



# Genomics of AML: Clinical Applications of Next-Generation Sequencing

John S. Welch<sup>1</sup> and Daniel C. Link<sup>1</sup>

<sup>1</sup>Division of Oncology, Department of Medicine, Washington University School of Medicine, St Louis, MO

In the past decade, a series of technological advances have revolutionized our ability to interrogate cancer genomes, culminating in whole-genome sequencing, which provides genome-wide coverage at a single base-pair resolution. As sequencing technologies improve and costs decrease, it is likely that whole-genome sequencing of cancer cells will become commonplace in the diagnostic workup of patients with acute myelogenous leukemia (AML) and other cancers. The unprecedented molecular characterization provided by whole-genome sequencing offers the potential for an individualized approach to treatment in AML, bringing us one step closer to personalized medicine. In this chapter, we discuss how next-generation sequencing is being used to study cancer genomes. Recent publications of whole-genome sequencing in AML are reviewed and current limitations of whole-genome sequencing are examined, as well as current and potential future clinical applications of whole-genome sequencing.

## Introduction

Since Nowell and Hungerford identified the t(9;22) translocation associated with chronic myeloid leukemia,<sup>1</sup> a wealth of data has accumulated showing that the karyotype and mutation status of certain genes provide important prognostic, and in some cases, therapeutic information for acute myelogenous leukemia (AML). However, standard cytogenetics has a limited resolution (3-5 megabases) to detect chromosomal abnormalities and genotyping of individual genes is expensive; therefore, as the number of mutations identified in AML increases, this method may soon become impractical. During the last decade, the technology to interrogate cancer genomes has rapidly advanced. The resolution for variant detection was improved first with array-based technologies using comparative genome hybridization (CGH) and single nucleotide polymorphism (SNP) analysis, and now with whole-genome sequencing, which provides unbiased, genome-wide coverage at a single base-pair resolution. Gene-expression profiling and CGH/SNP analysis of AML have provided important insights into AML classification and biology, and this has been recently reviewed elsewhere.<sup>2-4</sup> This chapter will focus on the use of next-generation sequencing technologies to characterize AML genomes.

In the past, the size and complexity of the human genome (3 billion base pairs) made the sequencing of human cancer genomes impractical. Two major advances helped to overcome these obstacles. First, the generation of the draft sequence of the human genome by the Human Genome Project in 2001 provided a road map of the human genome.<sup>5</sup> Second, technological advances in DNA sequencing dramatically reduced the cost and time required to sequence genomes (Figure 1). Whereas the Human Genome Project took more than 10 years and several billion dollars to sequence the first human genome, current estimates are 6 weeks and \$20 000 per human genome (\$40 000 for paired tumor/normal genomes). Therefore, we are rapidly approaching the time when sequencing the genomes of patients with cancer will be practical in the clinical setting.

## Next-generation sequencing of cancer

The development of massively parallel sequencing (termed next-generation or second-generation sequencing) revolutionized our

ability to analyze cancer genomes. Interested readers are referred to several recent reviews describing next-generation sequencing technologies.<sup>6,7</sup> In brief, massively parallel sequencing results in the generation of millions of short (50-100 nucleotides) DNA sequences simultaneously. These sequences are then mapped back to the human reference genome to generate a picture of the cancer genome. For studies of cancer, it is key to sequence both the tumor and the normal tissue (eg, skin tissue) from the individual. There are 3-4 million inherited sequence variants per human genome (and hundreds of copy number variants). Consequently, the majority of sequence variants identified in a cancer genome are inherited polymorphisms and not acquired mutations. Therefore, a comparison of a tumor genome with its paired normal genome is required to efficiently identify acquired (somatic) sequence variants.

