# Clonal Architecture of Secondary Acute Myeloid Leukemia

Matthew J. Walter, M.D., Dong Shen, M.D., Ph.D., Li Ding, Ph.D., Jin Shao, M.S.,
Daniel C. Koboldt, M.S., Ken Chen, Ph.D., David E. Larson, Ph.D.,
Michael D. McLellan, B.S., David Dooling, Ph.D., Rachel Abbott, B.S.,
Robert Fulton, M.S., Vincent Magrini, Ph.D., Heather Schmidt, B.S.,
Joelle Kalicki-Veizer, B.S., Michelle O'Laughlin, B.S., Xian Fan, M.S.,
Marcus Grillot, B.S., Sarah Witowski, M.A., Sharon Heath, John L. Frater, M.D.,
William Eades, Michael Tomasson, M.D., Peter Westervelt, M.D., Ph.D.,
John F. DiPersio, M.D., Ph.D., Daniel C. Link, M.D., Elaine R. Mardis, Ph.D.,
Timothy J. Ley, M.D., Richard K. Wilson, Ph.D., and Timothy A. Graubert, M.D.

## ABSTRACT

From the Department of Internal Medi-
cine, Division of Oncology (M.J.W., J.S.,
M.G., S.W., S.H., W.E., M.T., P.W., J.F.D.,
D.C.L., T.J.L., T.A.G.), the Departments
of Genetics (M.J.W., L.D., E.R.M., T.J.L.,
R.K.W.) and Pathology and Immunology
(J.L.F., T.A.G.), Siteman Cancer Center
(M.J.W., W.E., M.T., P.W., J.F.D., D.C.L.,
E.R.M., T.J.L., R.K.W., T.A.G.), and the
Genome Institute (D.S., L.D., D.C.K.,
K.C., D.E.L., M.D.L., D.D., R.A., R.F.,
V.M., H.S., J.K.-V., M.O., X.F., E.R.M.,
T.J.L., R.K.W.) — all at Washington Uni-
versity, St. Louis. Address reprint requests
to Dr. Graubert at Washington University
School of Medicine, Division of Oncology,
Stem Cell Biology Section, Campus Box
8007, 660 S. Euclid Ave., St. Louis, MO
63110, or at graubert@wustl.edu.

Drs. Walter, Shen, and Ding contributed
equally to this article.

This article (10.1056/NEJMoa1106968) was
published on March 14, 2012, at NEJM.org.

N Engl J Med 2012;366:1090-8.
*Copyright © 2012 Massachusetts Medical Society.*

### BACKGROUND

The myelodysplastic syndromes are a group of hematologic disorders that often evolve into secondary acute myeloid leukemia (AML). The genetic changes that underlie progression from the myelodysplastic syndromes to secondary AML are not well understood.

### METHODS

We performed whole-genome sequencing of seven paired samples of skin and bone marrow in seven subjects with secondary AML to identify somatic mutations specific to secondary AML. We then genotyped a bone marrow sample obtained during the antecedent myelodysplastic-syndrome stage from each subject to determine the presence or absence of the specific somatic mutations. We identified recurrent mutations in coding genes and defined the clonal architecture of each pair of samples from the myelodysplastic-syndrome stage and the secondary-AML stage, using the allele burden of hundreds of mutations.

### RESULTS

Approximately 85% of bone marrow cells were clonal in the myelodysplastic-syndrome and secondary-AML samples, regardless of the myeloblast count. The secondary-AML samples contained mutations in 11 recurrently mutated genes, including 4 genes that have not been previously implicated in the myelodysplastic syndromes or AML. In every case, progression to acute leukemia was defined by the persistence of an antecedent founding clone containing 182 to 660 somatic mutations and the outgrowth or emergence of at least one subclone, harboring dozens to hundreds of new mutations. All founding clones and subclones contained at least one mutation in a coding gene.

### CONCLUSIONS

Nearly all the bone marrow cells in patients with myelodysplastic syndromes and secondary AML are clonally derived. Genetic evolution of secondary AML is a dynamic process shaped by multiple cycles of mutation acquisition and clonal selection. Recurrent gene mutations are found in both founding clones and daughter subclones. (Funded by the National Institutes of Health and others.)

THE MYELODYSPLASTIC SYNDROMES, A HET-erogeneous group of diseases characterized by ineffective hematopoiesis, are the most common cause of acquired bone marrow failure in adults.[1] Secondary acute myeloid leukemia (AML) develops in approximately one third of persons with myelodysplastic syndromes.[2] Clinical discrimination between the myelodysplastic syndromes and secondary AML currently rests predominantly on cytomorphologic analysis, since patients with myelodysplastic syndromes have dysplastic hematopoiesis and a myeloblast count of less than 20%, whereas those with a myeloblast count of 20% or more have AML. Although considerable overlap exists between the spectrum of cytogenetic and molecular lesions seen in the two disorders, there remains uncertainty among patients, insurers, and funding agencies about whether the myelodysplastic syndromes are actually cancers.[3]

