ALL1 gene alterations in acute leukemia: biological and clinical aspects

GIUSEPPE CIMINO, MARIA CRISTINA RAPANOTTI, TERESA SPROVIERI, LOREDANA ELIA
Dipartimento di Biotecnologie Cellulari ed Ematologia, Università degli Studi “La Sapienza” di Roma, Italy

Abstract

Background and Objective. The ALL1 gene, also referred to as MLL, HRX or Htrx1, is interrupted in the vast majority of translocations involving the chromosome band 11q23. Alterations in this gene are reported in approximately 5-10% of acute leukemias (AL) and characterize different leukemic subtypes such as infant (<12 months of age) AL, topoisomerase II inhibitors-related (TR) AL and a small subset of de novo AML and ALL. Distinguishing features of ALL1 alterations include the striking heterogeneity of its recombinations, i.e. more than 30 chromosomes partners have been described in ALL1 re-arrangements, and the lack of association with a definite lineage. The objective of this article is to review the biological and structural properties of ALL1 gene and its various fusion proteins, and to discuss the clinical relevance of these lesions with special emphasis on their role in molecular diagnosis and monitoring of minimal residual disease.

Evidence and Information Sources. The material examined in the present review includes data published by the authors in this field, articles and abstracts published in journals covered by the Science Citation Index® and Medline®, as well as some more recent personal unpublished observations.

State of the Art. The ALL1 gene spans approximately 90 kb of DNA in length, and consists of 36 exons, ranging in size from 65 bp to 4249 bp. ALL1 codifies for a major transcript of ~15 kb. It encodes a protein of more than 3910 amino acids, containing three regions sharing sequence homology with the Drosophila trithorax gene. These homologies suggest that ALL1 is a transcription factor controlling development and/or differentiation of human cells. To date, twelve ALL1 partner genes have been characterized which are involved in the following translocations: t(4;11), t(9;11), t(6;11), t(11;19), t(1;11) t(10;11), t(11;16), t(11;17) and t(X;11). Since all these genes do not share relevant homologies among each other, their putative role in ALL1 activation still remains to be clarified. The analysis of ALL1 breakpoint cluster region (bcr) shows that several DNA motifs implicated in illegitimate recombination events are located within the bcr. Thus, mapping of breakpoints in the different subtypes of ALL1-ve leukemia may help in understanding the events leading to translocations in human ALs. In this respect, data on ALL1 breakpoint localization suggest that similar pathogenetic mechanisms may underlie infant and TR AL and that these events might differ from those occurring in de novo AL. The availability of this molecular marker provides a new tool for diagnostic purposes and characterization of ALs and for monitoring of minimal residual disease. To date, the prognostic value of ALL1 rearrangements has been clearly demonstrated for infant ALs only, whereas the clinical relevance of ALL1 rearrangements in the other leukemic subtypes needs further evaluation by future prospective studies on a larger number of patients homogeneously treated. As concerning studies on minimal residual disease, data on PCR monitoring of the ALL1/AF4 fusion transcript, resulting from the t(4;11) translocation, show the clinical relevance of this molecular test in predicting outcome and, as a consequence, in designing individual post-remission therapies.

Perspectives. It is expected that future studies will provide more detailed information regarding either the normal ALL1 function and/or the leukemogenic effect of ALL1 alterations, together with a better definition of the prognostic relevance of the hybrid proteins formed by this gene at diagnosis and during remission of disease.

©1998, Ferrata Storti Foundation

Key words: acute leukemia, oncogene, ALL1, molecular diagnosis, minimal residual disease

Acute leukemia (AL), like other human malignancies, is a progressive clonal disease driven by somatically acquired genetic changes. However, unlike most solid tumors, AL is more frequently characterized by specific genetic alterations such as translocations and inversions, which lead to the activation of proto-oncogene products usually through the creation of tumor specific fusion proteins.\(^1,^5\)

Increasing biological evidence indicates that different genetic lesions underlie the phenotypic variability observed in AL subtypes. For example, gene rearrangements involving the immunoglobulin superfamily genes usually occur in lymphoid tumors. More specifically, abnormalities involving the T-cell receptor chains are almost exclusively found in T-lymphoid malignancies.2,4 As for hybrid fusion genes, relevant examples include PML/RAR\(_\alpha\), AML1/ETO, CBFB/MYH11, and E2A/PBX, which are associated with...
FAB M3, M2, and M4 eos acute myeloid leukemia (AML) and with pre-B subtype of acute lymphoblastic leukemia (ALL), respectively.6-10 On the other hand, in apparent contrast to this paradigm, distinct rearrangements of the chromosomal band 11q23 are observed in a variety of leukemia and lymphoma subtypes.

