The advent of cytogenetics and the subsequent development of DNA resequencing technologies have provided seminal insights into the biology of AML and led to evolution in our understanding of AML pathogenesis. Initial genetic insights led to the development of the “two-hit model” of leukemogenesis, which postulates that the genetic events causative of AML occur in 2 broad classes of genes with distinct roles in leukemic transformation. These include mutations in type I genes such as FLT3, JAK2, and RAS, which encode signaling effectors that confer a proliferation and survival advantage to leukemic cells. The second class of mutations are in type II genes (RUNXI, CEBPA, PML/RARA), and mutations in these genes are thought to impair the normal myeloid differentiation cascade from stem cell to mature granulocytes. Recent candidate gene sequencing efforts have challenged this concept by revealing that not all patients with AML have mutations in type I or type II genes, suggesting that additional pathways are involved in AML pathogenesis. For example, alterations in genes involved in epigenetic modifications have been identified in AML, and subsequent research has revealed distinct patterns of epigenomic alteration in AML subtypes with specific somatic mutational spectra.

It has also long been recognized that the outcome of patients with AML is heterogeneous, and that some patients derive benefit from more intensive therapy, including allogeneic SCT, while others do not. Indeed, many studies have shown that cytogenetic alterations and a subset of somatic mutations can be used to predict overall patient outcomes, as well as to guide decisions regarding consolidation therapy. With the advent of next-generation sequencing (NGS) technologies, more disease-related alleles have been revealed in AML patients. Recent efforts

The use of genomic profiling in acute myeloid leukemia (AML) has led to an improved understanding of disease pathogenesis. Genomic profiling has given rise to fundamental observations about the biology of AML and has served to better define clinical outcomes for patients based on somatic mutational status. As additional mutations are identified with a known or postulated role in AML pathogenesis, the challenge ahead will be learning how to integrate these findings into clinical practice in such a way that they have a meaningful impact on patient care and, ultimately, on patient outcomes. Potential goals include using genomic information for refined risk stratification and clinical decision making, and to identify genetic lesions that guide the use of molecularly targeted therapies. The development of next-generation sequencing technologies has made genomic profiling a viable option for use in clinical practice because it can provide robust, high-coverage sequencing data for multiple genes in 1 assay, within a clinically reasonable time frame. The present article discusses recent candidate gene sequencing studies, the development of prognostic models based on these studies, and the current and potential future uses of next-generation sequencing technologies in making treatment decisions for patients with AML.

Semin Hematol 51:298–305. © 2014 Elsevier Inc. All rights reserved.

Integrating Genomics Into Prognostic Models for AML

Matias Sanchez, Ross L. Levine, and Raajit Rampal

The standard treatment for patients with acute myeloid leukemia (AML) has not changed substantially over the past 40 years, with the notable exception of stem cell transplantation (SCT). The use of induction regimens composed of 7-day continuous infusion cytarabine and 3-day daunorubicin was first described in 1973. This initial regimen remained largely unchanged until 2009 when a survival benefit was demonstrated for patients aged <60 years treated with dose-intensified daunorubicin. In this context, it is not surprising that patient outcomes have improved only marginally, much of which can be attributed to improved supportive care. However, the strategies for assignment of AML patients to specific subtypes at diagnosis and stratification to prognostic subgroups are evolving rapidly. In addition, numerous recurrent genetic lesions have been described over the past 20 years, some of which may be amenable to pharmacologic inhibition with emerging therapeutic modalities.

The advent of cytogenetics and the subsequent development of DNA resequencing technologies have provided...
are focused on integrating these findings into clinically useful models to guide therapy and to predict prognosis.

The present article discusses recent candidate gene sequencing studies, the development of prognostic models based on these studies, and the current and potential future uses of NGS technologies in making treatment decisions for patients with AML.

