Genomic profiling of B-progenitor acute lymphoblastic leukemia

Charles G. Mullighan, MD, PhD [Assistant Member]
Department of Pathology, St Jude Children’s Research Hospital, 262 Danny Thomas Place, Memphis, TN 38112

Abstract

Childhood acute lymphoblastic leukemia is comprised of multiple subtypes defined by recurring chromosomal alterations that are important events in leukemogenesis and are widely used in diagnosis and risk stratification, yet fail to fully explain the biology of this disease. In the last 5 years, genome-wide profiling of gene expression, structural DNA alterations and sequence variations has yielded important insights into the nature of submicroscopic genetic alterations that define novel subgroups of acute lymphoblastic leukemia (ALL) and that cooperate with known cytogenetic alterations in leukemogenesis. Importantly, several of these alterations are important determinants of risk of relapse and are potential targets for therapeutic intervention. Here, these advances and future directions in the genomic analysis of ALL are discussed.

Keywords

ALL; acute lymphoblastic leukemia; genomic profiling; B-progenitor; IKZF1; BCR-ABL1-like; PAX5; JAK1/2; CRLF2; CREBBP

Introduction

Acute lymphoblastic leukemia (ALL) is the commonest childhood malignancy, and despite cure rates now exceeding 80% [1], remains the leading cause of non-traumatic death in children and young adults [2]. Moreover, although ALL is less common than acute myeloid leukemia (AML) in the older adults, cure rates decline with increasing age. Consequently, there is intense interest in identifying genetic and biologic features that influence the pathogenesis of ALL and the risk of treatment failure [3]. In particular, the last 5 years have witnessed an explosion of activity in the use of high resolution genomic approaches to comprehensively identify all genetic alterations in leukemic cells [4–6].

Childhood ALL is characterized by recurring cytogenetic alterations including aneuploidy and recurring chromosomal translocations (Figure 1) [7, 8]. Hyperdiploidy with gain of at least 5 chromosomes is one of the most common types of ALL and is associated with excellent outcome; in contrast, hypodiploidy with less than 44 chromosomes is associated with a very poor prognosis [9–13]. Translocations in childhood ALL commonly disrupt
hematopoietic transcription factors or activate oncogenes and are key events in leukemogenesis. These include the (usually cryptic) t(12;21) ETV6-RUNX1 [14, 15], t(1;19) TCF3-PBX1, t(9;22) BCR-ABL1, and rearrangement of MLL. Similarly, T-lineage ALL is characterized by recurring rearrangements that commonly juxtapose regulatory elements of the T-cell antigen receptor loci with transcription factor or homeobox genes [16]. Recent findings from genomic analysis of T-lineage ALL are beyond the scope of this review and the reader is referred to recent review of the field [16–18]. These alterations are central events in leukemogenesis and identification of these rearrangements is vital in accurate risk stratification for therapy [3]. However, several observations indicate that these alterations are insufficient to explain the biologic basis of ALL and response to therapy. A substantial minority of children and a high proportion of adults lack one of these rearrangements [8]. Moreover, in experimental models, these alterations commonly do not alone induce leukemia, indicating that additional genetic or epigenetic alterations are required. Identification of these additional genetic alterations has been limited by conventional cytogenetic approaches, which typically can only detect gross rearrangements or structural alterations more than several megabases in size. Until recently, detection of additional genetic alterations has been reliant on candidate gene analysis, which has identified a small range of additional genetic changes, such as deletion of CDKN2A/CDKN2B (encoding INK4/ARF) and sequence mutation of NOTCH1 in T-ALL [19, 20].

**Genome-wide profiling of ALL**

The completion of the human genome project and the development of microarray technologies to profile gene expression and structural genetic alterations at high resolution have revolutionized our ability to identify genetic alterations in cancer genomes. These approaches have been applied to large cohorts of ALL cases by several groups [21–27] and have transformed our understanding of the genetic basis of this disease.

