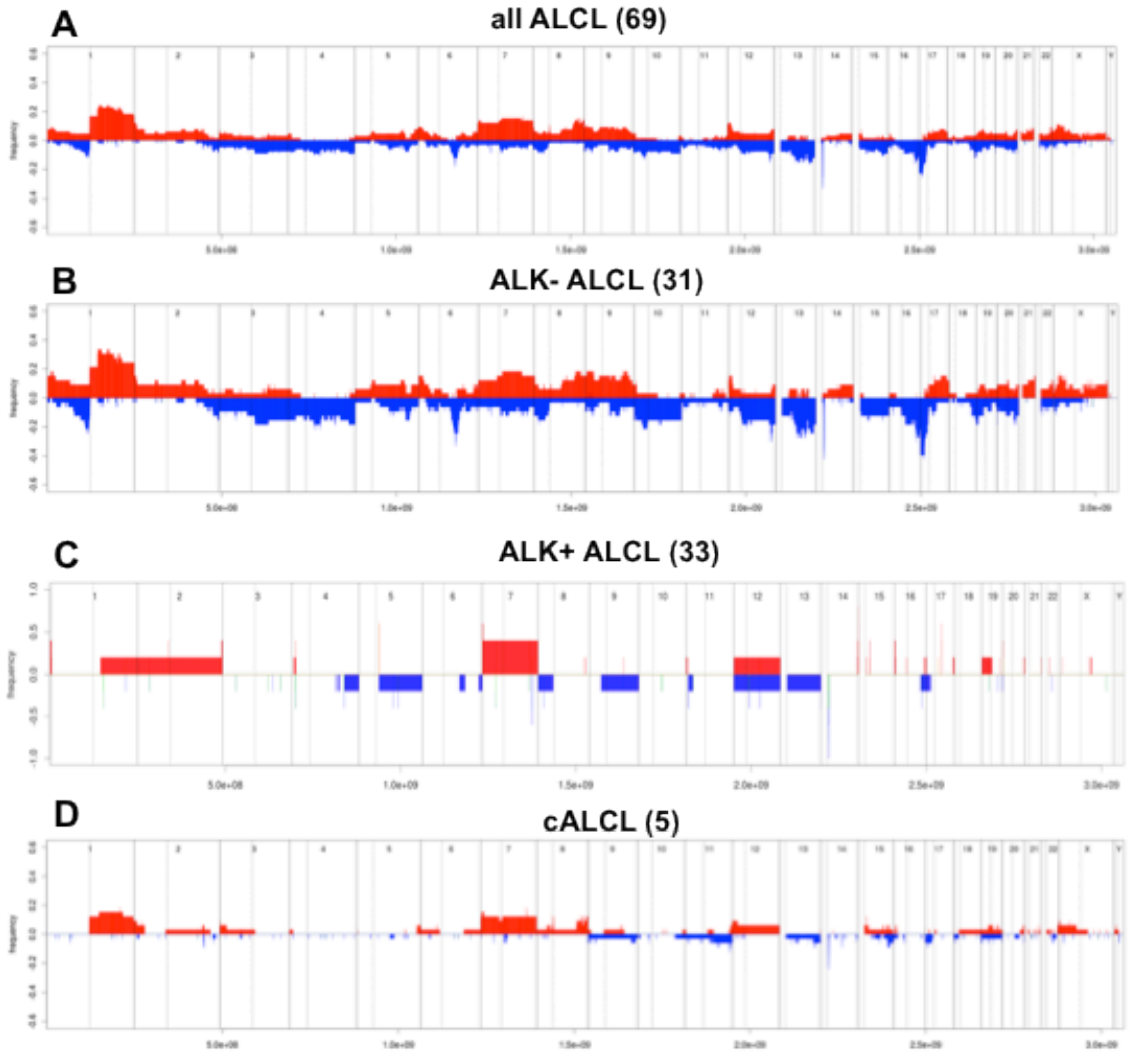


Figure 22. Frequency of DNA gains (up) and losses (down) observed in ALCL primary samples. *(A) 69 ALCL, 64 systemic ALCL and 5 cALCL; (B) 31 ALK- ALCL; (C) 33 ALK+ ALCL; (D) 5 cALCL samples. Red represents gains and blue represents losses. X-axis represents chromosome localization and physical mapping and Y-axis proportion of cases showing the aberrations.*

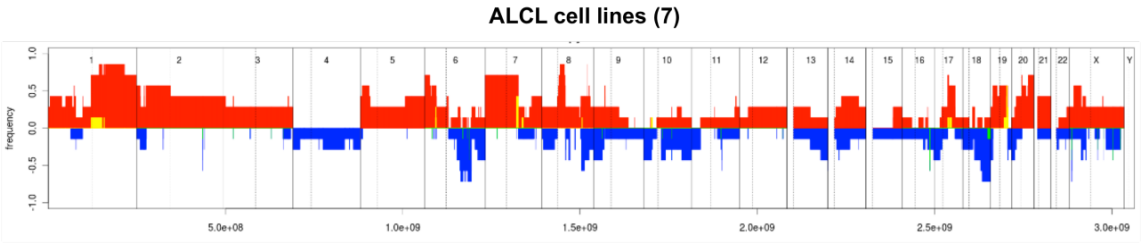


***Figure 23. Frequency of DNA gains (up) and losses (down) observed in 7 ALCL cell lines**. Red represents gains and blue represents losses. X-axis represents chromosome localization and physical mapping and Y-axis proportion of cases showing the aberrations.*

ALK+ and the ALK- ALCL clearly differed in terms of genomic profiles, as also indicated by the median number of aberrations per sample which was 14.4 among ALK+ and 80.2 among ALK- samples. In fact, ALK- ALCL subset shows a more complex genomic profiling respect to the ALK+ counterpart. Unfortunately the small number of cALCL samples did not permit us to perform further analysis, even if from the frequency plot obtained for these five samples we can observe some similarity with ALK- ALCL samples in terms of lesions (i.e. gain of 1q and chromosome 7).

In order to reduce the large amount of genomic profiling array data and to identify lesions which should encompass the loci likely to contain the genes implicated in the tumor pathogenesis, the minimal common regions (MCRs) were defined using the algorithm by Lenz et al [220]. Four distinct types of MCR were defined: short recurrent abnormality (SRA), long recurrent abnormality (LRA), abnormal chromosomal arm (ACA) and abnormal whole chromosome (AWC). With the aim to identify putative tumor suppressor genes or oncogenes, we focused our attention on the smallest MCRs defined : SRA infact generally involve a limited number of genes because of their small size. In this way the study of focused little lesions helps us in the identification of genes possibly having a role in the lymphomagenesis. MCRs were calculated for the whole series of 69 clinical ALCL specimens and then for the ALK- and ALK+ subgroups separately (Table 1 and 2).

Table 5. Most common MCRs observed in ALCL series. *Aberrations occurring in more than 15% of cases are reported.*

Hystotype	Lesion	Cytoband	Frequency, %	Start	Size_Kb	Genes
ALCL	*Losses*	17p13.3-p12	25%	526	11747.22	>20 (*TP53*)
		6q21	19%	105944900	951.75	*ATG5, PRDM1*
		13q32.3-q33.3	16%	99090837	9349.42	>20
		16q23.2	16%	78371498	1719.63	*MAF, WWOX*
	Gains	1q	23%	121344093	127880.29	>20
		1q25.1	23%	172126982	283.85	*C1orf105, DNM3, PIGC*
		1q32.2	21%	208620761	152.76	
		7q32.3	17%	130574632	236.34	*FLJ43663, LOC646329, MKLN1*
		7	16%	43259	159076.46	>20
		8q24.22	16%	133799280	919.4	*NDRG1, PHF20L1, SLA, ST3GAL1, TG, WISP1*

Table 6. Most common MCRs observed in ALK- ALCL and ALK+ ALCL subgroups.

Aberrations occurring in more than 20% of ALK-ALCL cases and in more than 15% of ALK+ALCL cases are reported.

Hystotype	Lesion	Cytoband	Frequency, %	Start	Size_Kb	Genes
ALK- ALCL	*Losses*	17p13.3-p12	42%	526	1747.22	>20 (*TP53*)
		6q16.3-q21	35%	102407043	4489.61	*ATG5, BVES, C6orf112, GRIK2, HACE1, LIN28B, POPDC3, PRDM1, PREP*
		16q23.2	29%	78371498	1719.63	*MAF, WWOX*
		12q24.31	27%	121535350	1194.43	>20
		1p13.2-p12	26%	108242856	12136.48	>20
		17p	24%	526	20221.6	>20
		13q21.31-q21.32	24%	62417944	3957.48	*OR7E156P*
		10	21%	6760531	128746.17	>20
		13q31.3-q34	21%	93233408	21874.99	>20
		4q22.1	21%	89066075	1030.38	*ABCG2, FAM13A, FAM13AOS, HERC3, HERC5, HERC6, NAP1L5, PIGY, PPM1K, TIGD2*
		10p12.1	21%	26345258	1172.92	*ABI1, ACBD5, ANKRD26, APBB1IP, GAD2, LOC731789, MASTL, MYO3A, NCRNA00202, NCRNA00264, PDSS1, YME1L1*
		13q31.3-q33.3	21%	93233408	15213.76	>20
		13q31.1-q31.3	21%	82447782	6567.01	*LOC642345, SLITRK1, SLITRK5, SLITRK6*
		13q21.2-q21.31	21%	60452343	189.98	*DIAPH3*
		13q22.3-q31.1	21%	77714441	141.16	*MYCBP2*
		16q23.1	21%	77171365	432.83	*ADAMTS18, MON1B, SYCE1L*
	Gains	1p11.1-q32.1	30%	121344093	77304.61	>20
		1q25.1	30%	172126982	283.83	*C1orf105, DNM3, PIGC*
		1q32.1-q41	27%	201721363	11687.98	>20
ALK+ ALCL	*Gains*	1q32.2	19%	208620761	152.76	
		7q32.3	19%	130574632	236.34	*FLJ43663, LOC646329, MKLN1*
		1q	16%	121344093	127880.29	>20
		7p22.3-p21.3	16%	1162084	8064.89	>20

The commonest lost region is observed at level of 14q11.2 cytoband, where genes of the T-cell receptor (TCR), are located. Since physiologically, the generation of TCR implies an extensive rearrangement of the locus, including interstitial losses, loss of 14q11.2 was so excluded from the subsequent analysis. Among all the ALCLs, the most common losses affected 17p13.3-p12 (25%), in which *TP53,* a well-characterized tumor suppressor gene, is located, and 6q21 (19%), region that encompass *PRDM1* and *ATG5*. Other common losses were identified at level of 13q32.3-q33.3 and 16q23.2 (16%). More than 20% of ALCL patients presented gains of different regions of the long arm of chromosome 1 (1q, 1q25.1,1q32.2), while chromosome 7 and 8q24.22 were gained in 14% and 17% of cases, respectively.

The analysis of the MCRs in the ALK- patients revealed that the most common aberrations were the same identified in the whole ALCL series, but at a higher frequency. Losses of 17p13.3-p12 and 6q21 were observed in 42% and 35% respectively, and losses of distal regions of chromosome 13, 12q and 10 were identified in more than 20% of the cases. Gains at level of 1q were observed in almost 30% of ALK- cases. The ALK+ subset presented less lesions or at a lower frequency than what seen in ALK- cases: only gains of 1q, 1q32.2, 7q32.3 and 7p22.3-p21.3 were observed in more than 15% of samples. Figure 24 shows the most common MCRs presenting significant differences between ALK- and ALK+ ALCL. As reported, loss of 17p13.3-p12 and loss of 6q21 were two of the genomic aberrations most differentially represented in the two subsets, being more frequent in ALK- ALCL cases.

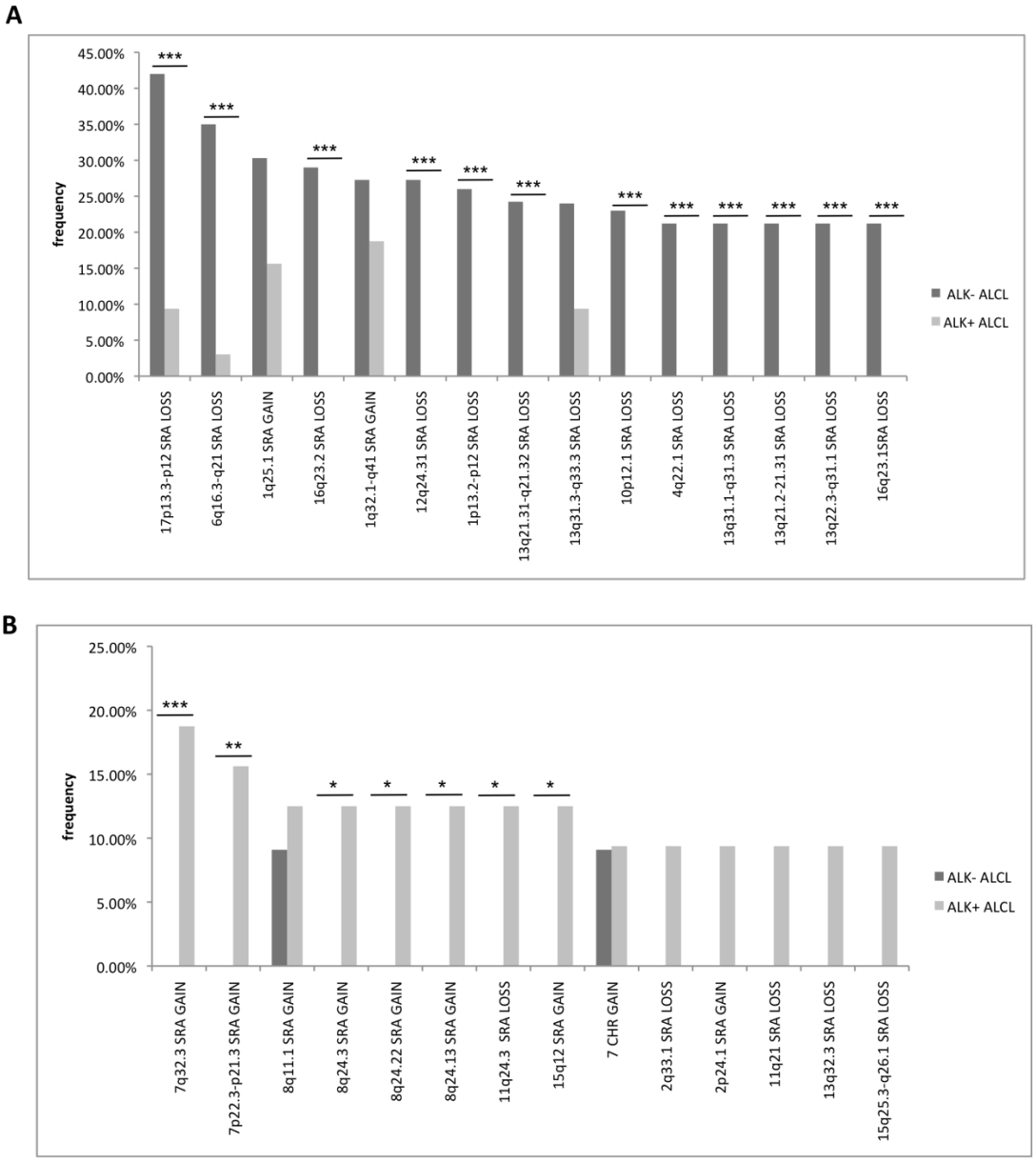


Figure 24. Comparison of the most common and significantly different MCRs between ALK- and ALK+ ALCL. *(A) more common MCRs in ALK- ALCL. (B) more common MCRs in ALK+ ALCL. The Y-axis reports the percentage of patients showing the aberration. Differences in genomic lesions between the two subsets were evaluated by applying Fisher's exact test (p-value). * p-value <0.05; ** p-value <0.01; *** p-value <0.001.*

To further define regions containing genes relevant for the disease, we also applied the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm, which identifies the aberrations occurring at a frequency higher than expected by chance, giving particular relevance to loci targeted by DNA amplifications or homozygous deletions [223]. Briefly, the algorithm identifies location and type of aberrations (amplification, gain, loss and homozygous deletion) and gives them a specific score on the basis of the frequency and the type of lesion. If the score of a specific aberration is higher than a significance threshold, calculated on the basis of the possibility that this event could be happened by chance, the lesion is thought to contain genes possibly involved in pathogenesis. Figure 25 shows an overview of GISTIC method.

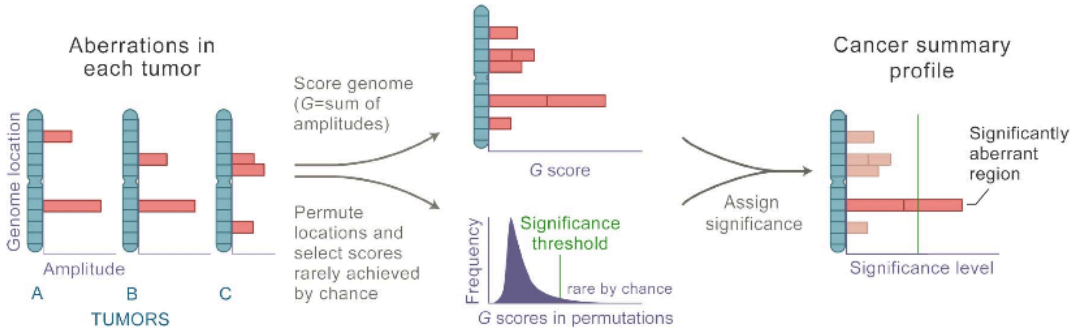


***Figure 25. Overview of the GISTIC method**. After identifying the locations and, in the case of copy-number alterations, magnitudes (as log_2 signal intensity ratios) of chromosomal aberrations in multiple tumors (Left), GISTIC scores each genomic marker with a G score that is proportional to the total magnitude of aberrations at each location (Upper Center). In addition, by permuting the locations in each tumor, GISTIC determines the frequency with which a given score would be attained if the events were due to chance and therefore randomly distributed (Lower Center). A significance threshold (green line) is determined such that significant scores are unlikely to occur by chance alone. Alterations are deemed significant if they occur in regions that surpass this threshold (Right). Adapted from Beroukhim R et al. PNAS 2007 [223].*

Figure 26, Table 7 and 8 show the regions of gains and of losses identified by GISTIC. Since ALK- ALCL pathogenetic events are still largely unknown, I focused my attention on ALK- ALCL GISTIC analysis, in order to understand if some significant regions could contain genes with a possible role in lymphomagenesis.

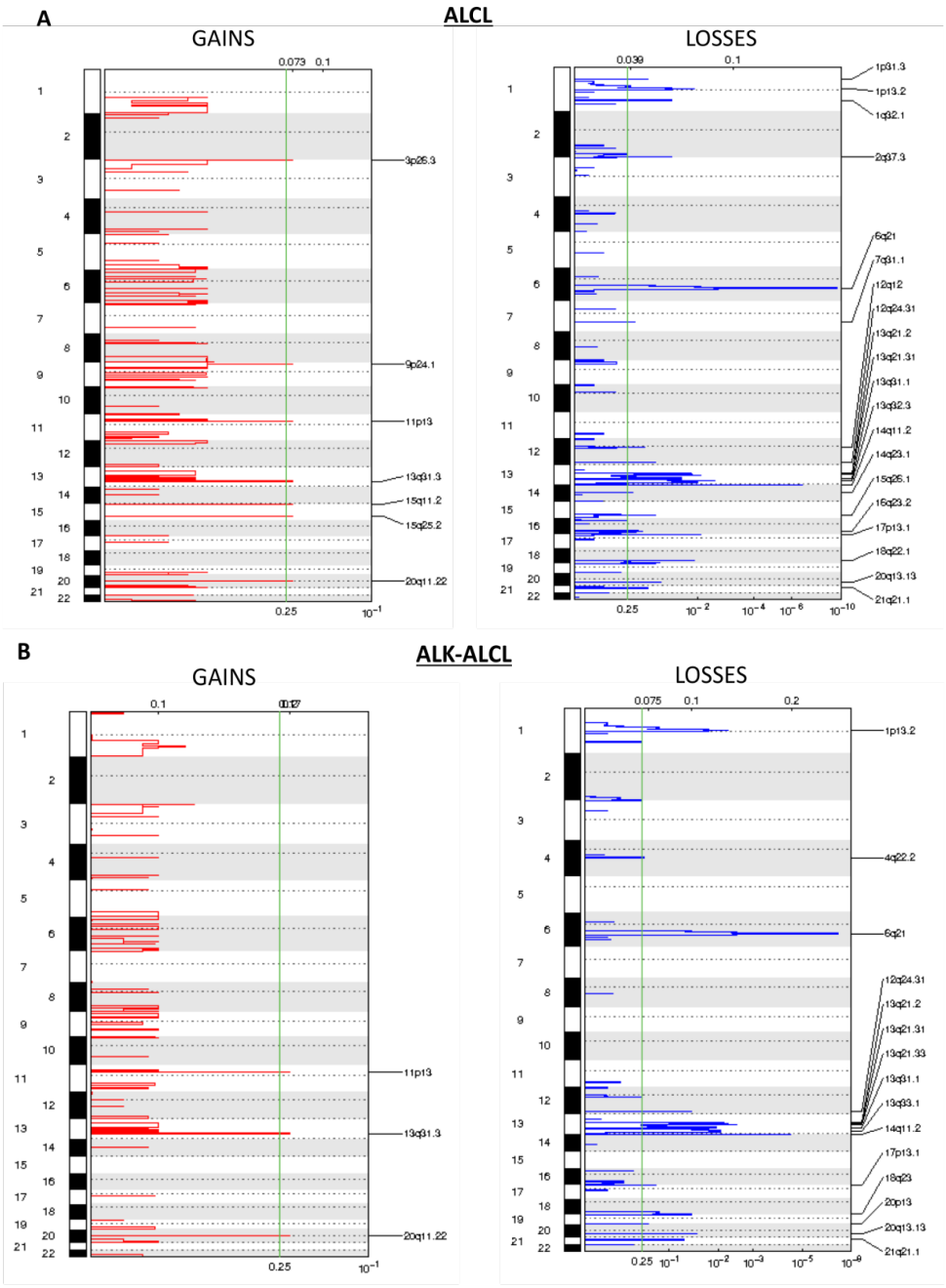


Figure 26. Focal aberrations in ALCL and ALK- ALCL samples. *Analysis via GISTIC of copy number gains (left) and losses (right) for the 64 ALCL samples (A) and only for the 31 ALK- samples (B). False-discovery rate q values are plotted along the X-axis with chromosomal position along the Y-axis. Altered regions with significance levels exceeding the vertical green line (significance threshold) were considered significant.*

The most significant aberrations were loss at 17p13.1 (*TP53*) and loss at 6q21 (*PRDM1* and *ATG5*), both among all ALCL samples and within the ALK- subset. Among the most significant losses, GISTIC highlighted also loss at 14q11.2 as expected, region in which genes involved in TCR rearrangements are located. As already reported, this region has been discarded for subsequently analysis because of the physiologically rearrangement occurring at this locus, including interstitial losses, for the generation of the TCR. Conversely, regions of gains identified by GISTIC are different among the two subsets: the most significant gain in ALCL is at 9p24.1 region (*JAK2*), while gains at 20q11.2 (*ASIP, AHYC*, *ITCH, MIR644*), at 11p13 (*APIP, PDHX*, *CD44*) and gain at 13q31.3 (*MIR17HG*) are identified in both ALCL and ALK-, being more significant in ALK- ALCL subtype.

Table 7. GISTIC in ALCL. *Significant regions identified with GISTIC algorithm are reported.*

Lesions	Cytoband	Frequency, %	Start	Size_Kb	q-value	Res q-value	Candidate gene
Gains							
	9p24.1	10%	5.51E+06	2.64E+06	0.23624	0.23624	*JAK2*
	3p26.3	7%	1.67E+06	1.14E+05	0.23624	0.23624	
	15q11.2	7%	2.53E+07	3.74E+04	0.23624	0.23624	*IPW*
	20q11.22	6%	3.28E+07	2.69E+05	0.23624	0.23624	*ASIP, AHCY, ITCH, MIR644*
	11p13	4%	3.48E+07	6.87E+05	0.23624	0.23624	*APIP, PDHX, CD44*
	13q31.3	3%	8.90E+07	4.22E+06	0.23624	0.23624	*MIR17HG*
	15q25.2	3%	8.45E+07	2.85E+04	0.23624	0.23624	*ADAMTSL3*
Losses							
	17p13.1	24%	6.09E+06	1.62E+06	0.081305	0.0099598	*TP53*
	6q21	17%	9.48E+07	2.12E+07	4.68E-11	2.81E-10	*ATG5, PRDM1*
	13q32.3	16%	9.32E+07	1.52E+07	0.0025428	0.011625	
	14q11.2	16%	2.22E+07	1.29E+06	6.57E-08	1.88E-07	*TCRA, TRAC*
	16q23.2	16%	7.88E+07	4.46E+03	1	0.1503	*WWOX*
	12q24.31	14%	1.22E+08	5.68E+05	0.052865	0.08445	
	13q21.31	14%	6.24E+07	3.97E+06	0.0092452	0.040311	
	13q31.1	14%	8.24E+07	6.66E+06	0.052865	0.098264	*SLITRK1, SLITRK6*
	1p13.2	11%	1.08E+08	1.09E+07	0.026836	0.014931	
	13q21.2	11%	6.05E+07	2.02E+05	1	0.18826	*DIAPH3*
	12q12	10%	4.44E+07	3.90E+05	0.57913	0.11838	*TMEM117*
	15q26.1	10%	8.92E+07	4.58E+04	0.052865	0.08445	*AEN, ISG20*
	2q37.3	9%	2.40E+08	1.50E+05	0.026836	0.041884	*HDAC4, MIR4269*
	18q22.1	7%	6.51E+07	3.95E+05	0.0067821	0.013671	*DSEL*
	20q13.13	7%	4.89E+07	7.76E+04	0.048594	0.069286	
	7q31.1	6%	1.11E+08	8.13E+04	0.14064	0.18826	*IMMP2L*
	1q32.1	4%	1.99E+08	3.31E+05	0.026836	0.041884	*PTPRC*
	14q23.1	4%	6.19E+07	1.41E+03	1	0.18826	*PRKCH*
	1p31.3	3%	6.67E+07	9.79E+04	0.081305	0.1503	*PDE4B*
	21q21.1	3%	2.26E+07	3.63E+03	1	0.11838	*NCAM2*

Table 8. GISTIC in ALK- ALCL. *Significant regions identified by GISTIC algorithm are reported.*

Lesions	Cytoband	Frequency, %	Start	Size_Kb	q-value	Res q-value	Candidate gene
Gains							
	20q11.22	12%	3.28E+07	2.69E+05	0.22719	0.22719	*ASIP, AHCY, ITCH, MIR644*
	11p13	6%	3.48E+07	6.87E+05	0.22719	0.22719	*APIP, PDHX, CD44*
	13q31.3	6%	8.90E+07	4.22E+06	0.22719	0.22719	*MIR17HG*
Losses							
	17p13.1	42%	7.47E+06	1.06E+04	0.99208	0.1585	*TP53*
	14q11.2	39%	2.22E+07	1.29E+06	6.74E-06	3.35E-05	*TCRA, TRAC*
	6q21	33%	9.48E+07	1.86E+07	1.43E-08	8.94E-08	*ATG5, PRDM1*
	12q24.31	27%	1.22E+08	5.68E+05	0.021862	0.039734	
	1p13.2	24%	9.85E+07	2.19E+07	0.0095216	0.0063173	
	13q21.31	24%	6.24E+07	3.97E+06	0.003499	0.066829	
	13q21.33	24%	6.64E+07	7.05E+06	0.0010479	0.25555	*PCDH9, KHLH1*
	13q33.1	24%	9.32E+07	1.52E+07	0.003499	0.047195	*SLITRK1, SLITRK6*
	13q21.2	21%	6.05E+07	2.02E+05	0.99208	0.14511	*DIAPH3*
	13q31.1	21%	8.24E+07	6.59E+06	0.021862	0.091229	
	20p13	18%	7.35E+04	1.93E+03	0.1217	0.21329	
	20q13.13	18%	4.89E+07	4.00E+08	0.019672	0.039734	
	4q22.2	15%	9.40E+07	1.39E+04	1	0.21329	*GRID2*
	18q23	15%	6.51E+07	7.06E+07	0.021862	0.039734	
	21q21.1	6%	2.26E+07	3.63E+03	1	0.047195	*NCAM2*

I then tried to combine the genomic data in order to understand if different subgroups of ALCL are distinguishable on the basis of their recurrent chromosomal aberrations. Figure 6 shows the heatmap of the ALCL cases sorted based on the most common lesions observed in our series. Among the ALK- cases two major groups of patients can be identified: a first one characterized by the presence of numerous genomic lesions, including 17p and 6q21 losses, detectable in a large fraction of samples; a second group in which very few and heterogeneous lesions were detectable. Among ALK+ ALCL samples two different groups were observed as well. In the majority of samples, no genomic lesions were identified at all, suggesting that the translocation involving ALK and the downstream activation of pro-survival pathways seem to be sufficient for the pathogenesis of lymphoma. On the contrary, other ALK+ cases were affected by very few lesions.

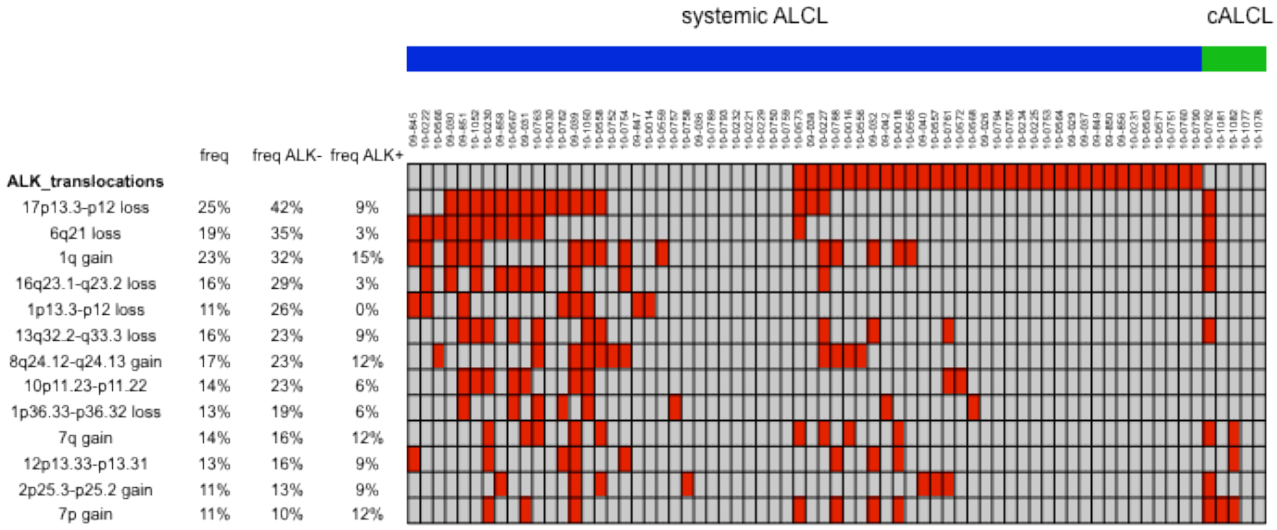


Figure 27. Heatmap with the relative distribution of the most common lesions detected in ALCL samples*. Each column is representative of one patient (systemic ALCL or cALCL); each line represent a common lesion identified trough the array-CGH method. The frequencies of each lesion in the whole cohort, in the ALK- and ALK+ subtypes are reported. The identification of the lesion concern in the patient examined is reported in red; the absence in grey.*

2. *PRDM1*/BLIMP1 gene is often inactivated in ALCL

Both MCRs and GISTIC analysis pointed to 6q21 loss as one of the most significantly recurrent lesion in ALCL, even if this lesion is more frequent in the ALK- samples. The smallest region involved in this aberration was 1,120 Kb of length and encompassed only *PRDM1* and *ATG5* genes. As reported in Figure 28, losses at 6q21 cytoband were observed in 11/31 (35%) ALK- samples, in 4/5 (80%) ALK+ cell lines, and in 1/33 (3%) ALK+ cases. The observed lesion in ALK- and ALK+ was statistically different (p-value <0.0001). Importantly, the only ALK+ case harbored a homozygous deletion of the *PRDM1* gene. Although most of the losses were relatively long, it was evident that the smallest ones and, especially, the homozygous deletion encompassed only *PRDM1* locus.

