blood

Prepublished online Jul 9, 2009;
doi:10.1182/blood-2008-08-173963

Identification and molecular characterisation of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1 -positive acute lymphoblastic leukaemia patients: on behalf of GIMEMA AL WP

Ilaria Iacobucci, Clelia Tiziana Storlazzi, Daniela Cilloni, Annalisa Lonetti, Emanuela Ottaviani, Simona Soverini, Annalisa Astolfi, Sabina Chiaretti, Antonella Vitale, Francesca Messa, Luciana Impera, Carmen Baldazzi, Pietro D'Addabbo, Cristina Papayannidis, Angelo Lonoce, Sabrina Colarossi, Marco Vignetti, Pier Paolo Piccaluga, Stefania Paolini, Domenico Russo, Fabrizio Pane, Giuseppe Saglio, Michele Baccarani, Robin Foa and Giovanni Martinelli

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published semimonthly by the American Society of Hematology, 1900 M St, NW, Suite 200, Washington DC 20036.

Identification and molecular characterisation of recurrent genomic deletions on 7p12 in the *IKZF1* gene in a large cohort of *BCR-ABL1*-positive acute lymphoblastic leukaemia patients: on behalf of GIMEMA AL WP

Ilaria Iacobucci[1], Clelia Tiziana Storlazzi[2], Daniela Cilloni[3], Annalisa Lonetti[1], Emanuela Ottaviani[1], Simona Soverini[1], Annalisa Astolfi[4], Sabina Chiaretti[5], Antonella Vitale[5], Francesca Messa[3], Luciana Impera[2], Carmen Baldazzi[1], Pietro D'Addabbo[2], Cristina Papayannidis[1], Angelo Lonoce[2], Sabrina Colarossi[1], Marco Vignetti[5], Pier Paolo Piccaluga[1], Stefania Paolini[1], Domenico Russo[7], Fabrizio Pane[6], Giuseppe Saglio[3], Michele Baccarani[1], Robin Foà[5] and Giovanni Martinelli[1]

[1]Department of Hematology/Oncology "L. and A. Seràgnoli" S.Orsola Malpighi Hospital, University of Bologna, Bologna; [2]Department of Genetics and Microbiology, University of Bari. [3] Department of Clinical and Biological Science, University of Turin at Orbassano, Turin; [4]Pediatric Oncology and Hematology "L. Seràgnoli", Bologna; [5]"La Sapienza" University, Dept. of Cellular Biotechnologies and Hematology, Rome; [6] CEINGE Biotecnologie Avanzate and Department of Biochemistry and Medical Biotechnology, University of Naples Federico II, Naples;
[7] Hematology and Bone Marrow Transplantation Unit, Spedali Civili Hospital, University of Brescia

Keywords: ALL, *BCR-ABL1*, Ikaros, SNP array.

Address for correspondence:
Prof. Giovanni Martinelli - Molecular Biology Unit
Department of Hematology/Oncology "Seràgnoli", University of Bologna,
Via Massarenti, 9 - 40138 Bologna, Italy.
Phone: 0039-051-6363829; Fax: 0039-051-6364037;
e-mail: giovanni.martinelli2@unibo.it

For a complete list of GIMEMA AL WP participants, see the Supplemental Appendix

Abstract

The *BCR-ABL1* fusion gene defines the subgroup of acute lymphoblastic leukaemia (ALL) with the worst clinical prognosis. To identify oncogenic lesions that combine with *BCR-ABL1* to cause ALL, we used Affymetrix Genome-Wide Human SNP arrays (250K NspI and SNP 6.0), FISH and genomic PCR to study 106 cases of adult *BCR-ABL1*-positive ALL. The most frequent somatic copy number alteration was a focal deletion on 7p12 of *IKZF1*, which encodes the transcription factor Ikaros and was identified in 80 of 106 patients (75%). Different patterns of deletions occurred but the most frequent were those characterised by a loss of exons 4 through 7 (*Δ4-7*) and by removal of exons 2 through 7 (*Δ2-7*). A variable number of nucleotides (patient-specific) were inserted at the conjunction and maintained with fidelity at the time of relapse. The extent of the *Δ4-7* deletion correlated with the expression of a dominant-negative isoform with cytoplasmic localisation and oncogenic activity, while the *Δ2-7* deletion resulted in a transcript lacking the translation start site. The *IKZF1* deletion was also identified in the progression of chronic myeloid leukaemia (CML) to lymphoid blast crisis (66%), but never in myeloid blast crisis or chronic phase CML, or in acute myeloid leukaemia patients. Known DNA sequence and structural features were mapped along the breakpoint cluster regions including heptamer recombination signal sequences (RSS) recognised by RAG enzymes during V(D)J recombination, suggesting that *IKZF1* deletions could arise from aberrant RAG-mediated recombination.

Introduction

Acute lymphoblastic leukaemia (ALL) is a heterogeneous disease characterised by multiple subtypes[1]. The Philadelphia chromosome (Ph) results from a reciprocal translocation fusing the *abelson* (*ABL1*) proto-oncogene from chromosome 9 with the *breakpoint cluster region* (*BCR*) sequences on chromosome 22, creating the Bcr-Abl fusion protein, a constitutively activated form of the Abl tyrosine kinase[2]. The Ph chromosome is considered the hallmark of chronic myeloid leukaemia (CML), but it is also the most frequent cytogenetic aberration associated with ALL, found in 20% to 40% of patients with ALL and in more than 50% of patients aged 50 years or older[3]. The presence of the *BCR-ABL1* rearrangement worsens the prognosis of ALL and represents the most significant adverse prognostic marker that influences disease outcome[4]. Ph-positive (Ph+) ALL is a more aggressive disease than CML, indicating that factors other than *BCR-ABL1* are involved in its development and progression[4-6]. Using high-resolution single nucleotide polymorphism (SNP) arrays, Mullighan et al.[7] recently described a deletion on 7p12 of *IKZF1*, which encodes the transcription factor Ikaros, in 83.7% of *BCR–ABL1* ALL cases, but not in chronic-phase CML, suggesting that loss of Ikaros[8-11] is an important step in the progression of *BCR-ABL1*-positive ALL. Ikaros is the prototypical member of the Kruppel-like zinc finger transcription factor subfamily, which is required for normal haematopoietic differentiation and proliferation, particularly in lymphoid lineages[8,12,13]. The Ikaros gene is transcribed as a number of isoforms due to alternative splicing, essentially altering the expression of exons 3 through 5 that encode the N-terminal DNA-binding domain. The long isoforms (Ik1 through Ik3), which have at least three zinc fingers, can bind efficiently to DNA, unlike the shorter isoforms (Ik4 to Ik8) that lack zinc fingers and instead behave as dominant negative isoforms upon homodimerisation, as well as heterodimerisation[14]. The most frequent alteration involving the *IKZF1* gene was a deletion of an internal subset of exons from 4 through 7 (Δ4–7) (Ikaros genomic organisation and exon designation is according to Kaufman et al[15]). Since heptamer recombination signal sequences (RSSs) recognised by RAG enzymes during V(D)J

recombination were located immediately internal to the deletion breakpoints[16], the authors suggested that the *IKZF1* Δ4–7 deletion was the result of aberrant RAG-mediated recombination.

Here, we carried out high resolution interrogation of genomic copy number alterations (Affymetrix GeneChip® Human Mapping 250K NspI and Genome-Wide Human SNP 6.0 array GeneChip microarrays), FISH, genomic polymerase chain reaction (PCR) analysis, cloning and direct sequencing of the genomic profile of adult chronic and acute leukaemias, including 106 de novo *BCR-ABL1*-positive ALL adult patients. Our aim was to identify oncogenic lesions that escape standard cytogenetic observations and combine with *BCR-ABL1* to induce ALL. The focal deletions on 7p12 in the *IKZF1* gene were the most frequent somatic copy number alterations in *BCR-ABL1*-positive ALL (75%), were predominantly monoallelic, and gave rise in the majority of cases to an Ikaros isoform with cytoplasmic localisation and oncogenic activity or a transcript that lacked the translation start site. Furthermore, by mapping the genomic breakpoints, we found that two major deletions primarily accounted for *BCR-ABL1*-positive ALL: *Δ4-7*, characterised by loss of exons 4 through 7 (55% of *IKZF1* deleted patients or 42% considering all patients), and *Δ2-7*, removing exons 2 through 7 (24% of *IKZF1* deleted patients or 18% considering all patients). In a small proportion of cases we also found deletions involving the *IKZF1* promoter region (4%). Finally, heptamer RSSs were mapped along the breakpoint cluster regions, confirming that *IKZF1* deletions could arise from aberrant RAG-mediated recombination.