Microarray gene expression profiling has demonstrated that known recurring cytogenetic alterations in ALL are associated with distinct gene expression profiles [28] and also identify novel subgroups of ALL lacking known structural genetic alterations [29, 30]. These studies have provided important insights into the pathways dysregulated in ALL, but are unable to directly identify novel or cooperating genetic alterations contributing to leukemogenesis. From 2006, several groups have used microarray approaches to identify structural genetic alterations at high resolution. These include microarray comparative genomic hybridization (array-CGH) that compares DNA content of tumor and references normal samples to identify regions of DNA loss or gain; and single nucleotide polymorphism (SNP) arrays that genotype millions of common inherited SNPs across the genome [31]. While originally designed primarily as a genotyping tool to facilitate genome-wide association studies (GWAS) that identify inherited genetic variants that are associated with a disease phenotype, SNP arrays also permit inference of DNA copy number state, and by comparing tumor and normal SNP genotypes, can detect copy-neutral loss-of-heterozygosity (LOH) that may indicate duplication of a mutated tumor suppressor gene or oncogene [32].

Several groups have performed detailed evaluation of DNA copy number alteration and LOH in large cohorts of childhood ALL. Notably, at this time, detailed analyses of large cohorts of adolescent, young adult, and older adult ALL cases are lacking [33]. Nonetheless, these studies have utilized array-CGH and SNP microarray platforms that interrogate up to 2 million markers across the genome, with the ability to detect genetic alterations at sub 5 kilobase resolution. In a study of 242 childhood B-progenitor ALL cases, we used Affymetrix SNP microarrays and detected over 50 regions of recurring genetic alteration [21]. Many of these were not evident on cytogenetic analysis, were less than one megabase
in size, and commonly involved only a single or few genes, directly implicating alteration of these genes in leukemogenesis. Many of the targets of alteration were logical candidates in leukemogenesis, including tumor suppressor and cell cycle regulatory genes (CDKN2A/B, PTEN, RB1), transcription factors and transcriptional coactivators (ETV6, ERG, TBL1XR1), as well as genes involved in lymphoid maturation and signaling that have not previously been studied (eg, BTLA/CD200, TOX) and genes involved in drug responsiveness (eg, the glucocorticoid receptor NR3C1). A notable finding was that genes encoding transcriptional regulators of B-lymphoid development were mutated by deletion, translocation of sequence mutation in over two-thirds of B-progenitor ALL cases (Table 1) [21, 23]. For a comprehensive listing of recurring regions of DNA copy number alteration in childhood ALL, the reader is referred to the Supplementary Materials of refs [21, 34].

An important observation was that the nature and frequency of individual lesions varied significantly between B-ALL subtypes. Notably, MLL-rearranged leukemias harbor few additional genetic alterations, consistent with the notion that MLL-rearrangement may be sufficient to induce leukemia [35]. In contrast, ETV6-RUNX1 (TEL-AML1) and BCR-ABL1 rearranged leukemias harbor 6–8 additional copy number alterations per case and may present years after the acquisition of the founding translocation, indicating that these additional genetic alterations are required to induce the full leukemic phenotype.

Genetic alterations perturbing lymphoid development in B-progenitor ALL

A key observation is that over two-thirds of B-progenitor ALL cases have genetic alterations that disrupt the normal process of lymphoid maturation [21, 34]. The development of the B-lymphoid lineage from hematopoietic stem cells is controlled by multiple transcription factors that control B lymphoid lineage commitment, the repression of alternate lineage fates, and lymphoid maturation. Common targets of alteration are PAX5 (paired box 5), IKZF1 (IKAROS), EBF1 (early B cell factor 1) and LEF1 (lymphoid enhancer factor 1). These genetic alterations are commonly heterozygous and include focal and broad deletions, sequence mutations, and translocations. The most common alteration is PAX5 (30% of cases), which is seen in multiple subtypes of B-ALL (Figure 2). IKZF1 alteration is less common, is more commonly deletion than sequence mutation, and is a hallmark of multiple subtypes of high-risk ALL (see below) [22, 36, 37]. Most B-ALL cases exhibit a block in maturation at the pro- to pre-B cell stage of maturation, and these transcriptional regulators control differentiation from the HSC to these early stages of B-cell maturation, suggesting that these genetic alterations directly contribute to the block in maturation that is a hallmark of B-ALL. Consistent with this, emerging studies using knockout mouse models to examine the role of Pax5 and Ikaros loss in leukemogenesis have shown that loss of these alleles accelerates the onset of ALL [38–41].

Genetic alterations in high-risk ALL

A notable finding from these genomic analyses is that alterations of PAX5 are common in ALL but do not influence treatment outcome [34], whereas alterations of IKZF1 are less common but strongly associated with poor outcome in multiple distinct subtypes of high-risk ALL—BCR-ABL1-positive lymphoid leukemia [22, 36] and a novel subtype of “BCR-ABL1-like” lymphoid leukemia [34, 42].