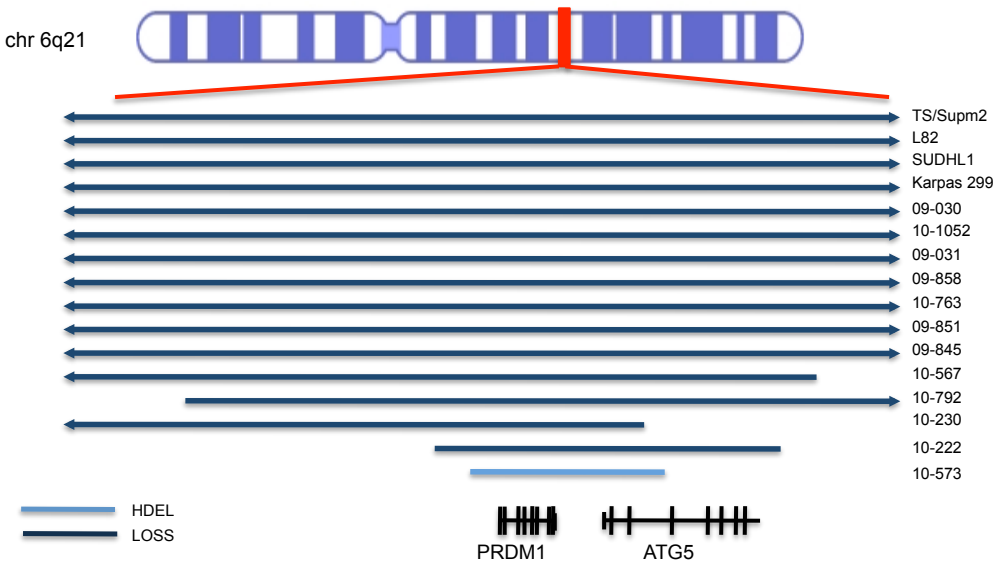


Figure 28. Lesions affecting PRDM1 locus. *Representation of the aberration, and their size, affecting PRDM1 locus (6q21); blue bars represent loss; light blue bar represents hdel.*

PRDM1, coding for BLIMP1 protein, is recurrently inactivated in Diffuse Large B-Cell Lymphomas (DLBCL) [49][50][217] and, as recently reported, also in NK-cell lymphomas [92][91]. In both cases, *PRDM1* deletion/inactivation seems to be important for the pathogenesis of lymphomas and it is considered a tumor suppressor gene. So, our arrayCGH data suggested that the loss of the gene could also be relevant for ALCL. Thus, I performed a mutational analysis of all the coding exons of *PRDM1* gene, in order to reveal if this gene could be inactivated also by somatic mutations. This analysis was performed for all ALK- cases, for all the ALCL cell lines and for 15/33 ALK+ samples. Two mutations were detected in two different ALK- patients, in which the nucleotide change leads to the formation of a stop codon. One of the patients showed a loss of the wild type allele (09-845), while the other patient (10-558) showed a copy neutral loss of heterozygosity (cnLOH), possibly indicating the formation of a truncated protein in both cases. A silent mutation was also observed in exon 4 of SUDHL1 cell line (Fig 29). A mutation was seen also in another ALK- patient, but it was present also in the matched normal DNA, suggesting it represented a new SNP, not already reported in the databases. In order to have a global overview of *PRDM*1 gene, I also analyzed the methylation profile of its promoter in 40/69 of our ALCL samples, but no methylation was identified. So, in our series *PRDM1* was inactivated by deletion, CN-LOH or somatic mutation.

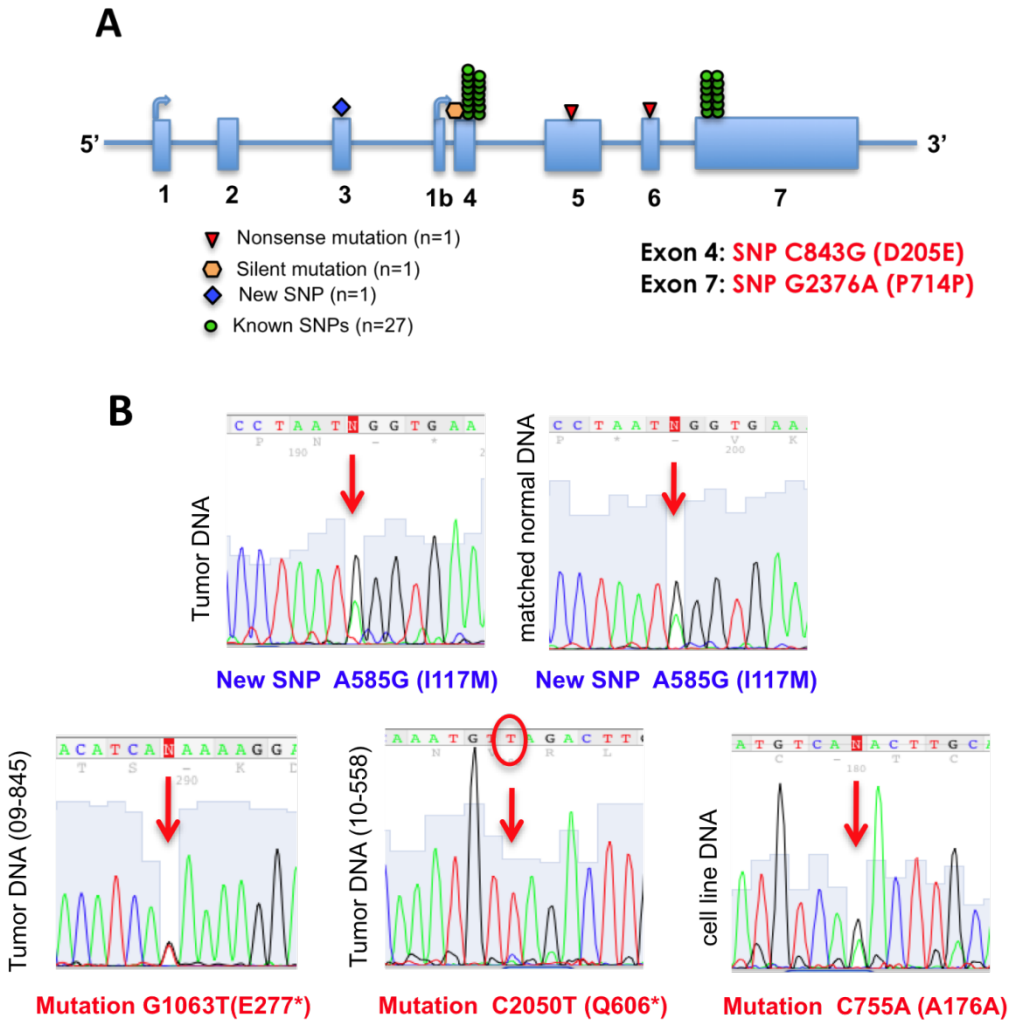


Figure 29. Sequencing and mutational analysis of PRDM1 coding exons. *(A) Results of sequencing and mutational analysis of PRDM1 coding exons in 46 ALCL samples (31 ALK- and 15 ALK+ samples) and in 7 ALCL cell lines (5 ALK+ and 2 ALK- cell lines). Green circles represent known SNPs, red triangles are nonsense mutations in two ALK- samples, orange exhagon is a mutation in SUDHL1 cell line (ALK+ ALCL) and blue rhombus is a new SNP (or germline mutation) observed in another ALK- sample. Arrows represent transcriptional start site for the long (exon 1) and the short (exon 1β) isoform. (B) View of the single nucleotide changes in different sequences. The new SNP has been observed in the tumoral and in the matched normal DNA (upper panel).*

Since *PRDM1* loss was frequently concomitant to *TP53* loss, located in the other genomic region most commonly lost, at 17p13.1 cytoband level, I investigated if *TP53* gene could be inactivated as well also by somatic mutations. I analyzed the mutational status of coding exons from 4 to 8, which are the most common affected by mutations in other lymphomas [224, 225]. I analyzed 11 ALK- cases, five harbouring *PRDM1* loss and *TP53* wild type and six other cases with both genes wild type. No mutations were identified in these cases. I performed the same analysis also for the cell lines and we confirmed the presence of mutations in Karpas299 (exon 8), SUDHL1 (exon 8) and L82 (exon 7) as already reported in literature [226]. Since in ALCL samples I couldn't detect any mutation, in ALCL *TP53* might not represent the second hit in the pathogenesis.
In summary, *PRDM*1 was inactivated, by deletion or by somatic mutations, in 12/31 (39%) and in 1/33 (3%) of clinical specimens derived from ALK- and ALK+ ALCL (p value <0.0001), respectively, in 4/7 (57%) ALCL cell lines, and in 1/5 (20%) cALK- ALCL (Fig 30).

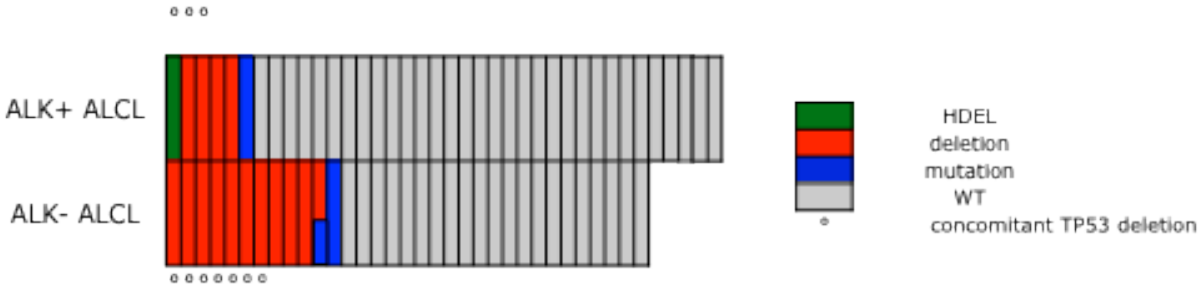


Figure 30. Summary of genomic lesions and mutations at PRDM1 level observed in the whole series of ALCL samples and cell lines*. Each bar is representative of one specimen. ALK+ and ALK- samples are reported separately. Green represents homozygous deletion, red represents deletion, blue represents mutation and grey represent a wild type status. Little circles identify samples in which concomitant TP53 deletion is present.*

In order to validate *PRDM1* expression in ALCL samples, I performed quantitative Real-Time PCR in a subset of 27 ALCL samples (14 ALK- ALCL and 13 ALK+ ALCL), already analyzed with the Affymetrix SNP 6.0 platform. For these samples we had collected the RNA material as well. As shown in Fig 31, 3/14 ALK- samples (10-222, 10-230, 10-761), carrying 6q21 loss, appeared to express moderately low levels of *PRDM1* transcript

(indicate by the black star). Interestingly, the ALK+ case showing an hdel of *PRDM1* (red star) showed undetectable mRNA levels of the gene. No differences in term of *PRDM1* expression levels can be observed between the two ALCL subsets.

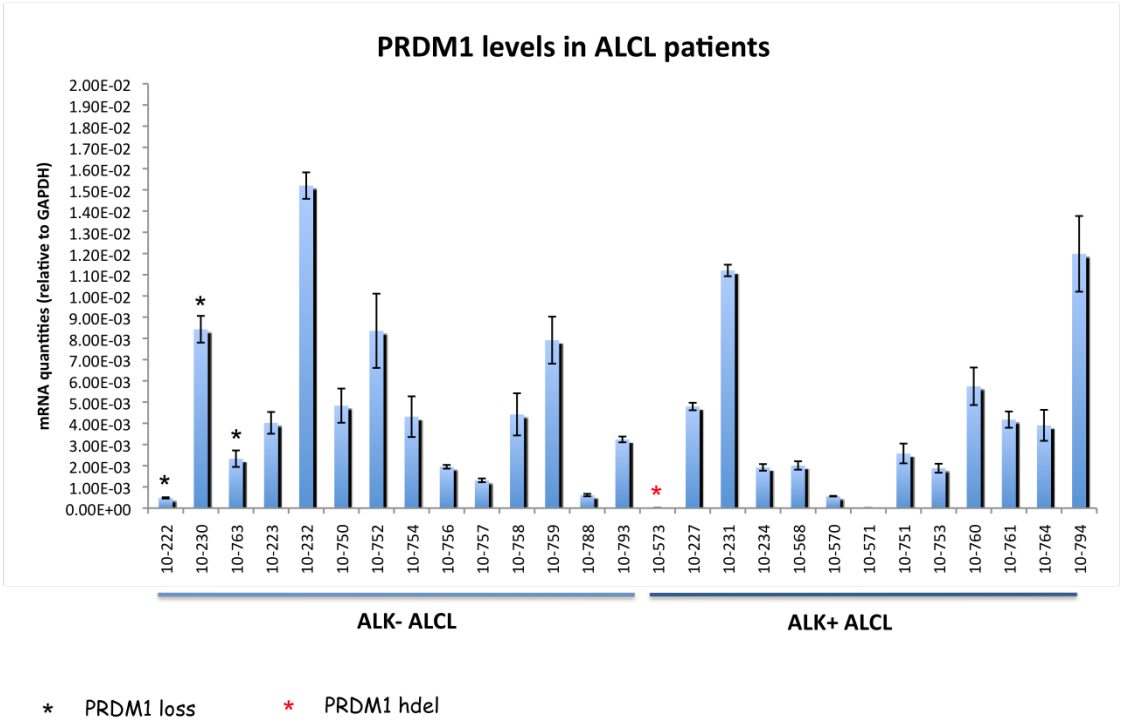


Figure 31. Validation of the expression of PRDM1 in a subset of 27 ALCL patients. *On the X-axis ALCL samples analyzed are listed; on the Y-axis the mRNA quantities relative to GAPDH housekeeping gene are reported. Black stars indicate samples bearing 6q21 loss, red star indicates the sample with hdel of PRDM1 gene.*

Expression of BLIMP1 protein was also evaluated by immunohistochemisty on an independent series of 38 clinical specimens of ALCL. No big difference was identified among the ALCL subtypes: in fact the protein was expressed at nuclear level in 61,9% of ALK+ and 47% of ALK- samples (Fig 32). Moreover, in two ALK- samples, the positivity for BLIMP1 was detected only at cytoplasmic level, and it was surprising, being PRDM1 a transcription factor and exerting its role in the nucleus. Interestingly, one of the two cases showing the positivity for BLIMP1 staining only at a cytoplasmic level (Fig 32, panel B) is

the same case (09-845) presenting a mutation, which lead to the formation of the stop codon, in BLIMP1 coding sequence accompanied by the deletion of the wild type allele. These observations suggest probably the presence of a truncated protein, due to the presence of a new stop codon that might lack the nuclear localization sequence, and the protein remain sequestered in the cytoplasm.

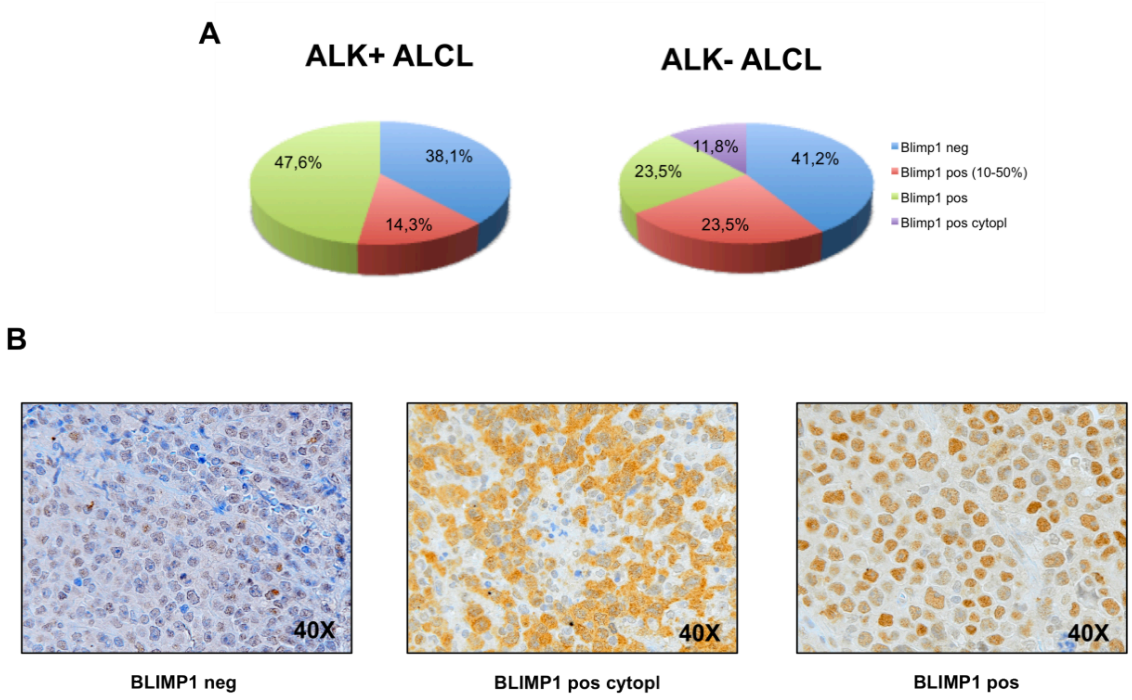


Figure 32. Immunohistochemistry results for BLIMP1 staining. *(A) Graphs reporting the percentage of ALK+ and ALK- cases, positive or negative for BLIMP1 staining in immunohistochemistry. BLIMP1 staining was evaluated on TMA obtained from an independent series of ALCL samples. TMA samples were considered negative or positive for BLIMP1 depending on the intensity of the staining. (B) A representative case for each condition, BLIMP1 negative (left panel), BLIMP1 positive only at cytoplasmic level (middle panel), BLIMP1 positive only at a nuclear level (right panel), is reported (Magnification 400X).*

3. PRDM1 is a tumor suppressor gene in an *in vitro* model of ALCL

To understand the role of *PRDM1*/BLIMP1 loss in ALCL, I studied the consequences of the re-expression of the gene in ALCL cell lines.

First of all, *PRDM1*/BLIMP1 mRNA and protein basal levels in our seven ALCL cell lines (Supm2 and TS cell lines, because one is a subclone of the other, are considered as a single cell line) was evaluated (Fig 33), five deriving from ALK+ ALCL tumors and two deriving from cALCL, commonly used as model for ALK-ALCL, since real ALK- ALCL cell lines do not exist.

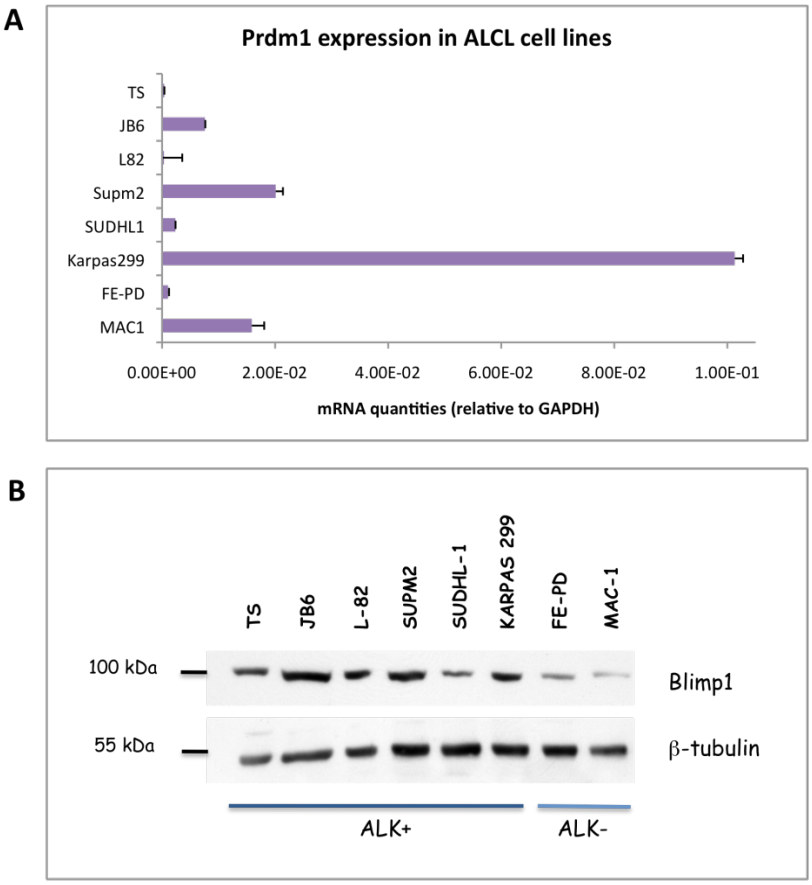


***Figure 33. PRDM1/BLIMP1 mRNA (A) and protein (B) levels in ALCL cell lines**. FE-PD and MAC1 are cALCL cell lines, TS, JB6, L82, SUPM2, SUDHL1, and Karpas299 are ALK+ ALCL.*

The presence of the same 6q21 loss observed in clinical specimens was seen in TS, ALK+ ALCL cell line as well. Considering also that PRDM1/BLIMP1 both mRNA and protein levels were very low in this cell line, I decided to use TS for subsequent functional experiments.

The first approach I used for BLIMP1 re-expression was a transfection with lipofectamine. I tried to introduce an expression vector for my gene of interest (pCMV-HA-blimp1) into TS cells, and I used an empty vector (pCMV-HA) as control (vectors are kindly provided by Laura Pasqualucci). Both expression vectors express a GFP tag, which allows the identification, and quantification of transduced cells through flow cytometry analysis. Different conditions for this transfection have been tried: different number of cells transfected for each conditions or leaving transfected cells in Optimem medium until the recovery of cells at different time point for RNA or protein extraction, or changing medium and replace it with RPMI with 10% FBS after 4h of transfection. Unfortunately, I could not observe any GFP positive cells, at fluorescence microscope and by FACS analysis, in none of the different conditions tried.

Then, I used electroporation for the re-expression of BLIMP1. Also in this case I tried different buffers and different programs in the Amaxa Nucleofactor machine as the manufacturer suggested, but I obtained again very low levels of GFP positive cells as a read out of the occurred transfection and, concomitantly, I observed high levels of dead cells. So, I thought that also this method for the transfection was not optimal to introduce BLIMP1.

Afterwards I tried to perform this gain of function experiment with a different technique. In order to obtain a stable and efficient transfection a lentiviral infection method using the pWPI-HA-BLIMP1 lentivirus was used. Also in this case, both empty vector (pWPI) and vector for BLIMP1 re-expression (pWPI-HA-BLIMP1) (lentiviral vectors were kindly provided by Laura Pasqualucci) express a GFP tag which allow the identification and quantification of transduced cells. After having confirmed the efficiency of transfection calculating the percentage of GFP positive cells at flow-cytometry (range 28-37% for BLIMP1 lentiviral vectors and range 62-71% for empty vector), RNA and total protein lysates from transduced cells were obtained and a good re-expression of BLIMP1 at both mRNA and protein level was observed (Fig 34, panel A). Interestingly, an increase of percentage of GFP positive cells infected with empty vector over time, with a concomitant parallele decrease of GFP positive cells after BLIMP1 re-expression was observed (Fig 34, panel B). These data suggested that re-expression of *PRDM1* expression in an ALK+

cell line, bearing *PRDM1* loss, exerted a negative selective pressure. Similar experiments in NK cells reported comparable results [92]. Since not in all experiments I observed a high efficiency transduction, I performed a pilot experiments in which all available ALCL cell lines have been infected with lentivirus, but I obtained highest efficiency with TS cell lines. For this reason, I performed subsequent functional experiments only with TS ALK+ ALCL cell line.

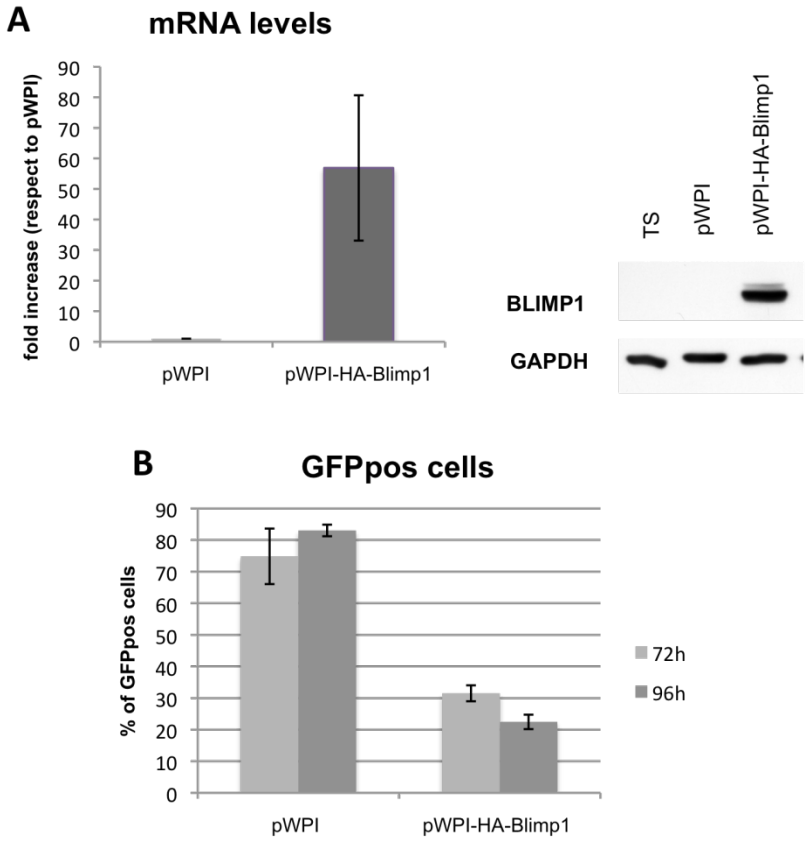


Figure 34. BLIMP1 reconstitution in an ALK+ cell line. *(A) PRDM1/BLIMP1 mRNA (average of three different experiments) and protein levels (one representative experiment) after infection with empty vector (pWPI) or vector for BLIMP1 re-expression (pWPI-HA-Blimp1). (B) Percentage of GFP positive cells after infection at 72h and 96h (average of three different experiments).*

I wondered if the re-expression of *PRDM1* gene could have an effect in terms of proliferation and apoptosis. In order to evaluate the effect of BLIMP1 re-expression in term of proliferation, the rate of cell proliferation over time was evaluated. Differently from

TS cells transduced with the empty vector, after the reconstitution of *PRDM1*, TS cells underwent proliferation arrest, which was more marked starting from the third day after infection (Fig 35).

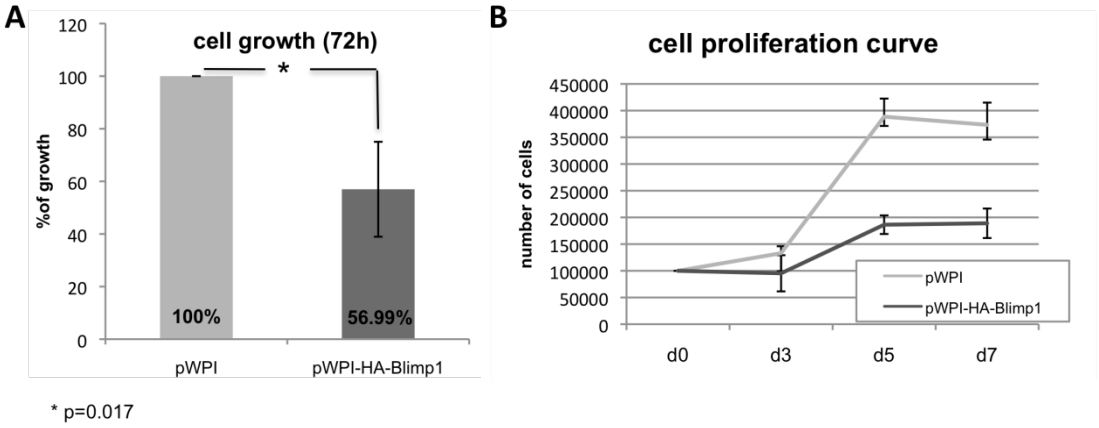


Figure 35. BLIMP1 re-expression leads to cell proliferation arrest. *(A) Percentage of growing cells after infection at 72h (average of three different experiments). The percentage of cell growing after BLIMP1 re-expression is expressed in comparison to cells infected with empty vector (pWPI). Difference in percentage of cells growth between the two cases was evaluated by applying Fisher's exact test (p-value). (B) Growth curve after infection, cells were counted at d3, d5, d7 after infection (one representative experiment).*

Concomitant to the proliferation arrest, I also observed a moderate increase in the percentage of apoptotic cells after BLIMP1 reconstitution (22.9% vs 29.7% at 72 h and 11.6% vs 17.1% at 96 h for control cells and BLIMP1 transduced cells respectively) and an arrest in cell cycle with an increase of the subG1 phase (8.06% vs 15.1% for control cells and BLIMP1 transduced cells respectively), confirming an increase of apoptotic cells. A change in the distribution of the cells in the different cell cycle phases (G1: 48.5% vs 25.1%; S: 19.3% vs 13.3% and G2-M: 20.6% vs 30.5%, after infection with empty vector or with vector for BLIMP1 reconstitution respectively) was observed as well (Fig 36). Thus, in accordance with the frequent inactivation observed in clinical specimens suggest, *in vitro* data suggested that *PRDM1* acts as a tumor suppressor gene in ALCL.

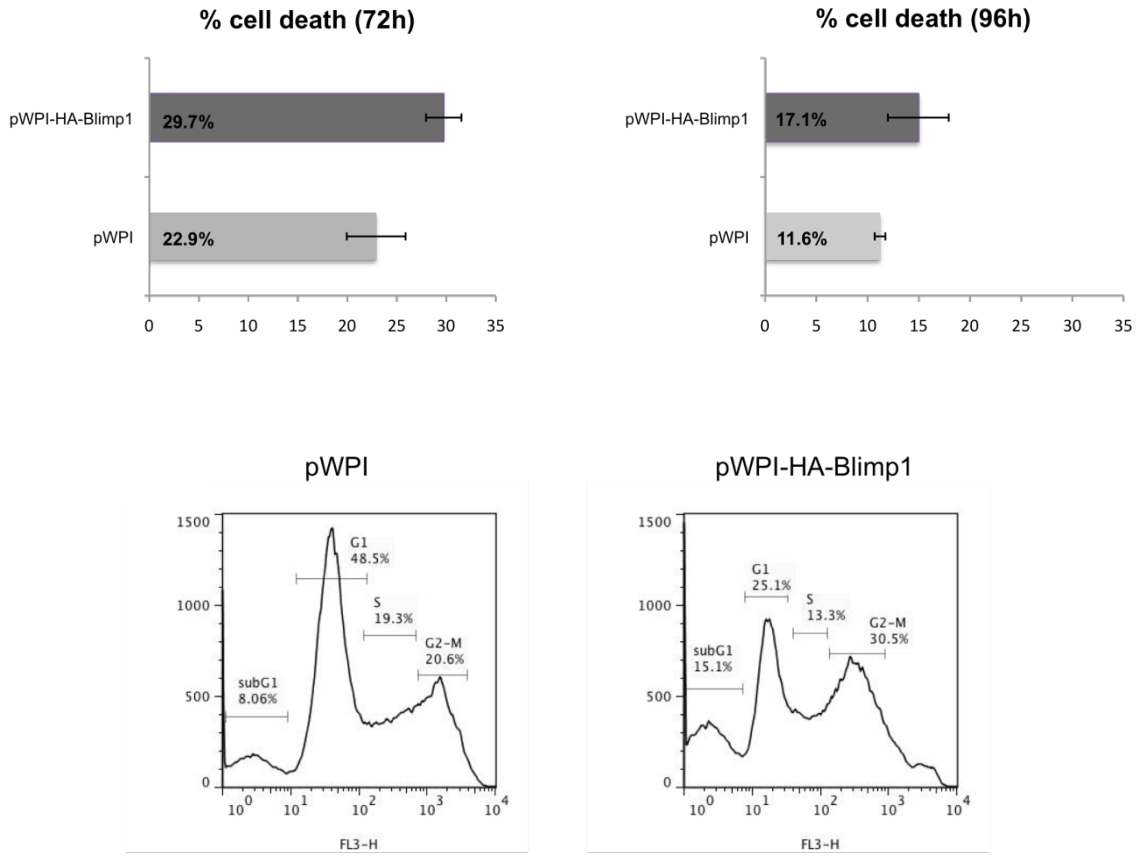


Figure 36. BLIMP1 re-expression leads to an increase of cell death and cell cycle arrest. *(A) Percentage of apoptotic cells after infection at 72h and 96h (average of three different experiments). (B) Cell cycle profile after infection (one representative experiment).*

4. *PRDM1* exerts a pro-apoptotic effect

With the aim to understand the transcriptional role of BLIMP1 after its re-expression in ALCL, gene expression profiling (GEP) analysis was carried out in TS ALCL cells, infected with lentivirus pWPI or pWPI-HA-Blimp1. Supervised analysis identified a selected number of genes specifically modulated after BLIMP1 re-exprexsion (Fig 37).

