

## MYELOID NEOPLASIA

## Acute myeloid leukemia ontogeny is defined by distinct somatic mutations

R. Coleman Lindsley,<sup>1</sup> Brenton G. Mar,<sup>2</sup> Emanuele Mazzola,<sup>3</sup> Peter V. Grauman,<sup>4</sup> Sarah Shareef,<sup>4</sup> Steven L. Allen,<sup>5</sup> Arnaud Pigneux,<sup>6</sup> Meir Wetzler,<sup>7</sup> Robert K. Stuart,<sup>8</sup> Harry P. Erba,<sup>9</sup> Lloyd E. Damon,<sup>10</sup> Bayard L. Powell,<sup>11</sup> Neal Lindeman,<sup>12</sup> David P. Steensma,<sup>1</sup> Martha Wadleigh,<sup>1</sup> Daniel J. DeAngelo,<sup>1</sup> Donna Neuberg,<sup>3</sup> Richard M. Stone,<sup>1</sup> and Benjamin L. Ebert<sup>4</sup>

<sup>1</sup>Department of Medical Oncology, Division of Hematological Malignancies, <sup>2</sup>Department of Pediatric Oncology, and <sup>3</sup>Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston MA; <sup>4</sup>Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; <sup>5</sup>Hofstra North Shore-LIJ School of Medicine, Hempstead, NY; <sup>6</sup>Hôpital Haut-Leveque, Centre Hospitalier Universitaire Bordeaux, Pessac, France; <sup>7</sup>Roswell Park Cancer Institute, Buffalo, NY; <sup>8</sup>Hollings Cancer Center, Medical University of South Carolina, Charleston, SC; <sup>9</sup>Division of Hematology and Oncology, University of Alabama, Birmingham, Birmingham, AL; <sup>10</sup>Helen Diller Family Comprehensive Cancer Center, The University of California, San Francisco, CA; <sup>11</sup>Comprehensive Cancer Center of Wake Forest University, Winston Salem, NC; and <sup>12</sup>Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

## Key Points

- The presence of a mutation in *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *ASXL1*, *EZH2*, *BCOR*, or *STAG2* is highly specific for secondary AML.
- Secondary-type mutations define an s-AML-like disease within t-AML and elderly de novo AML that underlies clinical heterogeneity.

**Acute myeloid leukemia (AML) can develop after an antecedent myeloid malignancy (secondary AML [s-AML]), after leukemogenic therapy (therapy-related AML [t-AML]), or without an identifiable prodrome or known exposure (de novo AML). The genetic basis of these distinct pathways of AML development has not been determined. We performed targeted mutational analysis of 194 patients with rigorously defined s-AML or t-AML and 105 unselected AML patients. The presence of a mutation in *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *ASXL1*, *EZH2*, *BCOR*, or *STAG2* was >95% specific for the diagnosis of s-AML. Analysis of serial samples from individual patients revealed that these mutations occur early in leukemogenesis and often persist in clonal remissions. In t-AML and elderly de novo AML populations, these alterations define a distinct genetic subtype that shares clinicopathologic properties with clinically confirmed s-AML and highlights a subset of patients with worse clinical outcomes, including a lower complete remission rate, more frequent re-induction, and decreased event-free survival. This trial was registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) as #NCT00715637. (*Blood*. 2015;125(9):1367-1376)**

## Introduction

Acute myeloid leukemia (AML) is a biologically heterogeneous disease that can be classified into 3 distinct categories based on clinical ontogeny: secondary AML (s-AML) represents transformation of an antecedent diagnosis of myelodysplastic syndrome (MDS) or myeloproliferative neoplasm (MPN), therapy-related AML (t-AML) develops as a late complication in patients with prior exposure to leukemogenic therapies, and de novo AML arises in the absence of an identified exposure or prodromal stem cell disorder. It is not known whether distinct somatic genetic lesions drive these different disease subtypes or whether ontogeny-defining mutations underlie relative differences in treatment outcomes.<sup>1-3</sup>

A central goal for the study of AML, and cancer more generally, is to elucidate organizing genetic principles that govern initiation and progression of the disease, and to link these genetic principles to clinical phenotype. Large-scale sequencing studies have revealed the remarkable complexity of genetic alterations that drive the pathogenesis of myeloid malignancies, including de novo AML, MDS, and MPN, as well as clonal diversity within individual

patients.<sup>4-10</sup> It is not clear at present whether the diverse genetic lesions in AML can be organized into a framework that reflects and informs our understanding of disease biology and development and could potentially be used to guide therapy and improve prognostic accuracy.

To investigate the genetic basis of AML ontogeny, we studied a cohort of s-AML and t-AML patients enrolled in a recent phase 3 clinical trial that represented the largest prospective evaluation of the role of induction therapy in these patient populations. Although no significant difference in outcome was observed between treatment arms (amionafide plus cytarabine vs daunorubicin plus cytarabine), the intrinsic therapy-resistance and prognostic adversity of s-AML and t-AML was confirmed, with an overall complete remission (CR) rate of 45% and median overall survival (OS) of 7 months.<sup>11</sup> The rigorous eligibility criteria, uniform treatment, and prospective data collection on this trial afforded a unique opportunity to evaluate the distinctive genetics and clinical associations of these high-risk and understudied leukemia subtypes.

Submitted November 5, 2014; accepted December 12, 2014. Prepublished online as *Blood* First Edition paper, December 30, 2014; DOI 10.1182/blood-2014-11-610543.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2015 by The American Society of Hematology

## Methods

### Patient samples

A total of 433 patients with s-AML ( $n = 216$ ) or t-AML ( $n = 217$ ) who were enrolled in the ACCEDE trial were considered for inclusion in this study. Patients were excluded only if they did not have available diagnostic bone marrow tissue. Reasons for samples unavailability include presence of only a decalcified core biopsy (unsuitable for sequencing) or reclamation of central review material by participating institutions. The clinical characteristics and geographic origins of the patients included in this study are similar to those of the entire trial cohort (supplemental Tables 1 and 2, available on the *Blood* Web site). In total, 194 patients from 81 sites in 22 countries were included in the study cohort: 93 had s-AML, defined by the histologic documentation of antecedent MDS or chronic myelomonocytic leukemia (CMML) according to World Health Organization (WHO) criteria at least 3 months before study entry; 101 patients had t-AML, 18 of whom had an interval diagnosis of therapy-related MDS (t-MDS), defined according to the protocol as AML developing any time after documented exposure to specific leukemogenic therapies for a nonmyeloid condition, including alkylating agents, platinum derivatives, taxanes, topoisomerase II inhibitors, antimetabolites, external beam radiotherapy to active marrow sites, and therapeutic systemic radioisotopes. Paired samples obtained at the time of MDS diagnosis and s-AML were available for 17 subjects, and paired samples from s-AML and CR were available for 16 subjects. All patients provided written informed consent with the approval of the appropriate ethics committees and in accordance with the Declaration of Helsinki. AML diagnosis was confirmed and treatment response assessed by central pathology review. Cytogenetic analysis was performed by a central laboratory and interpreted per International System for Human Cytogenetic Nomenclature (2013).<sup>12</sup> The median follow-up time, calculated from initiation of treatment, was 9.5 months. Remission samples used for serial analyses were obtained at protocol-specific time points (day  $37 \pm 4$  of induction or reinduction, earlier upon count recovery, or later in the case of bone marrow hypocellularity) and confirmed to represent morphologic remission (bone marrow blasts  $<5\%$ ) by local and central pathology review. The median time to remission was 38 days after start of cycle 1 (range 29-74). A validation cohort consisted of 105 unselected AML patients treated at Dana-Farber Cancer Institute (DFCI).

