Genome sequencing of lymphoid malignancies

Charles G. Mullighan
Our understanding of the pathogenesis of lymphoid malignancies has been transformed by next-generation sequencing. The studies in this review have used whole-genome, exome, and transcriptome sequencing to identify recurring structural genetic alterations and sequence mutations that target key cellular pathways in acute lymphoblastic leukemia (ALL) and the lymphomas. Although each tumor type is characterized by a unique genomic landscape, several cellular pathways are mutated in multiple tumor types—transcriptional regulation of differentiation, antigen receptor signaling, tyrosine kinase and Ras signaling, and epigenetic modifications—and individual genes are mutated in multiple tumors, notably TCF3, NOTCH1, MYD88, and BRAF. In addition to providing fundamental insights into tumorigenesis, these studies have also identified potential new markers for diagnosis, risk stratification, and therapeutic intervention. Several genetic alterations are intuitively “druggable” with existing agents, for example, kinase-activating lesions in high-risk B-cell ALL, NOTCH1 in both leukemia and lymphoma, and BRAF in hairy cell leukemia. Future sequencing efforts are required to comprehensively define the genetic basis of all lymphoid malignancies, examine the relative roles of germline and somatic variation, dissect the genetic basis of clonal heterogeneity, and chart a course for clinical sequencing and translation to improved therapeutic outcomes. (Blood. 2013;122(24):3899-3907)

Introduction

The lymphoid malignancies represent a diverse range of tumors characterized by variable stages of maturation ranging from pro-B or T cells in acute lymphoblastic leukemia (ALL) to cells representative of the lymph node in the non-Hodgkin lymphomas (NHLs) to mature plasma cells in myeloma and related disorders. These disorders have a diverse range of clinical manifestations, sites of organ involvement, and responsiveness to therapy. Somatic genetic alterations are a hallmark of lymphoid malignancies, and it is now known that each tumor type typically harbors a constellation of genetic alterations, including gross chromosomal alterations (aneuploidy and/or chromosomal rearrangements), structural genetic alterations (deletions and gains of DNA), and sequence mutations. Genomic techniques that have identified these alterations may be grouped into three broad categories: (1) cytogenetic studies that have identified structural chromosomal alterations, coupled with limited candidate gene sequencing or structural profiling studies; (2) first-generation genome-wide profiling using array-based comparative genomic hybridization or single nucleotide polymorphism microarrays and gene expression profiling; and (3) second-generation sequencing studies using whole-genome sequencing (WGS), transcriptome sequencing (mRNA-seq) and/or whole-exome sequencing (WES) to more comprehensively define the genomic landscape of these diseases. For each tumor type discussed below, key genetic alterations identified from cytogenetic and first-generation genomic profiling studies are presented to provide context for the results of next-generation sequencing (NGS) studies summarized in Table 1.
is highly dependent on tumor type. Factors such as exposure to environmental mutagens and mutations in *TP53* and other genes encoding mediators of DNA integrity are clearly important but do not explain the striking variation in the type and frequency of mutations between different childhood and adult tumors. Infant leukemia harbors few alterations apart from *MLL* rearrangement, but other early-onset leukemias harbor more mutations (eg, *ETV6*-RUNX1 and *BCR-ABL1*), and other childhood solid and brain tumors often have a remarkably high mutation frequency. More than 50 recurring regions of DNA copy number alteration have been identified, which are commonly focal deletions limited to one or few genes. These include transcriptional regulators of lymphoid development (*PAX5*, *IKZF1*, *EBF1*, *LEF1*), tumor suppressors (*CDKN2A, CDKN2B, RB1, TP53*), lymphoid signaling genes (*BTLA, CD200, TOX*), transcriptional regulators and coactivators (*TBL1XR1, ERG*), and regulators of chromatin structure and epigenetic regulators (*CTCF, CREBBP*). Several of these alterations are associated with high-risk ALL and an increased risk of treatment failure and relapse, notably deletions of *IKZF1* (IKAROS) and deletion or sequence mutation of *CREBBP*. Sanger sequencing studies have also identified recurring sequence mutations. In B-lineage ALL, these most commonly affect lymphoid development (*PAX5*), Ras signaling (*NRAS, KRAS, and NF1*), cytokine receptor signaling (*IL7R, JAK2*), and tumor suppression (*TP53*). Similarly, a number of targets of structural genetic alteration and/or sequence mutation have been identified in T-lineage ALL, including activating mutations of *NOTCH1*, deletion and/or mutation of *PTEN*, *WT1*, *FBXW7*, and amplification of *MYB*. The most extensive NGS studies in ALL include the St. Jude Children’s Research Hospital–Washington University Pediatric Cancer Genome Project (PCGP) and the Children’s Oncology Group–National Cancer Institute Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative (http://oacrg.cancer.gov/programs/target). Sequencing of the full spectrum of ALL subtypes is incomplete; however, the results of NGS analysis of several high-risk ALL subtypes have recently been reported. These data have shown that alteration of multiple cellular pathways, including cytokine receptor and Ras signaling, tumor suppression, lymphoid development, and epigenetic regulation, are hallmarks of multiple ALL subtypes.