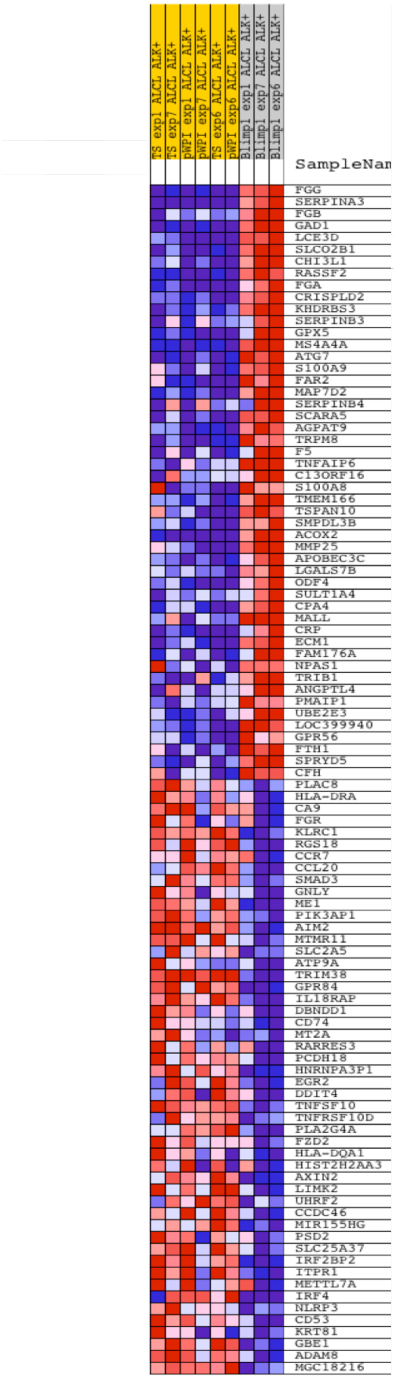


Figure 37. Hierarchical clustering of samples analyzed with GEP. *GEP analysis was performed in untreated TS cells (n=; Yellow), TS cells infected with empty vector (pWPI; n=3; Yellow) and TS cells infected with lentivirus for BLIMP1 reconstitution (pWPI-HA-Blimp1; n=3; Grey).*

Notably, more than 70% of differentially expressed genes were upregulated after BLIMP1 re-expression. The top 25 upregulated genes (Table 5) included transcripts involved in growth and cell proliferation inhibition (*SERPINA3, KHDRBS3*), apoptosis (*S100A9, PMAIP1*) and cell cycle arrest (*RASSF2*). Interestingly, among the most upregulated genes, fibrinogen genes component (*FGG, FGA, FGB*), located in a cluster region of chromosome 4, could be identified. Data on overexpression of *FGG, SERPINA3*, which codes for a growth inhibition protein, and *PMAIP1* gene, which codes for NOXA pro-apoptotic protein, were validated through qRT-PCR (Fig 38). The interestingly observation of the downregulation of mir155 (Fig 37) after BLIMP1 re-expression, and the knowledge that mir155 is a direct repressor of the pro-apoptotic factor SHIP1, lead us to analyze SHIP1 levels after BLIMP1 re-expression. As expected, SHIP1 mRNA levels were upregulated after BLIMP1 reconstitution (Fig 38). In this way we have validated the functional consequence of mir155 dowrwgulation, due to BLIMP1 re-expression, and provided another way through which BLIMP1 expression could exert its pro-apoptotic function. In Table 10, the top 25 downregulated genes after BLIMP1 re-expression were reported. Full lists of the most modulated transcripts and their functions were reported in the Appendix section.

Table 9. Top 25 upregulated genes after BLIMP1 re-expression. *Probe numbers on Illumina Whole-Genome Gene Expression BeadChip and corrisponding genes are reported. Fold-change is evaluated as the ratio between expression of the gene in cells after BLIMP1 reconstitution and the expression of the gene in cells infected with empty vector or not infected.*

Column ID	ILMN_GENE	CHROMOSOME	CYTOBAND	Fold-Change
ILMN_2315044	FGG	4	4q32.1	47.5345
ILMN_1678049	FGB	4	4q32.1	14.2424
ILMN_1788874	SERPINA3	14	14q32.13	9.82235
ILMN_2087656	SLCO2B1	11	11q13.4	9.64378
ILMN_1761199	SLCO2B1	11	11q13.4	9.22832
ILMN_2114972	FGB	4	4q32.1	8.65773
ILMN_2292646	GAD1	2	2q31.1	7.68648
ILMN_1737683	FGG	4	4q32.1	6.87751
ILMN_1718395	LCE3D	1	1q21.3	6.86135
ILMN_1709683	RASSF2	20	20p13	5.18258
ILMN_1703855	SERPINB3	18	18q21.33	4.91957
ILMN_1790978	ATG7	3	3p25.3-p25.2	4.88356
ILMN_1656487	FGA	4	4q32.1	4.25149
ILMN_1691747	KHDRBS3	8	8q24.23	4.17722
ILMN_1790689	CRISPLD2	16	16q24.1	3.69527
ILMN_1685703	ACOX2	3	3p14.3	3.58141
ILMN_2352303	RASSF2	20	20p13	3.40578
ILMN_1782716	SERPINB4	18	18q21.33	3.26909
ILMN_2370336	MS4A4A	11	11q12.2	3.23639
ILMN_2385410	GPX5	6	6p22.1	3.22234
ILMN_2385416	GPX5	6	6p22.1	3.18795
ILMN_2098446	PMAIP1	18	18q21.32	2.95458
ILMN_1751062	SCARA5	8	8p21.1-p21.1	2.75444
ILMN_1702973	TMEM166	2	2p12	2.75308
ILMN_2298159	PRDM1	6	6q21	2.74487

Table 10. Top 25 dowregulated genes after BLIMP1 re-expression. *Probe numbers on Illumina Whole-Genome Gene Expression BeadChip and corrisponding genes are reported. Fold-change is evaluated as the ratio between expression of the gene in cells after BLIMP1 reconstitution and the expression of the gene in cells infected with empty vector or not infected.*

Column ID	ILMN_GENE	CHROMOSOME	CYTOBAND	Fold-Change
ILMN_2101278	RGS18	1	1q31.2	-2.53
ILMN_1681301	AIM2	1	1q23.1-q23.2	-2.28624
ILMN_2394561	IRF2BP2	1	1q42.3	-2.16211
ILMN_2384139	KLRC1	12	12p13.2	-2.04685
ILMN_1766169	BCAT1	12	12p12.1	-2.00037
ILMN_1697971	TRIM38	6	6p22.2	-1.9861
ILMN_1744023	MGC18216	15	15q26.3	-1.7481
ILMN_1676528	BTN3A2	6	6p22.1	-1.65554
ILMN_1701114	GBP1	1	1p22.2	-1.62026
ILMN_1758543	CNIH	14	14q22.2	-1.58214
ILMN_1787511	THUMPD2	2	2p22.1	-1.54408
ILMN_2120624	ZIC4	3	3q24	-1.46575
ILMN_1799026	PCDH18	4	4q28.3	-1.46467
ILMN_1743939	C5ORF24	5	5q31.1	-1.40897
ILMN_2341254	STARD13	13	13q13.1	-1.40207
ILMN_1686553	INTS2	17	17q23.2	-1.32948
ILMN_1664047	CACNA1E	1	1q25.3	-1.3288
ILMN_1744795	TBL1X	X	Xp22.31-p22.2	-1.32597
ILMN_1677532	TARDBP	1	1p36.22	-1.31408
ILMN_1685930	DNAJC24	11	11p13	-1.29993
ILMN_1740493	TRAF5	1	1q32.3	-1.29973
ILMN_2335669	ZC3H14	14	14q31.3	-1.27453
ILMN_2137464	DVL3	3	3q27.1	-1.26506
ILMN_3230337	LOC100132707	7	7q36.2	-1.24259
ILMN_3245116	GOLIM4	3	3q26.2	-1.22227

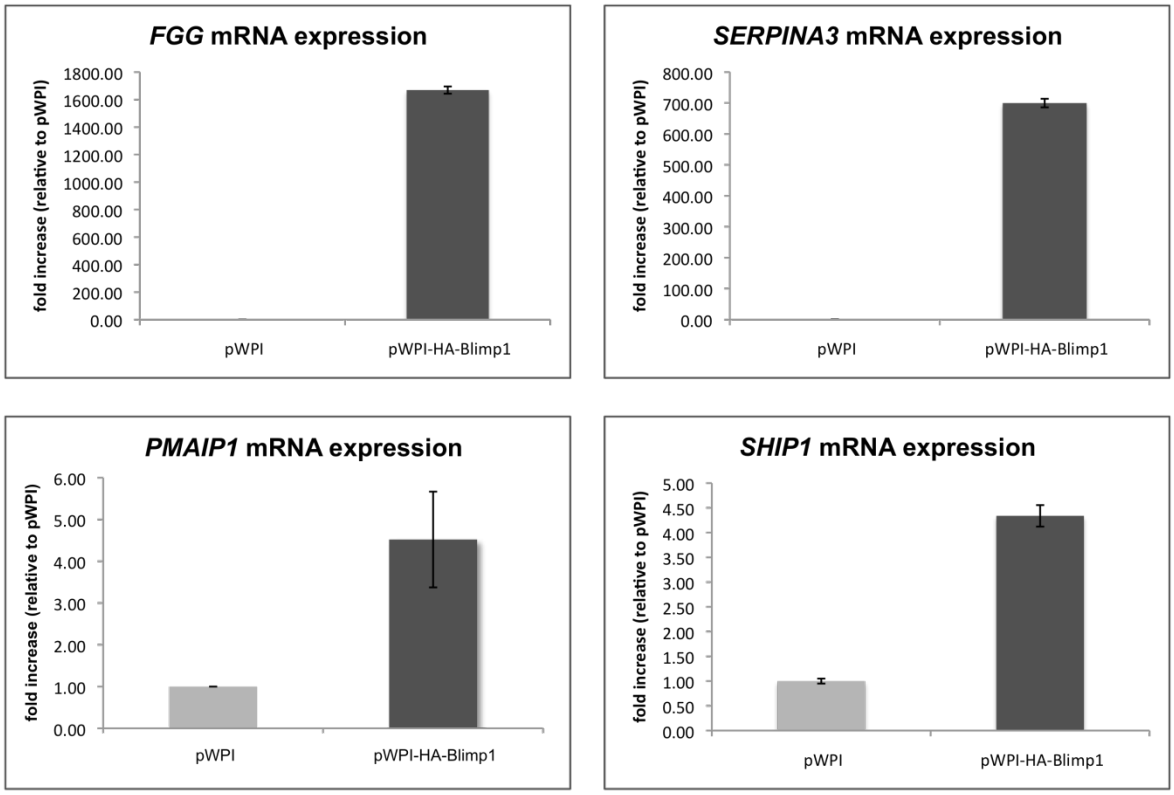


***Figure 38. Validation of GEP results**. FGG, SERPINA3, PMAIP1 and SHIP1 mRNA expression after infection with empty vector (pWPI) or with vector for BLIMP1 expression (pWPI-HA-Blimp1). On Y-axis fold change, relative tocells infected with empty vector (fold increase=1), are reported.*

The subsequent analysis of GEP data with Gene Set Enrichment Analyis (GSEA) or DAVID, revealed that the main pathways affected by BLIMP1 re-expression were those ones involved in prothrombin activation, complement and coagulation cascades and fybrinolysis.

Combining GEP experiment with previous BLIMP1 functional studies, it seems that BLIMP1 might exerts a pro-apoptotic effects if expressed in ALCL cells, acting on a series of genes that are involved in pathways regulating cell proliferation and survival.

5. PRDM1 is a tumor suppressor gene in vivo model of ALCL

In order to better understand the role of BLIMP1 in ALCL pathogenesis, *in vivo* experiments were performed. To this end, thanks to our collaborators at the University of Turin, pWPI-HA-Blimp1 lentivirus for the re-expression of our gene of interest and pWPI empty vector for control were prepared, as already done for *in vitro* experiments. Afterwards Supm2 cell line, a subclone of TS cell line used for *in vitro* functional experiments, was infected, obtaining about 50% of GFP positive cells as read out of the efficiency of transfection. Infected cells were directly injected in NSG mice: three mice were implanted with cells infected with lentivirus for BLIMP1 and three mice were implanted with cells infected with empty vector in both flanks. Tumors were explanted at day 20 after inoculum and analyzed by flow cytometry. Fig 39, panel A shows that the percentage of GFP positive cells in mice injected with cells BLIMP1 positive was dramatically reduced in respect to the percentage of GFP positive cells in control mice, suggesting that these cells had growth impairment *in vivo*. These results were also confirmed at protein level (Fig 39, panel B): levels of GFP tag protein is dramatically reduced in cells infected with pWPI-HA-Blimp1 lentivirus vector. Moreover, as expected, BLIMP1 expression is increased in tumors deriving from cells in which BLIMP1 has been re-expressed respect to the control. However, a basal level of BLIMP1 protein is observable, because the cell line used presents the loss of only one allele of the gene, which is probably sufficient for the protein production. Furthermore, xenograft growth curves of control (pWPI) and BLIMP1 transduced Supm2 cells in NSG mice indicate that latter xenograft grew less than controls, as previously observed in *in vitro* experiments (Fig 40).

As conclusion, BLIMP1 is frequently inactivated in ALCL clinical specimens, and it appeared a tumor suppressor gene in both *in vitro* and *in vivo* models for ALCL since its re-expression causes reduction in cell growth, increased cell death and cell cycle arrest.

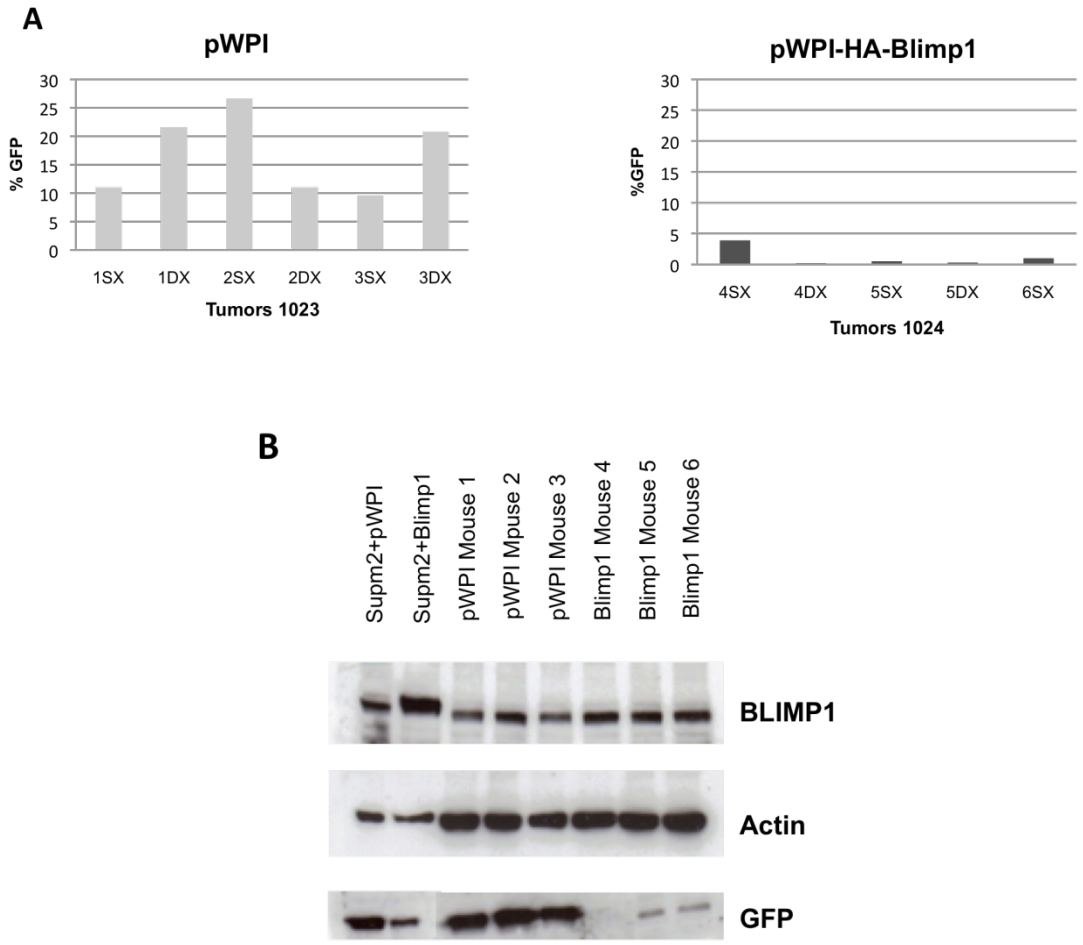


Figure 39. BLIMP1 re-expression in xenograft mouse models. *(A) Percentage of GFP positive cells in tumors explanted from mice inoculated with Supm2 cells infected with empty vector (pWPI) or with vector for BLIMP1 re-expression (pWPI-HA-Blimp1); data obtained from cells injected in both flanks (SX/DX) are reported. (B) Protein levels of BLIMP1 and GFP in tumors explanted from mice at d20.*

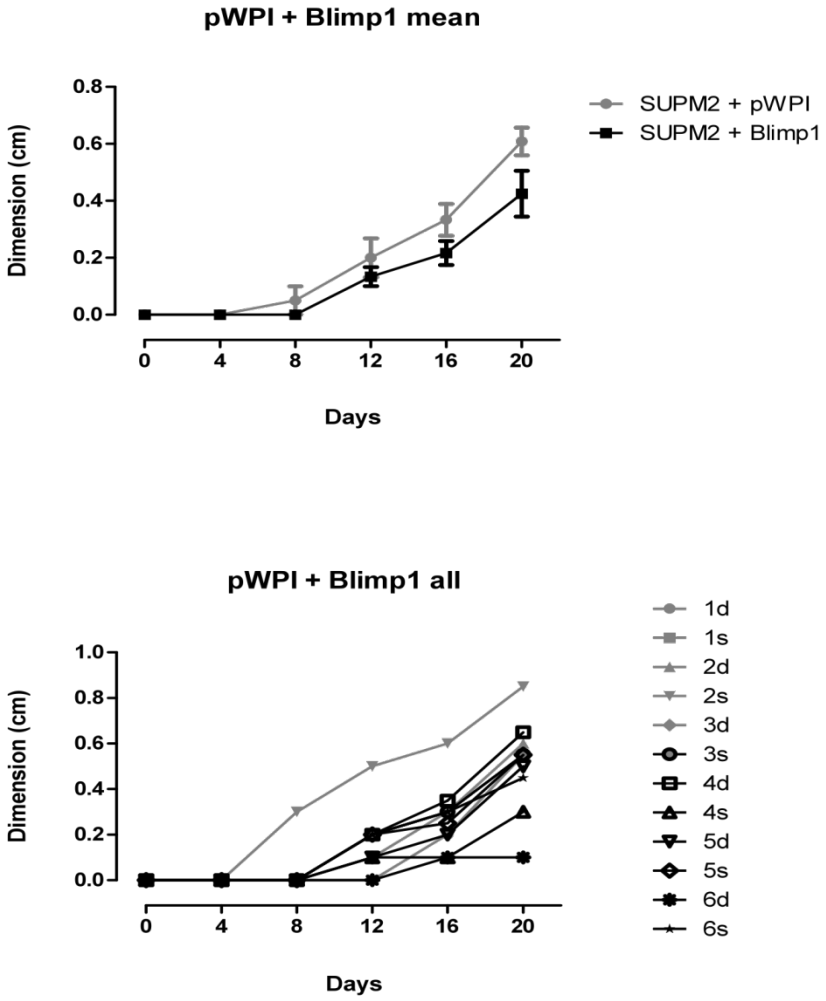


***Figure 40**. **Xenograft tumor growth curves of infected cells injected in NSG mice.** In the upper panel is reported the mean dimension of xenograft tumors, obtained from cells infected with empty vector or with cells injected with BLIMP1 lentiviral vectors, over time. In the lower panel dimensions of each single xenograft tumors over time are reported. Tumors 1, 2, 3 were obtained from cells infected with empty vector, while tumors 4, 5 and 6 were obtained from cells infected with BLIMP1 lentiviral vectors.*

II. MASSIVE RNASEQ DEFINES NOVEL TUMORIGENIC GENE FUSIONS OF ANAPLASTIC LARGE CELL LYMPHOMA

Data reported in the following section have been submitted for publication as: *Marchiolatti R, Messana K, Landra I, Abate F, Tabbo' F, Boi M et al. Massive RNAseq defines novel tumorigenic gene fusions of Anaplastic Large Cell Lymphoma.*

Massive RNAseq defines novel tumorigenic gene fusions of Anaplastic Large Cell Lymphoma

Rodolfo Machiorlatti[1*], Katia Messana[1*], Indira Landra[1*], Francesco Abate[*2,3], Fabrizio Tabbo'[1*], Michela Boi[*4], Maria Todaro[1], Antonella Barreca[1], Nicoletta Chiesa[1], Sabrina Aliberti[1], Filomena Di Giacomo[1], Elena Lasorsa[1], Luca Bessone[1], Marcello Gaudiano[1], Andrea Rinaldi[4], Maurilio Ponzoni[5], Dario Livio Longo[6], Silvio Aime[6], Mangeng Cheng[7], Bruce Ruggeri[7], Graham Tebb[8], Pier Paolo Piccaluga[9], Stefano Pileri[9], Enrico Tiacci[10], Brunangelo Falini[10], Leonard D. Shultz[11], Laurent Farinelli[12], Ivo Kwee[2,13,14], Roberto Piva[1,15], Enzo Medico[16], Raul Rabadan[4+], Francesco Bertoni[2,17+], Giorgio Inghirami[1,15+] and the European T-cell Lymphoma Study Group.

[*] These authors contributed equally

Corresponding authors:

[+]Giorgio Inghirami

[+]Francesco Bertoni

[+]Raul Rabadan

Running Title: novel tumorigenic gene fusions of ALCL

Keywords: anaplastic large cell lymphoma (ALCL), peripheral T-cell lymphoma (PTCL), mouse models, personalized medicine, targeted therapy, phenotype, immunohistochemistry, molecular biology, anaplastic lymphoma kinase (ALK), chimeric fusion proteins, signaling pathways.

ABSTRACT

T-cell lymphoma are heterogeneous neoplasms with a poor prognosis. The limited understanding of the underlying molecular mechanisms, coupled with a lack of animal models, has hampered the development of effective therapies. Here, we demonstrate the constitutive activation of STAT3, NF-κB, RAS and MYC pathways in distinct subsets of Anaplastic Large Cell Lymphoma patients. Searching for driving mutations, we applied a whole-transcriptome paired-end sequencing and identified novel NF1-MAZ and TRAF1-ALK gene fusion transcripts, leading to the constitutive activation of RAS or NF-κB and ALK. Growth inhibition by selective inhibitors in Patient Specific Tumorgrafts (PDT) verified their tumorigenic function. Our work shows that next generation sequencing coupled to PDT represent a powerful strategy to discover tumorigenic defects and to test tailored therapies.

RESULTS and DISCUSSION

(The Supplementary Figures and tables are reported in Section: Appendix II)

The current model of cell transformation predicts that an accumulation of genetic defects, "driver lesions", leads to the deregulated activation of a limited number of signalling pathways [1,2]. As ALK+ and ALK- ALCLs share molecular features [3] and ALK+ ALCLs are known to fire well defined pathways [4,5], we investigated whether specific signatures could cluster different entities within PTCLs and in particular within ALK- ALCLs [3], leading to the recognition of distinct subsets of ALK- ALCLs (Fig. 1A). The findings of a STAT3 signature were confirmed in an independent set by immunohistochemistry (IHC) (17 of 39 ALK- ALCLs were pSTAT3 +, Fig. 1B) and by testing the expression of STAT3-responsive genes in ALK+ ALCL cell lines and in primary ALK-/STAT3 + o – ALCL (Fig 1C, D) [3].

Previous experiments have implicated a number of other signalling pathways, such as NF-κB [6], c-Myc [7] and Ras [8], in distinct subsets of ALCLs and bioinformatics analysis confirmed that the pathways are constitutively activated in samples from the ALCLs (Fig 1D-F). In addition, several samples co-expressed multiple pathways (Fig. 1G), suggesting either that some driver lesions can simultaneously activate multiple pathways or that multiple activating mutations are present, each with specific effects on signalling.

To identify the lesions that lead to activation of these pathways *in vivo*, we took advantage of a newly established library of "Patient Derived Tumorgrafts" PDTs [9,10]. As

there are still no ALK- ALCL cell lines available, PDTs are highly innovative and currently represent the best models for this disease. We obtained evidence that ALCL-PDTs recapitulate the morphological, immunophenotypic and molecular features of their corresponding primary counterparts (Fig. S1A); in particular, they display expression profiles that closely mimic those of primary ALCLs (Figure S1A-B, S2); genome-wide DNA profiling shows a high concordance between ALCLs and PDTs (Fig. S3A and Table S1) and the ALCL-PDTs correlate well with the DNA and RNA from each patient (Fig. S3B).

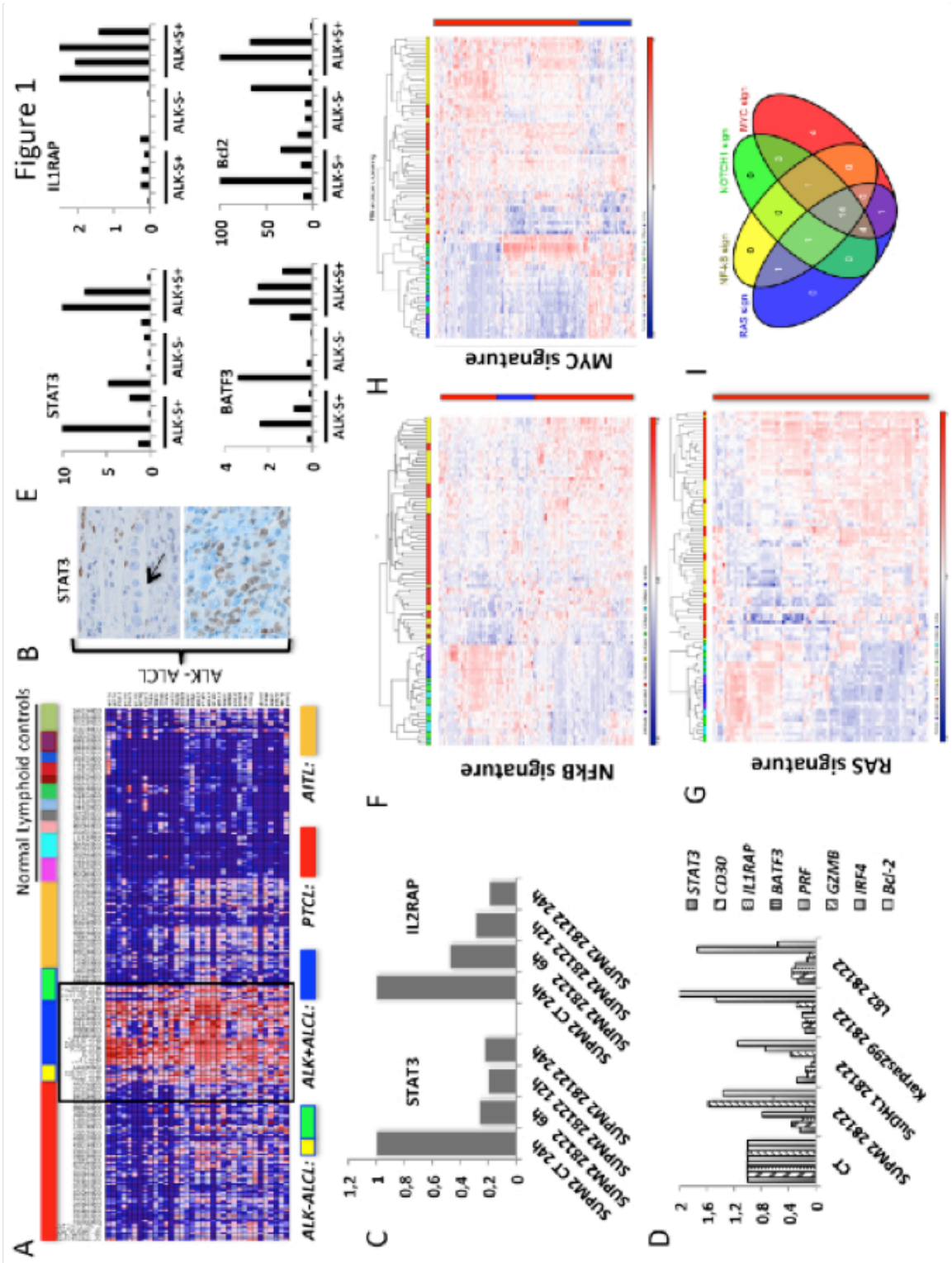


Figure 1
A
Normal lymphoid controls
ALK-ALCL:
ALK+ALCL:
PTCL:
AITL:
B
ALK- ALCL
STAT3
E
STAT3
BATF3
IL1RAP
Bcl2
ALK-S+
ALK-S-
ALK+S+
C
IL2RAP
STAT3
SUPM2 CT 2sh
SUPM2 28122 6h
SUPM2 28122 12h
SUPM2 28122 24h
F
NFkB signature
H
MYC signature
G
RAS signature
I
D
STAT3
CD30
IL1RAP
BATF3
PRF
GZMB
IRF4
Bcl-2
CT
SUPM2 28122
SUDHL1 28122
karpas299 28122
L82 28122

Fig. 1. ALCL and ALCL-PDT display constitutive activation of multiple signalling pathways.