### DNA sequencing and mutation analysis

**Library construction and sequencing.** Target regions of 82 genes, selected on the basis of their known or suspected involvement in the pathogenesis of myeloid malignancies or bone marrow failure, were enriched using the Custom SureSelect hybrid capture system (Agilent Technologies). Sequencing was focused on specific mutational hotspots in *NPM1*, *IDH1*, *IDH2*, *SRSF2*, *U2AF1*, *SF3B1*, *FLT3*, *KIT*, *CBL*, *CBLB*, *BRAF*, *CSF1R*, *JAK2*, *MPL*, *STAT3*, *GNAS*, and *SETBP1* and the entire coding regions, including canonical splice sites, of all other genes. A list of genes and the genomic coordinates of all target regions are provided in the supplemental information (supplemental Tables 3 and 4). Native genomic DNA extracted from bone marrow aspirate slides (supplemental Methods and supplemental Figure 1) was sheared and library constructed per manufacturer protocol. Libraries were pooled at 48 samples per lane in equimolar amounts totaling 500 ng of DNA. Each pool was hybridized to RNA baits, consisting of 8811 probes, spanning 311.266 kbp. Each capture reaction was washed, amplified, and sequenced on 2 lanes of an Illumina HiSeq 2000 100-bp paired-end run. Germline DNA was uniformly unavailable for analysis.

**Variant calling and annotation.** Fastq files were aligned to hg19 version of the human genome with BWA 0.6.2. Single-nucleotide and small indel calling was performed with samtools-0.1.18 mpileup and Varscan 2.2.3. *FLT3*-ITD analysis performed using Pindel 0.2.4 at genomic coordinates chr13:28 608 000-28 608 600 and confirmed by Sanger sequencing (5'-GC AATTTAGGTATGAAAGCCAGC-3' and 5'-CTTTCAGCATTTTGACG GCAACC-3'). Variants were annotated with cDNA and amino acid changes, number of reads supporting the variant allele, population allele frequency in 1000 Genomes release 2.2.2<sup>13</sup> and the Exome Sequencing Project,<sup>14</sup> and

presence in Catalogue of Somatic Mutations in Cancer, version 64.<sup>15</sup> Variants were excluded if there were fewer than 15 total reads; if they fell outside of the target coordinates, had excessive read-strand bias, or had an excessive number of variant calls in the local region; or caused synonymous changes. All mutations (supplemental Table 5) and variants of unknown significance (supplemental Table 6) are included in tabular form in the supplemental information. The distribution of gene-specific variant allele fractions is displayed in supplemental Figure 3.

Analysis of the publicly available exome- and genome-level data from The Cancer Genome Atlas (TCGA) de novo AML cohort was restricted to the same genomic coordinates as our focused analysis of the s-AML and t-AML cohorts, and variants were identified and annotated using the same criteria. For the DFCI cohort, DNA was extracted from fresh bone marrow aspirate or peripheral blood samples and coding regions of 275 cancer-associated genes were enriched using the Custom SureSelect hybrid capture system (Agilent Technologies) as previously described.<sup>16</sup> A list of sequenced genes is provided in the supplemental appendix (supplemental Table 7). All genes required to ascertain genetic ontogeny class were included in the validation platform.

### Statistical methods

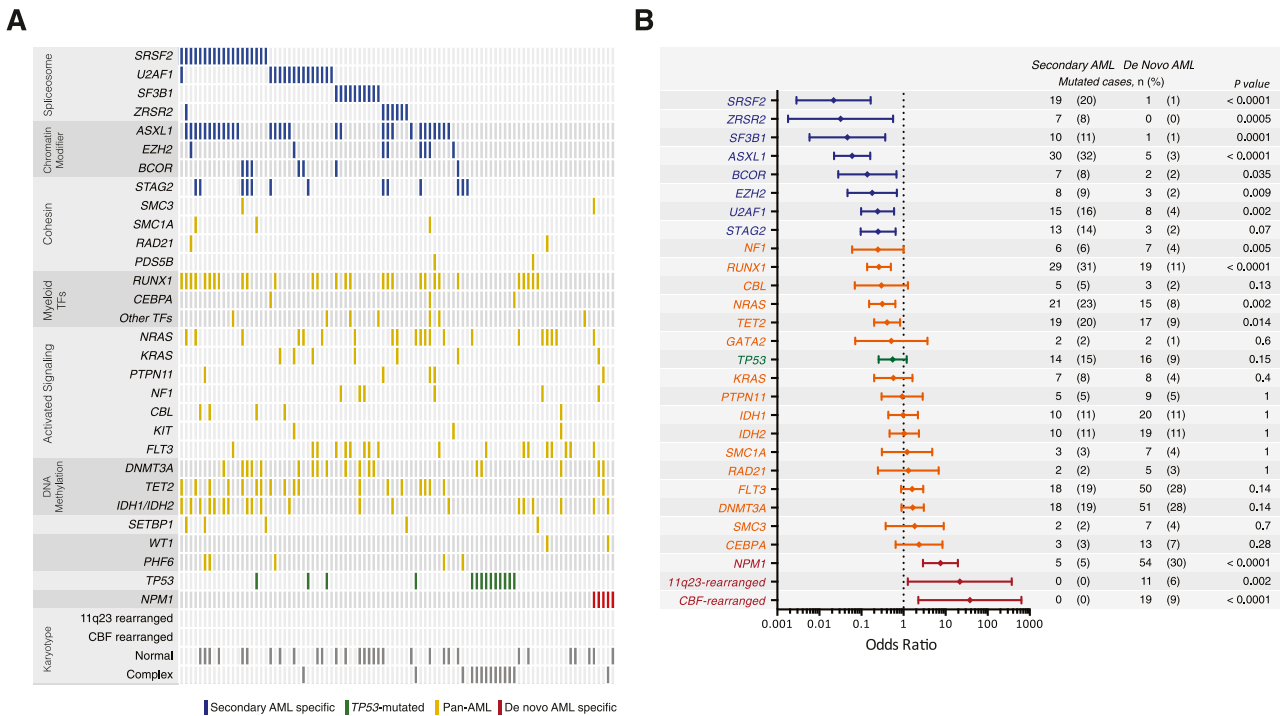
OS was calculated from the date of treatment initiation to the date of death. Surviving patients were censored at the date on which they were last known to be alive. OS curves were estimated using the Kaplan-Meier method, compared using a log-rank test, and analyzed using a univariable Cox model, and the significance of hazard ratio significance was assessed by Wald test. Associations of continuous measures between groups were assessed using a Wilcoxon rank-sum test and categorical variables were assessed using a Fisher exact test. *P* values are unadjusted, 2-sided, and considered significant at .05. Event-free survival was calculated from date of diagnosis to the date of death, relapse, or confirmation of no remission.

## Results

### Ontogeny-defining mutations in s-AML

To investigate the somatic genetic lesions in cases of AML that develop following an antecedent MDS or CMML, we isolated DNA from the diagnostic bone marrow aspirates of 93 rigorously defined s-AML patients enrolled on the ACCEDE trial and sequenced 82 genes that are recurrently mutated in myeloid malignancies or implicated in the biology of bone marrow failure. In total, we identified 353 single-nucleotide variants and small insertions or deletions affecting 40 genes (supplemental Figure 1), with at least 1 mutation detected in 96.8% (90/93) of cases (Figure 1A). The most frequently mutated genes were those involved in RNA splicing (55%), DNA methylation (46%), chromatin modification (42%), RAS pathway signaling (42%), transcriptional regulation (34%), and the cohesin complex (22%).