BCR-ABL1, encoded by the Philadelphia chromosome, is a hallmark of chronic myeloid leukemia (CML) and acute lymphoblastic leukemia. These two diseases differ markedly in phenotype and responsiveness to therapy. Genomic profiling of both diseases, including sequential CML samples obtained at diagnosis, remission, and transition to accelerated phase and myeloid or lymphoid blast crisis demonstrated that loss-of-function or dominant negative alterations of IKZF1 (IKAROS) are a hallmark of BCR-ABL1 lymphoid leukemia,
including de novo ALL and CML at the transition to lymphoid blast crisis [22, 43]. Alteration of IKZF1 is also associated with poor outcome in BCR-ABL1 positive adult ALL [37]. Accordingly, loss of IKZF1 or expression of dominant negative IKAROS isoforms (particularly the IK6 isoform that lacks coding exons 3 through 6, and the N-terminal zinc fingers) results in accelerated onset of ALL in experimental models [41] and resistance to tyrosine kinase inhibitors, providing a plausible mechanism as to why the responsiveness to treatment of BCR-ABL1 ALL is considerably inferior to that of chronic phase ALL. Ongoing studies are using genome wide analysis of sequence variation to address unanswered questions in this disease, including the often striking differences in outcome between childhood and adult BCR-ABL1 positive ALL, genetic alterations that may drive high leukocyte counts (and poor outcome) and the nature of genetic alterations in IKZF1-wild type BCR-ABL1 ALL.

**IKZF1 and BCR-ABL1-like ALL**

Several studies have shown that alteration of IKZF1 is also associated with poor outcome in BCR-ABL1-negative childhood ALL [34, 44–46]. The association between IKZF1 alterations is striking, in some studies tripling the risk of relapse. Moreover, the association of IKZF1 status and outcome is commonly independent of commonly used risk stratification features such as age, sex, white cell count, and levels of minimal residual disease, suggesting that testing for IKZF1 status at the time of diagnosis is warranted. This is being explored in prospective trials, but it should be emphasized that the variety of genetic alteration involving IKZF1 requires the use of multiple, complementary genetic approaches, including tests that detect the (commonly focal) deletions, as well as the less common sequence variations.

These IKZF1-mutated, high-risk cases commonly lack known recurring chromosomal rearrangements, but have a gene expression profile that is remarkably similar to that of BCR-ABL1-positive ALL [34, 42]. This suggests that IKZF1 alteration may directly influence the leukemic transcriptome, or that these cases harbor additional genetic alterations that results in activated kinase signaling analogous to that induced by expression of BCR-ABL1. Both hypotheses are true. IKZF1-altered ALL cases exhibit enrichment for hematopoietic stem cell genes, consistent with the notion that loss of IKAROS activity results in maturational arrest. In addition, BCR-ABL1-like ALL cases commonly harbor novel genetic alterations dysregulating cytokine receptor and tyrosine kinase signaling.

Up to 50% of BCR-ABL1-like ALL cases harbor rearrangements that dysregulate expression of CRLF2 (encoding cytokine receptor like factor 2, or thymic stromal lymphopoietin; Figure 3) [47–49]. CRLF2 is located in the pseudoautosomal region 1 (PAR1) of Xp/Yp, and the alterations include juxtaposition of Xp/Yp into the immunoglobulin heavy chain locus at 14q32, resulting in IGH@-CRLF2 rearrangement [48], or a focal deletion of PAR1 that juxtaposes the regulatory elements of the purinergic receptor gene P2RY8 to CRLF2, and a chimeric P2RY8-CRLF2 fusion [47]. Less commonly, a missense mutation in exon 6 of CRLF2, F232C, results in constitutive CRLF2 dimerization [49]. All three events result in overexpression of full length CRLF2 on the cell surface of leukemic cells harboring the genetic alterations [47]. The PAR1 deletion is more common than the IGH@-CRLF2 rearrangement, although this is somewhat cohort dependent. CRLF2 rearrangement is particularly common in Down syndrome-associated ALL and is present in approximately 55% of these cases.