(A) Hierarchical clustering of PTCL samples according to the expression of 34 probe sets (specific for 24 genes) [3]. Overexpressed STAT3 genes are seen in ALK+ and in a subset of ALK- ALCLs. Individual T-PTCL subsets are indicated with different colors. Mixed colors correspond to normal control cells (i.e. resting and activated CD4, CD8 etc.) [3,28]. (B) IHC stains performed with a specific pSTAT3 antibody show a strong nuclear reactivity in a subset of ALK- ALCLs (lower panel). Rare positive stromal elements are detected, representing intra-tissue positive control elements (upper panel). The arrow indicates the nuclei of pSTAT3 negative lymphoma cells. (C) Panel shows the normalized levels of RNA expression of STAT3 and representative STAT3-responsive genes in untreated or (anti-ALK) CEP28122 treated ALK+ ALCL cell lines at different intervals (200nM). Expression levels were determined using a qRT-PCR approach as previously described [3,28]. Data are depicted as 2^ddCt. (D) STAT3-response genes of ALK+ ALCL are efficiently down-regulated by anti-ALK CEP28122 treatment. Individual gene expression levels (dCt) of treated cells (24hr 200nM) were calculated in reference to those derived from untreated cells as 2^ddct. (E) STAT3-response genes are preferentially express in STAT3+ ALCL. The expression of representative STAT3-response (i.e. STAT3, IL1Rap and BATF3) and control genes (Bcl-2) were evaluated in a set of ALK-pSTAT3+, ALK-pSTAT3- and ALK+pSTAT3+ primary ALCL. Expression values were subtracted to the corresponding metagene values, generated using ALK+ ALCL cell lines (Supm2, SUDHL1, Karpas 299 and L82) and normalized for a CD30 expression (dCt target gene/dCt CD30). (F) Hierarchical clustering of PTCL-NOS, AILT, normal T cells and ALCL samples according to the expression of 91 probe sets (specific for 41 genes) regulated by activated NF-κB [29].(G) Hierarchical clustering of PTCLs according to 191 Ras-regulated probe sets (specific for 109 genes)[30]. Up- (red) and down-regulated (blue) transcripts are indicated based on the reference bar. ALK+ ALCL correspond to red color coded samples and ALK- to yellow color coded cases. (H) Hierarchical clustering of PTCL-NOS, AILT, normal T cells and ALCL samples according to the expression of 247 probe sets (specific for 142 genes) regulated by MYC [31,32]. (I) Venn diagram depicting ALCLs with overlapping signalling pathways.

Having established that the PDTs provide a true representation of the original tumors[11], we focused on the ALCL2-PDT line that carries lesions known to have a tumorigenic role: *JAK2* amplification and *NF1* deletion (Tab. S1 and Fig. 2A, B) [12,13]. In the absence of phosphorylated JAK2 (data not shown) we performed mRNA massive parallel sequencing (RNAseq) to examine whether there were additional defect(s) involving the NF1 locus and/or other lesions. We uncovered a novel chimeric transcript in which the N-terminus of the *NF1* gene (exons 1 to 46) was fused, out of frame, to the third exon of the *MAZ* gene (16p11.2) (Fig. 2C). The predicted outcome was a truncated NF1 (tNF1) protein, including the Ras Gap and Sec14 domains and able to activate the Ras pathway constitutively. The loss of heterozygosity at the *NF1* locus and the presence of a chimera were validated in both primary clinical and PDT samples (Fig. 2D). Constitutive Ras activation was confirmed by the high expression of genes, known to be regulated by Ras and by the relatively high levels of pErk1/2 (Fig. 2E). Importantly, when the ALCL-2 PDTs were treated *in vivo* with a selective MEK inhibitor (AZD6244), the levels of pERK1/2 were reduced and tumor growth was delayed (Fig. 2G). Taken together, these findings are consistent with a novel mechanism of activation of RAS in ALK- ALCL. The observation that a subset of ALK- and ALK+ ALCLs share signatures linked to RAS activation provides a strong rationale for targeted approaches in ALCL therapy[5]. MEK inhibitors control the cell growth of NF1-/- neurofibroma and malignant peripheral nerve sheath tumors, proving the feasibility of the strategy [14,15] and loss of NF1 has been documented as tumorigenic event in other hematopoietic neoplasms [13,16-18].

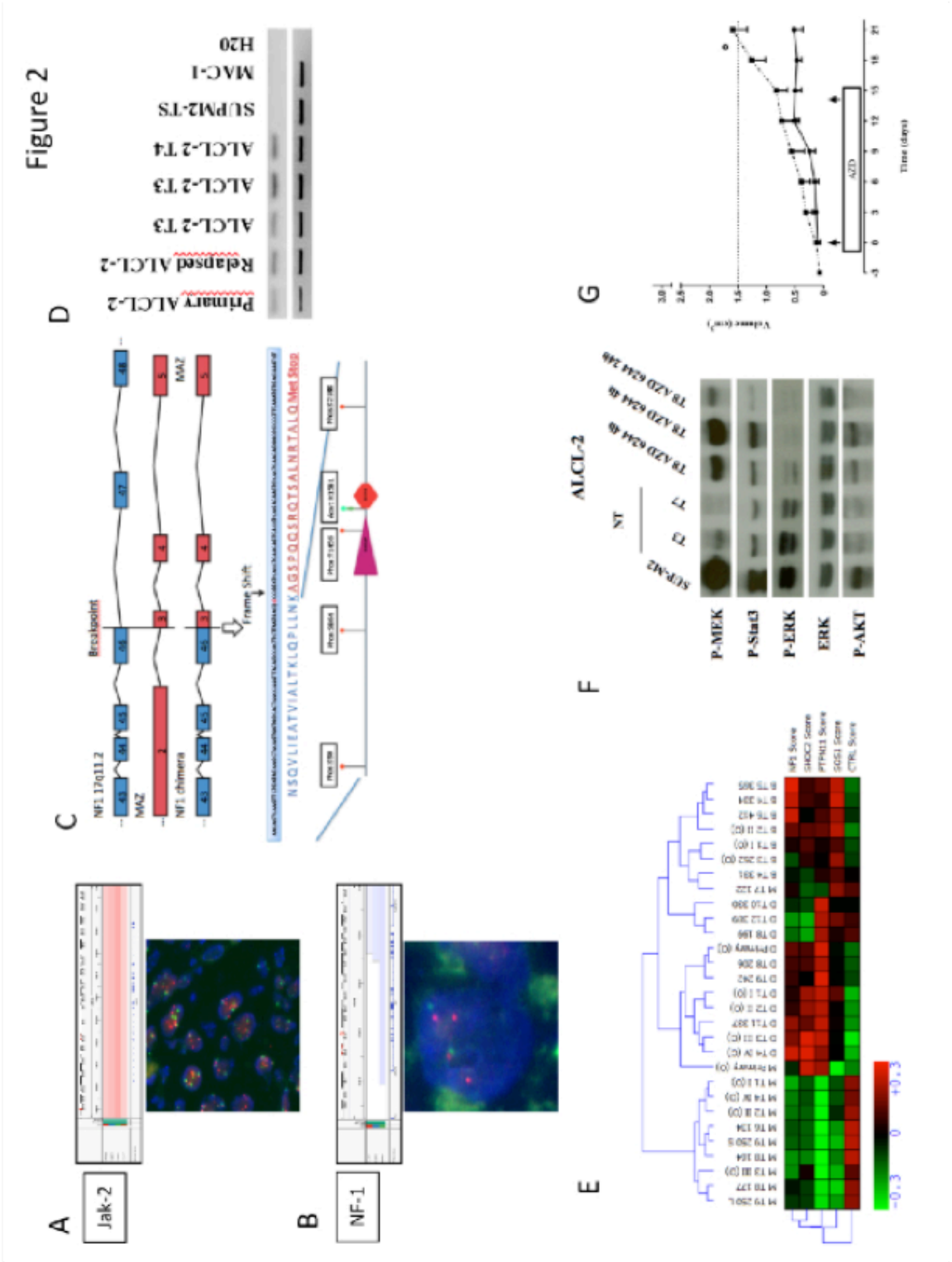


Figure 2
A
Jak-2
B
NF-1
C
NF1 17q11.2
MAZ
NF1 chimera
Breakpoint
Frame Shift
N S Q V L I E A T V I A L T K L Q P L L N K A G S P Q Q S R I Q T S A L N R T A L Q Met Stop
D
Primary ALCL-2
Relapsed ALCL-2
ALCL-2 T3
ALCL-2 T3
ALCL-2 T4
SU/PM2-TS
MAC-1
H20
E
F
ALCL-2
NT
T3
T7
TR AZD 624 4h
TR AZD 624 4h
TR AZD 624 24h
S(P)-M2
P-MEK
P-Stat3
P-ERK
ERK
P-AKT
G
Volume (cm³)
Time (days)
AZD

Fig. 2. ALCL-PDTs display multiple activated pathways and a novel NF-1-MAZ fusion transcript.

(A) SNP-array plots showing amplification in ALCL-2-PDT-T5 corresponding to the JAK2 gene locus at 9p24. FISH validation, red signals correspond to the JAK2 locus [RP11 39K24] and green signals to the centromere of chromosome 9 (Cep9 green). (B) SNP-array plots showing focal deletions in the primary and ALCL-2-PDT-T5 corresponding to the NF1 locus at 17q11.2. ALCL-2 FISH validation: Red signals correspond to flanking region [RP11 1107G21 (orange)/RP11 248L11 (green)]. (C) Schematic representation of the NF1 (blue) and MAZ (red) genes. (D) NF1-MAZ fusion products are detected in the primary and ALCL2-PDT samples by PCR. (E) Hierarchical clustering of primary and ALCL-PDTs according to expression of RAS-associated genes. (F) Basal protein expression in multiple ALCL-PDTs treated with AZD6244. ALK+ Supm2 cells were blotted as unrelated sample. Total ERK was used to assess protein loading. (G) In vivo growth of ALCL-2 PDTs is inhibited by AZD6244 treatment. Dotted lines: untreated (6 mice); full line treated samples (7 mice). The box indicates treatment schedule and the arrows starting and end points.

A second PDT (ALCL11-PDT) was derived from a patient with a very aggressive clinical course. Initially diagnosed as a nodal ALK- ALCL, the lymphoma rapidly became leukemic (Fig. S4). An initial workout demonstrated strong nuclear expression of p50 and p52 and the presence of phospho-STAT3 in the neoplastic cells, suggesting activation of the NF-κB and JAK-STAT pathways (Fig 3A). As ALK chimeras can target several pathways, we undertook a search for ALK chimeric transcripts. This demonstrated the presence of mRNA corresponding to the intracytoplasmic region of the molecule but none of the known partners of the ALK gene were found in the fusion protein. By means of RNAseq we identified a novel TRAF1-ALK chimera (Fig. 3B) encoding the first five exons of the *TRAF1* gene fused, in frame, to *ALK* exons 20-29. Validation in primary and relapsed samples as well as in ALCL11-PDT by Sanger DNA sequencing and RT-PCR confirmed the presence of TRAF1-ALK transcripts. FISH analysis and IHC proved that the ALK locus had been rearranged and that the ALK protein was present in the cytoplasm (Fig. 3D). To check whether other primary ALCL could display similar lesion(s), we tested all available ALCL with ALK cytoplasmic expression (total 7), failing to detected additional positive TRAF1-ALK cases. The TRAF1-ALK fusion interacts with known ALK adaptor molecules (i.e. Shp2 and STAT3), which undergo dephosphorylation in the presence of ALK inhibitors (Fig. 3E). As TRAF1 interacts with TRAF2 leading to NFkB activation[19], we tested whether TRAF1-ALK could bind to TRAF2 (Fig. 3 F-G). The intracytoplasmic region of ALK contains a docking site for TRAF2 [20], so we constructed a truncated form of TRAF1-ALK lacking this docking site (DTRAF-ALK) and tested whether DTRAF-ALK could heterodimerize with TRAF2. A similar molecule (DNPM-ALK) was used as a control [20]. As shown in Fig. 3H, DTRAF-ALK co-precipitated with TRAF2. Notably, ALCL-11 cells display a robust, constitutive activation of both canonical and non-canonical NFkB pathways (Fig. 3I). These findings represent the first evidence that the 5' portion of an ALK chimera has an independent pathogenic role, rather than merely providing heterodimerization domains.

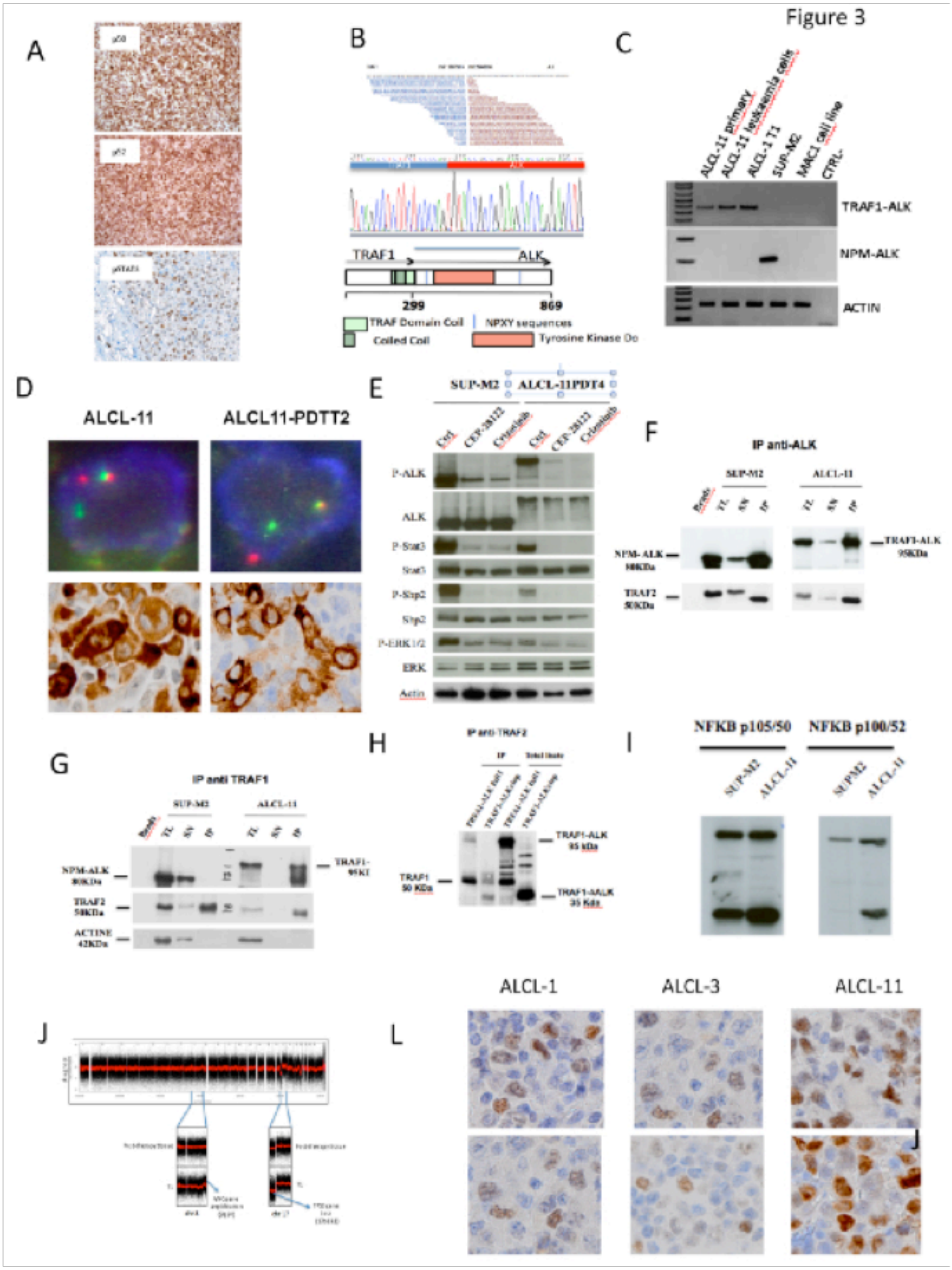


Figure 3
A
B
TRAF1
ALK
299
869
TRAF Domain Coil
NPXY sequences
Coiled Coil
Tyrosine Kinase Do
C
ALCL-11 primary
ALCL-11 leukaemia cells
ALCL-1 T1
SUP-M2
MAC1 cell line
CTRL-
TRAF1-ALK
NPM-ALK
ACTIN
D
ALCL-11
ALCL11-PDTT2
E
SUP-M2
ALCL-11PDT4
Ctrl
CEP-28122
Crizotinib
P-ALK
ALK
P-Stat3
Stat3
P-Shp2
Shp2
P-ERK1/2
ERK
Actin
F
IP anti-ALK
SUP-M2
ALCL-11
NPM-ALK 80KDa
TRAF2 50KDa
TRAF1-ALK 95KDa
G
IP anti TRAF1
SUP-M2
ALCL-11
NPM-ALK 80KDa
TRAF2 50KDa
ACTINE 42KDa
H
IP anti-TRAF2
TRAF1-ALK 95 kDa
TRAF1 50 KDa
I
NFKB p105/50
NFKB p100/52
SUP-M2
ALCL-11
J
L
ALCL-1
ALCL-3
ALCL-11

Fig. 3. ALCL11-PDT has a new TRAF1-ALK fusion protein and MYC gene amplification targetable by selective small molecules.

(A) Immunohistochemistry stains performed with the indicated antibodies demonstrated a a strong nuclear expression of NF-κB (p50 and p52) and pSTAT3 in the primary lymphoma (ALCL-11). (B) TRAF1–ALK gene fusion was defined by RNAseq (with 30 mate-pair overlapping sequences; pairing TRAF1 and ALK). The genomic coordinates are shown. TRAF1 exon 5 and ALK exon 20 reading frames were conserved (central panel). The TRAF1-ALK chimera includes the TRAF1 and Coiled-coil domains, fused to the ALK intracytoplasmic region (lower panel). (C) TRAF1-ALK fusion transcripts are detected in primary, leukemic and ALCL11-PDT samples by RT-PCR. ALK + (SUP-M2) and ALK- (Mac-1) lines were used as controls. Control lacks the cDNA template. (D) ALK break-apart probes are consistent with an ALK translocation (upper panels). An Anti-ALK Ab (D5F3) locates the ALK fusion proteins within the cytoplasmic (primary and ALCL11-PDT-T4 (lower panels). (E) ALK signalling in ALCL11-PDT is inhibited by anti-ALK molecules. PDT cells were treated in vitro (6hrs, Crizotinib or CEP28122, 200nM). Total cell lysates were immunoblotted with the indicated antibodies. (F-G) Primary ALCL-11 cells were immunoprecipitated (anti-ALK or anti-TRAF1) and blotted with specific Abs. (H) ΔT-ALK interacts with TRAF2. ΔT-ALK 293T cell lysates were immunoprecipated (TRAF2) and reacted with anti-TRAF1 Ab. (I) ALCL-11 cells show constitutive activation of the NFkB pathway. Primary ALCL-11 cells were immunoblotted with antibodies recognizing p105/p50 and p100/p52 NF-κB. NPM-ALK Supm2 cells were used for comparison. (J) SNP-array plots in ALCL11-PDT show a 17p13.1 deletion and MYC amplification. X-axis: chromosome localization and physical mapping; Y-axis: signal indicating copy number status at each locus. (L) IHC stains in primary and PDTs (ALCL-1-PDT-T2, ALCL-3-PDT-T8 and ALCL-11-PDT-T4) by anti-MYC mAb (Y69, Epitomics, Burlingame, CA) (Magnification 20x).

Although the majority of ALK+ ALCL patients reach remission, there are refractory and/or aggressive forms [21], but the cellular and molecular mechanisms underlying tumor progression are unclear [22]. However, *MYC* gene translocations have been documented in leukemic forms [23] and *MYC* mutations may be present in ALCL [7]. The genome analysis in the present study underlies the importance of the loss of 17p13 and a gain on 8q24 as secondary defects in leukemic samples of ALCL-11 (Fig. 3J and Table S3), suggesting the involvement of TP53 deficiency and *MYC* amplification. *In silico*, a MYC signature was seen in a leukemic ALCL-11 sample (Fig. S2) and IHC staining showed the presence of nuclear MYC in primary as well as in PDT lesions (Fig. 3L). We tested whether ALK signaling is critical to maintain the neoplastic phenotype [24] in the leukemic ALCL11-PDT model (Fig. 4A-B). A 14- or 21-day treatment with a selective ALK inhibitor [25] (CEP28122 100mg/kg/bid) led to significantly slower tumor progression. After the interruption of therapy, tumour progression again accelerated (Fig. 4A-B). Mice treated with CEP28122 survived longer (Fig. 4C) than controls, although both groups eventually succumbed to the disease. As ALCL-11 PDT harbour *MYC* amplification (Fig. 3J), we tested the effect of a novel BET-domain inhibitor targeting c-MYC expression (JQ1) *in vitro*[26]. When primary and ALCL-11 PDT cells were incubated with JQ1 or Bortezomib we observed a dose-dependent growth inhibition and down-regulation of c-myc protein expression in treated cells (Fig. 4D-E), suggesting that targeting the MYC and NF-κB pathways may provide a therapeutic benefit. This interpretation is supported by the more successful control of the leukemic progression by a protocol consisting of a short treatment with JQ1 (50 mg/kg IP daily, 1-7 days) in combination with CEP28122 (100mg/kg/BID OS, 1-14 days, Fig. 4F).

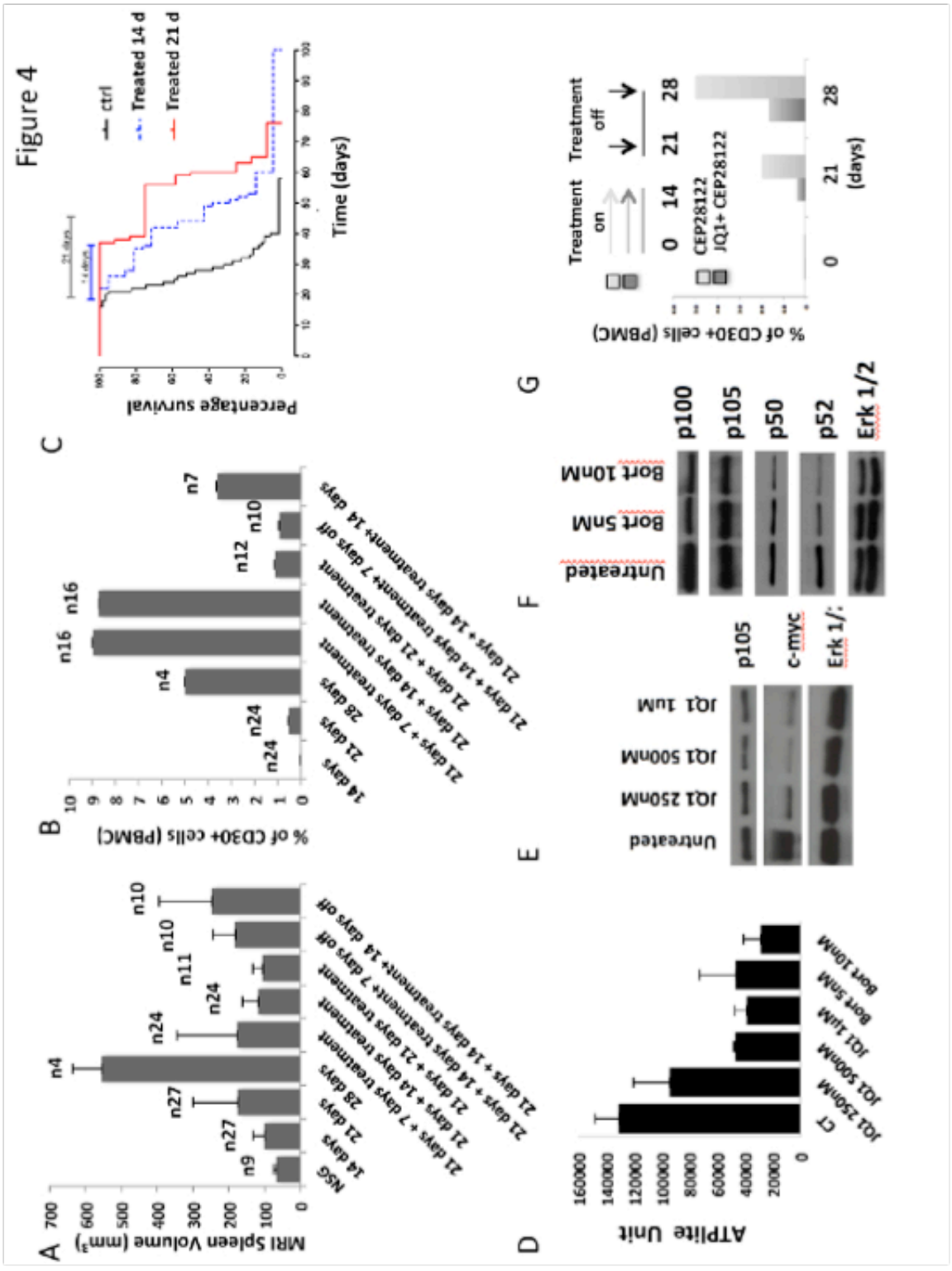


Figure 4
A
MRI Spleen Volume (mm³)
NSG
14 days
21 days
28 days
21 days + 7 days treatment
21 days + 14 days treatment
21 days + 21 days treatment
21 days + 14 days treatment+ 7 days off
21 days + 14 days treatment+ 14 days off
B
% of CD30+ cells (PBMC)
C
Percentage survival
Time (days)
ctrl
Treated 14 d
Treated 21 d
D
ATPlite Unit
CT
JQ1 250nM
JQ1 500nM
JQ1 1µM
Bort 5nM
Bort 10nM
E
Untreated
JQ1 250nM
JQ1 500nM
JQ1 1µM
p105
c-myc
Erk 1/2
F
Untreated
Bort 5nM
Bort 10nM
p100
p105
p50
p52
Erk 1/2
G
% of CD30+ cells (PBMC)
Treatment on
Treatment off
CEP28122
JQ1+ CEP28122
(days)

Fig. 4. Molecular-based therapies lead to positive responses in ALCL1-11-PDT.

(A) ALCL-11 NSG ($1x10^6$ lymphoma cells, i.v) were bled and CD45+CD30+ cells circulating in the peripheral blood determined by flow cytometry. Treatment of mice (CEP28122, 100mg/kg7BID) was started at day 21 from the i.v. inoculum and continued as described. (B) ALCL-11 NSG ($1x10^6$ lymphoma cells, i.v.) were imaged by MRI and spleen dimensions calculated. Treatment of mice (CEP28122, 100mg/kg7BID) was started at day 21 from the i.v. inoculum and continued as described. (C) Treatment with anti-ALK (CEP28122) prolongs overall survival of ALCL-11-PDT mice. Mice were randomized and treated as indicated (after 14 day from i.v. injection) with 100mg/kg/BID CEP28122. (D) ALCL-11-PDTs tumor growth is delayed in presence of JQ1 or Bortezomib molecules, in vitro. Highly enriched lymphoma cells (>95%, $1x10^5$/6 well plate) were incubated with different concentrations of drugs and cell viability was measured after 48hrs (ATPlite assay). (E) Protein expression of untreated and JQ1 treated ALCL-11-PDTs were determined at 36 hrs by Western Blotting with specific antibodies. (F) Protein expression of untreated and treated (Bortezomib) ALCL-11-PDTs was determined at 36 hrs by Western Blotting with specific antibodies. (G) The combination of anti-ALK (CEP28122) and anti-MYC (JQ1) therapy leads to better responses than a single anti-ALK treatment. Animals were treated with CEP28122 (100mg/kg/BID starting on day 10 from i.v. tumor injection ($5x10^6$ cells) for 14 days). JQ1 was given from day 14 to 21 in combination with CEP28122.

In summary, our integrated approach has led to the discovery of two novel translocations that cause the activation of several oncogenic pathways. *In silico*, ALCLs can be kept distinct from other PTCLs by virtue of their exclusive signatures (i.e. STAT3, MYC, NF-κB, RAS etc.), suggesting that they may be driven by a different set of lesions [including the t(2;x)(p23;x) and the t(6;7)(p25.3;q32.3)] [5,27], which fire a restricted number of pathways and share overlapping mechanisms of transformation. This hypothesis fits well with current models of the initiation and progression of cancer [1,2]. There is a growing perception that specific mutations may be present in small subsets of different types of human cancers. Indeed, it is becoming evident that human tumors harbor unique/patient-restricted combination(s) of rare mutations, which can best be treated by tailored therapy. The identification of pathways activated in individual patients should facilitate the implementation of personalized therapies, even in the absence of the identification of the associated lesions.

Our data have uncovered novel mechanisms of the activation of the RAS and NF-κB pathways that can drive the neoplastic phenotype, as demonstrated by the therapeutic efficacy of selective inhibitors *in vivo*. PDT-based models not only provide a tool to validate new tumorigenic defects but also allow the construction of ALCL classifiers able to predict clinical outcomes. This is critical to define individual subsets among heterogeneous groups, such as ALK- ALCLs or refractory ALCLs, which include some ALK+ ALCLs.

Methods

Patients and Case Selection

Fresh and/or viable cryopreserved samples (from 12 ALCL patients: 5 ALK+ and 7 ALK-) were obtained at the time of diagnosis, before treatment, or at relapse after single or multiple chemotherapy protocols. Informed consent was obtained following the procedures approved by the local ethical committees. Diagnoses were assigned according to the WHO classification by at least two experienced pathologists. Representative formalin-fixed tumor sections and/or tissue microarrays (TMAs) were processed for immunohistochemical analysis on a semi-automatedstainer (Piva et al., 2010).

Cell cultures and viability

Human ALCL cells TS-Supm2 and were cultured under standard conditions (37°C in humidified atmosphere, with 5% CO_2) in RPMI 1640 (Sigma-Aldrich, St Luis, MO, USA) supplemented with 10% fetal calf serum (Lonza, Rockland, ME, USA), 2 mM glutamine, 100 U/ml penicillin and 100µg/ml streptomycin (Eurobio Biotechnology, Les Ulis, France). Human embryonal kidney cells HEK-293T (ATCC, Manassas, VA) were cultured in Dulbecco modified Eagle medium (DMEM) as described above. Transfection of HEK-293T cells was performed with Effectene reagent (Qiagen, Valencia, CA), according to the manufacturer's instructions. DTRAF1-ALK and .DNPM-ALK cassettes were constructed by PCR including the first 299 aa of TRAF1 and 100aa of NPM link to a short peptide corresponding to the intracytoplamic region fo the ALK gene.
Cell viability was measured using the ATP-lite using Perkin Elmer luminometer.

Mice

Male and female NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG) mice were kindly provided by L. Shultz and bred within the Molecular Biotechnology Center (MBC) Animal Resource, under strict specific and opportunistic pathogen free (SOPF) conditions. The animal protocols were reviewed and approved by the Animal Committee of the University of Torino.

Tumor grafting and harvesting

From the surgical theater, fresh tissue fragments (pre-graft tissues) were stored in PBS and/or in complete media (RMI-1640 supplemented with 10% FCS and antibiotics), and rapidly delivered into the surgical pathology triage room. Dedicated tissue fragments were selected for tumorgraft implants only when sufficient material was available for routine diagnostic and molecular analyses. All samples were processed in sterile conditions. Tumor graft samples were cut into multiple 3x3x1 mm pieces (multiple pieces/specimen) in complete media and implanted fresh and/or cryopreserved in 10% DMSO-RPMI-1640 frozen media supplemented with 20% FCS.
Six- to eight-week-old NSG (Shultz et al., 2007) were first anesthetized (Rompun 0.05µl/g e Zoletil 1.6µg/g i.m.). Then, with animals lying on their ventral site, the dorsal region was sterilized (70% ethanol). A skin incision of 0.3 cm was subsequently made along the dorsal midline retro-nuchal region and a small pocket was created by blunt dissection. Multiple tumorgraft tissue fragments (2-4) were transferred into the

subcutaneous in each pocket using blunt-ended forceps with or without matrigel. The cut edges were sealed with a single metal clip. After the procedure, mice were regularly checked until they became vigil. Implanted animals were housed in same-sex groups. Similar frequencies were seen in both male and female mice. Implant growth was assessed by palpation and or by MRI scanning and harvested when tumor masses were required. Recipient animals were checked regularly and sacrificed at early sign of distress. At harvesting, mice were sacrificed in a CO_2 chamber and grafts were collected for histologic evaluation, re-grafting, or snap-freezing in liquid nitrogen.

