The Biology of B-Progenitor Acute Lymphoblastic Leukemia

Kathryn G. Roberts and Charles G. Mullighan

Department of Pathology, St. Jude Children’s Research Hospital, Memphis, Tennessee 38105, USA
Correspondence: charles.mullighan@stjude.org

Genomic analyses have revolutionized our understanding of the biology of B-progenitor acute lymphoblastic leukemia (ALL). Studies of thousands of cases across the age spectrum have revised the taxonomy of B-ALL by identifying multiple new subgroups with diverse sequence and structural initiating events that vary substantially by age at diagnosis and prognostic significance. There is a growing appreciation of the role of inherited genetic variation in predisposition to ALL and drug responsiveness and of the nature of genetic variegation and clonal evolution that may be targeted for improved diagnostic, risk stratification, disease monitoring, and therapeutic intervention. This review provides an overview of the current state of knowledge of the genetic basis of B-ALL, with an emphasis on recent discoveries that have changed our approach to diagnosis and monitoring.
tic targets that guide precision medicine approaches intended to improve the cure rate while reducing adverse treatment effects.

Here, we will review the genomic landscape of B-ALL with particular emphasis on new subtypes and prognosis and discuss both somatic and inherited variants that contribute to leukemogenesis. The role of the interaction between leukemic cells and the bone marrow microenvironment in disease development and response to treatment will also be discussed.

**RECURRENT CHROMOSOMAL ALTERATIONS AND PROGNOSIS**

The frequency of subtype-defining alterations varies with age (Table 1; Fig. 1). Secondary genetic alterations may be acquired or enriched during disease progression (Mullighan et al. 2007, 2008b; Moorman et al. 2012). Common targets include lymphoid transcription factors (IKZF1, PAX5, EBF1, ETV6), cell cycle regulators and tumor suppressors (CDKN2A/B, TP53, RB1), regulators of lymphoid signaling (BTLA and CD200), Ras pathway signaling (NRAS, KRAS, PTPN11), and chromatin modifiers (CREBBP, SETD2, WHSC1) (Kuiper et al. 2007; Mullighan et al. 2007). The prevalence, gene, and type of alteration vary between subtypes and have different prognostic relevance. Current risk stratification and treatment algorithms incorporate age, sex, presentation white blood cell count, established cytogenetic alterations, and response to initial therapy as measured by levels of minimal residual disease (MRD). Genomic alterations including composite copy number alterations have recently been proposed as important factors for determining prognosis (Hamadeh et al. 2019). Because MRD is such a central component of risk stratification, future clinical trials should aim to integrate new genomic information with response to therapy to develop a comprehensive relapse prediction model (O’Connor et al. 2018).

**Gross Chromosomal Abnormalities**

High hyperdiploidy (nonrandom gain of at least five chromosomes) is present in ∼25% of childhood ALL patients, but accounts for <5% of adolescents and young adults (16–39 yr old; AYA) and adults, and is associated with a favorable outcome. The patterns of chromosomal gain are nonrandom, and most commonly involve chromosomes 4, 10, 14, and 21 and the X chromosome. Mutations involving the Ras pathway (KRAS, NRAS, PTPN11) and epigenetic modifiers are frequent genetic events in hyperdiploid patients (Paulsson et al. 2015). Hyperdiploid ALL with less than 44 chromosomes comprises two principal subtypes with distinct transcriptional profiles and genetic alterations. Historically, hypodiploid ALL has been associated with an unfavorable prognosis (Harrison et al. 2004); however, the outcome is improved with contemporary studies utilizing MRD risk-stratified regimens, and transplantation provides no additional survival benefit compared to chemotherapy alone in MRD-negative patients (Mullighan et al. 2015; Pui et al. 2019). Patients with low hypodiploidy (31–39 chromosomes) commonly harbor deletion of IKZF2 and sequence mutations of TP53 that are frequently inherited (Holmfeldt et al. 2013). This subtype is rare in children (<1%), but increases with age, accounting for >10% of adults, and is associated with a very poor outcome (Moorman et al. 2007; Gu et al. 2019). Patients with near-haploid ALL (24–30 chromosomes) present at a younger age, accounting for ∼2% of childhood ALL (Holmfeldt et al. 2013). Frequent secondary alterations in this subtype include Ras-activating mutations and deletions of IKZF3 (Holmfeldt et al. 2013; Gu et al. 2019). Doubling of the hypodiploid clone (known as masked hyperdiploidy) is common in both near-haploid and low-hypodiploid ALL and results in a modal chromosome number in the hyperdiploid range. Given the markedly differing prognoses of hypodiploid and hyperdiploid ALL, it is important to distinguish masked hypodiploidy (which typically shows four copies of multiple chromosomes in the doubled clone, and copy-neutral loss of heterozygosity of multiple chromosomes) from hyperdiploidy (which typically has multiple trisomic chromosomes).

ALL with intrachromosomal amplification of chromosome 21 (iAMP21) defines a subtype
Table 1. Prevalence and prognosis of subtypes in B-ALL

