Molecular genetics of B-precursor acute lymphoblastic leukemia

Charles G. Mullighan

Department of Pathology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA.

B-precursor acute lymphoblastic leukemia (B-ALL) is the most common childhood tumor and the leading cause of cancer-related death in children and young adults. The majority of B-ALL cases are aneuploid or harbor recurring structural chromosomal rearrangements that are important initiating events in leukemogenesis but are insufficient to explain the biology and heterogeneity of disease. Recent studies have used microarrays and sequencing to comprehensively identify all somatic genetic alterations in acute lymphoblastic leukemia (ALL). These studies have identified cryptic or submicroscopic genetic alterations that define new ALL subtypes, cooperate with known chromosomal rearrangements, and influence prognosis. This Review describes these advances, discusses results from ongoing second-generation sequencing studies of ALL, and highlights challenges and opportunities for future genetic profiling approaches.

Introduction

Acute lymphoblastic leukemia (ALL) is a neoplasm of immature lymphoid progenitors that is most commonly of B cell lineage. Most childhood cases of B-precursor ALL (B-ALL) may be subclassified by the presence of either gross or submicroscopic genetic alteration. Approximately 75% of B-ALL cases exhibit aneuploidy or a recurring gross chromosomal rearrangement (1, 2) (Table 1 and Figure 1). Hyperdiploidy, with gain of at least 5 chromosomes, is one of the most frequent alterations in childhood ALL and is associated with favorable outcome (3). However, the biologic basis of the acquisition of multiple whole chromosomal gains is poorly understood. Conversely, hypodiploidy, with fewer than 44 chromosomes, is associated with dismal prognosis (4). A number of recurring chromosomal rearrangements are common in B-ALL and are critical events in leukemogenesis. These rearrangements commonly perturb genes encoding regulators of hematopoiesis, tumor suppressors, oncogenes, or tyrosine kinases but commonly require additional genetic hits to establish the full leukemic phenotype.

In the last five years, genome-wide profiling using microarrays, candidate gene, and second-generation sequencing have provided a number of key insights into the genetic basis of ALL. These studies have identified new subtypes of ALL and have uncovered recurring submicroscopic genetic alterations in known ALL subtypes. These include loss-of-function mutations involving genes regulating lymphoid development that contribute to the arrest in maturation characteristic of B-ALL, mutations that inactivate tumor suppressor and cell cycle regulatory proteins, and mutations that drive cytokine receptor and/or kinase signaling. Thus, as in AML, concomitant lesions disrupting hematopoietic development and tumor suppression as well as driving signaling and proliferation are hallmarks of many ALL subtypes. Importantly, several of these alterations are associated with specific subtypes of ALL defined by recurring chromosomal alterations.

This Review describes the genetic landscape of ALL, focusing on important subtypes and recently discovered genetic alterations of biologic and therapeutic relevance.

Conflict of interest: The author has declared that no conflict of interest exists.

Citation for this article: J Clin Invest. 2012;122(10):3407–3415. doi:10.1172/JCI61203.