Magnetic resonance imaging and data analysis and postprocessing

Prior to MRI examination animals were anesthetized by injecting tiletamine/zolazepam (Zoletil 100), 20 mg/kg, + xylazine (Rompun), 5 mg/ kg. MR images of NSG mice grafted with ALCL were acquired on a Bruker MRI scanner operating at 7 Tesla (Bruker Biospin, Ettlingen, Germany) using a 30 mm micro-imaging bird-cage coil. T_2-weighted anatomical images were acquired with a RARE sequence (TR/TE/RareFactor/NEX4000 ms/ 35.5 ms/16/3) covering 20 slices, thickness 1 mm, Field of View 30 mm and Matrix Size 256, for an in-plane resolution of 117 μm.

Whole-body MR images were acquired on the M2 Aspect 1T MRI scanner (Aspect Imaging, Shoham, Israel) equipped with a 30 mm solenoid RX/TX coil. T_2-weighted anatomical images were acquired with a Fast Spin Echo sequence (TR/TE/NEX 2800 ms/ 44 ms/2) covering 21 slices, thickness 1 mm, gap 0.1 mm, Field of View 100 mm and Matrix Size 256, for an in-plane resolution of 391 μm.

MRI images were manually segmented using 3D Doctor software to calculate the volume of target organs for 3D rendering.

Gene expression profiling

GEP was performed as described (Piva et al., 2010). Total RNA was extracted using the TRIZOL reagent (Invitrogen, Karlsruhe, Germany), and purified (Qiagen, Hilden, Germany). Synthesis of cDNA and biotinylated cRNAs was performed by Illumina Total Prep RNA Amplification Kit. Labeled cRNAs were hybridized on HumanHT-12 v4 Expression BeadChips (Illumina, San Diego, CA, USA). Bioinformatic analysis was performed as described (Fu and Medico, 2007).

RT-qPCR

RT-qPCR was performed with a Thermal iCycler (Bio-Rad) using the Bio-Rad iQ SYBR Green Supermix according to the manufacturer's instructions. The PCR cycling conditions were as follows: 95°C for 5 minutes, followed by 40 cycles at 94°C for 10 seconds and 60 or 62°C for 30 seconds. The oligonucleotide primer pairs used for RT-qPCR were designed with the PrimerBLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) to obtain amplicons of 70-110 bp. Primer sequences are available if requested. To confirm the amplification specificity, the PCR products were subjected to the analysis of melting curve, linearity, and slope of standard curve. All PCR assays were performed in triplicate. Gene-expression results were normalized to GAPDH and HUPO expressions and calculated using the DCt method according to the manufacturer's instructions.

The quality of cryopreserved and FFPE samples was assessed according to the following qPCR parameters: PTCL-NOS with GAPDH/HUPO Ct value D30, ALCL with either GAPDH/HUPO Ct value D30 or TNFRSF8 DCt value D5 were excluded from the analysis.

Metagenes were calculated using the mean of DCt established in reference cell lines (SUPM2, SuDHL1, Karpas 299 and L82). Target gene expression levels were determined substracting each individual DCt to the corresponding metagene (DDCt). Moreover, to normalize for the tumor content, target DDCt values were divided by CD30 DCt of each corresponding sample.

Genome-wide DNA profiling

DNA samples were obtained as previously reported (Rinaldi et al.). DNA integrity was verified by electrophoresis of 50 ng on a 1% agarose gel. Genomic profiles were obtained using Affymetrix Human Mapping GeneChip 6.0 arrays (Affymetrix, Santa Clara, CA, US). Initial array intensity data were obtained using AGCC Software with 270 Caucasian HapMap samples as a reference dataset for normalization (Mian et al., 2012).

Raw copy number (CN) was extracted from CEL files using the APT software package with standard parameters. Probes with an inter-probes distance of less than 180bp were discarded. The copy number for the X chromosome was centered to the median value. Genomic profiles were then segmented with the 'Fast first-derivative segmentation' (FFSEG) algorithm (Kwee, 2012). A multiple filtering strategy was then applied to minimize false positives. Segments containing less than 10 probes or with a

size smaller than 2,000 bp were discarded. Segments were filtered for copy number variation if they overlapped at least 50% with a known CNV according to the Database of Genomic Variants (Iafrate et al., 2004). Segments spanning the centromeres were filtered out. After CNV filtering, we employed a 'gap filling' (Lenz et al., 2008): small gaps of less than 500kb and flanked by aberrant regions larger than the gap size itself were closed; larger gaps between 500kb and 10Mb were filled if they were flanked by aberrant regions larger at least 1.5 times the gap itself. Finally, neighbouring segments were iteratively merged if their difference was smaller than 0.15 (log2ratio). CN segments were discretized as homozygous deleted, lost, normal, gained, or amplified using the thresholds -0.60, -0.10, 0.10 and 0.60 (log2ratio). As final quality control, filtered and merged segmented profiles, together with their raw profiles, were visually inspected. Profiles were considered as poor quality and discarded from further analyses if they showed severe over-segmentation or no aberrations at all as evaluated by two independent investigators (IK, FB).

Correlation between gene expression and copy number was computed by a one-sided Kendall's tau test. We calculated the moving average of the relative expression along the genome with a window of 20 genes. To show the local impact of copy number to gene expression, the local covariance of the smoothed relative expression (centered log2 values) and copy number (non-scaled) was calculated. Regions with high covariance corresponded to regions that showed concurrent local overexpression (or under-expression) together with copy number gain (or loss).

Raw data will be available at the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo) database.

Dual-color fluorescence in situ hybridization (FISH)

BAC clones spanning the gene loci were obtained from BACPAC Resources at http://bacpac.chori.org (Tab. S2). Commercial FITC centromeric probes were also used (Abbott S.p.A, Rome, IT). DNA was labeled by nick-translation using spectrum green or spectrum red dUTP fluorochromes (Abbott S.p.A, Rome, IT). FISH was performed by standard methods and at least 20 metaphase spreads or 200 interphase nuclei on DAPI-stained slides were scored.

RNA-Seq analysis

RNA-seq was performed as previously described (Blevins et al., 2011). Two microgram of RNA were retrotranscribed using random nucleotides primers and RNase H. cDNAs were then sheared to an average fragment size of 200 bases. In brief, polyA-containing mRNA was purified using oligo-dT beads from 10 ug of total RNAs for each sample and fragmented into small fragments using Zn cations under elevated temperature. Cleaved RNA fragments were reverse-transcribed into first strand cDNA using random primers (Invitrogen Inc.), followed by second-strand cDNA synthesis. After end-repair processing, a single 'A' base was added to cDNA fragments at 3′ end. cDNAs were subsequently ligated to adapters, purified by 2% agarose gel, and then enriched by PCR to create the final cDNA library. RNA double-end sequencing was performed using Illumina 2000 using the standard protocol. The cDNA library of each sample was loaded to a single lane of an Illumina flow cell. The data were then analyzed using dedicated tools (Bellerophontes, (Abate et al., 2012) deFuse, (McPherson et al., 2011) Chimerascan (Iyer et al., 2011) and TxFuse (Singh et al., 2012) and Pelagus (Abate F. et al. in preparation).

Mice treatment

The following drugs and routes were implemented: CEP28122 (Cephalon, 100 mg/kg body weight BID), AZD6244 (ChemieTek, 100mg/kg), SGN-35 (1mg/kg q4dx3 or 4) were administered by oral gavage, and intravenous (i.v.) injections, respectively.

Analysis of clinical data

Overall patient survival (OS) was calculated from the time of histologic transformation to the last follow-up or death of any cause. Log-rank test was used to investigate the impact on OS of categorical variables. The cumulative probability of OS was plotted according to the Kaplan-Meier method. Statistical analysis was performed with the SPSS software.

REFERENCES METHODS

Abate, F., Acquaviva, A., Paciello, G., Foti, C., Ficarra, E., Ferrarini, A., Delledonne, M., Iacobucci, I., Soverini, S., Martinelli, G., and Macii, E. (2012). Bellerophontes: an RNA-Seq data analysis framework for chimeric transcripts discovery based on accurate fusion model. Bioinformatics *28*, 2114-2121.

Blevins, T., Rajeswaran, R., Aregger, M., Borah, B. K., Schepetilnikov, M., Baerlocher, L., Farinelli, L., Meins, F., Jr., Hohn, T., and Pooggin, M. M. (2011). Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. Nucleic acids research *39*, 5003-5014.

Fu, L., and Medico, E. (2007). FLAME, a novel fuzzy clustering method for the analysis of DNA microarray data. BMC Bioinformatics *8*, 3.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nature genetics *36*, 949-951.

Iyer, M. K., Chinnaiyan, A. M., and Maher, C. A. (2011). ChimeraScan: a tool for identifying chimeric transcription in sequencing data. Bioinformatics *27*, 2903-2904.

Kwee, I., Rinaldi, A., De Campos, C., and Bertoni, F. (2012). Fast and Robust Segmentation of Copy Number Profiles Using Multi-Scale Edge Detection. Submitted.

Lenz, G., Wright, G. W., Emre, N. C., Kohlhammer, H., Dave, S. S., Davis, R. E., Carty, S., Lam, L. T., Shaffer, A. L., Xiao, W.*, et al.* (2008). Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proceedings of the National Academy of Sciences of the United States of America *105*, 13520-13525.

McPherson, A., Hormozdiari, F., Zayed, A., Giuliany, R., Ha, G., Sun, M. G., Griffith, M., Heravi Moussavi, A., Senz, J., Melnyk, N.*, et al.* (2011). deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS computational biology *7*, e1001138.

Mian, M., Rinaldi, A., Mensah, A. A., Rossi, D., Ladetto, M., Forconi, F., Marasca, R., Gattei, V., Zucca, E., Cavalli, F.*, et al.* (2012). Del(13q14.3) length matters: an integrated analysis of genomic, fluorescence in situ hybridization and clinical data in 169 chronic lymphocytic leukaemia patients with 13q deletion alone or a normal karyotype. Hematological oncology *30*, 46-49.

Piva, R., Agnelli, L., Pellegrino, E., Todoerti, K., Grosso, V., Tamagno, I., Fornari, A., Martinoglio, B., Medico, E., Zamo, A.*, et al.* (2010). Gene expression profiling uncovers molecular classifiers for the recognition of anaplastic large-cell lymphoma within peripheral T-cell neoplasms. J Clin Oncol *28*, 1583-1590.

Rinaldi, A., Mian, M., Kwee, I., Rossi, D., Deambrogi, C., Mensah, A. A., Forconi, F., Spina, V., Cencini, E., Drandi, D.*, et al.* (2011). Genome-wide DNA profiling better defines the prognosis of chronic lymphocytic leukaemia. British journal of haematology *154*, 590-599.

Shultz, L. D., Ishikawa, F., and Greiner, D. L. (2007). Humanized mice in translational biomedical research. Nature reviews Immunology *7*, 118-130.
Singh, D., Chan, J. M., Zoppoli, P., Niola, F., Sullivan, R., Castano, A., Liu, E. M., Reichel, J., Porrati, P., Pellegatta, S*., et al.* (2012). Transforming fusions of FGFR and TACC genes in human glioblastoma. Science *337*, 1231-1235.

Affiliations and Acknowledgments

[1]Department of Molecular Biotechnology and Health Science and Center for Experimental Research and Medical Studies (CeRMS), University of Torino, Torino, 10126 Italy; [2]Department of Control and Computer Engineering, Politecnico di Torino, 10129, Italy; [3]Department of Biomedical Informatics, Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10027 USA; [4]Lymphoma and Genomics Research Program, IOR Institute of Oncology Research, Bellinzona, 6500 Switzerland; [5]Pathology & Lymphoid Malignancies Units, San Raffaele Scientific Institute, Milan, 20132 Italy; [6]Molecular Imaging Center, Department of Chemistry IFM and Molecular Imaging Center, University of Torino, Torino, 10125 Italy; [7]Oncology, Cephalon Inc.,41 Moores Road Frazer, PA 19355 USA; [8]University of Veterinary Medicine Vienna, 1210 Vienna, Austria, [9]Institute of Hematology and Medical Oncology L. and A. Seràgnoli, S. Orsola-Malpighi Hospital, University of Bologna, Bologna, 40138 Italy; [10]Institute of Hematology, University of Perugia, Ospedale S. Maria della Misericordia, S. Andrea delle Fratte, Perugia, 06156 Italy; [11]The Jackson Laboratory, Bar Harbor, ME, 04609 USA; [12]Fasteris SA, Plan-les-Ouates/Geneva, 1028 Switzerland; [13]IDSIA Dalle Molle Institute for Artificial Intelligence, Manno, CH-6928 Switzerland; [14]SIB Swiss Institute of Bioinformatics,1015 Lausanne, Switzerland; [15]Department of Pathology, and NYU Cancer Center, New York University School of Medicine, New York, NY, 10016 USA; and [16]Departmentof Oncology, University of Torino and Institute for Cancer Research at Candiolo, 10060 Italy, [17]Lymphoma Unit, IOSI Oncology Institute of Southern Switzerland, 6500 Bellinzona, Switzerland.

The European T-Cell Lymphoma Study Group: *Italy*: Cristina Abele, Luca Bessone, Antonella Barreca, Nicoletta Chiesa, Ramona Crescenzo, Antonella Fienga, Marcello Gaudiano, Filomena di Giacomo, Giorgio Inghirami, Indira Landra, Elena Lasorsa,

Rodolfo Marchiorlatti, Barbara Martinoglio, Enzo Medico, Gian Battista Ferrero, Katia Messana, Elisabetta Mereu, Elisa Pellegrino, Roberto Piva, Irene Scafo', Elisa Spaccarotella, Fabrizio Tabbo', Maria Todaro, Ivana Ubezzi, Susanna Urigu (Azienda Ospedaliera Città della Salute e della Scienza di Torino, University of Torino); Francesco Abate, Elisa Ficarra, Andrea Acquaviva (Politecnico di Torino); Antonella Barreca, Domenico Novero, Annalisa Chiapella and Umberto Vitolo (ASO Molinette, and San Luigi Gonzaga Torino); Marco Chilosi and Alberto Zamó (University of Verona); Fabio Facchetti and Silvia Lonardi (University of Brescia); Anna De Chiara and Franco Fulciniti (National Cancer Institute, Napoli); Claudio Doglioni, Andrés Ferreri and Maurilio Ponzoni (San Raffaele Institute, Milano); Claudio Agostinelli, Pier Paolo Piccaluga and Stefano Pileri (University of Bologna); Brunangelo Falini and Enrico Tiacci (University of Perugia). *Belgium*: Peter Van Loo, Thomas Tousseyn, and Christiane De Wolf-Peeters (University of Leuven); *Germany*: Eva Geissinger, Hans Konrad Muller-Hermelink and Andreas Rosenwald, (University of Wuerzburg); *Spain*: Miguel Angel Piris and Maria E. Rodriguez (Hospital Universitario Marqués de Valdecilla, IFIMAV, Santander and Instituto de Investigaciones Biomédicas Alberto Sols, CSIC-UAM, Madrid)

Current address for Mangeng Cheng: In Vitro Pharmacology, Merck Research Laboratory, BMB11-138, 33 Avenue Louis Pasteur, Boston, MA 02115

GI is supported by the Italian Association for Cancer Research (AIRC) Special Program in Clinical Molecular Oncology, Milan (5x1000 No. 10007); Regione Piemonte (ONCOPROT, CIPE 25/2005); ImmOnc (Innovative approaches to boost the immune responses, Programma Operativo Regionale, Piattaforme Innovative BIO F.E.S.R. 2007/13, Asse 1 'Ricerca e innovazione' della LR 34/2004) and the Oncology Program of Compagnia di San Paolo, Torino. RR is supported by Partnership for Cure, NIH 1 P50 MH094267-01, NIH 1 U54 CA121852-05, NIH 1R01CA164152-01. FB sponsored by the Oncosuisse KLS-02403-02-2009 (Bern, Switzerland); Anna Lisa Stiftung (Ascona, Switzerland); Nelia and Amadeo Barletta Foundation (Lausanne, Switzerland); RP by Rete Oncologica del Piemonte e della Valle d'Aosta. LDS is supported by National Institutes of Health (USA) grant CA034196. M.B. is enrolled in the PhD program in Pharmaceutical Sciences, University of Geneva, Switzerland. We thank Drs. Vigliani C, Fioravanti A, and Mossino M for their technical support.

REFERENCES

1. Vogelstein, B. & Kinzler, K.W. Cancer genes and the pathways they control. *Nat Med***10**, 789-799 (2004).
2. Jones, S.*, et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science***321**, 1801-1806 (2008).
3. Piva, R.*, et al.* Gene expression profiling uncovers molecular classifiers for the recognition of anaplastic large-cell lymphoma within peripheral T-cell neoplasms. *J Clin Oncol***28**, 1583-1590 (2010).
4. Zamo, A.*, et al.* Anaplastic lymphoma kinase (ALK) activates Stat3 and protects hematopoietic cells from cell death. *Oncogene***21**, 1038-1047 (2002).
5. Tabbo, F., Barreca, A., Piva, R. & Inghirami, G. ALK Signaling and Target Therapy in Anaplastic Large Cell Lymphoma. *Front Oncol***2**, 41 (2012).
6. Eckerle, S.*, et al.* Gene expression profiling of isolated tumour cells from anaplastic large cell lymphomas: insights into its cellular origin, pathogenesis and relation to Hodgkin lymphoma. *Leukemia***23**, 2129-2138 (2009).
7. Inghirami, G.*, et al.* Molecular characterization of CD30+ anaplastic large-cell lymphoma: high frequency of c-myc proto-oncogene activation. *Blood***83**, 3581-3590 (1994).
8. d'Amore, F., Jantunen, E. & Relander, T. Hemopoietic stem cell transplantation in T-cell malignancies: who, when, and how? *Curr Hematol Malig Rep***4**, 236-244 (2009).
9. Shultz, L.D., Brehm, M.A., Garcia-Martinez, J.V. & Greiner, D.L. Humanized mice for immune system investigation: progress, promise and challenges. *Nat Rev Immunol***12**, 786-798 (2012).
10. Garber, K. From human to mouse and back: 'tumorgraft' models surge in popularity. *J Natl Cancer Inst***101**, 6-8 (2009).
11. Macor, P.*, et al.* An update on the xenograft and mouse models suitable for investigating new therapeutic compounds for the treatment of B-cell malignancies. *Current pharmaceutical design***14**, 2023-2039 (2008).
12. Green, M.R.*, et al.* Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood***116**, 3268-3277 (2010).
13. Boudry-Labis, E.*, et al.* Neurofibromatosis-1 gene deletions and mutations in de novo adult acute myeloid leukemia. *American journal of hematology* (2013).
14. Staser, K.*, et al.* Normal hematopoiesis and neurofibromin-deficient myeloproliferative disease require Erk. *J Clin Invest***123**, 329-334 (2013).
15. Chang, T.*, et al.* Sustained MEK inhibition abrogates myeloproliferative disease in Nf1 mutant mice. *J Clin Invest***123**, 335-339 (2013).

16. Biagi, C.*, et al.* Pediatric early T-cell precursor leukemia with NF1 deletion and high-sensitivity in vitro to tipifarnib. *Leukemia***24**, 1230-1233 (2010).
17. Kalender Atak, Z.*, et al.* High accuracy mutation detection in leukemia on a selected panel of cancer genes. *PloS one***7**, e38463 (2012).
18. Mavrakis, K.J.*, et al.* A cooperative microRNA-tumor suppressor gene network in acute T-cell lymphoblastic leukemia (T-ALL). *Nat Genet***43**, 673-678 (2011).
19. Pomerantz, J.L. & Baltimore, D. NF-kappaB activation by a signaling complex containing TRAF2, TANK and TBK1, a novel IKK-related kinase. *The EMBO journal***18**, 6694-6704 (1999).
20. Horie, R.*, et al.* The NPM-ALK oncoprotein abrogates CD30 signaling and constitutive NF-kappaB activation in anaplastic large cell lymphoma. *Cancer cell***5**, 353-364 (2004).
21. Savage, K.J.*, et al.* ALK- anaplastic large-cell lymphoma is clinically and immunophenotypically different from both ALK+ ALCL and peripheral T-cell lymphoma, not otherwise specified: report from the International Peripheral T-Cell Lymphoma Project. *Blood***111**, 5496-5504 (2008).
22. Barreca, A.*, et al.* Anaplastic lymphoma kinase in human cancer. *Journal of molecular endocrinology***47**, R11-23 (2011).
23. Monaco, S.*, et al.* Pediatric ALK+ anaplastic large cell lymphoma with t(3;8)(q26.2;q24) translocation and c-myc rearrangement terminating in a leukemic phase. *American journal of hematology***82**, 59-64 (2007).
24. Gambacorti-Passerini, C., Messa, C. & Pogliani, E.M. Crizotinib in anaplastic large-cell lymphoma. *N Engl J Med***364**, 775-776 (2011).
25. Cheng, L.Y.*, et al.* Anaplastic lymphoma kinase spares organ growth during nutrient restriction in Drosophila. *Cell***146**, 435-447 (2011).
26. Delmore, J.E.*, et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell***146**, 904-917 (2011).
27. Feldman, A.L.*, et al.* Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing. *Blood***117**, 915-919 (2011).
28. Agnelli, L.*, et al.* Identification of a 3-gene model as a powerful diagnostic tool for the recognition of ALK-negative anaplastic large-cell lymphoma. *Blood***120**, 1274-1281 (2012).
29. Brune, V.*, et al.* Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. *J Exp Med***205**, 2251-2268 (2008).
30. Loboda, A.*, et al.* A gene expression signature of RAS pathway dependence predicts response to PI3K and RAS pathway inhibitors and expands the population of RAS pathway activated tumors. *BMC medical genomics***3**, 26 (2010).
31. Bild, A.H.*, et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature***439**, 353-357 (2006).
32. O'Connell, B.C.*, et al.* A large scale genetic analysis of c-Myc-regulated gene expression patterns. *J Biol Chem***278**, 12563-12573 (2003).

Discussion

Anaplastic Large Cell Lymphoma (ALCL) is a T-cell lymphoma, which accounts for about 5% of all NHLs, and it is more frequent in children and young adults. For this reason, only few studies regarding the genomic analysis of ALCL, and performed only in cohorts composed by few samples, have been reported in literature. Moreover, in studies already published, the genomic analysis was performed with techniques that did not allow a high-resolution analysis. The array-CGH approach, through the use of the Affymetrix SNP 6.0 platforms, for the study of genomic aberrations in this ALCL series has allowed us to identify specific CN changes that might be critical for the pathogenesis of the lymphoma. Different chromosomal regions are affected by gain or loss of genetic material in ALCL patients: the most frequent alterations are the loss of the region on the short arm of chromosome 17 encompassing the well characterized tumor suppressor gene *TP53*, detected in 25% of patients, and the loss of 6q21 locus, in which only *PRDM1* and *ATG5* genes are located. Gains of the long arm of chromosome 1 and gain of genomic material at level of chromosomes 7 and 8 are identified in more than 15% of ALCL samples analyzed.

Since ALCL patients can be further subdivided into two main groups, ALK+ ALCL and ALK- ALCL, on the basis of Anaplastic Lymphoma Kinase (ALK) gene expression, due to the presence of a translocation that lead to the creation and activation of a chimeric fusion protein in which ALK is fused to another partner gene, we wondered if some of the genomic lesions detected could be characteristic of one subtype over the other. From the comparison of the genomic aberrations between ALK+ ALCL and ALK- ALCL we first noticed that the latter group showed a more complex kariotype, with a higher number of chromosomal aberrations, in term of both gains and losses. Despite the differences in term of approach and resolution applied, these results are in line with previously published studies [80, 152]. The analysis of ALK-ALCL samples revealed the very frequent loss of the region 17p13.3-p12 (*TP53*) and of the region 6q16.3-q21 (42% and 35% of cases, respectively). However, these lesions were not specifically identified only in ALK- ALCL patients, being present at low frequency also in ALK+ samples. Different losses at level of 1q and 13q, lost in a percentage of cases ranging from 20 to 26%, are characteristic of this subtype together with losses of 16q23.1-q23.2, 12q24.31, 10p21.1 and 4q22.1 identified in more than 20% of ALK- samples. In about 30% of ALK-ALCL gains at level of chromosome 1 (long recurrent aberration of 1p11.1-q32.1 and short recurrent aberrations of 1q25.1 and 1q32.1-q41) have been identified. Conversely, the ALK+ ALCL samples showed less genomic lesions and at lower frequency: gains at level

of 1q32.2 and 7q32.3 were identified in 19% of cases, while bigger gains of material at level of the same chromosomes were detected in 16% of cases. Gains of 7q32.3, 7p22.3-p21.3 and of different locus of the long arm of chromosome 8, together with gain of 15q12 and loss of 11q24.3 were the only lesions, more frequent in the ALK+ subset, significantly differentially detected among the two subtypes. Our results could confirm the presence of only some lesions identified in the previous study of Salaverria at al [152]. Discrepancies could be due to the different techniques used but also to the sample size of the studies, both not very large. However, the confirmation of the detection of loss of 6q21 locus, together with losses affecting 13q and 4q and gains at level of 1q in our ALK- series supports the idea that those are lesions in which genes critical for the pathogenesis of this ALCL subtype might be located. Conversely, our results were in line with most of the lesions detected by Zettl at al [80], identified with a conventional CGH approach. In that case, the analysis of ALCL samples, together with some PTCL, NOS samples, disclosed the presence of different genomic aberrations in each subtypes, supporting the idea of the distinction between ALK+ and ALK, but also between ALK- and PTCL, NOS, whose genetic relationship has been always controversial.
The differences between ALK+ and ALK- ALCL in terms of their clinical presentation, response to therapy and survival, suggest that they must have different pathogenetic mechanisms. Analysis at a genomic level and the identification of differences in the profile of certain chromosomal aberrations observed in this study, in confirm of previous ones, would support the idea that these lymphomas can be considered definitively different entities because we show that they are also different diseases at a genomic level.

After the analysis of the genomic aberrations in our series composed by 69 ALCL patients, I tried to understand if it would be possible to identify different groups of patients within the ALK- and ALK+ subtypes, specifically distinguishable by different chromosomal alterations. The design of an heatmap in which all the ALCL cases analyzed have been reported and sorted based on the most frequent lesions detected, allowed us to identify two main groups of patients within both ALK- and ALK+ subsets.
More than half of the ALK- cases showed the presence of multiple lesions, with a very high frequency of the concomitant losses at 17p (*TP53)* and 6q21 (*PRDM1/ATG5)*. Conversely, other ALK- samples showed no genomic lesions at all. The surprising observation that patients of the same subtypes present such differences in term of number of genomic aberrations identified, could have two different explanations: the first

one that at least some of ALK- patients for which no lesions have been detected, represent false negative cases, due to a high content of infiltrating normal cells, and the second one that probably in these patients some alternative lesions lead to the transformation of T cells. It is possible that other mechanisms, not detectable with the array-CGH approach, could have a role in the lymphomagenesis such as mutations at level of oncogenes or tumor suppressor genes, change in pattern of promoters methylation, translocations, alteration of expression pattern at level of microRNA that, in turn, might affect expression of other genes. For this reason, for example, important informations might came from the on-going whole exome sequencing project.

In the ALK+ series two different groups of patients could also be recognized on the basis of their genomic aberrations. Less than half of the ALK+ patients showed very few and heterogeneous chromosomal lesions, not specifically associated with the ALK status, being detectable also in ALK- samples. For the remaining ALK+ cases, array-CGH analysis was not able to detect any CN change. For these ALK+ cases we can exclude the possibility they do not represent ALK+ ALCL, since for the diagnosis of these samples, the presence of ALK translocations and the consequent ALK expression, have been evaluated. Although, similarly to what discussed for ALK- cases, some of these samples with no aberrations could be false negative due to high content of infiltrating normal cells. However, it is known that ALK+ cases can have very few lesions additional to ALK translocations [152]. It is conceivable that the constitutive activation of the chimeric fusion protein, generated from the ALK translocation, is sufficient for the neoplastic transformation of the cell through the activation of many downstream pathways involved in cell cycle progression, cell proliferation and survival. Moreover, also in this case, other types of lesions, such as somatic mutations and promoter methylation might occur.

The identification of the aberrations occurring at a frequency higher than expected by chance, giving particular relevance to loci targeted by DNA amplifications or homozygous deletions, has been possible through the application of the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm [223]. Among the lesions that are identified as aberrations likely to contain genes involved in pathogenesis, gain of 9p24.1 (*JAK2*) and losses of 17p13.1 (*TP53*) and 6q21 (*PRDM1* and *ATG5*) are the most significant in the whole ALCL series. The application of the GISTIC algorithm in the single ALK- cohort, highlights even more that losses of 17p13.1 and 6q21 locus contain genes

which are likely to be involved in ALCL pathogenesis, and in particular in ALK- subtype pathogenesis.

The analysis of less frequent aberrations, but probably involved in ALCL pathogenesis, revealed the presence of amplification of *JAK2* gene located at 9p24.1 and *MIR17HG*, which codes for the miR17-92 cluster, at level of 13q31.3 in few cases. Genes involved in these lesions might represent possible therapeutic target to be evaluated in combination with conventional therapy, since both of them are involved in the direct regulation of *MYC* oncogene, which is also gained in other ALCL cases. So, our data provide the rational of the evaluation of JAK2 inhibitors, AKT/mTOR inhibitors (leading to a block of the pathway that can be overactivated by miR17-92), and new drugs, such as JQ1, which are able to interfere with MYC activity, for the treatment of ALCL patients bearing specific genomic alterations.

The application of a genomic approach in order to study the pathogenesis of ALCL allowed us to recognize some specific alterations occurring in both ALK+ and ALK- subtypes, and hence identify locus in which critical genes might be located.

