ASBMT American Society for Blood and Marrow Transplantation

From Trees to the Forest: Genes to Genomics

Charles Mullighan,[1] Effie Petersdorf,[2] Stella M. Davies,[3] John DiPersio[4]

Crick, Watson, and colleagues revealed the genetic code in 1953, and since that time, remarkable progress has been made in understanding what makes each of us who we are. Identification of single genes important in disease, and the development of a mechanistic understanding of genetic elements that regulate gene function, have cast light on the pathophysiology of many heritable and acquired disorders. In 1990, the human genome project commenced, with the goal of sequencing the entire human genome, and a "first draft" was published with astonishing speed in 2001. The first draft, although an extraordinary achievement, reported essentially an imaginary haploid mix of alleles rather than a true diploid genome. In the years since 2001, technology has further improved, and efforts have been focused on filling in the gaps in the initial genome and starting the huge task of looking at normal variation in the human genome. This work is the beginning of understanding human genetics in the context of the structure of the genome as a complete entity, and as more than simply the sum of a series of genes. We present 3 studies in this review that apply genomic approaches to leukemia and to transplantation to improve and extend therapies.

Biol Blood Marrow Transplant 17: S52-S57 (2011)

KEY WORDS: Genomics, Acute lymphoblastic leukemia, Acute myeloid leukemia, HLA haplotypes

GENETIC ALTERATIONS IN B-PROGENITOR ACUTE LYMPHOBLASTIC LEUKEMIA (ALL)

Genome-Wide Profiling of ALL Samples Obtained at Diagnosis

Multiple studies have used microarray-based profiling to identify submicroscopic genetic alterations in ALL, predominantly in children. These studies have made several important observations. In contrast to many solid tumors, ALL is characterized by a relatively low number of genetic alterations—approximately 6 to 8 lesions per case [1,2]. These alterations commonly target genes and pathways with key roles in leukemogenesis. These include mutations targeting transcriptional regulators of lymphoid development (eg, *PAX5*, *IKZF1*, and *EBF1*), cell cycle regulators, and tumor suppressor genes (*CDKN2A*, *CDKN2B*, *RB1*, and *PTEN*), lymphoid signaling genes (*CD200*, *BTLA*, and *BLNK*), and drug-response genes (eg, the glucocorticoid receptor gene *NR3C1*). Genes regulating B lymphoid development are mutated in the majority of B-progenitor ALL cases, most commonly deletions, sequence mutations, or translocations of *PAX5* [1], deletion (and less commonly, sequence mutation) of *IKZF1* (IKAROS), and the IKAROS family members *IKZF2* (HELIOS) and *IKZF3* (AIOLOS), and deletion of *EBF1*. These mutations result in loss of function in vitro, and accelerate the onset of ALL in murine models [3-5].

Genomic Profiling of High-Risk ALL—A Central Role of IKZF1

The pattern of genetic alterations in ALL is strongly associated with disease subtype. *MLL*-rearranged ALL cases harbor fewer than 1 copy number alteration per case, suggesting few deletions are required to induce leukemia [1,6]. In contrast, *ETV6-RUNX1* and *BCR-ABL1* (Ph+) ALL cases harbor multiple distinct copy number alterations [1]. Deletion of *IKZF1* (IKAROS) is a hallmark of Ph+ de novo ALL and chronic myeloid leukemia (CML) at progression to lymphoid blast crisis [2,7]. IKAROS is a zinc-finger transcription factor with complex actions including transcriptional regulation and chromatin remodeling, and is required for the development of all lymphoid lineages. The *IKZF1*

From the [1]Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee; [2]Division of Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, Washington; [3]Division of Bone Marrow Transplantation and Immune Deficiency, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; and [4]Genomics of AML Program and the Genome Center, Washington University School of Medicine, St. Louis, Missouri.

Financial disclosure: See Acknowledgments on page S56.

Correspondence and reprint requests: Stella M. Davies, MB, BS, PhD, Division of Bone Marrow Transplantation and Immune Deficiency, Cincinnati Children's Hospital Medical Center, 333 Burnet Avenue, Cincinnati, OH 45229 (e-mail: stella.davies@echmc.org).