**CONVENTIONAL CYTOGENETICS AND MOLECULAR GENETICS**

The prognostication of AML has historically been based on cytogenetic data, which can be used to separate patients into favorable, intermediate, and unfavorable risk groups. Patients with t(8;21) or inv(16) are classified as having good risk disease, whereas patients with monosomy of chromosomes 5 and 7 or a complex karyotype are considered to have poor risk disease; there is a lack of consensus, however, regarding the full spectrum of cytogenetic alterations that confer an adverse prognosis. Allogeneic SCT has been historically reserved for patients with unfavorable risk and a subset of intermediate-risk patients. However, only 45% of patients have a clonal cytogenetic abnormality, making the choice of optimal consolidation therapy for intermediate-risk patients, most of whom have a normal karyotype, a continuing challenge.

Our understanding of prognostication in intermediate-risk AML changed substantively after a landmark study by Schlenk et al,9 who sought to further refine prognosis for patients with normal karyotype AML (NK-AML) through somatic mutational analysis. The investigators compared the outcome of 800 patients with NK-AML aged <60 years who had been treated with allogeneic SCT or with consolidation chemotherapy. The authors sequenced 5 genes involved in leukemogenesis:FLT3, nucleophosmin 1 (NPM1), CCAAT/enhancer binding protein alpha (CEBPA), the neuroblastoma RAS viral oncogene homologue (NRAS), and the runt-related transcription factor 1 gene (RUNX1). Two groups of patients were identified with a more favorable outcome: patients with mutations in CEBPA and patients with NPM1 mutations without a concurrent FLT3–internal tandem duplication (ITD) mutation. In addition, they demonstrated that allogeneic SCT improved outcome only in patients with FLT3-ITD mutations and in patients with the combination of unmutated NPM1, FLT3, and CEBPA. This was the first report of molecular genetic abnormalities being used for risk stratification of AML patients and has served as the paradigm for subsequent studies in the field.

**IDENTIFICATION OF NOVEL ALLELES IN AML**

The advent of whole-exome and whole-genome sequencing has allowed for the discovery of novel recurrent mutations in AML. In 2008, Ley et al10 reported the whole-genome sequence of a patient with NK-AML, which uncovered 8 new mutations in addition to NPM1 and FLT3-ITD. Notably, these mutations were present in both diagnostic and relapse samples. This study served as proof-of-principle that whole-genome sequencing could be used to interrogate the mutational spectrum of AML and, in a follow-up study, to evaluate clonal evolution of the disease.10,11

Subsequent research using similar approaches led to the discovery of recurrent IDH112 and DNMT3A mutations in AML,5,13 which have been biologically validated. More recent efforts from the Cancer Genome Atlas Research Network using whole-genome/exome sequencing of 200 patients with AML found 23 genes with a higher-than-expected mutational prevalence. Among these were genes known to have a role in AML, such as NPM1 and FLT3, as well as genes encoding factors in novel pathways, including RNA splicing (U2AF1), epigenetic modification (EZH2), and members of the cohesin complex (SMC1A, SMC3), the roles of which are still being evaluated in AML pathogenesis (Figure 1).

Candidate gene sequencing studies have also identified recurrent somatic mutations, some of which were first described in other hematologic malignancies. Mutations in PHF6, a plant homeodomain finger-containing protein, have been reported in 5% of AML patients.14 ASXL1, an epigenetic regulator, is mutated in 5% to 15% of AML patients with AML.15-18 TET2 mutations, initially described in myelodysplastic syndrome and myeloproliferative neoplasms,15,19 have been found to occur in ~10% of AML patients.17

These studies facilitated the discovery of new genes involved in leukemia pathogenesis and can also be used to refine prognosis. Importantly, compared with solid tumors, the mutational burden in AML is significantly lower, with an average of 13 somatic coding mutations per AML genome. Investigation into the relative contributions of mutations (“driver mutations” and “passenger mutations”) remains an important area of investigation. In this regard, Welch et al20 reported on whole-genome sequencing efforts comparing FAB M1 with M3 AML in an effort to uncover potential driver mutations. This study revealed that many mutations present in AML genomes are likely ancestral to leukemia initiation (ie, present in the hematopoietic cell that ultimately gives rise to the leukemic clone) and do not contribute directly to AML pathogenesis. Thus, the mutations that are rate-limiting in leukemic transformation may in fact be limited in scope. Further delineation of the relative contributions to the pathogenesis of variant alleles will require validation in animal models (as described by Perry and Attar in this issue of Seminars) and is an ongoing endeavor.

