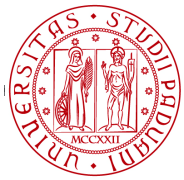


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Salute della Donna e del Bambino

Clinica di Oncoematologia Pediatrica

SCUOLA DI DOTTORATO DI RICERCA IN

MEDICINA DELLO SVILUPPO E SCIENZE DELLA PROGRAMMAZIONE

INDIRIZZO: EMATOONCOLOGIA, IMMUNOLOGIA E GENETICA

CICLO XXV

ULTRA-DEEP MUTATIONAL ANALYSIS OF NPM-ALK AND POSSIBLE IMPLICATIONS ON TARGET THERAPY IN ANAPLASTIC LARGE CELL LYMPHOMA OF CHILDHOOD

Direttore della Scuola e
Coordinatore d'indirizzo: Ch.mo Prof. GIUSEPPE BASSO

Supervisore: Prof. ANGELO ROSOLEN

Dottoranda: FEDERICA LOVISA

Giancarlo,

“A te che mi hai insegnato i sogni
E l’arte dell’avventura
A te che credi nel coraggio
E anche nella paura
A te che sei la miglior cosa
Che mi sia successa”.

Lorenzo Cherubini

INDEX

SUMMARY 1

RIASSUNTO 3

1. INTRODUCTION

1.1 CLINICAL AND BIOLOGICAL CHARACTERIZATION OF ALCL 5

1.2 STRUCTURE AND PATHOGENIC ROLE OF NPM-ALK 6

1.3 ALK AS A THERAPEUTIC TARGET 9

1.4 ULTRA-DEEP TARGETED SEQUENCING FOR LOW PREVALENCE MUTATIONS DETECTION IN CANCER 10

2. OBJECTIVES 13

3. MATERIALS AND METHODS

3.1 PATIENTS, SAMPLES AND CELL LINES 15

3.2 REAGENTS AND ANTIBODIES 15

3.3 STATISTICAL ANALYSIS 16

3.4 RT-PCR AND AMPLICON LIBRARY PREPARATION 16

3.5 NEXT-GENERATION ULTRA-DEEP SEQUENCING 18

3.6 DATA ANALYSIS AND DETECTION OF VARIANTS 18

3.7 HOMOLOGY MODELLING AND MOLECULAR DYNAMICS SIMULATION 19

3.8 SITE-DIRECTED MUTAGENESIS 19

3.9 SUBCLONING 20

3.10 TRANSIENT TRANSFECTION AND TREATMENTS 22

3.11 CELL LYSIS, IMMUNOBLOTTING AND IMMUNOPRECIPITATION 23

3.12 IMMUNOFLUORESCENCE 23

4. RESULTS

4.1 IDENTIFICATION OF NPM-ALK KD MUTATIONS 25

4.2 MOLECULAR MODELLING ANALYSIS OF NPM-ALK INDEL MUTATIONS 26

4.3 *IN VITRO* EXPRESSION AND ACTIVITY OF NPM-ALK INDEL MUTANTS 27

4.4 COESPRESSION OF NATIVE NPM-ALK AND INDEL MUTANTS 28

4.5 PATIENTS CHARACTERISTICS AND PROGNOSTIC IMPACT OF INDEL MUTATIONS 30

5. DISCUSSION 31

FIGURES AND TABLES 35

BIBLIOGRAPHY 49

RINGRAZIAMENTI 55

SUMMARY

Anaplastic Large Cell Lymphoma (ALCL) represents a distinct subset of aggressive T-cell non-Hodgkin lymphoma (NHL) accounting for about 3% of adult NHL and 10 to 15% of childhood lymphomas. In the vast majority of the cases, ALCL is associated to chromosomal translocations, the most frequent being the t(2;5)(p23;q35), involving the Anaplastic Lymphoma Kinase (*ALK*) gene, which lead to aberrant NPM-ALK protein expression and kinase activity. It has been extensively demonstrated that aberrant NPM-ALK expression contributes to the pathogenesis of ALK-positive ALCL, as it causes cell transformation through activation of several biological pathways related to cell proliferation, cell-cycle control and apoptosis. Although ALK-positive ALCL have a rather benign prognosis when treated with standard chemotherapy, the failure rate at two years is almost 30% for most of these regimens. Notably, most of relapses occur within the first year from the start of therapy, and long-term survival for relapsed disease is less than 50%. Aberrant ALK activity is one of the major oncogenic events not only in ALK-positive ALCL, but also in neuroblastoma, non-small cell lung cancer (NSCLC) and inflammatory myofibroblastic tumour (IMT) bearing ALK activating mutations/rearrangements, and the inhibition of ALK kinase activity was proven to substantially reduce cancer cell proliferation and invasiveness both *in vitro* and *in vivo.* Successful clinical experience with crizotinib further support the concept of ALK-specific inhibition as a valuable treatment strategy in ALK-positive ALCL, as well as in other ALK-addicted tumours. However, similarly to other inhibitors selectively targeting oncogenic kinases, data on relapse to crizotinib due to newly acquired secondary mutations were reported. In this context, although a robust clinical response of ALCL patients to an ALK inhibitor is expected, some of those patients are also anticipated to develop resistance, making the knowledge of NPM-ALK kinase domain (KD) mutational status a valuable and mandatory information to the rational design of ALK-targeted therapies.

To detect somatic tumour mutations with potential utility for predicting treatment response in ALK-positive ALCL patients, we performed ultra-deep sequencing analysis on ALK exons 22-25, corresponding to the entire KD coding region, in 37 ALCL pediatric patients. Two low frequent point mutations were identified in two distinct cases, corresponding to the R1275Q and R1231Q amino acid changes. The R1275Q mutation has been already reported as one of the most frequent activating mutations in

neuroblastoma, while the R1231Q amino acid substitution represents a novel ALK point mutation, which to our knowledge has never been reported neither in ALK receptor nor in other ALK-translocated kinases. The molecular implications of R1275Q and R1231Q point substitutions on NPM-ALK function and sensitivity to ALK-specific inhibition are still under our investigation. In addition to point mutation, oncokinase alternative spliced transcripts have been previously reported in patients with Bcr-Abl positive chronic myeloid leukemia and more recently ALK receptor isoforms were described in neuroblastoma. To our knowledge, however, NPM-ALK alternative splicing events have never been described. For the first time, we identified and characterized 9 NPM-ALK INDEL mutations, resulting from KD whole exons skipping or alternative canonical splicing sites recognition. To investigate the effect of INDEL mutations on the structure and activity of NPM-ALK, we performed molecular homology modelling and *in vitro* functional analysis. While all these mutants were shown to be kinase dead, we demonstrated that, when coespressed with wild-type NPM-ALK, these INDELs do interact with wild-type monomers and are likely to inhibit ALK kinase activity and increase sensitivity to treatment with crizotinib. This work demonstrates that NPM-ALK KD point mutations are extremely rare in newly diagnosed ALCL patients, but positive selection of mutated cells could not be excluded in case of an ALK-targeting therapy. Conversely, our results suggest that alternative splicing in NPM-ALK may represent a common event. A clear correlation between the presence of these variants and outcome could not be detected, possibly because of the restricted cohort of patients analyzed. However, we hypothesize that a significant impact of these mutations could be observed if an ALK-specific treatment is used. Patients bearing a consistent level of inactive NPM-ALK are therefore expected to respond differently to ALK kinase inhibitors; whether such a response will be increased or decreased, it remains to be elucidated.

RIASSUNTO

Il linfoma anaplastico a grandi cellule (ALCL) è un sottotipo di linfoma non-Hodgkin (LNH) aggressivo prevalentemente a fenotipo T e rappresenta circa il 3% dei LNH dell'adulto e il 10-15% dei linfomi pediatrici. Nella maggioranza dei casi, l'ALCL è associato alla presenza di traslocazioni cromosomiche che coinvolgono il gene della chinasi ALK (Anaplastic Lymphoma Kinase), la più frequente delle quali è la t(2;5)(p23;q35), che determina l'espressione aberrante della chinasi di fusione NPM-ALK, costitutivamente attiva. È stato ampiamente dimostrato che l'espressione aberrante di NPM-ALK contribuisce alla patogenesi dell'ALCL, esercitando la sua attività trasformante mediante l'attivazione di numerosi vie di trasduzione del segnale connesse alla proliferazione, al controllo del ciclo cellulare e all'apoptosi. Sebbene gli ALCL che esprimono ALK siano caratterizzati da una prognosi piuttosto favorevole, quando trattati secondo i protocolli terapeutici standard, la percentuale di fallimento di questo tipo di terapie si attesta comunque intorno al 30%. Inoltre, la maggior parte delle recidive si verifica entro un anno dall'inizio del trattamento e la sopravvivenza dei pazienti recidivati non raggiunge il 50%. La deregolazione di ALK rappresenta uno degli eventi oncogenici più importanti non soltanto nel ALCL, ma anche nel neuroblastoma, nel tumore a piccole cellule del polmone (NSCLC) e nel tumore miofibroblastico infiammatorio (IMT) che presentano mutazioni o riarrangiamenti del gene ALK. Studi *in vitro* ed *in vivo* hanno dimostrato che l'inibizione dell'attività chinasica di ALK riduce in modo significativo la proliferazione e l'invasività delle cellule tumorali. Inoltre, l'inibitore di ALK crizotinib si è dimostrato molto efficace per il trattamento di pazienti affetti da NSCLC, ad ulteriore dimostrazione che l'inibizione dell'attività catalitica di questa proteina potrebbe rappresentare una strategia terapeutica promettente anche nell'ALCL ALK-positivo, così come in altri tumori associati ad espressione aberrante di ALK. Tuttavia, analogamente a quanto è stato osservato con altri inibitori di tirosin chinasi oncogene, alcuni pazienti in corso di trattamento con crizotinib sono recidivati in seguito all'insorgenza di mutazioni a carico della chinasi bersaglio. In questo contesto, sebbene una terapia di questo tipo possa essere estremamente efficace per il trattamento dell'ALCL, è necessario tenere presente che alcuni pazienti potrebbero sviluppare resistenza. Risulta quindi di fondamentale importanza conoscere lo stato mutazionale del dominio chinasico di NPM-ALK per poter pianificare una strategia terapeutica il più possibile efficace. Allo scopo di identificare mutazioni somatiche di NPM-ALK potenzialmente associate alla risposta al trattamento

in pazienti con ALCL ALK-positivo, abbiamo sequenziato gli esoni 22-25 di NPM-ALK, corrispondenti all'intero dominio chinasico, in 37 pazienti con ALCL pediatrico, utilizzando un approccio di *Next-Generation sequencing*. In due pazienti distinti abbiamo identificato due mutazioni puntiformi a bassa frequenza, corrispondenti alle sostituzioni aminoacidiche R1275Q e R1231Q. La prima mutazione era già stata riportata nel neuroblastoma tra le più frequenti mutazioni attivanti, mentre la mutazione R1231Q non era nota né per ALK recettore, né in tumori in cui ALK è traslocato. Le implicazioni molecolari di queste mutazioni sull'attività di NPM-ALK e sulla sua sensibilità al trattamento con inibitori specifici sono oggetto di ulteriori studi in corso. Oltre alle mutazioni puntiformi, la presenza di trascritti derivanti da fenomeni di *splicing* alternativo è un fenomeno già noto per altre protein chinasi di fusione, come Bcr-Abl nella leucemia mieloide cronica, ma anche in ALK in pazienti affetti da neuroblastoma. Sulla base delle nostre conoscenze, tali fenomeni, tuttavia, non sono mai stati descritti per NPM-ALK. Per la prima volta, abbiamo identificato e caratterizzato 9 mutazioni INDEL di NPM-ALK, associate alla perdita di interi esoni o al riconoscimento di siti di *splicing* canonici alternativi. Allo scopo di comprendere quali potessero essere gli effetti di queste mutazioni sull'attività di NPM-ALK, abbiamo condotto delle analisi di *molecular homology modelling* e studi funzionali *in vitro*. Sebbene tutti questi mutanti siano risultati inattivi, abbiamo dimostrato che se vengono coespressi assieme alla loro controparte funzionale sono in grado di formare dei complessi con essa, riducendone l'attività catalitica e aumentando la sensibilità della chinasi al trattamento con crizotinib. Questo lavoro di tesi dimostra che mutazioni puntiformi del dominio chinasico di NPM-ALK sono piuttosto rare nei pazienti con ALCL prima dell'inizio della terapia, anche se una selezione positiva delle cellule con la mutazione potrebbe verificarsi nel caso in cui venisse utilizzato un trattamento a base di inibitori specifici di ALK. Al contrario, abbiamo dimostrato che gli eventi di *splicing* alternativo di NPM-ALK sono maggiormente frequenti, anche se non è stato possibile dimostrare una correlazione tra la presenza di queste varianti e il decorso della malattia, probabilmente a causa della bassa numerosità della nostra casistica. Tuttavia, riteniamo che queste mutazioni potrebbero avere un impatto significativo nel caso in cui i pazienti venissero trattati con inibitori specifici di ALK. In particolare, pazienti che esprimono livelli rilevanti di isoforme inattive di NPM-ALK potrebbero rispondere diversamente al trattamento. Resta da stabilire se questa differenza possa tradursi in una maggiore o minore efficacia di una terapia anti-ALK in questi pazienti.

1. INTRODUCTION

1.1 CLINICAL AND BIOLOGICAL CHARACTERIZATION OF ALCL

Anaplastic Large Cell Lymphoma (ALCL) was first described by Stein *et al.* in 1985 [1]. It represents a distinct subset of aggressive T-cell non-Hodgkin lymphoma (NHL), CD30-positive, accounting for about 3% of adult NHL and 10 to 15% of childhood lymphomas [2].

In the vast majority of the cases, ALCL is associated to chromosomal translocations involving the Anaplastic Lymphoma Kinase (*ALK*) gene, which lead to aberrant ALK protein expression and kinase activity (ALK-positive ALCL). The remaining cases, sharing similar morphologic and phenotypic features, but lacking the *ALK* rearrangement and expression (ALK-negative ALCL), are considered as a separate category [2].

Based on histological characteristics, ALCL are classified as "common" type when composed of large, pleomorphic, atypical tumour cells with an abundant grey-blue cytoplasm and large kidney-shaped nucleus, "small cell" variant when represented by a predominant population of medium and small cells with clear cytoplasm and irregular nuclei, and "lymphohistiocytic" when large numbers of histiocytes are mixed with the tumour cells. Other rarer types are the "giant cell-rich" subtype, with a large number of the tumour cells containing more than one nucleus, and the "sarcomatoid" subtype, which mimics soft-tissue tumours [3].

ALCL has been clinically subdivided into a primary form (*de novo*) and a secondary form (anaplastic transformation from another lymphoma), and among primary ALCLs, systemic and cutaneous subtypes can be distinguished [4]. Primary systemic ALCL is the most frequent category, accounting for 2% to 8% of NHL in adults and approximatevely 20-30% of large cell lymphoma in children [5, 6]. ALK-positive ALCL mostly occurs in young male adults (<30 yrs; male/female ratio 6.5), and frequently presents as an aggressive stage III or IV disease, usually associated to systemic symptoms and extranodal involvement (skin, bone, soft tissue, lung and liver). In contrast, almost invariably, ALK-negative ALCL is a primary cutaneous tumour, which accounts for approximately 9% of cutaneous lymphomas, and has a lower incidence of stage III/IV disease and extranodal involvement. ALK-negative ALCL affects older patients, and in almost 25% of the cases a partial or complete spontaneous regression is observed [7]. Several independent studies demonstrated a remarkably different prognosis between

ALK-positive and ALK-negative ALCL, with the former showing a far better survival rate (70-80%) than the latter one (15-45%) [8-10].

Approximatively 85% of ALK-positive ALCL are associated to the t(2;5)(p23;q35) chromosomal translocation, in which the *ALK* gene located at 2p23 juxtaposes to the *NPM1* housekeeping gene at 5q35 [11], leading to a novel fusion gene, *NPM-ALK*, that encodes a constitutively active oncogenic tyrosine kinase (NPM-ALK). In the remaining cases, cytoplasmic-restricted expression of truncated ALK is determined by alternative genetic abnormalities that result in the expression of variant ALK fusion proteins, such as TPM3-ALK (t(1;2)(q25;p23)), TFG-ALK (t(2;3)(p23;q21)), CLTCL-ALK (t(2;22)(p23;q21)) and ATIC-ALK (inv(2)(p23q35)) [12-14] (**Figure 1)**.

1.2 STRUCTURE AND PATHOGENIC ROLE OF NPM-ALK

The NPM-ALK fusion protein contains the N-terminal domain of NPM (aa 1-118) and the entire cytoplasmic region of ALK kinase (aa 119-680) (**Figure 2**). NPM1 is a protein of 23 kDa that shuttles continuously between the cytoplasm and the nucleus, and plays multiple roles inside the cell (ribosome biogenesis, regulation of cell division, transcription and DNA repair) [15]. ALK, in contrast, is a receptor tyrosine kinase (RTK) of 200 kDa (1620 aa) which contains an extracellular ligand-binding domain, a transmembrane-spanning domain and a cytoplasmatic catalytic region, and regulates important cellular functions upon ligand-induced dimerization, through phosphorylation and activation of several downstream target molecules. ALK belongs to the insulin receptor (IR) superfamily of receptor tyrosine kinases (RTKs) [16, 17] and its postnatal expression is restricted to few scattered cells in the nervous system [18]. The 1030 aa extracellular region contains 26 hydrophobic aa at its N-terminus, comprising a signal peptide sequence, and the binding site for two putative ligands, Pleiotrophin (PTN) and Midkine (MK) [19, 20]. The extracellular region also possesses several N-glycosylation consensus sites (Asn-X-Ser/Thr), 26 cysteine residues, an LDL-A domain, a MAM domain and a glycine-rich region. Next to this, a 28-aa transmembrane domain is joined to a 64-aa juxtamembrane region, which contains a binding site for phosphotyrosine-dependent interaction with the insulin receptor substrate-1 (IRS-1) and precedes the intracellular catalytic domain and activation loop (**Figure 3A,** PTK). By analogy with other RTKs, tyrosine residues 1278, 1282 and 1283 in the activation loop of ALK represent major autophosphorylation sites during kinase activation, occluding access to ATP when non

phosphorylated, while favouring nucleotide binding when phosphorylated [21] (**Figure 3B**).