ETV6-RUNX1 (TEL-AML1) B-ALL

The most common rearrangement in B-ALL is the t(12;21) (p13;q22) rearrangement that encodes ETV6-RUNX1 (5). This rearrangement is usually cryptic on cytogenetic analysis but is readily detected by fluorescent in situ hybridization and molecular techniques. ETV6 is a member of the ETS family of transcription factors that are frequently targeted by rearrangements and mutations in leukemia and other malignancies (6). With CBFB, RUNX1 forms the core binding factor transcription complex and is commonly rearranged in acute myeloid leukemia (7), and harbors sequence mutations in myeloid and lymphoid disorders (8–10). Both ETV6 and RUNX1 are required for normal definitive hematopoesis (7, 11), and the ETV6-RUNX1 protein may perturb expression of RUNX1-regulated genes, converting RUNX1 to a transcriptional repressor (12). ETV6-RUNX1 also causes overexpression of the erythropoietin receptor (EPOR) and activation of downstream JAK-STAT signaling (13). Expression of ETV6-RUNX1 promotes self-renewal in B cell progenitors but alone does not induce leukemia (14, 15). Furthermore, ETV6-RUNX1 is commonly detectable at birth (16), years prior to the onset of leukemia, suggesting that secondary genetic events are required to induce leukemia. This model is supported by genome-wide profiling studies of genetic alterations in ALL that have identified additional recurring submicroscopic genetic alterations in ETV6-RUNX1 ALL, including deletions of the B cell transcription factors PAX5 and EBF1 and deletion of the second copy of ETV6 (17–19). Moreover, monozygotic twins concordant for ETV6-RUNX1 ALL exhibit distinct submicroscopic DNA copy number alterations (CNAs), indicating that acquisition of the ETV6-RUNX1 is an early event in leukemogenesis (20). Submicroscopic genetic alterations are more common in ETV6-RUNX1 ALL than in many other ALL subtypes and include deletions targeting putative lymphoid signaling molecules (BTLA, TOX), transcriptional coactivators (TBLIX1), the glucocorticoid receptor gene NR3C1, and the putative apoptosis regulatory gene BTG1. Although these alterations may directly contribute to leukemogenesis, the genomic breakpoints of many of these deletions bear the hallmark of aberrant activity of the recombination activation genes. Thus, further work directly examining these alterations in leukemogenesis is required.
review series

Table 1

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Frequency (%)</th>
<th>Comment</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperdiploidy (≥50 chromosomes)</td>
<td>20–30</td>
<td>Excellent prognosis with antimitabolite-based therapy</td>
<td>3</td>
</tr>
<tr>
<td>Hypodiploidy (&lt;44 chromosomes)</td>
<td>1–2</td>
<td>Poor prognosis, high frequency of RAS pathway and IKAROS gene family mutations</td>
<td>4</td>
</tr>
<tr>
<td>t(12;21)(p13;q22) ETV6-RUNX1</td>
<td>15–25</td>
<td>Expression of myeloid antigens; excellent outcome</td>
<td>5</td>
</tr>
<tr>
<td>t(1;19)(q23;p13) TCF3-PBX1</td>
<td>2–6</td>
<td>Increased incidence in African Americans; generally excellent prognosis; association with CNS relapse</td>
<td>121–123</td>
</tr>
<tr>
<td>t(9;22)(q34;q11.2) BCR-ABL1</td>
<td>2–4</td>
<td>Historically dismal outcome, improved with addition of imatinib to intensive chemotherapy</td>
<td>39, 42</td>
</tr>
<tr>
<td>PAX5 rearrangement</td>
<td>–2%</td>
<td>Multiple partners, commonly from dic(7,9), dic(9;12) and dic(9;20); outcome unknown</td>
<td>17, 94</td>
</tr>
<tr>
<td>ABL1, PDGFRB, JAK2 rearrangements</td>
<td>2–5</td>
<td>Identified in BCR-ABL1–like ALL; multiple rearrangements encoding chimeric proteins, fusing 5' partners with 3' kinase domains; associated with IKZF1 alteration and very high leukocyte count; potentially amenable to tyrosine kinase inhibitor therapy</td>
<td>84</td>
</tr>
<tr>
<td>t(4;11)(q21;q23) MLL-AF4</td>
<td>1–2</td>
<td>Common in infant ALL (especially &lt;6 months of age); poor prognosis</td>
<td>45</td>
</tr>
<tr>
<td>MYC rearrangement t[(8;14)(q24;q32), t(2;8)(q12;q24)]</td>
<td>2</td>
<td>Favorable prognosis with short-term, high-dose chemotherapy</td>
<td>70</td>
</tr>
<tr>
<td>CRLF2 rearrangement (IGH@-CRLF2, PAR1 deletion, and P2RY8-CRLF2)</td>
<td>5–7</td>
<td>Extremely common in DS ALL (55%); association with IKZF1 deletion/mutation and JAK1/2 mutation; poor prognosis in non-DS ALL</td>
<td>71–75</td>
</tr>
<tr>
<td>ERG deletion</td>
<td>7</td>
<td>Subtype of B-ALL with a distinct gene expression profile; favorable outcome</td>
<td>124</td>
</tr>
</tbody>
</table>

TCF3-PBX1 (E2A-PBX1) and TCF3-HLF (E2A-HLF) B-ALL

The t(1;19)(q23;p13) translocation is present in up to 6% of childhood B-ALL cases (21). The rearrangement is commonly unbalanced, with duplication of 1q distal to PBX1.

