EDITORIALS

The Beginning of the End of the Beginning in Cancer Genomics

David P. Steensma, M.D.

This is the age of massive genome surveys — at least for a little while longer. Sixty years after Watson and Crick's discovery, and a decade after the completion of the Human Genome Project, large-scale sequencing efforts directed at human disease abound, especially for cancer and rare congenital syndromes.[1,2] International research teams supported by public funding agencies such as the National Institutes of Health and by private foundations such as the Wellcome Trust are rapidly enlarging the catalogue of genetic changes associated with neoplasia and other ailments, using ever faster, ever cheaper sequencing methods and heavy-duty bioinformatics.

Critics of big genomics projects have argued that such work is resource-intensive, is not hypothesis-driven, and amounts to little more than molecular philately.[3,4] But as discoveries from these projects stack up, and as terabytes of observational data yield new insights into disease biology and prompt the development of pathway-driven targeted therapies, the usefulness of such approaches is becoming undeniable. When the Cancer Genome Atlas (TCGA) wraps up its 8-year effort next year, it will have provided detailed information on 10,000 cancer genomes for less than the cost of a trio of F-22 Raptor stealth fighters.

Ley and colleagues in the Cancer Genome Atlas Research Network now report in the *Journal* findings from multiplatform genome analysis of 200 samples obtained from patients with de novo acute myeloid leukemia (AML).[5] (See Fig. 1 for an overview of some of the methods used.) The myeloid leukemias have long served as model systems for cancer biology, because leukemic cells are easily accessible for repeated sampling and leukemia diagnosis is usually straightforward. Fittingly, the first whole cancer genome that was sequenced was obtained from a patient with AML.[2] Chronic myeloid leukemia (CML), defined by fusion of the BCR and ABL1 genes through a translocation creating the Philadelphia chromosome (the first recurrent chromosomal abnormality linked to a human disease), proved that understanding the mutations that drive cancer can lead to effective, narrowly targeted therapies.[6] Treatment of CML with imatinib and other tyrosine kinase inhibitors also illuminated Darwinian competition among neoplastic subclones and mechanisms of resistance to targeted therapy that have been replicated in many other neoplasms.[7,8]

The AML-sequencing project has already generated several important observations. Discovery of recurrent somatic mutations affecting isocitrate dehydrogenase that confer neomorphic enzyme activity, together with similar findings in gliomas, reinvigorated the field of cancer metabolomics.[9] Mutations in *DNMT3A*, encoding a DNA methyltransferase, were found to be common in AML and have prognostic value.[10]

The new work confirms that AML genomes have fewer mutations than common epithelial cancers such as lung and ovarian cancer. This lack of complexity is relative, however. The clonal architecture of AML is dazzlingly intricate, especially in cases arising from the myelodysplas-