CRLF2 is known to have a role in T-lymphoid and dendritic cell development [50–52] and the pathogenesis of atopic disorders, but at present the role of CRLF2 in B-lymphoid neoplasms is relatively poorly understood. The downstream mediators of CRLF2 signaling
are incompletely characterized but are likely to involve the Janus kinase family, and up to 50% of CRLF2 rearranged ALL cases harbor concomitant activating mutations in JAK1 or JAK2 [47, 53, 54]. These are most commonly at or near R683 in the pseudokinase domain of JAK2 (although notably never at JAK2 V617 which is commonly mutated in myeloproliferative neoplasms, MPNs) [55], but are located throughout the kinase and pseudokinase domains of JAK1 and JAK2, most commonly at R683 in the pseudokinase domain of JAK2. The mutations in the pseudokinase domain of JAK2 are predicted to impair the negative regulatory effects of this domain on the catalytic activity of the kinase domain [56, 57]. The kinase domain mutations are predicted to impair interaction of the kinase domain with its substrates (including the pseudokinase domain) or are located in the catalytic site and are predicted to be directly activating [57]. In vitro studies indicate that, in contrast to the V617F mutations in MPNs that may alone transform hematopoietic cells, the JAK1/2 mutations observed in ALL usually require concomitant expression of CRLF2 in order to induce transformation, and moreover that mutant JAK2 and CRLF2 physically associate.

**Novel kinase-activating alterations in BCR-ABL-like ALL**

At least 50% of BCR-ABL1-like ALL cases lack known genetic alterations resulting in aberrant cytokine receptor and kinase signaling, including those involving CRLF2 and Janus kinases. Perhaps surprisingly, sequencing of the tyrosine kinase of BCR-ABL1-like ALL cases has failed to identify any novel kinase-activating sequence mutations apart from those involving the Janus kinases [58]. In contrast, transcriptome and whole genome sequencing of a small cohort of BCR-ABL1-like ALL cases has been highly revealing and has identified novel rearrangements, sequence mutations, and DNA copy number alterations activating tyrosine kinase signaling in the majority of BCR-ABL1-like cases [59]. In this approach, sequencing of cDNA libraries prepared from leukemic cells is performed using second generation sequencers (e.g., the Illumina GAIIx) and the paired reads from each cDNA molecule sequenced are mapped to the reference genome in order to identify chimeric cDNA sequences using a variety of bioinformatics algorithms [60, 61]. In a study of 12 BCR-ABL1-like cases, transcriptome sequencing identified identified rearrangements of ABL1, JAK2, PDGFRB and EPOR in 7 of 9 cases that lacked known rearrangements (the other three cases were known to have CRLF2 rearrangements) [59]. These rearrangements included the NUP214-ABL1 rearrangement previously identified in T-lineage ALL [62], BCR-JAK2, STRN3-JAK2, and EBF1-PDGFRB. Notably, several of these individual rearrangements and genes are also rearranged in myeloproliferative disorders [63–67], but have not previously been known to be rearranged in B-progenitor ALL. All BCR-ABL1-like B-ALL cases with these fusions harbor concomitant alterations of IKZF1, commonly the dominant negative IK6 deletion, suggesting that these two alterations drive proliferation of lymphoid cells as well as arresting lymphoid development, contributing to expression of an aggressive B-lymphoid malignancy. Several observations illustrate the complexity of the alterations that may result in aberrant kinase signaling. One case was predicted by RNA-seq to have a rearrangement of IGH@ to the erythropoietin receptor (EPOR) which has previously reported to be associated with a reciprocal t(14;19) translocation [68]. Fluorescence in situ hybridization for this rearrangement was negative, and careful mapping and manual inspection of RNA-seq read data indicated that this rearrangement was from a focal 7kb insertion of most of the EPOR locus into IGH@ adjacent to the promoter regions of this gene. An additional case lacked a putative fusion on RNA-seq analysis, but whole genome sequencing identified an activating mutation in the transmembrane domain of the interleukin-7 receptor alpha chain, which has recently been identified in both B- and T-ALL ([69] and Zhang et al, submitted). This case also harbored a concomitant focal inactivating deletion of SH2B3, encoding LNK, a negative regulator of JAK2 signaling [70], which is also mutated in JAK2 V617F-negative MPNs [71–73] and
whose inactivation accelerates the onset of myeloproliferative disease in murine models of MPN [74].

Each of these alterations is predicted to activate the tyrosine kinase domain of the rearranged gene or result in constitutive cytokine receptor signaling (eg, \textit{EPO}R and \textit{IL7R}). Existing experimental data support this notion. Murine cytokine-dependent hematopoietic Ba/F3 cells expressing \textit{NUP214-ABL1} or \textit{EBF1-PDGFRB} are transformed to factor-independent cell growth and exhibit activation of downstream signaling pathways, which is also evident on flow cytometric analysis of primary leukemic cells obtained from patients harboring these alterations. Moreover, this transformation and signaling is attenuated by low concentrations of currently available tyrosine kinase inhibitors such as imatinib and dasatinib, suggesting that these will be therapeutically useful in these patients.