The spectrum of genetic lesions in our s-AML cohort is notably different from previously described cases of de novo AML.<sup>1-5</sup> We therefore compared the genetic profile of our s-AML cases with 180 cases of non-M3 de novo AML reported in The Cancer Genome Atlas, analyzing all genetic data through the same computational pipeline. We identified 3 distinct and mutually exclusive patterns of mutations. First, 8 genes were mutated with  $>95\%$  specificity in s-AML compared with de novo AML, including *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *ASXL1*, *EZH2*, *BCOR*, and *STAG2*, hereafter named "secondary-type" mutations (Figure 1B). These 8 genes are also commonly mutated in MDS,<sup>2,4,6-9,17</sup> suggesting that they may primarily drive the dysplastic differentiation and ineffective hematopoiesis that is characteristic of MDS, without efficiently promoting development of frank leukemia.



**Figure 1. Spectrum and ontogeny specificity of myeloid driver mutations in s-AML.** (A) A mutation plot shows nonsynonymous mutations in individual genes, grouped into categories, as labeled on the left. Mutations are depicted by colored bars and each column represents 1 of the 93 sequenced subjects. Colors reflect ontogeny specificity of mutated genes, as described in (B). (B) Shown is the association between individual mutated genes and clinically defined s-AML or de novo AML ontogeny, as depicted by odds ratio on a log<sub>10</sub> scale. Colors indicate genes with >95% specificity for s-AML (blue), >95% specificity for de novo AML (red), or <95% specificity for s-AML or de novo AML (yellow or green). The number and frequency cases with mutations in each gene in s-AML and de novo AML cases are shown on the right.

Second, we identified 3 alterations that were significantly under-represented in s-AML compared with de novo AML, including *NPM1* mutations ( $P < .0001$ ), *MLL/11q23* rearrangements ( $P = .0002$ ), and *CBF* rearrangements ( $P < .0001$ ). We termed this set of lesions *de novo-type* alterations. *NPM1* mutations were identified in only 5.4% (5/93) of s-AML subjects, none of whom had concurrent secondary-type or *TP53* mutations.

Third, mutations in the *TP53* gene have been associated with a distinct and dominant clinical phenotype in myeloid malignancies, including a complex karyotype, intrinsic therapy resistance, and very poor survival.<sup>18,19</sup> In our s-AML cohort, cases with *TP53* mutations (n = 14, 15.1%) had more complex karyotypes (mean alterations per case = 10.3 vs 1.0,  $P < .0001$ ) and reduced OS (median OS = 4.0 vs 8.5 months, hazard ratio [HR] = 2.00,  $P = .044$ ) relative to s-AML without *TP53* mutations (supplemental Figures 3 and 4). All other mutations identified were not specific to either AML subtype and were thus labeled “pan-AML” mutations.

We therefore propose 3 distinct genetic ontogenies for AML defined by the presence of (1) secondary-type mutations, (2) de novo-type or pan-AML mutations, or (3) *TP53* mutations.

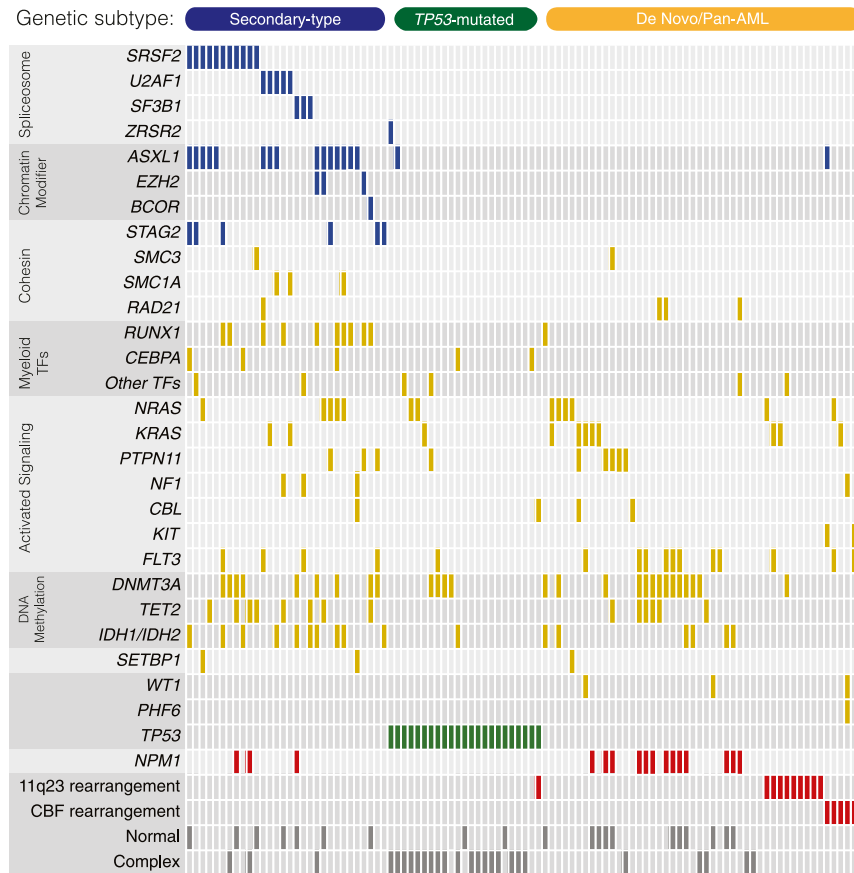
**Genetic classification of therapy-related AML**

We next applied our genetic ontogeny-based classification to t-AML, a heterogeneous disease unified only by a clinical history of exposure to leukemogenic therapy.<sup>20</sup> We reasoned that a genetic classifier could allocate t-AML patients into more uniform groups and inform our understanding of therapy-related leukemogenesis.<sup>21,22</sup> Therefore, we analyzed a cohort of 101 t-AML patients enrolled on the ACCEDE trial using the same sequencing platform as the s-AML cohort. In total, we identified 296 single-nucleotide variants and

small insertions or deletions affecting 43 genes (supplemental Figure 1), with at least 1 mutation detected in 97% (98/101) of cases (Figure 2).

Among subjects with clinically defined t-AML, 33% (34/101) harbored secondary-type mutations in *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *ASXL1*, *EZH2*, *BCOR*, or *STAG2*; 23% (23/101) of patients had *TP53* mutations; 47% (47/101) had only de novo or pan-AML alterations. Of note, 3 patients had both secondary-type and *TP53* mutations and are subsequently categorized in the *TP53* mutated subgroup only. An interval diagnosis of t-MDS was more common in t-AML patients with secondary-type than de novo/pan-AML mutations (29.4% vs 8.5%,  $P = .019$ ) (supplemental Figure 6). There were no differences in CR rate among t-AML patients based on ontogeny group (Figure 3A). However, patients with secondary-type (36.8%,  $P = .022$ ) and *TP53* mutations (54.5%,  $P = .004$ ) were significantly more likely to require multiple induction cycles than patients with de novo/pan-AML mutations (7.4%) (Figure 3B), suggesting relative chemoresistance in these groups.