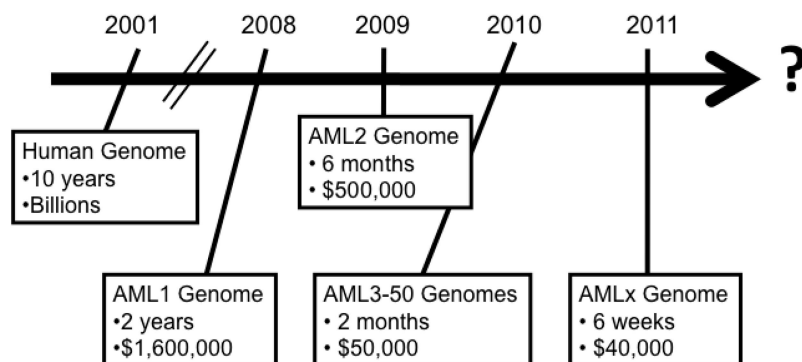
Currently, there are several different ways in which next-generation sequencing is being applied to study cancer genomes. The goals, advantages, and limitations of each approach are summarized below and in Table 1. Ultimately, combinations of approaches (eg, whole-genome and transcriptome sequencing) may be required to comprehensively study cancer cells.

## Whole-genome sequencing

The goal of whole-genome sequencing is to sequence the entire genome. Typically, this requires at least 30- to 40-fold haploid coverage of the genome (~100 Gbp of sequence) to achieve adequate diploid coverage for comprehensive mutation discovery. The advantages of whole-genome sequencing include: (1) the entire genome is surveyed, not just coding genes, and (2) structural variants, including deletions, amplifications, chromosomal translocations, and uniparental disomy, are readily identified. The major limitations are cost and the complexity of the data analysis. As cost and bioinformatic approaches to analyzing sequence data advance, whole-genome sequencing is likely to become the dominant platform for mutation detection.

## Exome sequencing

The goal of exome sequencing is to selectively sequence the 1%-2% of the genome containing coding genes, microRNAs, and other



**Figure 1. Timeline for sequencing of AML genomes.** Cost and time estimates are for paired leukemia/normal genomes and include data production, bioinformatic analysis, validation, and interpretation. Current cost and time estimates to sequence a human genome were provided by Dr Richard Wilson, Director, Genome Institute at Washington University (St Louis, MO).

noncoding RNAs. In brief, genomic DNA is hybridized to beads (or arrays) containing probes designed to capture the approximately 200 000 exons in the human genome. The DNA “captured” on the beads is then subjected to next-generation sequencing. The major advantages of exome sequencing are reduced cost and relatively deep sequence coverage because only 1%-2% of the genome is analyzed. However, exome sequencing will not detect mutations in regions outside of the exome (> 98% of the genome) and will not detect most structural variants, such as chromosomal translocations with intronic break points.

### Transcriptome sequencing

The goal of transcriptome sequencing is to sequence all transcribed genes,<sup>8</sup> including both coding and noncoding RNAs. The advantages of transcriptome sequencing include: (1) quantitative information about gene-expression levels are obtained, (2) posttranscriptional changes in gene expression such as alternative splicing are detected, and (3) fusion transcripts produced by chromosomal rearrangements can be detected. Similar to sequencing the exome, transcriptome sequencing does not detect mutations in noncoding regions of the genome. It cannot detect mutations that cause the loss of one or both copies of a gene or mutations that accelerate RNA turnover (eg, frameshift or nonsense mutations that cause “nonsense mediated decay”). Moreover, transcriptome sequencing is biased toward abundantly expressed transcripts and therefore the sequencing coverage of genes expressed at lower levels can be low or absent.

### Other sequencing applications

There are emerging next-generation sequencing applications to characterize genome-wide epigenetic modifications in cancer cells. This is particularly relevant to the study of AML, because genes regulating DNA methylation are frequently mutated in AML, including *DNMT3A*. Next-generation sequencing techniques have been developed to map genome-wide DNA methylation at a single-base resolution<sup>9</sup> and to assess chromatin structure by mapping the location of posttranslationally modified (eg, methylated and acetylated) histones on the genome.<sup>10</sup>

### Identification of novel somatic mutations in AML

Traditional approaches to mutation discovery in cancer rely on sequencing of a preselected group of genes. Typically, genes or gene family members already implicated in cancer are sequenced. This directed sequencing approach (in effect “looking under the lamppost”), while yielding important insights into cancer biology, is inherently limited. In contrast, genomic approaches study the entire cancer genome and thus provide a nonbiased way to look for mutations. Moreover, by simultaneously evaluating all genes, whole-genome sequencing is an ideal platform to assess cooperativity between gene mutations.