**PROGNOSTICATION USING INTEGRATED MOLECULAR GENETICS**

As the number of new recurrent mutations reported in the literature has increased over the last few years, the
challenge of how to incorporate these findings into the framework established by Schlenk et al.\(^9\) has become increasingly important. Several groups have performed either single-gene or multigene prognostic analyses, sometimes with conflicting results. Such analyses are often confounded by variability in patient characteristics and treatments and the lack of a validation cohort. Furthermore, many mutations occur at reported frequencies of <10%. As such, each mutation requires large cohort sizes to assess independent prognostic relevance.

In an attempt to integrate genomic findings with clinical variables and cytogenetics to determine the prognostic relevance of multiple genes in a large, homogeneously treated and clinically well-annotated cohort of AML patients, Patel et al.\(^{21}\) performed mutational profiling of 398 patients treated in the Eastern Cooperative Oncology Group E1900 study.\(^{22}\) This was a Phase III study of newly diagnosed AML patients between 17 and 60 years of age randomized to receive induction with daunorubicin at a dose of either 45 mg/m\(^2\) or 90 mg/m\(^2\), both with 7 days of cytarabine. They performed resequencing of 18 genes known to be recurrently mutated in AML in this cohort of patients who were randomized to receive induction chemotherapy with standard-dose or intensified-dose daunorubicin. The study found that in patients with unfavorable as well as favorable cytogenetics, prognosis was not affected by the presence or absence of specific mutations, suggesting that the driver fusion genes govern overall prognosis in favorable risk and adverse risk AML.

Consistent with previous observations, 63% of the patient population had intermediate-risk AML as defined by cytogenetics, and this patient subset demonstrated substantial genetic heterogeneity with 9 different genotypes.\(^{20}\) Mutations in ASXL1 and PHF6, as well as FLT3-ITD and partial tandem duplication of MLL (MLL-PTD) mutations, were associated with reduced survival. Conversely, the presence of IDH2 R140Q mutations was associated with improved survival. This study refined the previously reported outcomes for intermediate-risk patients with NPM1 mutations who lacked FLT3-ITD mutations: specifically, patients with concurrent NPM1 and IDH2/IDH1 mutations had an overall survival of >80% at 2 years, suggesting that this subgroup of patients will have a good outcome with high-dose chemotherapy and without SCT. Patients with NPM1 mutations but without IDH or FLT3-ITD mutations had an intermediate prognosis, suggesting that these patients may still need to be considered for more intense consolidative approaches, such as allogeneic SCT.

Most importantly, this study defined 2 subsets of patients with adverse outcomes based on mutational profiling.\(^{20}\) FLT3-ITD negative patients who had mutant TET2, ASXL1, PHF6, or MLL-PTD had a very poor outcome, consistent with previous reports from single gene studies. In addition to the relatively poor outcome of FLT3-ITD–mutant patients, FLT3-ITD–mutant patients with concurrent mutations in TET2, DNMT3A, an MLL-PTD, or trisomy 8 had worse outcome (3-year overall survival, 14.5%), compared with patients with FLT3-ITD

![Figure 1. Panoramic view of genetic events involved in the pathogenesis of acute myeloid leukemia. Nine different functional categories have been identified for each mutated pathway. Representative recurrently mutated genes in each pathway are depicted. The frequency of mutations detected in acute myeloid leukemia samples for each pathway is provided. TF = transcription factor.](image-url)
mutations without these concurrent mutations (3-year overall survival, 35.2%). Patients with FLT3-ITD and CEBPA mutations had a better outcome, with a 3-year survival of 42%. These data suggest that although FLT3-ITD mutations predict adverse outcome, the prognostic relevance can be further refined with additional mutational data.