Recently, we and others have cloned and characterized the ALL1 gene (also referred to as MLL, HRX, Htrx1), which is interrupted in the vast majority of translocations involving the 11q23 region.11-15 Alterations in this gene are detected in approximately 5-10% of AL cases, and characterize different leukemic subtypes including infant and topoisomerase II inhibitor-related (TR) ALs, and a small subset of de novo AML and ALL.16,17 Distinguishing features of ALL1 alterations include the striking heterogeneity of its recombinations, i.e. more than 30 chromosome partners have been described in ALL1 rearrangements, and the lack of association of these lesions with a definite lineage.18,19

In this article, we review the biological and structural properties of the ALL1 gene and its various fusing proteins. In addition, we discuss the clinical relevance of these lesions with special emphasis on their role in molecular diagnosis and monitoring of minimal residual disease (MRD).

**Cloning and structure of the ALL1 gene**

Using somatic cell-hybrids or fluorescent in situ hybridization (FISH), the chromosome 11q23 breakpoints were mapped into a region between the CD3γ/H9253 and porphobilinogen deaminase genes. Subsequently, a yeast artificial chromosome (YAC) containing the CD3γ/H9254 and CD3γ/H9253 genes was cloned and shown by FISH analysis to span the t(4;11), t(6;11), t(9;11) and t(11;19) chromosome translocation breakpoints (13). From a similar YAC, we obtained a DNA insert, which by Southern blot analysis detected rearranged bands in leukemic cells from patients with the t(1;11), t(4;11), t(6;11), t(9;11), t(10;11), or del(11q23), and we showed that breakpoints clustered in a small region of ≈ 8 kb named by us ALL1.11 This gene spans approximately 90 kb of DNA in length and consists of 36 exons, ranging in size from 65 bp to 4249 bp.20,21 ALL1 codifies for a major transcript of ≈ 15 kb. This latter encodes a protein of more than 3910 amino acids containing three regions homologous to sequences of the Drosophila trithorax gene, including cysteine-rich regions that can fold into six zinc finger-like domains. The trithorax gene in Drosophila acts to spatially maintain restricted expression patterns of Antennapoedia and Bithorax complexes during fruit fly development. Trithorax activates transcription of multiple genes of the two complexes and as such, counteracts the activity of polycomb group genes, which repress the transcription of the same genes.22-24 Based on what is known about the Drosophila homologous gene, very likely, ALL1 represents a transcription factor involved in the regulation of genes controlling human development and/or differentiation. In this respect, it is noteworthy that, at first, ALL1 sequences were isolated from a human fetal cDNA library, suggesting that this gene is highly expressed during fetal development.12 Furthermore, the ALL1 gene product possesses two other regions which would be directly or indirectly involved in the control of gene transcription. These are: 1) a region similar to the AT hook of highly-mobility-group-I, that binds to AT-rich regions of the minor groove of the DNA; and 2) a cysteine-rich region (CRR) homologous to the mammalian DNA methyltransferase double helix, which, favoring conformational DNA changes, facilitates the action of other regulatory genes14 (Figure 1).

**ALL1-partner genes in 11q23 translocations**

To date, twelve ALL1 partner genes have been characterized which are involved in the following translocations: t(4;11), t(9;11), t(6;11), t(11;19), t(1;11), t(10;11), t(11;16), t(11;17) and t(X;11).28-32,65

The impressive heterogeneity of ALL1 recombinations raises two important questions. Firstly, which of the two hybrid genes originated by translocations have more relevant oncogenic properties? Secondly,
do the partner genes play an active role either in determining the leukemic phenotype or in activating the ALL1 gene?