In a landmark study published in 2008, Ley et al reported the sequence of the first primary cancer genome, a case of FAB M1-AML.<sup>11</sup> Since then, the sequence of 3 additional AML genomes have been reported: an additional case of M1-AML (normal karyotype)<sup>11-13</sup> and 1 case each of M3-AML<sup>14</sup> and therapy-related

**Table 1. Summary of next-generation genomic-sequencing assays**

Sequencing target	Approach	Comment
Whole genome	Unselected genomic DNA is fragmented and sequenced	Comprehensive but expensive; detects point mutations, indels, translocations, deletions, and amplifications
Exome	Exonic DNA is enriched by capture and then sequenced	Only identifies mutations in exons, misses many structural variants
Transcriptome	RNA is reversed transcribed to generate cDNA and then sequenced	RNA expression and splicing are assessed; only identifies mutations in expressed genes
Methylome	gDNA is treated with bisulfite to mark methylated cytosines before sequencing*	Assay being optimized; provides genome-wide information about gDNA methylation
Histone-associated genome	ChIP using antibodies against modified histones followed by sequencing	Used to assess chromatin modifications, in particular genome-wide occupancy by modified histones

gDNA indicates genomic DNA.

\*Other methods to characterize the methylome also have been described previously.<sup>33</sup>

AML (complex karyotype).<sup>15</sup> In addition, exome sequencing was reported in 14 cases of M5-AML.<sup>16</sup>

These studies have begun to define several important global features of AML. First, there are only a relatively small number of somatically acquired single nucleotide variants (SNVs or “point mutations”) in AML cells. Within coding regions, the number of somatic SNVs ranged from 10-26 in the 4 sequenced AML genomes.<sup>11-15</sup> This suggests that AML does not result from genomic instability, but rather may be associated with a small number of oncogenic “driver” mutations. Second, most acquired AML-associated SNVs are private: of the 59 reported SNVs in the coding regions of these 4 genomes, no gene was mutated in 2 or more patients. This suggests that many of these mutations are likely to be personal “passenger” mutations, which are random and do not contribute to AML pathogenesis.

Despite the lack of recurring mutations within these 4 genomes, these studies identified 2 novel AML-associated genes that were found to be frequently mutated in AML and contribute to leukemogenesis. Mardis et al identified mutations in *IDH1* in the second AML genome published and found an additional 16 patients carrying *IDH1* mutations.<sup>13</sup> Mutations in *IDH1* and *IDH2* have since been identified in AML patients in several studies, occur at a frequency of 7%-16%, and appear to be mutually exclusive of one another and of acquired *TET2* mutations.<sup>17,18</sup> A recent study found that mutations in *IDH1* or *IDH2* disrupted *TET2* function and led to a hypermethylation phenotype with impaired hematopoietic differentiation,<sup>19</sup> although a second study found impaired methylation associated with *TET2* mutations.<sup>20</sup> Efforts are ongoing to develop targeted therapeutics that might abrogate these effects.

Acquired mutations in *DNMT3A*, a de novo DNA methyltransferase, were identified by next-generation sequencing of AML samples by 3 independent groups. Yamashita et al used targeted exome sequencing to identify acquired mutations in *DNMT3A* at amino acid R882 in 3 of 71 (4.2%) AML cases.<sup>21</sup> Ley et al identified a single base-pair deletion in *DNMT3A* by whole-genome sequencing and identified acquired mutations (including deletions, frame-shifts, nonsense, and missense mutations) in an additional 62 of 281 AML patients (22.1%).<sup>12</sup> They showed that *DNMT3A* mutations were more common in AML with normal cytogenetics (33.7%) or M5-AML cytomorphology (57.1%) and were associated with an adverse outcome. Finally, Yan et al used exome sequencing to identify missense mutations at R882 and at 3 other amino acid positions in 23 of 112 (20.5%) patients with AML-M5.<sup>16</sup> It is currently unclear whether some of these mutations (eg, the highly recurrent R882 mutations) might be associated with a gain of function. However, Challen et al found that *Dnmt3a*-deficient HSCs expanded significantly after serial transplantation, suggesting that the loss of *DNMT3A* could contribute directly to leukemogenesis.<sup>22</sup> A detailed description of the clinical utility of *IDH1*, *IDH2*, and *DNMT3A* mutations in the risk stratification and treatment of AML is provided in the companion chapter by Döhner and Gaidizik.