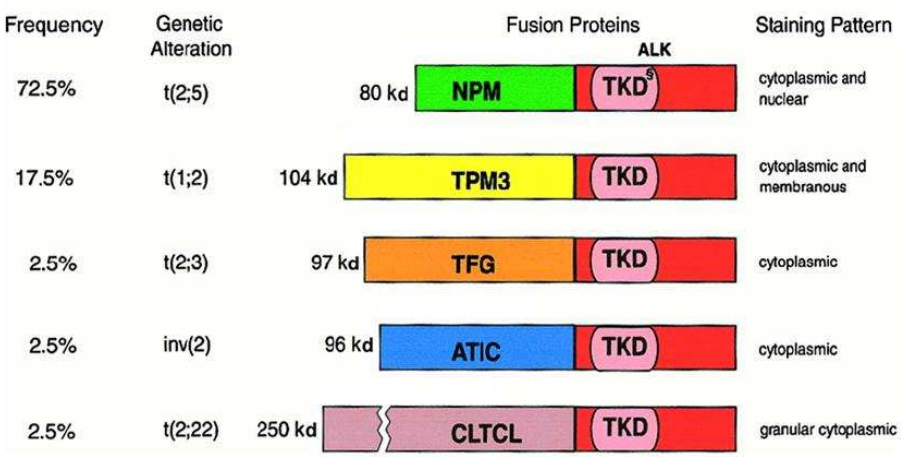


Figure 1 Features of the various ALK fusion proteins occurring in ALK-positive ALCL. §Tirosine Kinase Domain. (Stein *et al.*, 2000)

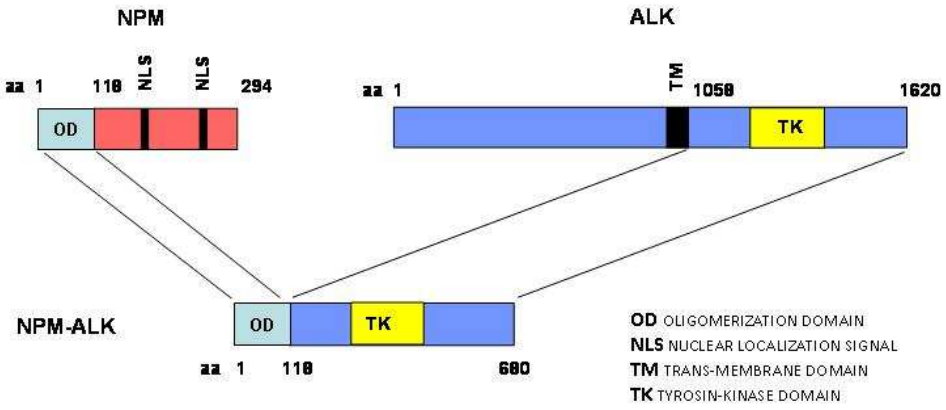


Figure 2 Schematic representation of NPM-ALK fusion protein.

The function of full-length ALK receptor is still poorly characterized, though PTN and MK have been proposed as putative ligands of ALK in mammals, and associated to ALK activity during neuron development and neurodegenerative diseases [19, 20]. However, a recent model has proposed that phosphorylation of ALK occurs through a PTN-dependent inactivation of the receptor protein tyrosine phosphatase Z1, rather than a direct interaction of PTN with ALK [22], as suggested by the observation that ALK is not phosphorylated by PTN when phosphatase Z1 lacks or has been inactivated [23-26]

Conversely, the presence of an oligomerization domain in the fused gene partners of ALK truncated variants promotes ligand-independent dimerization of this oncogenic kinase, and allows fusion proteins self-phosphorylation and activation in tumour cells [27].

Indeed, aberrant NPM-ALK expression contributes to the pathogenesis of ALK-positive ALCL, as it causes cell transformation through activation of several biological pathways related to cell proliferation, cell-cycle control and apoptosis [28-31]. Consistent with these findings, NPM-ALK has been shown to affect growth, adhesion and migration of various cell types, which frequently display an enhanced motility and invasiveness due to ALK-dependent dysregulation of proteins controlling both the intracellular plasticity and the extracellular microenvironment. To do so, NPM-ALK binds to and activates a host of important cellular signalling molecules, including several with mitogenic and anti-apoptotic activities (RAS, ERK1/2, AKT, STAT3 and JUNB) [32-38]

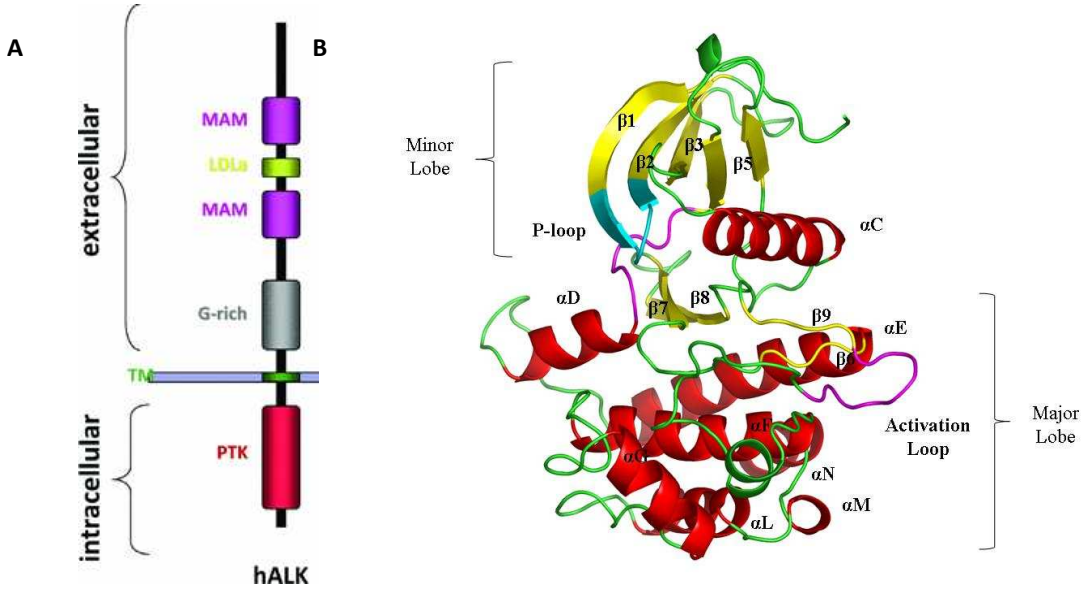


Figure 3 (**A**) **Domain structure of human ALK:** the N-terminal region comprises two MAM domains (aa 264-427 and 480-626), one LDLa domain (aa 453-471) and a glycin rich (G-rich) region (aa 816-940). A transmembrane(TM)-spanning segment connects the extracellular region with the protein tyrosine kinase (PTK) domain in the intracellular region (aa 1116-1383). Adapted from *Palmer et al.*, 2009.
(**B**) A representation of the apo crystal structure of ALK shown in orthogonal orientation.

1.3 ALK AS A THERAPEUTIC TARGET

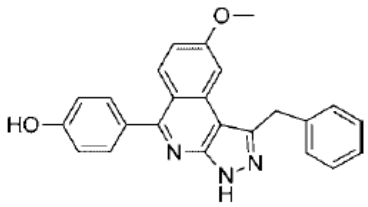
Similar to other T-cell lymphomas, the relative rarity of ALCL has limited the number of large prospective clinical trials to help optimize treatment, which remains, thus far, dependent on the use of combined intensive chemotherapy. Most European pediatric oncology groups use short-pulse chemotherapy regimes (high-dose methotrexate, cyclophospamide, vincristine, doxorubicine and corticosteroids) of 4-6 months duration to cure ALCL, and the failure rate at two years is almost 30% for most of these regimens. Notably, most of relapses (60%) occur within the first year from the start of therapy, and long-term survival for relapsed disease is less than 50% [39]. Therefore, although ALK-positive ALCL has a good prognosis compared to other childhood cancers, there is a continuous need for new therapeutic agents, particularly for relapsed cases.

Genetic aberrations of ALK are recurrent in several tumour types, such as neuroblastoma [40-43], non-small cell lung cancer (NSCLC) [44-47] and inflammatory myofibroblastic tumour (IMT) [48], and compelling preclinical data indicate that tumours harbouring gene amplification, translocations, or activating point mutations are partially or fully dependent upon ALK kinase activity for proliferation and survival [43, 44, 49]. Consistently, the inhibition of ALK catalytic activity yields potent anti-tumour efficacy *in vitro*, thus providing a sound rationale for clinical development of ALK small-molecule inhibitors (**Figure 4**) [38, 50-52].

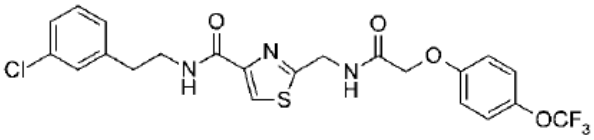
Proof of concept for this approach has been provided by clinical studies using Pfizer PF-2341066 inhibitor, also known as crizotinib, (ClinicalTrials.gov #NCT00585195), as ALK tyrosine kinase inhibitor (TKI), in refractory, heavily pre-treated NSCLC patients harbouring the EML4-ALK rearrangement. Among NSCLC patients, crizotinib was well tolerated, had rapid, durable response and correlated with improved survival [53, 54]. Consistent with these results, two recent case reports showed remarkable and brisk crizotinib activity in relapsed and refractory NPM-ALK positive ALCL as well [55], supporting the strong preclinical rationale of ALK inhibition in this type of tumour and in those others bearing ALK mutations/rearrangements.

However, data on acquired resistance to crizotinib due to secondary mutations in the ALK kinase domain (KD) have been also reported [56, 57], reflecting previous clinical experience with other TKIs. In this context, a wealth of clinical data demonstrate that resistance to drug inhibitory activity, like that observed for EGFR, c-Kit or Bcr–Abl oncogenic kinases, is often linked to secondary mutations in the KD, either when acquired *de novo* or resulting from selection of pre-existing, sub-dominant neoplastic

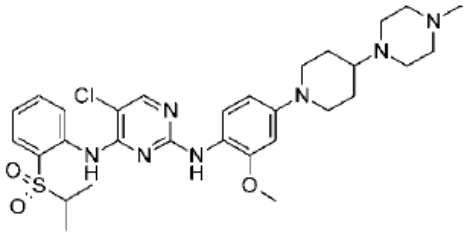
clones [58-61]. Of note, dominant mutations at relapse can be detected before the onset of treatment as well, suggesting that drug resistance may be predicted before treatment starts (intrinsic resistance) and risk of failures reduced. However, mutation frequency at diagnosis is often below the level of detectability of direct Sanger sequencing, requiring more sensitive methods for detection [62-64].



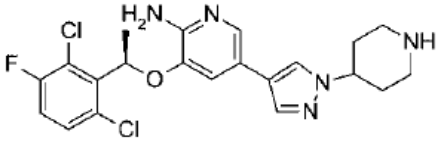
pyrazolo[3,4-c]isoquinoline



thiazoles



NVP-TAE684



PF-2341066 (crizotinib)

Fig.4 Chemical structures of some ALK inhibitors.

1.4 ULTRA-DEEP TARGETED SEQUENCING FOR LOW PREVALENCE MUTATIONS DETECTION IN CANCER

Sanger sequencing [65] has long been regarded as the gold standard for mutation detection since prior knowledge of mutations is not required and assay development is limited only by sequencing primer design and read length. This approach, however, is unreliable for detecting variants that account a small proportion (<20%) of the total population of genomes in a sample [66], as it generates only an average sequence of the PCR product. In addition, the analysis of samples with rarer heterogeneous insertion-deletion mutations remains challenging, since minor variant detection methods that rely on subcloning of PCR products, in conjunction with conventional sequencing, are limited by bacterial cloning artefacts and time constraints. On the other hand, nonsequencing-

based assays for minor variant detection such allele-specific PCR or probe-based methods [67] generally offer high sensitivity, but prior knowledge of the mutation of interest is required and linkage of the identified mutations is not possible. By contrast, within the last few years, massive parallel next-generation sequencing (NGS) technology has greatly enhanced the scope and the speed of genomic cancer research, making the detection of rare mutational events precise and possible in any tumour or normal tissue. The power of the new sequencing technologies and their utility for variant detection derive from the ability to sequence single molecules in massive amounts and give the frequency of a each single variant in a population of DNA molecules. In this process, each of the single molecules of an amplicon is clonally amplified and individually sequenced, allowing for the identification of rare variants over the whole read length. At present, the 454 pyrosequencing technology achieves the longest read lengths (400–500 bp reads) and enables the ultra-deep sequencing (UDS) of hundreds of thousands of molecules at high coverage [68]. The combination of high coverage and long read length has made the 454 pyrosequencing a promising tool for sensitive and quantitative detection of genomic variants, with potential applications in various clinically relevant research areas, including oncology and human genetics.

The number of somatic tumour mutations with potential utility for predicting treatment response is rapidly growing due to increasing numbers of targeted therapies. However, tumour samples can be heterogeneous due to invasion into stroma, infiltration by immune cells, and clonal evolution. Technological advances in DNA deep sequencing clearly offer an important solution to the problem of analyzing heterogeneous samples. One of the first studies on this field by Thomas et al. [69] reported the presence of low-abundance EGFR mutations in a relapsed lung adenocarcinoma specimen with low tumour content, for which conventional Sanger sequencing was not informative. Picotiter plate sequencing revealed the presence of a deletion mutation of 18 nt in exon 19 of EGFR at frequency of approximatively 3%, and a nucleotide substitution encoding the T790M mutation in exon 20 at a frequency of 2%, the latter one known to be associated with clinical resistance to EGFR inhibitors. The same 18 nt deletion was detected in a pretreatment sample in which tumour cells were exceptionally rare at a frequency of approximatively 0.3%, but no T790M mutation was revealed, consistently with the original erlotinib sensitivity and subsequent relapse. More recently, Choi et al. analyzed ALK kinase domain mutational status with a similar approach in a relapsed case of EML4-ALK-positive NSCLC, treated with crizotinib [56]. Two *de novo* point mutations were detected at frequencies of 41.8% and 14%, corresponding to the amino acid

changes C1156Y and L1196M, respectively. Ectopic expression of both mutant forms in BA/F3 cells determined a markedly reduced sensitivity to multiple ALK inhibitors, compared to cells expressing wild-type EML4-ALK, thus suggesting an association between these resistance mutations and patient relapse during clinical treatment with crizotinib.

In a more recent study on EGFR exon 19 deletions in NSCLC, UDS highlighted the presence of subpopulations of DNA molecules carrying EGFR deletions different from the main one in 43% of analyzed tumour samples, with a frequency between 0.1% and 17% [70]. These findings were correlated to a marked instability of a region in exon 19 of EGFR, to a different protein antigenicity and dismal outcome.

2. OBJECTIVES

Tyrosine kinase inhibitors are extremely effective drugs for the treatment of tumours associated to the expression and activity of mutated, oncogenic protein tyrosine kinases. However, clinical studies thus far indicate that a significant portion of patients acquire clinical resistance after achieving initial response, in most cases due to selection of cancer cells with mutations in the targeted PTK.

Based on these observations, we hypothesize that a good clinical response may be achieved in ALCL patients treated with selective small-molecule ALK inhibitors, but, similarly to other types of “kinase addicted cancers”, some of those patients are expected to develop drug resistance. Therefore, the identification of NPM-ALK mutants in ALCL primary tumours may be useful and mandatory to the rational design of more effective inhibitors.

The aim of this study is the ultra-deep mutational analysis of NPM-ALK kinase domain in a cohort of pediatric patients before the onset of treatment and the functional characterization of mutations identified, both in terms of oncogenic activity and sensitivity to ALK-specific inhibition.

OBJECTIVES

3. MATERIALS AND METHODS

3.1 PATIENTS, SAMPLES AND CELL LINES

A total of 37 tissue samples from ALK-positive ALCL patients enrolled between December 2000 and September 2010 in BFM-based AIEOP treatments protocols were included, after obtaining institutional board approval. Diagnosis was reviewed by the AIEOP (Associazione Italiana di Ematologia ed Oncologia Pediatrica) central panel of pathologist in all of the cases, and further characterized in our laboratory by means of RT-PCR using primers specific for t(2;5)(p23;q35) translocation. Reversed-transcribed RNA from a neuroblastoma patient carrying the R1275Q mutation was used as positive control (a kind gift of Dr. Luca Longo and Dr. Gian Paolo Tonini, Italian Neuroblastoma Foundation, Genova), while cDNA from human, non-pathologic brain tissue was included as negative control for amplicon library preparation and sequencing.

With respect to cell lines, neuroblastoma cells SK-N-SH, carrying the F1174L mutation, and monkey renal epitelian COS7 cells were grown in RPMI 1640 medium supplemented with 10% fetal calf serum (FCS) (Gibco, Life Technologies Co., Carlsbad, CA, USA), t(2;5)-positive ALCL Karpas299 cells were grown in 15% FCS RPMI 1640 medium, and the human embryonic kidney HEK-239T cells in DMEM medium enriched with 10% FCS. All culture media were supplemented with 2 mM glutamine (Gibco, Life Technologies Co., Carlsbad, CA, USA), 100U/ml penicillin and 100 µg/ml streptomycin (SIGMA-Aldrich Co., St. Louis, MO, USA)

3.2 REAGENTS AND ANTIBODIES

The dual ALK and c-MET inhibitor PF-02341066 (Crizotinib) was purchased from Selleckchem (Selleck Chemicals, Houston, TX, USA), dissolved in DMSO and stored at -20°C. Antibodies were purchased from Cell Signaling (anti-Phospho-STAT3^{Y705}, anti-Phospho-ERK1/2$^{T202/Y204}$, anti-Phospho-ALK^{Y1604}, anti-Phospho-ALK$^{Y1278/Y1282/Y1283}$, rabbit anti-*myc*) (Cell Signaling Technology, Inc., Danvers, MA, USA); Santa Cruz (anti-STAT3, anti-ERK1/2) (Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA); Invitrogen (rabbit anti-ALK, anti-V5, mouse anti-*myc*) (Invitrogen, Life Technologies Co., Carlsbad, CA, USA); SIGMA-Aldrich (anti-β-tubulin) (SIGMA-Aldrich Co., St. Louis, MO, USA). DAPI nucleic acid stain, and fluorophore-conjugated goat anti-rabbit Alexa488 and goat anti-mouse Alexa546 antibodies were bought from Molecular Probes (Molecular Probes, Life Technologies Co., Carlsbad, CA, USA). Horseradish peroxidase-conjugated sheep anti-

mouse, donkey anti-rabbit antibodies and protein G-sepharose Fast-Flow™ beads were purchased from GE Healthcare (GE Healthcare Life Sciences, Uppsala, Sweden). The BCA protein assay was from Pierce (Thermo Fisher Scientific, Inc., Waltham, MA, USA) while Western blot chemiluminescence reagents were purchased from PerkinElmer (PerkinElmer Inc., Waltham, MA, USA). Films were purchased from GE Healthcare, while nitrocellulose membranes were from Whatman (Whatman, GE Healthcare Life Sciences, Uppsala, Sweden). All other reagents, if not specified, were purchased from SIGMA-Aldrich.