This translocation fuses the transactivation domains of TCF3 with the C-terminal region of the homeobox gene PBX1. TCF3 encodes the basic helix-loop-helix (bHLH) transcription factors E12 and E47 that are required for early lymphoid development, and loss of E12/E47 promotes the development of T cell lymphoma (22, 23). PBX1, which encodes a member of the three-amino acid loop extension family of homeodomain proteins, is required for the development of lymphoid precursors (24). TCF3-PBX1 binds HOX proteins and likely interferes with hematopoietic differentiation by disrupting HOX-regulated gene expression (25).

The t(17;19)(q22;p13) translocation is a rare rearrangement that fuses the aminoterminial transactivation domains of TCF3 to the C-terminal DNA-binding and dimerization domains of HLF, a member of the PAR family of basic leucine zipper transcription factors (26). TCF3-HLF aberrantly regulates genes that control cell death in lymphoid progenitors, including LMO2 and BCL2 (27). TCF3-HLF–positive leukemia is associated with older age and very poor outcome (28).

BCR–ABL1 ALL

Breakpoint cluster region–ABL1 (BCR-ABL1) is generated by the der(22) of the t(9;22)(q34;q11) translocation, or Philadelphia (Ph) chromosome, which is present in over 95% of chronic myelogenous leukemia (CML) cases, 25% of adult ALL cases, and 3% to 5% of pediatric ALL cases (29). The t(9;22) fuses 5' sequences of breakpoint cluster region (BCR) to 3' sequences from ABL1, which encodes a tyrosine kinase. The breakpoints on chromosome 9 are scattered over a nearly 200-kb region within the first intron of ABL1, whereas the BCR breakpoints on chromosome 22 are clustered in two areas: a 5.8-kb major BCR (M-bcr) in CML and a minor BCR (m-bcr) in most cases of childhood Ph-positive ALL (30). Fusion genes created by breaks in M-bcr (CML-type break) encode a 210-kDa fusion protein (p210), whereas fusions that occur in m-bcr (ALL-type break) encode p190.

Both forms of BCR-ABL1 can transform hematopoietic cells in vitro and induce a syndrome similar to CML in mice (31). BCR-ABL1 activates multiple signaling pathways, increases cell proliferation, and deregulates differentiation and adhesion. Treatment of BCR-ABL1–positive leukemic cells with tyrosine kinase inhibitors results in the activation of the transcription factor BCL6, which may directly influence responsiveness to treatment with these agents (32).

Additional genetic alterations are critical determinants of the lineage and progression of BCR-ABL1 leukemia. Deletion of the early lymphoid transcription factor gene IKZF1 (IKAROS) is common in BCR-ABL1 lymphoid leukemia (CML at progression to lymphoid blast crisis) but is rarely present in CML at chronic phase (33, 34). IKAROS is a zinc finger transcription factor required for the development of all lymphoid lineages (35, 36). The IKZF1 deletions result either in haploinsufficiency or expression of dominant-negative isoforms. These alterations cooperate with BCR-ABL1 in the induction of lymphoblastic leukemia and promote resistance to therapy in recent experimental models of BCR-ABL1 ALL (37).

In childhood ALL, the presence of BCR-ABL1 is associated with older age, higher leukocyte count, and more frequent CNS leukemia at diagnosis (38). Historically, BCR-ABL1 ALL has been associated with extremely poor outcome (29, 39). Tyrosine kinase inhibitors
review series

such as imatinib mesylate (Gleevec) have transformed the treatment and outcome of patients with BCR-ABL1–positive leukemia (40, 41). The addition of imatinib to intensive chemotherapy in childhood BCR-ABL1–positive ALL results in a 4-year event-free survival rate of 84%, more than double that of historical controls (42). A proportion of patients become resistant to these agents, typically through acquired mutations in the ABL1 kinase domain (43).