The combination of SNP 6.0 profiling with the use of two different algorithms, the identification of Minimal Common Regions (MCRs), and GISTIC algorithm, pointed to 6q21 loss as one of the most frequent and significant lesion in ALCL, and in particular in ALK- ALCL samples. Since the MCR concerning the 6q21 loss encompass two different neighbor genes, *PRDM1* and *ATG5*, we performed a more careful analysis of the size of the loss in different samples. The observation that the smallest aberrations involved only the entire *PRDM1* gene, but not *ATG5*, and, especially, that the homozygous deletion, affecting one ALK+ sample, encompassed only *PRDM1* gene, suggest that the critical gene for ALCL pathogenesis could be *PRDM1*. Human *PRDM1* gene encodes two major isoforms designated BLIMP1α and BLIMP1β, which arise from alternative promoters [191]. While a specific role in plasma cells differentiation for the full-length BLIMP1α has been described [192], no clear functions have been attributed to BLIMP1β. Functional studies on BLIMP1 revealed that it is a key gene involved in both B and T cells development, differentiation and function: in the first case, BLIMP1 acts as a master regulator of plasma cell differentiation, playing a central role in the transcriptional network that controls B cell terminal differentiation [206, 208, 210, 211], while in T cells, BLIMP1 is mainly involved in the regulation of responsiveness and homeostatis, being more important for T cells function than development [189, 198, 205]. In our ALCL series,

PRDM1 was affected by the loss of one allele in most of the cases; only one ALK+ sample showed the loss of both alleles (hdel), suggesting the presumable complete loss of the protein. Interestingly, the same loss observed in the primary samples has been identified in ALCL cell lines as well. *PRDM1* gene, coding for BLIMP1 protein, has been reported to be inactivated in other types of lymphomas, such as the DLBCL and NK-cell lymphoma [49, 50, 91, 92, 216]. Numerous and different are the mechanisms of BLIMP1 inactivation in these cases. In 24% of ABC-DLBCL cases [49], and interestingly, in none GCB-DLBCL, BLIMP1 is inactivated as a result of several structural alteration such as gene truncations, nonsense mutations, frameshift deletions, splice site mutations, all of which lead to the formation of truncated BLIMP1 protein. In all cases, the inactivation is accompanied by deletion or by another mutations, supporting a two-hit hypotesis for the inactivation of the gene. The observation that BLIMP1 is commonly inactivated in a specific subtype of DLBCL suggested that it could be considered a *bona fide* tumor suppressor gene in ABC-DLBCL, whose inactivation may contribute to lymphomagenesis. Further works [50, 216] confirmed the tumor suppressor role of BLIMP1 in ABC-DLBCL, increasing the description of the mechanisms involved in such inactivation. Chromosomal inversions, single nucleotide mutations which lead to splicing alterations, hdel and transcriptional repression by constitutively active BCL6, have been reported as alternative mechanisms able to induce BLIMP1 inactivation in ABC-DLBCL cases. The idea that BLIMP1 could be specifically inactivated by a two-hit mechanism in a specific subtype of DLBCL, let us to think that probably, also for the ALK- subtype, BLIMP1 could have a tumor suppressor role. For this reason, I performed the sequencing and mutational analysis of the coding exons of *PRDM1* gene for almost all cases. Besides the identification of already known SNPs (D205E, P714P) in numerous cases, we observed a nucleotide change (A585G) in one sample, which could be interpreted as a new mutation. The analysis of the available matched normal DNA of the same sample revealed the presence of the same nucleotide change at the same level, making us to think to a new SNP, not already reported in databases, or to a germline mutation in the coding sequence of *PRDM1*, rather than a mutation acquired from the tumoral cells. Interestingly, we identified two new mutations, leading to the formation of a stop codon, into two different ALK- samples, one bearing the loss of the other allele, and one affected by copy neutral LOH, suggesting the formation of a truncated BLIMP1 protein in these cases. Unfortunately, for the concerned two cases, the matched normal DNA was not available for further analysis, but we are quite confident about the somatic nature of the

mutations we have found, because of the creation of a stop codon. Interestingly, one of the case in which a new stop codon has been identified, together with the loss of the wild type allele, shows a positivity for BLIMP1 staining only at a cytoplasmic level, which is surprising, being BLIMP1 a transcriptional factor exerting its role in the nucleus. This observation suggests that probably the creation of a stop codon due to a mutation, could lead to the formation of a truncated protein that could lack the nuclear localization sequence. The final effect could be the sequestration of the protein in the cytoplasm. Unfortunately we were not able to investigate BLIMP1 staining with immunohistochemistry for the other case presenting the formation of a new stop codon and the DNA of the other cytoplasmatic positive case was not available. It could be interesting to understand if a correlation among mutations, and consequent creation of truncated protein, and positive staining for BLIMP1 only at cytoplasmic level exists.
Moreover, the sequencing and mutational analysis of *PRDM1* exons of ALCL cell lines, allowed us to observe the presence of a silent nucleotide change, not previously reported, in the ALK+ cell line SUDHL1. In summary, *PRDM1* gene in the whole ALCL series analyzed, is inactivated by deletion or somatic mutation in 39% of ALK- ALCL samples and in only 3% of ALK+ patients. The huge difference of percentage of primary samples displaying a *PRDM1* inactivation between ALK- and ALK+ specimens highlights the idea that BLIMP1 could exert an important role in ALCL pathogenesis, in particular in ALK- cases. Differently from results obtained in DLBCL, we were not able to observe in most of the cases a double inactivation of the gene, tending to favor haploinsufficiency in lieu of a biallelic inactivation.

In order to better understand if BLIMP1 has a role in ALCL pathogenesis, and what type of function it exerts in this process, we thought that re-expression of *PRDM1* gene, through gain-of-function experiments, could help to answer our questions. All ALCL cell lines available underwent the genomic DNA profiling analysis and sequencing and mutational analysis of PRDM1 coding exons. The observation of the same 6q21 loss observed in primary samples, and the low basal levels of BLIMP1 mRNA and protein, induced us to choose TS ALK+ ALCL cell line as model for our functional experiments. The best model would have been an ALK- ALCL cell lines bearing 6q21 loss as observed in many ALK- primary samples, but, unfortunately, real ALK- ALCL cell lines are not available; there are only cell lines established from cALCL cases, which resemble ALK- ALCL, but showing some clinical and biologic differences. Since we observed 6q21 loss

also in ALK+ cases, even if at a very low percentage, we decided to use TS cell lines for our purposes.

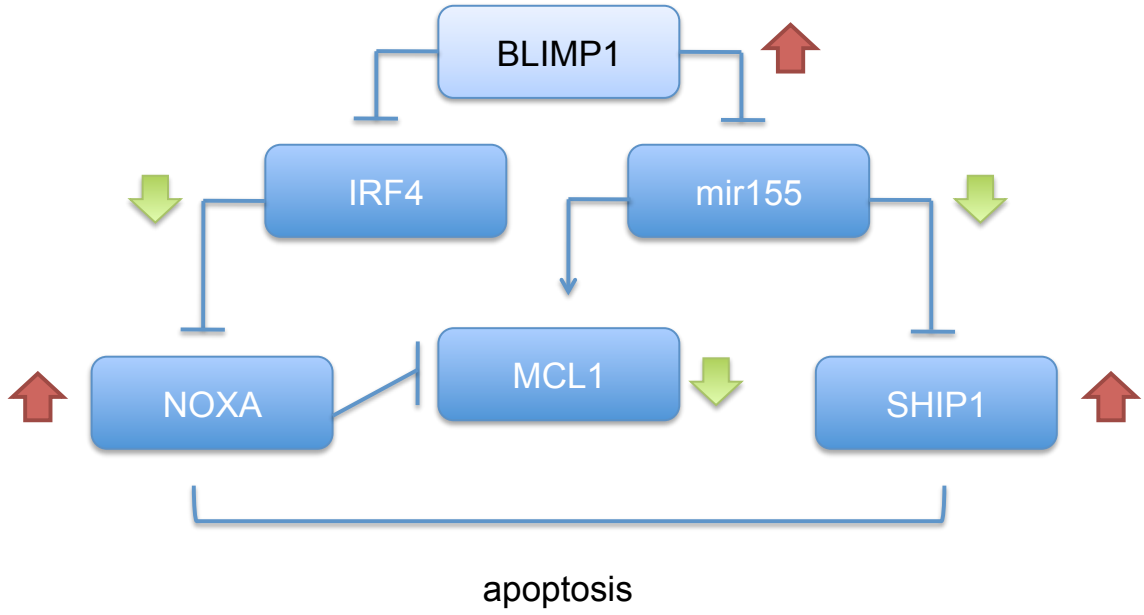
After having evaluated different options for the re-expression of *PRDM1* gene in our cellular model, we decided to use the lentiviral transduction method, which took the advantage of obtaining a stable and efficient transfection. Transfection of *PRDM1* gene in TS cell lines lead to a remarkable decline in term of cell growth and a block in cell proliferation. Concomitantly to this phenotype, cells in which BLIMP1 was reconstituted, started to die in a higher percentage more than control cells, infected with an empty vector. The distribution of cells in different cell cycle phases was affected after re-expression of BLIMP1as well, with a block in G2-M phase. Taken together these results suggested us that BLIMP1 could actually exert a critical role in ALCL pathogenesis, and that *PRDM1* gene can be considered a tumor suppressor gene in our *in vitro* model of ALCL. Since in our primary samples, we observed a specific loss of the locus encompassing *PRDM1* gene, we can speculate that the loss of the tumor suppressor gene *PRDM1* in ALCL cells leads to uncontrolled proliferation of ALCL cells and a block in the apoptotic process. Our results obtained after functional experiments are in line with those obtained in NK-cell lymphoma cell lines after BLIMP1 reconstitution [91, 92]. In NK-lymphoma primary samples, BLIMP1 inactivation as result of genomic loss, nonsense mutations and promoter methylation, has been observed as well. Reconstitution of BLIMP1 in NK-cell line leaded to suppression of cell proliferation, cell cycle arrest and increased apoptosis, just in our case, providing another example of the tumor suppressor role of BLIMP1 in lymphomas.

All the results obtained with gain-of-function experiments in an *in vitro* model of ALCL have been confirmed with in *in vivo* experiments. Cells infected for BLIMP1 re-expression and injected in NSG mice were analyzed after 20 days from inoculum in order to evaluate the percentage of cells infected grown in mice. The low percentage of infected cells in tumors explanted from mice, and the xenograft growth curves of infected cells in NSG mice, confirmed the idea that BLIMP1 reconstituted cells presented growth impairment in *in vivo* experiments as well.

After the identification of a critical role for BLIMP1 in ALCL cells, we wondered how BLIMP1 could exert its functional role. In order to have a global overview of what are the transcriptional changes happening after BLIMP1 reconstitution, we performed a GEP analysis. From the comparison among GEP obtained from cells infected for BLIMP1 re-

expression and GEP obtained from cells infected with an empty vector or not infected at all, we could discriminate genes that were affected by BLIMP1 re-expression. Interestingly, among the genes showing a higher up-regulation after the reconstitution we can recognize different genes involved in growth inhibition process, in cell cycle arrest and in apoptosis promoting processes. These results were in line with what we observed after BLIMP1 reconstitution in *in vitro* experiments. Surprisingly, genes that codes for the different chain of fibrinogen (*FGG, FGB, FGA*) are the ones more affected by the infection for the re-expression of BLIMP1. One explanation for this massive overexpression of fibrinogen genes could be the response of the cell to a stress due to the exogenous expression of a protein (R. Piva, personal communication).

Interestingly, both IRF4 and mir155 downregulation, after BLIMP1 re-expression, seems to be linked to the pro-apoptotic role exerted by BLIMP1. Recently, it has been reported that IRF4 is able to negatively regulate the expression of *PMAIP1* gene [227], which codes for the pro-apoptotic factor NOXA. Both IRF4 downregulation and PMAIP1 overexpression have been observed in our GEP analysis after BLIMP1 reconstitution, and so we think that the overexpression of a pro-apoptotic protein could in part explain the phenotype we observe in our *in vitro* experiments. Moreover, depletion of IRF4 expression has been demonstrated to lead to cell death, with a concomitant increase of the subG1 phase, and to a cell cycle block in a Multiple Myeloma (MM) model [228, 229]. But NOXA overexpression might also cause the inhibition of Mcl-1 function [230], which, on the contrary, is an anti-apoptotic factor. Mcl-1 has been reported to be a direct target of mir155 [231], suggesting us that both upregulation of NOXA and mir155 downregulation could have a role in the regulation of Mcl-1. The observed downregulation of mir155 could also exert a role in the direct upregulation of another pro-apoptotic factor SHIP1 [232]. The observation of increased levels of SHIP1 mRNA after BLIMP1 reconstitution let us to consider another way through which BLIMP1 might exert its pro-apoptotic function. It remains to be established how BLIMP1 acts to repress mir155 activity. Taken together these results might explain, at least in part, the complex network among BLIMP1 reconstitution and its pro-apoptotic effect.



Scheme 2. Proposed model through which BLIMP1 exherts its pro-apoptotic effect. *In a normal situation BLIMP1 is able to repress both IRF4 and mir155. IRF4, in turn, is able to repress the pro-apoptotic factor NOXA. mir155, on the other hand is a direct repressor of SHIP1, another pro.apoptotic factor. NOXA and mir155 can act on the expression on MCL1, a pro-apoptotic factor in an opposite way: NOXA is able to repress MCL1, while mir155 is able to activate MCL1. (Repression is shown through an arrow ending with a line; activation is shown through a simple arrow). In our experiments, after BLIMP1 re-expression, we have an increase of BLIMP1 levels (red arrow), which lead to a decrease of both IRF4 and mir155 (green arrow) , which in turn lead to increased levels of the pro-apoptotic factors NOXA and SHIP1, with a concomitant dowregulation of MCL1 (pro-apoptotic). In this way, BLIMP1 could exhert its pro-apoptotic role in our model.*

Our results suggest that 6q21 loss, observed both in ALK+ and ALK- specimens with higher frequency in the latter subtype, could be an early pathogenetic event. In fact, ALK+ and ALK- subsets shared many genomic aberrations, indicating that the two subsets could have a common origin, but the fact that ALK- samples show a much more complex genome could imply that ALK+ and ALK- specimens could encounter different pathogenetic events. Since that, we thought they could be considered as two different entities, as the 2008 WHO Classification suggests. Our observations indicate that BLIMP1, located in 6q21 locus, is inactivated by loss, hdel and somatic mutation, as previously reported for DLBCL [49, 50] and NK-T cell lymphoma [91, 92]. Experiments reintroducing BLIMP1 in an ALCL cell line presenting the same loss observed in patients, highlight the capacity of this gene to lead to cell proliferation impairment, cell cycle block and apoptosis. Even if BLIMP1 seems to have an important role in ALCL pathogenesis, we assume that it is not the only tumor suppressor gene implicated in the pathogenesis, at least in the ALK- subset. In fact, after the reconstitution of the gene we observe only a moderate increase in the percentage of apoptotic cells, suggesting that some other factors are suspected to act together with BLIMP1. However, re-expression of BLIMP1 in ALCL patients, bearing the loss of the gene, could be of great importance. Nevertheless, the development of therapeutic strategies for cancers associated with mutations in tumor suppressor genes is difficult, since this often required the restoration of the wild type gene function, an undoubtedly challenging task. Unfortunately, at the moment, there are no methods to direct regulate the expression of BLIMP1. In MCL has been reported that the treatment with Bortezomib, a proteasome inhibitor, lead to upregulation of BLIMP1 mRNA levels and the consequent overexpression and activation of the pro-apoptotic protein NOXA with a final initiation of the apoptotic process [233]. The findings that Bortezomib induces PRDM1 expression could provide the rational of evaluation of proteasome inhibitors in ALCL bearing 6q21 loss. Moreover, it could be interesting to evaluate the combination of these drugs with ALK inhibitors and SGN-35, already used for ALCL. Another way to induce BLIMP1 expression, in those patient for whom an haploinsufficient model is considered, might be the block of an upstream repressor of BLIMP1, such as BCL6, but at the moment there are not therapies available able to act in this way.

Appendix

Appendix 1

Transcripts modulated by BLIMP1 and their function

Table 11. Function of the top upregulated genes after BLIMP1 re-expression.

ILMN_GENE	FUNCTION
FGG	Gamma component of fibrinogen, a blood-borne glycoprotein comprised of three pairs of nonidentical polypeptide chains. Following vascular injury, fibrinogen is cleaved by thrombin to form fibrin which is the most abundant component of blood clots. In addition, various cleavage products of fibrinogen and fibrin regulate cell adhesion and spreading, display vasoconstrictor and chemotactic activities, and are mitogens for several cell types.
FGB	Beta component of fibrinogen.
SERPINA3	Plasma protease inhibitor and member of the serine protease inhibitor class. It can inhibit neutrophil cathepsin G and mast cell chymase, both of which can convert angiotensin-1 to the active angiotensin-2.
SLCO2B1	Mediates the Na(+)-independent transport of organic anions such as taurocholate, the prostaglandins PGD2,PGE1, PGE2, leukotriene C4, thromboxane B2 and iloprost.
GAD1	One of several forms of glutamic acid decarboxylase, identified as a major autoantigen in insulin-dependent diabetes. The enzyme encoded is responsible for catalyzing the production of gamma-aminobutyric acid from L-glutamic acid. Catalyzes the production of GABA.
LCE3D	Precursors of the cornified envelope of the stratum corneum.
RASSF2	Potential tumor suppressor. Acts as a KRAS-specific effector protein. May promote apoptosis and cell cycle arrest. Stabilizes STK3/MST2 by protecting it from proteasomal degradation.
SERPINB3	May act as a protease inhibitor to modulate the host immune response against tumor cells.
ATG7	Functions as an E1 enzyme essential for multisubstrates such as ATG8-like proteins and ATG12. Forms intermediate conjugates with ATG8-like proteins (GABARAP, GABARAPL1, GABARAPL2 or MAP1LC3A). PE-conjugation to ATG8-like proteins is essential for autophagy. Also acts as an E1 enzyme for ATG12 conjugation to ATG5 and ATG3.
FGA	Alpha component of fibrinogen.
KHDRBS3	RNA-binding protein that plays a role in the regulation of alternative splicing and influences mRNA splice site selection and exon inclusion. May play a role as a negative regulator of cell growth. Inhibits cell proliferation.
CRISPLD2	Promotes matrix assembly.
ACOX2	Oxidizes the CoA esters of the bile acid intermediates di- and tri-hydroxycholestanoic acids.
SERPINB4	May act as a protease inhibitor to modulate the host immune response against tumor cells.
MS4A4A	May be involved in signal transduction as a component of a multimeric receptor complex.
GPX5	Belongs to the glutathione peroxidase family. Protects cells and enzymes from oxidative damage, by catalyzing the reduction of hydrogen peroxide, lipid peroxides and organic hydroperoxide, by glutathione.
PMAIP1	Promotes activation of caspases and apoptosis. Promotes mitochondrial membrane changes and efflux of apoptogenic proteins from the mitochondria. Contributes to p53/TP53-dependent apoptosis after radiation exposure. Promotes proteasomal degradation of MCL1.
SCARA5	Ferritin receptor that mediates non-transferrin-dependent delivery of iron. Mediates cellular uptake of ferritin-bound iron by stimulating ferritin endocytosis from the cell surface with consequent iron delivery within the cell.
TMEM166	Acts as a regulator of programmed cell death, mediating both autophagy and apoptosis.
PRDM1	Transcriptional repressor that binds specifically to the PRDI element in the promoter of the beta-interferon gene. Drives the maturation of B-lymphocytes into Ig secreting cells.

Table 12. Function of the top downregulated genes after BLIMP1 re-expression.

ILMN_GENE	FUNCTIONS
RGS18	Inhibits signal transduction by increasing the GTPase activity of G protein alpha subunits thereby driving them into their inactive GDP-bound form.
AIM2	Tumor suppressor which may act by repressing NF-kappa-B transcriptional activity.
IRF2BP2	Acts as a transcriptional repressor. Acts as a transcriptional corepressor in aIRF2-dependent manner. This repression is not mediated at least in part by histone deacetylase activities.
KLRC1	Belongs to the killer cell lectin-like receptor family. Plays a role as a receptor for the recognition of MHC class I HLA-E molecules by NK cells and some cytotoxic T-cells.
BCAT1	Cytosolic form of the enzyme branched-chain amino acid transaminase. Catalyzes the first reaction in the catabolism of the essential branched chain amino acids leucine, isoleucine, and valine.
TRIM38	RING finger protein 15.
MGC18216	Receptor tyrosine kinase which mediates actions of insulin-like growth factor 1 (IGF1). Binds IGF1 with high affinity and IGF2 and insulin (INS) with a lower affinity. The activated IGF1R is involved in cell growth and survival control. IGF1R is crucial for tumor transformation and survival of malignant cell. Ligand binding activates the receptor kinase, leading to receptor autophosphorylation, and tyrosines phosphorylation of multiple substrates, that function as signaling adapter proteins including, the insulin-receptor substrates (IRS1/2), Shc and 14-3-3 proteins. Phosphorylation of IRSs proteins lead to the activation of two main signaling pathways: the PI3K-AKT/PKB pathway and the Ras-MAPK pathway. The result of activating the MAPK pathway is increased cellular proliferation, whereas activating the PI3K pathway inhibits apoptosis and stimulates protein synthesis.
BTN3A2	Butyrophilin, subfamily 3, member A2.
GBP1	Guanylate binding protein 1, interferon-inducible. Binds GTP, GDP and GMP.
CNIH	Involved in the selective transport and maturation of TGF-alpha family proteins.
THUMPD2	SAM-dependent methyltransferase.
ZIC4	Binds to DNA.
PCDH18	Potential calcium-dependent cell-adhesion protein.
C5ORF24	Chromosome 5 open reading frame 24.
STARD13	GTPase-activating protein for RhoA, and perhaps for Cdc42. May be involved in regulation of cytoskeletal reorganization, cell proliferation and cell motility. Acts a tumor suppressor in hepatocellular carcinoma cells.
INTS2	Component of the Integrator complex, a complex involved in the small nuclear RNAs (snRNA) U1 and U2 transcription and in their 3'-box-dependent processing. The Integrator complex is associated with the C-terminal domain (CTD) of RNA polymerase II largest subunit (POLR2A) and is recruited to the U1 and U2 snRNAs genes.
CACNA1E	Calcium channel, voltage-dependent, R type, alpha 1E subunit. oltage-sensitive calcium channels (VSCC) mediate the entry of calcium ions into excitable cells and are also involved in a variety of calcium-dependent processes, including muscle contraction, hormone or neurotransmitter release, gene expression, cell motility, cell division and cell death.
TBL1X	F-box-like protein involved in the recruitment of the ubiquitin/19S proteasome complex to nuclear receptor-regulated transcription units. Plays an essential role in transcription activation mediated by nuclear receptors. Probably acts as integral component of corepressor complexes that mediates the recruitment of the 19S proteasome complex, leading to the subsequent proteasomal degradation of transcription repressor complexes, thereby allowing cofactor exchange.
TARDBP	DNA and RNA-binding protein which regulates transcription and splicing.
DNAJC24	Stimulates the ATPase activity of several Hsp70-type chaperones.
TRAF5	TNF receptor-associated factor 5. dapter protein and signal transducer that links members of the tumor necrosis factor receptor family to different signaling pathways by association with the receptor cytoplasmic domain and kinases. Mediates activation of NF-kappa-B and probably JNK. Seems to be involved in apoptosis.
ZC3H14	Binds the polyadenosine RNA oligonucleotides.
DVL3	This gene is a member of a multi-gene family which shares strong similarity with the Drosophila dishevelled gene, dsh. May play a role in the signal transduction pathway mediated by multiple Wnt genes.
LOC100132707	Uncharacterized protein.
GOLIM4	Plays a role in endosome to Golgi protein trafficking; mediates protein transport along the late endosome-bypass pathway from the early endosome to the Golgi.

Appendix 2

Supplement Figures and Tables

Manuscript : *Marchiolatti R, Messana K, Landra I, Abate F, Tabbo' F, Boi M et al. "Massive RNAseq defines novel tumorigenic gene fusions of Anaplastic Large Cell Lymphoma".*

Figure S1

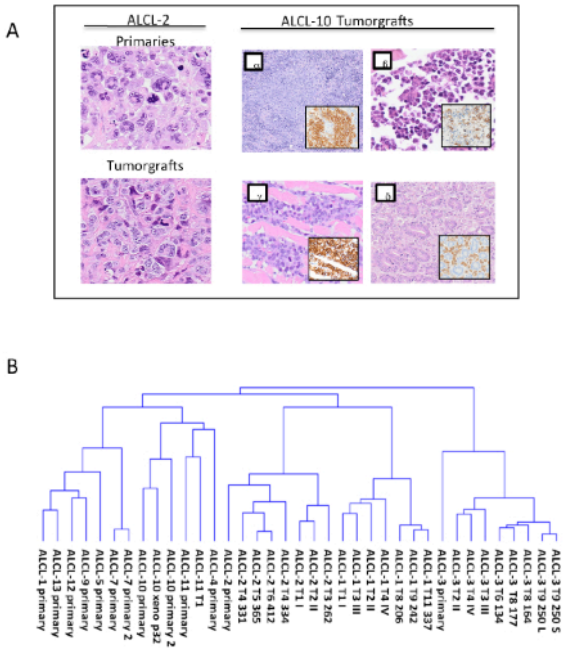


Figure S1. ALCL-PDT display histological and immunophenotypic features resembling primary ALCL lesions and be successfully followed by MRI*. (A) The neoplastic cells of ALCL-PDT display cytological features analogous to those of primary ALCL (ALCL-2, 40x and ALCL-10 40x). CD30+ ALCL-PDT infiltrate an area corresponding to the white pulp consistent with the T-cell zone ([α], 10x), bone marrow ([β], 20x), muscle, ([γ],20x) stomach ([δ] , 20x). Tumor cells were identified (inserts) by IHC and anti-CD30 (α,β,γ) or anti-ALK antibodies (δ). (B) GEP shows a close relationship between primary lesions and corresponding PDT. Unrelated primary ALCL clustered separately independently of ALK positivity. Samples are hierarchically clustered with maximum linkage based on cosine correlation.*

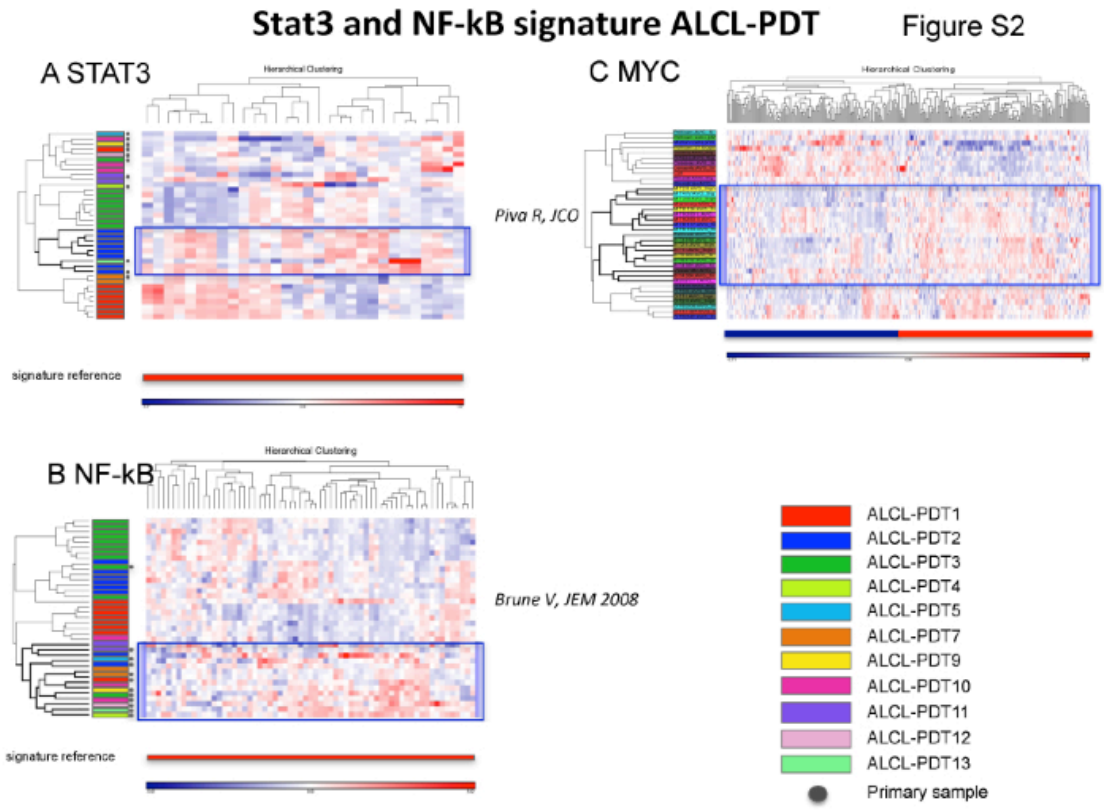


Figure S2. Hierarchical clustering of primary samples and tumorgrafts *according to the expression of 40 probe sets (specific for 21 genes) regulated by STAT3, of 91 probe sets (specific for 41 genes) regulated by activated NF-κB and of 247 probe sets (specific for 142 genes) regulated by MYC.*

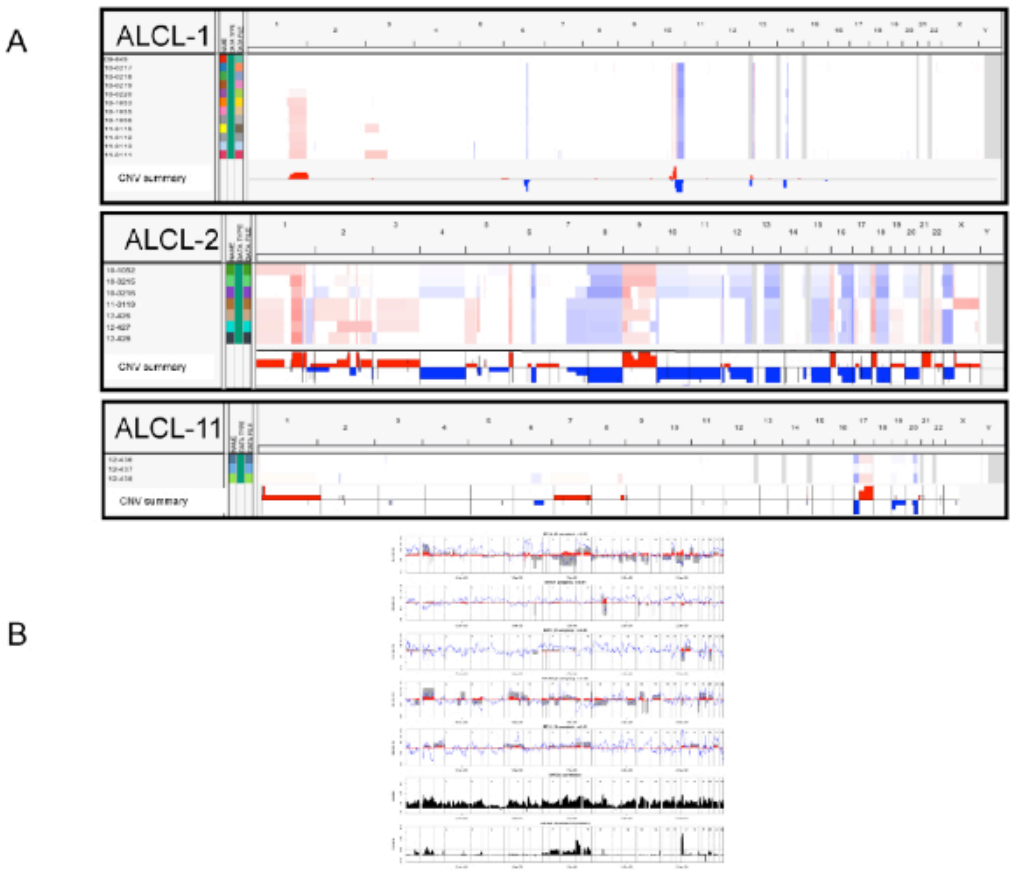


Figure S3. ALCL-PDT share a set of genomic lesions with their corresponding primary lymphomas. *(A) Chromosomal defects in primary and ALCL-PDT samples. Frequency plot of gains (red) and losses (blue) in primary samples (first lane in each ALCL panel) by SNP 6.0 array. CNV summary (last lane: X-axis, chromosome localization and physical mapping; Y-axis, percentage of cases bearing the aberration). (B) Median DNA/RNA correlation (in red). Y-axis: covariance; X-axis: genomic positions across the 22 non-sex chromosomes.*

Table S1. Representative genes in PTDs. *The status of some genes is reported for the primary sample (in red) of ALCL-2 and subsequent tumorgrafts. AMPL: amplification, WT wild type.*

TABLE S1 ALCL.2: Genomic configuration of selective genes in primary and corresponding PDTs

	ALCL-2	PRIMARIES 10-1052 6000/06	ALCL-2 PDT2	ALCL-2 PDT3	ALCL-2 PDT6
JAK2		AMPL	AMPL	AMPL	AMPL
NF1		BREAKPOINT	BREAKPOINT	BREAKPOINT	BREAKPOINT
RC3H2		GAIN	GAIN	GAIN	GAIN
BTBD7		WT	WT	WT	WT
SMAD6		LOSS	LOSS	LOSS	LOSS
CTR9		LOSS	LOSS	LOSS	WT
ZNF641		LOSS	LOSS	LOSS	WT
ZNF358		WT	WT	WT	WT
HSPA12B		LOSS	LOSS	LOSS	LOSS
SLITRK3		WT	WT	WT	WT
RXFP1		WT	WT	LOSS	WT
HEY2		WT	GAIN	WT	WT
ZNF271		LOSS	LOSS	LOSS	LOSS
IGFALS		GAIN	GAIN	GAIN	GAIN
GPR61		GAIN	GAIN	WT	WT
PAX2		LOSS	LOSS	LOSS	LOSS
WT1		WT	LOSS	LOSS	WT
EZH2		WT	WT	LOSS	WT
BCL11B		LOSS	LOSS	LOSS	LOSS
PTEN		WT	LOSS	LOSS	WT
HRAS		WT	LOSS	LOSS	WT
TP53		LOSS	LOSS	LOSS	LOSS
NOTCH 1		WT	GAIN	GAIN	GAIN