<table>
<thead>
<tr>
<th>ALL subtype</th>
<th>Median age (yr)</th>
<th>Prevalence</th>
<th>Genomic alterations</th>
<th>Clinical features</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperdiploid (&gt;50 chromosomes)</td>
<td>4</td>
<td>High in children (25%)</td>
<td>Ras pathway, epigenetic modifiers</td>
<td>Excellent prognosis with low WBC</td>
<td>Paulsson et al. 2015</td>
</tr>
<tr>
<td>Low-hypodiploid (31–39 chromosomes)</td>
<td>47</td>
<td>High in adults (10%–15%)</td>
<td>IKZF2 del, TP53 mut (commonly inherited)</td>
<td>Poor prognosis</td>
<td>Holmfeldt et al. 2013</td>
</tr>
<tr>
<td>Near-haploid (24–30 chromosomes)</td>
<td>5.4</td>
<td>&lt;3% in all ages</td>
<td>Ras pathway, IKZF3 del</td>
<td>Intermediate prognosis</td>
<td>Holmfeldt et al. 2013</td>
</tr>
<tr>
<td>iAMP21 Copy number gain</td>
<td>10</td>
<td>~3% in children and AYA</td>
<td>Complex structural alterations of chromosome 21</td>
<td>Good prognosis with intensive therapy</td>
<td>Harrison 2015</td>
</tr>
<tr>
<td>ETV6-RUNX1 t(12;21)(p13;q22)</td>
<td>4</td>
<td>High in children (25%)</td>
<td>ETV6 fusions and del, IKZF1 del</td>
<td>Poor prognosis with bortezomib or DOT1L inhibition</td>
<td>Lilljebjörn et al. 2015; Zaliova et al. 2017</td>
</tr>
<tr>
<td>DUX4-rearranged</td>
<td>14.3</td>
<td>~5% in children</td>
<td>ERG del, IKZF1 del, Ras pathway</td>
<td>Good prognosis, CNS relapse</td>
<td>Barber et al. 2007; Yamada et al. 2010</td>
</tr>
<tr>
<td>KMT2A-rearranged</td>
<td>40</td>
<td>High in infants (~50%) and adults (~15%)</td>
<td>Ras pathway (commonly subclonal)</td>
<td>Poor prognosis, sensitive to HDAC inhibition</td>
<td>Andersson et al. 2015</td>
</tr>
<tr>
<td>TCF3-PBX1 t(1;19)(q23;p13)</td>
<td>8</td>
<td>~5% in children, rarely in adults</td>
<td>DUX4-rearranged</td>
<td>Intermediate prognosis, sensitive to HDAC inhibition</td>
<td>Gu et al. 2018; Gu et al. 2017</td>
</tr>
</tbody>
</table>

Continued
### Table 1. Continued

<table>
<thead>
<tr>
<th>ALL subtype</th>
<th>Category</th>
<th>Median age (yr)</th>
<th>Prevalence</th>
<th>Genomic alterations</th>
<th>Clinical features</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCF3-HLF t(17;19) (q22;p13)</td>
<td>TF rearrangement</td>
<td>15</td>
<td>Very rare in all ages (&lt;1%)</td>
<td>TCF3 mut, PAX5 del, Ras pathway</td>
<td>Very poor prognosis, sensitive to Bcl2 inhibition</td>
<td>Fischer et al. 2015</td>
</tr>
<tr>
<td>PAX5alt</td>
<td>Other TF-driven</td>
<td>10</td>
<td>Highest in children (~11%)</td>
<td>PAX5 fusion, mut, amp</td>
<td>Intermediate prognosis</td>
<td>Li et al. 2018; Gu et al. 2019</td>
</tr>
<tr>
<td>PAX5 P80R</td>
<td>Other TF-driven</td>
<td>22</td>
<td>Highest in adults (~4%)</td>
<td>Ras pathway</td>
<td>Intermediate prognosis</td>
<td>Li et al. 2018; Gu et al. 2019</td>
</tr>
<tr>
<td>IKZF1 N159Y</td>
<td>Other TF-driven</td>
<td></td>
<td>Very rare in all ages (&lt;1%)</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Li et al. 2018; Gu et al. 2019</td>
</tr>
<tr>
<td>BCL2/MYC-rearranged</td>
<td>Other TF-driven</td>
<td>48</td>
<td>Almost exclusively in AYA and adults (~3%)</td>
<td>Unknown</td>
<td>Poor prognosis</td>
<td>Gu et al. 2019</td>
</tr>
<tr>
<td>Ph-like</td>
<td>Kinase-driven</td>
<td>21</td>
<td>Peaks in AYA (25%–30%)</td>
<td>Multiple kinase alterations, IKZF1 del and mut, CDKN2A/B del</td>
<td>Poor prognosis, amenable to TKI therapy</td>
<td>Roberts et al. 2014a, 2017a</td>
</tr>
<tr>
<td>BCR-ABL1 t(9;22) (q34;q11.2)</td>
<td>Kinase-driven</td>
<td>40–45</td>
<td>5% in children, highest in adults (40%–50%)</td>
<td>IKZF1 del and mut, CDKN2A/B del</td>
<td>Historically poor prognosis, improved with TKI</td>
<td>Mullighan et al. 2008a; Roberts et al. 2014a, 2017a</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>16</td>
<td>~5% in children, ~10% in AYA and adults</td>
<td>Unknown</td>
<td>Intermediate prognosis</td>
<td></td>
</tr>
</tbody>
</table>

(AYA) Adolescent and young adult, (amp) amplification, (B-ALL) B-progenitor acute lymphoblastic leukemia, (CNS) central nervous system, (del) deletion, (HDAC) histone deacetylase, (mut) sequence mutation, (TF) transcription factor, (TKI) tyrosine kinase inhibitor.
of ALL that is more common in older children (median age 10 yr), and is rarely observed in patients older than 30 yr (Harrison et al. 2014). The role of iAMP21 in leukemogenesis is unclear, but a common region of amplification includes ERG and DYRK1A with gain of at least two copies of RUNX1 (Li et al. 2014). Secondary events include the P2RY8-CRLF2 fusion and genetic alterations in kinase signaling, including IL7R and FLT3. Improved risk stratification and treatment with intensive therapy can rescue the poor outcome of these patients when treated as standard-risk (Moorman et al. 2013).