**MLL–rearranged B-ALL**

Mixed-lineage leukemia–rearranged (MLL-rearranged) leukemia is a unique entity notable for initiation in utero, for myeloid and lymphoid features, and for poor responsiveness to therapy (44, 45). Rearrangement of the MLL gene at 11q23 occurs in at least two-thirds of infants with ALL, 5% of AML, and 85% of secondary leukemias that occur in patients treated with topoisomerase II inhibitors (46, 47) but is less frequent in older individuals (48, 49). Over 80 partners of MLL rearrangement have been identified (50).

The most common MLL rearrangements are t(4;11)(q21;q23)/MLL-AFF1(AF4), found in approximately 50% of cases, followed by t(9;11)(p22;q23)/MLL-MLLT3(AF9), MLL-ENL, and t(10;11)(p13.15;q14-21)/MLL-MLLT10(AF10) (51, 52). The translocation breakpoints are located in an 8.5-kb cluster between exons 5 and 11 and juxtapose the A-T hook and methyltransferase domains to partner proteins. MLL regulates hematopoiesis through maintenance of normal homeotic gene expression (53), in part through transcriptional activation of HOX genes, mediated through the histone H3 lysine 4 (H3K4) methyltransferase activity of the SET domain (54). In contrast to other subtypes of B-ALL, additional genetic alterations are uncommon in MLL-rearranged leukemia (17, 55, 56).

**MLL-rearranged ALL** exhibits a distinct gene expression profile characterized by the upregulation of class I HOX genes (HOXA3, -A7, and -A9 and the HOX cofactor gene MEIS1) (57). Co-expression of HOXA genes and MEIS1 cooperates with MLL fusions in induction of leukemia and maintenance of a “stem cell–like” state of maturation (58). fms-related tyrosine kinase 3 (FLT3) is overexpressed in MLL-rearranged ALL and harbors point mutations in 10%–20% of cases (59). In animal models, FLT3 kinase inhibitors have shown potential as therapeutic agents for MLL-rearranged leukemias, and FLT3 inhibitors are currently being tested in clinical trials (59, 60). MLL-rearranged leukemias also exhibit a distinct epigenetic profile, with signatures of cytosine (61), microRNA (62) and H3K79 methylation (63) that differ from non-MLL leukemias.

Many of the MLL fusion proteins are located in protein complexes that regulate transcriptional elongation, which may be in part responsible for transcriptional dysregulation and leukemogenesis. These include polymerase-associated factor complex, which associates with the N terminus of MLL and regulates RNA polymerase II (Pol II), the pTEFb (CDK9/cyclin T) complex, which associates with the fusion partners MLLT1 (ENL), ELL, and AFF1 (AF4) to release stalled Pol II and stimulate transcriptional elongation, and the H3K79 histone methyltransferase DOT1L, which binds MLL fusion partners including AFF1, AF9, AF10, and ENL (64). H3K79 methylation is a mark of actively transcribed genes, and suppression or inactivation of DOT1L in human or murine leukemia cells leads to suppression of the MLL fusion protein–induced gene expression program (65), differentiation and/or apoptosis of leukemic cell lines, and suppression of leukemogenesis (65). DOT1L inhibition is thus being pursued as a therapeutic approach in MLL-rearranged ALL (66).

These findings highlight the potential importance of epigenetic dysregulation in other subtypes of ALL. Several studies profiling cytosine methylation have shown that different ALL subgroups exhibit distinct methylation profiles and that promoter methylation is associated with gene expression and outcome (67, 68). Additional studies are needed to comprehensively determine the relative importance of genetic and epigenetic alteration in the pathogenesis of ALL as well as the potential therapeutic role of epigenetic modifying drugs.

**Rearrangement of CRLF2 and other immunoglobulin heavy locus rearrangements in B-ALL**

Rearrangements of the immunoglobulin heavy locus (IGH@) at 14q32.33, resulting in juxtaposition of IGH@ enhancer elements with transcription factor and cytokine receptor genes, are observed in both B-ALL and T-lineage ALL (69). B-lineage leukemia/lymphoma with rearrangement of MYC is reviewed elsewhere (70).