### T-lineage ALL

T-ALL is characterized by an older age of onset than that for B-ALL, male sex preponderance, and inferior outcome in comparison with B-ALL. In the first NGS study of ALL, Ferrando and colleagues examined the male sex preponderance by performing targeted capture and NGS of X chromosome genes. This identified sequence mutations and, less commonly, deletion of *PHF6* in 16% and 38% of childhood and adult T-ALL cases, respectively. The *PHF6* alterations result in loss of *PHF6* expression and are associated with *TLX1*- and *TAL1*-rearranged ALL. The role of *PHF6* in leukemogenesis is poorly understood, but the loss-of-function alterations suggest that *PHF6* is a tumor suppressor.

Early T-cell precursor (ETP) ALL is an aggressive subtype of immature leukemia that accounts for a high proportion of T-ALL treatment failures. Various laboratory criteria have been proposed to identify these immature cases, but the original definition proposed by Campana and colleagues is based on immunophenotype: leukemic cells that express T-lineage markers (eg, cytoplasmic CD3) but lack expression of CD1a and CD8, have weak or negative CD5 expression, and exhibit aberrant expression of myeloid and/or stem cell markers. This pattern is reminiscent of the murine ETP, the earliest stage of thymic T-cell maturation that retains lineage plasticity.

In the first report of WGS of a lymphoid malignancy, the PCGP performed WGS of tumor and matched nontumor DNA of 12 ETP ALL cases, WES and mRNA-seq in selected cases, and mutation recurrence testing of selected genes in 94 additional ETP and non-ETP T-ALL cases. Unexpectedly, no common structural rearrangement or sequence mutation was identified. There was marked diversity in the frequency and nature of genetic alterations, with several cases exhibiting complex multichromosomal structural alterations with the hallmarks of chromothripsis, and other cases with no structural alterations whatsoever. However, the majority of cases harbored alterations in three pathways: loss-of-function mutations in genes encoding regulators of hematopoietic development (*ETV6*,

### Table 1. Key genetic alterations identified by NGS studies in lymphoid neoplasms

<table>
<thead>
<tr>
<th>Tumor type</th>
<th>Genetic alteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-ALL</td>
<td><em>PHF6</em>, <em>CNOT3</em>, <em>RPL5</em>, <em>RPL10</em></td>
</tr>
<tr>
<td>ETP ALL</td>
<td>Loss-of-function mutations in hematopoietic regulators (GATA3, <em>IKZF1, RUNX1, ETV6</em>)</td>
</tr>
<tr>
<td></td>
<td>Gain-of-function mutations in Ras, <em>FLT3, IL7R</em></td>
</tr>
<tr>
<td></td>
<td>Inactivating mutations in epigenetic regulators (EZH2, <em>SUZ12, EED, SETD2, DNMT3A</em>)</td>
</tr>
<tr>
<td>BCR-ABL1–like ALL</td>
<td>Rearrangement of <em>CRLF2</em> in 50% of cases; concomitant activating JAK mutations in 50% of <em>CRLF2</em>-rearranged cases</td>
</tr>
<tr>
<td>Hypodiploid ALL</td>
<td>Ras mutations (<em>NF1, PTPN11, NRAS, KRAS</em>) in near-haploid ALL</td>
</tr>
<tr>
<td></td>
<td><em>IKZF2</em> and <em>TP53</em> mutations in low-hypodiploid ALL; <em>TP53</em> mutations are commonly germline,</td>
</tr>
<tr>
<td>Relapsed ALL</td>
<td><em>CREBBP</em> mutations enriched at relapse</td>
</tr>
<tr>
<td>Familial ALL</td>
<td><em>NTSC2</em> mutations enriched at relapse</td>
</tr>
<tr>
<td>DLBCL and NHL</td>
<td><em>Lympoid signaling (CD79B)</em></td>
</tr>
<tr>
<td></td>
<td><em>NF-κB signaling (CARD11, MYD88)</em></td>
</tr>
<tr>
<td>SMZL</td>
<td><em>Histone modification (CREBBP/EP300, EZH2, MEF2B, MLL2/MLL3)</em></td>
</tr>
<tr>
<td>MCL</td>
<td><em>NOTCH1</em> mutations, associated with poor outcome</td>
</tr>
<tr>
<td>HL</td>
<td><em>CIITA</em> rearrangements</td>
</tr>
<tr>
<td>PMBCL</td>
<td><em>CIITA</em> rearrangements</td>
</tr>
<tr>
<td>BL</td>
<td><em>TCF3/I2D3</em> mutations in BL and other</td>
</tr>
<tr>
<td></td>
<td><em>MYC</em>-rearranged lymphomas</td>
</tr>
<tr>
<td></td>
<td><em>CCND</em> mutations</td>
</tr>
<tr>
<td>CLL</td>
<td><em>NOTCH1</em> mutations; associated with disease progression</td>
</tr>
<tr>
<td></td>
<td>mRNA splicing mutations (eg, <em>SF3B1</em></td>
</tr>
<tr>
<td></td>
<td>DNA damage and/or repair mutations (ATM, <em>POT1</em>)</td>
</tr>
<tr>
<td></td>
<td>Regulation of apoptosis (BIRC3)</td>
</tr>
<tr>
<td></td>
<td>Innate immunity (<em>MYD88, TLR2</em>)</td>
</tr>
<tr>
<td>HCL</td>
<td>Activating <em>BRAF</em> mutations</td>
</tr>
<tr>
<td>MM</td>
<td>Multiple targets of mutation, including <em>NRAS, KRAS, TP53, CCND1, DIS3, BRAF, NF-κB signaling</em> and histone modification</td>
</tr>
<tr>
<td>WM</td>
<td><em>MYD88</em> p.Leu265Pro in &gt;90% of cases</td>
</tr>
</tbody>
</table>