We next asked whether t-AML patients as a whole possessed a shared set of clinical characteristics or instead conformed to genetic ontogeny classes independent of prior therapy. t-AML patients with secondary-type mutations were significantly older (62.7 vs 53.4,  $P = .002$ ) and had significantly more recurrent driver mutations (4.1 vs 2.5,  $P < .0001$ ) than t-AML patients with de novo/pan-AML mutations (Figure 3C). In fact, t-AML with secondary-type mutations closely resembled clinically defined *TP53*-unmutated s-AML without prior exposure to leukemogenic therapy with regards to age (62.7 vs 62.3), male predominance (male = 65% vs 69%), number of recurrent myeloid driver mutations per case (4.1 vs 4.0), and frequency of chromosome 5 or



**Figure 2. Mutations in therapy-related AML.** A mutation plot shows nonsynonymous mutations in individual genes, as labeled on the left. Mutations are depicted by colored bars, and each column represents 1 of the 101 sequenced subjects. Colors reflect ontogeny specificity of mutated genes, as described in Figure 1. Genetic ontogeny groups are labeled on the top.

7 abnormalities (15.4% vs 16.2%). By contrast, t-AML patients with genetically defined de novo AML closely resembled clinically defined de novo AML, reflected by similar frequencies of mutations in specific genes, including *NPM1*, *FLT3*, *DNMT3A*, *TET2*, *IDH1/IDH2*, and *WT1*.<sup>2,4</sup>

Similar to *TP53*-mutated AML without prior therapy, *TP53* mutations in t-AML were associated with a highly complex (mean alterations 7.5 vs 1.7,  $P < .0001$ ), often monosomal karyotype, with frequent abnormalities of chromosomes 5 and 7.<sup>18</sup> The mean number of recurrent driver mutations in *TP53*-mutated t-AML cases was lower than in t-AML cases with secondary-type mutations (1.9 vs 4.1  $P < .0001$ ) (Figure 3C), and 39% harbored no additional point mutations in known myeloid driver genes, consistent with previous findings of *TP53*-mutated myeloid malignancies.<sup>19</sup>

Together, these data indicate that prior exposure to leukemogenic therapy does not define a genetically conforming “t-AML” ontogeny. Rather, t-AML can be separated into 3 groups that vary in prevalence across age groups (Figure 3D), each bearing more similarity to AML with the same genetic alterations and no leukemogenic exposure than to t-AML patients as a whole. Outside of the known association between exposure to topoisomerase-2 inhibitors and MLL rearrangements, we observed no associations between specific leukemogenic exposures and genetic ontogeny groups (Figure 3E).

#### Pan-AML mutations are gained at disease progression

Having organized mutations into secondary-type, *TP53*-defined, and de novo/pan-AML lesions, we asked whether genes in these

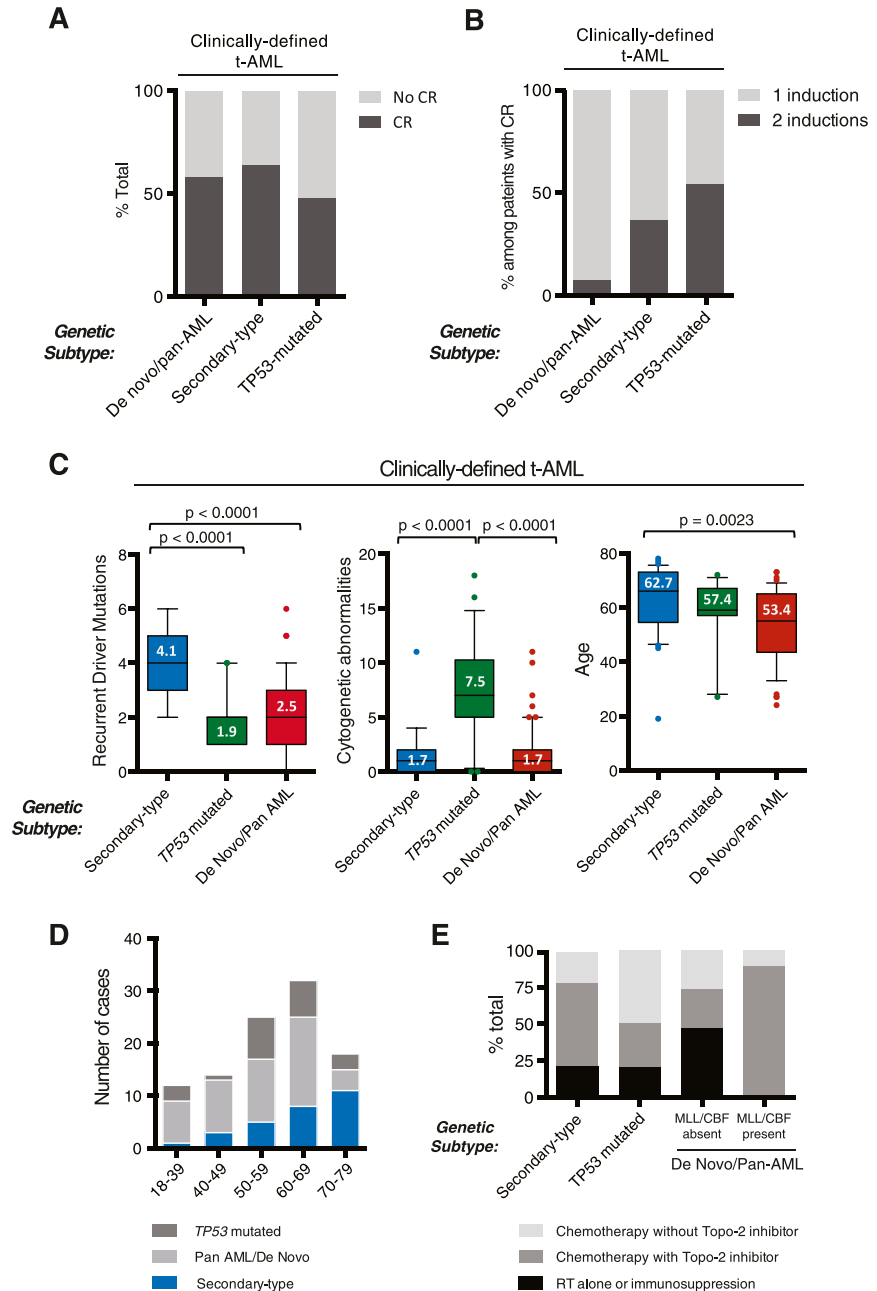
different ontogeny classes would have distinct functional roles during leukemic progression. To address this question, we sequenced 17 paired MDS and s-AML bone marrow aspirate samples from the clinical trial cohort. At the time of s-AML transformation, we detected at least 1 new recurrent driver mutation in 59% of cases (supplemental Table 8). Patients with *TP53* mutations at the time of MDS sampling (3/17) did not gain additional mutations at disease progression (0% vs 42%,  $P = .051$ ), providing further evidence that *TP53* mutations define a distinct class of myeloid malignancies.

All new mutations (100%, 18/18) fell within the pan-AML class of genes (Figure 4A and supplemental Figure 7). The most commonly acquired mutations involved genes encoding myeloid transcription factors (*RUNX1*, *CEBPA*, *GATA2*) and signal transduction proteins (*FLT3* or RAS pathway), together accounting for 78% of new mutations (Figure 4C). Notably, among 8 patients in our cohort with *CEBPA* mutations, only one had biallelic mutations. In this case, we demonstrate that the second *CEBPA* mutation was a subclonal progression event that occurred during the transition from MDS to s-AML.

#### Gross persistence of mutations during remission reveals reservoirs of therapy resistance

We hypothesized that poor response to chemotherapy in s-AML may be caused by the presence of secondary-type genetic lesions that are also present in MDS. To examine the relative sensitivity of different mutations to chemotherapy, we examined 16 paired pre- and posttreatment bone marrow specimens from s-AML patients who achieved a morphologic CR after induction chemotherapy.