Patients and Methods

Patients

One hundred and six *BCR-ABL1*-positive adult ALL patients were analyzed by Affymetrix Genome-wide Human SNP arrays (78 patients by 250K NspI and 28 patients by SNP 6.0; 2 patients by both). The median age was 53 years (range:18-76) and the median blast percentage at diagnosis was 90% (range: 18-99). The characteristics of the patients are shown in Table 1S. Diagnosis of all ALL cases was made on the basis of morphologic, biochemical and immunologic features of the leukemic cells. In addition, the human lymphoblastoid SD-1 and the human B-cell precursor leukaemia BV-173 cell lines were also included in the analysis. Human cell lines were obtained from DMSZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany) and maintained in culture following the DMSZ recommendations.

Single nucleotide polymorphism (SNP) microarray analysis

Genomic DNA was extracted using the DNA Blood Mini Kit (Qiagen, Valencia, CA) from mononuclear cells (MNCs) isolated from peripheral blood or bone marrow aspirates samples by Ficoll gradient centrifugation. DNA was quantified using the Nanodrop Spectrophotometer and quality was assessed using the Nanodrop and by agarose gel electrophoresis.

Samples were genotyped with GeneChip® Human Mapping 250K NspI and Genome-Wide Human SNP 6.0 array microarrays (Affymetrix, Santa Clara, CA) and according to the manufacturer's instructions. For all samples, 250 ng of DNA was digested with NspI (New England Biolabs, Boston, MA) and 250 ng with Sty enzymes for cases analyzed by SNP 6.0 array GeneChip. Digested DNA was adaptor-ligated and PCR-amplified using Clontech Titanium TAQ DNA polymerase (CELBIO) in four 100µl PCR reactions for NspI and three for StyI. PCR products from each set of reactions were pooled, purified and fragmented. Fragmented PCR products were then labeled, denatured and hybridized to the arrays. Arrays

were then washed using Affymetrix fluidics stations, and scanned using the Gene Chip Scanner 3000. Array image data was analyzed using Affymetrix GCOS 1.4 operating software and Genotyping Console 3.0 to derive .cel data files, which were also exported to the Partek Genomic Suite (Partek Inc, Saint Louise, MO) for further data visualization and analysis. Copy number aberrations were scored using the Hidden Markov Model and the segmentation approach available within the Partek software package as well as by visual inspection. For the analysis by HMM we set the following parameters: max probability (specify the max probability of retaining the same state between neighbouring observations) equal to 0.9; genomic decay (the decay parameter describes now quickly the HMM retention of state will decay towards the initial probability) equal to 0 and sigma (specify the gaussian bandwidth of the distribution from which observation are drawn equal to 0.8). All aberrations were calculated with respect to a set of 270 Hapmap normal individuals and a set of samples obtained from acute leukaemia cases in remission in order to reduce the noise of raw copy number data. When available, in order to exclude inherited copy number variants a comparison to paired constitutional DNA and to paired remission DNA was performed. Copy number aberrations involving *IKZF1* gene were also analyzed and eventually confirmed using Genotyping Console 3.0 (Affymetrix, Santa Clara, CA). SNP array files are available for free download at http://gebbama-prod.cineca.it/MicroarrayRepository/MicroarrayRepository.htm using the following username and password: *public public*.

IKZF1 genomic quantitative PCR (Q-PCR)

IKZF1 genomic quantitative PCR (Q-PCR) of all *IKZF1* coding exons was performed as described by Mullighan et al.[7] using a 7900 Real-Time PCR system and 7900 System Software (Applied Biosystems, Foster City, CA).

IKZF1 RT-PCR

Mononuclear cells were separated by Ficoll-Hypaque density gradient centrifugation and samples stored at -190°C in RPMI 1640 with 20% FBS and 10% dimethylsulfoxid or in GTC (guanidine thiocyanate) at -80°C, as needed. Total cellular RNA was extracted from cells using the RNeasy total RNA isolation kit (Qiagen, Valencia, CA) according to the instructions of the manufacturer, and 1µg of the total RNA sample was used for cDNA synthesis with Moloney murine leukaemia virus reverse transcriptase (Invitrogen) in the presence of dNTPs. Polymerase chain reaction (RT-PCR) using primers specific for exon 1 and exon 8 (in Table 2S A3 and A5, respectively) of *IKZF1* and nucleotide sequencing were performed to identify the specific Ikaros isoforms as previously described[17]. RNA integrity was confirmed by PCR amplification of the *GAPDH* mRNA, which is expressed ubiquitously in human hematopoietic cells.

IKZF1 gene expression analysis was performed using Syber Green PCR master Mix (Applied Biosystems, Foster City, CA) and primers B1 and B2 in Table 2S. Taqman Syber Green assays were performed using a 7900 Real-Time PCR system and 7900 System Software (Applied Biosystems, Foster City, CA). The quantitative PCR thermal protocol consisted of: 50°C for 2 minutes, followed by 95°C for 10 minutes, then 40 cycles of 95°C for 1 minute and 60°C for 1 minute. *GAPDH* was used as control gene, using primers E3-E4 in Table 2S.

Fluorescence in situ hybridization (FISH)

FISH analysis was performed as previously described[18]. Briefly, chromosome preparations from bone marrow cells were hybridized in situ with 1 µg of probe labeled by nick translation. Hybridization was performed at 37°C in 2X SSC, 50% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, 5 µg COT1 DNA (Bethesda Research Laboratories, Gaithersburg, MD, USA), and 3 µg sonicated salmon sperm DNA in a

volume of 10 µL. Post-hybridization washing was performed at 60°C in 0.1X SSC (three times). In cohybridization experiments, the probes were directly labeled with Fluorescein and Cy3. Chromosomes were identified by DAPI staining. Digital images were obtained using a Leica DMRXA epifluorescence microscope equipped with a cooled CCD camera (Princeton Instruments, Boston, MA). Cy3 (red; New England Nuclear, Boston, MA, USA), fluorescein (green; Fermentas Life Sciences, Milan, IT), and DAPI (blue) fluorescence signals, which were detected using specific filters, were recorded separately as gray-scale images. Pseudocoloring and merging of images were performed with Adobe Photoshop software.

Probes

The whole chromosome paints (WCP) used for chromosomes 7, derived from flow-sorted chromosomes, was a gift of the Sanger Centre (Dr Nigel Carter).

Fosmid probes specific for the *IKZF1* gene [G248P800745C8 (chr7:50,381,496-50,422,338) and G248P87926C7 (chr7:50,418,455-50,458,507)], as well as a bacterial artificial chromosome (BAC) probe specific for *BCR* [RP11-164N13 (chr22:21,897,904-22,091,572)], where properly selected accordingly to the latest release (March 2006) of the University Santa Cruz (UCSC) Human Genome Browser (http://genome.ucsc.edu/).