*PMBCL, primary mediastinal B-cell lymphoma.*
GATA3, IKZF1, RUNX1; see also Della Gatta et al21 and Van Vlierberghe et al22; activating mutations in cytokine receptor and Ras signaling (NRAS, KRAS, FLT3, JAK1, JAK3, and IL7R; see also Flex et al,25 Shochat et al,26 Zenatti et al,27 and Neumann et al28); and inactivating mutations targeting epigenetic regulators, most commonly components of the polycomb repressor complex 2 that mediates histone 3 lysine 27 (H3K27) trimethylation (EZH2, SUZ12, EED; see also Nizziachristos et al29), the histone 3 lysine 36 trimethylase gene SETD2, and the histone acetyltransferase and CREBBP homolog EP300 (p300). Subsequent studies have also identified recurring mutations in the DNA methyltransferase gene DNMT3A in adult ETP ALL.30 Additional new targets of mutation were identified, including DNMT2, ECT2L, and RELN, and several genes—and indeed specific somatic mutations—were identified that had previously been reported as germline mutations in inherited developmental disorders, notably those in the zinc finger domain of the hematopoietic transcription factor gene GATA3. Several genes, including RUNXI, ETV6, and IKZF1, were targeted by multiple types of genomic aberration, including sequence mutation, deletion, and translocation, indicating the need for WGS to completely define the spectrum of somatic genomic alterations in hematologic neoplasms. The mutational spectrum of ETP ALL is similar to that spectrum of somatic genomic alterations in hematologic neoplasms, indicating the need for WGS to completely define the repertoire of kinase-activating alterations in BCR-ABL1–like ALL and to develop clinical trials to direct patients with BCR-ABL1–like ALL to appropriate TKI therapy.

Transcriptome sequencing and WGS of 15 BCR-ABL1–like ALL cases, 12 of which lacked CRLF2 rearrangement, identified a strikingly diverse array of genetic alterations activating cytokine receptor and tyrosine signaling.32 These were the most commonly rearrangements resulting in chimeric fusion genes deregulating tyrosine kinases (NUP214-ABL1, ETV6-ABL1, RANBP2-ABL1, RCSD1-ABL1, BCR-JAK2, PAX5-JAK2, STRN3-JAK2, and EBF1-PDGFBR) and cytokine receptors (IGH-EPOR). Up to 20% of BCR-ABL1–like cases lack a chimeric fusion on mRNA-seq analysis, and sequence mutations (eg, activating mutations of FLT3 and IL7R) and structural alterations (eg, focal deletions of SH2B3, or LNK, which constrains JAK signaling43) that activate signaling have been identified in fusion-negative cases. These diverse genetic alterations activate a limited number of signaling pathways, notably ABL1 and PDGFBR (both of which may be inhibited with the TKIs imatinib and dasatinib) and JAK-STAT signaling. These rearrangements have been shown to activate signaling pathways in model cell lines and in primary leukemic cells, and xenografts of BCR-ABL1–like ALL are highly sensitive to TKIs in vivo. Moreover, a recent report of a child with refractory EBF1-PDGFBR–positive ALL that was exquisitely sensitive to imatinib emphasizes the potential clinical utility of TKI therapy in BCR-ABL1–like ALL.44 Ongoing studies are performing NGS of childhood and adult ALL to comprehensively define the repertoire of kinase-activating alterations in BCR-ABL1–like ALL and to develop clinical trials to direct patients with BCR-ABL1–like ALL to appropriate TKI therapy.

