Acute myelogenous leukemia (AML) frequently relapses after complete remission (CR), necessitating improved detection and phenotypic characterization of treatment-resistant residual disease. In this work, we have optimized droplet digital PCR to broadly measure mutated alleles of recurrently mutated genes in CR marrows of AML patients at levels as low as 0.002% variant allele frequency. Most gene mutations persisted in CR, albeit at highly variable and gene-dependent levels. The majority of AML cases demonstrated residual aberrant oligoclonal hematopoiesis. Importantly, we detected very rare cells (as few as 1 in 15,000) that were genomically similar to the dominant blast populations at diagnosis and were fully clonally represented at relapse, identifying these rare cells as one common source of AML relapse. Clinically, the mutant allele burden was associated with overall survival in AML, and our findings narrow the repertoire of gene mutations useful in minimal residual disease–based prognostication in AML. Overall, this work delineates rare cell populations that cause AML relapse, with direct implications for AML research directions and strategies to improve AML therapies and outcome.

Introduction
Acute myelogenous leukemia (AML) afflicts 21,000 new patients yearly in the US and causes 10,000 premature deaths. While most patients with AML achieve complete remission (CR) with induction chemotherapy, many will experience disease relapse. This clinical reality, together with findings of shared genomic characteristics and invariant clonal relatedness (1, 2) of AML at diagnosis and relapse, supports the notion that aberrant cells exist that are able to survive chemotherapy and persist in the bone marrow of patients in clinical CR, a phenomenon commonly referred to as minimal residual disease (MRD). Additional evidence connecting MRD with AML relapse includes the detection of cells with persistent, aberrant leukemia-associated immunophenotypes (LAIPs) in CR using multiparameter flow cytometry (3, 4) and of persistent gene mutations in bone marrow cells in CR that are found in matched relapsed specimens (5, 6). Additionally, quantitation of residual leukemia-associated traits in remission such as fusion transcripts (7–10), gene mutations (11–16), and LAIPs (3, 17–19) has been associated with patient survival, suggesting a prognostic utility of MRD monitoring in specific clinical contexts.

Various biological scenarios have been proposed with regard to the underlying mechanisms of AML relapse. One scenario invokes resistant preleukemic progenitors that, through continued molecular evolution, ultimately acquire properties of relapsing AML blasts; another more parsimonious scenario simply invokes surviving drug-resistant blasts; while yet another scenario invokes selection via chemotherapy of rare mutated cells of whatever cellular characteristic that can form AML blasts. There remain, then, significant limitations to the current understanding of the biology, clonal composition, and genomic landscape of residual leukemia during CR, as well as the very definition and prognostic relevance of residual disease. Most studies of MRD defined by genomic aberrations have focused on analysis of single mutated genes in isolation, notably nucleophosmin 1 (NPM1). Mutant NPM1 mRNA transcripts are frequently detected at less than 0.1% of total NPM1 mRNA in CR bone marrow using allele-specific quantitative PCR (qPCR), and its levels are associated with relapse-free survival (RFS) and overall survival (OS) (11, 14, 16). This suggests that the measurement of other mutated genes at this high sensitivity may be useful for characterizing and prognosticating residual AML, but for most genes, current qPCR approaches are not readily feasible (20).

Furthermore, concurrent measurements of multiple leukemia-associated gene mutations in CR samples may provide a critical genomic context that better delineates the burden and composition of MRD and refines its prognostic value. A few studies have addressed this by performing targeted next-generation DNA sequencing (NGS) of CR samples (12, 13, 21, 22), with the notable finding that most DNMT3A mutations and many TET2 and IDH1/2 mutations persist at high levels (>5% variant allele frequency [VAF]) in early CR. However, error rates inherent to current NGS platforms limited the detection of mutated alleles to a 2.5% or greater VAF, preventing the assessment of lower-level but nevertheless persistent mutations and thereby limiting determination of the clonal composition of residual hematopoiesis in CR.