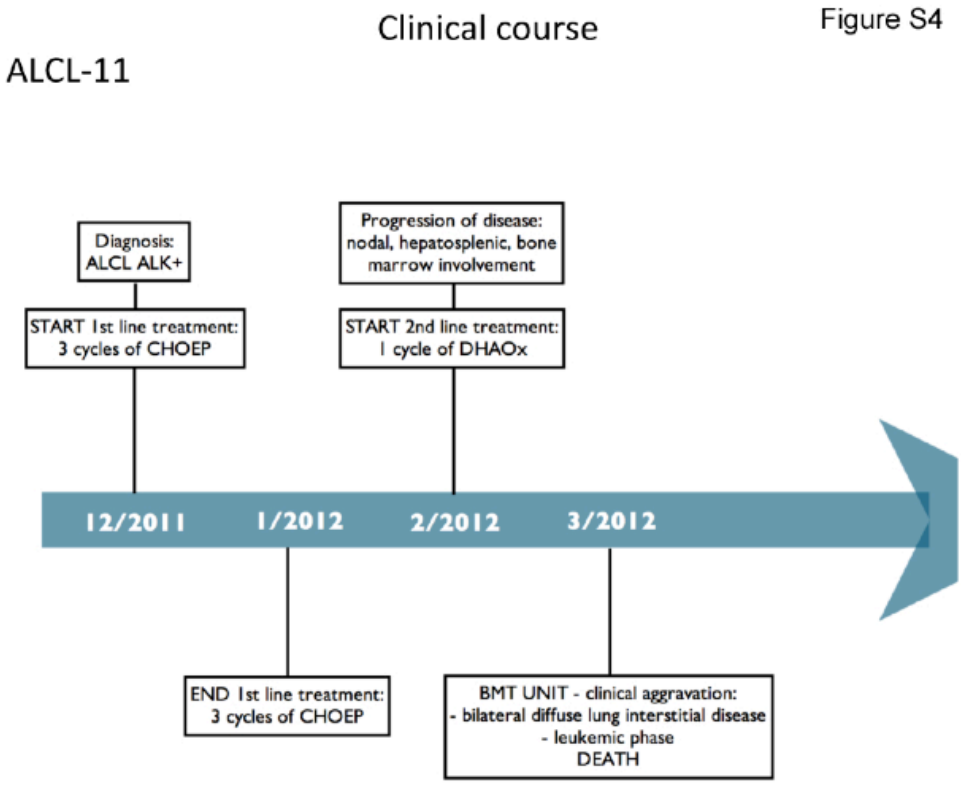


Figure S4. ALK– ALCL-11 clinical course. *Schematic representation of most relevant clinical information and therapeutic modalities.*

Table S2. Genomic lesions detected in the primary sample and ALCL10-PTDs. *Acquired and lost lesions detected in subsequent passages of tumorgrafts are reported.*

Table S2 ALCL-11: Genomic aberration in primary and corresponding PDTs

ALCL-11			
Genomic lesions			
	Primary Sample		
Samples	Primary lesion	Acquired lesion	Lost lesion
Primary sample	+1p36.21-p36.23; +2p11.2; -3p21.1; +8q24.21 -14q11.2; +14q32.32; -17p; +17q; -19; -20q11.23-13.13		
Post-therapy tissue			+1p36.21-p36.23; +2p11.2; -3p21.1; +8q24.21 -14q11.2; +14q32.32; -19;
T1		+1; -2q11.2; -6q13-q22.32; +7p22.3-q34 +8q24.12-q24.21	+1p36.21-p36.23; +2p11.2; -3p21.1; -19q

Bibliography

1. Swerdlow, S.H., et al., eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* 2008, IARC Press: Lyon, France. 439.
2. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
3. Collisson, E.A., R.J. Cho, and J.W. Gray, *What are we learning from the cancer genome?* Nat. Rev. Clin. Oncol., 2012.9(11): p. 621-30
4. Knudson, A.G., *Two genetic hits (more or less) to cancer.* Nat. Rev. Cancer, 2001. **1**(2): p. 157-62.
5. Frohling, S. and H. Dohner, *Chromosomal abnormalities in cancer.* N. Engl. J. Med., 2008. **359**(7): p. 722-34.
6. Balmain, A., *Cancer genetics: from Boveri and Mendel to microarrays.* Nat. Rev. Cancer, 2001. **1**(1): p. 77-82.
7. Nambiar, M., et al., *Amplification of chromosomal translocation junctions from paraffin-embedded tissues of follicular lymphoma patients.* Biomed Mater, 2008. **3**(3): p. 034103.
8. Rowley, J.D., *Chromosome translocations: dangerous liaisons revisited.* Nat. Rev. Cancer, 2001. **1**(3): p. 245-50.
9. Fodde, R. and R. Smits, *Cancer biology. A matter of dosage.* Science, 2002. **298**(5594): p. 761-3.
10. Gajduskova, P., et al., *Genome position and gene amplification.* Genome Biol., 2007. **8**(6): p. R120.
11. Cavenee, W.K., et al., *Expression of recessive alleles by chromosomal mechanisms in retinoblastoma.* Nature, 1983. **305**(5937): p. 779-84.
12. Kralovics, R., et al., *A gain-of-function mutation of JAK2 in myeloproliferative disorders.* N. Engl. J. Med., 2005. **352**(17): p. 1779-90.
13. Tuna, M., S. Knuutila, and G.B. Mills, *Uniparental disomy in cancer.* Trends Mol. Med., 2009. **15**(3): p. 120-8.
14. O'Keefe, C., M.A. McDevitt, and J.P. Maciejewski, *Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies.* Blood, 2010. **115**(14): p. 2731-9.
15. Stephens, P.J., et al., *Massive genomic rearrangement acquired in a single catastrophic event during cancer development.* Cell, 2011. **144**(1): p. 27-40.
16. Meyerson, M. and D. Pellman, *Cancer genomes evolve by pulverizing single chromosomes.* Cell, 2011. **144**(1): p. 9-10.
17. Cassidy, L.D. and A.R. Venkitaraman, *Genome instability mechanisms and the structure of cancer genomes.* Curr. Opin. Genet. Dev.., 2012. **22**(1): p. 10-3.
18. Eifert, C. and R.S. Powers, *From cancer genomes to oncogenic drivers, tumour dependencies and therapeutic targets.* Nat. Rev. Cancer., 2012.12(8): p. 572-8.
19. Feinberg, A.P. and B. Tycko, *The history of cancer epigenetics.* Nat. Rev. Cancer, 2004. **4**(2): p. 143-53.

20. Esteller, M., *Epigenetics in cancer.* N. Engl. J. Med., 2008. **358**(11): p. 1148-59.
21. Tysnes, B.B., *Tumor-initiating and -propagating cells: cells that we would like to identify and control.* Neoplasia, 2010. **12**(7): p. 506-15.
22. Hatziapostolou, M. and D. Iliopoulos, *Epigenetic aberrations during oncogenesis.* Cell Mol. Life Sci., 2011. **68**(10): p. 1681-702.
23. Smeets, D.F., *Historical prospective of human cytogenetics: from microscope to microarray.* Clin. Biochem., 2004. **37**(6): p. 439-46.
24. Trask, B.J., *Human cytogenetics: 46 chromosomes, 46 years and counting.* Nat. Rev. Genet., 2002. **3**(10): p. 769-78.
25. Murthy, S.K. and D.J. Demetrick, *New approaches to fluorescence in situ hybridization.* Methods Mol. Biol., 2006. **319**: p. 237-59.
26. Ventura, R.A., et al., *FISH analysis for the detection of lymphoma-associated chromosomal abnormalities in routine paraffin-embedded tissue.* J. Mol. Diagn., 2006. **8**(2): p. 141-51.
27. Schrock, E., et al., *Multicolor spectral karyotyping of human chromosomes.* Science, 1996. **273**(5274): p. 494-7.
28. Kallioniemi, A., et al., *Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.* Science, 1992. **258**(5083): p. 818-21.
29. Pinkel, D., et al., *High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.* Nat. Genet., 1998. **20**(2): p. 207-11.
30. Rubio-Moscardo, F., et al., *Mantle-cell lymphoma genotypes identified with CGH to BAC microarrays define a leukemic subgroup of disease and predict patient outcome.* Blood, 2005. **105**(11): p. 4445-54.
31. Kohlhammer, H., et al., *Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions.* Blood, 2004. **104**(3): p. 795-801.
32. Oostlander, A.E., G.A. Meijer, and B. Ylstra, *Microarray-based comparative genomic hybridization and its applications in human genetics.* Clin. Genet., 2004. **66**(6): p. 488-95.
33. Lockwood, W.W., et al., *Recent advances in array comparative genomic hybridization technologies and their applications in human genetics.* Eur. J. Hum. Genet., 2006. **14**(2): p. 139-48.
34. Liang, P. and A.B. Pardee, *Analysing differential gene expression in cancer.* Nat. Rev. Cancer, 2003. **3**(11): p. 869-76.
35. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins.* J. Biol. Chem., 1975. **250**(10): p. 4007-21.
36. Sargent, T.D., *Isolation of differentially expressed genes.* Methods Enzymol., 1987. **152**: p. 423-32.
37. Zimmermann, C.R., et al., *Molecular cloning and selection of genes regulated in Aspergillus development.* Cell, 1980. **21**(3): p. 709-15.

38. Liang, P. and A.B. Pardee, *Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction.* Science, 1992. **257**(5072): p. 967-71.
39. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-70.
40. Chee, M., et al., *Accessing genetic information with high-density DNA arrays.* Science, 1996. **274**(5287): p. 610-4.
41. Orsborne, C. and R. Byers, *Impact of gene expression profiling in lymphoma diagnosis and prognosis.* Histopathology, 2011. **58**(1): p. 106-27.
42. Kearney, L. and S.W. Horsley, *Molecular cytogenetics in haematological malignancy: current technology and future prospects.* Chromosoma, 2005. **114**(4): p. 286-94.
43. Hartmann, E.M., G. Ott, and A. Rosenwald, *Molecular biology and genetics of lymphomas.* Hematol. Oncol. Clin. North Am., 2008. **22**(5): p. 807-23, vii.
44. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503-11.
45. Rosenwald, A., et al., *The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma.* N. Engl. J. Med., 2002. **346**(25): p. 1937-47.
46. Tagawa, H., et al., *Comparison of genome profiles for identification of distinct subgroups of diffuse large B-cell lymphoma.* Blood, 2005. **106**(5): p. 1770-7.
47. Chen, W., et al., *Array comparative genomic hybridization reveals genomic copy number changes associated with outcome in diffuse large B-cell lymphomas.* Blood, 2006. **107**(6): p. 2477-85.
48. 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, 2005. **106**(9): p. 3183-90.
49. Pasqualucci, L., et al., *Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma.* J. Exp. Med., 2006. **203**(2): p. 311-7.
50. Mandelbaum, J., et al., *BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma.* Cancer Cell, 2010. **18**(6): p. 568-79.
51. Lenz, G., et al., *Oncogenic CARD11 mutations in human diffuse large B cell lymphoma.* Science, 2008. **319**(5870): p. 1676-9.
52. Offit, K., et al., *Rearrangement of the bcl-6 gene as a prognostic marker in diffuse large-cell lymphoma.* N. Engl. J. Med., 1994. **331**(2): p. 74-80.
53. Weiss, L.M., et al., *Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas.* N. Engl. J. Med., 1987. **317**(19): p. 1185-9.

54. Yunis, J.J., et al., *bcl-2 and other genomic alterations in the prognosis of large-cell lymphoma.* N. Engl. J. Med., 1989. **320**(16): p. 1047-54.
55. Hummel, M., et al., *A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling.* N. Engl. J. Med., 2006. **354**(23): p. 2419-30.
56. Garcia, J.L., et al., *Abnormalities on 1q and 7q are associated with poor outcome in sporadic Burkitt's lymphoma. A cytogenetic and comparative genomic hybridization study.* Leukemia, 2003. **17**(10): p. 2016-24.
57. Kanungo, A., et al., *Lymphoid neoplasms associated with concurrent t(14;18) and 8q24/c-MYC translocation generally have a poor prognosis.* Mod. Pathol., 2006. **19**(1): p. 25-33.
58. Dave, S.S., et al., *Molecular diagnosis of Burkitt's lymphoma.* N. Engl. J. Med., 2006. **354**(23): p. 2431-42.
59. Lossos, I.S., *Molecular pathogenesis of diffuse large B-cell lymphoma.* J. Clin. Oncol., 2005. **23**(26): p. 6351-7.
60. Barrans, S., et al., *Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab.* J. Clin. Oncol., 2010. **28**(20): p. 3360-5.
61. Horning, S.J., *Follicular lymphoma: have we made any progress?* Ann. Oncol., 2000. **11 Suppl 1**: p. 23-7.
62. McDonnell, T.J., et al., *bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation.* Cell, 1989. **57**(1): p. 79-88.
63. Liu, Y., et al., *BCL2 translocation frequency rises with age in humans.* Proc. Natl. Acad. Sci. U S A, 1994. **91**(19): p. 8910-4.
64. Viardot, A., et al., *Cytogenetic evolution of follicular lymphoma.* Semin. Cancer. Biol., 2003. **13**(3): p. 183-90.
65. Viardot, A., et al., *Clinicopathologic correlations of genomic gains and losses in follicular lymphoma.* J. Clin. Oncol., 2002. **20**(23): p. 4523-30.
66. Dave, S.S., et al., *Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells.* N. Engl. J. Med., 2004. **351**(21): p. 2159-69.
67. Li, J.Y., et al., *Detection of translocation t(11;14)(q13;q32) in mantle cell lymphoma by fluorescence in situ hybridization.* Am. J. Pathol., 1999. **154**(5): p. 1449-52.
68. Swerdlow, S.H. and M.E. Williams, *From centrocytic to mantle cell lymphoma: a clinicopathologic and molecular review of 3 decades.* Hum. Pathol., 2002. **33**(1): p. 7-20.
69. Salaverria, I., et al., *Specific secondary genetic alterations in mantle cell lymphoma provide prognostic information independent of the gene expression-based proliferation signature.* J. Clin. Oncol., 2007. **25**(10): p. 1216-22.
70. Fu, K., et al., *Cyclin D1-negative mantle cell lymphoma: a clinicopathologic study based on gene expression profiling.* Blood, 2005. **106**(13): p. 4315-21.

71. Gesk, S., et al., *A chromosomal translocation in cyclin D1-negative/cyclin D2-positive mantle cell lymphoma fuses the CCND2 gene to the IGK locus.* Blood, 2006. **108**(3): p. 1109-10.
72. Rinaldi, A., et al., *Genome-wide DNA profiling of marginal zone lymphomas identifies subtype-specific lesions with an impact on the clinical outcome.* Blood, 2011. **117**(5): p. 1595-604.
73. Dohner, H., et al., *Genomic aberrations and survival in chronic lymphocytic leukemia.* N. Engl. J. Med., 2000. **343**(26): p. 1910-6.
74. Kay, N.E., et al., *Chronic lymphocytic leukemia.* Hematology Am. Soc. Hematol. Educ. Program, 2002: p. 193-213.
75. Rosenwald, A., et al., *Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia.* J. Exp. Med., 2001. **194**(11): p. 1639-47.
76. Rinaldi, A., et al., *Genome-wide DNA profiling better defines the prognosis of chronic lymphocytic leukaemia.* Br. J. Haematol., 2011. **154**(5): p. 590-9.
77. O'Leary, H.M. and K.J. Savage, *Update on the World Health Organization classification of peripheral T-cell lymphomas.* Curr. Hematol. Malig. Rep., 2009. **4**(4): p. 227-35.
78. de Leval, L., et al., *The gene expression profile of nodal peripheral T-cell lymphoma demonstrates a molecular link between angioimmunoblastic T-cell lymphoma (AITL) and follicular helper T (TFH) cells.* Blood, 2007. **109**(11): p. 4952-63.
79. Boi, M., et al., *Genetic alterations in systemic nodal and extranodal non-cutaneous lymphomas derived from mature T cells and natural killer cells.* Cancer Sci., 2012. **103**(8): p. 1397-404.
80. Zettl, A., et al., *Genomic profiling of peripheral T-cell lymphoma, unspecified, and anaplastic large T-cell lymphoma delineates novel recurrent chromosomal alterations.* Am. J. Pathol., 2004. **164**(5): p. 1837-48.
81. Fujiwara, S.I., et al., *High-resolution analysis of chromosome copy number alterations in angioimmunoblastic T-cell lymphoma and peripheral T-cell lymphoma, unspecified, with single nucleotide polymorphism-typing microarrays.* Leukemia, 2008. **22**(10): p. 1891-8.
82. Streubel, B., et al., *Novel t(5;9)(q33;q22) fuses ITK to SYK in unspecified peripheral T-cell lymphoma.* Leukemia, 2006. **20**(2): p. 313-8.
83. Piccaluga, P.P., et al., *Gene expression analysis of peripheral T cell lymphoma, unspecified, reveals distinct profiles and new potential therapeutic targets.* J. Clin. Invest., 2007. **117**(3): p. 823-34.
84. Lepretre, S., et al., *Chromosome abnormalities in peripheral T-cell lymphoma.* Cancer Genet. Cytogenet., 2000. **117**(1): p. 71-9.
85. Piccaluga, P.P., et al., *Gene expression analysis of angioimmunoblastic lymphoma indicates derivation from T follicular helper cells and vascular endothelial growth factor deregulation.* Cancer Res., 2007. **67**(22): p. 10703-10.

86. Iqbal, J., et al., *Molecular signatures to improve diagnosis in peripheral T-cell lymphoma and prognostication in angioimmunoblastic T-cell lymphoma.* Blood, 2010. **115**(5): p. 1026-36.
87. Zhao, W.L., et al., *Vascular endothelial growth factor-A is expressed both on lymphoma cells and endothelial cells in angioimmunoblastic T-cell lymphoma and related to lymphoma progression.* Lab. Invest., 2004. **84**(11): p. 1512-9.
88. Siu, L.L., et al., *Consistent patterns of allelic loss in natural killer cell lymphoma.* Am. J. Pathol., 2000. **157**(6): p. 1803-9.
89. Nakashima, Y., et al., *Genome-wide array-based comparative genomic hybridization of natural killer cell lymphoma/leukemia: different genomic alteration patterns of aggressive NK-cell leukemia and extranodal Nk/T-cell lymphoma, nasal type.* Genes Chromosomes Cancer, 2005. **44**(3): p. 247-55.
90. Iqbal, J., et al., *Genomic analyses reveal global functional alterations that promote tumor growth and novel tumor suppressor genes in natural killer-cell malignancies.* Leukemia, 2009. **23**(6): p. 1139-51.
91. Karube, K., et al., *Identification of FOXO3 and PRDM1 as tumor-suppressor gene candidates in NK-cell neoplasms by genomic and functional analyses.* Blood, 2011. **118**(12): p. 3195-204.
92. Kucuk, C., et al., *PRDM1 is a tumor suppressor gene in natural killer cell malignancies.* Proc. Natl. Acad. Sci. U S A, 2011. **108**(50): p. 20119-24.
93. Huang, Y., et al., *Gene expression profiling identifies emerging oncogenic pathways operating in extranodal NK/T-cell lymphoma, nasal type.* Blood, 2010. **115**(6): p. 1226-37.
94. Ng, S.B., et al., *Dysregulated microRNAs affect pathways and targets of biologic relevance in nasal-type natural killer/T-cell lymphoma.* Blood, 2011. **118**(18): p. 4919-29.
95. Deleeuw, R.J., et al., *Whole-genome analysis and HLA genotyping of enteropathy-type T-cell lymphoma reveals 2 distinct lymphoma subtypes.* Gastroenterology, 2007. **132**(5): p. 1902-11.
96. Ko, Y.H., et al., *Enteropathy-associated T-cell lymphoma--a clinicopathologic and array comparative genomic hybridization study.* Hum. Pathol., 2010. **41**(9): p. 1231-7.
97. Cejkova, P., et al., *Amplification of NOTCH1 and ABL1 gene loci is a frequent aberration in enteropathy-type T-cell lymphoma.* Virchows Arch., 2005. **446**(4): p. 416-20.
98. Alonsozana, E.L., et al., *Isochromosome 7q: the primary cytogenetic abnormality in hepatosplenic gammadelta T cell lymphoma.* Leukemia, 1997. **11**(8): p. 1367-72.
99. Travert, M., et al., *Molecular features of hepatosplenic T-cell lymphoma unravels potential novel therapeutic targets.* Blood, 2012. **119**(24): p. 5795-806.

100. Hahtola, S., et al., *Clinicopathological characterization and genomic aberrations in subcutaneous panniculitis-like T-cell lymphoma.* J. Invest Dermatol., 2008. **128**(9): p. 2304-9..
101. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.* J. Mol. Biol., 1975. **94**(3): p. 441-8.
102. Reis-Filho, J.S., *Next-generation sequencing.* Breast Cancer Res., 2009. **11 Suppl 3**: p. S12.
103. Pareek, C.S., R. Smoczynski, and A. Tretyn, *Sequencing technologies and genome sequencing.* J. Appl. Genet., 2011. **52**(4): p. 413-35.
104. Schadt, E.E., S. Turner, and A. Kasarskis, *A window into third-generation sequencing.* Hum. Mol. Genet., 2010. **19**(R2): p. R227-40.
105. Stein, H., et al., *The expression of the Hodgkin's disease associated antigen Ki-1 in reactive and neoplastic lymphoid tissue: evidence that Reed-Sternberg cells and histiocytic malignancies are derived from activated lymphoid cells.* Blood, 1985. **66**(4): p. 848-58.
106. Schwab, U., et al., *Production of a monoclonal antibody specific for Hodgkin and Sternberg-Reed cells of Hodgkin's disease and a subset of normal lymphoid cells.* Nature, 1982. **299**(5878): p. 65-7.
107. Bitter, M.A., et al., *Morphology in Ki-1(CD30)-positive non-Hodgkin's lymphoma is correlated with clinical features and the presence of a unique chromosomal abnormality, t(2;5)(p23;q35).* Am. J. Surg. Pathol., 1990. **14**(4): p. 305-16.
108. Kaneko, Y., et al., *A novel translocation, t(2;5)(p23;q35), in childhood phagocytic large T-cell lymphoma mimicking malignant histiocytosis.* Blood, 1989. **73**(3): p. 806-13.
109. Le Beau, M.M., et al., *The t(2;5)(p23;q35): a recurring chromosomal abnormality in Ki-1-positive anaplastic large cell lymphoma.* Leukemia, 1989. **3**(12): p. 866-70.
110. Morris, S.W., et al., *Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma.* Science, 1994. **263**(5151): p. 1281-4.
111. Skinnider, B.F., et al., *Anaplastic large cell lymphoma: a clinicopathologic analysis.* Hematol. Oncol., 1999. **17**(4): p. 137-48.
112. Kinney, M.C., R.A. Higgins, and E.A. Medina, *Anaplastic large cell lymphoma: twenty-five years of discovery.* Arch. Pathol. Lab. Med., 2011. **135**(1): p. 19-43.
113. Lechner, M.G., et al., *Survival Signals and Targets for Therapy in Breast Implant-Associated ALK- Anaplastic Large Cell Lymphoma.* Clin. Cancer Res., 2012.18(17): p. 4549-59.
114. Falini, B., et al., *ALK+ lymphoma: clinico-pathological findings and outcome.* Blood, 1999. **93**(8): p. 2697-706.
115. Gascoyne, R.D., et al., *Prognostic significance of anaplastic lymphoma kinase (ALK) protein expression in adults with anaplastic large cell lymphoma.* Blood, 1999. **93**(11): p. 3913-21.

116. Savage, K.J., et al., *ALK- anaplastic large-cell lymphoma is clinically and immunophenotypically different from both ALK+ ALCL and peripheral T-cell lymphoma, not otherwise specified: report from the International Peripheral T-Cell Lymphoma Project.* Blood, 2008. **111**(12): p. 5496-504.
117. Piccaluga, P.P., et al., *Pathobiology of anaplastic large cell lymphoma.* Adv. Hematol., 2010: p. 345053.
118. Huang, W., et al., *Expression of ALK protein, mRNA and fusion transcripts in anaplastic large cell lymphoma.* Exp. Mol. Pathol., 2009. **86**(2): p. 121-6.
119. Medeiros, L.J. and K.S. Elenitoba-Johnson, *Anaplastic Large Cell Lymphoma.* Am. J. Clin. Pathol., 2007. **127**(5): p. 707-22.
120. Stein, H., et al., *CD30(+) anaplastic large cell lymphoma: a review of its histopathologic, genetic, and clinical features.* Blood, 2000. **96**(12): p. 3681-95.
121. Falini, B., et al., *CD30 (Ki-1) molecule: a new cytokine receptor of the tumor necrosis factor receptor superfamily as a tool for diagnosis and immunotherapy.* Blood, 1995. **85**(1): p. 1-14.
122. Wright, C.W., J.M. Rumble, and C.S. Duckett, *CD30 activates both the canonical and alternative NF-kappaB pathways in anaplastic large cell lymphoma cells.* J. Biol. Chem., 2007. **282**(14): p. 10252-62.
123. Feldman, A.L., et al., *PAX5-positive T-cell anaplastic large cell lymphomas associated with extra copies of the PAX5 gene locus.* Mod. Pathol., 2010. **23**(4): p. 593-602.
124. Piva, R., et al., *Gene expression profiling uncovers molecular classifiers for the recognition of anaplastic large-cell lymphoma within peripheral T-cell neoplasms.* J. Clin. Oncol., 2010. **28**(9): p. 1583-90.
125. Thompson, M.A., et al., *Differential gene expression in anaplastic lymphoma kinase-positive and anaplastic lymphoma kinase-negative anaplastic large cell lymphomas.* Hum. Pathol., 2005. **36**(5): p. 494-504.
126. Lamant, L., et al., *Gene-expression profiling of systemic anaplastic large-cell lymphoma reveals differences based on ALK status and two distinct morphologic ALK+ subtypes.* Blood, 2007. **109**(5): p. 2156-64.
127. Merkel, O., et al., *Identification of differential and functionally active miRNAs in both anaplastic lymphoma kinase (ALK)+ and ALK- anaplastic large-cell lymphoma.* Proc. Natl. Acad. Sci. U S A, 2010. **107**(37): p. 16228-33.
128. Webb, T.R., et al., *Anaplastic lymphoma kinase: role in cancer pathogenesis and small-molecule inhibitor development for therapy.* Expert Rev. Anticancer Ther., 2009. **9**(3): p. 331-56.
129. Stoica, G.E., et al., *Identification of anaplastic lymphoma kinase as a receptor for the growth factor pleiotrophin.* J. Biol. Chem., 2001. **276**(20): p. 16772-9.
130. Stoica, G.E., et al., *Midkine binds to anaplastic lymphoma kinase (ALK) and acts as a growth factor for different cell types.* J. Biol. Chem,. 2002. **277**(39): p. 35990-8.

131. Shi, X.L., X.W. Tang, and D.P. Wu, *Research progresses in the pathogenesis of anaplastic large cell lymphoma.* Chin. J. Cancer., 2011. **30**(6): p. 392-9.
132. Bowden, E.T., G.E. Stoica, and A. Wellstein, *Anti-apoptotic signaling of pleiotrophin through its receptor, anaplastic lymphoma kinase.* J. Biol. Chem., 2002. **277**(39): p. 35862-8.
133. Mourali, J., et al., *Anaplastic lymphoma kinase is a dependence receptor whose proapoptotic functions are activated by caspase cleavage.* Mol. Cell Biol., 2006. **26**(16): p. 6209-22.
134. Lu, K.V., et al., *Differential induction of glioblastoma migration and growth by two forms of pleiotrophin.* J. Biol. Chem., 2005. **280**(29): p. 26953-64.
135. Delsol, G., et al., *A new subtype of large B-cell lymphoma expressing the ALK kinase and lacking the 2; 5 translocation.* Blood, 1997. **89**(5): p. 1483-90.
136. Chiarle, R., et al., *The anaplastic lymphoma kinase in the pathogenesis of cancer.* Nat. Rev. Cancer., 2008. **8**(1): p. 11-23.
137. Li, R. and S.W. Morris, *Development of anaplastic lymphoma kinase (ALK) small-molecule inhibitors for cancer therapy.* Med. Res. Rev., 2008. **28**(3): p. 372-412.
138. Lamant, L., et al., *A new fusion gene TPM3-ALK in anaplastic large cell lymphoma created by a (1;2)(q25;p23) translocation.* Blood, 1999. **93**(9): p. 3088-95.
139. Hernandez, L., et al., *TRK-fused gene (TFG) is a new partner of ALK in anaplastic large cell lymphoma producing two structurally different TFG-ALK translocations.* Blood, 1999. **94**(9): p. 3265-8.
140. Touriol, C., et al., *Further demonstration of the diversity of chromosomal changes involving 2p23 in ALK-positive lymphoma: 2 cases expressing ALK kinase fused to CLTCL (clathrin chain polypeptide-like).* Blood, 2000. **95**(10): p. 3204-7.
141. Colleoni, G.W., et al., *ATIC-ALK: A novel variant ALK gene fusion in anaplastic large cell lymphoma resulting from the recurrent cryptic chromosomal inversion, inv(2)(p23q35).* Am. J. Pathol., 2000. **156**(3): p. 781-9.
142. Tort, F., et al., *Molecular characterization of a new ALK translocation involving moesin (MSN-ALK) in anaplastic large cell lymphoma.* Lab. Invest., 2001. **81**(3): p. 419-26.
143. Ma, Z., et al., *Fusion of ALK to the Ran-binding protein 2 (RANBP2) gene in inflammatory myofibroblastic tumor.* Genes Chromosomes Cancer, 2003. **37**(1): p. 98-105.
144. Soda, M., et al., *Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer.* Nature, 2007. **448**(7153): p. 561-6.
145. Cools, J., et al., *Identification of novel fusion partners of ALK, the anaplastic lymphoma kinase, in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor.* Genes Chromosomes Cancer, 2002. **34**(4): p. 354-62.