**BCR-ABL1-like ALL**

In 2009, two study groups identified a group of childhood B-progenitor ALLs lacking a recurring chromosomal alteration that exhibited a gene expression profile similar to BCR-ABL1–positive ALL, deletion of IKZF1 (also common in BCR-ABL1–positive ALL), and poor outcome.33,34 The similarity of the gene expression profile to BCR-ABL1 positive ALL in these BCR-ABL1–like or Philadelphia-like ALL cases suggested the presence of previously unrecognized genetic alterations that activate tyrosine kinase signaling pathways. BCR-ABL1–like ALL is common, comprising up to 10% to 15% of childhood B-ALL and a higher frequency of ALL in adolescents and young adults (unpublished data), and it is associated with poor outcome.35 Up to half of BCR-ABL1–like ALL cases harbor rearrangement of CRLF2 located at the pseudoautosomal region of Xp22.3/Yp11.3, either as a translocation to the immunoglobulin heavy chain enhancer region at 14q32.33 (IGH-CRLF2), or a focal deletion proximal to CRLF2 resulting in the expression of a P2RY8-CRLF2 fusion transcript.36–38 Both alterations result in overexpression of CRLF2 on the surface of lymphoid blasts that may be detected by immunophenotyping.36,38 Approximately half of CRLF2 rearranged cases harbor concomitant activating mutations of the Janus kinase genes JAK1 and JAK2, most commonly at p.Arg683 in the pseudokinase domain of JAK2 and, less commonly, in the kinase domains of JAK1 and JAK2. The JAK2 p.Val617Phe mutation commonly observed in myeloproliferative neoplasms39 does not occur in ALL, although the corresponding mutation in the pseudokinase domain in JAK1 (p.Val678Phe) is observed.11,40 These alterations result in activation of JAK-STAT signaling that may be amenable to therapy with JAK inhibitors such as ruxolitinib,41 and this is currently being explored as a therapeutic strategy.

**Hypodiploid ALL**

Hypodiploidy with less than 44 chromosomes is observed in up to 3% of ALL cases and is associated with poor prognosis. Two subtypes of hypodiploid ALL have been described according to the severity of aneuploidy: near-haploid cases with 24 to 31 chromosomes and low-hypodiploid cases with 32 to 39 chromosomes.45–47 However, the nature of additional genetic alterations driving leukemogenesis and poor outcome in hypodiploid ALL is unknown. Microarray and NGS analysis of a large cohort of more than 120 hypodiploid ALL samples has demonstrated that near-haploid and low-hypodiploid ALL have distinct transcriptomic signatures and submicroscopic genetic alterations.48 The majority of near-haploid cases harbor mutations activating Ras signaling, most commonly in NF1, and the IKAROS family gene IKZF3 (AIOLOS). In contrast, low-hypodiploid cases have near universal mutation of the tumor suppressor TP53 (p53), with the mutations present in the germline in

---

From bloodjournal.hematologylibrary.org by RAUL RIBEIRO on December 30, 2013. For personal use only.
approximately half the cases and inactivating mutations of a third IKAROS family member IKZF2 (HELIOS). Parallel analysis of primary hypodiploid xenografts (primagrafts) demonstrated activation of Ras-Raf-MEK-ERK and phosphatidylinositol-3-OH kinase (PI3K) signaling in the majority of hypodiploid primagrafts that was sensitive to PI3K and PI3K/mTOR inhibitors, but not MEK inhibitors, suggesting that PI3K inhibition represents a novel therapeutic approach. In addition, the identification of low-hypodiploid ALL as a manifestation of Li-Fraumeni syndrome indicates that all children with low-hypodiploid ALL should be tested for TP53 mutational status.49,50 Recent NGS studies have also identified other germline mutations in familial leukemia, such as a germline PAX5 p.Gly183Ser mutation in two kindreds with autosomal dominant pre-B ALL;51 thus, there is considerable interest in defining the full range of inherited genetic variants that influence susceptibility to ALL.