Translocations

ETV6-RUNX1, encoded by the t(12;21)(p13; q22) translocation, is another favorable cytoge-
nentic alteration with a high frequency in childhood ALL (25%) and <5% in AYAs and adults. Secondary DNA copy number alterations, notably PAX5 deletion, and mutation of WHSCI are frequent genetic events in patients harboring ETV6-RUNX1 (Mullighan et al. 2007; Jaffe et al. 2013; Papaemmanuil et al. 2014). KMT2A (MLL) rearrangements are a hallmark of infant ALL (age < 1 yr). They also account for a significant proportion of adults with ALL (~15%) and are associated with a poor prognosis in all ages (Hunger and Mullighan 2015). The reasons underlying the biphasic distribution in age are not well understood. In infant cases, KMT2A rearrangement is frequently acquired in utero (Ford et al. 1993), and patients harbor very few cooperating lesions, suggesting the rearrangement itself is sufficient to induce leukemia (Andersson et al. 2015). The most commonly perturbed pathways include PI3K and Ras signaling, with the majority of mutations present at a low tumor burden (Driessen et al. 2013; Andersson et al. 2015; Agraz-Doblas et al. 2019). Subclonal activating mutations of FLT3 were recently shown to accelerate disease onset in a mouse model of KMT2A-rearranged leukemia, suggesting these alterations can influence the rate of leukemogenesis even at low levels (Hyrenius-Wittsten et al. 2018).

TCF3-PBX1, encoded by the t(1;19)(q23; p13) translocation, is present in ~5% of children and less in AYAs and adults. Previously considered a high-risk subtype with a propensity to central nervous system relapse, it is now associated with a favorable outcome on contemporary ALL therapies (Barber et al. 2007; Burmeister et al. 2010). By contrast the t(17;19)(q22;p13) translocation, encoding the TCF3-HLF fusion gene, defines a rare subtype of ALL (<1% in all ages) with a distinct transcriptional profile that is typically associated with an overall survival of <2 yr from diagnosis (Inaba et al. 1992; Hunger 1996). Interestingly, primary leukemic cells harboring TCF3-HLF show sensitivity to the BCL2 inhibitor venetoclax (ABT-199), identifying a new therapeutic option for this fatal subtype (Fischer et al. 2015).

BCR-ABL1 ALL is uncommon in children (2%-5% of patients), but accounts for at least 25% of adults (Roberts et al. 2014a, 2017a). The addition of ABL1 tyrosine kinase inhibitors (TKIs) to chemotherapeutic regimens in both children and adults has significantly improved the survival of BCR-ABL1-positive patients (Ravandi et al. 2010; Schultz et al. 2014; Slayton et al. 2018). IKZF1 alterations (deletion or mutation) are a hallmark of kinase-driven ALL (Ph+ and Ph-like) and are associated with treatment failure and relapse, even in the era of TKI therapy (Mullighan et al. 2008a; Martinelli et al. 2009; Roberts et al. 2014a; Slayton et al. 2018). The co-occurrence of IKZF1 deletions with CDKN2A/B, PAX5, or PAR1 deletions in the absence of ERG deletions (termed IKZF1plus) detected by multiplex ligation probe amplification (MLPA) in childhood ALL confers a worse prognosis compared to patients with IKZF1 deletion who do not fulfill the criteria for IKZF1plus (Stanulla et al. 2018). Although technically straightforward, identification of IKZF1plus as a biomarker of poor outcome is limited by the inability of the MLPA approach to identify the full spectrum of IKZF1 alterations, cases with high-risk ALL that do not have IKZF1 alterations, and the lack of ERG deletion in approximately one-third of DUX4 cases that commonly have IKZF1 alterations and favorable outcome.

NEW SUBTYPES IN B-ALL

The application of comprehensive sequencing and integrative analyses continues to refine the genomic landscape of ALL, resulting in the identification of new entities with prognostic and therapeutic significance. Rearrangements in these new subtypes involve a diverse range of partners that converge on a single gene (e.g., MEF2D and ZNF384-rearranged ALL) or are cryptic by cytogenetic analysis (e.g., DUX4-rearranged ALL). Other subtypes may harbor alteration of a range of driver genes by diverse mechanisms (e.g., Ph-like ALL) or are initiated by sequence mutations (e.g., PAX5 P80R and IKZF1 N159Y). Additional groups have similar gene expression profiles to known subtypes with different genetic alterations (Ph-like and ETV6-RUNX1-like ALL).
Ph-like ALL: An Opportunity for Targeted Therapies

Philadelphia chromosome like (Ph-like or BCR-ABL1-like) ALL was incorporated as a provisional entity to the revision of the World Health Organization (WHO) classification of acute leukemia in 2016 (Arber et al. 2016). Leukemic cells from patients with Ph-like ALL have similar transcriptional profiles to Ph+ ALL but lack the BCR-ABL1 fusion gene (Den Boer et al. 2009; Mullighan et al. 2009). Similar to with Ph+ ALL, the incidence of Ph-like ALL increases with age, comprising 10%–15% of childhood ALL cases, >20% of adults, and peaking at 25%–30% in AYAs (Loh et al. 2013; Roberts et al. 2014a, 2017a, 2018; Jain et al. 2017a; Reshmi et al. 2017; Tasian et al. 2017a). In all ages, Ph-like ALL is associated with elevated MRD levels and/or higher rates of treatment failure compared to non-Ph-like ALL patients (Roberts et al. 2014a, 2017a; Tasian et al. 2017b). Thus, the inferior treatment outcomes in AYA and adults may be partly explained by the high prevalence of Ph-like ALL. In Children’s Oncology Group (COG) cohorts of National Cancer Institute (NCI) standard-risk (SR) ALL, Ph-like ALL is less common and confers a better prognosis compared to children with high-risk (HR) ALL (Roberts et al. 2018). Furthermore, children with Ph-like ALL treated on St. Jude Total XV studies had a favorable outcome with MRD risk-directed therapy intensification (Roberts et al. 2014b).