Cloning of deletion junctions

The information derived from the SNP-array experiments were utilized to design appropriate forward and reverse primers mapping in the distal and proximal breakpoint regions and used in long-range PCR experiments (Table 2S), aimed at sequencing the deletion junction regions. In case of positive results, nested PCR experiments were performed. The long-range PCR analyses were performed using TaKaRa LA Taq™ (Cambrex Bio Science, Milan, Italy), and the nested PCR products, containing the junction

fragments, were sequenced using the Big Dye Terminator Sequencing kits (Applied Biosystems). In the cases of breakpoints occurred in intron 1 (*Δ2-7* deletion) the primers C1/C2 and C5 were used in long PCR experiments and then we chose the primers C3/C4 (sense) and C5 (anti-sense) to obtain a smaller amplified. PCR was performed using 100 ng of genomic DNA in a 25-μL reaction with 2U of FastStart Taq DNA Polymerase (Roche Diagnostics, Mannheim, Germany). The long-range PCR cycling parameters were as follows: 95°C for 5 min (complete denaturation), followed by 40 cycles at 95°C for 30 s, 62°C for 30 s, extension at 72°C for 90 s, with an additional extension at the end for 7 min. This protocol was used to easily screen for *IKZF1* deletions all patients and confirm SNP results. In the case of *Δ4-7* deletion, we used the primers D1 and D2 for genomic PCR experiments and to perform an extensive screening. The PCR protocol was the same before reported. Reference genome sequence data were obtained from the UCSC browser (http://genome.ucsc.edu/cgi-bin/hgGateway?db=hg18; March 2006 release), and sequence comparison was performed using the BLAST software tool (www.ncbi.nlm.nih.gov/BLAST/).

Western Blotting

Cells were lysed with sample buffer (2%SDS in 125 mM Tris HCL, pH 6.8). Cell lysates were subjected to SDS-PAGE on 12% gels and then transferred to nitrocellulose membranes (Amersham Biosciences). The blots were incubated for 60 min in Odyssey blocking buffer before incubation overnight (4°C) with polyclonal anti-Ikaros antibody (Santa Cruz). Blotted proteins were detected and quantified using the Odyssey infrared imaging system LI-COR.

Subcellular Localization Studies Using Confocal Laser Scanning Microscopy

The subcellular localization of Ikaros protein(s) was examined by immunofluorescence and confocal laser scanning microscopy as previously described[17].

Bioinformatic analysis

Known motifs were searched within the *IKZF1* breakpoint regions (intron1_2 and intron7_8), using the following algorithms:

EMBOSS/fuzznuc (http://emboss.sourceforge.net/apps/release/5.0/emboss/apps/fuzznuc.html), in order to search for the regular expression pattern match of the motifs listed in Supplementary Table 3; EMBOSS/marscan (http://emboss.sourceforge.net/apps/release/5.0/emboss/apps/marscan.html), to find out Matrix/Scaffold Attachment Regions (MAR/SAR) sites; MAR-Wiz (http://www.futuresoft.org/MAR-Wiz/), to disclose the occurrence of DNA patterns that have been known to occur in the neighbourhood of MARs; RepeatMasker (http://www.repeatmasker.org/), to detect the presence of known human interspersed repeated elements.

Results

SNP microarray analysis detected frequent and recurrent deletions in IKZF1 gene

To identify oncogenic lesions that combine with *BCR-ABL1* to induce ALL, by high resolution SNP array we profiled the genomes of 106 *BCR-ABL1*-positive ALL patients. We identified regions of high level amplification and homozygous deletion in all patients, with a mean of 10 somatic copy number alterations per case (range, 22-2), with 2 gains (range, 1-10) and 6.5 losses (range, 2-20). Deletions outnumbered amplifications almost 3:1. Lesions varied from loss or gain of complete chromosome arms (trisomy 4, monosomy 7, loss of 9p, 10q, 14q, 16q and gain of 1q and 17q) to microdeletions and microduplications targeting genomic intervals. Deletions of known leukaemia-associated genes occurred with a high frequency, as well as deletions in genes involved in the G1/S cell cycle checkpoint and B-cell lineage differentiation (e.g., *CDKN2A* in 42%, *BTG1* in 19%, *PAX5* in 29%, *EBF1* in 5%). A detailed list of the alterations is shown in Table 4S. The most frequent somatic copy number alteration was a deletion

on 7p12 involving the *IKZF1* gene (80/106 adult *BCR-ABL1* ALL cases, 75%) encoding the transcription factor Ikaros, which is required for the earliest stages of lymphoid lineage commitment and acts as a tumour suppressor in mice[8]. Deletions involving the *IKZF1* gene were mono-allelic in 85% of cases and were limited to the gene in all patients, except for one patient (BCR-ABL1-#47) and patients with monosomy of chromosome 7 (Fig. 1 and Table 5S).

IKZF1 deletions were confirmed by FISH analysis and quantitative PCR

SNP array data were confirmed using the overlapping fosmid clones G248P800745C8 (chr7:50,381,496-50,422,338), G248P82319E12 (chr7:50,400,870-50,443,154), and G248P87926C7 (chr7:50,418,455-50,458,507) in co-hybridisation with a probe encompassing the *BCR* gene (BAC clone RP11-164N13, chr22:21,897,904-22,091,572), in order to specifically identify Ph-positive cells.

Patients BCR-ABL1-#24 and BCR-ABL1-#28 showed a heterozygous deletion for the fosmid G248P800745C8 in Ph-positive cells only (Fig. 2A). Patients BCR-ABL1-#17, BCR-ABL1-#30 and BCR-ABL1-#50 showed homozygous *IKZF1* deletions (Fig. 2B). Interestingly, patient BCR-ABL1-#17 showed two deletions with a different proximal breakpoint (Fig. 2A and 2B) in Ph-positive cells; Ph-negative cells were normal. FISH analysis not only confirmed the SNP array results but also demonstrated that *BCR-ABL1* rearrangements and *IKZF1* deletions were in the same clones and, further, that all Ph+ metaphases also had the *IKZF1* deletion.

Real-time genomic quantitative PCR (Q-PCR) of promoter and all *IKZF1* exons was performed in order to confirm SNP results and to characterise the focal deletions. A big concordance between the extension of deletions previously identified by SNP array and the Q-PCR results was found. Furthermore, Q-PCR allowed us to determine the extension of homozygous deletions, suggesting that in some cases different proximal and distal breakpoints can occur on the two alleles (Table 6S). Q-PCR on Ikaros mRNA demonstrated that the deletion of *IKZF1* correlated with a down-modulation at the transcript level of the

Ikaros DNA-binding isoforms (Fig. 1S).

Two types of deletions in the IKZF1 gene account for most cases of BCR-ABL1-positive acute lymphoblastic leukaemia

The information derived from the SNP-array, genomic quantitative PCR and FISH experiments was utilised to design appropriate forward and reverse primers mapping in the distal and proximal breakpoint regions, which were used in genomic PCR and sequencing experiments aimed at defining the deletion junction regions and at investigating the mechanisms responsible for *IKZF1* deletions. This combined approach (genomic PCR and direct sequencing) has been demonstrated to be a power tool to identify *IKZF1* deletions in those cases in which SNP array analysis has failed.

The ⊿4-7 deletion was identified in 42% of patients with BCR-ABL1-positive acute lymphoblastic leukaemia

In 44/106 patients (42%), the deletion of *IKZF1* was limited to exons 4-7 (*⊿4-7*), with breakpoints occurring in the region ranging from 50380371 to 50380388 and from 50431123 to 50431167 in introns 3 and 7 (Fig. 2S), respectively, on chromosome 7p12. A variable number of patient-specific nucleotides were inserted at the conjunction, and cloning analysis showed that the breakpoints were the same in each *BCR-ABL1*-positive cell (Tables 7S and Fig. 3A-B). As previously shown[7], the extent of the deletion correlated with the expression of a dominant-negative isoform, Ik6, with cytoplasmic localisation and oncogenic activity (Fig. 3S). The Ik6 transcript and protein, which lack exons 4–7 and all zinc finger motifs, respectively, were exclusively observed in cases harbouring the *IKZF1* Δ4–7 deletion. This finding validates the concept that the expression of non-DNA-binding Ikaros isoforms is due to *IKZF1* genomic abnormalities and not aberrant post-transcriptional splicing induced by *BCR-ABL1*[19].

The Δ2-7 deletion (Type B) was identified in 18% of patients with BCR-ABL1-positive acute lymphoblastic leukaemia

In 19/106 patients (18%), the *IKZF1* deletion involved exons 2-7 (*Δ2-7*), with a variable pattern of breakpoints in intron 1 (from 50317112 to 50317936) not previously reported (Tables 8S and Fig. 4A-B), while the breakpoints in intron 7 were in the same region of those of the Δ4-7 deletion. There was a correlation between the extension of this deletion and the expression of an aberrant transcript containing only exons 1 and 8, which lacked the translation start (Fig. 4S).