To more fully characterize residual aberrant clonal hematopoiesis in CR and assess its prognostic impact, we optimized droplet digital PCR (ddPCR) (23–26) to allow high-sensitivity detection
of both hotspot and private mutations within most recurrently mutated genes in AML. Mutations identified in pretreatment blasts in 72 patients with AML were measured in CR bone marrow specimens by ddPCR to determine residual post-therapy VAFs with a limit of detection as low as 0.002% VAF. Additionally, in 20 of these AML patients for whom paired relapse specimens were available, we determined the VAF of gene mutations before therapy, during CR, and at relapse. These efforts resulted in the identification of very rare cells with genomic similarity to the dominant pretreatment blast clones as a common cellular source of AML relapse. Finally, we determined the association of ultrasensitive genomic MRD measurements with RFS and OS, thereby proving a justification for in-depth study of ddPCR-based mutation analysis as a clinical tool in AML management.

Results

Characteristics of the AML patients studied. Seventy-two patients were studied on the basis of availability of specimens procured after treatment and after induction of therapy from a single-center AML clinical translational trial (UMCC 2004.072). All but 2 patients were treated with high-intensity induction chemotherapy regimens involving either an anthracycline combined with cytarabine or high-dose cytarabine, and all achieved CR after induction. Sixty-five percent (47 of 72) of patients ultimately relapsed, while twenty-five percent (18 of 72) remained in a durable remission, defined as remaining relapse free for more than one thousand days. Ten percent (7 of 72) of patients died from nonrelapse-related causes before the one thousand–day point. Twenty-eight percent (20 of 72) of patients received an allogeneic stem cell transplantation (alloSCT) during first CR (CR1), while an additional fourteen percent (10 of 72) underwent alloSCT at a time point after the first relapse. The median time to CR bone marrow sampling was 36 days after initiation of induction chemotherapy (range, 26–84 days), referred to hereafter as initial CR samples. Targeted DNA sequencing of 49 recurrently mutated genes in AML in pre-therapy samples identified 168 gene mutations across 72 patients (median 3 mutations/patient; range, 0–6). Whole-exome sequencing (WES) was performed for 19 patients with 1 or fewer mutations found in the 49-gene panel (see Methods), identifying 2 or more gene mutations for all 19 patients. Patients’ characteristics are listed in Table 1 and Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI91964DS1. Gene mutation frequencies detected in pre-therapy samples with Sanger sequencing and targeted NGS are listed in Table 2, and additional gene mutations detected with WES in pre-therapy samples and subsequently assayed in CR are denoted in Supplemental Table 2.

Most pretreatment gene mutations are detectable in AML CR. A total of 183 mutations identified in 58 genes detected at diagnosis were measured in paired initial CR samples using ddPCR (median 2.5 measurements/patient; range, 1–5). Eighty-four percent (153 of 183) of these mutations were detectable at VAFs ranging from 0.002% to 42.6%, and at least one persistent gene mutation was detected in 90% of AML remission marrows (65 of 72); the remaining 30 mutations were not detected. The range of VAFs of mutated genes in remission marrows varied widely across most multiply measured genes including NPM1 (not detected [ND], 3.198%); IDHI (ND, 14.563%); IDHI2 (ND, 42.457%); ASXL1 (ND, 57.381%); TET2 (ND, 15.047%); PHH6 (ND, 5.229%); RUNXI (0.002%–8.814%); DNMT3A (0.017%, 34.607%); and U2AF1 (0.1%–7.570%) (Figure 1, Supplemental Figure 1, and Supplemental Table 2). Notably, a subset of mutated genes frequently demonstrated persistence at high VAF levels (>1%), including 85% (11 of 13) of mutations in DNMT3A (range, 1.631%–34.607%), 62% (8 of 13) in ASXL1 (2.907%–37.381%); 66% (6 of 9) in IDHI2 (2.070%–42.457%); 58% (7 of 12) in RUNXI1 (1.689%–9.950%); 56% (5 of 9) in IDHI (1.110%–14.563%); and 44% (7 of 16) in TET2 (1.151%–15.047%). Conversely, another group of genes was typically reduced to low levels (<1%), including 100% (6 of 6) of mutations in NRB3A (range ND, 0.685%); 94% (16 of 17) in FLT3 (ND, 0.837%); and 94% (15 of 16) in NPM1 (ND, 0.476%). Overall, VAFs of mutations fell between greater than 1%, greater than 0.1% to 1%, greater than 0.01% to 0.1%, greater than 0.002% to 0.01%, and ND in 36%, 22%, 13%, and 16% of measurements, respectively.