As concerns the former issue, several lines of evidence suggest that the chimeric gene is transcribed from the derivative 11 containing at the 5' side the ALL1 portion and at the 3' side the partner gene sequences. In fact, in cases with complex translocations involving the 11q23 cytogenetic band, the junction leading to the fusion of the ALL1 exons encoding the NH2 portion of the ALL1 is formed, whereas the distal part of 11q is translocated onto another chromosome.33-35 Similarly, the 5' portion of the reciprocal gene can be juxtaposed to a gene other than ALL1, as recently reported for the AF10 at 10p12 band in a AL case with an invins(10;11) (pl2;q23,q12) in which the 5' portion of the AF10 was fused to a gene named HEAB at the 11q12.36 Furthermore, the ALL1 portion telomeric to the breakpoint site is deleted in about 20-30% of cases with 11q23 abnormalities.37,38,39,40 Finally, in ALL cases with the t(4;11) the ALL1/AF4 transcript is present whereas the reciprocal AF4/ALL1 chimeric mRNA is inconsistently detected.38

To date, no significant associations have been found between the leukemic phenotype and the most common ALL1 chimeric genes except for the ALL1/CBP, recently identified in the t(11;16)(q23;p13), which is strictly linked with therapy-related myelodysplastic syndrome.33 To answer the question whether genes fused to ALL1 provide functionally common domains, or simply supply initiation or termination signals for translation of the disrupted ALL1 coding region, several investigators have searched structural similarities among the different ALL1 partner genes. As shown in Table 1, with the exception of AF9 with ENL, of AF10 with AF17, and of AFX with AF6q21, sequence analysis did not reveal structural or functional similarities among the different ALL1 partner genes. Thus, it is unlikely that these genes could play a role in the function of the chimeric ALL1 partner products by simply providing transcriptional modulation (activation or repression) domains. By contrast, it has been suggested that they might supply dimerization domains, which could activate the ALL1 chimeric genes. This hypothesis is supported by the following observations: 1) epsil5 and AF6 are cytoplasmic proteins showing structural similarities with the rod-like region of various myosin chains, which in turn are involved in protein-protein interactions; 2) AF17 and AF10 contain the dimerizing leucine zipper motifs, commonly regarded as structures controlling protein-protein interactions; 3) a newly described genetic mechanism, named self-fusion, leads to the tandem duplication of an internal fragment of the ALL1 gene originating as an abnormal protein, which could be functionally equivalent to a dimer of the NH2 portion of the ALL1.39

To date, the physiological functions of both the wild-type ALL1 and its chimeric products are still unknown. To study the role of this gene in hematopoiesis, Hess et al. have recently examined the effects of the haplo-insufficiency or absence of ALL1 gene on the in vitro differentiation of yolk salk progenitor cells, showing that ALL1 is required for generation of normal numbers of hematopoietic progenitors and their proper differentiation, especially along the granulocyte and macrophage lineages.40 Furthermore, anti-ALL1 polyclonal antibodies showed that proteins translated by ALL1 chimeric genes are localized exclusively to the nucleus with bodies smaller and more numerous than those usually detected in normal