Relapsed ALL

Although several subtypes of ALL are associated with a high risk of treatment failure, relapse occurs across the spectrum of ALL subtypes and is associated with very poor outcome. Moreover, it has long been recognized that ALL genomes are not static but exhibit acquisition of chromosomal abnormalities over time.52 Single nucleotide polymorphism microarray profiling studies of matched diagnosis-relapse ALL samples demonstrated that the majority of ALL cases show changes in the patterns of structural genomic alterations from diagnosis to relapse5 and that many relapse-acquired lesions, including those targeting genes associated with high-risk ALL (IKZF1, IKZF2, CDKN2A, and CDKN2B), are present at low levels at diagnosis, suggesting that genetically determined tumor heterogeneity is a key determinant of treatment failure and relapse.53,54 Sequencing of 300 genes in matched diagnosis-relapse samples recapitulated these findings and identified mutations in the transcriptional coactivator and histone acetyltransferase CREBBP (CREB-binding protein, or CBP) as a relapse-acquired lesion in up to 20% of relapsed ALL samples.55,56 CREBBP mutations are also observed at diagnosis in NHL, particularly diffuse large B-cell lymphoma (DLBCL),57 and they impair histone acetylation.56,57 CREBBP in part mediates the transcriptional response to glucocorticoids, and histone deacetylase inhibitors were effective in killing steroid-resistant ALL cell lines. Two study groups recently identified relapse-acquired mutations in NT5C2 that encode a 5′-nucleotidase enzyme responsible for the inactivation of nucleoside-analog drugs.57,58 Thus, mutations that confer resistance to drugs commonly used to treat ALL represent a key mechanism of treatment failure and resistance.

NHL

Genomic profiling studies, including NGS, have shed important light on the genetic basis and molecular heterogeneity of NHL, which comprises a diverse range of entities with distinct pathologic and genetic features, including DLBCL, follicular lymphoma, mantle cell lymphoma (MCL), splenic marginal zone lymphoma (SMZL), chronic lymphocytic leukemia (CLL), and hairy cell leukemia (HCL).59 As with ALL, there is an extensive literature describing microarray-based profiling of transcriptional signatures, structural genetic alterations, and candidate gene sequencing in NHL. These studies have identified important genetic alterations that are beyond the scope of this review,60 but key findings from NGS studies are reviewed below.

DLBCL

DLBCL is the most common form of NHL. Gene expression profiling studies identified at least three subtypes of DLBCL with distinct transcriptional profiles reflective of their cell of origin: activated B-cell, germinal center B-cell, and primary mediastinal B-cell lymphoma, which is thought to be of thymic B-cell origin. Multiple studies have used candidate gene, WGS, WES, and mRNA-seq to identify recurring genetic alterations and rearrangements in DLBCL and to examine their association with clinical features, transcriptional subtype, and outcome.56,61-68 Notably, mutations affecting B-cell receptor signaling (CD79B) and activating nuclear factor kB (NF-kB) signaling (CARD11 and MYD88) are enriched in activated B-cell DLBCL, and genes encoding histone modifiers (CREBBP, EP300, EZH2, MEF2B, MLL2/3) are more commonly mutated in the germinal center B-cell subtype. The high frequency of mutations in histone-modifying genes reinforces the central role of these alterations in the pathogenesis of a range of lymphoid malignancies but also highlights distinct roles of different alterations in different tumors. For example, the mutations in the histone acetyltransferase and transactivator CREBBP are similar to those observed in ALL but, in contrast to ALL, they are observed in the predominant clone at diagnosis and may act by impairing acetylation of TP53 and BCL6. Similarly, although EZH2 is mutated in both NHL and ALL, the type of EZH2 alteration and predicted functional consequences are different. In DLBCL and follicular lymphoma, the most common alterations are a missense mutation at p.Tyr641, which is a gain-of-function mutation that enhances histone 3 lysine 27 trimethylation and transcriptional repression.69,70 In contrast, the EZH2 alterations in ETP ALL are loss-of-function deletions and sequence mutations that are predicted to have the opposite effect.

These findings have been extended by a recent NGS study that performed WGS of 40 DLBCL samples and 13 cell lines, coupled with transcriptome sequencing and DNA copy number alterations analysis of 96 samples.71 This study expanded the number of significantly mutated genes to 74, more than 40 of which had not been previously reported, including cell surface receptor genes (CD70, CD83), purinergic receptor genes (P2RY8, P2RX5) and G-protein–coupled receptors. In contrast to precursor B-cell neoplasms, recurring chromosomal rearrangements leading to the expression of chimeric fusions are uncommon in DLBCL, with the exception of those involving TP63.72

SMZL and MCL

NGS studies have also identified recurring targets of mutation associated with other subtypes of NHL, including NOTCH2 mutations in approximately 25% of SMZLs.73 Transcriptome sequencing of MCL has identified NOTCH1 mutations in 12% of cases, most commonly in the C-terminal domain in the recognition sequence for FBXW7.74 This pattern is similar to that observed in CLL, and the NOTCH1 mutations are associated with poor outcome in MCL. Additional recurring targets of mutation in MCL include ATM, CCND1, TP53, and the E3 ligase UBR5.75
Hodgkin lymphoma