Ph-like ALL is genetically heterogeneous and is characterized by rearrangements, copy number alterations, and sequence mutations that activate tyrosine kinase or cytokine receptor signaling. Despite this complexity, most alterations can be divided into a limited number of distinct subgroups based on the activated kinase and signaling pathways. These include rearrangements or, less commonly, sequence mutations of CRLF2 (IGH-CRLF2, P2RY8-CRLF2), fusions involving ABL-class genes (ABL1, ABL2, CSF1R, LYN, PDGFRα, PDGFRβ), alterations activating JAK-STAT signaling (including rearrangements of JAK2, EPOR or TYK2) and mutations/deletions of IL7R, SH2B3, JAK1, JAK3, TYK2, IL2RB), Ras signaling pathways (NRAS, KRAS, PTPN11), and less common fusions (FLT3, FGFR1, NTRK3, PTK2B) (Fig. 2A; Roberts et al. 2014a, 2017a; Reshmi et al. 2017). The frequency of each kinase subgroup varies with age, particularly with respect to CRLF2 rearrangements, in which IGH-CRLF2 accounts for almost 50% of Ph-like ALL in AYAs and adults but is less common in children. ABL-class fusions are most prevalent in children with HR ALL (Fig. 2B). Fewer kinase alterations are identified in Ph-like ALL patients with SR ALL (Roberts et al. 2018). A small subset of children harboring rearrangement of CRLF2—most commonly P2RY8-CRLF2 and with Down syndrome a hallmark of this transcriptional signature (Gu et al. 2019).

The majority of Ph-like alterations can be targeted effectively in preclinical models using a combinatorial approach of chemotherapy with ABL1 (e.g., dasatinib) or JAK inhibition (e.g., ruxolitinib) (Roberts et al. 2017b), and a number of case reports demonstrate efficacy of ABL1 TKI treatment in Ph-like ALL patients with refractory disease (Lengline et al. 2013; Weston et al. 2013; Kobayashi et al. 2015; Schwab et al. 2016). This approach is currently being tested in frontline studies of patients treated at St. Jude Children’s Research Hospital (Total XVII, NCT03117751) (Inaba et al. 2017) and on COG protocols (AALL1131, NCT01406756; AALL1521, NCT02723994) (Tasian et al. 2017b).

ETV6-RUNXI-like ALL

Analogous to Ph-like ALL, ETV6-RUNXI-like ALL is defined by having a gene expression profile and immunophenotype (CD27 positive, CD44 low to negative) similar to ETV6-RUNX1 ALL, but lacking the ETV6-RUNX1 fusion (Lilljebjörn et al. 2016; Zaliova et al. 2017). Unsurprisingly, like ETV6-RUNX1 ALL, ETV6-RUNXI-like ALL is almost exclusively identified in children (~3%) and confers a favorable prognosis. This subtype is associated with alternate lesions (gene fusions or copy number alterations) in ETV6, IKZF1, or TCF3, suggesting global deregulation of lymphoid development is a hallmark of this transcriptional signature (Gu et al. 2019).
Figure 2. (A) Kinase alterations and signaling pathways dysregulated in Philadelphia chromosome-like (Ph-like) ALL. The majority of kinase and cytokine receptor alterations converge on two pathways that activate JAK-family member signaling or ABL signaling. Alterations that activate JAK-STAT signaling can be targeted with JAK and PI3K inhibitors. ABL-class alterations can be targeted with ABL-inhibitors such as dasatinib. Other kinase alterations and those that activate Ras signaling can be targeted with specific inhibitors including those that inactivate TRK, FLT3, FGFR1, and MEK for the MAPK pathway. (B) Distribution of kinase subtypes in Ph-like ALL within each age group (Roberts et al. 2014a, 2017a, 2018; Reshmi et al. 2017). Combined prevalence of Ph-like ALL subtypes in childhood National Cancer Institute (NCI) standard-risk (SR; age 1–9.99 yr and WBC < 50,000/µL), NCI high-risk (HR; age 10–15 yr or WBC ≥ 50,000/µL), adolescent and young adults (16–39 yr), and adults (≥ 40 yr). Genomic subtypes include IGH-CRLF2, P2RY8-CRLF2, and ABL-class fusions (ABL1, ABL2, CSF1R, LYN, PDGFRα, and PDGFRβ); JAK2 and EPOR rearrangements and other mutations in JAK–STAT signaling (JAK1/3, IL7R, SH2B3, TYK2, and IL2RB); and other kinase alterations (FLT3, FGFR1, NTRK3), Ras mutations (KRAS, NRAS, NF1, PTPN11, BRAF, and CBL), and unknown alterations.
DUX4-rearranged ALL

An interesting subtype of B-ALL with a very distinctive gene expression profile and immunophenotype (CD2 and CD371 positive) is characterized by genetic alterations and deregulation of the transcription factor genes DUX4 (double homeobox 4) and ERG (ETS-related gene) (Yeoh et al. 2002; Harvey et al. 2010; Lilljebjörn et al. 2016; Yasuda et al. 2016; Zhang et al. 2016; Schinnerl et al. 2019). DUX4 is located in microsatellite D4Z4 repeat domains in the subtelomeric region of chromosome 4 that is duplicated on chromosome 10q and is normally exclusively expressed in germinal tissues (Gatica and Rosa 2016). In DUX4-rearranged ALL, translocation or insertion of DUX4 to IGH is the initiating event that results in overexpression of a 3′ truncated isoform of DUX4 not normally expressed in B cells. The aberrantly expressed DUX4 binds to an intragenic region of ERG, resulting in gross transcriptional deregulation of ERG, and, commonly, expression of ERGalt, a transcript that utilizes a noncanonical first exon that encodes a truncated carboxy-terminal ERG protein. ERGalt retains the DNA-binding and transactivating domain of ERG, inhibits the transcriptional activity of wild-type ERG, and is transforming in mouse models of B-ALL (Zhang et al. 2016). This subtype accounts for 5%–10% of B-ALL, with a slight peak in AYAs. Of clinical relevance, DUX4-rearranged ALL is associated with an excellent prognosis in both children and adults (Gu et al. 2019), even despite the presence of secondary genetic alterations otherwise associated with poor outcome, such as IKZF1 deletions, which are present in ~40% of DUX4-rearranged ALL (Zhang et al. 2016).