1083-8791/$36.00
doi:10.1016/j.bbmt.2010.10.026

deletions in ALL typically result either in loss of expression of 1 copy of the gene and haploinsufficiency, or internal deletion of the gene and expression of an internally truncated isoform, IK6. This isoform lacks the N-terminal zinc fingers of IKAROS and cannot bind DNA, but retains the C-terminal zinc fingers and can act in a dominant-negative fashion. Existing data have indicated a role for IKAROS alterations in tumorigenesis. Mice harboring a dominant-negative mutation in the *Ikzf1* gene develop aggressive T-lineage lymphoproliferative disease [8], and deletion of *Ikzf1* accelerates leukemogenesis in a murine model of Ph+ ALL. Recent studies have also shown that inherited variations in the *IKZF1* gene are associated with an increased risk of ALL, further supporting a role for this gene in leukemogenesis [9]. Alterations of *IKZF1* are also associated with poor outcome in Ph−ALL. The U.S. and Dutch Children's Oncology Group have both found that *IKZF1* alteration was associated with a near tripling of the risk of treatment failure [10,11]. In addition, profiles of serial ALL samples have identified substantial differences in the genetic alterations present at diagnosis and relapse [12]. However, *IKZF1* alterations are almost always preserved from diagnosis to relapse, and may also be acquired as a new lesion at relapse [13,14].

The gene expression profile of poor outcome, *IKZF1*-altered B-progenitor ALL is strikingly similar to that of Ph+ ALL [10,14]. This suggests that these "*BCR-ABL1*-like" (Ph-like) cases may harbor alternative rearrangements or sequence mutations that activate signaling cascades downstream of cytokine receptors or tyrosine kinases, and this is indeed the case.

Genetic Characterization of *BCR-ABL1*-Like, Ph Negative ALL

Sequencing of tyrosine kinases in high-risk B-progenitor ALL (the COG P9906 cohort) [10] identified somatic mutations in *JAK1*, *JAK2*, and *JAK3* in approximately 10% of cases [15]. The mutations were most commonly at or near R683 in the pseudokinase domain of JAK2, but were also found in the kinase domain of JAK2 and the pseudokinase domain of JAK1. Strikingly, the V617F mutation commonly observed in the myeloproliferative disorders has not been identified in B-progenitor ALL, although the homolog of JAK2 V617F, JAK1 V658F, has been identified [16]. The presence of JAK mutations was associated with *IKZF1* mutations, a *BCR-ABL1*-like gene expression profile, and poor outcome. *JAK2* mutations (most commonly at R683) are also present in one-quarter of cases of B-progenitor ALL associated with Down syndrome [17-19]. The *JAK1* and *JAK2* mutations observed in ALL are transforming in vitro, conferring cytokine-independent growth and constitutive Jak-Stat activation when introduced into Ba/F3 cells expressing the erythropoietin or thrombopoietin receptors [15,18].

The identification of distinct JAK mutations in myeloproliferative diseases and ALL suggested that different mutated JAK alleles might interact with different downstream signaling pathways and influence the disease lineage. Recent studies have shown that JAK mutations in ALL are associated with chromosomal rearrangements resulting in overexpression of the cytokine receptor CRLF2 (cytokine receptor-like factor 2, or TSLPR, thymic stromal lymphopoietin receptor). These rearrangements are either a focal interstitial deletion involving the pseudoautosomal region 1 (PAR1) of Xp/Yp, or rearrangement of *CRLF2* into the immunoglobulin heavy chain locus (*IGH@-CRLF2*) [16]. Both represent a form of promoter/enhancer dysregulation. The PAR1 deletion results in a novel fusion transcript, *P2RY8-CRLF2*, in which the first, noncoding exon of *P2RY8* is fused to the entire coding region of *CRLF2*. *P2RY8* is a member of a family of purinergic receptor genes that is expressed in hematopoietic cells, including leukemic blasts, and has previously been identified as a rare target of translocation to *SOX5* in lymphoma.

CRLF2 alterations in B-progenitor ALL have been subsequently confirmed and identified by multiple groups, including adult ALL [16,20]. *CRLF2* is rearranged in 5% to 7% of B-progenitor childhood ALL cases. Both alterations result in increased cell surface expression of *CRLF2* by leukemic cells, and flow cytometric analysis of *CRLF2* expression may be used to detect *CRLF2*-rearranged cases. Less commonly, *CRLF2* is rearranged to other, as yet unknown partner genes or harbors presumed activating mutations, most commonly F232C. A striking observation is that *CRLF2* alteration, most commonly the PAR1 deletion, is present in over 50% of ALL associated with Down syndrome (DS-ALL) [16], in which other chromosomal rearrangements characteristic of childhood ALL are uncommon. In both DS- and non-DS-ALL, *CRLF2* rearrangement is significantly associated with the presence of activating Janus kinase mutation. Over half of *CRLF2*-rearranged cases harbor activating *JAK1* or *JAK2* mutations, and conversely, nearly all JAK-mutated cases have *CRLF2* rearrangements, suggesting that these lesions together contribute to leukemogenesis. Importantly, in non-DS-ALL, *CRLF2* alteration and JAK mutations are associated with the presence of *IKZF1* alterations. Several studies have observed strong associations between CRLF2/JAK alterations and very poor outcome [20], suggesting that JAK inhibition may be a useful therapeutic approach in these cases that commonly fail therapy.