Prior to the advent of NGS, few structural genetic alterations had been described in Hodgkin lymphoma (HL). A remarkable finding of mRNA-seq of HL cell lines and tumor samples was a high frequency of rearrangement of CIITA to a range of fusion partners including the uncharacterized gene BX648577, CD274, CD273, and RALGDS, commonly resulting in expression of chimeric fusion transcripts in HL (15% of cases) and the phenotypically related disorder primary mediastinal B-cell lymphoma (38%), whereas the fusions were rare in DLBCL. CIITA is an important transactivator of class II major histocompatibility complex expression. The role of the fusions in lymphomagenesis is incompletely understood, but modulation of HLA expression and escape from immunosurveillance has been postulated.

Burkitt lymphoma

Burkitt lymphoma (BL) is a highly aggressive B-lymphoid malignancy that, like DLBCL, is thought to arise from the germinal center. Deregulation of MYC by translocation into the immunoglobulin loci enhancer regions, most frequently, the IGH, is a hallmark of BL. BL is observed in multiple contexts, including sporadic BL in developed countries, Epstein-Barr virus–associated BL typically observed in the developing world, and HIV-associated BL. Three study groups recently described the first NGS studies of BL and used a combination of WGS, WES, and/or mRNA-seq and, in one study, high-throughput RNA interference. These studies identified at least 70 recurrently mutated genes in BL, the most common of which were alterations affecting TCF3 activity in 70% of sporadic BL cases, either due to mutations of TCF3 or, more commonly, loss-of-function mutations or deletions of ID3, which encodes a negative regulator of TCF3. ID3 alterations were observed in 35% to 68% of BL, 13% of other MYC-rearranged lymphoma, but not DLBCL. These alterations activated B-cell receptor and PI3K signaling. In addition, 38% of sporadic BLs harbored oncogenic mutations of the CCND3, encoding cyclin D3, which promotes cell cycle progression. These remarkable findings highlight two new avenues for potential therapeutic exploitation in this highly aggressive malignancy and suggest that further sequencing efforts to examine the genetic differences between the different forms of BL will be of considerable interest.

CLL

CLL is the most common leukemia, and it is characterized by monoclonal proliferation of CD5+ B cells and variable lymphadenopathy, splenomegaly, cytopenias, and autoimmune phenomena. The clinical course of CLL is variable, and genetic features previously associated with prognosis include the extent of somatic hypermutation of immunoglobulin genes and cytogenetic alterations, including deletions of 11q23, 17p13, and 13q14. The presence of or lack of hypermutation of Ig genes may reflect the cell of origin of CLL—naïve B cells in cases lacking hypermutation and germinal center B cells in cases with somatic hypermutation. Several recent studies have reported WGS and WES of CLL, which have provided insights into the mutational landscape and clonal heterogeneity of this disease. Compared with other lymphoid malignancies, few genes are mutated in more than 10% of cases. In contrast, multiple genes are mutated at lower frequency. In approximately one third of cases, no recurring targets of mutation have been identified. Mutations are enriched in several pathways, which are likely driver events, including NOTCH1 signaling (NOTCH1 and the U3 ubiquitin ligase gene FBW7); mRNA splicing, transport, and processing (DDX3X, SF3B1, SFRS1, U2AF2, and XP01); DNA damage, repair, and telomere maintenance (ATM, POT1, and TP53); regulation of apoptosis (BIRC3); and innate immunity and inflammation (MAPK1, MYD88, and TLR2). Recurrent structural alterations leading to chimeric fusion genes are uncommon in CLL. The frequency of mutations has varied across studies. This may in part reflect the nature of the cohorts studied (early stage vs advanced disease), the enrichment for cases with unmutated or mutated immunoglobulin genes, and possibly the sequencing approach used, but the most common targets of mutation are NOTCH1, SF3B1, TP53, and MYD88. The identification of mutations targeting the spliceosome machinery is one of the most notable novel findings from these studies. These mutations may result in aberrant splicing, intron retention, and expression of aberrant protein isoforms that may potentially deregulate a diverse range of cellular pathways. Several of these alterations, notably NOTCH1 mutations, are associated with poor survival. Mutational profiling has been integrated with cytogenetic analysis to refine outcome prediction in CLL, with TP53 and BIRC3 associated with poor survival; NOTCH1, SF3B1, and/or del(11q22) with intermediate risk; trisomy 12 or normal karyotype with low risk; and del (13q14) with very low risk. Mutations in SF3B1 and NOTCH1 are commonly identified as subclonal events early in the disease course and may drive the emergence of resistant clones and disease progression.