The genetic events underlying the progression from cellular dysplasia to cancer have been studied extensively in epithelial tissues (e.g., from the colon and oropharynx),[4-6] but less is known about the genetic progression of the myelodysplastic syndromes to secondary AML. Candidate-gene resequencing has identified several genes that have recurrent mutations during the evolution from the myelodysplastic syndromes to secondary AML (e.g., *FLT3*, *NPM1*, *RUNX1*, *TP53*, and *NRAS*),[7-15] but our understanding of the total number and clonal distribution of mutations in this disease is limited. We used whole-genome sequencing to discover somatic mutations in bone marrow samples obtained from seven subjects with secondary AML and determined whether these mutations were present in paired samples obtained during the antecedent myelodysplastic-syndrome stage. We used this information to define the proportion of clonal cells and genetic architecture at the time of diagnosis of a myelodysplastic syndrome and progression to secondary AML.

## METHODS

Methods are described in detail in the Supplementary Appendix, available with the full text of this article at NEJM.org. In brief, bone marrow biopsy specimens were obtained from seven subjects, all of whom provided written informed consent on a form that contained specific language authorizing whole-genome sequencing. Myeloblasts were enriched by means of cell sorting. Paired-end DNA libraries from bone marrow samples (for secondary AML) and normal skin samples were sequenced on the Illumina HiSeq 2000 and Genome Analyzer IIx. Aligned reads were analyzed to detect putative somatic mutations, as described previously.[16] Somatic-mutation predictions were validated with the use of solid-phase capture, followed by deep sequencing of skin, myelodysplastic-syndrome, and secondary-AML samples. Mutations in one subject (UPN266395) have been reported previously.[17] Array-based profiling of DNA copy number and gene expression were performed, as described previously.[18] Sequence and microarray data have been deposited in the database of Genotypes and Phenotypes (dbGaP) of the National Center for Biotechnology Information (accession number, phs000159.v3.p2).

## RESULTS

### WHOLE-GENOME SEQUENCING AND CAPTURE VALIDATION

We performed whole-genome sequencing using paired-end reads generated from DNA libraries prepared from secondary-AML samples and matched skin samples from seven subjects with antecedent de novo myelodysplastic syndromes (Table 1). Diploid genome coverage was more than 95% for the secondary-AML samples and 97% for the skin samples (Table 1 in the Supplementary Appendix). Putative somatic mutations were called and prioritized into nonoverlapping tiers, as described previously.[16] To validate somatic mutations and measure mutant allele frequencies, we designed custom solid-phase long-oligonucleotide arrays for each subject and used them to capture regions of the genome containing putative single-nucleotide variants (SNVs, or point mutations) and insertions or deletions (indels). For each subject, a trio of samples was analyzed (normal skin, bone marrow obtained during the antecedent myelodysplastic-syndrome stage, and bone marrow obtained during the secondary-AML stage). The captured DNA was sequenced to provide an average of 640 reads at the sites of validated mutations (Table 1 in the Supplementary Appendix). In each genome, we validated 304 to 872 somatic point mutations and 0 to 2 indels in tier 1 (i.e., the tier consisting of changes in the amino acid coding regions of annotated exons, consensus splice-site regions, and RNA genes) (Tables 2 and 3 in the Supplementary Appendix). The enrichment of myeloblasts by

The New England Journal of Medicine

**Table 1. Characteristics of Seven Subjects with Myelodysplastic Syndromes (MDS) and Progression to Secondary Acute Myeloid Leukemia (sAML).**

| UPN* | Sex | Age at Diagnosis | IPSS Score† | Bone Marrow Myeloblasts | | Cytogenetic Analysis | | MDS to sAML‡ | Survival since MDS Diagnosis | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | MDS | sAML | MDS | sAML | | Disease-free | Overall |
| | | yr | | % | | | | | months | |
| 461282 | M | 65 | 2.0 | 6 | 69 | Complex | Complex | 22 | 58 | 67 |
| 667720 | F | 66 | 1.0 | 10 | 29§ | Normal | Normal | 22 | 21 | 24 |
| 610184 | F | 45 | 1.0 | 0 | 38 | Complex | Complex | 5 | 10 | 20 |
| 182896 | M | 75 | 1.0 | 0 | 51 | Complex | Complex | 32 | 35 | 40 |
| 266395 | M | 65 | 1.0 | 7 | 66 | Normal | Normal | 1 | 3 | 11 |
| 288033 | F | 31 | 1.0 | 13 | 43 | Normal | Normal | 1 | 1 | 12 |
| 298273 | M | 26 | 2.0 | 18 | 35 | Normal | Normal | 4 | 4 | 18 |

* UPN denotes unique patient number.
† The International Prognostic Scoring System (IPSS) ranges from 0 to 3.5, with higher scores indicating a worse prognosis.
‡ Values are the intervals between tissue banking of MDS samples and sAML samples.
§ This analysis was performed by means of flow cytometry.

means of cell sorting had no major effect on measurement of the mutant allele burden (Fig. 1 in the Supplementary Appendix).