Existing data suggest that aberrant CRLF2/JAK signaling contributes to leukemogenesis. CRLF2 forms a heterodimeric receptor with interleukin-7

receptor alpha (IL7RA) for the cytokine TSLP (thymic stromal lymphopoietin). TSLP/CRLF2 signaling has a role in dendritic cell development, T cell responses, allergic inflammation, and promotes the proliferation of normal and leukemic B cells. Expression of either *CRLF2* or mutant JAK alleles alone in Ba/F3 cells lacking expression of EpoR/TpoR usually does not result in transformation [16]. A notable exception is JAK1 V658F, the homolog of JAK2 V617F, which transforms this cell line irrespective of cytokine receptor coexpression. Prior to the identification of *CRLF2* alterations in ALL, JAK mutations in ALL were shown to transform Ba/F3-EpoR cells to cytokine-independent growth and result in constitutive Jak-Stat activation [15,18], suggesting that interaction of Jak mutants with a cytokine receptor scaffold is required for transformation. Subsequent studies have shown that coexpression of JAK mutations and *CRLF2* in Ba/F3 cells is transforming, and that this transformation is inhibited by either pharmacologic JAK inhibition or short hairpin RNA–mediated knockdown of *CRLF2* expression [16]. Similarly, studies using primary murine hematopoietic progenitors have shown that enforced expression of *CRLF2* alone promotes lymphoid expansion, but this is insufficient to result in the development of leukemia ([21] and unpublished data). Ongoing studies modeling *CRLF2* dysregulation and JAK mutations will be of interest to determine the role of these alterations in leukemogenesis. In addition, such studies will provide preclinical models of ALL that faithfully recapitulate human leukemia in which to test the efficacy of pharmacologic JAK inhibitors.

Future Directions for Genomic Profiling in High-Risk ALL

Our understanding of the genetic basis of high-risk ALL remains incomplete. Almost one-half of *CRLF2*-rearranged cases lack an activating JAK mutation, yet may have a *BCR-ABL1*-like gene expression profile, suggesting that additional cooperating or kinase-activating lesions remain to be identified. Moreover, many "*BCR-ABL1*-like" cases lack *CRLF2* alterations, and the genetic alterations driving these leukemias remain unknown. Similarly, a substantial proportion of ALL cases remain that lack known cytogenetic alterations and fail therapy, and the frequency of these cases rises with increasing age. There is a relative lack of detailed, high-resolution genomic profiling data from adolescent and adult ALL, which has a markedly inferior outcome to that of childhood ALL. This is a critical issue and an area of active enquiry. Furthermore, several high-risk subtypes of leukemia have either not been studied in detail (eg, ALL with low hypodiploidy) or have few structural genetic alterations on microarray analysis (eg, *MLL*-rearranged leukemia). Moreover, microarray platforms do not directly detect structural rearrangements or DNA sequence alterations.

Ongoing and future genomic profiling studies will entail analysis of less well-studied cohorts, and the application of next-generation sequencing approaches interrogating genetic and epigenetic changes. Next-generation sequencing of either tumor DNA or RNA has identified new targets of mutation in acute myelogenous leukemia (AML), T-lineage ALL, and lymphoma, and has identified new targets of rearrangement in cancer. Accordingly, preliminary studies in Ph-like B-lineage ALL have identified novel rearrangements dysregulating *ABL1* and *JAK2*. It is likely that as the time and cost requirements of these methods decline, sequencing-based approaches will assume greater importance in interrogating cancer genomes and may supplant array-based methodologies.