3.3 STATISTICAL ANALYSIS

Survival analysis was performed according to Kaplan–Meier method. The 5-years overall survival (OS) was calculated from the date of diagnosis to the date of death from any cause or to the date of last follow-up. The 5-years progression-free survival (PFS) was calculated from the date of diagnosis to the date of the first event (tumor progression or relapse) or to the date of last follow-up. *P*-values are two-sided, with a type I error rate fixed at 0.05. Statistical analysis was performed by using the SAS statistical program (SAS-PC, version 9.3; SAS Institute Inc., Cary, NC, USA).

3.4 RT-PCR AND AMPLICON LIBRARY PREPARATION

Total RNA was isolated using TRIzol reagent (Invitrogen, Life Technologies Co., Carlsbad, CA, USA), following the manufacturer's instructions. An amount of 1 µg of total RNA was reverse transcribed using SuperScript II reverse transcriptase (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) and random hexamers. Briefly, samples were denatured at 75°C for 3 min and cooled in ice for 5 min. Subsequently, 10 µl of reaction mixture were added, containg 1X buffer, 2 mM of each dNTP (GE Healthcare Life Sciences, Uppsala, Sweden), 250 nM random hexamers and 40 U of RNAse inhibitor (Roche Applied Science, Penzberg, Germany). After 10 min at 20°C, 10 mM DDT and 200 U of SuperScript II reverse transcriptase were added. Retrotranscription was performed as follows: 60 min at 37°C, 5 min at 99°C, 5 min at 4°C. ALK KD coding region, corresponding to *ALK* exons 22-25, was amplified using special fusion primers, consisting in a 18-19 nt target-specific sequence on the 3'-end and a 19 nt fixed sequence on the 5'-end (adapter A on forward primer and adapter B on reverse primer), according to manufacturer's guidelines. A tag sequence (Multiplex Identifier, MID) of 10 nt specific

for each primer pair was added between the adapter and template specific sequences, to allow pooling of 22 different samples in the same sequencing run (**Figure 5**). Primer sequences are reported in **Table 1**. Each reaction mixture contained 1 µl of cDNA, 1X Buffer, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 200 nM of each primer and 1 U of Taq Gold (Applied Biosystems, Life Technologies Co., Carlsbad, CA, USA), in a final volume of 50 µl. PCR reaction consisted of initial denaturation at 95°C for 7 min, followed by 10 cycles of 94°C for 15 s, 65°C for 15 s, 72°C for 1 min, 25 cycles of 94°C for 15 s, 60°C for 15 s, 72°C for 1 min, and a final extension at 72°C for 10 min. Negative and positive controls used were a commercially available human brain cDNA library (Human MTC Panel I, Clontech Laboratories Inc., Mountain View, CA, USA) and Karpas299, SK-N-SH and a neuroblastoma tissue sample, respectively. PCR products were analyzed by 2% agarose gel electrophoresis and visualized under UV trans-illumination by SYBR® Safe staining (Invitrogen, Life Technologies Co., Carlsbad, CA, USA), using MassRuler™ Low Range DNA ladder (Thermo Fisher Scientific Inc., Waltham, MA, USA) as molecular weight standard. Amplicon products were, thus, quantified using the Quantity One software (Bio-Rad Laboratories Inc., Hercules, CA, USA) and pooled at an equimolar ratio. To remove excess of primers and other reaction components, the amplicon pool was run on agarose gel, and the sample band, after being excised, was purified using the QIAquick gel extraction kit (Qiagen Co., Hilden, Germany). Finally, the sample was diluted to a final concentration of 10^7 PCR fragment molecules/µl.

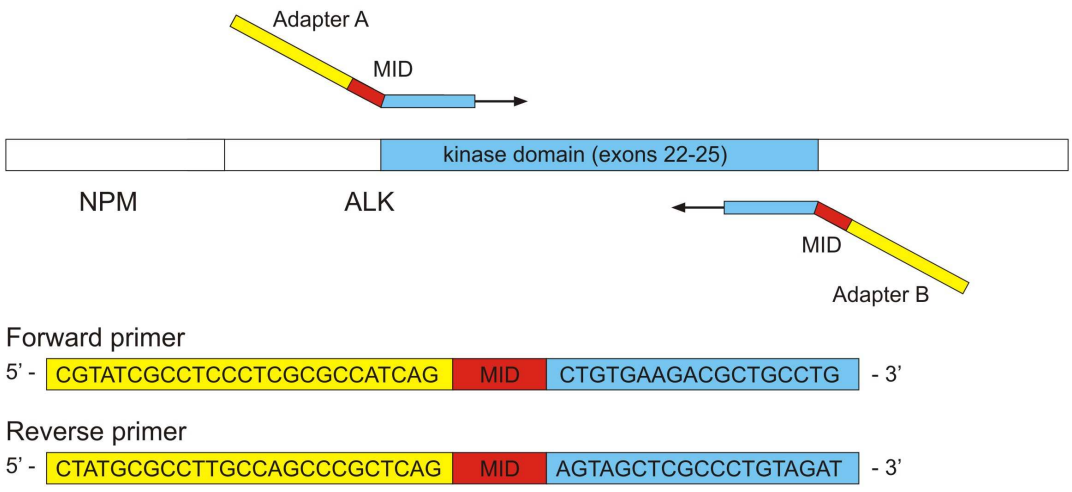


Figure 5 Schematic representation of fusion primers for the amplification and sequencing of ALK kinase domain.

3.5 NEXT-GENERATION ULTRA-DEEP SEQUENCING

Amplicon ultra-deep sequencing was performed at BMR Genomics (Padova, Italy), using the Roche Genome Sequencer (GS) FLX 454 (Roche Applied Science, Penzberg, Germany). The samples were analyzed in two separate runs of 22 amplicons each, using 1/8 lane of the GS FLX PicoTiterPlate. The amplicon-PCR-derived fragments were annealed to carrier beads and clonally amplified by emulsion PCR (emPCR), according to the manufacturer's protocol. Briefly, the beads were isolated and compartmentalized into droplets of an aqueous PCR reaction buffer in oil emulsion. Subsequently, the emulsions were broken by isopropanol to facilitate collection of the amplified fragments bound to their specific beads. The beads carrying single-stranded DNA templates were enriched, counted and deposited into the PicoTiterPlate for sequencing. A CCD camera-based imaging assembly was used to capture the pyrosequencing-derived light signal. A nucleotide sequence output read ("flowgram") was generated by integrating the light signal for each flown nucleotide and correlating the measured signal to dynamically determined thresholds for mono- and multi-mers of each nucleotide.

3.6 DATA ANALYSIS AND DETECTION OF VARIANTS

All data were generated using the GS Sequencer software version 2.5.3 (Roche Applied Science). Image processing and amplicon pipeline analysis was performed using default settings of the GS RunBrowser software version 2.5.3 (Roche Applied Science). Sequence alignments and variant detection was performed using the GS Amplicon Variant Analyzer (AVA) version 2.7 (Roche Applied Science), in combination with a blast-based pipeline for low frequent large insertions/deletions (INDELs) (CRIBI Genomics, University of Padova, Padova, Italy). AVA software filters were set to display sequence variances represented even by a single read.

The Human nucleophosmin-anaplastic lymphoma kinase fusion protein mRNA sequence (GenBank U04946) was used as reference sequence. Point mutations (single nucleotide polymorphisms, SNPs) were considered when present with a frequency of at least 0.5% in both forward and reverse reads, while INDELs were considered when validated by both software, regardless of frequency. To confirm calls made by both software, consensus alignments and individual flowgrams were manually reviewed.

INDEL consensus sequences were analyzed using the mRNA-to-genomic alignment program Spidey **(http://www.ncbi.nlm.nih.gov/spidey).**

3.7 HOMOLOGY MODELLING AND MOLECULAR DYNAMICS SIMULATION

ALK mutants were analysed through the MOE Protein Align tool, which is a modified version of the Needleman methodology. In this approach, substitution matrix (BLOSUM 62 in our case) are applied to pairs of aligned residues to produce residue similarity scores used to optimize sequence alignments; gap penalties of one sequence in respect to the other are also taken into consideration. The homology models of R308Ins8 and R308Ins12 were obtained through the MOE homology modeling tool, a method used to build and refine 3D models for protein sequences, based on homologous template protein structures. The two sequences were aligned as described before to the crystallographic template of human ALK (PDB code: 3LCT). The models have been generated using AMBER99 forcefield, in the presence of the ADP docked to the template active site. For each model a small database of intermediate models has been created; the final models were chosen through energetic parameters (GB/VI and Contact Energy) and minimized only to relieve steric strain (Medium Minimization).

The protonation state of the two ALK R308Ins8 and R308Ins12 tyrosine kinase models were evaluated with Protonate3D (T=300K pH=7) within MOE and Protonate within AmberTools 1.5. Subsequently, we used tLeap and Amber FF99SB to parameterized the protein models. We solvated them in TIP3P water boxes and added counterions (Na^+; Cl^-) to assure the neutrality of the molecular systems. ClickMD has been used as molecular dynamic platform for NAMD 2.9 minimization (100000 step, conjugated-gradient method), equilibration (0.5 ns, alpha carbon positional restrains) and production phase (100 ns NVT, P = 1 atm, T = 300 K) of the molecular systems through 100000 conjugated gradients method. ACEMD v2728 has been used as molecular dynamics engine on nVidia GeForce GTX680 computational platform. Finally, the analysis of the resulting trajectories was based on RMSD overtime, RaibowRMSD, heatmaps and distance analysis employing VMD 1.9.1, RMSD Trajectory Tools 2.01, RAINBOWRMSD and NRGPLOT.

3.8 SITE-DIRECTED MUTAGENESIS

The pcDNA3 plasmids containg wild-type and K210A NPM-ALK has been previously obtained from the original pSRα-tkneo-NPM-ALK plasmid (a kind gift from Dr. S. Morris, S. Jude Research Hospital, Memphis, TN, USA). Site-directed mutagenesis was performed using the Phusion site-directed mutagenesis kit (Thermo Fisher Scientific Inc.,

Waltham, MA, USA) and the pcDNA3-NPM-ALK as a template. Deletional (Del) and insertional (Ins) mutants were generated using specific HPLC-purified 5'-phosphorilated oligonucleotides and 500 pg of template DNA, according to the manufacturer's instructions. PCR cycling conditions were the following: 30 s at 98°C, followed by 25 cycles of 10 s at 98°C, 30 s at the specific annealing temperature, 3 min at 72°C, and a final extension of 10 min at 72°C. Primer sequences and respective annealing temperatures for each INDEL mutation are listed below.

Del(923-924):	FW	5'P-GACATTGCTGCCAGAAACTGC-3'	55°C
	REV	5'P-GGTGGATGAAGTGGTTTTCCTCC-3'	
Del(826-924):	FW	5'P-TATTTGGAGGAAAACCACTTCATCCAC-3'	55°C
	REV	5'P-CGGGCGAGGGCGGGT-3'	
Del(696-825):	FW	5'P-AGCCAGCCCTCCTCCCTGGCC-3'	57°C
	REV	5'P-CTGATGATCAGGGCTTCCATGAGG-3'	
Ins24(924):	FW	5'P-**CTTTCTTCCCAG**AGACATTGCTGCCAGAAACTGC-3'	62°C
	REV	5'P-**GAAATGCATTTC**CGGTGGATGAAGTGGTTTTCCT-3'	
Ins36(924):	FW	5'P-**CATTTCCTTTCTTCCCAG**AGACATTGCTGCCAGAAACTG-3'	65°C
	REV	5'P-**CATTTCCTAATTTTATCC**CGGTGGATGAAGTGGTTTTCCT-3'	

An amount of 2 µl of the PCR product was circularized with Quick T4 DNA Ligase, according to the manufacturer's recommendations, in a final volume of 10 µl. Competent *E. coli* DH5α cells (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) were transformed with 2 µl of the ligation reaction product, and serial dilutions of transformed bacteria (200 µl, 100 µl and 50 µl) were plated on LB agar plates containing 50 µg/ml ampicillin. LB agar plates were incubated overnight at 37°C, and at least 4 colonies for each transformation were isolated and grown overnight in LB containg 100 µg/ml ampicillin, with constant shaking (225 rpm). Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (Qiagen Co., Hilden, Germany) according to the manufacturer's protocol, and stored at -20°C until used. All INDEL mutations were confirmed by sequencing of both strands.

3.9 SUBCLONING

The pBudCE4.1 expression vector (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) was chosen to express simultaneously wild-type and mutated NPM-ALK. Wild-type (WT) NPM-ALK was subcloned into the EF-1α multiple cloning site (MCS) and fused to the V5

epitope, while NPM-ALK mutants were subcloned into the CMV MCS and fused to *myc* tag (**Figure 6**). However, to do that, a silent point mutation was introduced at nucleotide 1887 (GT**A** → GT**T**) of wild-type NPM-ALK to remove an internal KpnI restriction site (GGTACC) and allow subcloning into pBudCE4.1 expression vector, using 700 pg of template DNA and the Phusion site-directed mutagenesis kit. Oligonucleotides were the following: FW 5'P-CAATATGAAGGAGGT**T**CCTCTGTTCAGGC-3'; REV 5'P-GCAGTCAGCGAAGAGGGCTCTA-3'. Amplification reaction was performed as described previously, with an annealing temperature of 69°C. In addition, stop codons and untranslated sequences of out-of-frame (OOF) deletional mutants were removed before cloning into pBudCE4.1 plasmid, using the same mutagenesis approach. A common forward primer was used (5'P-ATCGATACCGTCGACCTCGAG-3'), whereas 3 different reverse primers were designed (5'P-GGGCCCAGGCTGGTTCATGCT-3' for WT, Del(826-924) and Ins24(924); 5'P-TAGTAGCTCGCCCTGTAGATGTCTCG-3' for Del(923-924); 5'P-AGAGGCAGTTTCTGGCAGCAAT-3' for Del(696-825)) in the same amplification conditions: 30 s at 98°C, 25 cycles of 10 s at 98°C, 30 s at 68°C, 3 min at 72°C, and a final extension of 10 min at 72°C. Ligation, transformation and plasmid preparation were performed as indicated in paragraph 3.7. All mutations were confirmed by bidirectional sequencing. Briefly, digestion was performed using 2 µg each of pcDNA3-WT and pBudCE4.1 expression vectors, 20 U of KpnI and 40 U of XhoI (New England Biolab Inc., Ipswich, MA, USA), in a final volume of 20 µl. Linearized pBudCE4.1 was directly purified with the QIAquick PCR purification kit (Qiagen Co., Hilden, Germany) and eluted in 50 µl of TE buffer, whereas digested pcDNA3-WT was loaded in a 2% agarose gel, the insert band excised, purified with the QIAquick Gel extraction kit and then recovered with 30 µl of TE buffer. The ligation reaction contained 1X buffer, 2 µl of T4 DNA ligase (New England Biolabs Inc., Ipswich, MA, USA) and a molar ratio of 1:3 of insert to vector, in a final volume of 20 µl. Samples were incubated overnight at 16°C and used to transform competent *E. coli* DH5α, following the manufacturer's protocol. An amount of 200 µl, 100 µl and 50 µl of each transformation were plated on Low Salt LB agar plates containing 25 µg/ml zeocin (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) and incubated overnight at 37°C. At least 4 colonies for each transformation were isolated and grown overnight in Low Salt LB medium supplemented with zeocin (25 µg/ml) and under constant shaking (225 rpm). Correct insertions were verified by restriction enzyme digestion with KpnI and XhoI.

An additional step of site-directed mutagenesis was performed on pBudCE4.1/WT to establish the correct reading-frame between NPM-ALK and the V5 tag. PCR amplification

was done as previously described, using specific primers (FW 5'P-CCGTCGACCTCGAGAGATCTG-3'; REV 5'P-GGGCCCAGGCTGGTTCATGCT-3') and an annealing temperature of 69°C.

The mutated NPM-ALK inserts were excised with 20 U of HindIII and 20 U of XbaI (New England Biolab Inc., Ipswich, MA, USA) and the insert bands purified from agarose gel as previously indicated. Empty pBudCE4.1 and pBudCE4.1/WT were digested and linearized as reported above. Each mutated insert was ligated into both plasmids, following the same ligation and transformation protocol as for WT, and validated by restriction enzyme digestion with XbaI and HindIII.