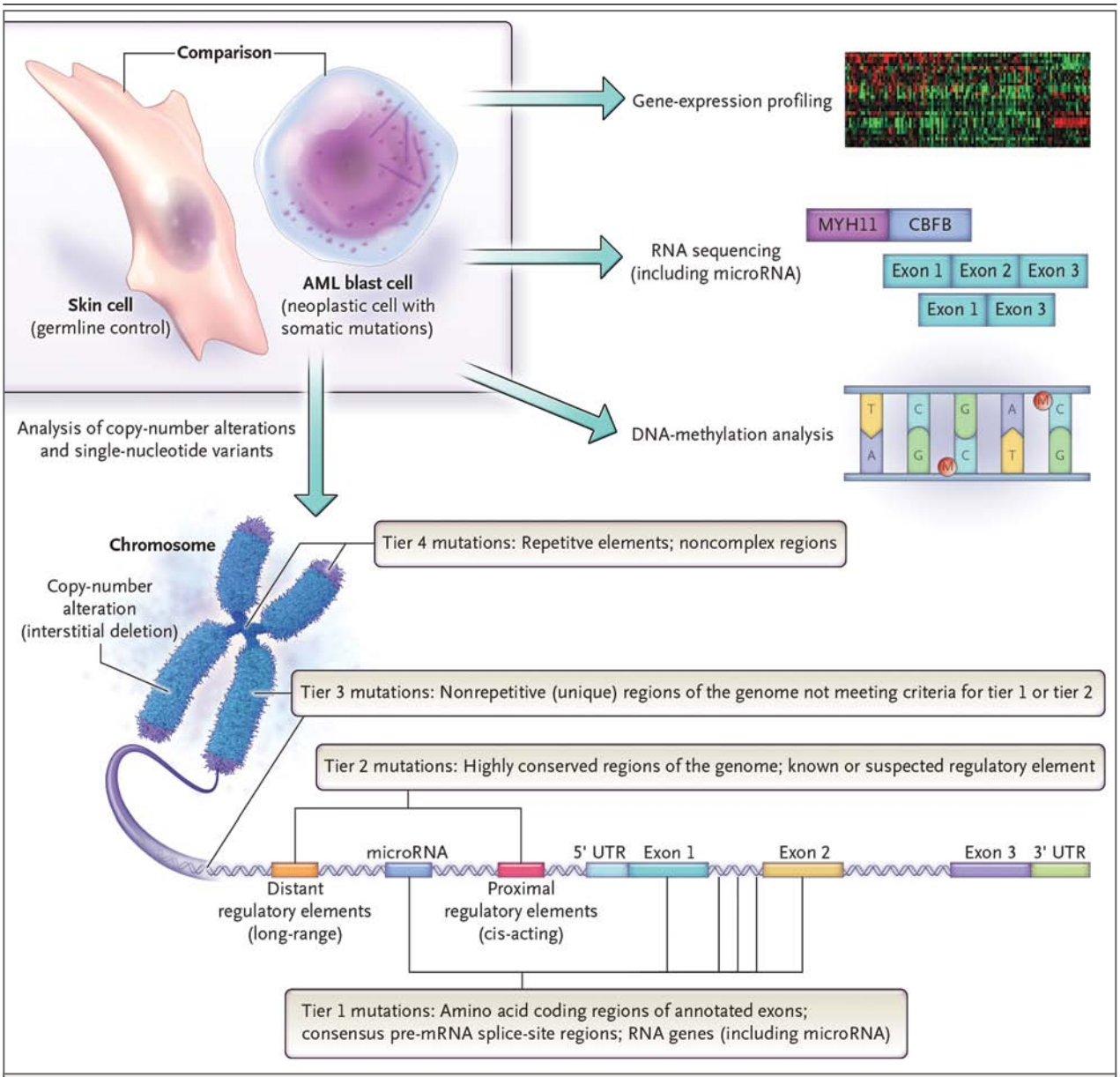

Figure 1. A Comprehensive Analysis of Acute Myeloid Leukemia (AML).

Large-scale projects to catalogue the cancer genome, such as those that are part of the Cancer Genome Atlas, characterize the genomes of tumor cells and nonmalignant cells obtained from the same patient. In the study of AML reported by Ley et al.,[5] skin cells served as the matched control cells, which are presumed to retain a germline configuration. Comparison of genetic variants in tumor cells with those of control cells permits detection of tumor-specific acquired (somatic) changes. Ley et al. used several high-throughput techniques to detect genomic features at different levels. These included conventional karyotyping, exome and genome sequencing, RNA sequencing — which allows for detection of gene fusions and altered messenger RNA splicing, comparison of transcript expression levels, and assessment of microRNAs and other noncoding RNAs — and a pangenomic assay of DNA cytosine methylation. DNA single-nucleotide variants observed in cancer cells can be classified into four tiers, ranked in descending likelihood of pathological significance. The abbreviation mRNA denotes messenger RNA, and UTR untranslated region.

tic syndromes, with some subclones becoming extinct over time and others achieving dominance, unpredictably.[11] Patterns of interaction among mutations, including cooperation (i.e., mutations seen together more commonly than would be expected statistically, such as *FLT3* with *DNMT3A* and *NPM1*) and mutual exclusivity (i.e., mutations seen together less frequently than predicted by overall prevalence, such as mutations in *FLT3* and in genes encoding other tyrosine kinases), point to elaborate biologic relationships that deserve additional study.