GENOMICS OF UNRELATED DONOR HEMATOPOIETIC CELL TRANSPLANTATION

Worldwide clinical experience demonstrates that HCT from unrelated donors (URDs) can cure both malignant and nonmalignant blood disorders. Since its discovery, the HLA system has served as the cornerstone of transplantation genetics. HLA genes reside within the major histocompatibility complex (MHC), the most comprehensively studied multimegabase region of the human genome. The importance of MHC resident variation in transplantation, autoimmunity, and infectious diseases has provided the necessary platform for investigation of the HLA region in human disease. More than 421 loci are now identified within the 7.6-Mb extended MHC region, an estimated 30% of which are involved in immune function including inflammation, leukocyte maturation, complement cascade, nonclassical class I, immunoregulation, stress response, and the immunoglobulin superfamily [22]. Of the 252 expressed genes, over 20% reside within the class III region, the most gene-dense region in the genome. The classical HLA genes, *HLA-A*, *B*, *C*, *DR*, *DQ*, and *DP* encode polypeptides that are critical in controlling T cell recognition and determining histocompatibility. Complete and precise URD matching for 10 alleles at the 5 genetic loci, HLA-A, B, C, DRB1, DQB1, is associated with lowered risks of graft rejection, acute and chronic graft-versus-host disease (aGVHD, cGVHD) and superior survival compared to HLA mismatching [23]. Despite 10/10 allele matching of URDs, transplant-related complications are still higher than that observed in sibling donor transplantation [24].

A hallmark of HLA genes is their extensive degree of polymorphism and the high degree of nonrandom association of alleles at 2 or more HLA loci, or linkage

disequilibrium (LD). Recently, a worldwide effort by the MHC Sequencing Consortium, the HapMap Project, The Human Genome Project, and The SNP Consortium has led to a dense map of over 36,000 MHC region single nucleotide polymorphisms (SNPs) and common haplotypes and important information on the organization of MHC region variation. High LD across the MHC opens the possibility that some of the over 250 expressed genes with immune function contribute to posttransplant risks after transplantation. Efforts to identify functional resident MHC genes may provide new insight into the pathogenesis of GVHD, and offer new approaches to the prevention of posttransplant complications [25].

GENOMICS OF AML AND MYELODYSPLASTIC SYNDROME (MDS)

Identifying genetic events that lead to the development of AML and MDS will not only provide insight into the biology of these diseases but also identify key pathways for therapeutic interventions. There has been an evolution and revolution in cancer genetics and genomics over the past 50 years culminating in novel and robust DNA sequencing technologies and array-based platforms that have been used to identify microdeletions, point mutations, amplifications, and translocations that are associated with these diseases. In addition to routine cytogenetic analysis and fluorescence in situ hybridization (FISH), which have been used to identify commonly mutated genes in AML such as *FLT3*, *NPM1*, *N-ras*, *CEBPA*, and *MLL*, our group has explored comparative genomic hybridization (CGH) and SNP array-based platforms to help identify small (<5 Mb) subcytogenetic amplifications and deletions in the AML genome. These studies have demonstrated that there are very few recurrent acquired copy number changes (CNVs) in most AML genomes, and that these techniques may not provide significantly more information and prognostic valve than routine cytogenetics and FISH [26]. In order to overcome the inherent variability of RNA profiling and the lack of validated technologies for precisely and reproducibly defining genome-wide methylation an effort began almost 10 years ago to perform massive parallel sequencing of AML genomes using modern platforms in order to take an unbiased view of both mutations associated with AML and allelic variants that may predispose to the development of both MDS and AML.

A large percentage of AML genomes are cytogenetically normal, and it is this particular group of AML patients for which optimal therapy has not yet been defined. In addition, the study of AML genomes with a array-based high-resolution comparative genomic hybridization approaches has revealed that many have no detectable copy number alterations at the 35-kb resolution. From preliminary studies, investigators determined that cytogenetically normal (diploid) genomes contain many fewer mutations than cancer genomes that are highly aneuploid. The likelihood of identifying mutations that might be pathogenetically relevant was improved by first focusing on these less complex cancer genomes. In addition, a number of mutations that have been now confirmed in hematologic malignancies have been seen in solid tumor malignancies (*RAS*, *ALK*).

In order to take a comprehensive and unbiased look at the AML and MDS genomes, next-generation sequencing has been applied, despite its significant cost and manpower requirements. A collaborative effort between multiple departments at Washington University School of Medicine and the Washington University Genome Center has completed the sequencing of over 100 AML and MDS genomes [27]. The technologies developed for the sequencing of AML genomes have been extended to other diseases including breast cancer, multiple myeloma, lung cancer, glioblastoma multiforme and genitourinary malignancies [28]. Finally, access to germline DNA (skin biopsies or buccal rinses) has been absolutely essential and has allowed us to attribute sequence changes in AML genomes to either polymorphisms or somatic mutations.