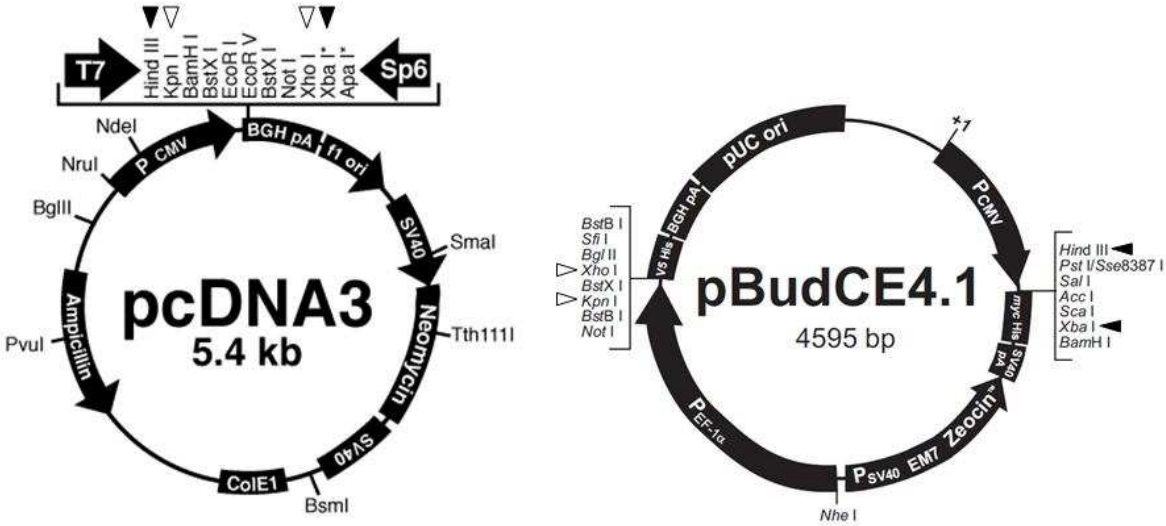


Figure 6 Map of the expression vectors pcDNA3 and pBudCE4.1.
Restriction sites for wild-type insert are indicated with open arrowheads, restriction sites for mutated inserts with closed arrowheads.

3.10 TRANSIENT TRANSFECTION AND TREATMENTS

To evaluate the effects of NPM-ALK INDEL mutations generated, HEK-239T cells were transiently transfected with wild-type NPM-ALK and/or INDEL mutants. Three days before transfection, 0.8 x 10^6 cells were plated in 100 mm petri dishes and let grown under standard cell culture conditions. Exponentially growing cells were then transfected with 20 µg of each plasmid and 50 µl of Lipofectamine™ 2000 (Invitrogen, Life Technologies Co., Carlsbad, CA, USA) following the manufacturer's instructions, before lysis and protein detection. NPM-ALK expression and activity was analyzed in the presence or absence of the ALK-specific inhibitor crizotinib, in a time- (0, 30 min, 1, 2, 4

and 6 h) and dose- (0.5, 1.0, 5.0 μM) dependent manner. To assess localization of wild-type and mutant NPM-ALK proteins, COS7 cells (0.2 x 10^5) were plated on 8-well chamber slides, transfected with 0.5 μg of the above reported plasmids plus 1.25 μl of Lipofectamine™ 2000 and processed for immuoflorescence.

3.11 CELL LYSIS, IMMUNOBLOTTING AND IMMUNOPRECIPITATION

The cells were washed twice in ice-cold 1X phosphate-buffered saline (PBS) and lysed by addition of TritonX-100 sample buffer (10 mM Tris-HCl [pH 7.5]; 130 mM NaCl; 1% TritonX-100; 5 mM EDTA; 1 mg/ml BSA; 20 mM sodium phosphate [pH 7.5]; 10 mM sodium pyrophosphate [pH 7.0]; 25 mM glycerophosphate; 1 mM sodium orthovanadate; 10 mM sodium molybdate; 1 mM PMSF; 20 μg/ml leupeptin; 20 μg/ml aprotinin). Cell lysates were clarified by centrifugation (at 4°C) at 14000 × g for 30 min and protein concentrations were determined by the BCA assay using BSA as a standard. Binding of wild-type NPM-ALK to each mutant was performed by incubating protein extracts with 1–2 μg of specific antibodies (anti-V5, anti-*myc* and anti-Phospho-ALK) at 4°C overnight. Immunocomplexes were adsorbed to 20 μl of Protein G-Sepharose beads for 2 h at 4°C. The immunoadsorbed pellets were washed 4 times with 1% Triton X-100 lysis buffer and heated at 95°C in 1X reducing Laemmli loading buffer. Aliquots of cell lysates (50 μg) and immunoprecipitates were diluted in 5X reducing Laemmli loading buffer [2% SDS, 100 mM DTT, 60 mM Tris (pH 6.8), 0.01% bromophenyl blue, and 10% glycerol], fractionated by 10% SDS-PAGE and electrotransferred to nitrocellulose membranes. Proteins were visualized by chemiluminescence. Films were scanned and analyzed by using image analysis software (NIH Image, National Institute of Health, Bethesda, MD, USA)

3.12 IMMUNOFLUORESCENCE

Transfected HEK-293T cells were fixed in 3.7% paraformaldehyde, permeabilized in 0.2% TritonX-100 and blocked in 100 mM glycin, followed by incubation in 10% FCS-PBS. Permeabilized cells were incubated for 1 h at 37°C with the specific primary antibodies (mouse anti-V5 and rabbit anti-*myc*) and, washed in PBS, incubated with fluorophore-conjugated secondary goat anti-rabbit Alexa488 and goat anti-mouse Alexa546 antibodies, at a 1:1000 dilution of 2 mg/ml stock. Cells were then washed in PBS, mounted onto slides with 1:1 PBS/glycerol, with the addition of DAPI nucleic acid stain

(1:1000). The cells were observed at 63X/0.75 NA in a Leica DMBL microscope. Images were acquired with a Leica DC 300F digital camera and prepared for reproduction with Leica IM1000 software (Leica Mycrosystem Ltd., Germany).

4. RESULTS

4.1 IDENTIFICATION OF NPM-ALK KD MUTATIONS

To detect somatic tumour mutations that may predict treatment response in ALK-positive ALCL patients, *ALK* kinase domain mutational status was investigated in 37 tumour specimens by 454 amplicon ultra-deep sequencing, using a healthy human brain cDNA library and neuroblastoma tumour cells as negative and positive controls, respectively. About 99,000 sequences aligned with *ALK* exons 22-25 were obtained (mean of 2,234 ± 734 sequences per sample) and an overall of 686 sequence variants were detected (300 SNPs and 386 INDELs). To ensure maximum detection sensitivity, filters were set to display sequence variances represented even by a single read. However, to discard false positive variants, SNPs consensus alignments and individual flowgrams were manually inspected for proper variant calling and validated if represented with the same frequency (≥0.5%) on both forward and reverse reads, whereas INDELS were considered if validated by at least two independent alignment algorithms. In this context, a total of 7 SNPs were confirmed, 5 of which represented silent point mutations, while two were missense mutations at R1275 and R1231 positions in patients 17 and 36, respectively (**Figure 7A** and **B**). As expected, ALK point mutations were found in neuroblastoma tumour cells (ALK F1174L in SK-N-SH NB cell line, and ALK R1275Q in a NB tumour sample) but not in the healthy human brain cDNA library.

Since most of dedicated programs are able to rapidly analyze all the reads reporting SNPs and small INDELs, but they often fail to manage large insertions/deletions, we validated INDELs by combining Roche AVA software and a high sensitive blast-based pipeline developed by CRIBI Genomics. In this respect, among the 386 INDELs detected by Roche AVA software, 65 were large gaps, including 9 deletions and 3 insertions that were detected by the CRIBI Genomics pipeline as well. Of these 12 INDELS, a deletion of 130 nt corresponding to the entire *ALK* exon 23, Del(696-825), was found in all patients with a frequency ranging from 1 to 12.8%, while the remaining 11 were expressed in a variable number of patients at low frequency (~0.5%) (**Table 2)**. Of note, a common INDEL, previously reported in neuroblastoma tumor cells [71], (Del(923-924), freq. 1.4-4%), was detected by Roche AVA software only. However, further sequencing analysis demonstrated that this deletion was present in the healthy human brain cDNA library as

well (13.9 %), supporting the hypothesis of a potential splicing error rather than a novel somatic variant in ALCL transformed cells. Indeed, to determine if INDEL mutations detected could represent NPM-ALK alternative spliced transcripts, each consensus was aligned to the respective genomic sequence using the mRNA-to-genomic alignment program Spidey, and in 9 of 13 variants, including Del(923-924), canonical splicing sites were identified at the boundaries of deleted or inserted regions. In 5 of these, the mutation resulted in out of frame (OOF) transcript products, corresponding to truncated proteins.

4.2 MOLECULAR MODELLING ANALYSIS OF NPM-ALK INDEL MUTATIONS

To investigate the effect of INDEL mutations detected *in vivo* on the structure and activity of NPM-ALK, molecular homology modelling analysis was performed. In 10 of 13 variants, the mutations were most likely disruptive of protein function, being associated to extensive protein truncation or deletion of crucial regions of the kinase domain, while in 3 variants the structural rearrangements were compatible with ALK enzymatic activity. Therefore, we decided to concentrate our studies on the two deletions that were most common in our patients (Del(923-924) and Del(696-825)) and on the 3 in frame mutations that conserved most of the residues and motifs critical for ATP-binding and hydrolysis (Del(826-924), Ins24(924) and Ins36(924)). As for the common deletions, Del(923-924) caused a translational frame shift at the catalytic Asp309 (D309H) and originated an inactive protein of 342 aa (D309H-OOF) (**Figure 8** and **9A**), while Del(696-825) resulted in a truncated protein of 273 aa (S232R-OOF) which presented an extensive deletion in correspondence of both the activation and the catalytic loops (**Figure 9B)**. In addition, the DFG motif, a catalytic site responsible for the activation of the ALK kinase, was replaced or lost in the two D309H-OOF and S232R-OOF NPM-ALK mutants, respectively, as the 3 major autophosphorylation sites (Tyr338, Tyr342 and Tyr343) were not conserved. In contrast, Del(826-924), Ins24(924) and Ins36(924), were characterized by more restrained structural modifications, which were predicted to maintain ALK kinase activity. In particular, Del(826-924) corresponded to an internal deletion of 23 aa, which, though relevant for the proper kinase conformation, did not affect adjacent regions participating in catalytic processes, (**Figure 9C**), while Ins24(924) and Ins36(924) corresponded to the addition of 8 or 12 aa between Arg308 and Asp309, respectively (R308Ins8/12, **Figure 8** and **9D**). These two insertions occur within the

activation loop, but maintained the catalytic residues (Asp309) and motifs (glycin-rich and activation loop) that are important for ensure ALK active conformation.

4.3 *IN VITRO* EXPRESSION AND ACTIVITY OF NPM-ALK INDEL MUTANTS

To confirm molecular modelling analysis and to test the effects of the mutations on NPM-ALK kinase activity, we introduced either insertions (D309Ins8, D309Ins12) or deletions (Δ276-299) into NPM-ALK cDNA and expressed the mutant proteins in HEK-293T cells. Mutated constructs were generated by site-directed mutagenesis on pcDNA3-NPM-ALK and used to transfect exponentially growing cells. NPM-ALK mutants were therefore assessed for protein expression and activity, and compared to both wild-type and kinase-dead NPM-ALK (NPM-ALK WT and K210A, respectively) (**Figure 10**). Notably, although both insertion and deletion mutants were effectively expressed, none of these showed detectable levels of phosphorylation (p-NPM-ALK) or were capable of signalling through the activation of downstream STAT3 and ERK1/2 effector proteins (**Figure 10,** p-STAT3 and p-ERK1/2), demonstrating that even the most conservative INDEL mutations were kinase inactive. These results were unexpected, particularly for insertional mutations, which were shown to be structurally well tolerated. Thus, to provide a molecular explanation to these findings, we performed molecular dynamics simulation on R308Ins8 and R308Ins12 NPM-ALK mutants, and compared results with the native ALK kinase. Insights from kinase domain modelling were generated by comparing time-dependent conformational changes of ALK kinase core (β-sheet; α-helix; hinge region; ATP-binding site) and catalytic sites (Glycin-rich ATP-binding loop, or P-loop; Activation loop, or A-loop; DFG motif) during protein folding. While wild-type P-loop showed several conformational changes during the selected time frame (**Figure 11A**), R308Ins8 exhibited a different profile, with an A-loop conformational change simultaneous to the "Ins8" rearrangement (**Figure 11B**). A novel polar interaction between R316 of the inserted sequence (Ins8) and D338 (DFG motif) of the ALK catalytic loop was described (**Figure 11D**), and an inhibitory influence on the whole catalytic process was proposed. As for R308Ins12, the largest conformational changes during ALK "Ins12" MD simulation were in the P-loop and in the "Ins12" moieties (**Figure 11C**). Two novel interactions that "lock" the P-loop and make more difficult the entrance of the ATP were demonstrated, the formation of a salt-bridge interaction between D220 (αC-helix) and K313 ("Ins12"), and a H-bond between the backbone of A186 (P-loop) and

the backbone of K313 (**Figure 11E**), suggesting that retention of intron sequences as a result of splicing errors may have negative effects on ALK kinase activity as much as skipping of exons or part of them.

4.4 COESPRESSION OF NATIVE NPM-ALK AND INDEL MUTANTS

In cells coexpressing native and catalytically inactive mutant NPM-ALK, the formation of heterocomplexes may have an impact on oncogene transforming activity, as well as on the sensitivity of NPM-ALK to specific inhibition. To functionally test this hypotesis, we coexpressed wild-type and the inactive mutants previously described: D309H-OOF and S232R-OOF truncated proteins, Δ276-299 and R308Ins8 in frame INDELs.

To ensure the same expression efficiency, we subcloned wild-type NPM-ALK and each mutant protein into the same pBudCE4.1 expression vector, which contains two promoters for high-level, constitutive, independent protein expression and a C-terminal tag peptides (*myc* or V5 epitopes) for protein detection and purification in single cells (**Figure 12A**). We performed immunoflorescence microscopy analysis to assess native and mutant NPM-ALK localization, which confirmed that in trasfecfected cells wild-type and mutant NPM-ALK proteins colocalize (**Figure 12B**). With coexpression, a kinase inactive mutant may exhibit a dominant-negative effect on the active allele, or may interfere with protein's sensitivity to drug-induced inhibition. Therefore, mutant NPM-ALK constructs were expressed alone in HEK-293T cells, and protein steady-state was assessed in the presence or absence of active NPM-ALK kinase (**Figure 13**). As shown in figure, *myc*-tagged INDELs were expressed in HEK-293T cells independently of NPM-ALK expression (**Figure 13A**, *arrowheads*), and differences between mutants were mostly related to the severity of the modifications carried. A lower expression efficiency was indeed observed for both D309H-OOF and S232R-OOF mutants compared to Δ276-299 and R308Ins8, which were detected at higher levels either alone or in combination with wild-type NPM-ALK (**Figure 13A**, *left panel*). Truncated protein S232R-OOF, which lacks almost entirely the tyrosine kinase domain, was even barely detectable when expressed alone, while it was evident when coexpressed with constitutively active NPM-ALK (**Figure 13A**, *right panel*, samples 11 and 17, respectively). Notably, the mutants did not shown detectable tyrosine phosphorylation both when expressed alone or in combination with wild-type NPM-ALK, but exerted a dominant-negative effect with respect to wild-type NPM-ALK autophosphorylation (**Figure 13B**, *right panel*, α-pALK). This did not perturb protein steady-state (**Figure 13B**, *left panel*,

α-V5), but significantly affected NPM-ALK signalling (p-STAT3) in transfected HEK-293T cells (**Figure 13C**, *right panel*).

To test whether the reduction of NPM-ALK activity was due to the formation of inactive WT/INDEL heterocomplexes, fresh cell lysates were immunoprecipitated using anti-V5 and -*myc* antibodies, and reciprocal immunoblottings were performed. As expected, *myc*-tagged NPM-ALK mutants coprecipitated with V5-tagged native NPM-ALK (**Figure 14A**, *left panel*, arrowheads), likewise wild-type NPM-ALK bound to *myc*-tagged INDELS (**Fig.14A**, *right panel*, arrowhead). However, when a site-specific anti-phosphoALK antibody was used, we found that purified immunocomplexes contained only V5-tagged wild-type NPM-ALK proteins and not *myc*-tagged INDELs (**Figure 14B**, *right* and *left panel*, respectively). Besides, higher levels of phosphorylated NPM-ALK was found in cells expressing only the native protein (**Figure 14B**, *lower panel*, samples 13-18), indicating that there was a dominant-negative effect of INDEL mutants coexpression with respect to NPM-ALK activation and function.

Finally, to investigate whether NPM-ALK mutant proteins may affect NPM-ALK sensitivity to inhibition (**Figure 15A**), HEK-293T cells expressing wild-type NPM-ALK, either alone or in combination with NPM-ALK INDELs, were exposed to increasing concentrations of the ALK-specific inhibitor crizotinib (0.5, 1 and 5 μM) for 6 h, and steady-state of both total and phosphorylated NPM-ALK was assessed (**Figure 15B**, NPM-ALK and p-NPM-ALK, respectively). Unexpectedly, immunoblotting experiments showed a significantly higher sensitivity of wild-type NPM-ALK to crizotinib in cells cotransfected with mutated NPM-ALK constructs, as coexpression led to a complete inhibition of NPM-ALK autophosphorilation (p-NPM-ALK) even at the lowest dose of 0.5 μM, while some NPM-ALK activity was still detectable in cells expressing the functional protein alone (**Figure 15B**). To further elucidate the kinetics of NPM-ALK inhibition, transfected cells were thus treated with crizotinib and time-dependent NPM-ALK inhibition was assessed. As shown in **Figure 15C**, a markedly reduced NPM-ALK phosphorylation was observed after 30' of exposure to the inhibitor (*graph*). However, a progressive recovery of kinase activity was noticed in HEK-293T cell lysates containing WT or WT and S232R-OOF NPM-ALK protein, providing further evidence of a stronger ALK kinase activity when NPM-ALK is expressed alone or together with unstable mutants (**Figure 15C**, *immunoblotting*). In the same experimental conditions, NPM-ALK-induced phosphorylation of STAT3 was also reduced by crizotinib (**Figure 15D**, *graph* and *immunoblotting*). However, when INDEL mutants were coexpressed, drug-dependent inhibition of NPM-ALK signalling was

more evident (50% p-STAT3 inhibition for WT; 20% for WT/mutant), most likely because of the less abundant formation of functional homodimers.