These results informed the development of a model based on which treatment decisions can be made (Table 1). Indeed, although convergent data exist regarding the prognostic implications of many of the mutations described in this study, such as FLT3-ITD and DNMT3A, divergent outcomes have been reported for patients with IDH mutations. Several groups have reported that IDH1/2 mutations confer a worse prognosis in patients with NPM1 mutations without concomitant FLT3-ITD, whereas others have reported a favorable prognostic outcome of IDH2 mutations, particularly in the presence of concurrent NPM1 mutations. It is important to note that treatment type, intensity, and age are likely confounding variables from the different studies; most studies in younger adults treated with aggressive therapy found that NPM1/IDH mutant disease is chemosensitive and favorable, whereas studies in older adults suggest a worse prognosis for these elderly patients. As such, further validation in different prospective patient cohorts is needed to confirm these findings.

Several other groups have generated prognostic models by integrating cytogenetic and molecular data. Grossman et al developed a hierarchical model that identified 5 distinct prognostic subgroups: very favorable (PML-RARA, CEBPA double mutation), favorable (t[8;21], inv[16], NPM1 mutant without FLT3-ITD), intermediate (no mutations present in other categories), unfavorable (MLL-PTD, RUNX1 mutant, ASXL1 mutant), and very unfavorable (TP53 mutant). The European LeukemiaNet has published treatment recommendations that take into account molecular and cytogenetic data. Four prognostic groups were identified by this analysis (Table 2). Subsequent research has validated the prognostic significance of these subgroups but has also demonstrated that the percentage of younger (defined as age <60 years) versus older patients differ among these subgroups and may represent a confounding variable. These studies raise the question of how patient age should be incorporated into prognostic models. Recent work by Pastore et al sought to develop a prognostic model for patients with NK-AML for both overall and relapse-free survival by using clinical data, patient characteristics (including age and performance status), and mutational status (CEBPA, NPM1, FLT3-ITD). This research established a scoring system that divides patients with NK-AML into 3 prognostic groups based on the aforementioned factors and can be applied to patients of all ages.

Although it is clear that substantial progress in the ability to predict outcome and to use this information to guide therapy has been achieved over the last 20 years, reconciling the different prognostic models from different studies into a clinically robust approach applicable to patients in routine clinical practice remains a significant challenge. For example, recurrent mutations in splicing

<table>
<thead>
<tr>
<th>Table 1. Revised AML Risk Stratification in Patients and Treatment Recommendations Based on Integrated Mutational Profiling</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cyogenetic Classification</strong></td>
</tr>
<tr>
<td>Favorable Normal karyotype or intermediate-risk cytogenetic</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Unfavorable</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

AML, acute myeloid leukemia; HSCT, hematopoietic stem cell transplantation; ITD, internal tandem duplication; PTD, partial tandem duplication.
factors, epigenetic modifiers, RAS pathway components, and the cohesin complex have been described in secondary AML, de novo AML, and are currently not integrated into the prognostic models described earlier (Table 3). This is further complicated by the continued identification of novel pathogenic mutations of uncertain prognostic significance, heterogeneity in patient treatment across cohorts (which may be expected to increase as more clinical trial data are reported), intrapatient and interpatient heterogeneity, and variability in sequencing technologies used in different studies. Meta-analyses in which robust data from different cohorts are assessed en bloc are needed to develop more robust, universally applicable prognostic models.

### USING MOLECULAR GENETICS TO GUIDE THERAPY

To date, molecular and cytogenetic data have largely been used to define prognosis and to guide the use of allogeneic SCT as a consolidative strategy. However, molecular data are increasingly being used to gauge response to therapy and, most importantly, to guide novel targeted therapy. The initial report from the Eastern Cooperative Oncology Group E1900 study demonstrated that dose-intensified daunorubicin improved outcomes overall in younger adults with AML but did not define specific subsets that benefit versus ones that do not benefit from more intensified induction therapy. However, post hoc genomic analysis of this trial