Thus, collectively, these data have identified two subtypes of high-risk BCR-ABL-like ALL; one with CRLF2 rearrangement and JAK1/2 mutation, and a second with a diverse array of rearrangements, sequence and structural mutations constitutively activating kinase signaling. While CRLF2 rearrangements may be readily detected by genetic or flow cytometric techniques, detection of a diverse array of kinase gene rearrangements—not all of which have been currently identified and many of which are cytogenetically cryptic—poses major challenges. However, our existing data suggest leukemic cells harboring these alterations may be identified by focused gene expression profiling or flow cytometric analysis of primary leukemic cells (eg, for phosphorylated CRKL or tyrosine), suggesting that these may be useful diagnostic approaches. A challenge for the future is defining the nature and frequency of these alterations in adolescents and adults with ALL, for whom the outcome of therapy is often poor. These studies are ongoing.

Other novel subgroups of B-progenitor ALL

Cytogenetic analyses identify recurring chromosomal alterations in approximately 75% of childhood ALL; until the advent of microarray genome-wide profiling, the underlying genetic alterations of the remaining 25% were unknown. The majority of these cases may now be accounted for by the genetic alterations described above: CRLF2 rearrangement (5%–7% of cases), the heterogeneous alterations in the remaining BCR-ABL1-like cases (5%–7%), as well as several additional diagnostic groups, including intrachromosomal amplification of chromosome 21 and alterations of \textit{ERG} (see below). It is likely that several additional groups defined by recurring chromosomal alterations, submicroscopic structural alterations, and sequence variations will be identified, especially with ongoing studies of ALL in older individuals, in which the recurring alterations common in childhood ALL are uncommon or absent (eg, hyperdiploidy and \textit{ETV6-RUNX1}).

Intrachromosomal amplification of chromosome 21

Intrachromosomal amplification of chromosome 21 (iAMP21) defines up to 2% of B-ALL cases [75, 76]. The region of amplification is typically large but variable, however, but always includes gain of at least three copies of \textit{RUNXI}, commonly with accompanying deletion of the subtelomeric regions of chromosome 21. iAMP21 is characterized by commonly complex patterns of chromosomal rearrangement suggestive of breakage-fusion-bridge cycle [77, 78]. iAMP21 frequently occurs in older children or adolescents and is associated with a distinct gene expression profile [79] and poor outcome in the UK childhood ALL trials [80, 81].
**ERG deletions in ALL**

Focal deletions of ERG occur exclusively in cases lacking known chromosomal rearrangements and are a hallmark of a novel subtype of B-ALL with a distinct gene expression profile. The ERG deletions involve an internal subset of exons resulting in loss of the central inhibitory and pointed domains, and expression of an aberrant C-terminal ERG fragment that retains the ETS and transactivation domains, and functions as a competitive inhibitor of wild-type ERG [82]. These findings are in contrast to the rearrangements of ERG observed in carcinoma of the prostate, which result in in marked overexpression of the most of the ERG protein [83].

**Genetic alterations in relapsed ALL**

The studies described above have focused primarily on genetic alterations identified by profiling of the “bulk” leukemic sample obtained at relapse. There is now burgeoning evidence that many newly diagnosed ALL cases are in fact composites of multiple subclones that harbor distinct genetic alterations (including both copy number alterations and sequence mutations) that may influence treatment responsiveness and the risk of relapse.