Persistence of oligoclonal hematopoiesis is common in AML CR. Two or more distinct mutations were assayed by ddPCR in 68 AML

### Table 1. Clinical characteristics of patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No.</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. of patients</td>
<td>72</td>
<td>100</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median age</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>21–78</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>41</td>
<td>57</td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td>43</td>
</tr>
<tr>
<td>WHO 2008 classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AML-RGA: (11q23)</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>AML-RGA: (16p13q22)</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>AML-RGA: t(8;21)</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>AML-MRC</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>AML-TR</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>AML-NOS</td>
<td>42</td>
<td>58</td>
</tr>
<tr>
<td>Cytogenetic prognostic class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favorable</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Intermediate</td>
<td>50</td>
<td>69</td>
</tr>
<tr>
<td>Unfavorable</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>Normal cytogenetics</td>
<td>35</td>
<td>49</td>
</tr>
<tr>
<td>alloSCT in first CR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>20</td>
<td>28</td>
</tr>
<tr>
<td>No</td>
<td>52</td>
<td>72</td>
</tr>
<tr>
<td>alloSCT at any time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>30</td>
<td>42</td>
</tr>
<tr>
<td>No</td>
<td>42</td>
<td>58</td>
</tr>
<tr>
<td>Relapse status after CR sampling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relapsed</td>
<td>47</td>
<td>65</td>
</tr>
<tr>
<td>Durable remission</td>
<td>18</td>
<td>25</td>
</tr>
<tr>
<td>Early nonrelapse death</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

*AML-RGA, AML with recurrent genomic abnormalities; AML-MRC, AML with myelodysplasia-related changes; AML-TR, therapy-related AML; AML-NOS, AML not otherwise specified; Durable remission, alive and free from AML relapse for more than 1,000 days from the date of diagnosis; Early nonrelapse death, death without evidence of AML relapse before 1,000 days from the date of diagnosis.
patients. In 44% (30 of 68) of these patients, all mutations were detected at a similar, concordant VAF, defined as a log10 difference of less than 1 (and often much less) in VAF between mutations (Figure 2). However, in 56% (38 of 68) of cases, the VAFs of various mutations were discordant (>1 log10 difference in VAF), indicating the presence of multiple coexisting clones in CR bone marrow. Within discordant cases, the mutated genes most commonly demonstrating the highest allele frequencies included DNMT3A (10 of 38); ASXL1 (6 of 38); TET2 (3 of 38); IDH1 (3 of 38); RUNX1 (2 of 38); and WT1 (2 of 38), along with single instances of genes including STAG2, BCOR, SRSF2, U2AF2, and TP53 (Figure 3). Seventy-nine percent (30 of 38) of these mutations showed high-level persistence with greater than 1% VAF. The genes most commonly demonstrating the lowest allele frequencies in discordant cases included FLT3 (9 of 38); NPM1 (6 of 38); RUNX1 (5 of 38); TET2 (3 of 38); IDH1 (3 of 38); IDH2 (3 of 38); and NRAS (2 of 38). In most patients (74%, 28 of 38), these mutations were present at a VAF of less than 0.1%. AML often relapses from very minor residual cell populations genomically resembling blasts present in AML at diagnosis. To identify which mutations detected in CR marrows tracked subclones that were ultimately present in the dominant relapse clones, we first determined the mutational status and VAF of 49 genes recurrently mutated in AML in 20 cases using available FACs–sorted, high-purity (>95% tumor) paired pre-therapy and relapsed AML blasts using a targeted NGS panel. We identified 65 gene mutations in the pre-therapy specimens (median of 3 mutations/patient, range 1–7). The VAFs for 77% (50 of 65) of these predominantly heterozygous mutations were greater than 40%, indicating near or full clonal representation before therapy. The VAFs of the remaining 23% (15 of 65) of mutations ranged from 7.4% to 38.5%, representing major subclones in the pre-therapy blasts. Comparison of the mutational spectrum in paired relapsed specimens showed that nearly all mutations (96%, 48 of 50) with more than 40% VAF present at diagnosis were present at relapse. Details of these mutations are listed in Supplemental Table 3 and depicted in Supplemental Figure 2.