### RECURRENT GENE MUTATIONS

We first focused on tier 1 mutations with predicted translational consequences (i.e., whole-gene deletions, indels, missense, nonsense, frameshift, or splice-site mutations). There were 17 to 32 validated point mutations or indels (mean, 24) per secondary-AML genome in 168 genes among the seven samples (Table 3 in the Supplementary Appendix). Most of these genes did not have recurrent mutations, suggesting that many of the mutations were randomly acquired and not causally related to the pathogenesis of the myelodysplastic syndromes. Two recurrently mutated genes were detected in two samples each: loss-of-function mutations in a known myeloid-tumor suppressor, *RUNX1*, and two somatic missense mutations in *UMODL1* (Table 2). *UMODL1* was recently reported to be mutated in patients with multiple myeloma and those with ovarian cancer.[19,20] *UMODL1* messenger RNA (mRNA) was expressed in normal CD34+ progenitor cells and in secondary-AML cells from our seven subjects, and the two mutations occurred in regions of *UMODL1* encoding conserved domains (T533P in a calcium-binding epidermal-growth-factor–like domain and V882M in a sea-urchin sperm protein, enterokinase, and agrin [SEA] domain).[21]

To extend these results, we compared the mutations with translational consequences in myelodysplastic-syndrome and secondary AML samples with mutations identified in 200 de novo AML samples (50 subjected to whole-genome sequencing and 150 to whole-exome sequencing). We identified 10 genes that were mutated in one secondary-AML sample and in at least 3 of 200 AML samples (>1%) (Table 2). Seven of these genes are known to have recurrent mutations in AML. Four genes with recurrent mutations (including *UMODL1*) have not previously been implicated in the myelodysplastic syndromes or AML. A specific codon in *U2AF1* harbored missense mutations in multiple AML samples, suggesting that these mutations may cause gain of function. Supporting this hypothesis is our recent report[17] that the S34F substitution in *U2AF1* (affected by a missense mutation) enhances alternative mRNA splicing in vitro. The X-chromosome gene *STAG2* was also recurrently mutated. All *STAG2* mutations that we observed in this study are predicted to cause protein truncation (H738fs in secondary AML and all 4 nonsense or frameshift mutations in AML). In addition, *STAG2* is deleted in AML and other cancers.[22,23] Taken together, these results suggest that *STAG2* loss of function probably contributes to the pathogenesis of the myelodysplastic syndromes and AML. Although *STAG2* inactivation was recently reported to cause aneuploidy in glioblastoma and

1092

N ENGL J MED 366;12 NEJM.ORG MARCH 22, 2012

The New England Journal of Medicine
Downloaded from nejm.org on December 31, 2012. For personal use only. No other uses without permission.
Copyright © 2012 Massachusetts Medical Society. All rights reserved.

colorectal-cancer cell lines,[23] the subjects with *STAG2* mutations in our study had normal karyotypes.

Singleton mutations (i.e., mutations that are not recurrent) may also be important in the pathogenesis of the myelodysplastic syndromes. Overall, the nonrecurrent tier 1 mutations implicate mutant proteins in 1 of 11 biologic pathways[24] that are relevant for cancer pathogenesis, and nearly all these pathways were affected in all seven samples (Fig. 2 in the Supplementary Appendix). Tier 1 mutations were not substantially enriched in any single pathway (data not shown).[24]

### CLONALITY OF MYELODYSPLASTIC SYNDROMES AND SECONDARY AML

The clinicopathological diagnosis of the myelodysplastic syndromes requires a finding of less than 20% myeloblasts (morphologically defined malignant cells) in the bone marrow. However, data from single-nucleotide-polymorphism (SNP) array and single-gene resequencing studies have suggested that the clonal population of cells in the myelodysplastic syndromes can be greater than 20%.[18,25,26] In this study, we used capture sequencing data to accurately measure the prevalence of clonal cells in myelodysplastic-syndrome samples. We calculated the mutant allele frequencies for all validated somatic SNVs (adjusted for chromosome copy number) and performed an unsupervised clustering analysis to define mutation clusters.[27] Each myelodysplastic-syndrome and secondary-AML genome contained a founding clone of cells, defined as the mutation cluster (containing 182 to 660 somatic SNVs) that had the highest mutant allele burden at the myelodysplastic-syndrome and secondary-AML stages (cluster 1 in Fig. 1A, and Fig. 5 in the Supplementary Appendix). We estimated maximum tumor clonality as twice the average mutant allele frequency of the founding clone (since all mutations were present once per diploid genome after copy-number correction). In all cases, the majority of unfractionated bone marrow cells (up to 92.7%) in the myelodysplastic-syndrome samples were clonal and indistinguishable from the cells in the secondary-AML samples, even with a myeloblast count of less than 5% (Fig. 1B). Although these results confirm previous observations that the myeloblast count can underestimate the size of the clonal population in myelodysplastic syndromes,[18,25,26] they also suggest that when the entire genome is interrogated for mutations, clon-