### Table 2. European LeukemiaNet Risk Stratification and Survival

<table>
<thead>
<tr>
<th>Genetic Group</th>
<th>Cytogenetics</th>
<th>Molecular Genetics</th>
<th>Age &lt; 60 y</th>
<th>Age ≥ 60 y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favorable</td>
<td>t(8;21)</td>
<td>Mutated NPM1 without FLT3-ITD or mutated CEBPA</td>
<td>96% 5.5 11.5</td>
<td>83% 1.1 1.6</td>
</tr>
<tr>
<td></td>
<td>inv(16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t(16;16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal karyotype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate, I</td>
<td>Normal karyotype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t(9;11)</td>
<td>Mutated NPM1 and FLT3-ITD; or, Wildtype NPM1 and FLT3-ITD</td>
<td>79% 1.2 2.1</td>
<td>63% 0.7 0.9</td>
</tr>
<tr>
<td>Intermediate, II</td>
<td>Cytogenetic abnormalities</td>
<td>not classified as favorable or adverse</td>
<td>50% 0.6 0.8</td>
<td>39% 0.5 0.5</td>
</tr>
<tr>
<td>Adverse</td>
<td>Inv(3) or t(3;3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t(6;9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>t(v;11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5 or del(5q)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abn(17p)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complex karyotype</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| CR, complete remission; DFS, disease-free survival; ITD, internal tandem duplication; OS, overall survival; inv, inversion; del, deletion; Abn, abnormality; t, translocation.

### Table 3. Recurrent Alterations in AML Requiring Integration Into Large-Scale Prognostic Models

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Gene</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromatin modifiers</td>
<td>EZH2, KDM6A</td>
<td>35, 39</td>
</tr>
<tr>
<td>Cohesin complex</td>
<td>STAG1, STAG2, SMC1A, SMC3, RAD21</td>
<td>20, 38, 39</td>
</tr>
<tr>
<td>Splicing factors</td>
<td>SRSF2, ZRSF2, SF3B1, SF3A1, U2AF1</td>
<td>33, 34, 39</td>
</tr>
<tr>
<td>AML, acute myeloid leukemia.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
determined that patients who had mutations in NPM1, DNMT3A, or MLL fusions derived benefit from daunorubicin dose intensification, whereas patients with other genotypes did not. This finding suggests that molecular evaluation at baseline might be useful to identify patients most likely to benefit from dose intensification and to identify patients who are unlikely to derive benefit; they can thus be spared the risks of cardiotoxicity and increased myelosuppression.

The challenge will be to extend this to other AML therapies. The hypomethylating agents azacitidine and decitabine are widely used, particularly for older patients or patients deemed to be unfit for standard induction chemotherapy. Anecdotal reports indicate that DNMT3A mutations and TET2 mutations may affect response to decitabine and azacitidine, respectively. Although these findings will need to be confirmed in randomized, prospective studies, they highlight the potential to use genomic profiling to guide initial therapies for patients with AML. The identification and characterization of novel recurrent mutations in AML provide an opportunity to use this information to design rationally targeted therapies for AML. To date, this has been most pronounced with the development of FLT3 inhibitors, several of which have been evaluated in early-phase trials. Novel inhibitors of MLL (NCT01684150), IDH (NCT01915498), and BET (NCT01943851) bromodomains, as well as inhibitors of downstream mediators of the RAS pathway (NCT01449058), are currently in clinical trials for patients with AML. It will be important to ascertain whether the presence or absence of additional mutations affects the response to these targeted therapies, and if complex genotypes can be used to inform combinational therapies for AML patients moving forward.

IMPLEMENTATION OF MASSIVELY PARALLEL SEQUENCING IN THE CLINICAL SETTING

Testing for mutations in NPM1, CEBPA, KIT, and FLT3-ITD has been incorporated into daily clinical practice for risk stratification and is largely considered standard of care. The implementation of analysis of larger gene sets has been limited thus far by: (1) lack of studies validating the use of the newer mutations in large numbers of uniformly treated patients; (2) the lack of incorporation of other clinical variables or biomarkers into molecular-based risk models; (3) the limited number of studies which demonstrate that biomarkers affect therapeutic decisions, including the use of SCT or chemotherapy; and (4) the lack of availability of fast and simple assays to apply genomic profiling for daily clinical use.