Cytokine receptor-like factor 2 (CRLF2; also known as thymic stromal-derived lymphopoietin [TSLP] receptor) is located at the pseudoautosomal region (PAR1) at Xp22.3/Yp11.3 and is rearranged or mutated in up to 7% of B-ALL cases (71, 72). With the IL-7Rt chain, CRLF2 forms a heterodimeric receptor for TSLP. The rearrangements are either rearrangement into IGH@-CRLF2 or focal deletion immediately upstream of CRLF2, resulting in a chimeric fusion of the purinergic receptor gene P2RY8 to the entire coding region of CRLF2. Both rearrangements result in overexpression of CRLF2 on the cell surface (71). Less commonly,
Intrachromosomal amplification of chromosome 21
Intrachromosomal amplification of chromosome 21 (iAMP21) occurs in up to 2% of B-progenitor ALL cases (85). The region of amplification is typically large but variable, but includes gain of at least three copies of RUNX1, commonly with deletion of the subtelomeric regions of chromosome 21. iAMP21 is characterized by commonly complex patterns of chromosomal rearrangement suggestive of a breakage-fusion-bridge cycle (86). iAMP21 frequently occurs in older children or adolescents and has been reported to be associated with poor outcome (87), although this association may not be independent of other established prognostic variables such as levels of minimal residual disease (88).

BCR-ABL1–like ALL
Up to 15% of childhood B-ALL cases exhibit a gene expression profile similar to that of BCR-ABL1–positive ALL, often have deletion/mutation of IZKF1, and have a very poor outcome (81–83). CRLF2 is rearranged in up to 50% of BCR-ABL1–like ALL cases, but until recently the lesions responsible for kinase activation in the remaining cases were unknown. Second-generation mRNA sequencing and whole genome sequencing of 15 BCR-ABL1–like ALL cases identified rearrangements, sequence mutations, and DNA CNAs activating kinase signaling in all cases (84), including rearrangements of PDGFRB, ABL1, JAK2, and EPOR as well as deletion/mutation of SH2B3 and IL7R. Several of these alterations have been shown to result in activation of downstream (e.g., JAK/STAT) signaling pathways, and cellular transformation was attenuated with JAK or ABL1/PDGFRB inhibitors. Thus, defining kinase-activating alterations in this subtype of ALL is critical to identifying and treating patients with this subtype of leukemia.

Hypodiploid ALL
Hypodiploidy, with fewer than 44 chromosomes, is present in up to 3% of childhood ALL cases and is associated with a high risk of treatment failure (89). Hypodiploid ALL may be subclassified by severity of aneuploidy into near-haploid (NH; 24–31 chromosomes) and low-hypodiploid (LH; 32–44 chromosomes) cases, but until recently the concomitant submicroscopic genetic alterations were poorly characterized. NH ALL has a very high frequency of deletions and sequence mutations that activate RAS signaling, and NH and LH ALL have a high frequency of inactivating alterations of IKAROS genes IZKF2 (HELIOS) and IZKF3 (AIOLOS) that are otherwise rare in ALL (90). Moreover, LH cases have a very high frequency of loss-of-function mutations in TP53 (encoding p53) and RB1, which are otherwise uncommon at diagnosis in B-ALL.

Submicroscopic genetic alterations in B-ALL
The observation that up to 25% of children with ALL lack a recurring chromosomal alteration (2) and the identification of translocation-encoded fusions at birth or in years prior to the onset of leukemia (16) indicates that additional genetic alterations cooperate in leukemogenesis. The advent of array-comparative genomic hybridization and SNP microarrays has enabled interrogation of structural genetic alterations at very high resolution (91). These approaches, coupled with candidate gene sequencing and integrated gene expression profiling (92), have identified a number of critical new targets of mutation in B-ALL that often involve genes that regulate lymphoid development, cell cycle, tumor suppression, and a variety of other key cellular pathways (Table 2). The mean number of DNA CNAs is lower than in many solid tumors (an average of 6–8 lesions per case) (17, 19, 33), but over 50 recurring CNAs have been identified. The nature and frequency of these CNAs are significantly associated with ALL subtype, with few lesions in MLL-rearranged ALL (17, 55) but 6–8 lesions per case in BCR-ABL1 and ETV6-RUNX1 ALL (33, 34).