4.5 PATIENTS CHARACTERISTICS AND PROGNOSTIC IMPACT OF INDEL MUTATIONS

Survival analysis was conducted on 36 ALK-positive ALCL patients, 23 males and 13 females, with median age at diagnosis of 9 years (range 3.6 months to 17.5 years). Patient 26 was affected by ALCL as a secondary lymphoma, and therefore excluded from the analysis. Most of the cases were common type ALCL (41%) and, based on St Jude classification, 92% of them were stage III-IV. The median follow-up of the patients was 5.4 years (range 1.4-11.8). Twelve patients suffered from disease progression or relapse after a median time from diagnosis of 6.4 months (range 1.5-16.8 months). Four patients died, one due to disease progression, two from complications arising following blood stem cell transplantation as second line treatment, and one from septicemia. The 5-year PFS and OS of the whole study group were 66% (s.e. 8%) and 88% (s.e. 6%), respectively. **Table 3** summarizes the main clinical characteristics of the study population, reporting the frequencies of the 2 common INDELs, Del(923-924) (median 0.71%, range 0.03-3.97%) and Del(696-825) (median 2.35%, range 0.14-12.83%), and the absolute number of INDELs (median 4, range 2-8) identified in each patient. ALK mutational status was investigated at both diagnosis and relapse in 3 patients (19, 22 and 27). The prognostic impact of Del(923-924) and Del(696-825) was evaluated using the median frequency for clinical risk stratification. The 5-year PFS was 76% (s.e. 10%) and 56% (s.e. 12%) for patients with high (> 0.7%) and low (≤ 0.7%) frequency of Del(923-924), respectively (P = 0.19, **Figure 16A**). As for Del(696-825), the 5-year PFS was 65% (s.e. 12%) and 66% (s.e. 11%) for patients with high (> 2.4%) and low (≤ 2.4%) frequency of exon 23 deletion, respectively (P = 0.84, **Figure 16B**). In comparison, based on the absolute median number of INDELs, PFS was 74% (s.e. 9%) for patients with 4 or less INDELs and 50% (s.e. 14%) for patients with more than 4 INDELs (P = 0.11, **Figure 16C**). Alternative splicing variants associated to non-canonical splicing sites were not included in this analysis. In summary, no clear relationship between presence of NPM-ALK INDELs and relapse was observed. However, since patients were not treated with an anti-ALK single-agent therapy, it remains unknown whether the propensity of ALCL cells to generate splice variants of the *NPM-ALK* gene may influence prognosis of ALCL treated with ALK inhibitors.

5. DISCUSSION

Although ALK-positive ALCL have a rather benign prognosis, around 30% of patients do not respond or relapse after standard therapy, with a long-term survival of less than 50% in relapsed cases. Moreover, intensive multi-agent chemotherapy is associated with considerable toxicity, a problem particularly relevant in pediatric patients [72-75].

It has been extensively demonstrated that aberrant ALK activity is one of the major oncogenic events non only in ALK-positive ALCL [28-31], but also in neuroblastoma, NSCLC and IMT bearing ALK activating mutations/rearrangements[40-48], and the inhibition of ALK kinase activity was proven to substantially reduce cancer cell proliferation and invasiveness both *in vitro* and *in vivo*[38, 50-52]. Consistently, successful clinical experience in NSCLC carrying the EML4-ALK rearrangement treated with crizotinib [53, 54] suggests that ALK-specific inhibition would be a valuable treatment strategy in ALK-positive ALCL, as well as in other ALK-addicted tumours. However, treatments failures after crizotinib therapy due to newly acquired secondary mutations have been reported in both NSCLC and IMT, due to selection of cancer cells with mutations in the catalytic kinase domain of ALK that impair drug binding affinity [56, 57]. Therapy resistance due to mutations was previously observed in the KD of Bcr-Abl, c-Kit or EGFR, in patients treated with small-molecule inhibitors [58-61, 76-79]. In this context, although a robust clinical response of ALCL patients to an ALK inhibitor is expected, some of those patients are also anticipated to develop resistance. Therefore, the knowledge of NPM-ALK mutational status in ALCL primary tumours may represent an important information to design efficacious anti-ALK inhibitors, as KD mutations present at low levels in treatment-naive patients may be selected as dominant mutations at relapse, if conferring a growth advantage to targeted therapy [62-64].

With the aim to discover functionally relevant NPM-ALK mutations, we performed ultra-deep sequencing analysis on ALK exons 22-25, corresponding to the entire KD coding region, in 37 ALCL pediatric patients. Two point mutations were identified in two distinct cases: a G→A transition at nucleotide 1004 of NPM-ALK cDNA and a G→A substitution at nucleotide 872, which corresponded to the R1275Q and R1231Q mutations, respectively. R1275Q has been already reported as one of the most frequent activating mutations in neuroblastoma [40, 41, 43] and the molecular mechanism behind its constitutive activation has been recently elucidated [80]. The introduction of such an activating mutation into a constitutively activated kinase like NPM-ALK may not be

critical for both protein's transforming activity and sensitivity to specific inhibition, as proven recently in neuroblastoma and EML4-ALK-positive NSCLC [81, 82]. In contrast, the R1231Q amino acid substitution represents a novel ALK point mutation, which to our knowledge has never been reported neither in full-length ALK nor in the various truncated variants. This mutation is localized on the subdomain VIA of the C-terminal lobe, close to the A1234T mutation reported in neuroblastoma [43]. Differently from R1275Q and F1174L substitutions, which are associated to ligand-independent ALK activation and signalling, A1234T is a point mutation that, although affecting critical domains of ALK, is not oncogenic "*per se*" and needs to be validated as potential therapeutic target. Consistently, ALCL patients harbouring R1275Q or R1231Q point mutations were not characterized by worse prognosis. However, these patients were treated according to a standard chemotherapy regimen, and drug efficacy is expected to be independent from oncogenic ALK mutational status. Consequently, selection of cancer cells bearing these subtle point mutations can not be excluded if any ALK-specific therapy is used, and thus the molecular implications of both R1275Q and R1231Q point mutations on NPM-ALK function and sensitivity to inhibition need to be further investigated.

Although our principal aim was the identification of KD mutations potentially relevant for acquired drug resistance, our approach of cDNA-based ultra-deep sequencing allowed us to obtain information on NPM-ALK alternative-splicing events as well. Alternative spliced transcripts have been extensively reported in patients with Bcr-Abl-positive chronic myeloid leukemia [83-93] and, more recently, for ALK receptor in neuroblastoma cells [71, 94], while NPM-ALK alternative splicing events have never been reported so far. We identified, for the first time, 9 NPM-ALK INDEL mutations, resulting from KD whole exons skipping or partial intron retention. In particular, Del(923-924), identified in all of the patients, was determined by the recognition of the first two bases of exon 25 as intron 24 acceptor splicing site. Assuming there are no other alterations outside the sequenced domain, this deletion resulted in a truncated NPM-ALK protein of 342 aa, which was shown to be catalytically inactive. Consistently, this deletion has been reported in a neuroblastoma patient, alongside the whole deletion of exon 27 [71], but any obvious correlation between the genetic aberration and the protein expression was not found. However, we detected Del(923-924) in a healthy brain tissue pool, where ALK kinase activity is thought to have a role in the developing central and peripheral nervous system [95]. We therefore suggest that

misrecognition of intron 24 acceptor splicing site may represent a frequent event, whose functional implications are not of significant relevance for both physiological and oncogenic ALK function. The second INDEL mutation reported in all the patients is a whole exon 23 deletion, Del(696-825) which leads to a truncated inactive protein of 273 aa. The same deletion has been reported in neuroblastoma, both as solitary deletion or combined to exon 25 skipping [71]. We found that this deletion corresponded to a truncated protein product, S232R-OOF, that was clearly detectable when coexpressed with wild-type NPM-ALK, but not alone. Whether such a different behaviour may be related or not to protein stability it was not established, but the importance of the structural modifications carried supported this hypothesis. The great majority of the remaining variants identified were catalytically inactive as well, since extensively deleted or truncated. However, in 3 cases we hypothesized that some catalytic activity could be preserved. The first was a 69 nucleotide deletion, resulting from a splicing event in which nucleotides 893-894 AG were recognized as an alternative acceptor splice site. This mutation was associated to an in-frame translated protein (Δ276-299), lacking 23 aa of the subdomain VIA on the C lobe. The other two were both characterized by a partial intron retention (24 and 36 nucleotides of intron 24, respectively), resulting in an insertion of 8 and 12 aa close to the catalytic Asp309. However, all these mutants were catalytically inactive, and the molecular mechanism leading to their inactivation was elucidated by molecular dynamics observations. Therefore, many different NPM-ALK alternative transcripts were identified, but all were associated to the loss of kinase activity. This is consistent with the observation that NPM-ALK INDELs have not a prognostic impact on ALCL patient's risk stratification, and a correlation between the presence of INDELs and outcome coulc not be detected. However, it's worth nothing that INDELs have a clear dominant-negative effect with respect to ALK kinase's activity and drug sensitivity, as they associate and form nonfunctional complexes with their wild-type counterpart. This may have profound implication with respect to ALCL drug responsiveness, as INDELs can reduce the overall kinase activity and increase sensitivity to ALK specific inhibition, otherwise they might bind and prevent ATP-competitive molecules from inhibiting functional NPM-ALK kinase and ALCL cells growth. Indeed, when coexpressed with wild-type NPM-ALK, INDELs do interact with potentially active monomers and prevent NPM-ALK activation.

In summary, our results demonstrate that NPM-ALK KD point mutations are extremely rare in newly diagnosed ALCL patients, but may become predominant if positive

selection of mutated cells is induced by novel ALK-targeting therapies. Similarly to Bcr–Abl in CML, alternative splicing of *NPM-ALK* fusion gene represents a common event in ALCL, though the majority of patients are likely to express kinase dead fusion variants. We were unable to detect a correlation between the presence of this variants and the survival, however a significant impact of these mutations on ALCL growth and survival is likely if any ALK-specific treatment is used. Patients bearing a consistent level of inactive NPM-ALK are therefore expected to respond differently to ALK kinase inhibitors; whether such a response will be increased or decreased it remains to be elucidated.

TABLES AND FIGURES

TABLE 1

Sequences of 22 fusion primer pairs for the ALK KD amplification.
Target-specific sequences are represented in bold.

KD_MID2_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ACGCTCGACA **CTGTGAAGACGCTGCCTG**-3′
KD_MID2_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ACGCTCGACA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID3_F	5′-CGTATCGCCTCCCTCGCGCCATCAG AGACGCACTC **CTGTGAAGACGCTGCCTG**-3′
KD_MID3_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG AGACGCACTC **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID4_F	5′-CGTATCGCCTCCCTCGCGCCATCAG AGCACTGTAG **CTGTGAAGACGCTGCCTG**-3′
KD_MID4_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG AGCACTGTAG **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID5_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ATCAGACACG **CTGTGAAGACGCTGCCTG**-3′
KD_MID5_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ATCAGACACG **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID11_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TGATACGTCT **CTGTGAAGACGCTGCCTG**-3′
KD_MID11_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TGATACGTCT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID16_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TCACGTACTA **CTGTGAAGACGCTGCCTG**-3′
KD_MID16_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TCACGTACTA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID19_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TGTACTACTC **CTGTGAAGACGCTGCCTG**-3′
KD_MID19_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TGTACTACTC **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID22_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TACGAGTATG **CTGTGAAGACGCTGCCTG**-3′
KD_MID22_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TACGAGTATG **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID23_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TACTCTCGTG **CTGTGAAGACGCTGCCTG**-3′
KD_MID23_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TACTCTCGTG **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID28_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ACTACTATGT **CTGTGAAGACGCTGCCTG**-3′
KD_MID28_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ACTACTATGT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID37_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TACACACACT **CTGTGAAGACGCTGCCTG**-3′
KD_MID37_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TACACACACT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID40_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TACGCTGTCT **CTGTGAAGACGCTGCCTG**-3′
KD_MID40_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TACGCTGTCT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID46_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TGACGTATGT **CTGTGAAGACGCTGCCTG**-3′
KD_MID46_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TGACGTATGT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID47_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TGTGAGTAGT **CTGTGAAGACGCTGCCTG**-3′
KD_MID47_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TGTGAGTAGT **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID52_F	5′-CGTATCGCCTCCCTCGCGCCATCAG AGTATACATA **CTGTGAAGACGCTGCCTG**-3′
KD_MID52_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG AGTATACATA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID54_F	5′-CGTATCGCCTCCCTCGCGCCATCAG AGTGCTACGA **CTGTGAAGACGCTGCCTG**-3′
KD_MID54_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG AGTGCTACGA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID62_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TACGTCATCA **CTGTGAAGACGCTGCCTG**-3′
KD_MID62_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TACGTCATCA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID65_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TATGCTAGTA **CTGTGAAGACGCTGCCTG**-3′
KD_MID65_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TATGCTAGTA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID73_F	5′-CGTATCGCCTCCCTCGCGCCATCAG TGTCGTCGCA **CTGTGAAGACGCTGCCTG**-3′
KD_MID73_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG TGTCGTCGCA **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID74_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ACACATACGC **CTGTGAAGACGCTGCCTG**-3′
KD_MID74_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ACACATACGC **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID75_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ACAGTCGTGC **CTGTGAAGACGCTGCCTG**-3′
KD_MID75_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ACAGTCGTGC **AGTAGCTCGCCCTGTAGAT**-3′
KD_MID76_F	5′-CGTATCGCCTCCCTCGCGCCATCAG ACATGACGAC **CTGTGAAGACGCTGCCTG**-3′
KD_MID76_R	5′-CTATGCGCCTTGCCAGCCCGCTCAG ACATGACGAC **AGTAGCTCGCCCTGTAGAT**-3′

TABLE 2

INDEL mutations of NPM-ALK KD identified in 37 ALCL patients.

	Variant name	*Description*	*Protein change*	*Amino acids*	*Patients*	*Frequency*
1	Del(923-924)*	exon 25 first 2 bases as 3' splice site	D309H OOF	342	37/37 (100%)	1.4 - 4% **
2	Del(696-825)	exon 23 skipping	S232R OOF	273	37/37 (100%)	1 - 12.8%
3	Del(696-923)	exons 23-24 skipping	Δ232-307	604	16/37 (43%)	0.04 - 0.6%
4	Del(826-923)	exon 24 skipping	P276R OOF	310	18/37 (49%)	< 0.5%
5	Del(733-825)	exon 23 partial deletion, alternative 5' splice site	Δ245-275	649	7/37 (19%)	< 0.5%
6	Del(826-894)	exon 23 partial deletion alternative 3' splice site	Δ276-299	657	4/37 (11%)	< 0.5%
7	Ins24(924)	intron 24 partial retention alternative 3' splice site	R308Ins8	688	10/37 (27%)	< 0.5%
8	Ins36(924)	intron 24 partial retention alternative 3' splice site	R308Ins12	692	4/37 (11%)	< 0.5%
9	Ins106(924)	intron 24 partial retention alternative 3' splice site	D309H OOF	312	10/37 (27%)	0.04 - 0.6%
10	Del(727-766)	exons 23 partial deletion *non-canonical splice sites*	I243W OOF	303	5/37 (14%)	< 0.5%
11	Del(765-799)	exon 23 partial deletion *non-canonical splice sites*	L256P OOF	331	7/37 (19%)	< 0.5%
12	Del(751-841)	exons 23-24 partial deletion *non-canonical splice sites*	Δ251-281	649	5/37 (14%)	< 0.5%
13	Del(771-854)	exons 23 partial deletion *non-canonical splice sites*	Δ257-284	652	6/37 (16%)	< 0.5%

Variant name indicates the position of deleted nucleotides on the reference sequence or the number of inserted nucleotides, following the corresponding starting nucleotide on the reference. *Description* indicates the alternative splicing mechanism associated to each variant, *Protein change* and *Amino acids* indicate the corresponding protein mutation and the protein length, respectively. For each variant, the number of patients carrying the mutation and the respective ranges of frequency are indicated.

* this variant was detected only by the AVA software

** frequencies determined according to the AVA software

TABLE 3

Main clinical characteristics of the study population and INDEL distribution.

Patient	Age (years)	Gender	Histology	Disease Stage (St Jude)	% reads Del(923-924)	% reads Del(696-825)	# INDELs	Event	F.U.
1	8,1	M	Small Cell	III	4,0	9,9	7		alive
2	9,9	M	Common type	III	0,2	2,7	4		alive
3	5,5	M	Mixed with Small Cell Component	III	0,8	3,6	5		alive
4	3,5	F	Common Type	III	1,6	6,3	6		alive
5	5,4	F	ND	III	0,5	1,4	3		alive
6	12,6	M	Mixed with Lymphohistiocytic	III	0,5	3,9	4		alive
7	0,6	M	Lymphohistiocycitic	IV	0,9	1,1	2	DP	dead
8	2,6	F	Mixed with Small Cell Component	III	1,1	6,0	5	REL	dead
9	10,8	M	Common Type	III	1,4	2,6	3		alive
10	6,8	F	Common Type	II	2,0	8,0	5		alive
11	6,0	F	Mixed with Small Cell Component	III	0,7	5,3	4	REL	alive
12	10,8	F	Common Type	IV	3,7	2,2	3		alive
13	9,3	M	Common Type	III	2,3	2,5	5		alive
14	11,2	M	Common Type	III	0,8	8,9	4		alive
15	11,1	M	Mixed	III	0,7	4,8	2		alive
16	6,5	M	ND	III	0,3	12,8	8		alive
17	1,6	F	Common Type	III	0,8	1,4	3		alive
18	10,2	M	Hodgkin-like	III	1,4	5,8	4		alive
19*	5,6	M	Lymphohistiocycitic	III	0,3 (0,2)	5,1 (0,2)	3 (3)	REL	alive
20	11,2	F	Common Type	I	0,2	1,0	2		alive
21	0,3	M	Large Cell	III	2,1	3,1	5	REL	alive
22*	1,5	M	Mixed with Small Cell Component	IV	0,9 (0)	3,3 (0)	4 (0)	REL	alive
23	1,7	M	Lymphohistiocycitic	III	2,0	2,8	5	REL	alive
24	4,8	F	Lymphohistiocycitic	III	0,5	0,1	2		alive
25	11,4	M	Common Type	III	1,0	2,2	5	REL	dead
27*	9,8	M	Mixed with Lymphohistiocytic	III	0,4 (1,0)	0,4 (0,6)	5 (2)	REL	alive
28	12,7	F	Lymphohistiocycitic	II	0,8	1,2	6	REL	alive
29	9,2	F	ND	IV	0,2	0,7	3		alive
30	12,0	M	Common Type	III	0,3	0,3	3		alive
31	11,4	F	Common Type	III	0,4	0,5	4		dead
32	2,5	F	ND	III	0,0	0,3	3	REL	alive
33	7,0	M	Small Cell	III	0,4	0,6	2	REL	alive
34	15,6	M	Common Type	IV	0,6	0,4	2		alive
35	15,9	M	ND	III	0,2	0,6	2		alive
36	17,5	M	Common Type	III	0,4	1,0	2		alive
37	8,9	M	Common Type	III	0,2	0,6	4		alive

DP = disease progression, REL = relapse.