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Location</th>
<th>Transcript (kb)</th>
<th>Protein (aa)</th>
<th>Putative function</th>
<th>Shared homologies</th>
<th>References (No.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF4 or FEL</td>
<td>4q21</td>
<td>12 &amp; 15</td>
<td>1200</td>
<td>Transcription factor</td>
<td>No</td>
<td>(25,26)</td>
</tr>
<tr>
<td>AF9</td>
<td>9p22</td>
<td>5</td>
<td>578</td>
<td>Transcription factor</td>
<td>ENL</td>
<td>(25)</td>
</tr>
<tr>
<td>ENL</td>
<td>19p13</td>
<td>4.7</td>
<td>559</td>
<td>Transcription factor</td>
<td>AF9</td>
<td>(14,61)</td>
</tr>
<tr>
<td>AF6</td>
<td>6q27</td>
<td>8</td>
<td>1612</td>
<td>Cytoplasmic prot.</td>
<td>No</td>
<td>(27)</td>
</tr>
<tr>
<td>AF1p</td>
<td>1p32</td>
<td>4.4</td>
<td>896</td>
<td>Cytoplasmic prot.</td>
<td>No</td>
<td>(32)</td>
</tr>
<tr>
<td>AF10</td>
<td>10p12</td>
<td>5.5</td>
<td>1027</td>
<td>Cytoplasmic prot.</td>
<td>AF17</td>
<td>(34)</td>
</tr>
<tr>
<td>AF17</td>
<td>17q21</td>
<td>7.5 &amp; 5</td>
<td>1093</td>
<td>Dimerization prot.</td>
<td>AF10</td>
<td>(29)</td>
</tr>
<tr>
<td>ELL</td>
<td>19p13.1</td>
<td>4.4</td>
<td>621</td>
<td>Transcription factor</td>
<td>No</td>
<td>(28)</td>
</tr>
<tr>
<td>AF1q</td>
<td>1q21</td>
<td>1.7</td>
<td>90</td>
<td>Growth factor</td>
<td>No</td>
<td>(31)</td>
</tr>
<tr>
<td>AFX1</td>
<td>Xq13</td>
<td>Not determined</td>
<td>Not determined</td>
<td>Transcription factor</td>
<td>AF6q21</td>
<td>(30)</td>
</tr>
<tr>
<td>CBP</td>
<td>16p13</td>
<td>9</td>
<td>2491</td>
<td>Transcriptional coactivator</td>
<td>No</td>
<td>(33)</td>
</tr>
<tr>
<td>AF6q21</td>
<td>6q21</td>
<td>6</td>
<td>403</td>
<td>Transcription factor</td>
<td>AFX</td>
<td>(65)</td>
</tr>
</tbody>
</table>
cells, while the wild type ALL1 protein is localized to both the cytoplasm and nucleus. This observation suggests that ALL1 chimeric proteins are activated by deocalization within the cell.41

The leukemogenic activity of the ALL1 fusion gene has been recently demonstrated in a study that showed the development of AML in a mouse carrying the ALL1/AF9 abnormality.42

**Putative mechanisms for origin of ALL1 recombinations**

A critical unsolved question in AL pathogenesis concerns mechanisms undergoing illegitimate recombination events leading to translocations. In this respect it is interesting to note that several DNA motifs implicated in DNA recombination machinery have recently been identified within the ALL1 bcr. These include: 1) recombinase signal sequences (heptamers and nonamers); 2) scaffold attachment regions (SARs); 3) high affinity topo-II consensus binding sites, including a strong site in exon 9, and 4) Alu-sequences.43-45 By comparing ALL1 rearrangements in de novo versus TR AL, Strissel-Broeker et al. have reported statistically significant differences in breakpoint distribution between the two groups. In particular, they found that in TR AL, breakpoints clustered in the telomeric portion of the ALL1 bcr, characterized by the presence of SARs and high affinity topo II binding sites. By contrast, de novo AL most frequently showed breakpoints in the centromeric or 5' bcr portion, which is high prone to recombinogenic events for the presence of a high number of Alu sequences. Based on these observations, the authors suggest that translocations mechanisms in de novo and TR AL might be different.43 This conclusion also has important implications to explain the etiology and pathogenesis of ALL1+ve AL occurring during early infancy. Since molecular analyses showed that in these ALs ALL1 alterations arise during pregnancies,46 it has been suggested that critical event(s) in utero might involve exposure to topo II inhibitors, and several such natural and medicinal substances are available as candidate agents.48-50 To test this hypothesis, we recently compared ALL1 gene breakpoint distribution in a large group of infant leukemias (26 pts) with those of de novo childhood or adult AL (29 pts) and TR AL (5 pts). We found that, compared to de novo AL, infant ALs showed the same biased distribution of ALL1 gene breaks as topoisiomerase II-related secondary AL.51 These data lend further credence to the hypothesis that exposure to topo II inhibiting chemicals in pregnancy may be causally related to infant AL pathogenesis.