Other deletions may occur in the IKZF1 gene at low frequency

Other deletions in the *IKZF1* gene were observed at low frequency. Interestingly, three patients had a homozygous *IKZF1* deletion with two different proximal and distal breakpoints, giving rise to Δ4-7 and Δ4-8 deletions (BCR-ABL1-#17), and Δ2-7 and Δ4-7 (BCR-ABL1-#50 and BCR-ABL-#95) (Table 9-10S). In one patient (BCR-ABL-#36), we identified a deletion of the entire *IKZF1* gene and part of *GRB10*, and in two patients (BCR-ABL1-#47 and BCR-ABL1-#71), we identified a smaller deletion spanning the whole promoter and first exon of *IKZF1*. Another uncommon deletion was that spanning from promoter to exon 7 (Table 5-6S). Overall, deletions involving the *IKZF1* gene were identified in 68/106 (64%) *BCR-ABL1*-positive ALL patients. It is important to consider that in 13 patients, we had a monosomy 7; in these cases, therefore, there was a deletion of the entire *IKZF1* gene on one allele. In only one case, we also found the *Δ4-7* deletion on the other allele. Since monosomy 7 results in Ikaros haploinsufficiency – like many focal deletions – if we include these cases in our analysis, the rate of patients with *IKZF1* loss and/or haploinsufficiency increases to 75% (80/106).

IKZF1 deletion is not detected in other leukaemias except for lymphoid blast crisis CML

To investigate the possibility that *IKZF1* alteration is a frequent event across different leukaemias, we

used high resolution SNP arrays and genomic PCR to examine 30 CML chronic phase patients, 10 CML in blast crisis patients (3 lymphoid and 7 myeloid), 28 adult AML patients with different FAB classification, and 2 hypereosinophilic syndrome patients (Table 11S). No focal deletions were identified in these leukaemias, except for two patients with CML in lymphoid blast crisis. In particular, we found the *Δ4-7* deletion in one case (CML-ly-BC#1) and the *Δ2-7* deletion in another case (CML-ly-BC#3), suggesting a role for *IKZF1* genomic alterations in the progression from chronic phase to blast crisis in CML, given that the deletion was not detected at diagnosis (Table 10S).

Recombination signal sequences (RSS) flanked IKZF1 genomic breakpoints

To investigate the mechanisms responsible for the generation of the *IKZF1 Δ4-7* and *Δ2-7* deletions in *BCR-ABL1*-positive leukaemia, we searched within the *IKZF1* breakpoint regions for regular expression pattern matches of the motifs listed in Table 3S such as AID (activation induced cytidine deaminase) motif (DGYW/WRCH), V(D)J RSS, topoisomerase II binding site, translin binding site, Chi-like sequence, PurI binding site, Eukaryotes replication origin sequence, ARS-S cerevisiae, Putative triple helices, Pyrimidine trait, Human minisatellite core sequence, and Satellite III core. As mentioned[7], RSSs recognised by RAG enzymes during V(D)J recombination were located immediately internal to the deletion breakpoints, and a variable number of additional nucleotides (patient-specific) were present between the consensus intron 3 and 7 sequences, suggestive of the action of terminal deoxynucleotidyl transferase (TdT). Interestingly, the breakpoints in the *IKZF1* gene and the additional nucleotides inserted at the junction were maintained with fidelity at the time of relapse and could be exploited for monitoring minimal residual disease. We also noted that AID consensus sequences (DGYW/WRCH corresponding to AGTA and TGTT or TGTA in our cases)[18] flanked the genomic breakpoints in both intron 3 and intron 7 (Table 6S). Furthermore, AID consensus sequences were found very close to the breakpoints in intron 1 in the case of the *Δ2-7* deletion (Fig. 5S). Since there was no evidence of

nucleotide changes to suggest AID activity and since the AID consensus sequence is highly degenerate and may likely occur very frequently by chance, we hypothesised that aberrant RAG activity could be responsible for the *IKZF1* deletions.

Discussion

The Philadelphia chromosome encodes the oncogenic Bcr-Abl kinase and defines a subgroup of ALL with a particularly unfavourable prognosis[1]. The reasons for the aggressive nature of *BCR-ABL1*-positive ALL are still under investigation and have not yet been elucidated. In order to identify additional oncogenic lesions involved in the generation of *BCR-ABL1*-positive ALL and responsible for its poor outcome and biological difference from CML, we performed a high resolution genomic study of copy number alterations in a large cohort of adult patients with acute and chronic leukaemias, including 106 adult patients with *BCR-ABL1*-positive leukaemia. We found that homozygous or heterozygous deletions in the *IKZF1* gene frequently occur only in *BCR-ABL1*-positive ALL (75%) patients or lymphoid blast crisis CML (66%), as previously demonstrated by Mullighan et al.[7] in 21 paediatric and 22 adult *BCR–ABL1* ALL patients.

IKZF1 encodes a zinc finger protein required for lymphoid lineage differentiation, proliferation and function[19,20]. Ikaros contains two separate regions with zinc-finger domains: 4 DNA-binding zinc fingers (ZnFs) near the N-terminus and 2 ZnFs for protein–protein interactions near the C-terminus. Alternative splicing can generate multiple functionally different Ikaros isoforms that lack variable numbers of internal exons. Isoforms that lack the N terminal ZnFs are unable to bind transcriptional targets normally but retain the C-terminal ZnFs and the ability to dimerise and act as dominant negative inhibitors of Ikaros function. Ikaros transgenic and mutant mouse models have clearly demonstrated the important role of Ikaros in both normal haematopoiesis and tumour suppression. Mutant Ikaros −/− mice have severe lymphoid cell defects[21,22], but heterozygous Ikaros DN +/− mice invariably develop T-cell malignancies. During the last decade, several groups have reported the expression of aberrant Ikaros isoforms due to alternative splicing in different acute leukemias[11,23-25]. In 2006, Klein et al. suggested that the expression of aberrant Ikaros isoforms occurred at a post-transcriptional level as a result of the action of the Bcr-Abl fusion protein. This hypothesis was reversed by the findings of Mullighan et al.[7]

and by our study, which strongly demonstrated that intragenic deletions in the *IKZF1* gene were responsible for the generation of different aberrant isoforms. Here, we characterised and mapped all breakpoints, demonstrating that two frequent deletions occur in *BCR-ABL1*-positive lymphoid leukaemias: *Δ4-7* (42%), removing exons 4 through 7, and *Δ2-7* (18%), resulting in a transcript with only exons 1 and 8.

An elevated frequency of genomic aberrations could be directly caused by an abnormally high incidence of DNA double-strand breaks. In normal cells, DNA lesions are detected and repaired by sophisticated physiological machinery and a system of cell cycle checkpoints, preventing cells that have sustained DNA damage from proliferating further. In this study, we performed a bionformatic analysis in order to determine the mechanisms responsible for the generation of *IKZF1* deletions. Known motifs associated with deletions were searched within the *IKZF1* breakpoint regions using different algorithms, and known DNA sequence and structural features were mapped along the breakpoint cluster regions, including heptamer RSSs, suggesting that *IKZF1* deletions could arise from aberrant RAG -mediated recombination.

In conclusion, our results confirmed, as previously showed by Mullighan et al., that *IKZF1* deletion is the most frequent somatic copy number alteration in Ph+ ALL, and they provided new details on other common breakpoints and on the mechanisms involved in the rearrangement. It is likely that Ikaros loss combines with *BCR–ABL1* to induce lymphoblastic leukaemia, arresting B-lymphoid maturation.

Acknowledgments: This work was supported by**:** AIL, European LeukaemiaNet, AIRC, Fondazione Del Monte di Bologna e Ravenna, FIRB 2006, Ateneo 60% grants and Gimema Onlus Working Party ALL and CML. We thank Formica Serena (University of Bologna), Ferrari Anna (University of Bologna) and Messina Monica (University of Rome) for helpful in performing SNP arrays.