* Patients 19, 22 and 27 were analyzed at both diagnosis and relapse. Frequencies of Dels and number of INDELs at relapse are indicated between brackets.

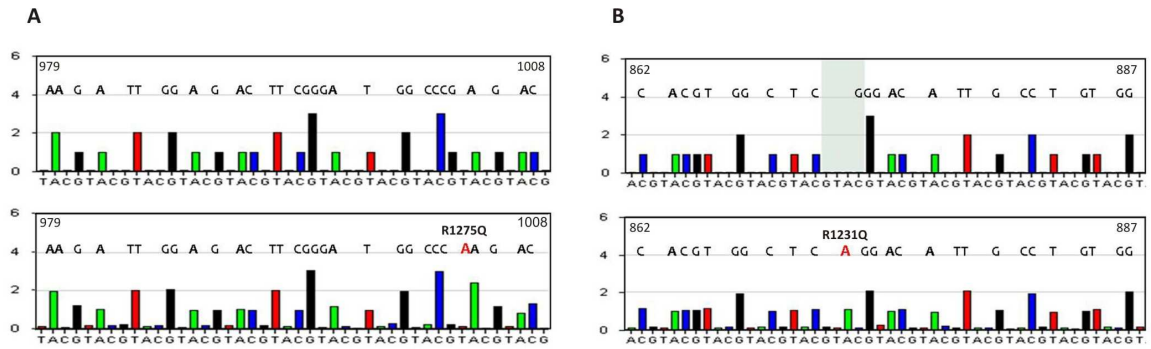


Figure 7 Raw flowgrams from individual reads showing relative luminescence signal (y-axis) obtained with each sequentially flowed nucleotide (x-axis).
(**A**) Raw flowgrams of unmutated sequence (*upper panel*) and showing the R1275Q mutation (*lower panel*) in patient 17. (**B**) Raw flowgrams of unmutated sequence (*upper panel*) and showing the R1231Q mutation (*lower panel*) in patient 36.

```
                181                                                   240
NPM-ALK WT      GLGHGAFGEVYEGQVSGMPNDPSPLQVAVKTLPEVCSEQDELDFLMEALIISKFNHQNIV
R308Ins8        ..............................................................
R308Ins12       ..............................................................
Δ276-299        ..............................................................
D309H-OOF       ..............................................................
S232R-OOF       ..............................................RASPPPWPC

                241                                                   310
NPM-ALK WT      RCIGVSLQSLPRFILLELMAGGDLKSFLRETRPRPSQPSSLAMLDLLHVARDIACGCQYL
R308Ins8        ..............................................................
R308Ins12       ..............................................................
Δ276-299        ..................................----------------------..
D309H-OOF       ..............................................................
S232R-OOF       WTFCTWLGTLPVAVSIWRKTTSSTETLLPETAS*

                311                                                   680
NPM-ALK WT      EENHFIHR-----------DIAARNCLLTCPGPGRVAKIGDFGMARDIYRASYYR - -
R308Ins8        ........----KCISFLPR..........................................
R308Ins12       ........DKIRKCISFLPR..........................................
Δ276-299        ........------------..........................................
D309H OOF       ........HCCQKLPLDLSRPWKSGQDWRLRDGPRHLQGELL*
S232R OOF
```

Figure 8 Amino acid sequence alignments of NPM-ALK wild-type (black) and 4 alternative splicing variants: R308Ins8 (red), R308Ins12 (orange), Δ276-299 (green), D309H OOF (blue) and S232R OOF (purple). Sequence differences caused by frame shift mutations are shown. Premature stop codons are indicated by asterisks.

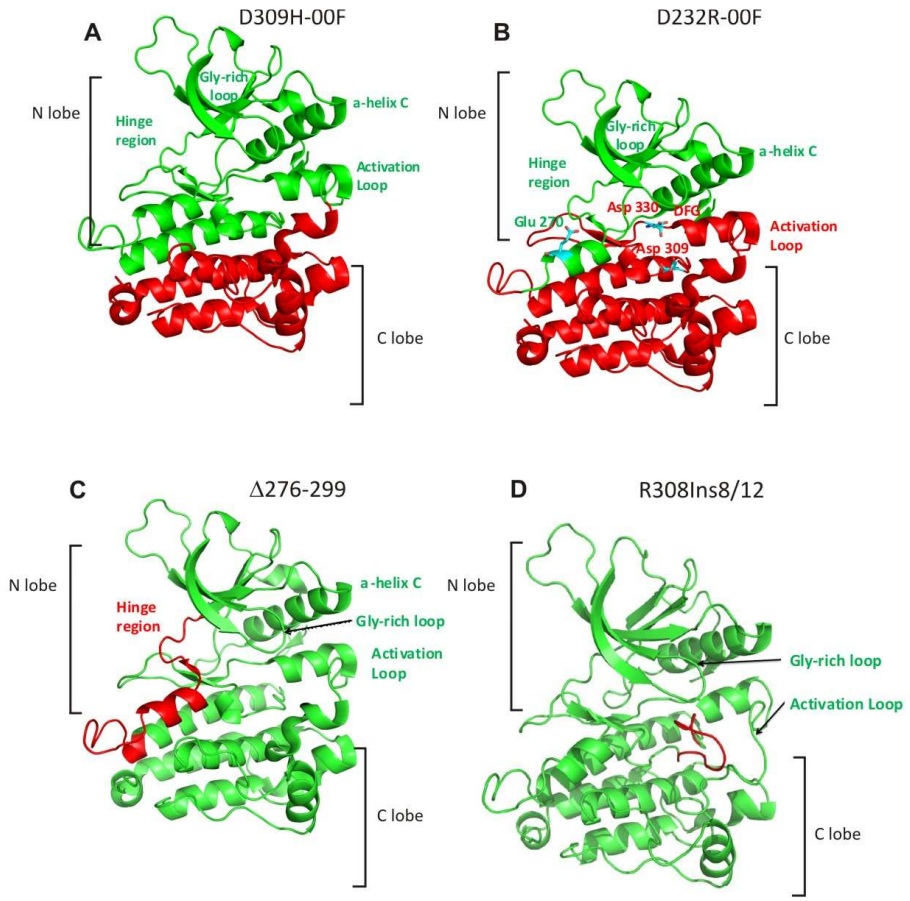


Figure 9 Homology modelling structure of D309H OOF [Del(923-924)] (**A**), S232R OOF [Del(696-825)] (**B**), Δ276-299 [Del(826-894)] (**C**) and R308Ins8/12 [Ins24/36(924)] (**D**). Deleted or inserted regions are indicated in red.

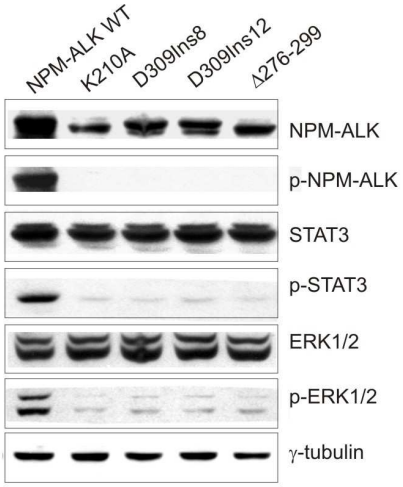


Figure 10 Western blot analysis of HEK-293T cells transiently transfected with wild-type NPM-ALK, K210A, and 3 INDEL mutants, showing the effect of R308Ins8/12 and Δ276-299 expression on protein activation (p-NPM-ALK) and function (p-STAT3 and p-ERK1/2). γ-tubulin was used as loading control.

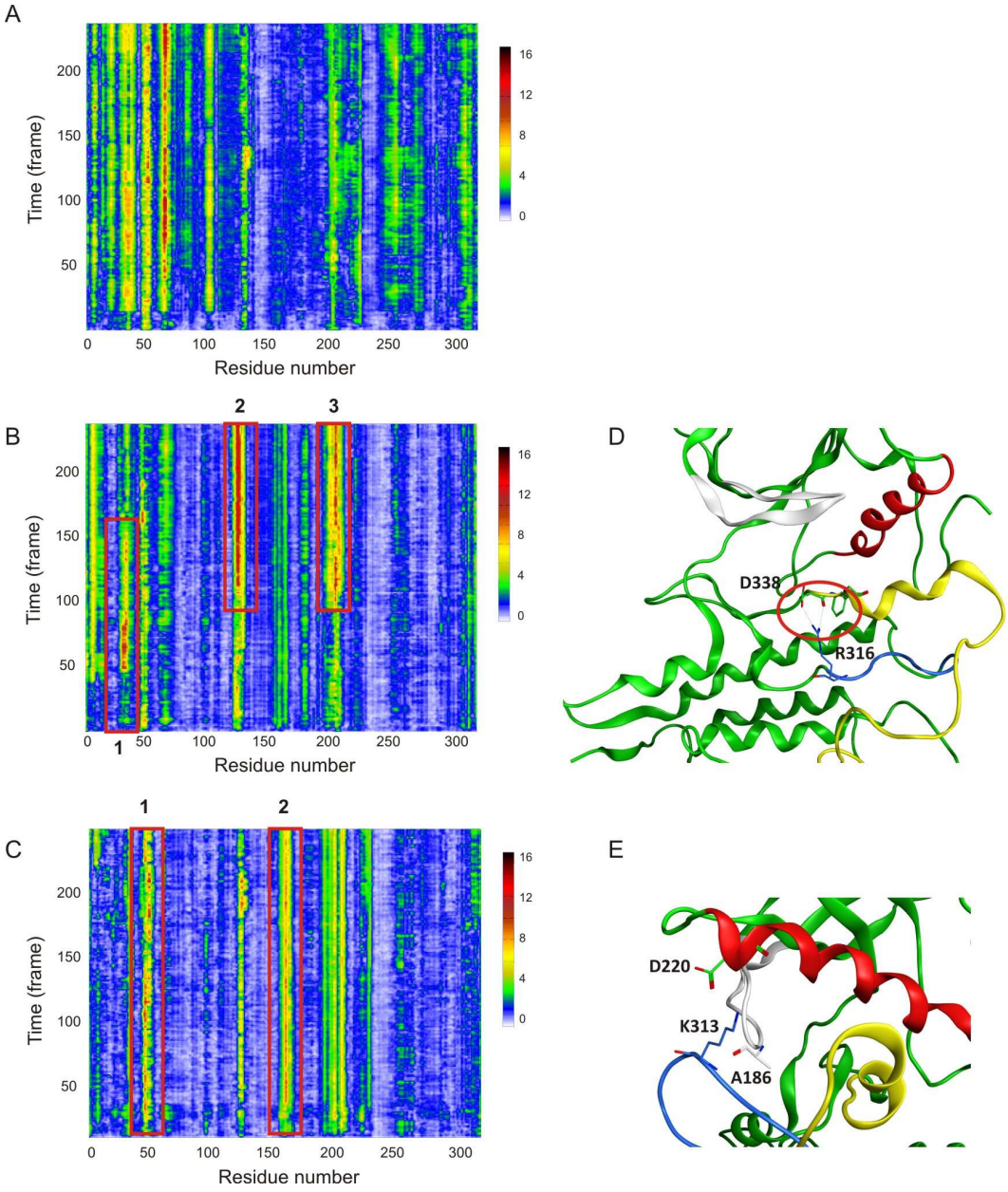


Figure 11 Molecular dynamics simulation of R308Ins8 and R308Ins12, in comparison to ALK native kinase. Panels **A**, **B** and **C** depict the alpha carbon RainbowRMSD of the WT, R308Ins8 and R308Ins12 tyrosine kinase domain, respectively, during the MD run. The largest conformational changes are circled with red boxes and correspond to the P-loop (*Box 1*), "Ins8" (*Box 2*) and to the A-loop (*Box 3*) in panel **B**, "Ins12" (*Box 1*) and the P-loop (*Box 2*) in panel **C**. Panels **D** and **E** represent snapshots of R308Ins8 and R308Ins12 ALK kinase domain (green cartoon) conformation after 40 and 100 ns of simulation, respectively. The P-loop is colored in white, the αC-helix in red, the Ins8/12 in blue and the A-loop in yellow. The side chains of R316, D338, A186, D220, and K313 are represented in sticks.

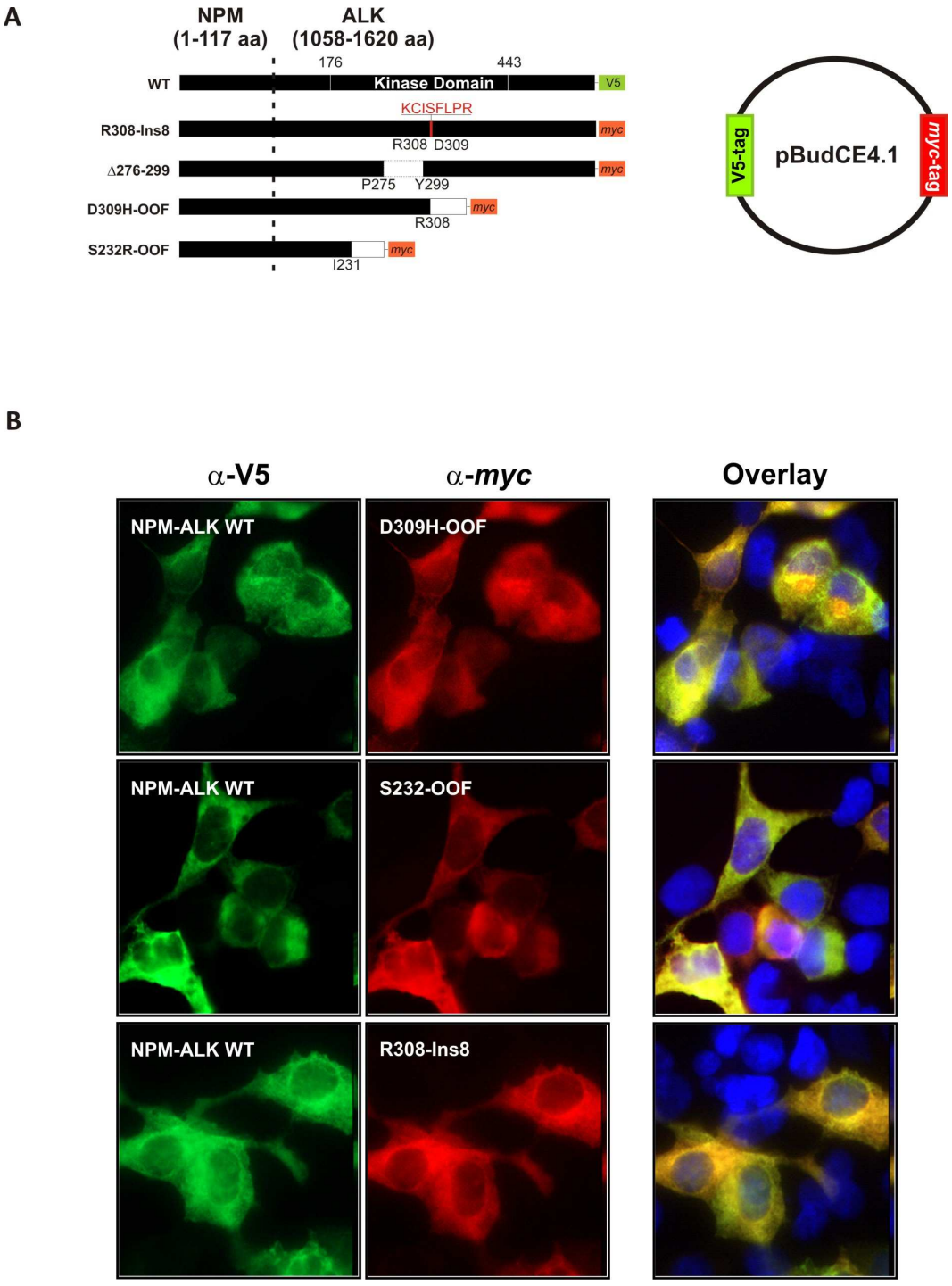


Figure 12. Coexpression strategy for native and mutant (D309H-OOF, S232R-OOF, and R308-Ins8) NPM-ALK proteins HEK-293T cells. (**A**) Schematic representation of expressed fusion constructs and pBudCE4.1 plasmid. (**B**) Subcellular localization of wild-type (α-V5, Alexa488 green) and mutant (α-*myc*, Alexa546 red) NPM-ALK proteins in co-transfected HEK-293T cells, by immunofluorescence microscopy analysis. Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) dye (blue).