### Identification of cancer susceptibility genes in AML

Cancer susceptibility is an important consideration in individuals presenting with cancers at an early age, with multiple primary cancers, or with a suggestive family history. Determining the genetic basis for cancer susceptibility is desirable because it may affect the prevention, early detection, and treatment of associated neoplasms. There are more than 100 genes implicated in cancer susceptibility.<sup>23</sup> Genetic testing for all of these genes is cost-

prohibitive, impractical, and frequently unrevealing, because many of the genes contributing to cancer susceptibility have not yet been identified. Whole-genome sequencing provides a comprehensive, unbiased approach to identify mutations in genes contributing to cancer susceptibility.

We recently reported the results of whole-genome sequencing in a woman who presented at age 37 with breast cancer, at age 39 with ovarian cancer, and at age 42 with rapidly fatal therapy-related AML.<sup>15</sup> Although her family history of cancer was unremarkable, the early onset of multiple cancers suggested a cancer susceptibility syndrome. Commercial genetic testing for mutations in *BRCA1* and *BRCA2* was negative. Consistent with current guidelines for assessing familial cancer risk, no further genetic testing was performed. Whole-genome sequencing revealed a novel, heterozygous 3-kb deletion removing exons 7-9 of *TP53* in the patient's normal skin DNA, which was homozygous in the leukemia DNA as a result of uniparental disomy. Genetic testing of the patient's mother and the absence of a history of cancer in the patient's extended family suggested that this mutation arose spontaneously in the patient. The finding of the germline *TP53* mutation has important clinical implications for the patient's children. Carriers of a deleterious *TP53* germline mutation have an approximately 90% lifetime risk of developing cancer, and current guidelines from the National Comprehensive Cancer Network (NCCN) recommend annual screening for certain cancers.<sup>24</sup> Therefore, although whole-genome sequencing did not affect the management of this patient's AML, it did provide clinically important, actionable information for her family.

### Characterization of complex genomes and cryptic translocations in AML

Cytogenetic translocations such as t(15;17), t(8;21), and inv(16) have become an essential part of AML diagnostics and are used to predict response to therapy. This is perhaps best illustrated by the use of all-trans-retinoic acid (ATRA) therapy in patients with t(15;17) and the associated *PML-RARA* fusion oncogene. However, standard cytogenetics or FISH analysis do not detect all translocations producing *PML-RARA*. Cryptic insertions and/or complex translocations that produce functional *PML-RARA* transcripts have been described.<sup>25-27</sup> We recently described a case in which whole-genome sequencing was used clinically to identify an actionable cryptic translocation in a patient with M3-AML.<sup>14</sup>

### Case history

A 39-year-old woman presented with morphologic and clinical features of M3-AML. However, chromosomes 15 and 17 appeared to be normal by metaphase cytogenetics, but multiple chromosomal alterations were identified, including unusual translocations and the loss of whole chromosomes, consistent with high-risk complex cytogenetics. FISH was not diagnostic for *PML-RARA*. The patient entered complete remission after induction therapy with cytarabine and idarubicin, and her sister was found to be an HLA match. This presented a diagnostic conundrum with divergent treatment options: does she have APL (which should be treated with ATRA consolidation) or does she have complex cytogenetic AML (which should be treated with allogeneic stem cell transplantation in first remission)?

To address this issue, whole-genome sequencing was performed. The patient was empirically treated with 2 cycles of arsenic consolidation, which imposed a deadline of 8 weeks to complete the sequencing, perform the bioinformatic analysis, and validate the findings. This deadline was met, and whole-genome sequencing

revealed that a 77-kb fragment from the *LOXLI-PML* locus on chromosome 15 was cleanly inserted into the *RARA* locus on chromosome 17, creating an in-frame *PML-RARA* fusion allele. Admittedly, alternative strategies may have identified this insertion. RT-PCR to detect the *PML-RARA* fusion transcript was not initially performed and clinical samples were not available later for analysis. Other research-based approaches could be envisioned (eg, long-range PCR or nested PCR using a linker-ligated library). However, each of these requires personalized design and iterative trouble-shooting. In summary, this case illustrates that, already in 2011, whole-genome sequencing can be clinically valuable in appropriately selected patients. The diagnostic information obtained from whole-genome sequencing allowed this patient to avoid the hazards and expense of an allogeneic stem cell transplantation in favor of ATRA consolidation.