**Figure 3. Ontogeny-based genetic classification defines clinically distinct t-AML subgroups.** (A-B) Induction outcomes in clinically defined t-AML patients according to genetic ontogeny group. (A) Morphologic CR outcomes according to genetic ontogeny group among clinically defined t-AML patients receiving standard induction chemotherapy. (B) Shown is the number of induction cycles among t-AML patients achieving CR. (C) Within clinically defined t-AML, genetic classification identifies subgroups with distinct characteristics, including number of recurrent driver mutations per case, number of cytogenetic abnormalities per case, and age. In box plots, center lines show the median value, box limits indicate the 25th and 75th percentiles, whiskers extend to the 10th and 90th percentiles, and outliers are represented by dots. (D) Distribution of genetic ontogeny groups in t-AML patients according to age group. (E) History of prior chemotherapy or radiation exposure based on genetic ontogeny class.



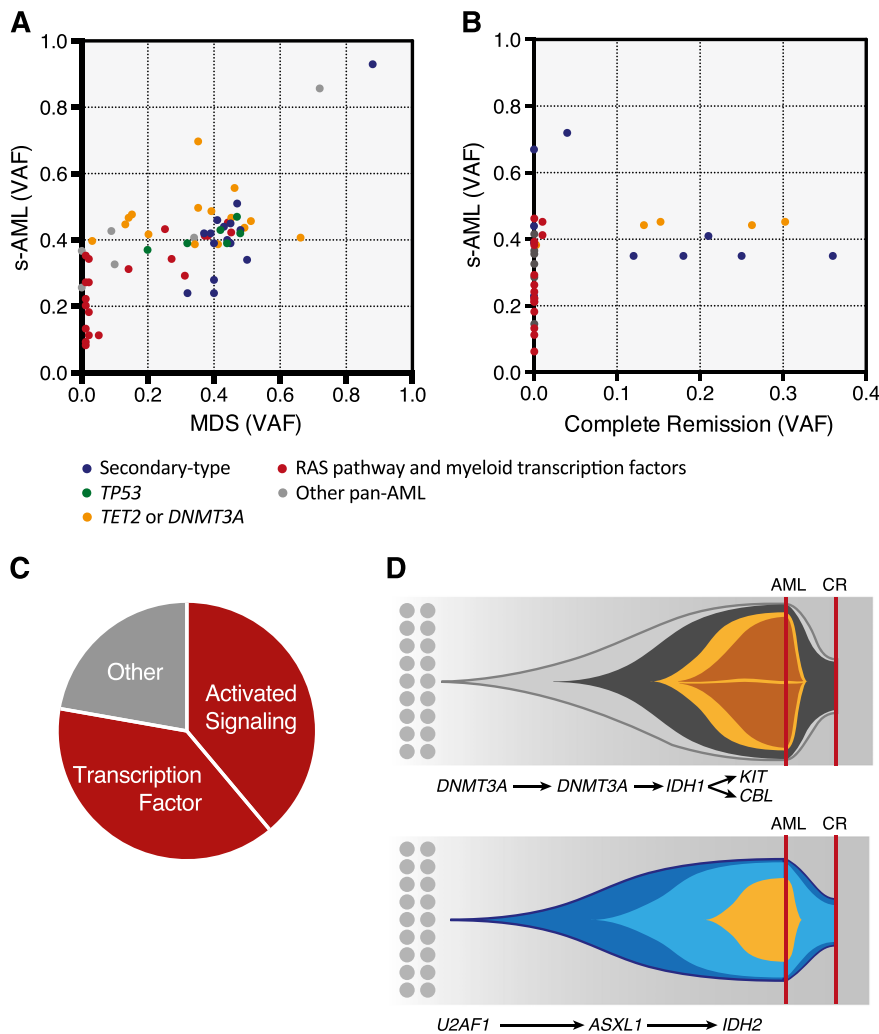
Strikingly, 69% (11/16) of s-AML remission samples had measurable persistence of disease-driving mutations, despite achieving a morphologic CR (Figure 4B). In nearly half of these patients, we observed selective clearance of a subset of mutations alongside a relatively chemoresistant founder clone, suggesting the presence of genetic subclones with differential chemosensitivity. The same pan-AML mutations that are commonly acquired at the time of s-AML transformation, including those affecting myeloid transcription factors and signal transduction proteins, were preferentially lost in the setting of morphologic remission with selective elimination of a genetic subclone (67%, 10/15). By contrast, mutations in *TP53*, *DNMT3A*, *TET2*, and genes involved in RNA splicing or chromatin modification were rarely acquired at disease progression and were preferentially retained with a high mutant allele fraction.

In aggregate, our analyses of serial samples demonstrate that secondary-type mutations, along with *TET2* and *DNMT3A*, are

acquired during the MDS phase of disease, are rarely gained at leukemic progression, and are preferentially retained in clonal remission after induction chemotherapy. By contrast, other pan-AML mutations are less common in MDS, are frequently gained at disease progression, and are more likely to be lost in the context of morphologic remission (supplemental Table 8).

**Genetic classification of an unselected cohort of AML patients**

We next asked whether our ontogeny-based genetic classifier could resolve unrecognized clinical heterogeneity in an unselected cohort of AML patients. We collected a sequential cohort of 105 patients with AML who were treated at our institution over the past year and whose leukemia had been subjected to prospective sequencing of 275 cancer-associated genes (supplemental Methods). Based on clinical history, 64% of cases were diagnosed with de novo AML,



**Figure 4. Analysis of serial samples.** Scatter plots showing variant allele fractions (VAF) at time of paired samples at (A) MDS and s-AML and (B) diagnosis and morphologic CR. Colors show ontogeny specificity of mutated genes including secondary-type (blue); *TP53* (green); and de novo/pan-AML subsets, including RAS pathway and myeloid transcription factors (red), *TET2* and *DNMT3A* (yellow), and other pan-AML (gray). Mutations were categorized as new if they were detected in s-AML but present below 1% allele frequency in MDS (85%), or if their allele frequency increased from  $\leq 10\%$  in MDS to  $>30\%$  in the subsequent s-AML (15%). Mutations were labeled as selectively lost only if variant allele fraction was  $<0.5\%$  at remission. (C) Pie chart showing s-AML progression mutations by functional class. (D) Representative fish plots from 2 cases with subclonal remissions showing clonal architecture at diagnosis (indicated by a red line—AML) and after treatment at time of morphologic CR (indicated by a red line—CR). In both cases, clonal remission is characterized by disappearance of progression mutations and relative persistence of founder mutations despite the absence of bone marrow myeloblasts.

30% with s-AML and 6% with t-AML (Figure 5). Consistent with our previous findings, patients with secondary-type mutations were older and had more genetically complex disease compared with patients with de novo/pan-AML or *TP53* mutations (Figure 6A). The presence of a *TP53* mutation was associated with karyotype complexity (mean cytogenetic alterations 6.7 vs 0.9,  $P < .0001$ ) and reduced OS (median OS = 5.4 vs 10.9 months; HR 3.35,  $P = .0002$ ) (supplemental Figure 4).