**Table 2. Genes with Recurrent Mutations in De Novo AML Detected in MDS and sAML Genomes.**

| Gene and Mutation* | Present in MDS Sample | Present in sAML Sample |
|---|---|---|
| *CDH23* (1235insL)† | No | Yes |
| *NPM1* (W288fs)‡ | Yes | Yes |
| *PTPN11* (G60R)‡ | No | Yes |
| *RUNX1* | | |
|   G170fs | No | Yes |
|   del(21) (q22.11) | Yes | Yes |
| *SMC3* (e8-1 splice)† | Yes | Yes |
| *STAG2* (H738fs) | Yes | Yes |
| *TP53* (V272M) | Yes | Yes |
| *U2AF1* (S34F)‡ | Yes | Yes |
| *UMODL1*† | | |
|   T533P | Yes | Yes |
|   V882M | Yes | Yes |
| *WT1* (D436E) | No | Yes |
| *ZSWIM4* (P818A)† | Yes | Yes |

\* The mutation is listed according to the amino acid position. The abbreviation sAML denotes secondary AML.
† This mutation has not previously been reported in MDS or AML.
‡ The same codon was affected in more than one subject.

al hematopoiesis involving most of the bone marrow appears to be the rule even in early-stage myelodysplastic syndromes.

To assess the clonality of copy-number alterations, we designed capture probes spanning heterozygous SNPs that were present in the skin samples and deleted or retained in the bone marrow samples on the basis of SNP arrays and data from whole-genome sequencing (Table 4 in the Supplementary Appendix). We determined the proportion of read counts containing alleles present in the skin samples that were retained in the myelodysplastic-syndrome and secondary-AML samples. The prevalence of the retained alleles in the skin samples was approximately 50%, as expected for heterozygous SNPs. The prevalence of these alleles in the myelodysplastic-syndrome and secondary-AML samples diverged from 50% and was proportionate to the percentage of cells harboring a deletion (Table 5 in the Supplementary Appendix). Clustering of the data regarding copy-number alterations largely recapitulated the SNV clusters for these subjects (Fig. 1C, and Fig. 7 in the Supplementary Appendix). However, unique clones (i.e.,
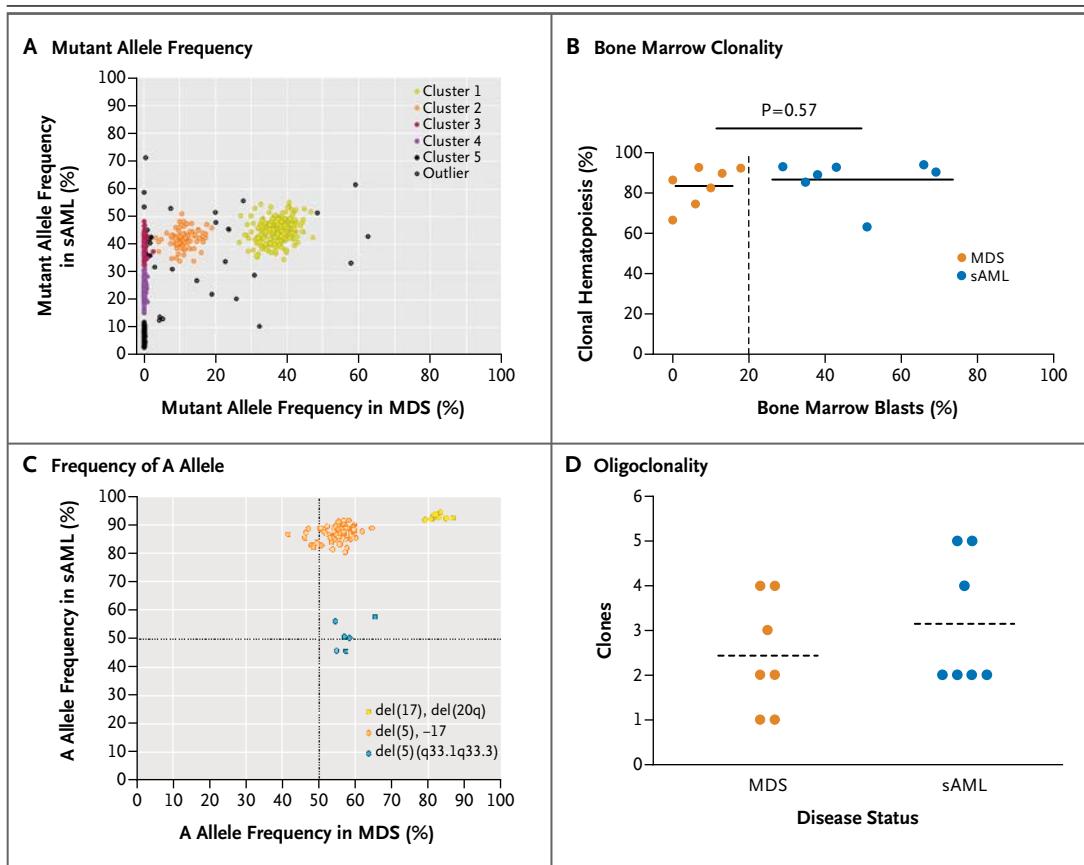
**Figure 1. Oligoclonality of Genomes for Myelodysplastic Syndrome (MDS) and Secondary Acute Myeloid Leukemia (sAML) in One Subject.**