### Current limitations of whole-genome sequencing

Although clearly a valuable and promising approach to analyzing AML (and other cancers), there are some technical limitations of whole-genome sequencing. The ability of next-generation sequencing to detect sequence variants is dependent on the “read-depth,” the number of unique times a single nucleotide is sampled. Whole-genome sequencing typically aims to cover each nucleotide with an average of 30–40 separate reads. However, at this coverage, nearly 50% of sequence variants identified are false positives due to mapping errors, polymerase errors, and the low frequency of somatic variants within the sample. Therefore, secondary validation of variants remains a necessary part of all sequencing projects. The detection of small insertions and deletions (indels) using current algorithms is problematic. In addition, mapping sequence reads to the approximately 50% of the human genome that contains repetitive sequences is imperfect, raising the possibility that mutations are being missed. The incident AML case highlighted these technical issues<sup>11</sup>; a single base-pair deletion in *DNMT3A* was missed during short-read analysis (30- to 40-bp reads, which are difficult to align to the reference genome if they contain an indel), but was later identified when the genome was resequenced with paired-end reads (pairs of 50- to 75-bp reads, which align more robustly). These problems will be improved as bioinformatics can draw on additional sequenced human genomes and as decreased costs permit deeper sequencing, but efficiently mapping sequences to repetitive regions of the genome will likely require technical improvements that permit longer sequence read lengths.

There are inherent limitations of whole-genome sequencing that need to be considered when analyzing cancers. First, each AML genome carries approximately 500–1000 somatic mutations.<sup>12</sup> It is likely that the majority of these mutations do not contribute to leukemic transformation. Identifying the “driver” mutations in each AML genome is currently challenging. Second, whole-genome sequencing does not provide information about epigenetic modifications and alterations in gene expression. Therefore, to fully characterize acquired genomic changes in an AML patient, assays to measure RNA expression (eg, transcriptome sequencing) and epigenetic changes (eg, bisulfate sequencing to define the methylation) in combination with whole-genome sequencing are required.

Genome sequencing remains expensive, and the requisite infrastructure, expertise, and time to complete sequence analysis are significant barriers to the routine use of whole-genome sequencing of AML in the clinical setting. However, as the number of sequenced AML genomes increases, the technical and bioinformatic infrastructure required for clinical analysis will become more accessible.

Current cost estimates at the Washington University Genome Institute for whole-genome sequencing of a paired normal/leukemia sample is \$40 000. This is a fully loaded cost that includes sequencing, bioinformatic analysis, and validation. The current time to sequence and analyze a paired tumor and normal genome is 6 weeks. The cost of whole-genome sequencing needs to be considered in the context of the increasingly complex diagnostic evaluation for AML. A common molecular evaluation consists of cytogenetics; FISH for t(15;17), t(8;21), inv(16), and t(9;22); and molecular testing for *FLT3*, *KIT*, and *NPM1* mutations. The list price of this “standard of care” molecular evaluation at Quest Diagnostics (a national laboratory) is currently \$4575 and does not include the cost of molecular testing for other recurring mutations (eg, *NRAS*, *IDH1*, *IDH2*, *TET2*, *DNMT3A*, *RUNX1*, *ASXL1*, and *CEBPA*) that are likely to become routine. As the cost of sequencing continues to decrease and the number of molecular assays continues to increase, we will rapidly reach the inflection point where whole-genome sequencing will be the most cost-effective diagnostic tool for AML patients.

Finally, there are important practical and ethical considerations concerning the clinical application of whole-genome sequencing. Based on the cancer genomes analyzed to date, it is clear that a tremendous number of genetic variants will be identified in each AML genome, of which the majority will be inherited variants. Some of these will be in genes known to be involved in cancer susceptibility, and some will be associated with unsuspected genetic diseases (eg, hereditary hemochromatosis). Best practices to communicate this information to patients and their families are being considered and developed by several groups of investigators.