146. Lamant, L., et al., *Non-muscle myosin heavy chain (MYH9): a new partner fused to ALK in anaplastic large cell lymphoma.* Genes Chromosomes Cancer, 2003. **37**(4): p. 427-32.
147. Takeuchi, K., et al., *KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer.* Clin. Cancer. Res., 2009. **15**(9): p. 3143-9.
148. Debelenko, L.V., et al., *Identification of CARS-ALK fusion in primary and metastatic lesions of an inflammatory myofibroblastic tumor.* Lab. Invest., 2003. **83**(9): p. 1255-65.
149. Panagopoulos, I., et al., *Fusion of the SEC31L1 and ALK genes in an inflammatory myofibroblastic tumor.* Int .J. Cancer, 2006. **118**(5): p. 1181-6.
150. Maes, B., et al., *The NPM-ALK and the ATIC-ALK fusion genes can be detected in non-neoplastic cells.* Am. J. Pathol., 2001. **158**(6): p. 2185-93.
151. Ambrogio, C., et al., *The anaplastic lymphoma kinase controls cell shape and growth of anaplastic large cell lymphoma through Cdc42 activation.* Cancer Res., 2008. **68**(21): p. 8899-907.
152. Salaverria, I., et al., *Genomic profiling reveals different genetic aberrations in systemic ALK-positive and ALK-negative anaplastic large cell lymphomas.* Br. J. Haematol., 2008. **140**(5): p. 516-26.
153. Youssif, C., et al., *Genomic profiling of pediatric ALK-positive anaplastic large cell lymphoma: a Children's Cancer and Leukaemia Group Study.* Genes Chromosomes Cancer, 2009. **48**(11): p. 1018-26.
154. Ott, G., et al., *Chromosomal abnormalities in nodal and extranodal CD30+ anaplastic large cell lymphomas: infrequent detection of the t(2;5) in extranodal lymphomas.* Genes Chromosomes Cancer, 1998. **22**(2): p. 114-21.
155. Tsukasaki, K., et al., *Comparative genomic hybridization analysis in adult T-cell leukemia/lymphoma: correlation with clinical course.* Blood, 2001. **97**(12): p. 3875-81.
156. Ferreri, A.J., et al., *Anaplastic large cell lymphoma, ALK-positive.* Crit. Rev. Oncol. Hematol., 2012. **83**(2): p. 293-302.
157. Feldman, A.L., et al., *Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing.* Blood, 2011. **117**(3): p. 915-9.
158. Feldman, A.L., et al., *Recurrent translocations involving the IRF4 oncogene locus in peripheral T-cell lymphomas.* Leukemia, 2009. **23**(3): p. 574-80.
159. Pham-Ledard, A., et al., *IRF4 gene rearrangements define a subgroup of CD30-positive cutaneous T-cell lymphoma: a study of 54 cases.* J. Invest. Dermatol., 2010. **130**(3): p. 816-25.
160. Desjobert, C., et al., *MiR-29a down-regulation in ALK-positive anaplastic large cell lymphomas contributes to apoptosis blockade through MCL-1 overexpression.* Blood, 2011. **117**(24): p. 6627-37.

161. Armitage, J.O., *The aggressive peripheral T-cell lymphomas: 2012 update on diagnosis, risk stratification, and management.* Am. J. Hematol., 2012. **87**(5): p. 511-9.
162. Schmitz, N., et al., *Treatment and prognosis of mature T-cell and NK-cell lymphoma: an analysis of patients with T-cell lymphoma treated in studies of the German High-Grade Non-Hodgkin Lymphoma Study Group.* Blood, 2010. **116**(18): p. 3418-25.
163. Tilly, H., et al., *Primary anaplastic large-cell lymphoma in adults: clinical presentation, immunophenotype, and outcome.* Blood, 1997. **90**(9): p. 3727-34.
164. Brugieres, L., et al., *Single-drug vinblastine as salvage treatment for refractory or relapsed anaplastic large-cell lymphoma: a report from the French Society of Pediatric Oncology.* J. Clin. Oncol., 2009. **27**(30): p. 5056-61.
165. Le Deley, M.C., et al., *Vinblastine in children and adolescents with high-risk anaplastic large-cell lymphoma: results of the randomized ALCL99-vinblastine trial.* J. Clin. Oncol., 2010. **28**(25): p. 3987-93.
166. Forde, P.M. and C.M. Rudin, *Crizotinib in the treatment of non-small-cell lung cancer.* Expert Opin. Pharmacother., 2012. **13**(8): p. 1195-201.
167. Zou, H.Y., et al., *An orally available small-molecule inhibitor of c-Met, PF-2341066, exhibits cytoreductive antitumor efficacy through antiproliferative and antiangiogenic mechanisms.* Cancer Res., 2007. **67**(9): p. 4408-17.
168. Merkel, O., et al., *Novel therapeutic options in anaplastic large cell lymphoma: molecular targets and immunological tools.* Mol. Cancer Ther., 2011. **10**(7): p. 1127-36.
169. Mesaros, E.F., et al., *Novel 2,3,4,5-tetrahydro-benzo[d]azepine derivatives of 2,4-diaminopyrimidine, selective and orally bioavailable ALK inhibitors with antitumor efficacy in ALCL mouse models.* Bioorg. Med. Chem. Lett., 2011. **21**(1): p. 463-6.
170. Neckers, L. and P. Workman, *Hsp90 molecular chaperone inhibitors: are we there yet?* Clin. Cancer Res., 2012. **18**(1): p. 64-76.
171. Galkin, A.V., et al., *Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK.* Proc. Natl. Acad. Sci. U S A, 2007. **104**(1): p. 270-5.
172. Sabbatini, P., et al., *GSK1838705A inhibits the insulin-like growth factor-1 receptor and anaplastic lymphoma kinase and shows antitumor activity in experimental models of human cancers.* Mol. Cancer Ther., 2009. **8**(10): p. 2811-20.
173. Chiarle, R., et al., *The anaplastic lymphoma kinase is an effective oncoantigen for lymphoma vaccination.* Nat. Med., 2008. **14**(6): p. 676-80.
174. Passoni, L., et al., *ALK as a novel lymphoma-associated tumor antigen: identification of 2 HLA-A2.1-restricted CD8+ T-cell epitopes.* Blood, 2002. **99**(6): p. 2100-6.
175. Changelian, P.S., et al., *The specificity of JAK3 kinase inhibitors.* Blood, 2008. **111**(4): p. 2155-7.

176. Han, Y., et al., *Restoration of shp1 expression by 5-AZA-2'-deoxycytidine is associated with downregulation of JAK3/STAT3 signaling in ALK-positive anaplastic large cell lymphoma.* Leukemia, 2006. **20**(9): p. 1602-9.
177. Siddiquee, K., et al., *Selective chemical probe inhibitor of Stat3, identified through structure-based virtual screening, induces antitumor activity.* Proc. Natl. Acad. Sci. U S A, 2007. **104**(18): p. 7391-6.
178. Oltersdorf, T., et al., *An inhibitor of Bcl-2 family proteins induces regression of solid tumours.* Nature, 2005. **435**(7042): p. 677-81.
179. Tse, C., et al., *ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor.* Cancer Res., 2008. **68**(9): p. 3421-8.
180. Sebti, S.M. and A.D. Hamilton, *Farnesyltransferase and geranylgeranyltransferase I inhibitors and cancer therapy: lessons from mechanism and bench-to-bedside translational studies.* Oncogene, 2000. **19**(56): p. 6584-93.
181. Lindsley, C.W., *The Akt/PKB family of protein kinases: a review of small molecule inhibitors and progress towards target validation: a 2009 update.* Curr. Top Med. Chem., 2010. **10**(4): p. 458-77.
182. Falini, B., et al., *Response of refractory Hodgkin's disease to monoclonal anti-CD30 immunotoxin.* Lancet, 1992. **339**(8803): p. 1195-6.
183. Ansell, S.M., et al., *Phase I/II study of an anti-CD30 monoclonal antibody (MDX-060) in Hodgkin's lymphoma and anaplastic large-cell lymphoma.* J. Clin. Oncol., 2007. **25**(19): p. 2764-9.
184. Forero-Torres, A., et al., *A Phase II study of SGN-30 (anti-CD30 mAb) in Hodgkin lymphoma or systemic anaplastic large cell lymphoma.* Br. J. Haematol., 2009. **146**(2): p. 171-9.
185. Younes, A., et al., *Brentuximab vedotin (SGN-35) for relapsed CD30-positive lymphomas.* N. Engl. J. Med., 2010. **363**(19): p. 1812-21.
186. Keller, A.D. and T. Maniatis, *Identification and characterization of a novel repressor of beta-interferon gene expression.* Genes Dev., 1991. **5**(5): p. 868-79.
187. Turner, C.A., Jr., D.H. Mack, and M.M. Davis, *Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells.* Cell, 1994. **77**(2): p. 297-306.
188. Huang, S., *Blimp-1 is the murine homolog of the human transcriptional repressor PRDI-BF1.* Cell, 1994. **78**(1): p. 9.
189. Martins, G. and K. Calame, *Regulation and functions of Blimp-1 in T and B lymphocytes.* Annu. Rev. Immunol., 2008. **26**: p. 133-69.
190. Mock, B.A., et al., *The B-lymphocyte maturation promoting transcription factor BLIMP1/PRDI-BF1 maps to D6S447 on human chromosome 6q21-q22.1 and the syntenic region of mouse chromosome 10.* Genomics, 1996. **37**(1): p. 24-8.
191. Gyory, I., et al., *Identification of a functionally impaired positive regulatory domain I binding factor 1 transcription repressor in myeloma cell lines.* J. Immunol., 2003. **170**(6): p. 3125-33.

192. Shapiro-Shelef, M., et al., *Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells.* Immunity, 2003. **19**(4): p. 607-20.
193. Keller, A.D. and T. Maniatis, *Only two of the five zinc fingers of the eukaryotic transcriptional repressor PRDI-BF1 are required for sequence-specific DNA binding.* Mol. Cell Biol., 1992. **12**(5): p. 1940-9.
194. Kuo, T.C. and K.L. Calame, *B lymphocyte-induced maturation protein (Blimp)-1, IFN regulatory factor (IRF)-1, and IRF-2 can bind to the same regulatory sites.* J. Immunol., 2004. **173**(9): p. 5556-63.
195. Ren, B., et al., *PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins.* Genes Dev., 1999. **13**(1): p. 125-37.
196. Yu, J., et al., *Transcriptional repression by blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase.* Mol. Cell Biol., 2000. **20**(7): p. 2592-603.
197. Vincent, S.D., et al., *The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse.* Development, 2005. **132**(6): p. 1315-25.
198. Kallies, A., et al., *Transcriptional repressor Blimp-1 is essential for T cell homeostasis and self-tolerance.* Nat. Immunol., 2006. **7**(5): p. 466-74.
199. Chang, D.H., G. Cattoretti, and K.L. Calame, *The dynamic expression pattern of B lymphocyte induced maturation protein-1 (Blimp-1) during mouse embryonic development.* Mech. Dev., 2002. **117**(1-2): p. 305-9.
200. Kallies, A., et al., *Plasma cell ontogeny defined by quantitative changes in blimp-1 expression.* J. Exp. Med., 2004. **200**(8): p. 967-77.
201. Messika, E.J., et al., *Differential effect of B lymphocyte-induced maturation protein (Blimp-1) expression on cell fate during B cell development.* J. Exp. Med., 1998. **188**(3): p. 515-25.
202. Chang, D.H., C. Angelin-Duclos, and K. Calame, *BLIMP-1: trigger for differentiation of myeloid lineage.* Nat. Immunol., 2000. **1**(2): p. 169-76.
203. Nishikawa, K., et al., *Blimp1-mediated repression of negative regulators is required for osteoclast differentiation.* Proc. Natl. Acad. Sci. U S A, 2010. **107**(7): p. 3117-22.
204. Chan, Y.H., et al., *Absence of the transcriptional repressor Blimp-1 in hematopoietic lineages reveals its role in dendritic cell homeostatic development and function.* J. Immunol., 2009. **183**(11): p. 7039-46.
205. Martins, G.A., et al., *Transcriptional repressor Blimp-1 regulates T cell homeostasis and function.* Nat. Immunol., 2006. **7**(5): p. 457-65.
206. Fairfax, K.A., et al., *Plasma cell development: from B-cell subsets to long-term survival niches.* Semin. Immunol., 2008. **20**(1): p. 49-58.
207. Klein, U. and R. Dalla-Favera, *Germinal centres: role in B-cell physiology and malignancy.* Nat .Rev. Immunol., 2008. **8**(1): p. 22-33.
208. Shapiro-Shelef, M. and K. Calame, *Regulation of plasma-cell development.* Nat. Rev. Immunol., 2005. **5**(3): p. 230-42.

209. Fearon, D.T., P. Manders, and S.D. Wagner, *Arrested differentiation, the self-renewing memory lymphocyte, and vaccination.* Science, 2001. **293**(5528): p. 248-50.
210. Klein, U., et al., *Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination.* Nat. Immunol., 2006. **7**(7): p. 773-82.
211. Sciammas, R., et al., *Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation.* Immunity, 2006. **25**(2): p. 225-36.
212. Kallies, A. and S.L. Nutt, *Terminal differentiation of lymphocytes depends on Blimp-1.* Curr. Opin. Immunol., 2007. **19**(2): p. 156-62.
213. Nutt, S.L., K.A. Fairfax, and A. Kallies, *BLIMP1 guides the fate of effector B and T cells.* Nat. Rev. Immunol., 2007. **7**(12): p. 923-7.
214. Barnes, M.J. and F. Powrie, *Regulatory T cells reinforce intestinal homeostasis.* Immunity, 2009. **31**(3): p. 401-11.
215. Kusam, S., et al., *Inhibition of Th2 differentiation and GATA-3 expression by BCL-6.* J. Immunol., 2003. **170**(5): p. 2435-41.
216. Tam, W., et al., *Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas.* Blood, 2006. **107**(10): p. 4090-100.
217. Calado, D.P., et al., *Constitutive canonical NF-kappaB activation cooperates with disruption of BLIMP1 in the pathogenesis of activated B cell-like diffuse large cell lymphoma.* Cancer Cell, 2010. **18**(6): p. 580-9.
218. Rancoita, P.M., et al., *Bayesian DNA copy number analysis.* BMC Bioinformatics, 2009. **10**: p. 10.
219. Olshen, A.B., et al., *Parent-specific copy number in paired tumor-normal studies using circular binary segmentation.* Bioinformatics, 2011. **27**(15): p. 2038-46.
220. Lenz, G., et al., *Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways.* Proc. Natl. Acad. Sci .U S A, 2008. **105**(36): p. 13520-5.
221. Mermel, C.H., et al., *GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers.* Genome Biol., 2011. **12**(4): p. R41.
222. Koressaar, T. and M. Remm, *Enhancements and modifications of primer design program Primer3.* Bioinformatics, 2007. **23**(10): p. 1289-91.
223. Beroukhim, R., et al., *Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma.* Proc. Natl. Acad. Sci. U S A, 2007. **104**(50): p. 20007-12.
224. Rossi, D., et al., *The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness.* Clin. Cancer Res., 2009. **15**(3): p. 995-1004.
225. Xu-Monette, Z.Y., et al., *Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies.* Blood, 2012. **119**(16): p. 3668-83.

226. Drakos, E., et al., *The therapeutic potential of p53 reactivation by nutlin-3a in ALK+ anaplastic large cell lymphoma with wild-type or mutated p53.* Leukemia, 2009. **23**(12): p. 2290-9.
227. Piya, S., et al., *Suppression of IRF4 by IRF1, 3, and 7 in Noxa expression is a necessary event for IFN-gamma-mediated tumor elimination.* Mol. Cancer Res., 2011. **9**(10): p. 1356-65.
228. Lopez-Girona, A., et al., *Lenalidomide downregulates the cell survival factor, interferon regulatory factor-4, providing a potential mechanistic link for predicting response.* Br. J. Haematol., 2011. **154**(3): p. 325-36.
229. Shaffer, A.L., et al., *IRF4 addiction in multiple myeloma.* Nature, 2008. **454**(7201): p. 226-31.
230. Ploner, C., R. Kofler, and A. Villunger, *Noxa: at the tip of the balance between life and death.* Oncogene, 2008. **27 Suppl 1**: p. S84-92.
231. Kim, J.H., W.S. Kim, and C. Park, *Epstein-Barr virus latent membrane protein-1 protects B-cell lymphoma from rituximab-induced apoptosis through miR-155-mediated Akt activation and up-regulation of Mcl-1.* Leuk. Lymphoma, 2012. **53**(8): p. 1586-91.
232. Babar, I.A., et al., *Nanoparticle-based therapy in an in vivo microRNA-155 (miR-155)-dependent mouse model of lymphoma.* Proc. Natl. Acad. Sci. U S A, 2012. **109**(26): p. E1695-704.
233. Desai, S., et al., *PRDM1 is required for mantle cell lymphoma response to bortezomib.* Mol. Cancer. Res., 2010. **8**(6): p. 907-18.

Michela Boi

Viale A. Gramsci 87, 10095 Grugliasco, Turin – Italy
•michelaboi84@gmail.com • +393498240384

PERSONAL INFORMATION

Name: Michela Boi
Birth Date: 14th May 1984
Birth Place: Turin
Nationality: Italian

EDUCATION

University of Geneva, Switzerland
Ph.D. student in Pharmaceutical Sciences — scheduled for July 2013

- preparation of thesis on the study of Anaplastic Large Cell Lymphoma (ALCL) at a genomic and gene expression level (supervisor at the University of Geneva: Prof. Leonardo Scapozza; leonardo.scapozza@unige.ch)

University of Turin, Italy
Master Degree in Medical Biotechnology (2 years) — 09/2008

- Thesis title: "Telomere length is an independent prognostic factor in chronic lymphocytic leukemia: analysis on 401 patients"
- Grade: 110/110 *summa cum laude*

Bachelor Degree in Biotecnology - Medical Curriculum- (3 years) — 10/2006
- Thesis title: "The array-CGH technique in onco-haematology"
- Grade: 104/110

Liceo Scientifico "Marie Curie" Grugliasco, Turin, Italy — 07/2003
B.A. Scientific Lyceum (grade: 90/100)

PROFESSIONAL EXPERIENCE

Department of Molecular Biotechnology and Health Science and Center for Experimental Research and Medical Studies (CeRMS), University of Torino, Torino, Italy 5/2013-today

Fellowship financed by “Fondazione Gigi Ghirotti”: “Identification of new risk factors for the improvement of onco-hematology pathologies treatment”

- Research topic: Discover of new biological and molecular markers associated with the refractory to conventional chemotherapy of Diffuse Large B-Cell Lymphoma through the study of genomic and transcriptional imprint (WES and RNAseq) of these samples. The putative pathogenetic lesions identified are validated *in vivo* using Patient Derived Tumorgrafts (PDT, models derived from implant of primary lesions) immunodeficient mice. This model is used to study in depth the lesion and to evaluate different therapeutic strategy.

Institute of Oncology of Southern Switzerland (IOSI) / Institute of Oncology Research (IOR), Bellinzona, Switzerland 9/2009-3/20013

Ph.D. student, Functional genomics and lymphoma group, Francesco Bertoni, MD. via V. Vela 6, Bellinzona, Switzerland; frbertoni@mac.com

- Research topic: study of lymphomas pathogenesis through genomic and gene expression profiling analyisis; characterization of Anaplastic Large Cell Lymphoma (ALCL) pathogenesis through functional studies; study of new drugs function in the lymphoma field.
-

University of Turin, Italy 11/2008-07/2009

Internship,Molecular Biology Laboratory, Haematology Division Prof. Boccadoro, University of Turin.Marco Ladetto, MD. Via Genova 3, Turin, Italy; marco.ladetto@unito.it

Scholarship financed by Piedmont Region funds: “ Molecular and flow-cytometric evaluation of the immunologic reconstitution of the T cellular compartment after hematopoietic stem cell allogenic transplant”

- Research topic: analysis of the presence or absence of typical chromosomal translocation of lymphomas (Blc2/IgH- Bcl1/IgH), mutational analysis of Ig heavy chain in CLL and MM; study on telomere length

11/2005-09/2008

Visiting Student, Molecular Biology Laboratory, Haematology Division Prof. Boccadoro, University of Turin. Marco Ladetto, MD. Via Genova 3, Turin, Italy; marco.ladetto@unito.it

- Research topic: training for the preparation of the thesis master degree on the study of telomere length in Chroonic Lymphocytic Leukemia (CLL) and training for the preparation of the thesis bachelor degree on the study of cytogenetic techniques in onco-haematology

RESEARCH AREA AND SCIENTIFIC ACTIVITY

My research activity is mainly centered on the study of the pathogenesis of lymphomas. To this end, I used genomic and gene expression profiling data in order to understand what are the genes, located in aberrated regions of the genome, which could have a role in lymphomagenesis. An important part of my research is the validation of this data *in vitro*, through functional studies and phenotypic analysis. Meanwhile, I am testing new drugs on lymphoma cell lines in order to characterize their function *in vitro*.

LABORATORY SKILLS

MOLECULAR BIOLOGY
Experienced with many molecular biology techniques on DNA, RNA, and protein, such as nucleic acid extractions, PCR, RealTime-qPCR; RT-PCR; mutational analysis; array-CGH; Methylation assay; Western Blot; basic skills on cloning.

CELLULAR BIOLOGY
Experience with many lymphoma cell lines (HEK293T, B-cell lymphoma cell lines, ALCL cell lines); cellular toxicity assay (MTT assay); stable and transient transfection techniques; lentivirus production and lentiviral infection; primary cell isolation (PBMC isolation from blood, bone marrow and lymphonode tissue); basic techniques of cytofluorimetry.

BIOINFORMATICS
Basic skills on bioinformatic analysis: Partek, GSEA, DAVID

LANGUAGES AND OTHER SKILLS

Italian: native; English: fluent; French: basic.
IT: MSOffice, Adobe Illustrator, Endnote, 4Peaks and Chromas, FlowJo, ClustalW, Blast and IgBlast, Knowledge of the main scientific browsers and databases (PubMed, NCBI, Ensembl, UCSD Genome Browser, etc).
Other Skills: Good communication skills; Training; Continuous Improvement.

POSTERS

1. Bonetti P, Boi M, Bernasconi E, Rinaldi A, Kwee I, Gaudio E, Ponzoni M, Tibiletti MG, Stathis A, Riveiro E, Inghirami G, Zucca e, Bertoni F. The BRD-inhibitor OTX015 affects proliferation and gene expression in pre-clinical models of mature lymphoid tumors AACR Annual Meeting Abstracts 2013

2. Boi M, Bonetti P, Ponzoni M, Tibiletti MG, Stathis A, Cvitkovic E, Inghirami G, Zucca E, Bertoni F. The Brd-Inhibitor OTX015 Shows Pre-Clinical Activity in Anaplastic Large T-Cell Lymphoma (ALCL). ASH Annual Meeting Abstracts 2012; 120: 4872-.

3. Bonetti P, Boi M, Ponzoni M, Tibiletti MG, Stahis A, Inghirami G, Noel K, Zucca E, Bertoni F. The Brd-Inhibitor OTX015 Is Active in Pre-Clinical Models of Mature B-Cell Lymphoid Tumors. ASH Annual Meeting Abstracts 2012; 120: 1657-.

4. Boi M, Rinaldi A, Piva R, Rancoita PMV, Bonetti P, Matolcsy A, Tousseyn T, Rodríguez-Pinilla SM, Piris M, Beà S, Campo E, Bhagat G, Swerdlow SH, Rosenwald A, Ponzoni M, Young KH, PiccalugaPP, Pileri S, Neri A, Medico E, Zucca E, Kwee I, Inghirami G and Bertoni F. "Genomic Profiling of Anaplastic Large Cell Lymphoma (ALCL)". PhD day, Geneve 15th June 2012

5. Boi M, Rinaldi A, Piva R, Rancoita PMV, Bonetti P, Matolcsy A, Tousseyn T, Rodríguez-Pinilla SM, Piris M, Beà S, Campo E, Bhagat G, Swerdlow SH, Rosenwald A, Ponzoni M, Young KH, PiccalugaPP, Pileri S, Neri A, Medico E, Zucca E, Kwee I, Inghirami G and Bertoni F. "BLIMP1 is commonly inactivated in Anaplastic Large Cell Lymphoma". Abstract submission for Congress of the American Society of Hematology 2011; Abstract submission for USGEB, Losanna, february 2012.

6. Ferrero S, Capello D, Svaldi M, Boi M, Gatti D, Drandi D, Monitillo L, Bussano S, Barbiero S, Lobetti Bodoni C, Bertoni F, Gattei V, Palumbo A, Boccadoro M, Cortelazzo S, Gaidano G and Ladetto M. “Multiple Myeloma (MM) immunoglobulin sequence show no intra-disease clustering but are occasionally related to repertoires from normal lymphocytes and other B-cell tumors.” Abstract submission for 15th Congress of the European Hematology Association 2010.

7. Lobetti Bodoni C,. Genuardi E, Passera R, Sia D, Crisà E, Giai V, Rocci A, Monitillo L, Drandi D, Zanni M, Boi M, Isaia G, Bernocco E, Lunghi M, Abruzzese E, Radaelli F, Pini M, Pregno P, Gaidano G, Boccadoro M, Carlo-Stella C, Ferrero D and Ladetto M. “Severe telomeric loss occurs in Ph-negative hematopoiesis emerging after successful treatment of chronic myeloid leukemia and associates with acquired cytogenetic lesions and low hemoglobin levels.”Abstract submission for 15th Congress of the European Hematology Association 2010.

8. Boi M, Kwee I, De Campos CP, Greiner TC, Chan WC, Gaidano G, Bhagat G, Ponzoni M, Zucca E and Bertoni F. “High-resolution genomic profiling of 533 B-cell lymphomas defines distinct tumor signatures, genomic aberrations correlated with outcome and pathogenetic subgroups.” ECCLU 2010.

9. Ferrero S, Capello D, Svaldi M, Boi M, Drandi D, Rossi D, Pagliano G, Monitillo L, Lobetti Bodoni C, Genuardi E, Critelli R, Bussano S, Barbero D, Mantoan B, Bernocco E, Palumbo A, Boccadoro M, Gaidano G, Cortelazzo S and Ladetto M. "Characteristics of the immunoglobulin heavy chain (IgH) rearrangement in multiple myeloma (MM): an analysis on 308 cases." ASH Annual Meeting Abstracts 2009.

10. Rossi D, Lobetti Bodoni C, Genuardi E, Monitillo L, Cerri M, Deambrogi C, Drandi D, Rocci A, Boi M, Ferrero S, Capello D, De Paoli L, Bergui L, Omedè P, Massaia M, Boccadoro M, Gaidano G and Ladetto M. "Short telomere length is an independent predictor of survival, progression, Richter's syndrome transformation and recurrent infections: an analysis on 421 CLL patients including blinded validation on 230 cases". 13th Congress of the European Hematology Association, June 15, 2008.

11. Ladetto M, Drandi D, Santo L, Mantoan B, Boi M, Dell'Aquila M, Monitillo L, Lobetti Bodoni C, Palumbo A, Astolfi M and Boccadoro M. "Immunoglobulin (Ig) repertoire in Multiple Myeloma: high frequency of recurrent aminoacid substitutions in the FR2 and CDR2". ASH Annual Meeting Abstracts 2006.

12. Drandi D, Ladetto M, El Ali M, Santo L, Ricca I, Dell'Aquila M, Monitillo L, Li Bergolis S, Boi M, Rocci A, Lobetti Bodoni C, Astolfi M, Ambrosini MT, Omedè P, Palumbo A and Boccadoro M. "Array-CGH detects frequent recurrent imbalances in plasma cells disorder (PCD)". 11th Congress of the European Hematology Association 2006.

ORIGINAL REPORTS

1. Boi M, Rinaldi A, Kwee I, Bonetti P, Todaro m, Tabbo' F, Piva R, Rancoita PMV, Matolcsy A, Timar B, Tousseyn T, Rodriguez-Pinilla SC, Piris MA, Beà S, Campo E, Bhagat G, Swerdlow SH, Rosenwald A, Ponzoni M, Young KH, Piccaluga PP, Tzankov A, Dummer R, Pileri S, Zucca E, Inghirami G, Bertoni F. "BLIMP1/PRDM1 is commonly inactivated in Anaplastic Large T-cell Lymphoma". Submitted.

2. Marchiolatti R, Messana K, Landra I, Abate F, Tabbo' F, Boi M, Todaro M, Barreca A, Chiesa N, Aliberti S, Di Giacomo F, Lasorsa E, Bessone L, Gaudiano M, Rinaldi A, Ponzoni M, Longo DL, Aime S, Cheng M, Ruggeri B, Tebb G, Piccaluga PP, Pileri S, Tiacci E, Falini B, Shultz LD, Farinelli L, Kwee I, Piva R, Medico E, Rabadan R, Bertoni F, Inghirami G and the European T-cell Lymphoma Study Group. "Massive RNAseq defines novel tumorigenic gene fusions of Anaplastic Large Cell Lymphoma". Submitted

3. Agnelli L, Mereu E, Pellegrino E, Limongi T, Kwee I, Bergaggio E, Ponzoni M, Zamò A, Iqbal J, Piccaluga PP, Neri A, Chan WC, Pileri S, Bertoni F, Inghirami G, Piva R; European T-Cell Lymphoma Study Group. "Identification of a 3-gene model as a powerful diagnostic tool for the recognition of ALK-negative anaplastic large-cell lymphoma." Blood 2012, 120(6):1274-81

4. Lobetti-Bodoni C, Ferrero D, Genuardi E, Passera R, Bernocco E, Sia D, Grignani G, Crisà E, Monitillo L, Rocci A, Drandi D, Giai V, Zanni M, Boi M, Isaia G, Barbero D, Lunghi M, Abruzzese E, Radaelli F, Pini M, Pregno P, Carlo-Stella C, Gaidano G, Boccadoro M and Ladetto M."Telomere loss in Philadelphia-negative hematopoiesis after successful treatment of chronic myeloid leukemia: Evidence for premature aging of the myeloid compartment", Mech Ageing Dev. 2012 Jun 9

5. Tabbó F, Barreca A, Piva R, Inghirami G; The European T-Cell Lymphoma Study Group. "ALK Signaling and Target Therapy in Anaplastic Large Cell Lymphoma." Front Oncol. 2012;2:41.

6. Boi M, Stathis A, Zucca E, Inghirami G and Bertoni F. "Genetic alterations in systemic nodal and extranodal non-cutaneous lymphomas derived from mature T cells and natural killer cells".Cancer Sci. 2012 May 9.

7. Ferrero S, Capello D, Svaldi M, Boi M, Gatti D, Drandi D, Rossi D, Barbiero S, Mantoan B, Mantella E, Zanni M, Ghione P, Larocca A, Passera R, Bertoni F, Gattei V, Forconi F, Laurenti L, Del Poeta G, Marasca R, Cortelazzo S, Gaidano G, Palumbo A, Boccadoro M and Ladetto M. "Multiple myeloma shows no intra-disease clustering of immunoglobulin heavy chain genes". Haematologica. 2012 Jun;97(6):849-53.

8. Rossi D, Lobetti Bodoni C, Genuardi E, Monitillo L, Drandi D, Cerri M, Deambrogi C, Ricca I, Rocci A, Ferrero S, Bernocco E, Capello D, De Paoli L, Bergui L, Boi M, Omedè P, Massaia M, Tarella C, Passera R, Boccadoro M, Gaidano G and Ladetto M."Telomere length is an independent predictor of survival, treatment requirement and Richter's syndrome transformation in chronic lymphocytic leukemia." Leukemia. 2009 Jun;23(6):1062-72.

I hereby grant permission to use my personal data in accordance with Art.23, Legislative Decree n. 196/03 -Personal Data Protection Code.

Very Truly Yours,
Michela Boi

Acknowledgements

I would like to thank all people introducing me to the research and I would like to express my deepest appreciation to all those who provided me the possibility to complete this thesis. I especially wish to thank my advisor, Dr Francesco Bertoni, for his guidance during my research and study at the Institute of Oncolgy Research (IOR) and for the useful comments, remarks and engagement through the learning process of this PhD thesis.

Besides my advisor, I would like to thank my thesis reviewers: Prof. Leonardo Scapozza at the University of Geneva for living me the opportunity to be enrolled in the PhD program of the University of Geneva and for his encouragement, insightful comments, and stimulating questions

My sincere thanks also goes to Prof. Giorgio Inghirami, for giving me the opportunity to perform my PhD project in collaboration with him and his group at the University of Turin His perpetual energy and enthusiasm in research gave me the stimulus to proceed in this long and usually insidious way.

I am indebted to all the people working with me, my colleagues and friends for providing a stimulating and pleasant environment in which to learn and grow. I am especially grateful to Paola Bonetti, my collegue and friend from whom I have learnt a lot, and Lara Brambilla and Laura Curti, my special roommates, friends, collegues and companions of adventures.

Last but not the least, I would like to extend a special thanks to my family and my friends for their continuous support.