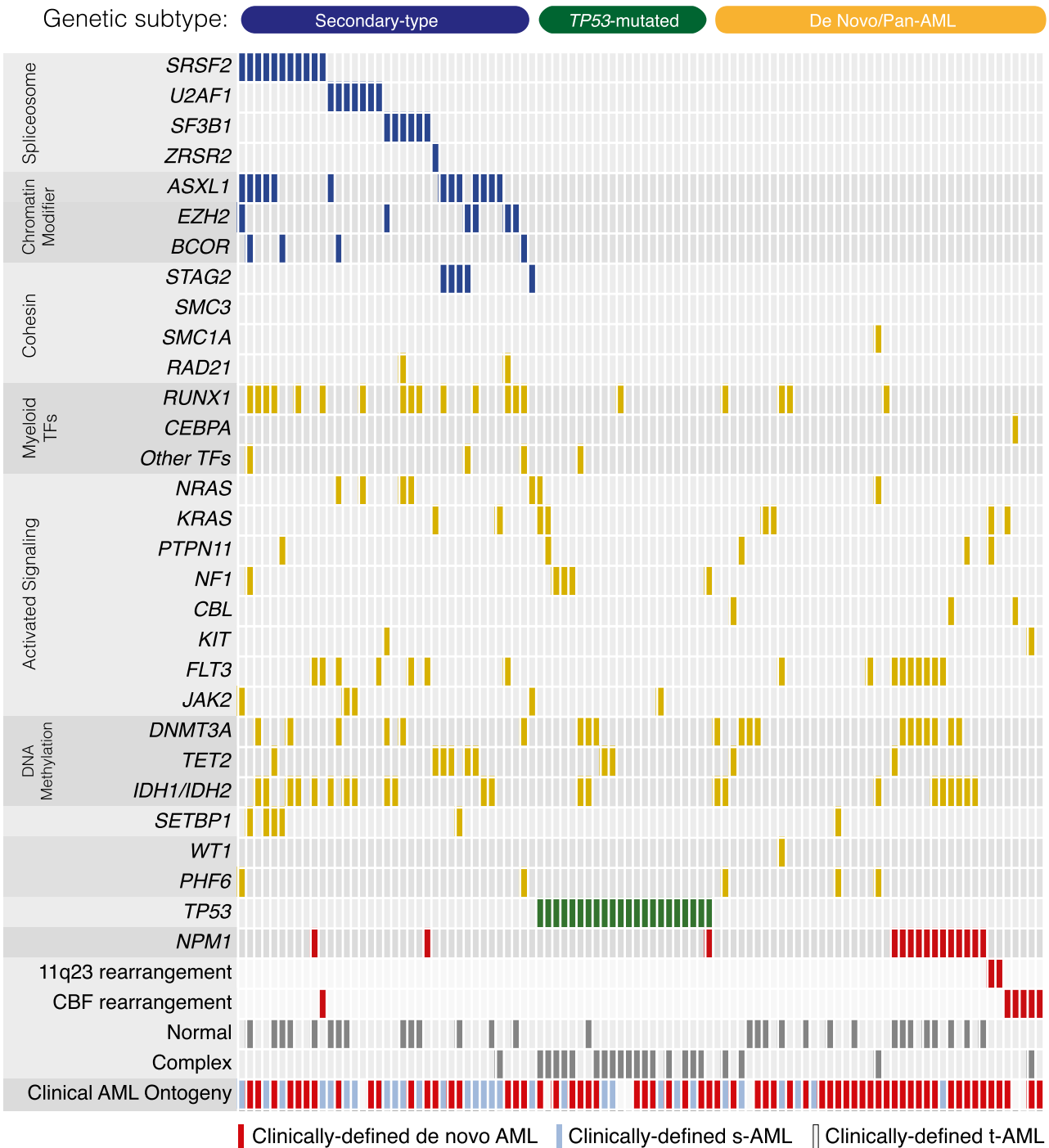
In this cohort, 42 of 67 (63%) de novo AML patients were  $\geq 60$  years old at the time of diagnosis, consistent with the epidemiology of AML and highlighting a group of patients with a lower rate of remission with standard induction chemotherapy.<sup>23-26</sup> Among these older de novo AML patients, 33.3% (14/42) had secondary-type mutations, 21.4% (9/42) had *TP53* mutations, and 45.2% (19/42) had de novo/pan-AML mutations (supplemental Figure 9). To determine whether genetic ontogeny could identify subsets of older de novo AML patients with distinct clinical outcomes, we evaluated the rate of CR in evaluable patients who received standard induction regimens ( $n = 30$ , 71.4%). Among patients with secondary-type mutations, 50% (6/12) achieved CR and 50% (3/6) of these CRs required 2 induction cycles. The CR rate among elderly de novo AML patients with secondary-type mutations closely mirrored the CR rate among evaluable patients with clinically confirmed s-AML who received standard induction therapy in this independent cohort (53%, 10/20). By

contrast, 92% (11/12) of older patients with de novo/pan-AML mutations achieved CR ( $P = .069$ ), and only 18% (2/11) of these remissions required reinduction (Figure 6B). Event-free survival among older de novo AML patients with secondary-type (4.2 months,  $P = .059$ ) or *TP53* mutations (6.2 months  $P = .039$ ) was shorter than those de novo/pan-AML mutations (15.7 months) (Figure 6C).

These data suggest that genetic ontogeny highlights a subset of patients with secondary-type mutations who, based on age, genetic characteristics, and induction outcomes, may have had an unrecognized period of antecedent myelodysplasia before AML diagnosis. In elderly AML, genetic ontogeny, even more than clinical ontogeny, may account for relative differences in intrinsic chemosensitivity.

## Discussion

AML is currently classified into 3 clinical ontogeny groups that are, in practice, defined by the ability to document an antecedent MDS phase (s-AML), previous leukemogenic exposures (t-AML), or the absence of both (de novo AML). Disease classification is thus intrinsically inexact and depends on the availability of prior clinical data rather than objective criteria at the time of diagnosis. Genomic