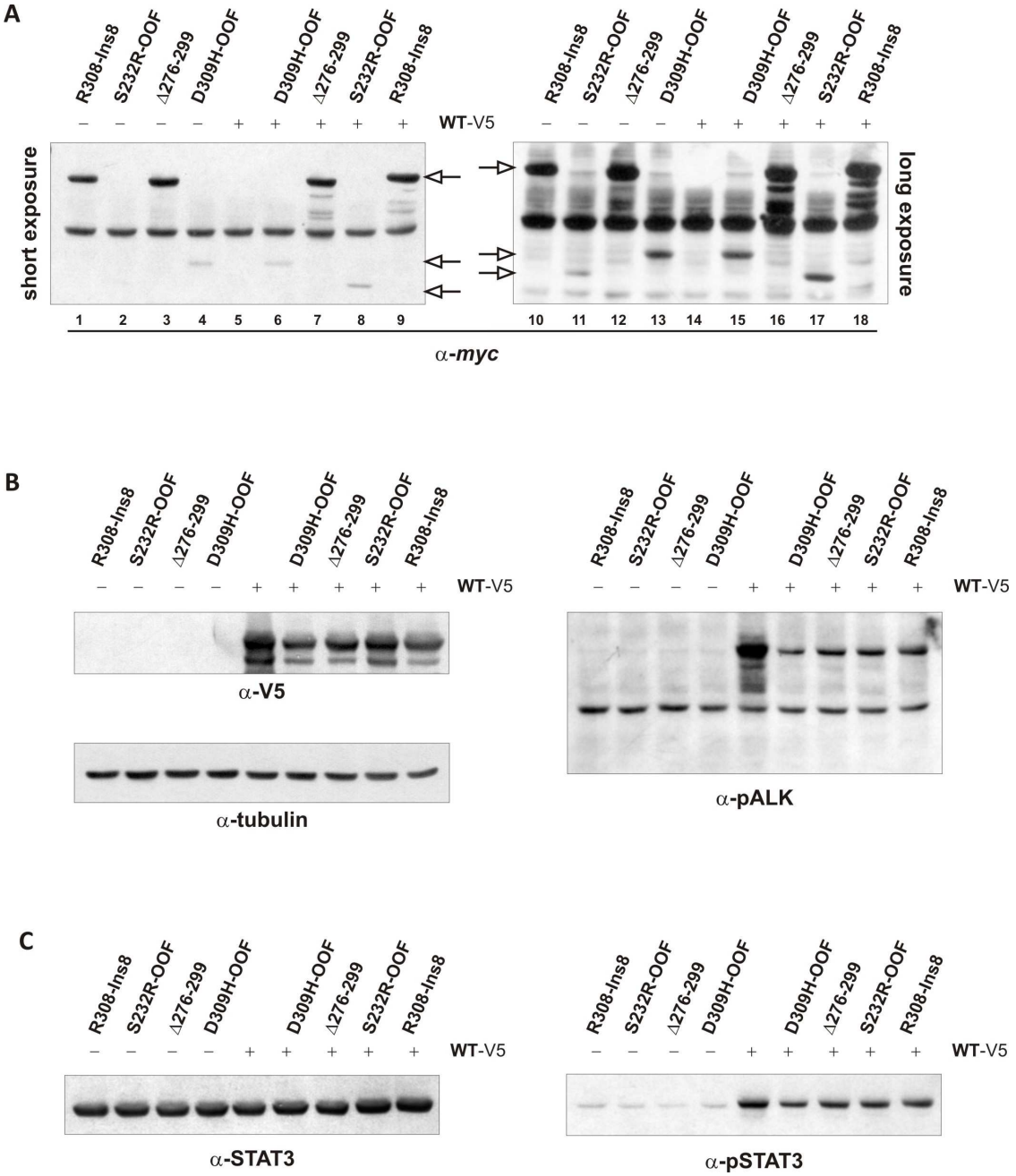


Figure 13. INDELs expression and activity in HEK293-T cells. (**A**) To measure INDELs protein expression in the presence (+) or absence (–) of wild-type NPM-ALK, western blot analysis was performed using an anti-*myc* specific antibody (α-*myc*). In addition, for proper visualization of mutants expression different time exposures were included (short exposure, samples 1-9; long exposure, samples 10-18). (**C**) Similarly, total and phosphorylated NPM-ALK expression was detected using an anti-V5-tag (α-V5) and phospho-ALK antibody (α-pALK), respectively, whereas NPM-ALK kinase activity was indirectly measured by assessing phosphorylation status of downstream STAT3 client protein (p-STAT3). Tubulin was included as loading control.

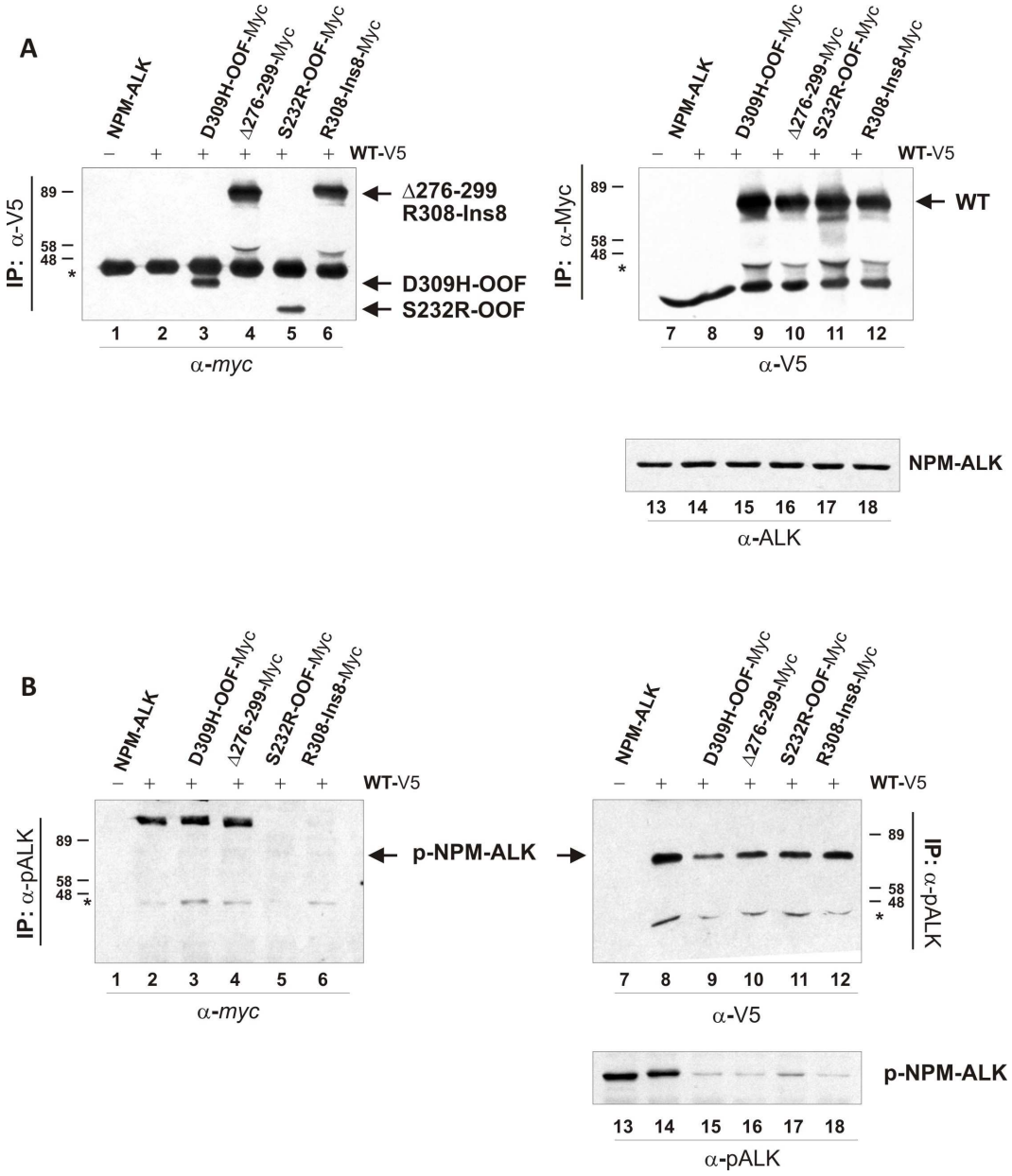


Figure 14 ALK *myc*-tagged INDELs bind and inactivate V5-tagged NPM-ALK kinase. (**A**) NPM-ALK WT/INDEL complex formation was determined in HEK293-T cotransfected cells by reciprocal co-immunoprecipitation, using anti-V5 (samples 1-6) and anti-*myc* (samples 7-12) specific antibodies. The corresponding immunocomplexes were then detected by Western blotting using anti-*myc* and anti-V5 antibodies, respectively. Untagged wild-type NPM-ALK (samples 1 and 7, –) was also included as a blank. The lower panel (samples 13-18) shows NPM-ALK expression in corresponding whole-cell lysates. INDELs were able to form holocomplexes with wild-type NPM-ALK. (**B**) Similarly, immunoprecipitation was performed using an anti-phopsho site-specific ALK antibody (α-pALK), and nitrocellulose blots were probed against *myc*- (samples 1-6, α-*myc*) or V5-tagged (samples 7-12, α-V5) proteins. Compared to V5-tagged NPM-ALK, INDELs are neither phosphorylated nor associated with active NPM-ALK kinase. The lower panel shows phosphorylation status of NPM-ALK in HEK293-T whole cell lysates (samples 13-18). Asterisk indicate nonspecific immunoglobulin (IgG) heavy chain.

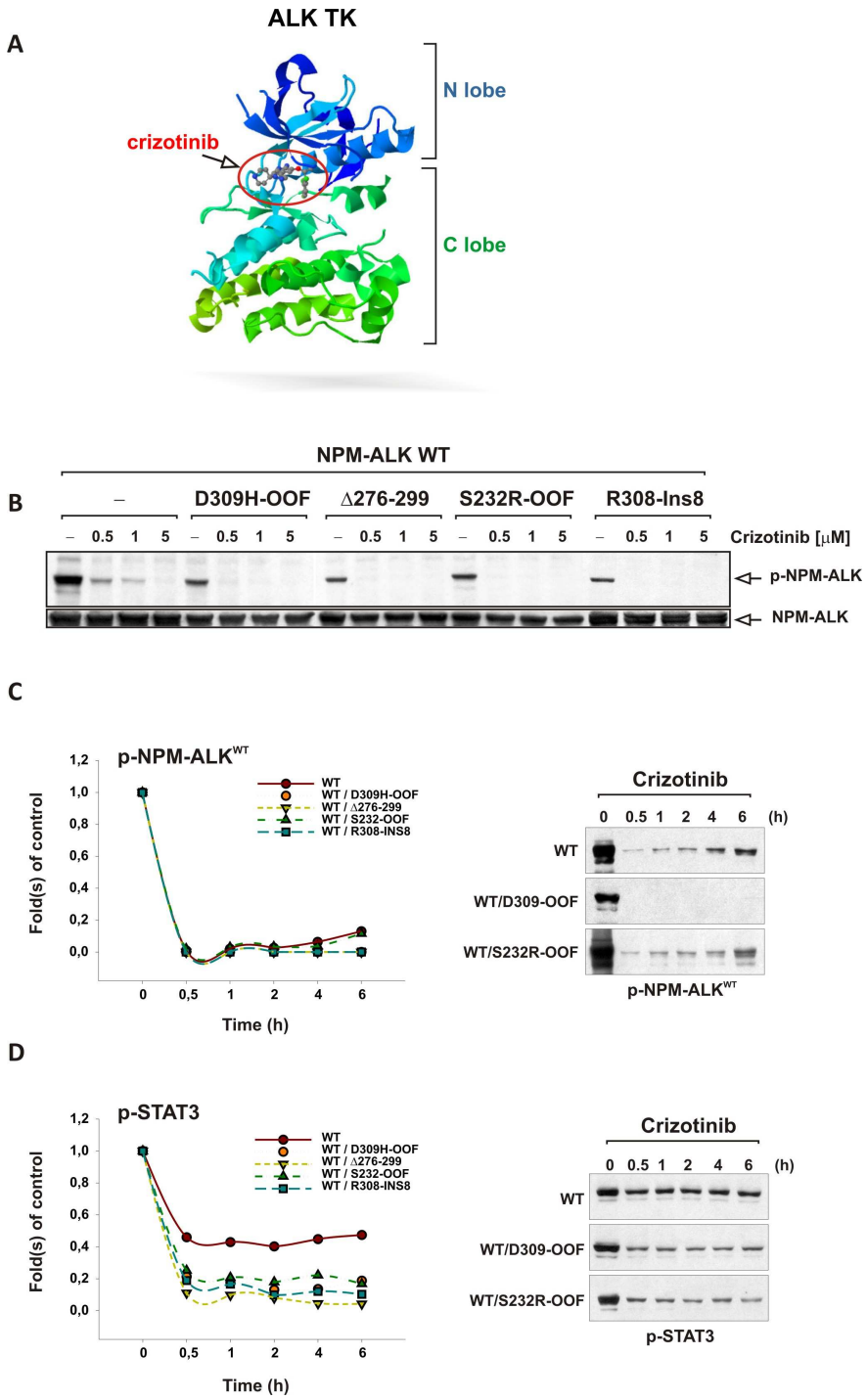


Figure 15 Effects of INDEL protein expression on NPM-ALK sensitivity to crizotinib. (**A**) Structure of the human ALK tyrosine kinase domain (TK) in complex with crizotinib (PDB code 2XP2). (**B**) To inhibit NPM-ALK kinase activity, HEK-293T cells were transfected with wild-type NPM-ALK alone (–) or in combination with D309H-OOF, Δ276-299, S232R-OOF and R308-Ins8 mutants, and total or phosphorylated NPM-ALK levels were measured after exposure to crizotinib (0.5, 1 and 5 μM) for 6 hours. Crizotinib-induced inhibition of NPM-ALK phosphorylation and signaling was also assessed in a time-dependent manner, by exposing transfected and cotransfected cells for different time intervals (0.5, 1, 2, 4 and 6 hours). Phosphorylated proteins (**C**) NPM-ALK (*p-NPM-ALK*) and (**D**) STAT3 (*p-STAT3*) were detected by Western blotting (*right*), and band densities, measured with dedicated image software, were expressed in graph (*left*) as fold control (untreated samples).

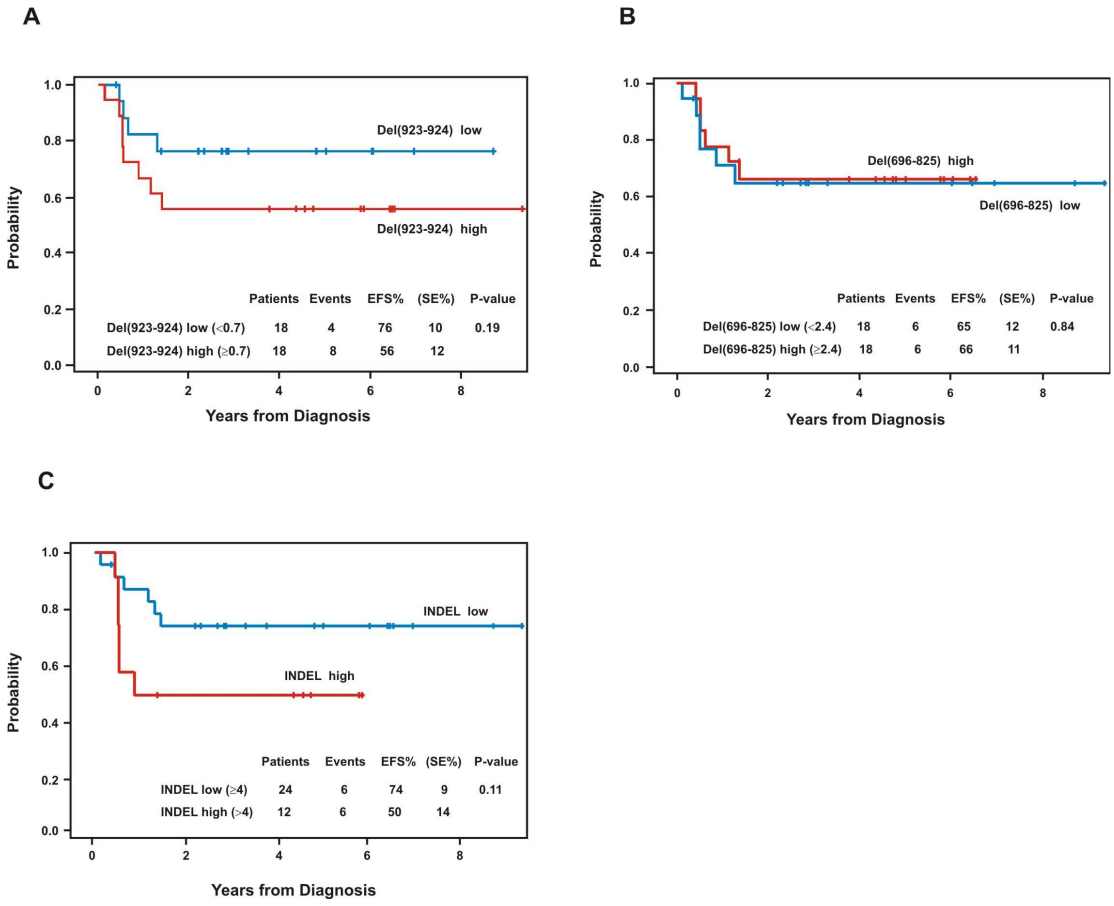


Figure 16 Five-years PFS according to the frequency of Del(923-924) (**A**) and Del(696-825) (**B**) and the absolute number of INDELs (**C**).