Author Contributions

I.I. designed research, performed SNP array analysis and drafted the manuscript;

C.T.S., L.I., A.L. and C.B. performed molecular and conventional cytogenetic analysis;

P.D., performed bioinformatic analysis;

E.O. and A.A. contributed to SNP array analysis;

S.S., A.L., C.P., S.C., P.P.P., S.P., D.R., F.P., G.S., M.B. contributed in the development of the study;

S.C., A.V., M.V, R.F. contributed in data collection and data analysis;

D.C. and F.M. contributed in proteomic analysis and data analysis;

G.M. designed research and gave final approval to the manuscript.

Authors' Disclosures of Potential Conflicts of interest

The authors have no conflicts of interest to disclose

References

1. Faderl S, Jeha S, Kantarjian HM. The biology and therapy of adult acute lymphoblastic leukemia. Cancer. 2003;98:1337-1354.
2. Sawyers CL, Gishizky ML, Quan S, Golde DW, Witte ON. Propagation of human blastic myeloid leukemias in the SCID mouse. Blood. 1992;79:2089-2098.
3. Larson ED, Maizels N. Transcription-coupled mutagenesis by the DNA deaminase AID. Genome Biol. 2004;5:211.
4. Faderl S, Kantarjian HM, Thomas DA, et al. Outcome of Philadelphia chromosome-positive adult acute lymphoblastic leukemia. Leuk Lymphoma. 2000;36:263-273.
5. Radich JP. Philadelphia chromosome-positive acute lymphocytic leukemia. Hematol Oncol Clin North Am. 2001;15:21-36.
6. Cornelissen JJ, Carston M, Kollman C, et al. Unrelated marrow transplantation for adult patients with poor-risk acute lymphoblastic leukemia: strong graft-versus-leukemia effect and risk factors determining outcome. Blood. 2001;97:1572-1577.
7. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008;453:110-114.
8. Georgopoulos K, Bigby M, Wang JH, et al. The Ikaros gene is required for the development of all lymphoid lineages. Cell. 1994;79:143-156.
9. Molnar A, Georgopoulos K. The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. Mol Cell Biol. 1994;14:8292-8303.
10. Kirstetter P, Thomas M, Dierich A, Kastner P, Chan S. Ikaros is critical for B cell differentiation and function. Eur J Immunol. 2002;32:720-730.
11. Sun L, Goodman PA, Wood CM, et al. Expression of aberrantly spliced oncogenic ikaros isoforms in childhood acute lymphoblastic leukemia. J Clin Oncol. 1999;17:3753-3766.
12. Molnar A, Wu P, Largespada DA, et al. The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse. J Immunol. 1996;156:585-592.
13. Westman BJ, Mackay JP, Gell D. Ikaros: a key regulator of haematopoiesis. Int J Biochem Cell Biol. 2002;34:1304-1307.
14. Georgopoulos K. Haematopoietic cell-fate decisions, chromatin regulation and ikaros. Nat Rev Immunol. 2002;2:162-174.
15. Kaufmann C, Yoshida T, Perotti EA, Landhuis E, Wu P, Georgopoulos K. A complex network of regulatory elements in Ikaros and their activity during hemo-lymphopoiesis. Embo J. 2003;22:2211-2223.
16. Soulas-Sprauel P, Rivera-Munoz P, Malivert L, et al. V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining. Oncogene. 2007;26:7780-7791.
17. Iacobucci I, Lonetti A, Messa F, et al. Expression of spliced oncogenic Ikaros isoforms in Philadelphia-positive acute lymphoblastic leukemia patients treated with tyrosine kinase inhibitors: implications for a new mechanism of resistance. Blood. 2008.
18. Iacobucci I, Ottaviani E, Astolfi A, et al. High-resolution genomic profiling of Ph-positive acute lymphoblastic leukaemia (ALL) identified recurrent copy number anomalies in genes regulating the cell cycle and the B-cell differentiation. Haematologica. 2008;93:1084.
19. Klein F, Feldhahn N, Herzog S, et al. BCR-ABL1 induces aberrant splicing of IKAROS and lineage infidelity in pre-B lymphoblastic leukemia cells. Oncogene. 2006;25:1118-1124.
20. Rogozin IB, Diaz M. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C

bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. J Immunol. 2004;172:3382-3384.
21. Smale ST, Dorshkind K. Hematopoiesis flies high with Ikaros. Nat Immunol. 2006;7:367-369.
22. Gomez-del Arco P, Maki K, Georgopoulos K. Phosphorylation controls Ikaros's ability to negatively regulate the G(1)-S transition. Mol Cell Biol. 2004;24:2797-2807.
23. Wang JH, Nichogiannopoulou A, Wu L, et al. Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity. 1996;5:537-549.
24. Winandy S, Wu P, Georgopoulos K. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. Cell. 1995;83:289-299.
25. Ishimaru F. Expression of Ikaros isoforms in patients with acute myeloid leukemia. Blood. 2002;100:1511-1512; 1512-1513.
26. Sun L, Crotty ML, Sensel M, et al. Expression of dominant-negative Ikaros isoforms in T-cell acute lymphoblastic leukemia. Clin Cancer Res. 1999;5:2112-2120.
27. Sun L, Heerema N, Crotty L, et al. Expression of dominant-negative and mutant isoforms of the antileukemic transcription factor Ikaros in infant acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 1999;96:680-685.

Figure legend:

Figure 1. Representation from Genotyping Console 3.0 (Affymetrix, Santa Clara, CA) of *IKZF1* deletions or monosomy 7 in 4 *BCR-ABL1*-positive ALL patients. Patients #23, #105, #103 have a mono-allelic deletion (DNA copy number state = 1) of different regions of *IKZF1*. Patient #95 shows a bi-allelic deletion. Normal diploid DNA content is shown as a continuous line of copy number state = 2. Red segment reports indicate the genomic extent of the *IKZF1* deletions. The opalescent vertical box indicates the gap (chr7:50338126-50378125) inside the *IKZF1* gene.

Figure 2. (A) **Map of the deletion proximal breakpoints within the *IKZF1* gene.**
Purple arrowheads indicate the breakpoint regions mapped by FISH experiments. Particularly, the left arrow corresponds to both Δ4-7 and Δ2-7 deletion proximal breakpoint regions, while the right arrow indicates a novel proximal breakpoint region identified in patient #17 (see below).
(B) **FISH**. FISH results obtained in some of the cases under study, showing homozygous (cases #17, 30, and 50) or heterozygous (cases #24 and 28) deletion of the fosmid clone G248P800745C8 (red) only in Ph positive cells (middle column), as shown by the FISH pattern of the RP11-164N13 (*BCR*) probe (in yellow). Clone G248P87926C7 (green) is always retained on deleted chromosomes 7. Notably, in patient #17, one of the two green signals is fainter than the other, revealing the occurrence of a partial deletion of G248P87926C7 in one of the two chromosome 7 homologs.

Figure 3. *IKZF1* Δ4-7 breakpoints. (A) Schematic representation of *Δ4-7* deletion with arrows indicating the region in which the breakpoints occur. In the graphs are shown the chromosome positions of breakpoints in the proximal (intron 3) and distal (intron 7) regions. The most frequent breakpoints occur at 50380384 and 50380388 in the proximal region and at 50431125 and 50431128 in the distal region. Primers used were D1 and D2 (Table 2S); products were then directly sequenced to characterize

the sequence flanking deletion breakpoints. **B**. Pherograms of sequencing of *IKZF1* Δ4-7 breakpoints. Regions matching the reference genomic *IKZF1* sequence are shown by arrows, separated by additional nucleotides not matching the consensus sequence.

Figure 4. *IKZF1* Δ2-7 breakpoints. (A) Schematic representation of *Δ2-7* deletion with arrows indicating the region in which the breakpoints occur. In the graphs are shown the chromosome positions of breakpoints in the proximal (intron 1) and distal (intron 7) regions. The most frequent breakpoints occur at 50317927 and 50317933 in the proximal region and at 50431128 in the distal region. Primers used were C3/C4 and C5 (Table 2S); products were then directly sequenced to characterize the sequence flanking deletion breakpoints. **B**. Pherograms of sequencing of *IKZF1* Δ2-7 breakpoints. Regions matching the reference genomic *IKZF1* sequence are shown by arrows, separated by additional nucleotides not matching the consensus sequence.