Most notably, genes encoding transcriptional regulators of B-lineage commitment and differentiation are mutated in over two-thirds of B-ALL. These include PAX5 (encoding paired box 5, mutated in over 30% of cases), EBFI (early B cell factor 1), IZKF1–3, and lymphoid enhancer factor 1. PAX5 harbors broad deletions predicted to result in haploinsufficiency, focal internal deletions that result in the expression of truncated PAX5 transcripts, sequence mutations that disrupt DNA binding or truncate the transactivating domain (17, 93), and translocations that create chimeric fusion proteins (17, 93, 94). These lesions usually only involve a single copy of PAX5, and expression of key PAX5 targets, such as CD19, in ALL cells is frequently normal, suggesting that PAX5 haploinsufficiency contributes to leukemogenesis.

Despite their high frequency of occurrence, PAX5 alterations are not associated with outcome in ALL (81, 95). In contrast, IZKF1 alterations are less frequent (15% of B-ALLs) but are associated with adverse outcome in both BCR-ABL1–positive and negative ALL cases. IZKF1 alterations are present in up to one-third of high-risk B-ALL cases and triple the risk of treatment failure (81, 96). IZKF1, which is required for lymphoid development, binds DNA and forms homodimers or heterodimers with other IKAROS family members. IZKF1 has multiple functions including transcriptional activation, repression, and chromatin remodeling, in part mediated via the nucleosome remodeling and deacetylase complex that also contains IKZF3. The alterations observed in ALL include broad and focal deletions of IZKF1, which result in loss of func-
### Table 2
Key novel genetic alterations in B-ALL

<table>
<thead>
<tr>
<th>Gene</th>
<th>Alteration</th>
<th>Frequency</th>
<th>Consequences of alteration</th>
<th>Clinical relevance</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PAX5</strong></td>
<td>Focal deletions, translocations, sequence mutations</td>
<td>31.7% of B-ALL cases</td>
<td>Transcription factor required for B-lymphoid development; mutations impair DNA binding and transcriptional activation</td>
<td>Over 80% of BCR-ABL1 ALL cases and 66% of CML cases in lymphoid blast crisis</td>
<td>17, 19, 33</td>
</tr>
<tr>
<td><strong>IKZF1</strong></td>
<td>Focal deletions or sequence mutations</td>
<td>15% of pediatric B-ALL cases</td>
<td>Transcription factor required for development of HSCs to lymphoid precursors; deletions and mutations result in loss of function or dominant-negative isoforms</td>
<td>Approximately 30% of high-risk BCR-ABL1-negative ALL cases</td>
<td>17</td>
</tr>
<tr>
<td><strong>JAK1/2</strong></td>
<td>Pseudokinase and kinase domain mutations</td>
<td>18%–35% of DS-ALL cases</td>
<td>Constitutive JAK/STAT activation; transforms hematopoietic cell lines to be growth factor independent</td>
<td>10.7% of high-risk BCR-ABL1 ALL cases</td>
<td>76, 77</td>
</tr>
<tr>
<td><strong>CRLF2</strong></td>
<td>IGH@-CRLF2 or P2RY8-CRLF2, resulting in CRLF2 overexpression</td>
<td>5%–16% of pediatric and adult B-ALL cases; &gt;50% of DS ALL</td>
<td>Associated with mutant JAK in up to 50% of cases; CRLF2 mutations and JAK mutations co-transform in Ba/F3 cells and result in constitutive JAK/STAT activation</td>
<td>14% of pediatric high-risk ALL cases</td>
<td>71–73, 75</td>
</tr>
<tr>
<td><strong>CREBBP</strong></td>
<td>Focal deletion and sequence mutations</td>
<td>19% of relapsed ALL cases, especially ALL with high hyperdiploidy; also mutated in non-Hodgkin lymphoma</td>
<td>Mutations result in impaired histone acetylation and transcriptional regulation</td>
<td>Mutations selected for at relapse and associated with glucocorticoid resistance</td>
<td>112, 114, 115</td>
</tr>
</tbody>
</table>
review series

Genetics of relapse in ALL

ALL genomes are not static, frequently acquiring secondary karyotypic and genetic alterations during disease progression (101–103). Genome-wide profiling has enabled key questions in the genetic basis of treatment failure to be investigated, including the nature of genetic alterations that are acquired with disease progression, the genetic relationship between the diagnosis and relapsed leukemic clones, and the degree of clonal heterogeneity present at the time of diagnosis.