BIBLIOGRAPHY

1. 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.
2. Ferreri, A.J., et al., *Anaplastic large cell lymphoma, ALK-positive.* Crit Rev Oncol Hematol. **83**(2): p. 293-302.
3. Benharroch, D., et al., *ALK-positive lymphoma: a single disease with a broad spectrum of morphology.* Blood, 1998. **91**(6): p. 2076-84.
4. 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.
5. *A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project.* Blood, 1997. **89**(11): p. 3909-18.
6. Mora, J., et al., *Large cell non-Hodgkin lymphoma of childhood: Analysis of 78 consecutive patients enrolled in 2 consecutive protocols at the Memorial Sloan-Kettering Cancer Center.* Cancer, 2000. **88**(1): p. 186-97.
7. Willemze, R., et al., *EORTC classification for primary cutaneous lymphomas: a proposal from the Cutaneous Lymphoma Study Group of the European Organization for Research and Treatment of Cancer.* Blood, 1997. **90**(1): p. 354-71.
8. Falini, B., et al., *ALK+ lymphoma: clinico-pathological findings and outcome.* Blood, 1999. **93**(8): p. 2697-706.
9. Shiota, M., et al., *Anaplastic large cell lymphomas expressing the novel chimeric protein p80NPM/ALK: a distinct clinicopathologic entity.* Blood, 1995. **86**(5): p. 1954-60.
10. 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.
11. 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.
12. Wlodarska, I., et al., *The cryptic inv(2)(p23q35) defines a new molecular genetic subtype of ALK-positive anaplastic large-cell lymphoma.* Blood, 1998. **92**(8): p. 2688-95.
13. Rosenwald, A., et al., *t(1;2)(q21;p23) and t(2;3)(p23;q21): two novel variant translocations of the t(2;5)(p23;q35) in anaplastic large cell lymphoma.* Blood, 1999. **94**(1): p. 362-4.
14. 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.
15. Adachi, Y., et al., *Nucleolar targeting signal of Rex protein of human T-cell leukemia virus type I specifically binds to nucleolar shuttle protein B-23.* J Biol Chem, 1993. **268**(19): p. 13930-4.
16. Iwahara, T., et al., *Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system.* Oncogene, 1997. **14**(4): p. 439-49.
17. Morris, S.W., et al., *ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK).* Oncogene, 1997. **14**(18): p. 2175-88.

18. Pulford, K., et al., *Detection of anaplastic lymphoma kinase (ALK) and nucleolar protein nucleophosmin (NPM)-ALK proteins in normal and neoplastic cells with the monoclonal antibody ALK1.* Blood, 1997. **89**(4): p. 1394-404.
19. 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.
20. 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.
21. Palmer, R.H., et al., *Anaplastic lymphoma kinase: signalling in development and disease.* Biochem J, 2009. **420**(3): p. 345-61.
22. Perez-Pinera, P., et al., *Anaplastic lymphoma kinase is activated through the pleiotrophin/receptor protein-tyrosine phosphatase beta/zeta signaling pathway: an alternative mechanism of receptor tyrosine kinase activation.* J Biol Chem, 2007. **282**(39): p. 28683-90.
23. Motegi, A., et al., *ALK receptor tyrosine kinase promotes cell growth and neurite outgrowth.* J Cell Sci, 2004. **117**(Pt 15): p. 3319-29.
24. Moog-Lutz, C., et al., *Activation and inhibition of anaplastic lymphoma kinase receptor tyrosine kinase by monoclonal antibodies and absence of agonist activity of pleiotrophin.* J Biol Chem, 2005. **280**(28): p. 26039-48.
25. Miyake, I., et al., *Activation of anaplastic lymphoma kinase is responsible for hyperphosphorylation of ShcC in neuroblastoma cell lines.* Oncogene, 2002. **21**(38): p. 5823-34.
26. Mathivet, T., P. Mazot, and M. Vigny, *In contrast to agonist monoclonal antibodies, both C-terminal truncated form and full length form of Pleiotrophin failed to activate vertebrate ALK (anaplastic lymphoma kinase)?* Cell Signal, 2007. **19**(12): p. 2434-43.
27. Chiarle, R., et al., *The anaplastic lymphoma kinase in the pathogenesis of cancer.* Nat Rev Cancer, 2008. **8**(1): p. 11-23.
28. Wellmann, A., et al., *The activated anaplastic lymphoma kinase increases cellular proliferation and oncogene up-regulation in rat 1a fibroblasts.* FASEB J, 1997. **11**(12): p. 965-72.
29. Kuefer, M.U., et al., *Retrovirus-mediated gene transfer of NPM-ALK causes lymphoid malignancy in mice.* Blood, 1997. **90**(8): p. 2901-10.
30. Chiarle, R., et al., *NPM-ALK transgenic mice spontaneously develop T-cell lymphomas and plasma cell tumors.* Blood, 2003. **101**(5): p. 1919-27.
31. Ambrogio, C., et al., *p130Cas mediates the transforming properties of the anaplastic lymphoma kinase.* Blood, 2005. **106**(12): p. 3907-16.
32. Fujimoto, J., et al., *Characterization of the transforming activity of p80, a hyperphosphorylated protein in a Ki-1 lymphoma cell line with chromosomal translocation t(2;5).* Proc Natl Acad Sci U S A, 1996. **93**(9): p. 4181-6.
33. Bai, R.Y., et al., *Nucleophosmin-anaplastic lymphoma kinase of large-cell anaplastic lymphoma is a constitutively active tyrosine kinase that utilizes phospholipase C-gamma to mediate its mitogenicity.* Mol Cell Biol, 1998. **18**(12): p. 6951-61.
34. Zamo, A., et al., *Anaplastic lymphoma kinase (ALK) activates Stat3 and protects hematopoietic cells from cell death.* Oncogene, 2002. **21**(7): p. 1038-47.
35. Bai, R.Y., et al., *Nucleophosmin-anaplastic lymphoma kinase associated with anaplastic large-cell lymphoma activates the phosphatidylinositol 3-kinase/Akt antiapoptotic signaling pathway.* Blood, 2000. **96**(13): p. 4319-27.
36. Rassidakis, G.Z., et al., *Inhibition of Akt increases p27Kip1 levels and induces cell cycle arrest in anaplastic large cell lymphoma.* Blood, 2005. **105**(2): p. 827-9.

37. Piva, R., et al., *Ablation of oncogenic ALK is a viable therapeutic approach for anaplastic large-cell lymphomas.* Blood, 2006. **107**(2): p. 689-97.
38. Wan, W., et al., *Anaplastic lymphoma kinase activity is essential for the proliferation and survival of anaplastic large-cell lymphoma cells.* Blood, 2006. **107**(4): p. 1617-23.
39. 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. **28**(25): p. 3987-93.
40. Mosse, Y.P., et al., *Identification of ALK as a major familial neuroblastoma predisposition gene.* Nature, 2008. **455**(7215): p. 930-5.
41. Janoueix-Lerosey, I., et al., *Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma.* Nature, 2008. **455**(7215): p. 967-70.
42. Chen, Y., et al., *Oncogenic mutations of ALK kinase in neuroblastoma.* Nature, 2008. **455**(7215): p. 971-4.
43. George, R.E., et al., *Activating mutations in ALK provide a therapeutic target in neuroblastoma.* Nature, 2008. **455**(7215): p. 975-8.
44. 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.
45. Choi, Y.L., et al., *Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer.* Cancer Res, 2008. **68**(13): p. 4971-6.
46. Koivunen, J.P., et al., *EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer.* Clin Cancer Res, 2008. **14**(13): p. 4275-83.
47. Soda, M., et al., *A mouse model for EML4-ALK-positive lung cancer.* Proc Natl Acad Sci U S A, 2008. **105**(50): p. 19893-7.
48. Coffin, C.M., et al., *ALK1 and p80 expression and chromosomal rearrangements involving 2p23 in inflammatory myofibroblastic tumor.* Mod Pathol, 2001. **14**(6): p. 569-76.
49. Butrynski, J.E., et al., *Crizotinib in ALK-rearranged inflammatory myofibroblastic tumor.* N Engl J Med. **363**(18): p. 1727-33.
50. McDermott, U., et al., *Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors.* Cancer Res, 2008. **68**(9): p. 3389-95.
51. 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.
52. Christensen, J.G., et al., *Cytoreductive antitumor activity of PF-2341066, a novel inhibitor of anaplastic lymphoma kinase and c-Met, in experimental models of anaplastic large-cell lymphoma.* Mol Cancer Ther, 2007. **6**(12 Pt 1): p. 3314-22.
53. Camidge, D.R., et al., *Activity and safety of crizotinib in patients with ALK-positive non-small-cell lung cancer: updated results from a phase 1 study.* Lancet Oncol. **13**(10): p. 1011-9.
54. Shaw, A.T., et al., *Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis.* Lancet Oncol. **12**(11): p. 1004-12.
55. Gambacorti-Passerini, C., C. Messa, and E.M. Pogliani, *Crizotinib in anaplastic large-cell lymphoma.* N Engl J Med. **364**(8): p. 775-6.
56. Choi, Y.L., et al., *EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors.* N Engl J Med. **363**(18): p. 1734-9.
57. Sasaki, T., et al., *The neuroblastoma-associated F1174L ALK mutation causes resistance to an ALK kinase inhibitor in ALK-translocated cancers.* Cancer Res. **70**(24): p. 10038-43.

58. Shah, N.P., et al., *Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia.* Cancer Cell, 2002. **2**(2): p. 117-25.
59. Tamborini, E., et al., *A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient.* Gastroenterology, 2004. **127**(1): p. 294-9.
60. Carter, T.A., et al., *Inhibition of drug-resistant mutants of ABL, KIT, and EGF receptor kinases.* Proc Natl Acad Sci U S A, 2005. **102**(31): p. 11011-6.
61. Kobayashi, S., et al., *EGFR mutation and resistance of non-small-cell lung cancer to gefitinib.* N Engl J Med, 2005. **352**(8): p. 786-92.
62. Willis, S.G., et al., *High-sensitivity detection of BCR-ABL kinase domain mutations in imatinib-naive patients: correlation with clonal cytogenetic evolution but not response to therapy.* Blood, 2005. **106**(6): p. 2128-37.
63. Pfeifer, H., et al., *Kinase domain mutations of BCR-ABL frequently precede imatinib-based therapy and give rise to relapse in patients with de novo Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL).* Blood, 2007. **110**(2): p. 727-34.
64. Soverini, S., et al., *Philadelphia-positive acute lymphoblastic leukemia patients already harbor BCR-ABL kinase domain mutations at low levels at the time of diagnosis.* Haematologica. **96**(4): p. 552-7.
65. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors.* Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
66. Schuurman, R., et al., *Worldwide evaluation of DNA sequencing approaches for identification of drug resistance mutations in the human immunodeficiency virus type 1 reverse transcriptase.* J Clin Microbiol, 1999. **37**(7): p. 2291-6.
67. Milbury, C.A., J. Li, and G.M. Makrigiorgos, *PCR-based methods for the enrichment of minority alleles and mutations.* Clin Chem, 2009. **55**(4): p. 632-40.
68. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-80.
69. Thomas, R.K., et al., *Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing.* Nat Med, 2006. **12**(7): p. 852-5.
70. Marchetti, A., et al., *Complex mutations & subpopulations of deletions at exon 19 of EGFR in NSCLC revealed by next generation sequencing: potential clinical implications.* PLoS One. **7**(7): p. e42164.
71. van Gaal, J.C., et al., *Anaplastic lymphoma kinase aberrations in rhabdomyosarcoma: clinical and prognostic implications.* J Clin Oncol. **30**(3): p. 308-15.
72. Reiter, A., et al., *Successful treatment strategy for Ki-1 anaplastic large-cell lymphoma of childhood: a prospective analysis of 62 patients enrolled in three consecutive Berlin-Frankfurt-Munster group studies.* J Clin Oncol, 1994. **12**(5): p. 899-908.
73. Brugieres, L., et al., *CD30(+) anaplastic large-cell lymphoma in children: analysis of 82 patients enrolled in two consecutive studies of the French Society of Pediatric Oncology.* Blood, 1998. **92**(10): p. 3591-8.
74. Seidemann, K., et al., *Short-pulse B-non-Hodgkin lymphoma-type chemotherapy is efficacious treatment for pediatric anaplastic large cell lymphoma: a report of the Berlin-Frankfurt-Munster Group Trial NHL-BFM 90.* Blood, 2001. **97**(12): p. 3699-706.
75. Williams, D.M., et al., *Anaplastic large cell lymphoma in childhood: analysis of 72 patients treated on The United Kingdom Children's Cancer Study Group chemotherapy regimens.* Br J Haematol, 2002. **117**(4): p. 812-20.

76. Shah, N.P. and C.L. Sawyers, *Mechanisms of resistance to STI571 in Philadelphia chromosome-associated leukemias.* Oncogene, 2003. **22**(47): p. 7389-95.
77. O'Hare, T., C.A. Eide, and M.W. Deininger, *Bcr-Abl kinase domain mutations, drug resistance, and the road to a cure for chronic myeloid leukemia.* Blood, 2007. **110**(7): p. 2242-9.
78. Engelman, J.A. and P.A. Janne, *Mechanisms of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non-small cell lung cancer.* Clin Cancer Res, 2008. **14**(10): p. 2895-9.
79. Liu, J., et al., *BCR-ABL mutants spread resistance to non-mutated cells through a paracrine mechanism.* Leukemia, 2008. **22**(4): p. 791-9.
80. Epstein, L.F., et al., *The R1275Q neuroblastoma mutant and certain ATP-competitive inhibitors stabilize alternative activation loop conformations of anaplastic lymphoma kinase.* J Biol Chem. **287**(44): p. 37447-57.
81. Bresler, S.C., et al., *Differential inhibitor sensitivity of anaplastic lymphoma kinase variants found in neuroblastoma.* Sci Transl Med. **3**(108): p. 108ra114.
82. Zhang, S., et al., *Crizotinib-resistant mutants of EML4-ALK identified through an accelerated mutagenesis screen.* Chem Biol Drug Des. **78**(6): p. 999-1005.
83. Gruber, F.X., et al., *A novel Bcr-Abl splice isoform is associated with the L248V mutation in CML patients with acquired resistance to imatinib.* Leukemia, 2006. **20**(11): p. 2057-60.
84. Laudadio, J., et al., *An intron-derived insertion/truncation mutation in the BCR-ABL kinase domain in chronic myeloid leukemia patients undergoing kinase inhibitor therapy.* J Mol Diagn, 2008. **10**(2): p. 177-80.
85. Curvo, R.P., et al., *A recurrent splicing variant without c-ABL Exon 7 in Imatinib-resistant patients.* Leuk Res, 2008. **32**(3): p. 508-10.
86. Sherbenou, D.W., et al., *Characterization of BCR-ABL deletion mutants from patients with chronic myeloid leukemia.* Leukemia, 2008. **22**(6): p. 1184-90.
87. Lee, T.S., et al., *BCR-ABL alternative splicing as a common mechanism for imatinib resistance: evidence from molecular dynamics simulations.* Mol Cancer Ther, 2008. **7**(12): p. 3834-41.
88. Khorashad, J.S., D. Milojkovic, and A.G. Reid, *Variant isoforms of BCR-ABL1 in chronic myelogenous leukemia reflect alternative splicing of ABL1 in normal tissue - letter.* Mol Cancer Ther. **9**(7): p. 2152.
89. Sorel, N., et al., *Comprehensive characterization of a novel intronic pseudo-exon inserted within an e14/a2 BCR-ABL rearrangement in a patient with chronic myeloid leukemia.* J Mol Diagn. **12**(4): p. 520-4.
90. Gaillard, J.B., et al., *Exon 7 deletion in the bcr-abl gene is frequent in chronic myeloid leukemia patients and is not correlated with resistance against imatinib.* Mol Cancer Ther. **9**(11): p. 3083-9.
91. Meggyesi, N., et al., *Characterization of ABL exon 7 deletion by molecular genetic and bioinformatic methods reveals no association with imatinib resistance in chronic myeloid leukemia.* Med Oncol. **29**(3): p. 2136-42.
92. Gruber, F.X., et al., *BCR-ABL isoforms associated with intrinsic or acquired resistance to imatinib: more heterogeneous than just ABL kinase domain point mutations?* Med Oncol. **29**(1): p. 219-26.
93. Ma, W., et al., *Three novel alternative splicing mutations in BCR-ABL1 detected in CML patients with resistance to kinase inhibitors.* Int J Lab Hematol. **33**(3): p. 326-31.
94. Okubo, J., et al., *Aberrant activation of ALK kinase by a novel truncated form ALK protein in neuroblastoma.* Oncogene. **31**(44): p. 4667-76.
95. Pulford, K., et al., *The emerging normal and disease-related roles of anaplastic lymphoma kinase.* Cell Mol Life Sci, 2004. **61**(23): p. 2939-53.

RINGRAZIAMENTI

Ringrazio il prof. Angelo Rosolen per avermi dato l'opportunità di svolgere questo lavoro di tesi, che mi ha permesso di acquisire nuove importanti competenze scientifiche e una sempre maggiore autonomia nelle gestione di un progetto di ricerca.

Ringrazio di tutto cuore il dr. Paolo Bonvini, perché con instancabile determinazione mi ha supportato nei momenti decisivi di questo percorso. Grazie per aver cercato di insegnarmi l'amore per la ricerca e per avermi sostenuta quando sembrava così difficile decidere come andare avanti. Per me, Paolo, sarai sempre un punto di riferimento.

Grazie alla dr.ssa Marta Pillon, alla dr.ssa Lara Mussolin, a Elisa Carraro, per il prezioso contributo per la selezione dei pazienti e le analisi statistiche.

Ringrazio il dr. Alessandro Albiero e le dr.sse Francesca Dal Pero, Sara Todesco, Barbara Simionati e tutto lo staff di BMR genomics, che si sono occupati del sequenziamento e dell'analisi dei dati.

Grazie di cuore al prof. Stefano Moro e ai suoi collaboratori dr. Giorgio Cozza e dr. Andrea Cristiani, per la loro competenza e disponibilità.

Un grazie anche ai miei colleghi di laboratorio Elena Poli, Angelica Zin, Elisa Tosato, Simona Primerano, che mi hanno incoraggiato e sopportato nella frenesia delle ultime settimane.

Grazie a Giancarlo, a pieno titolo co-autore di questa tesi, per essermi stato sempre vicino, anche quando avrebbe fatto bene ad evitarmi… e per essersi preso cura di me con amorevole dedizione. Grazie tesoro.

Grazie come sempre alla mia famiglia, papà Gino, mamma Paola e sorella Laura, perché anche se non siamo più così vicini siete sempre presenti nel momento del bisogno. Ora avete un motivo in più per chiamarmi "dottoressa".

Grazie alla mia meravigliosa amica Martina, alla piccola Giulia, a Giovanni, grazie a Federica e Antonio.

Grazie a Mariafrancesca, perché qualcosa di straordinario è stato veramente possibile, e questa tesi lo dimostra.