Fifty of these mutations across the twenty patients were measured using ddPCR in intervening initial CR specimens, and all but four were detected (range, 0.003%–42.457%). Two of the undetected mutations in CR involved FLT3 internal tandem duplications (ITDs), which were subsequently detected again at relapse, suggesting their presence in subclones below the detection limit. However, the other 2 undetected mutations—a fully clonal RUNX1 mutation (49.0% VAF) and a subclonal STAG2 mutation (10.1% VAF) in pre-therapy specimens—were not detected at relapse. Importantly, however, all mutations detected in CR marrows, even at levels of less than 0.01% VAF, were present in the dominant clone or major subclones at relapse (VAF 36.6%–99.0%). Furthermore, in 14 cases in which multiple genes were concurrently measured in CR marrow specimens, genes with the lowest VAF in each specimen were fully clonally represented at relapse and had equivalent VAF with other mutations present (Figure 4). This finding suggests that the cell population comprising the small clone demarcated by the low-level residual mutation is responsible for AML disease relapse rather than clones persisting at relatively high levels in CR which lack that mutation. These findings therefore support the persistence of fully transformed leukemic cells (possibly AML blasts) as one common source of relapse. Furthermore, these findings qualify the concept of preleukemic cells defined as those requiring acquisition of additional changes as a common source of clinical AML relapse, since such acquisitions would be characterized by more stochastic mutational spectra and frequent divergence from diagnosis patterns. The relationship of pre-therapy, CR, and relapse VAFs is displayed in Figure 4 and Supplemental Figure 3, with details listed in Supplemental Table 3.

**DNMT3A mutations persist at a high VAF, even in cases of durable remission.** We next evaluated the allelic frequency of mutated genes with respect to categories of relapse status defined as relapse less than 1 year from diagnosis, relapse 1 year or more after diagnosis, or durable remission (relapse free for more than 1,000 days [range, 1,425–3,300 days]). High-level persistence (>1% VAF) of 1 or more gene mutations was noted in 54% (39 of 72) of patients (Figure 5A). Notably, even in patients with durable remissions, recurrent high-level persistence of DNMT3A mutations was observed as well as persistence of ASXL1 and STAG2 mutations, each in a single case (11.50% and 19.98% VAF, respectively) (Figure 5A).

To elucidate the clinical and biological implications of the persistence of gene mutations with high remission VAFs, we performed parallel clinical outcome analyses. When analyzed using the mutations with the lowest detected VAF for each patient, the geometric VAF means for patient subgroups categorized by relapse within less than 1 year, relapse after more than 1 year, or durable remission were 0.10%, 0.02%, and 0.005%, respectively, with a significant, 10-fold
difference between the relapse before 1 year and durable remission groups \( (P < 0.01) \) (Figure 5B). This difference in depth-of-mutation persistence in the relapsed groups versus the durable remission group was even more distinct when comparing geometric means of the highest measured VAF for each group, excluding DNMT3A, which, as noted, persisted at high levels across all groups. Specifically, the geometric VAF means for patient subgroups categorized by relapse before 1 year, relapse after 1 year, and durable remission were 1.10\%, 0.68\%, and 0.02\%, respectively, showing a significant, 10-fold difference between both relapse groups compared with the durable remission group \( (P < 0.01 \text{ for both}) \) (Figure 5C).