Microarray profiling has shown that relapse samples exhibit striking differences in the patterns of genomic alterations from the matched diagnosis samples (104–108), with the majority of relapse samples exhibiting the loss of CNAs present at diagnosis and the acquisition of new genetic alterations. Importantly, most paired diagnosis-relapse samples share a common clonal origin. Moreover, lesions detected at relapse are often detectable in diagnostic samples, indicating that the predominant relapse clone(s) are present at low levels at diagnosis. Genetically unrelated second leukemias (at least at the level of DNA CNA analysis by SNP array) are infrequent but may be more common in late relapse (109). Approximately one-third of cases show a pattern of linear clonal evolution (with the acquisition of new genetic changes by the relapse clone in addition to those seen at relapse), and over 50% show a complex picture with both loss of diagnosis CNAs and the acquisition of new lesions, as well as evidence of low levels of relapse clones at diagnosis (Figure 2).

Together, these data suggest that a pre-diagnosis “ancestral” clone harboring one or more genetic alterations (such as a founding translocation) undergoes divergent evolution into multiple clones that acquire different genetic alterations and emerge as the predominant clones at diagnosis and relapse. This is supported by studies of monozygotic twins concordant for ALL, showing different secondary CNAs in each twin (20), and recent xenotransplantation studies that have modeled clonal evolution and heterogeneity by engrafting ETV6-RUNX1 and BCR-ABL1 samples into immunocompromised mice. These studies have confirmed clonal heterogeneity in the majority of cases at the time of diagnosis and have shown that specific genetic alterations (e.g., deletion of CDKN2A/CDKN2B) influence the efficiency and tempo of engraftment (110, 111).

Genome-wide sequencing is required to comprehensively identify all sequence and structural variants contributing to the pathogenesis of ALL, and a number of sequencing efforts are ongoing. However, a recent study performed Sanger sequencing of 300 genes in 23 matched diagnosis-relapse samples, with recurrence testing of over 300 paired and unpaired ALL samples (112). This study identified a number of uncharacterized mutations in ALL, most notably deleterious mutations in CREBBP (encoding CREBBinding protein) in over 20% of relapse ALL samples (112). The mutations were present at relapse and were detected in either the predominant clone or a subclone at diagnosis. CREBBP is a transcriptional coactivator, histone and non-histone acetylase, and ubiquitin ligase (113). Mutations were common in the histone acetyltransferase (HAT) domain and resulted in impaired acetylation of histone targets (112). CREBBP mutations are highly enriched...
in hyperdiploid ALL (114) and are common at diagnosis in B cell lymphoma samples (115). Importantly, the HAT mutations impair the normal CREBBP-mediated transcriptional response to glucocorticoids, which are widely used in ALL therapy. Thus, CREBBP mutations may impair the response of leukemic cells to glucocorticoids and influence treatment responsiveness, and therapeutic approaches modifying acetylation may reverse glucocorticoid resistance. Mutations in TTP3, which are otherwise infrequent in ALL, are also enriched in relapsed ALL (116).

**Inherited genetic variation and risk of ALL**

Several candidate gene studies have implicated inherited genetic variation, for example in the folate metabolic pathway, in the risk of developing leukemia (117). Multiple genome-wide association studies have reproducibly identified variations at several genomic loci that are associated with ALL risk, including IKZF1, ARID5B (encoding the AT-rich interactive domain 5B transcription factor), and CEBPE (encoding the transcription factor CCAAT/enhancer-binding protein, epsilon) (118, 119). All three genes encode proteins considered important for normal lymphoid development and/or lymphoid leukemogenesis. CEBPE is a target of translocations in ALL, and Arid5B knockout mice exhibit defects in the B lymphoid compartment (120). Although the functional effects of these inherited variants are poorly understood, the IKZF1 variants may influence level of IKZF1 gene expression (119), suggesting a direct role for these variants in disease susceptibility.