Evolution or alterations in cytogenetic abnormalities from diagnosis to relapse in ALL has been recognized for many years [84], and several studies have performed limited candidate gene sequencing studies and shown that relapsed ALL samples may acquire sequence mutations (eg, NRAS, TP53) not present in the bulk diagnosis leukemic clone [85–87]. To look more broadly at the patterns of genetic alteration, several groups including our own performed genome-wide analysis of DNA copy number alterations in cohorts of matched diagnosis and relapsed ALL samples and demonstrated that in the majority of cases, the diagnosis and relapse samples share a common pre-diagnosis or “ancestral” clonal origin, but exhibit differences in the nature of genetic alterations [88–91]. In less than 10% of cases is a completely distinct leukemia identified bearing no similarity whatsoever to the diagnosis clone, although this frequency may be higher in certain ALL subgroups, such as late-relapsing T-ALL [92]. Similarly, relatively few samples are genetically identical at diagnosis and relapse. In contrast, approximately one third of cases show the acquisition of new genetic alterations in addition to those seen at diagnosis (ie, simple clonal evolution). In over half of cases, the pattern is more complex, with sharing of some lesions between diagnosis and relapse, loss of some lesions present at diagnosis, and the acquisition of new genetic alterations. In the majority of cases, “new” genetic alterations first detected at relapse may be detected at diagnosis using sensitive molecular methods, indicating that the relapse clones are present at low levels at diagnosis. Notably, many of the new lesions emerging at relapse are associated with aggressive disease or treatment failure, including IKZF1 and CDKN2A/B deletions.

Recent studies have coupled these genomic analyses with xenotransplantation studies of diagnosis and relapsed leukemia cells in mice, in which complementary genomic profiling of the xenograft samples is compared to primary leukemic samples. These studies have confirmed the observations made from the comparative genomic profiling of diagnosis and relapse samples. Specifically, diagnosis samples are comprised of multiple subclones that selectively engraft different recipient mice transplanted with the same primary tumor; that the clones that engraft mice may be genetically identical to the clones that emerge in patients at the time of relapse, and that specific genetic alterations (eg, deletion of CDKN2A/CDKN2B is associated with aggressive growth in mice) [93–95].

A logical extension of these studies is to perform genome-wide analysis of sequence variation in diagnosis and relapsed leukemia samples. To that end we performed extensive Sanger resequencing of 300 genes in a cohort of 23 matched diagnosis-relapse samples [96].
The findings from this study mirrored those of the SNP microarray analysis of DNA copy number alterations, namely that matched diagnosis and relapse samples share some mutations, but are dissimilar at other loci. A striking finding was that 19 percent of relapsed ALL samples harbor loss-of-function mutations (either deletions or sequence mutations) of CREBBP, encoding CREB-binding protein or CBP. CREBBP is a large protein with multiple roles including transcriptional coactivation and acetylation of histone and non-histone targets [97–99]. Rearrangements of CREBBP and the homologous gene EP300 (p300) are rare events in acute leukemia [100], and focal deletions of CREBBP are observed (uncommonly) in B-ALL, but sequence mutations of this gene in cancer have not previously been reported.

The sequence mutations of CREBBP observed in relapsed ALL included missense, nonsense, and frameshift mutations throughout the gene, with clustering of missense mutations in the histone acetyltransferase (HAT) domain. CREBBP is also mutated in the inherited developmental disorder Rubinstein-Taybi syndrome [101], and notably, a HAT domain mutation observed as a somatic event in relapsed ALL, Q1500P is also observed as an inherited mutation in Rubinstein-Taybi syndrome. Several lines of evidence indicate that the mutations in CREBBP are deleterious and are likely to influence treatment responsiveness in ALL. In contrast to other common targets of sequence mutation (eg, NRAS, KRAS, NF1), which when present at diagnosis may no longer be evident at relapse, the CREBBP mutations are (1) preserved at relapse when present at diagnosis; (2) acquired at relapse but present in minor subclones at diagnosis; or (3) reduplicated to homozygosity at relapse when heterozygous at diagnosis. Structural modeling of the HAT domain mutations (using the crystal structure of the highly homologous EP300 HAT domain) shows that the missense mutations are likely to interfere with binding to the acetylation substrates of CREBBP. Moreover, introduction of the mutations variable attenuates acetylation of histone H3 lysine 18 (H3K18, a critical CREBBP acetylation substrate) in Crebbp−/−Ep300−/− murine embryonic fibroblasts. Moreover, CREBBP is an important mediator of the transcriptional response to glucocorticoids, which are widely used in therapy of ALL, and the mutations directly impair the transcriptional response to glucocorticoids.