The degree of residual mutation burden correlates with RFS and OS in AML. To further determine the extent to which absolute levels of residual mutation burden correlated with patient survival, we first divided patients into groups on the basis of logarithmic partitions \( (\text{VAF} < 0.01\%, 0.01\%–0.1\%, 0.1\%–1\%, \text{and} >1\%) \) using the lowest VAF measurement for each. We identified a significant trend associating increasing median RFS with each 10-fold decrease in residual mutation burden for remission bone marrow VAFs of less than 0.01\% \( (1,187 \text{ days}, n = 30); 0.01\%–0.1\% \ (328 \text{ days}, n = 13); 0.1\%–1\% \ (169 \text{ days}, n = 17); \text{and greater than} 1\% \ (306 \text{ days}, n = 12) \ (P < 0.01) \) (Figure 6A). This trend was maintained in the same analysis evaluating OS for VAFs of less than 0.01\% \( (2,445 \text{ days}, n = 30) \); 0.01\%–0.1\% \ (413 \text{ days}, n = 13); 0.1\%–1\% \ (330 \text{ days}, n = 17); \text{and greater than} 1\% \ (544 \text{ days}, n = 12) \ (P < 0.01) \) (Figure 6C). On the basis of the difference in median RFS and OS between the group with a VAF of less than 0.01\% and the other groups, patients were consolidated into groups of VAF of less than 0.01\% and VAF of 0.01\% or greater. Univariate analysis showed a significant and clinically relevant difference in median RFS for a VAF of less than 0.01\% \( (1,187 \text{ days}, n = 30) \) versus a VAF of 0.01\% or greater \( (264 \text{ days}, n = 42) \) \( (P < 0.01) \) (Figure 6B) and in median OS for a VAF of less than 0.01\% \( (2,445 \text{ days}, n = 30) \) versus a VAF of 0.01\% or greater \( (408 \text{ days}, n = 42) \) \( (P < 0.01) \) (Figure 6D). Five-year OS was 50\% for a VAF of less than 0.01\% versus 19\% for a VAF of 0.01\% or greater.

Next, we performed the same analyses using the highest measured VAF for each patient. Partitioning patients into logarithmic groups revealed no significant differences for RFS or OS \( (P = 0.25 \text{ and } 0.15, \text{respectively}) \). However, given the high-level persistence of DNMT3A mutations across the cohort as well as recent studies reporting a lack of prognostic value of DNMT3A mutation persistence in CR \( (21, 27) \), we repeated the analysis excluding DNMT3A. Using this refined approach, we identified a significant inverse correlation that associated VAFs of less than 0.01\% and 0.01\%–0.1\% with substantially improved median RFS \( (\text{not reached} \ [\text{NR}], n = 15); \text{and} \ NR, n = 8, \text{respectively}) \) compared...
The highest VAF, regardless of transplant status, and was significant or near-significant for RFS and OS when we measured the lowest VAF again, regardless of transplant status.

Discussion

This study represents to our knowledge the first broad quantitative measurement of recurrent gene mutations in any malignancy using high-sensitivity rare variant detection able to detect 1 in 50,000 mutated cells. The data provide several conclusions regarding the biology of residual AML and preleukemia in early CR. First, the vast majority of gene mutations identified in bulk pre-therapy AML blasts persist in initial CR marrows, including the 26% of mutations detected at very low levels of less than 0.1% VAF. Second, 36% of persistent gene mutations were detected at unexpectedly high levels of greater than 1% VAF (indicating mutation presence in at least 2% of marrow cells), particularly mutations in DNMT3A, TET2, ASXL1, RUNX1, and IDH1/2. Third, oligoclonal hematopoiesis was present in CR in most informative cases, suggesting the common scenario of persistence of both low-level residual leukemic cell populations and high-level pre- or nonleukemic clonal cell populations resistant to chemotherapy. Fourth, and importantly, low-level mutations detected in CR (0.003%–0.58% VAF) that were present in the dominant pretreatment blast populations or major subclones were also fully clonally represented at relapse, implying that the cell populations in which they reside, rather than numerically much larger, chemo-resistant cell populations, are a common source of AML relapse. Fifth, DNMT3A mutations and frequent ASXL1 and with VAFs of 0.1%–1% and greater than 1% (281 days, n = 16; and 279 days, n = 33, respectively) (P < 0.01) (Figure 6E). We observed a similar association with median OS for VAFs of less than 0.01% (NR, n = 15); 0.01%–0.1% (1,588 days, n = 8); 0.1%–1% (398 days, n = 16); and greater than 1% (534 days, n = 33) (P < 0.01) (Figure 6G). On the basis of the separation of median RFS between VAFs of less than 0.01% and 0.01%–0.1% versus VAFs of 0.1%–1% and greater than 1%, patients were consolidated into groups of VAFs of less than 0.1% and 0.1% or greater. Univariate analysis of the highest measured VAF in each sample had a significant and clinically relevant difference in median RFS for a VAF of less than 0.1% (NR, n = 23) versus a VAF of 0.1% or greater (279 days, n = 49) (P < 0.01) (Figure 6F) and in median OS for a VAF of less than 0.1% (NR, n = 23) versus a VAF of 0.1% or greater (413 days, n = 49) (P < 0.01) (Figure 6H). Five-year OS was 65% for a VAF of less than 0.1% versus 16% for a VAF of 0.1% or greater.