### Potential clinical applications (current and future)

Whole-genome sequencing provides a method to identify all genetic changes in a cancer genome, including single-nucleotide mutations, deletions, amplifications, chromosomal translocations, and uniparental disomy, at a single-base resolution. There are several ways that this unprecedented ability to characterize cancer genomes may affect the clinical management of patients with AML.

#### Classification of AML

Whole-genome sequencing may lead to a better classification of AML. There is evidence that a molecular classification of AML based on karyotype and specific gene mutations is superior to a morphology-based classification. As additional mutations are discovered in AML, it is likely that the molecular classification of AML will be further defined. Having access to all of the genetic changes in an AML genome through whole-genome sequencing would facilitate this process.

#### Optimizing therapy of AML

A detailed knowledge of mutations present in the AML genome may help to optimize the therapeutic plan. For example, mutations in tyrosine kinase genes may suggest the use of specific small-molecule inhibitors. In addition, disruption of specific pathways in the AML genome may be exploited therapeutically. For example, mutations in *BRCA1* or *BRCA2*, by inhibiting DNA repair through homologous recombination, may render cancer cells more sensitive to inhibitors of protein poly(ADP)-ribose polymerase.<sup>28</sup> Finally, the discovery of new mutations in AML may lead to the development of novel targeted therapeutics.



## Pharmacogenomics

Pharmacogenomics is the study of how an individual's genes influence his or her response to drugs. Genetic variants can influence the metabolism, toxicity, and therapeutic efficacy of a drug. By sequencing the "normal" and leukemia genomes in patients with AML, genetic variants known to effect the efficacy and toxicity of chemotherapeutic agents will be identified and this information may affect the choice and dose of chemotherapy used. For example, germline polymorphisms in the UDP glucuronosyl-transferase 1 family (*UGT1A1*) increase the toxic effects of irinotecan,<sup>29</sup> and germline polymorphisms in thiopurine S-methyltransferase (*TMPT*) increase sensitivity to mercaptopurine through decreased metabolism.<sup>30</sup> As more genomes are sequenced, the list of genetic variants influencing response to chemotherapy likely will expand.

## Cancer susceptibility

As discussed above, whole-genome sequencing of the "normal" genome in patients with AML may identify genetic variants contributing to cancer susceptibility. Not only will this information be important for genetic counseling, it may affect the patient's treatment and prognosis. Whole-genome sequencing will provide genotyping data for the more than 100 genes already implicated in cancer susceptibility. Causal associations between germline variants and AML have been firmly established for only a few genes, such as *TP53*<sup>31</sup> and *RUNX1*.<sup>32</sup> However, it is likely that the list of genetic variants contributing to AML susceptibility will increase in the near future as additional genomes are sequenced.

## Conclusions

The Human Genome Project and advances in sequencing technologies have revolutionized our ability to characterize cancers at a molecular level. Using whole-genome sequencing, it is now possible to interrogate cancers genome wide at a single-base-pair resolution. Application of whole-genome sequencing to AML has already yielded important discoveries, including the identification of common gene mutations in AML (eg, *IDH1* and *DNTM3A*). Ongoing whole-genome sequencing of additional AML samples likely will identify the majority of relevant mutations in AML. In the near term, sequencing of a panel of AML-related medically actionable mutations could be performed economically using next-generation sequencing and custom-capture arrays. However, with continued improvements in sequencing, it is likely that whole-genome sequencing will become a routine part of the diagnostic workup of patients with AML and other cancers.

## Disclosures

Conflict-of-interest disclosure: J.S.W. has received research funding from Eisai Pharmaceuticals. D.C.L. declares no competing financial interests. Off-label drug use: None disclosed.

## Correspondence

Daniel C. Link, MD, Division of Oncology, Department of Medicine, 660 S Euclid Ave, Campus Box 8007, St Louis, MO 63110; Phone: (314) 362-8771; Fax: (314) 362-9333; e-mail: dlink@dom.wustl.edu.