Finally, T-ALL cell lines harboring CREBBP mutations that are resistant to dexamethasone are highly sensitive to the histone deacetylase inhibitor vorinostat. Together, these data suggest that CREBBP alterations directly influence the response of leukemic cells to chemotherapy, and that modulation of CREBBP activity, such as with histone deacetylase inhibitors, may represent an important therapeutic approach in high-risk and relapsed ALL [102]. Several caveats, should, however be stated. While the in vitro evidence for a role of CREBBP mutations adversely impairing corticosteroid responses is compelling, these data are derived from non-leukemic cells (murine embryonic fibroblasts) and as CREBBP exerts pleiotropic effects, the exact manner in which CREBBP alterations influence leukemogenesis and responsiveness to therapy remains to be defined. Indeed, as the sequence mutation in CREBBP are located in different CREBBP domains (that interact with different transcription factors and partner genes), it is possible that different mutations have quite distinct effects. Moreover, contemporaneous work profiling genetic alterations in non-Hodgkin lymphoma identified inactivating mutations of CREBBP and EP300 in over one third of diffuse large-cell non-Hodgkin lymphoma (NHL) and follicular lymphoma in samples obtained at diagnosis [103, 104]. CREBBP may directly acetylate BCL6, which is of key importance in the pathogenesis of NHL, and thus it is likely that the alterations of CREBBP may exert distinct effects in lymphoma pathogenesis and responsiveness to treatment in ALL (in which mutations in these genes are rare at diagnosis, in contrast to NHL).
Future directions

The studies described above have provided clear evidence of the power of genome-wide profiling approaches to identify genes and pathways of central importance in establishment of the leukemic clone, and also in responsiveness to therapy. It is likely that much remains to be learned. These studies have primarily used microarray-based technology and candidate gene sequencing that provide only a sampling of the leukemic genome. Efforts are now well advanced to perform whole genome-sequencing of ALL, which will provide a comprehensive view of the landscape of genetic alterations in this disease, which has proven so informative in acute myeloid leukemia [6, 105, 106]. In addition to providing a catalog of structural and sequence variants affecting coding genes, these efforts will also allow detailed examination of non-coding genomic variation, which is likely important but at present relatively unexplored. It is also emphasized that many of the insights described above have only been possible with the detailed examination of large patient cohorts with comprehensive clinical data, and by the systematic integration of multiple modalities of genomic data including structural variants, sequence variation and gene expression. This will remain important with advances into the “whole-genome” age.

From the clinical perspective, the question remains as to how best integrate these findings into patient care. While there is great interest in performing microarray profiling and indeed, whole genome sequencing on patients at the time of diagnosis, a note of caution is warranted. We are still determining the functional and clinical relevance of many of the recurring lesions identified by these approaches, and indeed for next generation sequencing, still optimizing analytical approaches to sensitively and robustly identify all key genetic alterations [107]. Moreover, as the studies from transcriptomic sequencing of BCR-ABL1-like ALL attest, the genomic alterations driving otherwise homogenous subtypes of acute leukemia may be complex and polygenic and entail expensive, multimodal genomic analyses that are poorly suited to rapid clinical testing. Thus, the translation of findings from next generation sequencing to more conventional clinical tests such as flow cytometry, candidate gene sequencing, or focused gene expression profiling are likely to remain valuable approaches. Nonetheless, the genomic analysis of ALL is at a critical and tremendously exciting time, and the next few years are likely to lay bare the genetic aberrations of this disease and provide critical insights for future mechanistic analyses and clinical advances.

Acknowledgments

The author thanks colleagues at St Jude Children’s Research Hospital, the Children’s Oncology Group, the National Cancer Institute, and the Institute of Medical and Veterinary Science, Adelaide, Australia, who have contributed to these studies. The work described herein is supported by the American and Lebanese Syrian Associated Charities of St Jude Children’s Research Hospital, the Haematology Society of Australia and New Zealand, and grants from the National Cancer Institute and the TARGET (Therapeutically Applicable Research to Generate Effective Treatments) Initiative of the NCI. The author is a Pew Scholar in the Biomedical Sciences, and a St. Baldrick’s Scholar.

References


76. Soulier J, Trakhtenbrot L, Najfeld V, Lipton JM, Mathew S, Ajet-Loisau H, et al. Amplification of band q22 of chromosome 21, including AML1, in older children with acute lymphoblastic