New Transcription Factors: MEF2D and ZNF384

Recurrent rearrangements of MEF2D and ZNF384 account for ~4% and 5% of children and up to 7% and 10% in AYA patients, respectively. Accordingly, both subtypes are associated with older age of onset (median age 14 and 15 yr) (Gu et al. 2016; Liu et al. 2016; Suzuki et al. 2016).

Multiple 3′ partners have been identified for MEF2D (encoding myocyte enhancer factor 2D), including BCL9, CSF1R, DAZAP1, FOXJ2, HNRNPUL1, HNRNPH1, and SS18 (Gu et al. 2016; Ohki et al. 2019). All fusions preserve the MEF2D MADS-box domain that mediates DNA binding, resulting in enhanced transcriptional activity and deregulation of MEF2D targets (Gu et al. 2016). An exception is MEF2D-CFS1R, which displays the Ph-like gene expression profile (Roberts et al. 2014a). MEF2D-rearranged ALL is associated with an aberrant immunophenotype (CD10 negative, CD38 positive) and an intermediate to poor outcome (Gu et al. 2016; Suzuki et al. 2016; Ohki et al. 2019). Alterations of PHF6, recurrently mutated in T-cell ALL, were the most frequent cooperating lesions identified by targeted sequencing (Ohki et al. 2019). Deregulation of MEF2D also results in the overexpression of HDAC9 (histone deacetylase 9), which can be targeted therapeutically using HDAC inhibitors (Gu et al. 2016).

Rearrangements of ZNF384 (encoding zinc finger 384) define a subtype of acute leukemia that transcends immunophenotypic classification and may manifest as classical pre-B ALL without lineage aberrancy, B-ALL with expression of the myeloid markers (CD13/33), or B/myeloid mixed phenotype acute leukemia (Alexander et al. 2018). To date, 11 different 5′ fusion partners, usually involving a transcriptional regulator or chimeratin modifier, have been identified for ZNF384: ARID1B, BMP2K, CLTC, CREBBP, EP300, EWSR1, NIPBL, SMARCA2, SYNRG, TAF15, and TCF3 (Liu et al. 2016; Shago et al. 2016; Yasuda et al. 2016; Hirabayashi et al. 2017). An intermediate prognosis has been described in small pediatric cohorts (Liu et al. 2016). The rearrangements are also distinctive, usually involving the entire coding region of ZNF384, resulting in the expression of wild-type ZNF384 in a lineage inappropriate manner, as well as the chimeric fusion protein. Studies of hematopoietic progenitor cells from primary leukemia samples, as well as xenografting of immunophenotypically multilocal populations, has shown that ZNF384 rearrangements are acquired in a subset of hematopoietic stem cells and prime leukemic cells.
for lineage plasticity (Alexander et al. 2018). More recently, cases harboring rearrangement of the zinc finger ZNF362 to SMARCA2 and TAF15 were shown to cluster with ZNF384-rearranged ALL, indicating deregulation of similar downstream targets (Li et al. 2018).

REDEFINING “OTHER” B-ALL

Despite the advances made in refining the classification of B-ALL, until recently, almost one-quarter of cases across the age spectrum lacked a subtype defining lesion and were collectively known as “Other.” These cases were excluded from risk stratification, commonly relapsed, and lacked targeted therapeutic approaches. To systematically define the frequency and prognostic significance of subtypes across the age spectrum, two groups recently performed an integrated large scale genomic analysis of 1223 and 1988 B-ALL cases, respectively, using transcriptional profiling to refine subtype classification (Li et al. 2018; Gu et al. 2019). In addition to known groups, including those defined by aneuploidy, up to five new subtypes were identified with distinct gene expression signatures, accounting for an additional 15% of B-ALL. As such, >90% of ALL cases may be classified into distinct genetic subtypes using these algorithms.

PAX5-Driven Subtypes

PAX5 is largely considered to function as a haploinsufficient tumor suppressor in ALL, with secondary heterozygous deletions and loss-of-function mutations present in one-third of all patients with B-ALL across a range of subtypes (Kuiper et al. 2007; Mullighan et al. 2007). In mouse models, Pax5 heterozygosity cooperates with constitutive activation of the JAK-STAT pathway to promote B-ALL development, supporting its role as a tumor suppressor (Dang et al. 2015). PAX5 translocations are reported in 2%–3% of B-ALL (Nebral et al. 2009; Coyaud et al. 2010). Recent analyses identified two PAX5 subtypes defined by distinct transcriptional profiles and genetic alterations. The first subtype, referred to as PAX5-altered (PAX5alt), comprises cases with diverse PAX5 rearrangements (most commonly to ETV6 or NOLAL), sequence mutations or intragenic amplification (Schwab et al. 2017), with the highest prevalence observed in children and AYA (10% each vs. 7% in adults) (Gu et al. 2019). The second group of PAX5-driven ALL is defined by the presence of the PAX5 P80R mutation, which is homozygous in almost all cases because of deletion or frameshift mutation of the wild-type PAX5 allele, suggesting that loss of both PAX5 alleles drives the unique gene expression profile of this subtype (Fig. 3; Li et al. 2018; Gu et al. 2019; Passet et al. 2019). The prevalence of PAX5 P80R increases with age, accounting for almost 5% of adults. This subtype confers an intermediate to favorable prognosis in both children and adults (Bastian et al. 2019; Gu et al. 2019; Passet et al. 2019; Zaliova et al. 2019). Cooperating lesions identified in PAX5 P80R patients include a high frequency of signaling mutations, particularly in the Ras, JAK-STAT, and other kinase signaling pathways (FLT3, PIK3CA), highlighting the potential for targeted therapies (Gu et al. 2019; Passet et al. 2019). Notably, heterozygous Pax5P80R/+ or homozygous Pax5P80R/P80R knock-in mice develop B-progenitor ALL that is transplantable, and tumors that arise in Pax5P80R/+ mice genetically inactivate the wild-type Pax5 allele by deletion or truncation, recapitulating the loss of wild-type PAX5 observed in human ALL (Gu et al. 2019). In a mouse model of B-ALL, PAX5-ETV6 activated distinct transcriptional pathways including pre-B cell receptor signaling and migration/adhesion, confirming its role as an oncoprotein rather than simply acting as a competitive inhibitor of the wild-type PAX5 protein (Smeenk et al. 2017). The identification of PAX5 subtypes as distinct entities highlights the importance of this gene in regulating B-cell differentiation, and confirms PAX5 alterations as founding lesions in B-lymphoid leukemogenesis as opposed to secondary cooperating events as previously thought.