In Panel A, the allele frequencies of all validated mutations in Subject UPN461282 are shown at the MDS stage and after progression to sAML. Mutant allele frequencies were adjusted for chromosomal copy number. Unsupervised clustering of individual mutations identified five distinct mutation clusters. Panel B shows similar total bone marrow clonality (the proportion of cells containing the founding mutations) in MDS and sAML (P=0.57), despite myeloblast counts ranging from 0 to 69%. In Panel C, the frequency of reads supporting the retained (A) allele in regions of copy-number alteration is shown in MDS and sAML. Cells containing two clusters of clonal copy-number changes increase in number between MDS and sAML (yellow and orange data points), whereas a subclone containing a distal 5(q) deletion is lost during progression from MDS to sAML (blue data points). The founding clone contains del(17), del(20q), and cluster 1 single-nucleotide variants (SNVs) (yellow data points in Panels A and C), whereas a subclone also contains del(5), monosomy 17, and cluster 2 SNVs (orange data points in Panels A and C). Dashed lines indicate the expected A allele frequency for heterozygous SNPs that are not in deleted segments. Panel D shows that MDS and sAML are oligoclonal, with a mean of 2.4 clones in MDS and 3.1 in sAML (P=0.047). The abbreviation del(17) denotes telomeric chromosome 17 deletions, del(20q) denotes del(20)(q11.21q13.13), and del(5) denotes interstitial deletions on the short and long arms of chromosome 5.

unique populations of cells defined by the mutation clusters they contained) were identified by both approaches, suggesting that these types of data provide complementary views of clonal evolution. Our methods do not capture changes in DNA methylation, which are known to be present in myelodysplastic-syndrome genomes,[28] suggesting that there may be additional layers of clonal complexity that we have not detected.

All seven samples were oligoclonal at the sec-ondary-AML stage and were monoclonal in only two cases at the myelodysplastic-syndrome stage (Fig. 1D). In the first subject (UPN461282), we identified five distinct mutation clusters that collectively defined the clonal genetic progression of this tumor (Fig. 1A). Two clones were present in the myelodysplastic-syndrome sample. A single cell in clone 2, which contained cluster 1 and 2 mutations, moved forward and by acquisition of cluster 3, 4, and 5 mutations generated three new

clones that were present only in the secondary-AML sample. Collectively, the data suggest that the evolution of this tumor proceeded sequentially from cells with cluster 1 mutations to cells containing cluster 1 to 5 mutations, with each new clone carrying forward all the preexisting pathogenic and nonpathogenic mutations (Fig. 2A). The other six tumors also conformed to linear models of clonal evolution (Fig. 5 in the Supplementary Appendix), although analysis at the single-cell level may reveal more complex patterns in some cases.