The projected cost of using traditional Sanger-based sequencing for human genomes would be approximately $90 million per genome, but with the advent of next-generation technology, the cost has already fallen to between $50,000 and $100,000 per cancer-germline pair. The expense is justified in light of the fact that humans have 3-4 million sequence variations and hundreds of copy number variations. Because of the large number of variations, sequencing of both tumor and matched normal tissue needs to be performed. Because most somatic mutations will be heterozygous, both alleles must be sequenced at every position in the genome to obtain adequate coverage for mutation discovery, whereas only 6-fold coverage (ie, 18 billion base pairs of sequence for a 3 billion base-paired genome) is required to solve a genome's primary structure; at least 4 times that or 25-fold coverage (72 billion bases pairs of sequence) is required to achieve adequate diploid coverage for comprehensive mutation discovery.

Currently, a shotgun sequencing method using "paired end reads" on massively parallel sequencing devices is the method of choice in most genome centers. Libraries of DNA fragments from AML and MDS tumor specimens can now be made with very limited amounts of DNA (as little as 100 ng), which is critical because sample abundance is very limiting for many tumors. Our initial work on sequencing cancer genomes focused on unidirectional short-fragment reads of 30-35 base pairs in length. Because of these

short reads, many could not be unambiguously mapped back to the referenced genome and longer reads, 100 base pairs in length, are now being read bidirectionally in so-called paired end reads. This dramatically reduces the cost and increases the overall efficiency of the sequencing reactions. In addition, paired end reads now dramatically increases our ability to identify structural variants such as deletions, amplifications, translocations, insertions-deletions (INDELs), and inversions. As the size of the fragments in these libraries increases, there is a dramatically increased chance of identifying translocations, which was almost impossible using libraries made from shorter DNA fragments.

Exon-capture methods are being developed and are being used widely. There are advantages, including large numbers of patients that could potentially be screened and the deep read counts for each captured exon ensuring high sensitivity for mutation discovery. Disadvantages are that these approaches are biased toward predefined regions of the genome, and the efficiency of capture method is still far from perfect. This must be considered still a work in progress.

The time now required to complete the sequencing of both tumor and germline DNA for simple diploid AML genomes is less than 1 week, and the cost is approximately $50,000/$100,000. This represents a 10- to 100-fold reduction in cost and time required to complete compared to the first genome sequenced.

The first genome sequenced was chosen specifically in a patient with de novo AML and the most common FAB classification of AML (M1). The patient was <60 years and had no exposure to chemotherapy and had no history of MDS. This patient had normal cytogenetics and no obvious microdeletions, amplifications, or INDELs using high-resolution and array-based technologies. Finally, this patient was also chosen because she eventually relapsed, which would allow us to identify relapse-specific mutations, which we have already completed. Surprisingly, in this first genome, we found only 10 nonsynonymous mutations within coding regions. By performing deep read counts of variant allele frequency for each mutation, we were able to establish that all of these nonsynonymous mutations were present in virtually all tumor cells. Only 4 of these mutations were found in the 187 additional AML genomes that were tested, strongly suggesting that they are indeed important for pathogenesis [27]. A second AML genome was sequenced [29]. There were only 12 mutations in coding regions and 54 mutations that were not in coding regions but did fall in highly conserved or regulatory regions of the genome.

The challenge is to sift through all of the noise in order to identify which mutations are recurrent and which of the recurrent mutations are driver mutations. To develop preclinical and animal models to test the biology of these mutations in biological systems will also be challenging and may require the development of completely novel in vitro and in vivo models. The final frontier will be to correlate the presence of the specific mutations with outcomes. Of note, in the second AML genome sequenced, we identified recurrent mutations, IDH1 and IDH2, which occur quite frequently in glioblastoma and were present in 8% to 9% of de novo AML patients. We also showed, which was confirmed by other groups, that these mutations were associated with a poor outcome.

MDS genomes are complex compared to de novo AML, and current studies are starting to document mutations. Additional or recurrent mutations that have important predictive impact on outcomes in patients with AML and MDS will also be identified in these studies. Identification of relapse-specific mutations that occur not only after standard antileukemic chemotherapy, but also after less aggressive therapies such as hypomethylating therapies and after maximally aggressive therapies such as allogeneic stem cell transplantation, will be of clinical importance.