IKZF1 N159Y

Another uncommon subtype (accounting for <1% of ALL) defined by a single mutation in a lymphoid transcription factor includes cases
Figure 3. (A) Genetic alterations of PAX5, including gene rearrangements (PAX5r), sequence mutations (PAX5mut), and focal intragenic amplifications (PAX5amp, pink in PAX5 CNA) observed in the PAX5alt cohort. (B) Protein domain plot of PAX5 showing the mutations detected in PAX5alt and other B-ALL subtypes (top panel) and in the PAX5 P80R subtype (bottom panel). (CNA) Copy number alteration.
NUTM1 Rearrangements

An additional subtype present exclusively in 1% of childhood ALL (median age 3 yr) involves fusion of almost all the coding region of NUTM1 (nuclear protein in testis midline carcinoma family 1) to six different 5’ partners—ACIN1, BRD9, CUX1, IKZF1, SLC12A6, and ZNF618—resulting in increased expression of NUTM1 (Li et al. 2018; Gu et al. 2019). NUTM1 is normally expressed in the testis and acts as a chromatin modifier by recruiting EP300 (p300) to increase local histone acetylation (Alekseyenko et al. 2015). Fusions of NUTM1 (commonly BRD4-NUTM1) are a hallmark of NUT midline carcinoma (NMC), an aggressive and fatal subtype of squamous cell carcinoma that also arises frequently in children (French 2014). BRD4-NUTM1 acts to repress differentiation in NMC by recruiting histone acetyltransferases and other transcriptional cofactors to regions of chromatin that are actively transcribing pro-proliferative and antidifferentiation genes, including MYC (French 2014). Thus, fusions such as BRD9-NUTM1 in ALL may have a similar mechanism of action, although experimental studies are required to elucidate the role of NUTM1 in leukemogenesis. In contrast to NMC, ALL patients with NUTM1 rearrangements have an excellent prognosis. Given the involvement of BRD9, bromodomain or HDAC inhibitors would be a logical targeted therapeutic approach for these patients.

MIXED PHENOTYPE ALL

Mixed phenotype acute leukemia (MPAL) is characterized by expression of cell surface proteins characteristic of multiple lineages, most commonly B and myeloid (B/M MPAL) or T and myeloid (T/M MPAL) markers, either in a single (biphenotypic) or multiple (bilineal) immunophenotypic subpopulations. Prior studies had identified rearrangements of KMT2A (MLL) or the BCR-ABL1 fusion in a minority of cases, but until recently the genetics of MPAL had been poorly understood. However, this is of great interest given the phenotypic plasticity and poor prognosis of this form of leukemia. Genomic analyses have shown that T/M and B/M are genetically distinct, with T/M leukemia characterized by founder mutations or rearrangements in transcription factors and chromatin modifiers (WT1, ETV6, RUNX1, CEBPA) and the majority of B/M cases to harbor rearrangements of ZNF384 (Alexander et al. 2018; Takahashi et al. 2018; Xiao et al. 2018). The phenotypic plasticity and characteristic of MPAL (that has bedeviled the selection of appropriate therapy) is largely independent of genetic variegation and, rather, is due to the
acquisition of founding lesions in very early hematopoietic progenitors. Thus, MPAL forms part of a spectrum of immature/stem cell leukemias (for T/M MPAL, like early T cell precursor ALL) (Zhang et al. 2012), and future studies are integrating ALL-directed therapy and genomic analysis to further refine optimal diagnostic and classification approaches (Hrusak et al. 2018).

INHERITED VARIANTS IN ALL

Genome-wide association studies (GWASs) have identified risk loci with common genetic polymorphisms that are associated with a modest increase in ALL susceptibility, including *IKZF1* (7p12.2), *CDKN2A/CDKN2B* (9p21), *PIP4K2A* (10p12.2), *GATA3* (10p14), *ARID5B* (10q21.2), *CEBPE*, and *ERG* (14q11.2) (Moriyama et al. 2015b; Qian et al. 2019). Associations with several of these loci exhibit a degree of ALL subtype specificity—for example, *GATA3* with Ph-like ALL (Perez-Andreu et al. 2013; Jain et al. 2017b) and *ERG* with TCF3-PBX1—suggesting an interplay of germline and somatic alterations in leukemogenesis. More recently, studies of families with multiple individuals with ALL and complementary examinations of large cohorts of patients with presumed sporadic ALL have identified deleterious germline variants in genes that are also targets of somatic mutation in ALL, including *PAX5*, *ETV6*, *IKZF1*, *TP53*, and *ERG*.