As discussed earlier, prognostic models continue to evolve, and the ideal set of molecular markers needed for accurate prognostication in the clinic remains uncertain. The implementation of newer sequencing technologies in the clinical setting also remains a challenge (as reviewed by Graubert and Stone in this issue of the Journal). Use of high-throughput technologies allows the identification of large numbers of genetic alterations that can help predict outcome and serve as potential drug targets. However, the bioinformatics infrastructure and expertise needed to rapidly analyze sequencing data limit the general applicability of whole-genome sequencing in the clinical setting at the current time. Another constraint is the precision and specificity of clinical NGS assays in development. For example, gene size and GC content can hinder the applicability of the assay, and sequencing depth needs to be high enough to ensure that mutations are not missed and artifacts are excluded in the entire coding sequence of all clinically relevant genes.

Many academic institutions have implemented and developed clinical sequencing technologies that are currently used in the clinic. In addition, broader commercial platforms, often involving hundreds of genes, are also gradually being incorporated into the clinic. Most recently, RNA-based platforms have been developed. These techniques allow for high confidence detection of translocations and other genomic rearrangements not easily detected with capture-based sequencing (as reviewed by Mardis in this issue of the Journal). Next-generation Sequencing Standardization Working Group Guidelines have been proposed to deal with the plethora of new sequencing platforms that are emerging, and these will have to be modified to deal with the rapidly evolving exigencies of clinical sequencing.

CONCLUSIONS

The advent of new techniques in molecular genetics and, specifically, of NGS is rapidly expanding our knowledge base of mechanisms of leukemogenesis and the nature of clonal evolution that occurs during the natural history of AML and in response to therapy. Patients with AML can be stratified into different risk subgroups according to genetic data that segregate patients according to biologic, prognostic, and therapeutic relevance. The available technologies provide a new window of opportunity to identify the patient population who will and will not benefit from standard treatments. In addition, the identification of novel mutations continues to inform the development of targeted therapies, many of which are currently in clinical trials.

Important challenges must be overcome to fully realize the utility of molecular genetics to improve clinical practice. The question of how to best utilize molecular data to guide patient therapy is a challenge that is not completely elucidated. Although several prognostic models have been developed to deal with such questions, important limitations on the applicability of these models remain. Furthermore, these models continue to evolve as more pathogenic mutations are uncovered in AML.
The implementation of sequencing technology in the clinic remains fraught with challenges. These relate to the cost of these technologies, appropriate bioinformatics support, and the balance between the urgency to initiate AML therapy and the turnaround time required to generate a report. Indeed, the last problem has important clinical implications, as the identification of mutations at baseline has the potential to not only alter long-term patient management decisions such as whether to use SCT or chemotherapy after remission but, more urgently, it may be used to tailor induction therapy. Moreover, rapid, robust genomic profiling will allow accrual to appropriate clinical trials using targeted therapy most appropriate to the genotype of each patient’s disease. Because successful treatment of adults with AML represents a major medical challenge, the incorporation of molecular genetics in to risk stratification and selection of targeted therapies for these patients could have a significant impact on response to treatment and long-term outcomes.

REFERENCES

26. Boissel N, Nibourel O, Reman O, Renna N, Renneville A, Gardin C, Ramma G, et al. Clinical implications, as the identification of mutations at baseline has the potential to not only alter long-term patient management decisions such as whether to use SCT or chemotherapy after remission but, more urgently, it may be used to tailor induction therapy. Moreover, rapid, robust genomic profiling will allow accrual to appropriate clinical trials using targeted therapy most appropriate to the genotype of each patient’s disease. Because successful treatment of adults with AML represents a major medical challenge, the incorporation of molecular genetics in to risk stratification and selection of targeted therapies for these patients could have a significant impact on response to treatment and long-term outcomes.