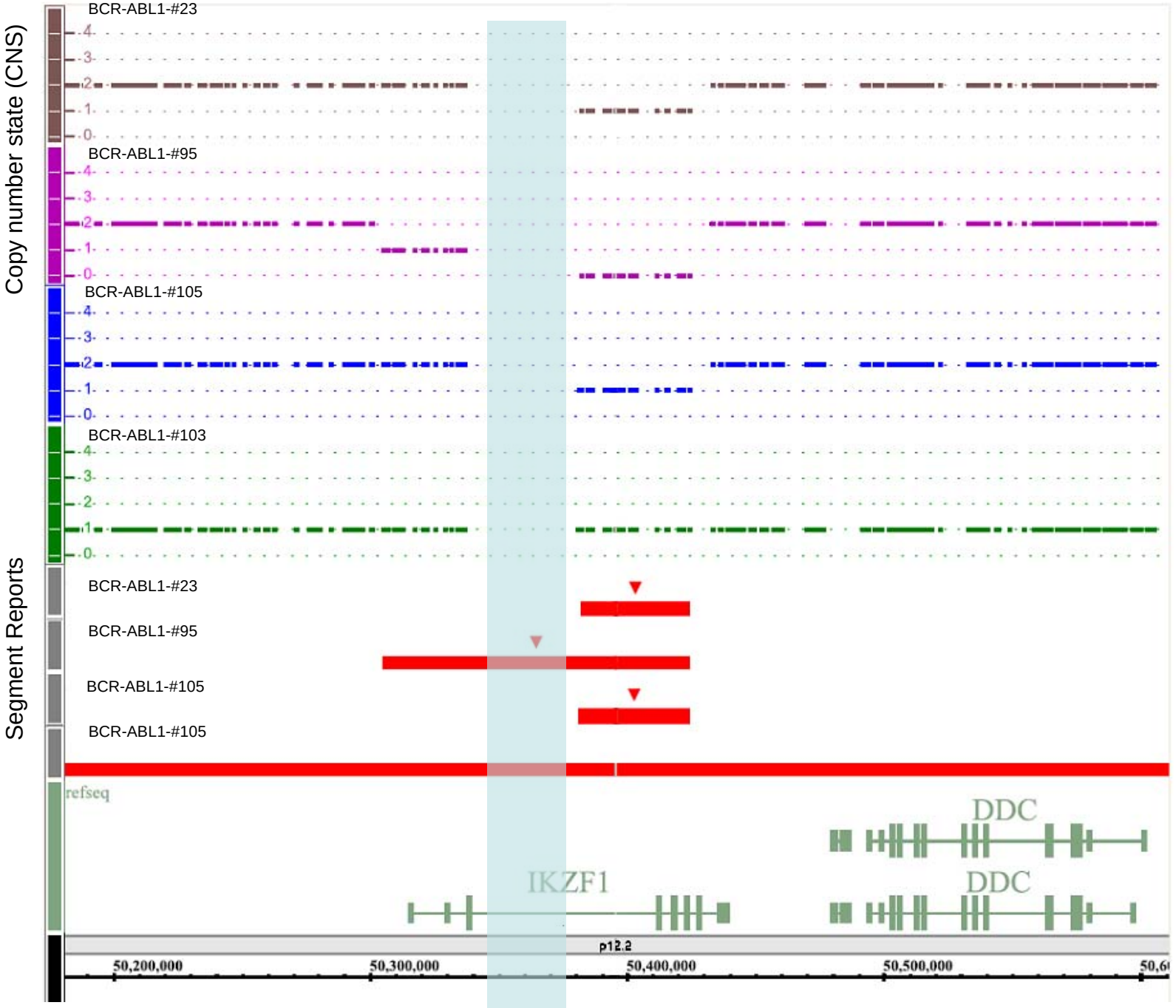

Copy number state (CNS)
BCR-ABL1-#23
BCR-ABL1-#95
BCR-ABL1-#105
BCR-ABL1-#103
Segment Reports
BCR-ABL1-#23
BCR-ABL1-#95
BCR-ABL1-#105
BCR-ABL1-#105
refseq
DDC
IKZF1
DDC
p12.2
50,200,000
50,300,000
50,400,000
50,500,000

Figure 1

Figure 2

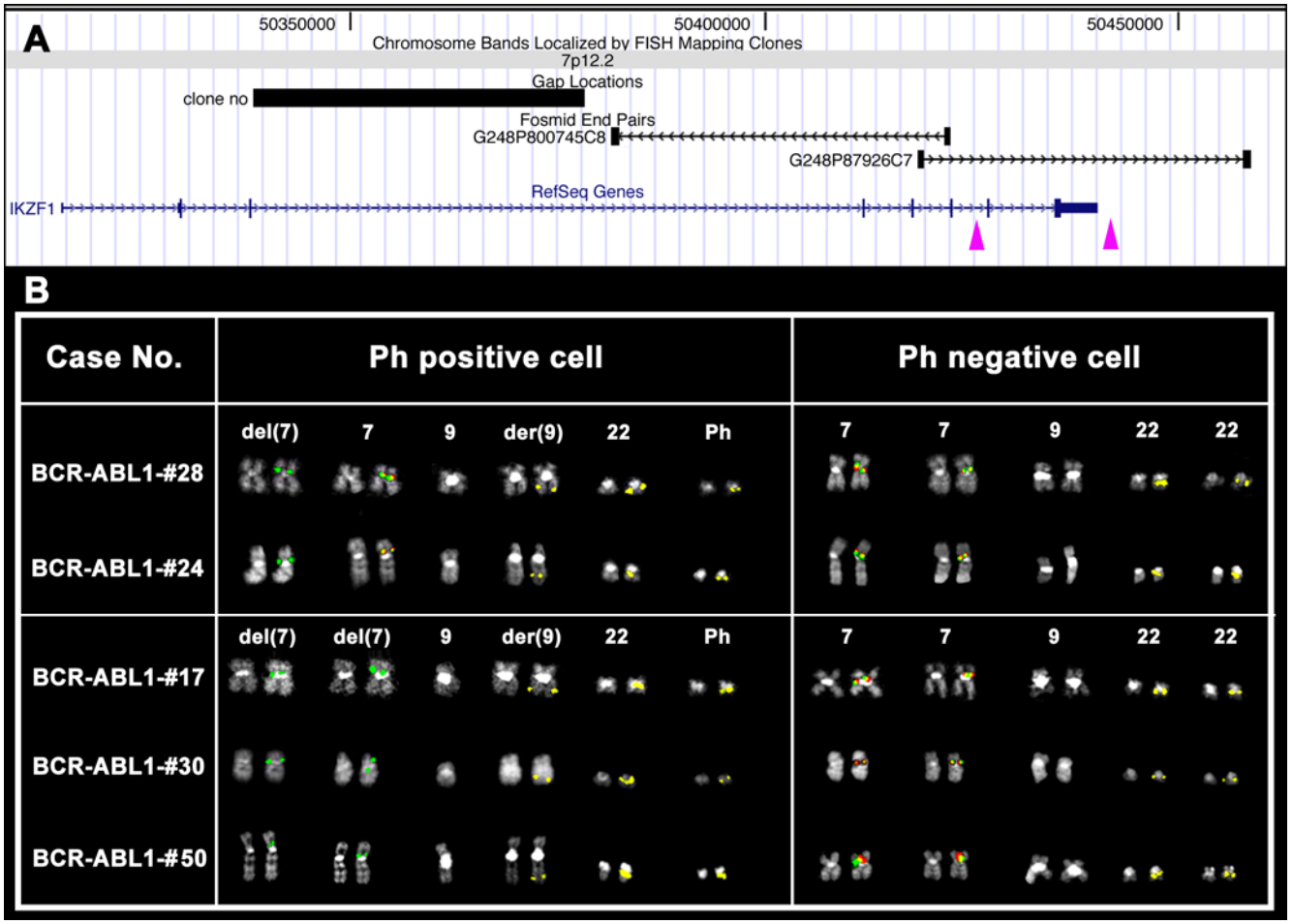

A
50350000
50400000
50450000
Chromosome Bands Localized by FISH Mapping Clones
7p12.2
Gap Locations
clone no
Fosmid End Pairs
G248P800745C8
G248P87926C7
RefSeq Genes
IKZF1
B
Case No.
Ph positive cell
Ph negative cell
del(7)
7
9
der(9)
22
Ph
BCR-ABL1-#28
BCR-ABL1-#24
BCR-ABL1-#17
BCR-ABL1-#30
BCR-ABL1-#50