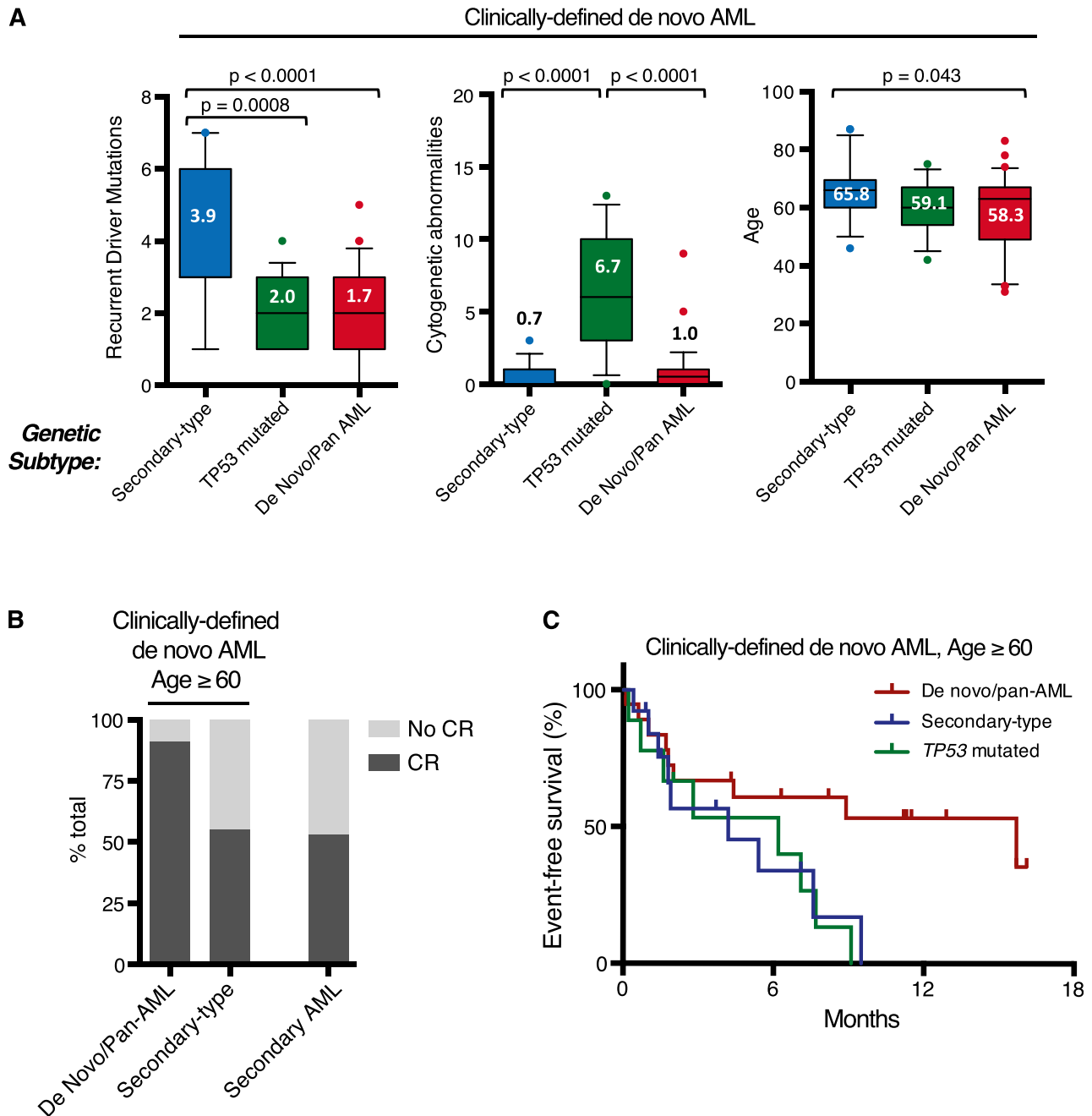
**Figure 5. Mutations in an unselected cohort of AML patients.** A mutation plot shows nonsynonymous mutations in individual genes, grouped into categories, as labeled on the left. Mutations are depicted by colored bars, and each column represents 1 of the 105 sequenced subjects. Colors reflect ontogeny specificity of mutated genes, as described in Figure 1.

discovery efforts have revealed remarkable genetic heterogeneity in AML, but genetic findings have not been broadly linked to leukemia ontogeny or clinical classification. By studying a cohort of rigorously defined s-AML cases, we defined a core set of mutations that is highly specific to cases of AML arising after MDS. Based on the presence of secondary-type, *TP53*, or de novo/pan-AML-type mutations, we defined 3 genetic ontogeny groups that facilitate an objective classification of AML development that is agnostic to clinical history.

In our initial cohorts, as well as in an independent, unbiased collection of AML cases, genetic ontogeny was associated with

characteristic phenotypes, irrespective of clinical ontogeny assignment. AML with secondary-type mutations tends to occur in older individuals and has approximately 4 mutations in myeloid driver genes per case, whereas AML with de novo/pan-AML alterations is more common in younger individuals and has fewer co-occurring driver alterations. AML with *TP53* mutations has highly distinctive characteristics, including marked karyotype complexity with multiple monosomies, a paucity of driver mutations, and very short survival.

We applied our genetic ontogeny-based classifier to 2 heterogeneous AML cohorts: a well-defined group of t-AML cases and



**Figure 6. Ontogeny-based genetic classification defines clinically distinct de novo AML subgroups.** (A) Within clinically defined de novo AML, genetic classification identifies subgroups with distinct characteristics, including number of recurrent driver mutations per case, number of cytogenetic abnormalities per case, and age. Box plots are described in Figure 3. (B) Proportion of patients achieving CR after intensive induction chemotherapy based on genetic subtype among older de novo AML patients (left) and clinically defined s-AML patients (right). (C) Event-free survival in clinically defined de novo AML patients age  $\geq 60$  years according to genetic ontogeny group. Curves show patients with de novo/pan-AML (red), secondary-type (blue), and *TP53* (green) mutations.

a collection of sequential, unselected AML cases. In both cohorts, we resolved well-recognized clinical heterogeneity into clearly distinct genetic subgroups. We found that one-third of clinically defined t-AML cases display both genetic and clinical characteristics that are indistinguishable from clinically defined s-AML. Similarly, in elderly AML patients without a known prior diagnosis of MDS, one-third of cases have secondary-type genetics and clinically resemble s-AML, indicating that many older patients with apparent de novo AML actually transit through an unrecognized MDS prodrome. Our results also suggest that intrinsic chemoresistance in older AML patients may be enriched in patients

with secondary-type mutations, consistent with a long-postulated explanation for lower response rates in these patients. By contrast, older de novo AML patients who do not harbor secondary-type or *TP53* mutations, those with true biologic de novo AML, may have marked chemosensitivity, similar to younger patients with de novo disease and thus defining a group of older patients with better than expected clinical outcomes.

Genome-level analysis has demonstrated that clinical transformation of MDS to s-AML is associated with subclonal disease progression, marked by genetic evolution with acquisition of numerous recurrent and nonrecurrent somatic mutations.<sup>9</sup> Detailed



interrogation of MDS clonal architecture has further shown that no single gene is uniformly mutated in the disease-founding clone, suggesting a need for identifying more generalized principles of disease initiation and progression.<sup>17</sup> By performing serial assessments before and after disease progression, as well as before and after induction chemotherapy, we resolved s-AML into a composite set of genetic lesions with distinct associations to ontogeny groups.

In our series of 17 MDS/s-AML pairs, we observed that newly acquired s-AML mutations were restricted to the pan-AML ontogeny group and most commonly involved genes encoding myeloid transcription factors or members of the RAS/tyrosine kinase signaling pathway. In contrast, secondary-type lesions and *TP53* mutations were commonly seen in MDS and were not newly gained at leukemic transformation. By evaluating these data in the context of our genetic ontogeny classes, we provide a general framework for functional interpretation of myeloid driver alterations, whereby mutations can be grouped into those that tend to occur as early events (*TET2*, *DNMT3A*, *TP53*, and secondary-type mutations) and those that tend to drive progression subclones (tyrosine kinase/RAS pathway, myeloid transcription factors). As such, mutations affecting TK/RAS signaling and myeloid transcription factors may be informative for detecting reemergence or retransformation of AML in s-AML patients who have remitted to clonal hematopoiesis. Our findings are consistent with recent reports demonstrating the presence of somatic mutations affecting a restricted subset of myeloid driver genes in apparently healthy individuals with normal blood counts.<sup>27,28</sup>

Our results in 16 s-AML diagnosis/remission pairs directly demonstrate the phenomenon of clonal remission, revealing the persistence of disease-driving somatic mutations in remission bone marrow.<sup>29-31</sup> By evaluating persistent mutations in the context of genetic ontogeny class, we show that pan-AML progression mutations are preferentially lost at the time of morphologic remission, whereas secondary-type mutations preferentially persist. We demonstrate a residual disease state, characterized by disappearance of blast-associated genetic subclones and persistence of a primordial neoplastic clone occupying 10% to 60% of the remission bone marrow cellularity. These results provide a plausible biological basis for the poor outcomes of s-AML patients after chemotherapy and provide a powerful rationale for systematic, sequencing-based investigation of remission clonality.