Finally, the same survival analyses were performed on the subset of patients with intermediate-risk cytogenetics (n = 50). The association of VAF of gene mutations with RFS and OS remained substantial and highly significant (Supplemental Figure 4). In addition, we evaluated the effect of residual mutation VAF on RFS in the subset of patients who received alloSCT in CR1 and on OS in patients who received alloSCT at any time (Supplemental Figure 5, A–D) as well as RFS in patients who did not receive alloSCT in CR1 and OS in patients who never received alloSCT (Supplemental Figure 5, E–H). The association of gene mutation VAF with RFS and OS remained highly significant when we measured the highest VAF, regardless of transplant status.
**TET2** mutations persist at high levels in initial CR marrows, even in cases in which durable remissions clearly demarcate a non- or preleukemic cellular compartment. These genes are therefore unlikely to be of value in clinical MRD applications. Finally, the allele frequency level of residual mutated genes detected in initial CR marrows is strongly and significantly associated with RFS and OS. Therefore, the much-improved level of detection of variant alleles and subclones harboring them achieved with ddPCR may provide more clinically meaningful risk stratification than that afforded by current genomic MRD platforms.

Previous studies into the clinical relevance of genomic MRD detection have focused either on high-sensitivity detection of single mutated genes, such as qPCR of **NPM1** mutations, or less sensitive surveys of large gene panels, such as targeted NGS approaches (11–15, 21, 22). The application of competitive allele-specific ddPCR to MRD detection in this study demonstrates the technical feasibility of detecting single nucleotide variants and insertion-deletion mutations in the vast majority of recurrently mutated genes in AML in clinical samples at levels comparable to or exceeding those detected with qPCR. Furthermore, the use of genomic DNA (gDNA) rather than cDNA favored in established qPCR assays of MRD allows a precise comparison of the relative abundance of mutant alleles of a measured gene and hence the proportion of cells containing those mutations.

The results of this broad, ultrasensitive measurement of residual mutant alleles in AML indicate that “complete remission” is a misnomer, since induction chemotherapy failed to clear the bone marrow of leukemia-associated gene mutations in almost all AML patients studied here. Conversely, not all mutant alleles detected should be considered evidence of residual disease but rather of residual clonal aberrant hematopoiesis of unclear clinical significance. In particular, **DNMT3A** mutations, which typically persisted at high levels in remission and obscured the prognostic value of VAF levels of other mutated genes when analyzed in aggregate, offer no clear value for MRD monitoring. The individual prognostic value of other genes such as **TET2**, **ASXL1**, **RUNX1**, and **IDH1/2** that occasionally persist at high levels in remission will need to be refined in much larger patient cohorts.

The uncertainty regarding the prognostic value of mutations persisting at high levels and the degree to which they mark leukemic or preleukemic cell populations is underscored by the key finding of this study: the cellular reservoir containing mutations with the lowest VAF is fully clonally represented at relapse. This observation, made possible by high-sensitivity detection of multiple mutated genes within each specimen, argues against stochastic reacquisition of mutations in more abundant cell populations marked by high VAF mutations as the source of AML relapse and rather indicates a persistence of cells competent to generate leukemia possibly simply through regrowth.