HCL

One of the most remarkable findings from recent NGS studies of lymphoid neoplasms is the identification of near universal BRAF mutations in HCL. HCL is characterized by splenomegaly, cytopenias, and the presence of malignant cells with distinctive villous projections (ie, hairy cells) in the marrow and spleen. Although it is often indolent and amenable to nucleoside analog therapy, the molecular pathogenesis has remained elusive. WES of an HCL genome identified a somatic BRAF p.Val600Glu mutation that was also identified as a universal phenomenon in a recurrence cohort of 47 cases. BRAF p.Val600Glu is observed in more than half of the malignant melanomas, papillary carcinoma of the thyroid, and Langerhans cell histiocytosis, and infrequently in myeloma, but it is uncommon in other lymphoid neoplasms. The presence of this mutation in nearly all cases of HCL has been confirmed by other studies. The p.Val600Glu mutation lies in the activation loop of the kinase domain of BRAF and explains the activation of the Raf-MEK-ERK pathway characteristic of HCL. Detection of BRAF p.Val600Glu has rapidly entered the clinical arena, as a diagnostic test for HCL and to monitor response to therapy. Moreover, reports of responsiveness of treatment-refractory HCL to the mutant BRAF inhibitor vemurafenib have spurred trials testing BRAF inhibitors in HCL. Although comprehensive definition of the genomic landscape of HCL by WGS is awaited, these findings indicate that mutant BRAF is a key driver and rational therapeutic target in this disease.
Multiple myeloma, the plasma cell dyscrasias, and related disorders

Multiple myeloma (MM) is characterized by proliferation of malignant plasma cells in the bone marrow, bony destruction, and end organ dysfunction, either directly from plasma cell infiltration or from the effects of the secreted monoclonal Ig (paraprotein). MM is one disorder in a spectrum of plasma cell dyscrasias that include monoclonal gammapathy of uncertain significance (MGUS), smoldering MM, MM and plasma cell leukemia. Chromosomal alterations include hyperdiploidy and translocations, commonly re-arrangements of CCDKN1, CCND3, and WHSC1 into the IGH locus. The first genome-wide mutational data on 38 MM cases studied by WGS or WES found 10 genes to be significant recurring targets of mutation, including NRAS, KRAS, TP53, CCND1, and six genes not previously found to be mutated in cancer: FAM46C, DIS3, ALOX12B, HLA-A, and MAGED1. Notably, DIS3 (RRP44; mutated in 11% of cases) encodes an RNA exoribonuclease; the mutations identified are predicted to be loss of function, implicating altered RNA processing in MM. In addition, mutations were identified in genes regulating protein folding and translation, including XBP1. Recurring mutations were identified in BRAF, including p.Val600Glu and p.Lys601Asn. Mutations were also observed in genes regulating NF-κB signaling and histone modification (MLL, MLL2, MLL3, KDM6A [UTX], WHSC1, and WHSC1L1), and in sites of somatic mutation in the noncoding region (eg, 5’ and within BCL6). Thus, these results have implicated new pathogenic pathways in MM, but additional analysis and sequencing of MM is required to fully define the mutational landscape of MM for several reasons. MM genomes are structurally complex, and accurate resolution of structural genetic alterations in MM has not yet been comprehensively performed. Moreover, the nature of mutations that define each stage of the plasma cell disorders has not been fully investigated. Consistent with sequencing studies of other leukemias and lymphomas, MM is characterized by substantial clonal heterogeneity that is likely to be an important determinant of disease progression. Accordingly, reports are emerging of mutations associated with aggressive, extramedullary disease, and sequencing of serial cohorts of MGUS, smoldering MM, and MM and plasma cell lymphoma have begun to elucidate the subclonal structure of MM and its relationship to disease evolution and progression.  

The genetic basis of Waldenström macroglobulinemia (WM), a lymphoplasmacytoid lymphoma with secretion of monoclonal IgM, was unknown until recently. Remarkably, WGS of 40 WM samples identified a somatic mutation of MYD88, p.Leu265Pro, in 91% of IgM- and non-IgM-secreting lymphoplasmacytic lymphomas. MYD88 mutations were uncommon in MGUS or SMZL—disorders that must be distinguished from WM—and thus mutation detection may be useful at diagnosis. Other recurring mutations included WHSC1L1, perturbation of these core pathways is critical in many tumor types. Signaling; and epigenetic alterations. These findings indicate that perturbation of these core pathways is critical in many tumor types. Conversely, several genes are mutated in multiple tumor types, including BRAF in HCL and myeloma (as well as solid tumors such as melanoma), NOTCH1 in T-ALL and CLL, and MYD88 in multiple types of leukemia. This underscores the importance of these mutations both in pathogenesis and in the deployment of targeted therapies across multiple tumor types. These studies should be considered works in progress, and several important areas for future work are envisaged. Sequencing of many additional cases of each tumor type using complementary NGS modalities is needed. Most studies have examined relatively small numbers of tumors, and it is clear that the different types of NGS are complementary in defining the full landscape of genetic alterations. Moreover, several types of genetic alterations, including complex sequence mutations and structural rearrangements, are often difficult to robustly identify from the short reads generated by NGS.