In each cohort (s-AML, t-AML, DFCI), ~80% of patients with antecedent clinical MDS diagnoses had either secondary-type or *TP53* mutations, raising the questions of whether the remaining 20% of cases are biologically distinct in some way. These cases may represent biological overlap between clinically defined “MDS” and “de novo” AML, may reflect inadequate specificity of the current WHO criteria for MDS, or may highlight variability in the diagnostic application of existing MDS pathologic criteria. Unfortunately, we do not have information regarding the extent of dysplasia or the number of involved lineages on prior bone marrow assessment of s-AML patients on this study. Nor do we have information regarding the tempo of disease or severity of cytopenias during the antecedent MDS phase. A systematic evaluation of MDS cases

with secondary-type vs *TP53* vs de novo/pan-AML mutations that integrates genetics with clinicopathologic data, disease latency, and kinetics would inform interpretation of these results.

Although a large number of somatic mutations with tremendous combinatorial diversity can drive AML pathogenesis, we have identified classes of mutations that define 3 clinicopathologically distinct subgroups. Integration of molecular genetic principles into future classification systems can refine existing clinical heuristics to support a more biologically precise disease classification.

## Acknowledgments

The authors thank Rui Chen, Hui Wang, Yumei Li, Ilene Galinsky, Adriana Penicaud, Susan Buchanan, Sarah Cahill, Shannon Millillo, David Yudovich, Bill Lundberg, Venugopal Parameswaran, Michael Maris, Dominik Selleslag, Jean Khoury, Tamas Masszi, Tibor Kovacs, Olga Frankfurt, Krzysztof Warzocha, David Claxton, and the ACCEDE investigators for their generous assistance.

This work was supported by the Friends of Dana-Farber Cancer Institute (R.C.L.), the Edward P. Evans Foundation (R.C.L.), a Harvard Catalyst KL2/CMeRIT Award (R.C.L.), the Lady Tata Memorial Trust (R.C.L.), the National Institutes of Health (National Cancer Institute grants T32CA00917237 and P01 CA108631, National Institute of General Medical Sciences grant T32GM007753, and National Heart, Lung, and Blood Institute grant R01HL082945), the Gabrielle’s Angel Foundation (B.L.E.), Leukemia and Lymphoma Society Scholar and SCOR Awards (B.L.E.), Flames/Pan Mass Challenge (R.M.S.), and the Ted Rubin Foundation (R.M.S.).

## Authorship

Contribution: R.C.L., B.L.E., and R.M.S. designed the study, reviewed data analysis, and wrote the manuscript; R.C.L. analyzed the sequencing data, curated clinical data and variants, and performed bioinformatics analysis; B.G.M., S.S., and P.V.G. developed variant calling algorithms and sequence data processing pipelines; D.N. and E.M. curated clinical data and performed statistical analysis; N.L. oversaw clinical sequencing platform and analyzed data; S.L.A., A.P., M.W., R.K.S., H.P.E., L.E.D., B.L.P., D.P.S., M.W., D.J.D., and R.M.S. diagnosed patients and prepared samples; and all authors reviewed the manuscript during its preparation.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Richard M. Stone, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215; e-mail: richard\_stone@dfci.harvard.edu; and Benjamin L. Ebert, Brigham and Women’s Hospital, 1 Blackfan Circle–Karp CHR 5.211, Boston, MA 02115; e-mail: bebert@partners.org.

## References

- Churpek JE, Larson RA. The evolving challenge of therapy-related myeloid neoplasms. *Best Pract Res Clin Haematol*. 2013;26(4):309-317.
- Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1079-1089.
- Pfirschmann M, Ehninger G, Thiede C, et al; Study Alliance Leukaemia (SAL). Prediction of post-remission survival in acute myeloid leukaemia: a post-hoc analysis of the AML96 trial. *Lancet Oncol*. 2012;13(2):207-214.
- Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074.
- Kantarjian H, O’Brien S, Cortes J, et al. Results of intensive chemotherapy in 998 patients age 65 years or older with acute myeloid leukemia or high-risk myelodysplastic syndrome: predictive prognostic models for outcome. *Cancer*. 2006; 106(5):1090-1098.

6. Papaemmanuil E, Gerstung M, Malcovati L, et al; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013;122(22):3616-3627, quiz 3699.
7. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014;28(2):241-247.
8. Itzykson R, Kosmider O, Renneville A, et al. Clonal architecture of chronic myelomonocytic leukemias. *Blood*. 2013;121(12):2186-2198.
9. Walter MJ, Shen D, Ding L, et al. Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1090-1098.
10. Itzhar N, Dessen P, Toujani S, et al. Chromosomal minimal critical regions in therapy-related leukemia appear different from those of de novo leukemia by high-resolution aCGH. *PLoS ONE*. 2011;6(2):e16623.
11. Stone RM, Mazzola E, Neubergh D, et al. A phase III open label randomized study of cytarabine in combination with amonafide L-malate or daunorubicin as induction therapy for patients with secondary acute myeloid leukemia (AML). *J Clin Oncol*. In press.
12. Shaffer LG, McGowan-Jordan J, Schmid M, eds. ISCN 2013: An International System for Human Cytogenetic Nomenclature. Basel, Switzerland: S. Karger; 2013.
13. 1000 Genomes Project Consortium. Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65.
14. Andreasen C, Nielsen JB, Refsgaard L, et al. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. *Eur J Hum Genet*. 2013;21(9):918-928.
15. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet*. 2008;57:10.11.1-10.11.26.
16. Wagle N, Berger MF, Davis MJ, et al. High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov*. 2012;2(1):82-93.
17. Walter MJ, Shen D, Shao J, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia*. 2013;27(6):1275-1282.
18. Rucker FG, Schlenk RF, Bullinger L, et al. TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. *Blood*. 2012;119(9):2114-2121.
19. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*. 2011;364(26):2496-2506.
20. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937-951.
21. Pedersen-Bjergaard J, Christiansen DH, Desta F, Andersen MK. Alternative genetic pathways and cooperating genetic abnormalities in the pathogenesis of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia*. 2006;20(11):1943-1949.
22. Pedersen-Bjergaard J, Pedersen M, Roulston D, Philip P. Different genetic pathways in leukemogenesis for patients presenting with therapy-related myelodysplasia and therapy-related acute myeloid leukemia. *Blood*. 1995;86(9):3542-3552.
23. Juliusson G, Antunovic P, Derolf A, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood*. 2009;113(18):4179-4187.
24. Appelbaum FR, Gundacker H, Head DR, et al. Age and acute myeloid leukemia. *Blood*. 2006;107(9):3481-3485.
25. Anderson JE, Kopecky KJ, Willman CL, et al. Outcome after induction chemotherapy for older patients with acute myeloid leukemia is not improved with mitoxantrone and etoposide compared to cytarabine and daunorubicin: a Southwest Oncology Group study. *Blood*. 2002;100(12):3869-3876.
26. Goldstone AH, Burnett AK, Wheatley K, Smith AG, Hutchinson RM, Clark RE; Medical Research Council Adult Leukemia Working Party. Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical Research Council AML11 trial. *Blood*. 2001;98(5):1302-1311.
27. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
28. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.
29. Fearon ER, Burke PJ, Schiffer CA, Zehnbaauer BA, Vogelstein B. Differentiation of leukemia cells to polymorphonuclear leukocytes in patients with acute nonlymphocytic leukemia. *N Engl J Med*. 1986;315(1):15-24.
30. Gale RE, Wheaton H, Goldstone AH, Burnett AK, Linch DC. Frequency of clonal remission in acute myeloid leukaemia. *Lancet*. 1993;341(8838):138-142.
31. Jacobson RJ, Temple MJ, Singer JW, Raskind W, Powell J, Fialkow PJ. A clonal complete remission in a patient with acute nonlymphocytic leukemia originating in a multipotent stem cell. *N Engl J Med*. 1984;310(23):1513-1517.