The nature of chemoresistant, high-level, persistent mutated clones remains unclear. Do these represent preleukemic stem cells or progenitors that have acquired some but not all gene mutations of the dominant pre-therapy AML clone and lack the potential to relapse (28, 29), or are they remnant bystanders of clonal hematopoiesis of indeterminate potential (30–33)? Additional genomic and functional studies of various bone marrow cellular compartments will be necessary to determine the identity of these high-level cell populations and, more importantly, of the low-level cell populations implicated in AML relapse. A critical clinical need to better characterize these residual cells is underscored by this study: patients who achieved a reduction of VAF to less than 0.01% for all measured genes per remission (apart from **DNMT3A**) were far less likely to

---

**Figure 3. Discordant mutant gene VAFs in CR bone marrow samples from 38 AML patients with 2 or more measured genes.** Patients with discordant mutant gene VAFs demonstrating the common presence of oligoclonal aberrant hematopoiesis. **DNMT3A**, **ASXL1**, **TET2**, and **IDH1** frequently persisted at greater than 1% VAF in CR. The mean of duplicate measurements is depicted. The VAF percentage is on a log10 scale.
relapse by 3 years than those who did not (33% vs. 81%), but nevertheless, a third of these patients relapsed. Similarly, while patients who achieved a reduction of VAF to less than 0.01% in any 1 measured gene had a superior RFS at 3 years compared with those who did not (50% vs. 82%), half of these patients still relapsed. A deeper understanding of the nature of these resistant populations may lead to more effective strategies to maintain remission and eliminate the cell populations responsible for relapse.

Although this is the first study to our knowledge to broadly interrogate CR marrows for low-level residual mutations, there are some limitations. First, most, but not all, mutations identified in pre-therapy samples were measured in CR, raising the possibility that for a given case, there may be additional discordant VAFs and therefore an even higher overall frequency of oligoclonal hematopoiesis than we report. Next, we focused on detecting mutations in CR that were near or fully clonal in pre-therapy disease and so cannot comment on minor subclonal mutations present at diagnosis and their behavior in remission and at relapse. Finally, while the number of cases studied demonstrated a significant association of residual mutation VAF with survival, it was insufficient to perform multivariate survival analyses; additional studies with larger patient cohorts will be necessary to demonstrate an independent effect of ultrasensitive mutation detection on survival and to assess the effect of specific mutated genes on survival.

Methods

Patients. Between March 2005 and June 2011, 223 patients with AML (all non-M3) were enrolled in a translational study at the University of Michigan Comprehensive Cancer Center. Of the 223 patients, 72 had peripheral blood and bone marrow samples taken shortly before initiating induction chemotherapy and again after achieving morphological CR. Specimens were also collected at the time of relapse for 20 of these 72 patients. Buccal swabs were obtained from all patients upon enrollment in the study to provide a source of normal DNA.

Cell isolation and preparation of sample DNA. Peripheral blood or bone marrow mononuclear cells from AML patients were isolated using Ficoll gradient separation followed by cryopreservation. To eliminate the influence of nonmalignant cell contamination on VAF measurement in pretreatment and relapse samples, thawed AML cells were highly enriched for blasts using a combination of magnetic column and fluorescence-activated cell sorting, as previously described, to achieve a 90% or greater pure blast population, confirmed by cytopsin examination (34). DNA was extracted from this highly purified AML blast prepa-
and exon 4 of IDH2. For the 50 cases with either fewer than 2 mutations identified in these 13 genes or for which matched relapse samples were available, an additional 36 recurrently mutated genes in AML were sequenced using droplet-based amplicon library preparation (RainDance Technologies) and paired-end NGS on a MiSeq sequencer using MiSeq Reagent Kits, v2 (both from Illumina). The resulting FASTQ files underwent quality filtering, end trimming, reference genome alignment, and variant calling using NextGENe, version 2.4.1 (SoftGenetics). The lower limit of detection using this method was approximately 3% VAF. By including this NGS panel, ratio and from paired buccal mucosal cell preparations, as previously described (34). For CR specimens, DNA was extracted directly from thawed bone marrow mononuclear cells without additional enrichment.