Figure 3A

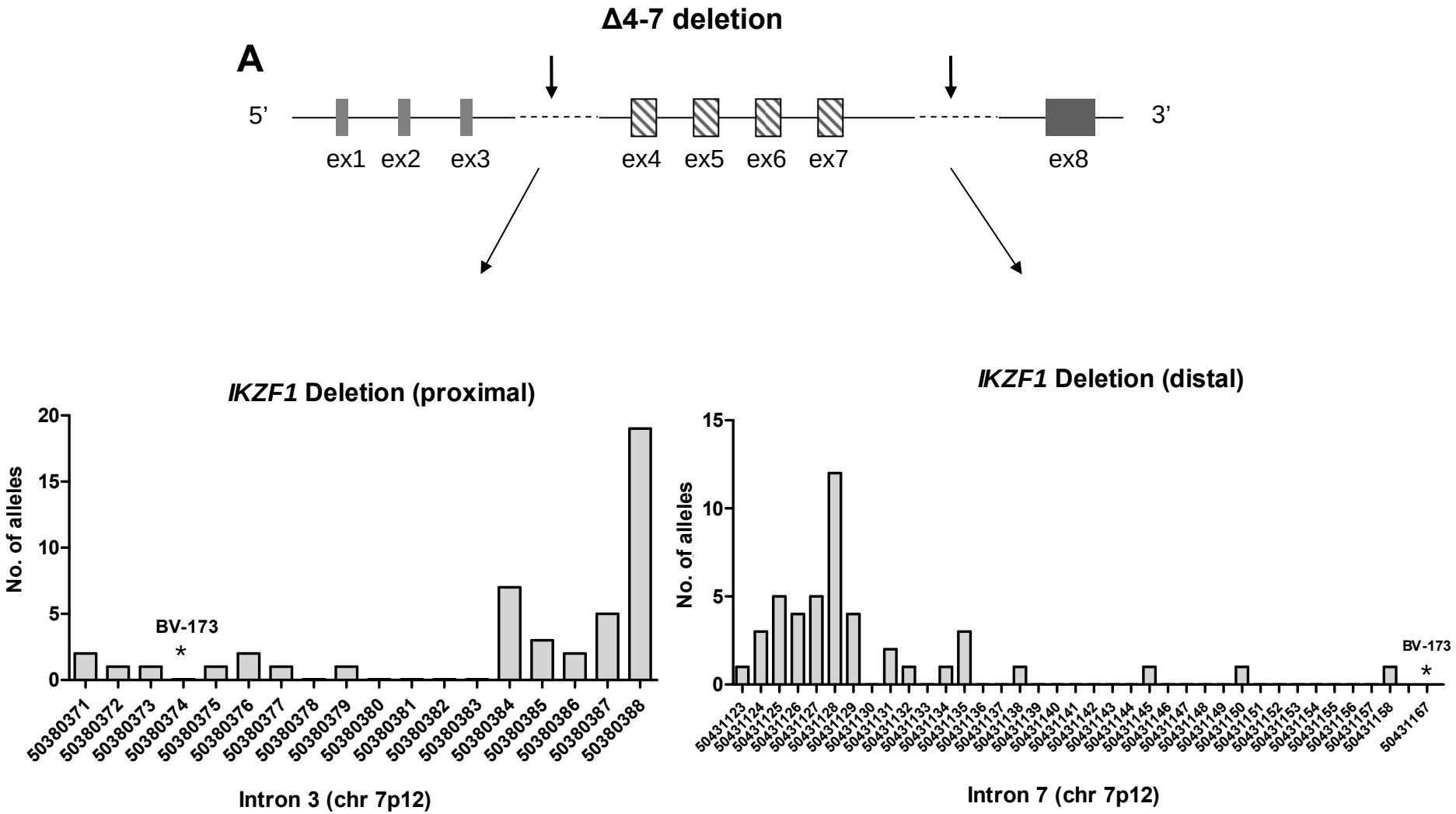
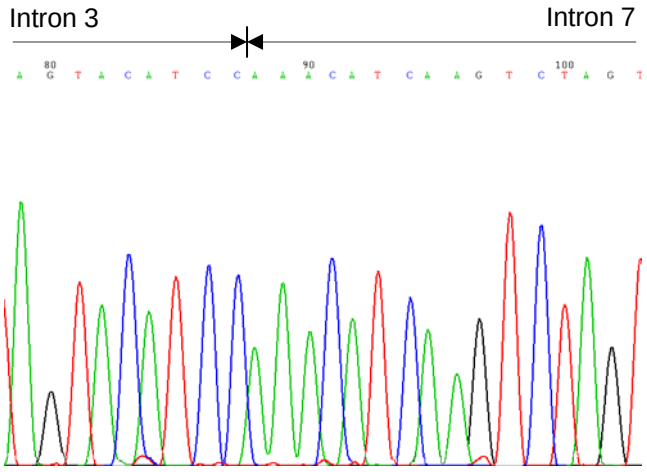
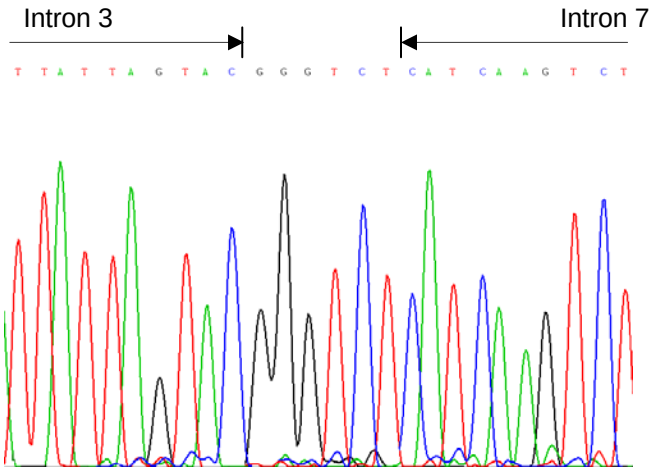

Δ4-7 deletion
A
5'
3'
ex1
ex2
ex3
ex4
ex5
ex6
ex7
ex8
IKZF1 Deletion (proximal)
No. of alleles
BV-173
Intron 3 (chr 7p12)
IKZF1 Deletion (distal)
No. of alleles
BV-173
Intron 7 (chr 7p12)