## References

1. Nowell P, Hungerford DA. A minute chromosome in human chronic granulocytic leukemia. *Science*. 1960;132:1497.

2. Godley LA, Cunningham J, Dolan ME, et al. An integrated genomic approach to the assessment and treatment of acute myeloid leukemia. *Semin Oncol*. 2011;38(2):215-224.
3. Suela J, Alvarez S, Cigudosa JC. DNA profiling by arrayCGH in acute myeloid leukemia and myelodysplastic syndromes. *Cytogenet Genome Res*. 2007;118(2-4):304-309.
4. Watt CD, Bagg A. Molecular diagnosis of acute myeloid leukemia. *Expert Rev Mol Diagn*. 2010;10(8):993-1012.
5. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
6. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet*. 2010;11(10):685-696.
7. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010;11(1):31-46.
8. Blencowe BJ, Ahmad S, Lee LJ. Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev*. 2009;23(12):1379-1386.
9. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315-322.
10. Neff T, Armstrong SA. Chromatin maps, histone modifications and leukemia. *Leukemia*. 2009;23(7):1243-1251.
11. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008;456(7218):66-72.
12. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363(25):2424-2433.
13. Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med*. 2009;361(11):1058-1066.
14. Welch JS, Westervelt P, Ding L, et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA*. 2011;305(15):1577-1584.
15. Link DC, Schuettpelz LG, Shen D, et al. Identification of a novel TP53 cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML. *JAMA*. 2011;305(15):1568-1576.
16. Yan XJ, Xu J, Gu ZH, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*. 2011;43(4):309-315.
17. Damm F, Heuser M, Morgan M, et al. Integrative prognostic risk score in acute myeloid leukemia with normal karyotype. *Blood*. 2011;117(17):4561-4568.
18. Schnittger S, Haferlach C, Ulke M, Alpermann T, Kern W, Haferlach T. IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutated NPM1 status. *Blood*. 2010;116(25):5486-5496.
19. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010;18(6):553-567.
20. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010;468(7325):839-843.
21. Yamashita Y, Yuan J, Suetake I, et al. Array-based genomic resequencing of human leukemia. *Oncogene*. 2010;29(25):3723-3731.
22. Challen GA, Berg JS, Goodell MA. Loss of de novo DNA

- methylation causes expansion of the mouse hematopoietic stem cell pool [abstract]. *Blood*. 2010;116(21):835.
23. Garber JE, Offit K. Hereditary cancer predisposition syndromes. *J Clin Oncol*. 2005;23(2):276-292.
  24. Hwang SJ, Lozano G, Amos CI, Strong LC. Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk. *Am J Hum Genet*. 2003;72(4):975-983.
  25. Abe S, Ishikawa I, Harigae H, Sugawara T. A new complex translocation t(5;17;15)(q11;q21;q22) in acute promyelocytic leukemia. *Cancer Genet Cytogenet*. 2008;184(1):44-47.
  26. Grimwade D, Biondi A, Mozziconacci MJ, et al. Characterization of acute promyelocytic leukemia cases lacking the classic t(15;17): results of the European Working Party. Groupe Francais de Cytogenetique Hematologique, Groupe de Francais d'Hematologie Cellulaire, UK Cancer Cytogenetics Group and BIOMED 1 European Community-Concerted Action "Molecular Cytogenetic Diagnosis in Haematological Malignancies". *Blood*. 2000;96(4):1297-1308.
  27. Hiorns LR, Swansbury GJ, Catovsky D. An eight-way variant t(15;17) in acute promyelocytic leukemia elucidated using fluorescence in situ hybridization. *Cancer Genet Cytogenet*. 1995;83(2):136-139.
  28. Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med*. 2009;361(2):123-134.
  29. Iyer L, King CD, Whittington PF, et al. Genetic predisposition to the metabolism of irinotecan (CPT-11). Role of uridine diphosphate glucuronosyltransferase isoform 1A1 in the glucuronidation of its active metabolite (SN-38) in human liver microsomes. *J Clin Invest*. 1998;101(4):847-854.
  30. Weinshilboum RM, Wang L. Pharmacogenetics and pharmacogenomics: development, science, and translation. *Annu Rev Genomics Hum Genet*. 2006;7:223-245.
  31. Birch JM, Alston RD, McNally RJ, et al. Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene*. 2001;20(34):4621-4628.
  32. Song WJ, Sullivan MG, Legare RD, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet*. 1999;23(2):166-175.