**ALL1 gene alterations in acute leukemias**

As we have seen, it is generally assumed that different genetic lesions underlie the phenotypic variability observed in acute leukemia subtypes. By contrast, ALs with ALL1 abnormalities are not associated with a definite lineage. A paradigmatic example of the lineage promiscuity observed in ALL1 rearranged malignancies is represented by the subset of infant AL. This is a distinct AL entity with consistent biological and clinical features and with an extremely poor response to treatment and survival. Furthermore, 60-70% of these patients present an alteration of ALL1 gene. Leukemic cells frequently reveal myelomonocytic or monocytic (FAB M4-M5) features within myeloid types, and CD19+/CD10- markers within lymphoid cases. In addition, hybrid B lymphoid/myelomonocytic phenotypes are frequently observed either simultaneously or as a consequence of lineage switch during disease evolution.32-36

Similar findings can be observed in the small subset of childhood and adults with de novo ALL1+ve ALs. In fact, in this last group of patients, ALL1 gene is altered in 50% of pre-pre B ALL,37 in 25-30% of AML classified as FAB M4/M5,38,39 and in 10% of M0-M1 AML.40 Furthermore, as observed in infant ALs, myeloid and lymphoid markers may be present on leukemic blasts either simultaneously or as a consequence of lineage switch at relapse. All together, these observations suggest that ALL1 rearranged leukemias originate from an early hemopoietic precursor with bipotential lymphoid-myelocytic differentiation capability. This hypothesis is furtherly supported by in vitro evidence that lymphoid cell lines bearing 11q23 abnormalities may acquire monocytic features after stimulation in culture.41 Finally, the demonstration of identical ALL1 rearranged bands on DNA samples from lymphoid and myelocytic leukemic blasts in patients who presented a lineage switch at relapse, provides compelling evidence that the phenotypically different leukemic cells arise from the same hemopoietic precursor (Figure 2).

Figure 2. ALL1 genic configuration in a 65-year-old female patient with ALL who presented at relapse a lineage switch to monoblastic features. DNA samples, digested to completion with Bgl II (lanes 1 and 3) and Bam HI (lanes 2 and 4), showed identical ALL1 rearranged bands (arrows) at diagnosis (lanes 1 and 2) and at relapse (lanes 3 and 4).
Diagnostic and prognostic relevance of ALL1 alterations in acute leukemias

The availability of this marker provides a new molecular tool for diagnostic and characterization of ALs. Because of the great number of partner genes and of diverse genetic mechanisms leading to ALL1 interruptions, Southern blot analysis is today's gold standard in studying ALL1 gene abnormalities. In fact, FISH analysis is uninformative in cases with ALL1 self-fusion or in cases with interstitial genetic insertion into the ALL1, as recently described for AF6 and AF10. However, cytogenetic and/or FISH are crucial techniques for the identification of ALL1 partner genes.

ALL1 breakpoints cluster in a small region of 8.3 kb encompassed by two Bam H1 sites, explored by the B859 cDNA insert, which includes ALL1 exons 5-11 sequences (Figure 3).12 Recently, reviewing Southern blot data achieved in 70 AL patients with ALL1 alterations, we found abnormal restriction fragments in 98%, 85%, 79% and 40% of DNA samples digested with Bam H1, Hind III, Bgl II and Xba I endonucleases, respectively. In this series, two DNA samples showed ALL1 rearrangements when digested with Bgl II endonuclease only. Thus, we recommend that, for routine analysis of ALL1 rearrangements, DNA samples should be digested with Bam H1 and Bgl II restriction enzymes and hybridized with the B859 cDNA probe.

To date, the prognostic value of ALL1 rearrangements in ALs has been clearly demonstrated in infant ALs only. Following preliminary observations in a limited series of cases in which an altered ALL1 gene configuration was associated to unfavorable clinical characteristics such as age < 6 months, hyperleukocytosis, organomegaly, hybrid phenotype, etc.,52 several studies on large series of infant patients have demonstrated the prognostic value of ALL1 configuration.53-56 In particular, in a group of 45 infant AL aged between 0 and 18 months, we reported an actuarial event free survival of 57% and 9% for those patients with germline or rearranged ALL1 configuration, respectively (p < 0.001). In this study, a multivariate analysis with a model including age, sex, WBC count, FAB classification and ALL1 status confirmed the independent value of this molecular alteration as a prognostic factor. These results on 11q23/ALL1 rearrangements were furthermore confirmed by Pui et al. in a series of 30 infants with ALL.55

In our opinion, two main considerations can be drawn from these results: 1) they are one of the first examples about the possibility to molecularly stratify ALs for prognosis; 2) they distinguish different risk categories within the apparently homogenous group of infant AL, a disease subset overall considered as bearing unfavorable outcome. Based on these observations, it is advisable that ALL1 gene status be taken into account to choose the most adequate, risk-adapted therapeutic strategy in this subset. Thus, extremely aggressive therapy should be considered for infants with ALL1 rearrangements, while the risk of very aggressive protocols, frequently associated with life-threatening and/or permanent complications, could be spared for those infants with a potentially curable disease for having a germline ALL1 configuration.