A role for *PAX5* in autosomal dominant predisposition to B-ALL was identified by the description of three unrelated families who harbored a germline *PAX5* c.547G > A mutation in the octapeptide domain (*PAX5 G183S*) that resulted in moderate attenuation of transcriptional activity in vitro (Shah et al. 2013; Auer et al. 2014). Notably, all affected individuals had somatic loss of the wild-type allele, suggesting that biallelic inactivation of *PAX5* is also important for B-cell leukemogenesis in this context.

Deleterious germline variants within the DNA-binding domain of *ETV6* are present in 1% of sporadic B-ALL and affect transcriptional repression either by abrogating binding to ETS-containing DNA sequences or through altered intracellular localization (Moriyama et al. 2015a; Noetzli et al. 2015; Topka et al. 2015; Zhang et al. 2015). Multiple subsequent reports suggest that *ETV6* sequence mutations may be the most common germline alterations predisposing to ALL (Feurstein and Godley 2017; Hock and Shimamura 2017; Duployez et al. 2018). Moreover, a focal germline *ETV6* splice site deletion resulting in exon skipping and protein truncation has been reported in a highly penetrant family (Rampersaud et al. 2019). Another report identified a constitutional translocation disrupting *ETV6* (Jarviah et al. 2019). These studies indicate that careful analysis of germline structural variants is required to describe the full repertoire of deleterious germline alterations in ALL.

Churchman et al. reported inherited germline variants in *IKZF1* that impair its function in a similar manner to somatic mutations. In contrast to somatic *IKZF1* alterations that are most commonly deletions or mutations in the amino-terminal (DNA-binding) or carboxy-terminal (dimerizing) zinc fingers (Churchman et al. 2015), the germline variants are scattered throughout the gene in regions of poorly characterized function and were not predicted to be deleterious by in silico analyses, but were highly deleterious in more sophisticated cellular assays including subcellular mislocalization, cell–cell adhesion, and cell stromal adhesion in vivo (Churchman et al. 2018).

*TP53* alterations are a hallmark of low-hypodiploid ALL, with almost half occurring in the germline, suggesting that low-hypodiploid ALL is another manifestation of Li–Fraumeni syndrome (Holmfeldt et al. 2013). In a large cohort of childhood ALL, 49 nonsilent rare *TP53* coding variants were identified in 77 patients, of which 22 variants were classified as pathogenic (Qian et al. 2018). Children with *TP53* pathogenic variants presented at an older age, had inferior outcomes to children with wild-type *TP53*, and were more likely to develop second malignancies. This study also confirmed the association of inherited *TP53* variants with hypodiploid ALL (Qian et al. 2018). A recent GWAS identified novel susceptibility variants at the *ERG* locus that were enriched in Hispanics (Qian et al. 2019), providing additional insight...
into the relationship of germline genetic variation in racial occurrence and outcomes in ALL (Yang et al. 2011; Karol et al. 2017). Together, these studies highlight the importance of these genes in both de novo and familial ALL.

RELAPSED ALL

Relapsed ALL remains a leading cause of childhood cancer death (Curtin et al. 2016) and is associated with high rates of treatment failure and death in older individuals (Fielding et al. 2007; Stock 2010; Frey and Luger 2015). The main curative approach for adults is an allogenic stem cell transplant; however, survival rates for relapsed ALL are improving with the implementation of new immunotherapeutic approaches including blinatumomab (CD19/CD3 bispecific T-cell engager), inotuzumab ozagamicin (anti-CD22 antibody conjugated to calicheamicin), and CAR T cells (chimeric antigen receptor) (Davila et al. 2014; Kantarjian et al. 2016, 2017; Maude et al. 2018; Park et al. 2018).

Genomic studies in childhood ALL show that predominant clones at diagnosis are often eradicated, and relapse arises from a minor clone that already harbors and/or acquires additional genomic alterations that drive resistance in a drug-specific or -agnostic manner (Fig. 4; Mullighan et al. 2008b, 2011; Li et al. 2015; Ma et al. 2015; Oshima et al. 2016; Tzoneva et al. 2018). Mutations in genes encoding epigenetic regulators and chromatin modifiers are recurrent events in relapsed ALL and can directly influence response to treatment (Mullighan et al. 2011; Mar et al. 2014; Ma et al. 2015). In particular, mutations in the transcriptional coactivator and acetyl transferase