ACKNOWLEDGMENTS

Financial disclosure: The authors have nothing to disclose.

REFERENCES

1. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007;446:758-764.
2. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*. 2008;453:110-114.
3. Miller CB, Mullighan CG, Su X, et al. Pax5 haploinsufficiency cooperates with BCR-ABL1 to induce acute lymphoblastic leukemia. *Blood*. 2008;112. abstract 293.
4. Dang J, Mullighan CG, Phillips LA, Mehta P, Downing JR. Retroviral and chemical mutagenesis identifies Pax5 as a tumor suppressor in B-progenitor acute lymphoblastic leukemia. *Blood*. 2008;112. abstract 1798.
5. Collins-Underwood JR, Miller CB, Downing JR, Mullighan CG. Ikzf1 haploinsufficiency contributes to the pathogenesis of BCR-ABL1 positive acute lymphoblastic leukemia. *Blood*. 2009;114. abstract 678.
6. Bardini M, Spinelli R, Bungaro S, et al. DNA copy-number abnormalities do not occur in infant ALL with t(4;11)/MLL-AF4. *Leukemia*. 2010;24:169-176.
7. Iacobucci I, Storlazzi CT, Cilloni D, et al. Identification and molecular characterization of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1-positive acute lymphoblastic leukemia patients: on behalf of Gruppo Italiano Malattie Ematologiche dell'Adulto Acute Leukemia Working Party (GIMEMA AL WP). *Blood*. 2009;114:2159-2167.
8. Winandy S, Wu P, Georgopoulos K. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. *Cell*. 1995;83:289-299.
9. Prasad RB, Hosking FJ, Vijayakrishnan J, et al. Verification of the susceptibility loci on 7p12.2, 10q21.2, and 14q11.2 in precursor B-cell acute lymphoblastic leukemia of childhood. *Blood*. 2010;115:1765-1767.

10. Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009;360:470-480.
11. Kuiper RP, Waanders E, van der Velden VH, et al. IKZF1 deletions predict relapse in uniformly treated pediatric precursor B-ALL. *Leukemia*. 2010;24:1258-1264.
12. Mullighan CG, Phillips LA, Su X, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science*. 2008;322:1377-1380.
13. Yang JJ, Bhojwani D, Yang W, et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood*. 2008;112: 4178-4183.
14. Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol*. 2009;10:125-134.
15. Mullighan CG, Zhang J, Harvey RC, et al. JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*. 2009;106:9414-9418.
16. Mullighan CG, Collins-Underwood JR, Phillips LA, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet*. 2009;41: 1243-1246.
17. Malinge S, Ben-Abdelali R, Settegrana C, et al. Novel activating JAK2 mutation in a patient with Down syndrome and B-cell precursor acute lymphoblastic leukemia. *Blood*. 2007;109: 2202-2204.
18. Bercovich D, Ganmore I, Scott LM, et al. Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down's syndrome. *Lancet*. 2008;372:1484-1492.
19. Kearney L, Gonzalez De Castro D, Yeung J, et al. A specific JAK2 mutation (JAK2R683) and multiple gene deletions in Down syndrome acute lymphoblastic leukaemia. *Blood*. 2008; 113:646-648.
20. Harvey RC, Mullighan CG, Chen IM, et al. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood*. 2010; 115:5312-5321.
21. Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood*. 2009;114:2688-2698.
22. Marsh SG, Albert ED, Bodmer WF, et al. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens*. 2010;75: 291-455.
23. Spellman S, Setterholm M, Maiers M, et al. Advances in the selection of HLA-compatible donors: refinements in HLA typing and matching over the first 20 years of the National Marrow Donor Program registry. *Biol Blood Marrow Transplant*. 2008;9 (Suppl):37-44.
24. Walter RB, Pagel JM, Gooley TA, et al. Comparison of matched unrelated and matched related donor myeloablative hematopoietic cell transplantation for adults with acute myeloid leukemia in first remission. *Leukemia*. 2010;24:1276-1282.
25. Petersdorf EW, Hansen JA. New advances in hematopoietic cell transplantation. *Curr Opin Hematol*. 2008;15:549-554.
26. Walter MJ, Graubert TA, Dipersio JF, et al. Next-generation sequencing of cancer genomes: back to the future. *Per Med*. 2009;6:653.
27. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008;456:66-72.
28. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. 2010;464: 999-1005.
29. Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med*. 2009;361:1058-1066.