Figure 1. Spectrum of recurring chromosomal rearrangements in childhood ALL
Representation of comment recurring numerical and structural genetic alterations in childhood B-progenitor and T-lineage ALL, including approximate frequencies. Alterations specific to T-lineage ALL are shown at the bottom of the pie chart in magenta. Data are adapted from [3, 113].
Figure 2. PAX5 alterations in ALL
A. 62 childhood ALL cases with deletions involving PAX5. Many deletions are focal and are not evident on inspection of chromosome-wide data. B. Same data at the PAX5 locus. Data taken from ref [21]. C. PAX5 sequence mutations in high risk B-progenitor ALL, showing missense (▼), frameshift (◆), and splice site (▶) mutations. PD, paired domain; O, octapeptide domain; H, homeodomain; TD, transactivating domain; A, activating; I, inhibitory. D. Structural modeling of the location of PAX5 paired domain mutations. The DNA double helix is blue, the PAX5 paired domain yellow, and ETS-1, which interacts with and increases the affinity of DNA binding of PAX5, is green. Each mutation disrupts normal interaction of PAX5 with DNA and/or ETS-1. Adapted from ref [34].
Figure 3. CRLF2 alterations in ALL
A. Focal deletions in the pseudoautosomal 1 region of Xp/Yp. SNP 6.0 log ratio copy number data for 6 Down syndrome ALL cases with matched tumor (T) and normal (N) data are shown. Light blue shows focal PAR1 deletion is present, the extent of which is identical in all cases. B. Mapping of the PAR1 deletion (probe level data are shown as vertical red lines) to between intron 1 of **P2RY8** and upstream of CRLF2. C. The PAR1 deletion results in a **P2RY8-CRLF2** fusion containing the entire CRLF2 open reading frame. D. The **P2RY8-CRLF2** fusion results in CRLF2 overexpression that may be detected by immunophenotyping of leukemic cells.
Table 1

Selected novel recurring genetic alterations in B-progenitor ALL

<table>
<thead>
<tr>
<th>Gene</th>
<th>Alteration</th>
<th>Frequency</th>
<th>Pathway and consequences of alteration</th>
<th>Clinical relevance</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX5</td>
<td>Focal deletions, translocations, sequence mutations</td>
<td>~32% of B-ALL</td>
<td>Transcription factor required for B-lymphoid development. Mutations impair DNA binding and transcriptional activation</td>
<td>No association with outcome</td>
<td>[21–23]</td>
</tr>
<tr>
<td>IKZF1</td>
<td>Focal deletions or sequence mutations</td>
<td>15% of childhood B-ALL cases. Rare in T-ALL</td>
<td>Transcription factor required for development of HSC to lymphoid precursor. Deletions and mutations result in loss of function or dominant negative isoforms</td>
<td>Over 80% BCR-ABL1 ALL and 66% CML in lymphoid blast crisis</td>
<td>[21]</td>
</tr>
<tr>
<td></td>
<td>Inherited variants</td>
<td>Over 80% BCR-ABL1 ALL and 66% CML in lymphoid blast crisis</td>
<td></td>
<td>One-third of high-risk BCR-ABL1 negative ALL</td>
<td>[22, 36, 37]</td>
</tr>
<tr>
<td></td>
<td>Inherited variants</td>
<td>Over 80% BCR-ABL1 ALL and 66% CML in lymphoid blast crisis</td>
<td></td>
<td></td>
<td>[34, 42, 45]</td>
</tr>
<tr>
<td>JAK1/2</td>
<td>Pseudokinase and kinase domain mutations</td>
<td>18%–35% DS-ALL and 10.7% high-risk BCR-ABL1-ALL</td>
<td>Constitutive JAK-STAT activation. Transforms mouse Ba/F3-EpoR lymphoid hematopoietic cell line</td>
<td>Associated with CRLF2 alteration, IKZF1 deletion/sequence mutation, and poor outcome</td>
<td>[56, 57, 110, 111]</td>
</tr>
<tr>
<td>CRLF2</td>
<td>Rearrangement as IGH@-CRLF2 or P2RY8-CRLF2; or F232C mutation resulting in overexpression</td>
<td>5%–16% pediatric and adult B-ALL, and &gt;50% DS-ALL</td>
<td>Associated with mutant JAK in up to 50% of cases. CRLF2 mutations and JAK mutations cotransforming in Ba/F3 cells and results in constitutive STAT activation</td>
<td>Associated with poor outcome</td>
<td>[47, 49, 53, 54]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14% pediatric high-risk ALL</td>
<td></td>
<td></td>
<td>[48, 112]</td>
</tr>
<tr>
<td>CREBBP</td>
<td>Focal deletion and sequence mutations</td>
<td>19% of relapsed ALL. 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>[96, 103]</td>
</tr>
</tbody>
</table>

ALL, acute lymphoblastic leukemia; CIR, cranial irradiation; CML, chronic myeloid leukemia; DS-ALL, Down syndrome ALL; HSC, hematopoietic stem cell;