It is likely that all the common, recurrent genetic lesions in AML — the molecular equivalents of major causes of death, such as stroke and heart disease — are now described. In individual cases, rare genetic events may occur, akin to uncommon causes of death, such as falling down a well or being struck by space debris. Within 2 years, the door to major new genetic findings will also close for most other common neoplasms and even for some rarer tumors.

Even though one era of cancer genomics is coming to a close, young investigators should not despair that all the best fruit has already been plucked. Numerous functional studies are needed, especially of mutant proteins with poorly characterized roles in cancer, such as the cohesins or spliceosome components in AML. The mechanism by which many of the other mutations described by Ley and colleagues contribute to leukemogenesis is far from obvious. (Definitions of mutation tiers are provided in Fig. 1.) Furthermore, many tier 2 mutations, and perhaps even some of those classified as tier 3 mutations, are likely to be pathologically relevant; researchers have thus far focused almost exclusively on coding mutations (which fall into the tier 1 category). For clinical trialists, the genomic complexity of AML and other cancers means that radical changes are needed in how new drugs are designed, studied, regulated, and marketed. Perhaps academic centers can regain a greater role in pharmaceutical development as blockbusters give way to boutiques.

As large-scale genome sequencing moves from the laboratory to the clinic, the process of making molecular data rapidly available to treating physicians is a major logistical hurdle that must be overcome. For instance, the presence of *DNMT3A*, *NPM1*, or *MLL* mutations influences dose response to daunorubicin, which is used to treat AML.[12] However, since patients with AML often require urgent initiation of therapy, results of genetic testing are not available to clinicians until long after daunorubicin treatment is completed.

In 1803, a few years before the inaugural issue of the *Journal*, Thomas Jefferson commissioned Meriwether Lewis and William Clark to survey the vast unknown American frontier. Lewis and Clark departed from St. Louis, where Ley et al. initiated the AML genome survey. Less than a century later, the western frontier was declared "closed," but land surveyors did not disappear; today, they focus on construction projects and property boundaries. Likewise, although the initial epic AML genomic survey that began in St. Louis is now largely complete and surveys of other neoplasms will soon conclude, the use of genomics in quotidian practice is just beginning.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Division of Hematologic Malignancies, Dana–Farber Cancer Institute and Harvard Medical School, Boston.

This article was published on May 1, 2013, at NEJM.org.

1. Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. Science 2003;300:286-90.
2. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 2008;456:66-72.
3. Weinberg R. Point: hypotheses first. Nature 2010;464:678.
4. Kaiser J. A skeptic questions cancer genome projects. Science Insider. April 23, 2010 (http://news.sciencemag.org/scienceinsider/2010/04/a-skeptic-questions-cancer-genom.html).
5. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 2013;368:2059-74.
6. Druker BJ, Talpaz M, Resta DJ, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med 2001;344:1031-7.
7. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 2001;293:876-80.
8. Sierra JR, Cepero V, Giordano S. Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy. Mol Cancer 2010;9:75.
9. Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med 2009;361:1058-66.
10. Ley TJ, Ding L, Walter MJ, et al. *DNMT3A* mutations in acute myeloid leukemia. N Engl J Med 2010;363:2424-33.
11. Walter MJ, Shen D, Ding L, et al. Clonal architecture of secondary acute myeloid leukemia. N Engl J Med 2012;366:1090-8.
12. Patel JP, Gonen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. N Engl J Med 2012;366:1079-89.

DOI: 10.1056/NEJMe1303816