Figure 4. Commonly altered pathways and stepwise progression of B-progenitor acute lymphoblastic leukemia (B-ALL). Common genetic polymorphisms (IKZF1, CDKN2A/B, PIP4K2A, GATA3, ARID5B, CEBPE, and ERG) and deleterious nonsilent inherited variants (PAX5, IKZF1, TP53, and ERG) increase the risk of ALL susceptibility. Driving or founding lesions of ALL define genomic subtypes: aneuploidy and other chromosomal abnormalities (hyperdiploid, low-hypodiploid, near-haploid, iAMP21), rearrangements deregulating transcription factors (ETV6-RUNX1, ETV6-RUNX1-like, KMT2A, TCF3-PBX1, DUX4, ZNF384, MEF2D, NUTM1, TCF3-HLF, PAX5, BCL2/MYC) or kinase genes (Ph-like, BCR-ABL1), and specific mutations in lymphoid transcription factors (PAX5 P80R, IKZF1, N195Y). Deletion and loss of lymphoid transcription factors (e.g., IKZF1, PAX5, EBF1) coupled with the alteration of tumor suppressors and cell cycle regulators (CDKN2A/B, TP53), kinase signaling pathway genes (e.g., NRAS, KRAS, FLT3), other transcriptional regulators (e.g., ETV6, ERG), or epigenetic regulators (e.g., CREBBP, WHSC1, CTCF) result in the accumulation of immature lymphoid blasts and presentation at diagnosis. During treatment, the predominant diagnosis clone is commonly eradicated and relapse arises from a minor clone that already harbors and/or acquires additional genetic alterations that drive resistance. Pathways that are enriched at relapse include those involving epigenetic regulators (e.g., CREBBP, SETD2, KDM6A), the glucocorticoid response (e.g., CREBBP, NR3C1), and thiopurine metabolism (e.g., NT5C2, MSH6).
CREBBP occur in up to 20% of relapsed ALL and impair sensitivity to glucocorticoid therapy (Mullighan et al. 2011). Mutations in NT5C2 (5′-nucleotidase catalytic enzyme II) confer increased resistance to purine analogs at the cost of impaired leukemia cell growth and leukemia-initiating cell activity (Meyer et al. 2013; Tzonneva et al. 2018). In addition, loss of MSH6, a major component of the mismatch repair (MMR) system, results in intrinsic chemoresistance to thiopurines because of an inability to recognize thioguanine nucleotide mismatching and failure to initiate MMR. Thus, cells defective for MSH6 do not undergo cell cycle arrest or apoptosis and continue to proliferate in the presence of thiopurine (Evensen et al. 2018). Other recurrent somatic alterations in relapsed ALL include deletions of the glucocorticoid receptor NR3CI and mutations in the H3K36 trimethyltransferase SETD2, the lysine-specific demethylase KDM6A, and the epigenetic regulator MLL2 (Mar et al. 2014; Ma et al. 2015). Enhancing our knowledge of relapse-enriched or acquired alterations is important for initial risk stratification and has implications for molecular monitoring given the increasingly widespread application of deep sequencing approaches to identify low levels of MRD.

ROLE OF THE MICROENVIRONMENT IN ALL
Most studies of mechanisms of leukemogenesis and treatment response have focused on leukemic cell-intrinsic features, but it is increasingly apparent that tumor cell-extrinsic factors, including the nature of nonleukemic hematopoietic cells, and the interaction of leukemic cells with the bone marrow microenvironment, are important determinants of response to therapy and may also be directly influenced by genetic alterations of the leukemic cell. This is exemplified by the finding that alterations of IKZF1 (Ikaros) in kinase-driven (Ph+ and Ph-like) ALL drive high-risk disease by derepressing expression of adhesion molecules that result in acquisition of a hematopoietic stem cell like phenotype and aberrant leukemic intercellular and cell-stromal adhesion (Joshi et al. 2014; Churchman et al. 2015). This leads to perturbed bone marrow mislocalization and resistance to therapy that may be circumvented, at least in this context, by retinoids (that bind to retinoid X receptor α, which is also derepressed by loss of Ikaros) that result in differentiation and up-regulation of wild-type Ikaros. Another approach is focal adhesion kinase (FAK) inhibitors, which inhibit FAK signaling downstream of integrin activation (Churchman et al. 2016), an approach that is entering the clinic for the treatment of solid tumors (Lee et al. 2015) and in conjunction with immunotherapy (Jiang et al. 2016).

Although there is extensive evidence that remodeling of, and interaction with, the bone marrow hematopoietic niche has an important role in the survival of acute myeloid leukemia cells (Tabe and Konopleva 2014), the nature and importance of the ALL cell microenvironment interaction is less well studied, but is likely important in light of findings that disruption of CXCR4-CXCL12-mediated interaction can improve drug responsiveness in experimental models of B-ALL and T-cell acute lymphoblastic leukemia (T-ALL) (Pitt et al. 2015; Randhawa et al. 2016). Our recent data indicate that interaction of leukemic cells with bone marrow stromal cells results in profound deregulation of adhesion, signaling cascades, and epithelial to mesenchymal transition–like phenotype in ALL cells and accompanying drug resistance that can potentially be exploited therapeutically (Yoshihara et al. 2018).

CONCLUSIONS
Within the last decade, integrated genomic analyses of large cohorts of childhood ALL, and more recently AYA and adult ALL, has revolutionized our understanding of the genetic basis of ALL by identifying new subtypes, dysregulated pathways, and therapeutic targets that have led to improved risk stratification and treatment strategies. Despite these advances, a proportion of ALL cases cannot be categorized into any of the currently established subtypes, and ongoing discovery studies are required to fully define the genomic landscape. Recent discoveries have already had substantial impact on diagnosis and management of the disease.
For example, targeted approaches are being tested in multiple trials of Ph-like ALL, and the appreciation that accurate classification and risk stratification requires genomic approaches that detect complex structural events in addition to sequence alterations has led to the increasingly widespread adoption of RNA sequencing and, to sequence alterations has led to the increasing...  

ACKNOWLEDGMENTS  
The authors thank colleagues at St. Jude, the Children’s Oncology Group, and the multiple centers and leukemia cooperative study groups that have contributed samples and expertise to many of the studies described in this review, including Joshua Stokes from Biomedical Communications at St. Jude. The authors are supported by a National Institutes of Health Outstanding Investigator Award, a St. Baldrick Communications at St. Jude. The authors are supported by a National Institutes of Health Outstanding Investigator Award, a St. Baldrick...  

REFERENCES  
B-Progenitor Acute Lymphoblastic Leukemia


Inaba T, Roberts WM, Shapiro LH, Jolly KW, Raimondi SC, Jain N, Roberts KG, Jabbour E, Patel K, Eterovic AK, Chen K, Hyrenius-Wittsten A, Pilheden M, Sturesson H, Hansson J, Cite this article as

K.G. Roberts and C.G. Mullighan


Cite this article as Cold Spring Harb Perspect Med doi: 10.1101/cshperspect.a034835

© 2021 by Cold Spring Harbor Laboratory Press. All rights reserved.