Concerning the group of childhood and adult de novo AL1+ve AL, two retrospective studies on AML failed to demonstrate significant prognostic differences between patients with or without ALL1 alterations.58-59 However, the above discussed data on infant AL and cytogenetic studies on adult patients with AL strongly indicate that ALL1 prognostic value should be examined in prospective studies involving large numbers of patients homogeneously treated.

Molecular monitoring of minimal residual disease

The cloning of several ALL1 partner genes has allowed the development of reverse transcriptase polymerase chain reaction (RT-PCR) strategies to specifically amplify the different fusion products for rapid diagnosis of translocations and sensitive monitoring of MRD.62-63 With respect to t(4;11) ALL, the ALL1/AF4 was amplified in diagnostic samples of all patients carrying a cytogenetically detectable t(4;11), whereas patients in long-term complete remission (CR) were...
found PCR negative. In a recent study, we used a RT-PCR based strategy to evaluate the presence of MRD in 12 patients with ALL1/AF4 positive ALL (7 infants and 5 adults). Eleven patients were treated with high-dose intensive induction and consolidation chemotherapy, without bone marrow transplantation, while the remaining case received a less intensive protocol because of her poor performance status. The sequential monitoring of the ALL1/AF4 hybrid transcripts showed that the 5 long-term survivors in continuous complete remission for a median time of 40 months (range 28-60 months) tested persistently PCR negative. By contrast, ALL1/AF4 products were persistently found in the remaining seven patients, including the four cases in complete hematological remission who later relapsed. These data were subsequently confirmed by the molecular monitoring of 3 additional cases with t(4;11) ALL (1 infant and 2 adults). A persisting molecular and hematological CR was observed in the infant case at 7+ months after induction of treatment, while the remaining two cases tested persistently PCR negative. These two patients relapsed within 12 months from hematological CR, and furthermore consolidated in one patient with an allogeneic BMT. Figure 4 graphically depicts PCR monitoring of MRD in the entire group of 15 patients.

These data provide convincing evidence that PCR positive tests during clinical CR are predictive of disease evolution and anticipate the occurrence of hematological relapse. The clinical value of PCR monitoring is also strengthened by results observed in long-term survivors that tested persistently PCR negative and remained in CCR. These observations also suggest that ALL1/AF4 positive ALL, at least in some instances, is a potentially curable leukemia, and that a repeatedly PCR negative status could be considered the optimal therapeutical goal, while intensive consolidation programs including BMT should be considered for those patients who did not achieve a molecular remission after initial treatment. Should these findings be confirmed in larger prospective studies, the PCR methods would be of utmost importance towards designing individual post-remission consolidation therapies.

**Future perspectives**

Studies on the ALL1 gene have provided important new information on the molecular mechanisms involved in leukemogenesis. Moreover, novel tumor-specific markers are now available for identifying previously unknown genetic subsets within the heterogeneous spectrum of AL. This should allow a better definition of prognostic AL groups and monitoring of MRD with the aim to define tailored induction and consolidation therapeutical programs.

**Contributions and Acknowledgments**

GC was responsible for the conception of this review article, the critical analysis of published data, and the writing of the paper. MCR, TS and LE contributed to the analysis of the literature and writing of the paper.
Funding
This work was supported in part by grants from ROMAIL (Associazione Italiana contro le Leucemie, Sezione di Roma).

Disclosures
Conflict of interest: none.
Redundant publications: no substantial overlapping with previous papers.

Manuscript processing
Manuscript received October 15, 1997; accepted January 13, 1998.

References
32. Bernard OA, Mauchauffe M, Mecucci C, Van Der Berghe H, Berger R. A novel gene, AF1q, fused to HRX in t(1;11) (p32;q23), is not related to AF1, AF9 nor ENL. Oncogene 1994; 9:1039-45.
34. Chaplin T, Ayton P, Olivier AB, et al. A novel class of zinc finger/leucine zipper genes identified from the
50. Buckley JD, Robinson LL, Swotosky R, et al. Occupa-