**Conclusions**

The use of contemporary technologies to identify genetic alterations in B-ALL has been tremendously informative, but clearly much work remains to be done. Microarray and candidate gene sequencing approaches are incapable of identifying mutations at nucleotide-level resolution, and initial second-generation sequencing of ALL genomes has identified multiple new genes targeted in ALL. An additional high-priority area of investigation is ALL in older children, adolescents, and adults, who typically have a substantially inferior prognosis. Profiling of ALL in these age groups, as well as in additional uncommon but high-risk ALL subtypes, is ongoing and likely to yield additional insights and therapeutic approaches.

**Acknowledgments**

The author thanks members of his laboratory and colleagues at St. Jude Children’s Research Hospital, the Children’s Oncology Group, and the National Cancer Institute Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Initiative, who have contributed to the studies described in this Review, and apologizes to the many investigators whose important studies could not be cited due to space restrictions. The author is supported by the American Lebanese Syrian Associated Charities of St. Jude Children’s Research Hospital and by funding from the National Cancer Institute, the American Society of Hematology, the American Association of Cancer Research, and Stand Up to Cancer. The author is a Pew Scholar in the Biomedical Sciences and a St. Baldrick’s Scholar.

Address correspondence to: Charles G. Mullighan, Department of Pathology, St. Jude Children’s Research Hospital, 262 Danny Thomas Place, Mail Stop 342, Room 4047G, Memphis, Tennessee 38105, USA. Phone: 901.595.3387; Fax: 901.595.5947; E-mail: charles.mullighan@stjude.org.

Inherited genetic variation and risk of ALL

Several candidate gene studies have implicated inherited genetic variation, for example in the folate metabolic pathway, in the risk of developing leukemia (117). Multiple genome-wide association studies have reproducibly identified variations at several genomic loci that are associated with ALL risk, including IKZF1, ARID5B (encoding the AT-rich interactive domain 5B transcription factor), and CEBPE (encoding the transcription factor CCAAT/enhancer-binding protein, epsilon) (118, 119). All three genes encode proteins considered important for normal lymphoid development and/or lymphoid leukemogenesis. CEBPE is a target of translocations in ALL, and Arid5B knockout mice exhibit defects in the B lymphoid compartment (120). Although the functional effects of these inherited variants are poorly understood, the IKZF1 variants may influence level of IKZF1 gene expression (119), suggesting a direct role for these variants in disease susceptibility.

Conclusions

The use of contemporary technologies to identify genetic alterations in B-ALL has been tremendously informative, but clearly much work remains to be done. Microarray and candidate gene sequencing approaches are incapable of identifying mutations at nucleotide-level resolution, and initial second-generation sequencing of ALL genomes has identified multiple new genes targeted in ALL. An additional high-priority area of investigation is ALL in older children, adolescents, and adults, who typically have a substantially inferior prognosis. Profiling of ALL in these age groups, as well as in additional uncommon but high-risk ALL subtypes, is ongoing and likely to yield additional insights and therapeutic approaches.

Acknowledgments

The author thanks members of his laboratory and colleagues at St. Jude Children’s Research Hospital, the Children’s Oncology Group, and the National Cancer Institute Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Initiative, who have contributed to the studies described in this Review, and apologizes to the many investigators whose important studies could not be cited due to space restrictions. The author is supported by the American Lebanese Syrian Associated Charities of St. Jude Children’s Research Hospital and by funding from the National Cancer Institute, the American Society of Hematology, the American Association of Cancer Research, and Stand Up to Cancer. The author is a Pew Scholar in the Biomedical Sciences and a St. Baldrick’s Scholar.

Address correspondence to: Charles G. Mullighan, Department of Pathology, St. Jude Children’s Research Hospital, 262 Danny Thomas Place, Mail Stop 342, Room 4047G, Memphis, Tennessee 38105, USA. Phone: 901.595.3387; Fax: 901.595.5947; E-mail: charles.mullighan@stjude.org.
review series


103. Irving JA, Minto L, Bailey S, Hall AG. Loss of heterozygosity and somatic mutations of the glucocorticoid receptor gene are rarely found at relapse in pediatric acute lymphoblastic leukemia but may occur in a subpopulation in the early disease course. Cancer Res. 2005;65(21):9712–9718.