Figure 3B

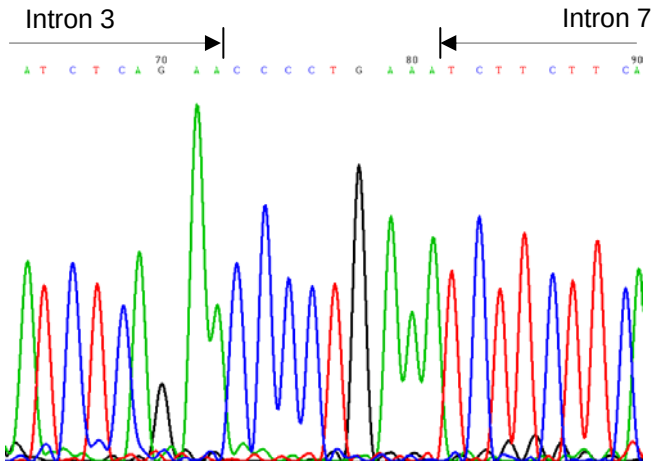
B . BCR-ABL1- #6



BCR-ABL1- #24



BCR-ABL1- #13



BCR-ABL1- #15

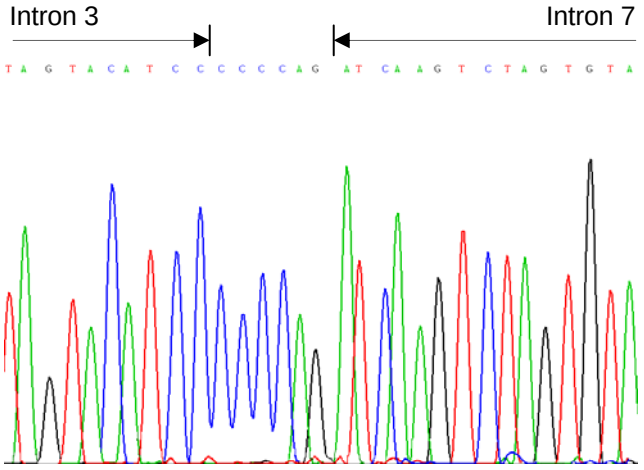


Figure 4A

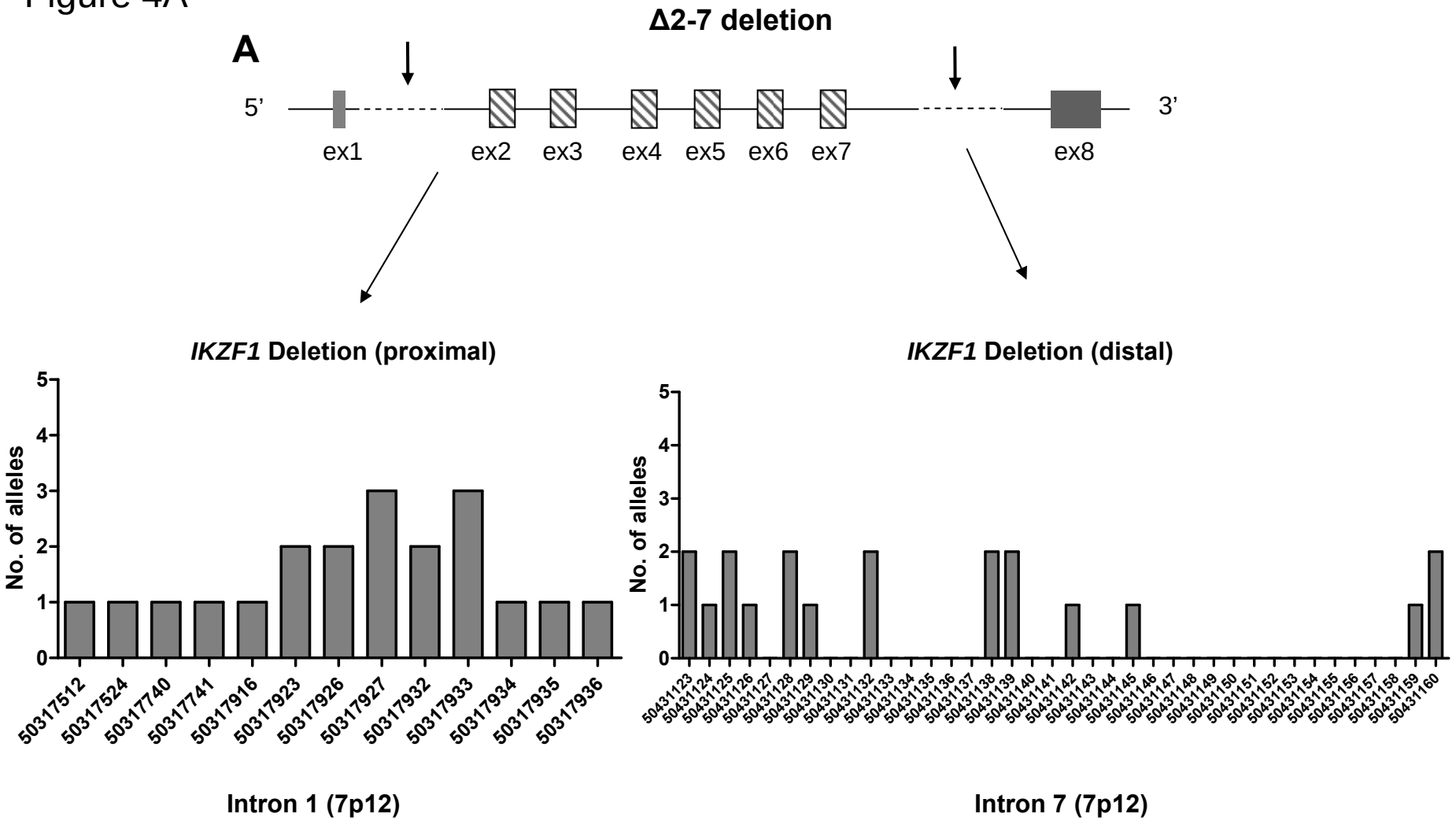
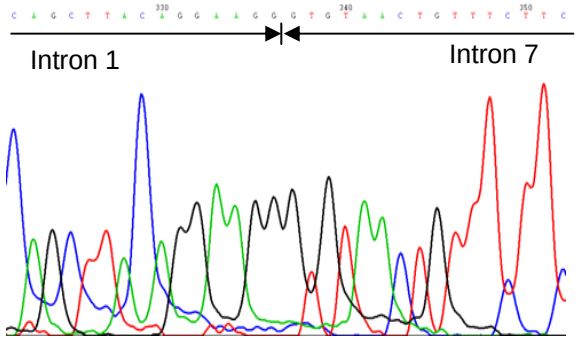

A
Δ2-7 deletion
5’
3’
ex1
ex2
ex3
ex4
ex5
ex6
ex7
ex8
IKZF1 Deletion (proximal)
IKZF1 Deletion (distal)
No. of alleles
0
1
2
3
4
5
50317512
50317524
50317740
50317741
50317916
50317923
50317926
50317927
50317932
50317933
50317934
50317935
50317936
Intron 1 (7p12)
Intron 7 (7p12)

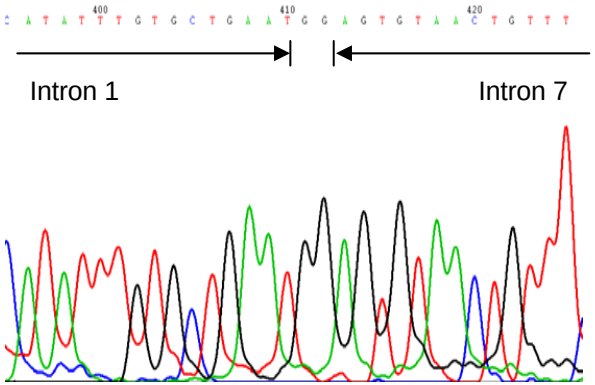
Figure 4B

B

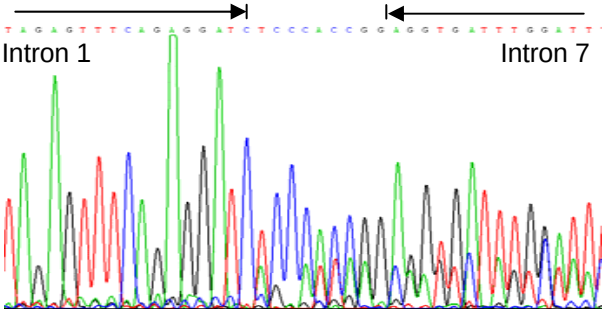
BCR-ABL1- #5



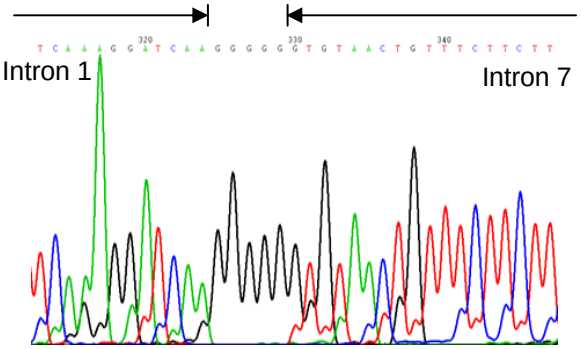
BCR-ABL1- #29



BCR-ABL1- #30



BCR-ABL1- #22
Allele 1



BCR-ABL1- #22
Allele 2