The secondary-AML genomes contained 304 to 872 somatic SNVs (Table 2 in the Supplementary Appendix). It is likely that most of these mutations are irrelevant to the pathogenesis of the disease. Consistent with this notion is the observation that the number of somatic SNVs per genome tier tended to be proportionate to the tier size (as expected by chance), suggesting that the vast majority of mutations were random background mutations (Fig. 3A in the Supplementary Appendix). We observed that most of the mutations in each secondary-AML sample were present in the paired myelodysplastic-syndrome sample, and there was a subset of secondary-AML–specific mutations. As expected, the proportion of secondary-AML–specific mutations was smaller in the four subjects who had rapid progression to secondary AML (<6 months) than in the three subjects with slower progression (>20 months). An average of 6.7% of all mutations were specific to secondary AML in the subjects with rapid progression, as compared with 37.8% of secondary-AML–specific mutations in the subjects with slow progression (P<0.05) (Fig. 3B in the Supplementary Appendix). The spectrum of transition and transversion mutations in the myelodysplastic-syndrome founding clone was similar in all seven subjects (Fig. 4 in the Supplementary Appendix). The most common substitution was a C·G→T·A transition, as seen in other cancer genomes.[29,30] Two subjects (who had been treated with decitabine for 4 to 11 months, after the diagnosis of a myelodysplastic syndrome and before the progression to secondary AML) had a significant increase in the frequency of secondary-AML–specific C→G transversions in their secondary-AML samples, whereas two subjects who were not treated with decitabine did not have a similar increase (Fig. 4 in the Supplementary Appendix). We could not identify a sequence-specific context for these transversions.
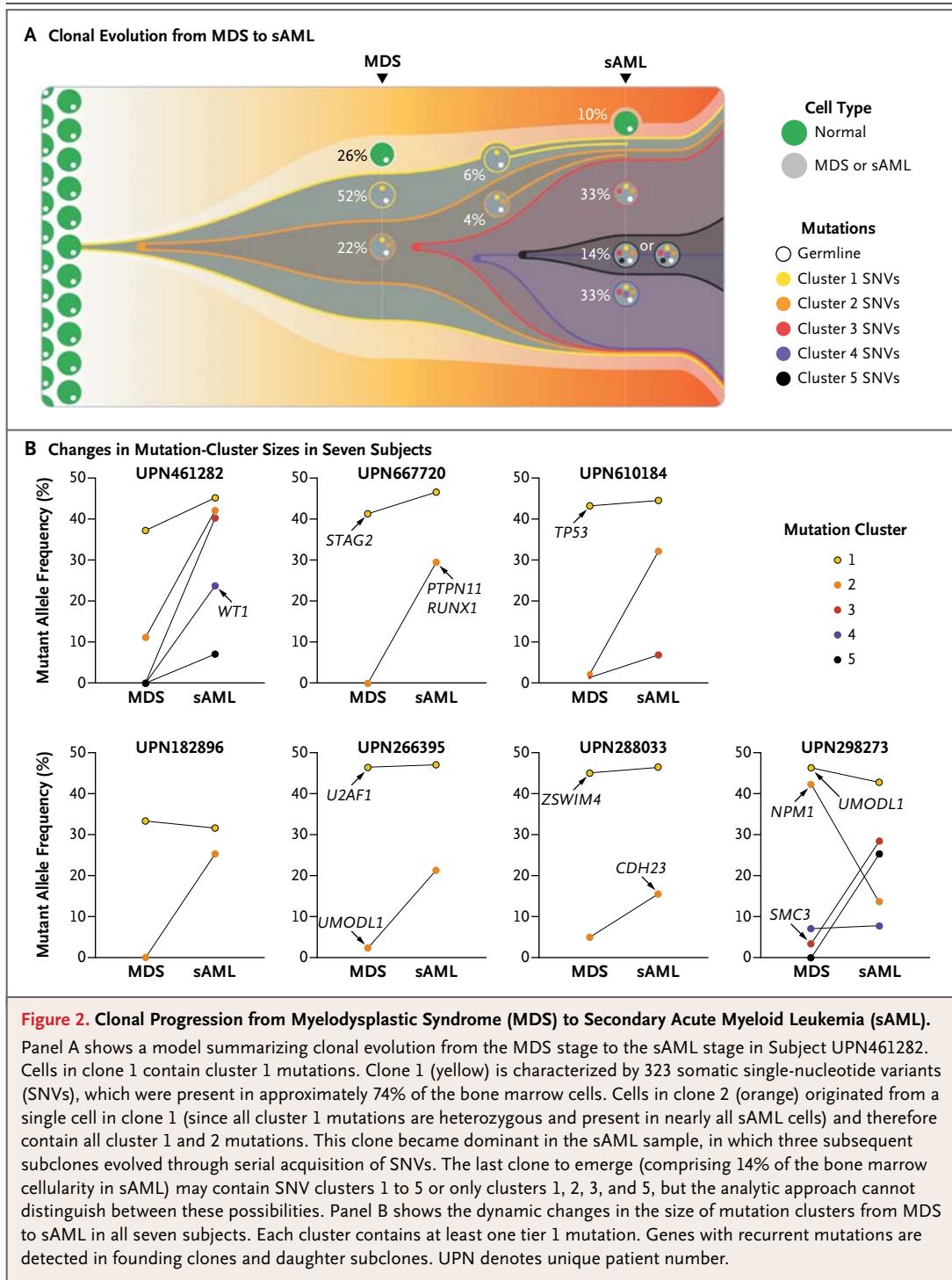
Even though the mutations driving clonal outgrowth at each stage are not known, the founding clone in the seven myelodysplastic-syndrome samples (containing 182 to 660 mutations) persisted in all seven secondary-AML samples and included at least one tier 1 mutation. The founding clones acquired at least one new tier 1 mutation with predicted translational consequences during the generation of secondary-AML–specific clusters (Fig. 5 in the Supplementary Appendix). Genes with recurrent mutations were detected in both founding clones and daughter subclones (Fig. 2B).

## DISCUSSION

Using next-generation sequencing, we have found that the proportion of neoplastic bone marrow cells is indistinguishable in myelodysplastic-syndrome and secondary-AML samples, suggesting that the myelodysplastic syndromes are as clonal as secondary AML, even with a myeloblast count of zero. Although clonality is not sufficient to define malignant transformation, it is a cardinal manifestation of most human cancers, and our findings suggest that the myelodysplastic syndromes and secondary AML are both highly clonal hematologic cancers.[31] Analysis of the proportion of mutant cells in samples obtained from the same subject before and after progression to secondary AML allowed us to compare the clonal architecture and genes that were mutated in order to gain insight into the genetics of these diseases. Robust detection of mutation clusters was possible because hundreds of mutations per genome were identified by whole-genome sequencing, and the allele burdens were quantified by deep resequencing at two time points. If we had analyzed only the secondary-AML samples or only the tier 1 (i.e., exomic) variants, this complexity could not have been elucidated. In the samples from all seven subjects, the secondary-AML genomes were oligoclonal. The preexisting myelodysplastic-syndrome founding clone always persisted in secondary AML, although it was outcompeted by daughter subclones in some cases. With the acquisition of each new set of mutations, all the preexisting mutations were carried forward, resulting in subclones that contained increasing numbers of mutations during evolution. On the basis of our experimental design, we cannot exclude the possibility that there were additional subclones in the myelodysplastic-syndrome samples that were not present in the secondary-AML samples, and one genome (UPN298273) suggests that this could occur.