Clinical implementation

It is clear that although additional sequencing of expanded cohorts of samples is required to fully define the genomic landscape of each tumor type, existing studies have identified genetic alterations that are immediately attractive for diagnosis, risk stratification, and therapeutic targeting. The most suitable sequencing approach is dependent on the type of alterations, and it should also be highlighted that clinical sequencing is rapidly evolving, both in terms of technological approaches and implementing rapid sequencing and analysis in a regulated environment. It is envisaged that diagnostic approaches will range from conventional molecular diagnostics targeting one or few genetic alterations (eg, BRAF or MYD88 alterations in HCL and WM; reverse transcriptase polymerase chain reaction for recurring fusions identified in high-risk ALL) to panels of targets for disease subtypes with multiple types of mutation, through to genome-wide sequencing approaches (WES, transcriptome sequencing, and/or WGS) for entities driven by multiple types of genetic alteration. Although NGS is likely to rapidly enter the clinic, a note of caution is warranted because, although many recurring alterations are of considerable biologic interest, many are not intuitively actionable or druggable. Moreover, the identification of potential druggable mutations in a novel disease context is increasingly common, and evidence to recommend the use of a targeted agent is often limited or absent (for a more detailed review on clinical sequencing, see Simon and Roychowdhury).

Conclusions

NGS studies reported in the last 3 years have advanced our understanding of the molecular basis of the lymphoid neoplasms by implicating new genes and pathways, refining classification schema, and identifying new targets for therapeutic intervention. As described above, a remarkable finding of these studies is the identification of similarly mutated genes and pathways in different hematopoietic and nonhematopoietic tumors: notably, transcriptional regulation of lymphoid development; lymphoid antigen receptor signaling; cytokine receptor, tyrosine kinase, and Ras signaling; and epigenetic alterations. These findings indicate that perturbation of these core pathways is critical in many tumor types. Conversely, several genes are mutated in multiple tumor types, including BRAF in HCL and myeloma (as well as solid tumors such as melanoma), NOTCH1 in T-ALL and CLL, and MYD88 in multiple types of leukemia. This underscores the importance of these mutations both in pathogenesis and in the deployment of targeted therapies across multiple tumor types.
advances in sequencing and computational approaches will yield additional insights. Most studies have sequenced bulk tumor populations, and these have inferred extensive intratumoral heterogeneity. Since it is now clear that subclonal mutations may be important determinants of disease progression and treatment failure, sequencing of purified cell populations, including single-cell analysis, will be required to fully dissect this clonal complexity and determine the relationship to disease behavior. Most studies have focused on somatic genetic alterations by directly comparing tumor genomes to matched nontumor genomes. The role of inherited variations in tumorigenesis has not been thoroughly investigated but is clearly important in view of recent studies identifying pathogenic mutations in familial and sporadic tumors and in the literature describing associations between common inherited variants and susceptibility to ALL.

Finally, most studies have focused on cataloging genetic alterations in the protein-coding genome. The role of genetic alterations in the noncoding genome and the relationship between structural, transcriptional, and epigenetic profiles remains to be fully explored and is particularly important in view of the high frequency of mutations in epigenetic regulators in many lymphoid neoplasms.

Acknowledgments

The author thanks the many colleagues and collaborators, particularly those from St. Jude Children’s Research Hospital, the PCGP, the Children’s Oncology Group, and the TARGET initiative, who have made key contributions to the studies described and apologizes to those whose work could not be cited for space limitations.

This work was supported by the American Lebanese Syrian Associated Charities of St. Jude Children’s Research Hospital; C.G.M. and is a Pew Scholar in the Biomedical Sciences, and a St. Baldrick’s Foundation Scholar.

Authorship

Contribution: C.G.M. wrote the manuscript.

Conflict-of-interest disclosure: The author declares no competing financial interests.

Correspondence: Charles G. Mullighan, Department of Pathology, Mail Stop 342, 262 Danny Thomas Place, Memphis TN 38105; e-mail: charles.mullighan@stjude.org.

References