A unique aspect of the biology of leukemia is that hematopoietic cells freely mix and recirculate

**Figure 2. Clonal Progression from Myelodysplastic Syndrome (MDS) to Secondary Acute Myeloid Leukemia (sAML).**

Panel A shows a model summarizing clonal evolution from the MDS stage to the sAML stage in Subject UPN461282. Cells in clone 1 contain cluster 1 mutations. Clone 1 (yellow) is characterized by 323 somatic single-nucleotide variants (SNVs), which were present in approximately 74% of the bone marrow cells. Cells in clone 2 (orange) originated from a single cell in clone 1 (since all cluster 1 mutations are heterozygous and present in nearly all sAML cells) and therefore contain all cluster 1 and 2 mutations. This clone became dominant in the sAML sample, in which three subsequent subclones evolved through serial acquisition of SNVs. The last clone to emerge (comprising 14% of the bone marrow cellularity in sAML) may contain SNV clusters 1 to 5 or only clusters 1, 2, 3, and 5, but the analytic approach cannot distinguish between these possibilities. Panel B shows the dynamic changes in the size of mutation clusters from MDS to sAML in all seven subjects. Each cluster contains at least one tier 1 mutation. Genes with recurrent mutations are detected in founding clones and daughter subclones. UPN denotes unique patient number.

between the peripheral blood and the bone marrow. Clones that persist and grow over time must retain the capacity for self-renewal. Mutations in new clones must confer a growth advantage for them to successfully compete with ancestral clones. The result is that these secondary-AML samples are not monoclonal but are instead a mosaic of several genomes with unique sets of mutations; this mosaic is shaped by the acquisition of serial mutations and clonal diversification. Simi-

larly, recent analysis of de novo AML samples with the use of whole-genome sequencing showed that relapse after chemotherapy is associated with clonal evolution and acquisition of new mutations.[32] Analysis of individual cancer cells may reveal additional layers of genetic complexity. Recent studies of B-cell acute lymphoblastic leukemia have shown that serial acquisition of cytogenetic abnormalities in that disease most often occurs through a branching hierarchy and only rarely follows a simple linear path.[33,34] Extending this work to include the full complement of mutations discovered by whole-genome sequencing will be a major goal for future studies of cancer genetics.

Our study has several clinical implications. First, the distinction between the myelodysplastic syndromes and secondary AML currently relies on manual enumeration of bone marrow myeloblasts, a standard that is subject to interobserver bias but nonetheless drives major decisions about treatment for patients with small differences in myeloblast counts. Ultimately, identifying the patterns of pathogenic mutations and their clonality in bone marrow samples from patients with myelodysplastic syndromes should lead to greater diagnostic certainty and improved prognostic algorithms. On this note, two of the subjects in our study had progression from myelodysplastic syndromes to secondary AML in 1 month. This progression was based on an increase in the myeloblast count from 7% to 66% in one subject and from 13% to 43% in the other, despite an absence of change in the number of clones (two in each case) and only minor increases in point mutations (<2% were gained

during progression to secondary AML in these cases) when the same specimens were analyzed by means of next-generation sequencing. Second, our finding that the dominant secondary-AML clone was derived from a myelodysplastic-syndrome founding clone in all cases suggests that therapies targeted to these early mutations might be the most effective strategy for eliminating disease-propagating cells and improving the rate of response to traditional chemotherapy for patients with secondary AML.[35,36] Finally, it is possible that disease progression in patients with myelodysplastic syndromes is driven not only by the presence of recurrent mutations, which have recently been shown to have prognostic value,[37] but also by the clone (i.e., founding vs. daughter) in which they arise. Coupling genotyping of myelodysplastic-syndrome samples for prognostically important mutations with analysis of the clonal architecture may yield more informative biomarkers and a better understanding of the pathogenesis of the myelodysplastic syndromes.

**REFERENCES**

1. Cogle CR, Craig BM, Rollison DE, List AF. Incidence of the myelodysplastic syndromes using a novel claims-based algorithm: high number of uncaptured cases by cancer registries. Blood 2011;117:7121-5.
2. Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood 1997;89:2079-88. [Erratum, Blood 1998;91:1100.]
3. Steensma DP. Are myelodysplastic syndromes "cancer"? Unexpected adverse consequences of linguistic ambiguity. Leuk Res 2006;30:1227-33.
4. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-32.
5. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759-67.

6. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996;56:2488-92.
7. Shih LY, Huang CF, Wang PN, et al. Acquisition of FLT3 or N-ras mutations is frequently associated with progression of myelodysplastic syndrome to acute myeloid leukemia. Leukemia 2004;18:466-75.
8. Schnittger S, Bacher U, Haferlach C, et al. Characterization of NPM1-mutated AML with a history of myelodysplastic syndromes or myeloproliferative neoplasms. Leukemia 2011;25:615-21.
9. Chen CY, Lin LI, Tang JL, et al. RUNX1 gene mutation in primary myelodysplastic syndrome — the mutation can be detected early at diagnosis or acquired during disease progression and is associated with poor outcome. Br J Haematol 2007;139:405-14.

10. Lai JL, Preudhomme C, Zandecki M, et al. Myelodysplastic syndromes and acute myeloid leukemia with 17p deletion: an entity characterized by specific dysgranulopoiesis and a high incidence of P53 mutations. Leukemia 1995;9:370-81.
11. Jasek M, Gondek LP, Bejanyan N, et al. TP53 mutations in myeloid malignancies are either homozygous or hemizygous due to copy number-neutral loss of heterozygosity or deletion of 17p. Leukemia 2010;24:216-9.
12. van Kamp H, de Pijper C, Verlaan-de Vries M, et al. Longitudinal analysis of point mutations of the N-ras proto-oncogene in patients with myelodysplasia using archived blood smears. Blood 1992;79:1266-70.
13. Horiike S, Misawa S, Nakai H, et al. N-ras mutation and karyotypic evolution are closely associated with leukemic trans-

formation in myelodysplastic syndrome. Leukemia 1994;8:1331-6.

14. Bacher U, Haferlach T, Kern W, Haferlach C, Schnittger S. A comparative study of molecular mutations in 381 patients with myelodysplastic syndrome and in 4130 patients with acute myeloid leukemia. Haematologica 2007;92:744-52.

15. Dicker F, Haferlach C, Sundermann J, et al. Mutation analysis for RUNX1, MLL-PTD, FLT3-ITD, NPM1 and NRAS in 269 patients with MDS or secondary AML. Leukemia 2010;24:1528-32.

16. 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.

17. Graubert TA, Shen D, Ding L, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet 2011;44:53-7.

18. Graubert TA, Payton MA, Shao J, et al. Integrated genomic analysis implicates haploinsufficiency of multiple chromosome 5q31.2 genes in de novo myelodysplastic syndromes pathogenesis. PLoS One 2009; 4(2):e4583.

19. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474:609-15.

20. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. Nature 2011; 471:467-72.

21. Di Schiavi E, Riano E, Heye B, Bazzicalupo P, Rugarli EI. UMODL1/Olfactorin is an extracellular membrane-bound molecule with a restricted spatial expression in olfactory and vomeronasal neurons. Eur J Neurosci 2005;21:3291-300.

22. Walter MJ, Payton JE, Ries RE, et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. Proc Natl Acad Sci U S A 2009;106:12950-5.

23. Solomon DA, Kim T, Diaz-Martinez LA, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 2011;333:1039-43.

24. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4(5):P3.

25. Heinrichs S, Kulkarni RV, Bueso-Ramos CE, et al. Accurate detection of uniparental disomy and microdeletions by SNP array analysis in myelodysplastic syndromes with normal cytogenetics. Leukemia 2009;23: 1605-13.

26. Walter MJ, Ding L, Shen D, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia 2011;25:1153-8.

27. Fraley C, Raftery A. Model-based clustering, discriminant analysis and density estimation. J Am Stat Assoc 2002;97:611-31.

28. Figueroa ME, Skrabanek L, Li Y, et al. MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. Blood 2009;114:3448-58.

29. Greenman C, Stephens P, Smith R, et al. Patterns of somatic mutation in human cancer genomes. Nature 2007;446: 153-8.

30. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature 2010; 464:999-1005.

31. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144:646-74.

32. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 2012;481:506-10.

33. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature 2011;469:356-61.

34. Notta F, Mullighan CG, Wang JC, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. Nature 2011;469:362-7. [Erratum, Nature 2011; 471:254.]

35. Larson RA. Is secondary leukemia an independent poor prognostic factor in acute myeloid leukemia? Best Pract Res Clin Haematol 2007;20:29-37.

36. Bello C, Yu D, Komrokji RS, et al. Outcomes after induction chemotherapy in patients with acute myeloid leukemia arising from myelodysplastic syndrome. Cancer 2011;117:1463-9.

37. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med 2011;364:2496-506.

*Copyright © 2012 Massachusetts Medical Society.*

---

**NEJM 200TH ANNIVERSARY INTERACTIVE TIMELINE**
Explore a special 200th Anniversary interactive timeline,
History of Medical Discoveries, 1812–2012, at the NEJM 200th Anniversary
website, **http://NEJM200.NEJM.org**. Each image is linked to historical
information and to articles published in the *Journal*.

---