**Genomic MRD target identification in pretreatment samples.** Sanger sequencing was performed on pretreatment gDNA isolated from highly pure blast populations for commonly mutated genes in AML (35, 36), including all coding exons and adjacent intronic sequences of CEBPA, ASXL1, TET2, DNMT3A, RUNX1, and BCORL1, along with exon 12 of NPM1, exons 13–15 and 20 of FLT3, exons 2–10 of TP53, exons 2–3 of KRAS and NRAS, exon 2 of IDH1,
ddPCR was then performed in duplicate on DNA from 72 samples extracted from Ficoll-enriched bone marrow aspirates obtained in CR to determine the VAF for all mutations. Each ddPCR run also included a positive control derived from a pre-therapy patient’s sample harboring the mutation of interest as well as a WT-only control derived from the Ly18 cell line and was performed as follows: a master mix containing 2× TaqMan Genotyping Master Mix, 40× Probe Mix, 25× Droplet Stabilizer (RainDance Technologies), and 50× Restriction Endonuclease (HaeIII or MseI; New England Bio-Labs) (38) was added to separate wells containing 500 ng gDNA, for a total reaction volume of 25 μL, and then incubated at room temperature for 15 minutes. Samples were then pipetted into a RainDrop Source Chip (RainDance Technologies), and 3.5–4.5 million droplets were generated with the RainDrop Source machine. The resulting emulsion was amplified using the previously derived optimal T9 with the following settings: 95°C for 10 minutes, and then 45 cycles at 95°C for 15 seconds and T9 for 60 seconds, and then 98°C for 10 minutes. Endpoint fluorescence was then measured with a RainDrop Sense machine and analyzed with RainDrop Analyst software. The WT and mutant signals of each pre-therapy sample were used to draw positive gates to then analyze VAFs of the diagnosis samples and WT-only controls. Of note, for mutations in which both ddPCR and NGS measurements of pretreatment-positive control specimens were available, the resulting VAFs were highly concordant (95% of measurements with less than a 5% difference in VAF percentage), with the exception of FLT3-ITD mutations. NGS variant calling of FLT3-ITD underestimated the VAF by approximately 2-fold compared with ddPCR (data not shown). Because of this discrepancy, NGS was used to quantify the VAF of all mutations in paired pre-therapy and relapse samples, except for FLT3-ITD, for which the values represent ddPCR measurements.

The general level of blank (LOB) across all probes was 0.0082% VAF and 0.0009% VAF for NPM1 probes in particular. Because the LOB could not be determined for probe sets used in a single case, the presence of a mutant signal was conservatively defined as the presence of 5 or more droplets in the mutant gate and a 5-fold higher number of mutant droplets in patients’ samples compared with the concurrently assayed WT control. Additional details on the assay parameters and LOB for all probe sets used in 2 or more cases are listed in Supplemental Table 4.

Statistics. Pair-wise assessment of differences in median VAF levels between various relapse groups was performed using the Mann-Whitney U test, corrected for multiple hypotheses testing with Bonferroni’s method. OS was defined as the time between AML diagnosis and the patient’s death. RFS was defined as the time between achievement of CR (documented by bone marrow examination and review of complete blood counts) and subsequent relapse. For patients still alive or relapse free, respectively, at the time of data analysis, the date of censoring was June 1, 2015. The log-rank test for trend was used to evaluate the significance of the trend of Kaplan-Meier estimates of improved RFS and OS with sequentially decreasing orders of magnitude of measured VAF. Univariate analyses were based on Kaplan-Meier estimates of survivor functions. Median survival times were estimated directly from the survivor function estimates. Significance levels for group-wise comparisons in the univariate analyses assessed whether the hazard ratio between groups differed from 1 in a Cox proportional hazards.
model. All statistical tests were 2ailed, and a P value of less than 0.05 was considered significant.

Study approval. This study was approved by the IRB of the University of Michigan (IRB0004-1022), and written informed consent was obtained from all participants prior to their enrollment in the study.

Author contributions
BP, ALJ, and QK performed the laboratory research. MT and ADR provided reagents and reviewed the manuscript. BP and SNM conceived the study, analyzed the data, and wrote the manuscript. SNM supervised the work.

Acknowledgments
SNM is a Scholar of the Leukemia and Lymphoma Society of America and is supported by the Hartman MDS fund and NIH grant R01-CA171972. BP is a Scholar of the American Society of Hematology and is supported by the Elsa U. Pardee Foundation.

Address correspondence to: Sami N. Malek, Professor, Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, 1500 E. Medical Center Drive, Ann Arbor, Michigan 48109-0936, USA. Phone: 734.763.2194; Email: smalek@med.umich.edu.

ALJ’s present address is: Department of Radiation Oncology, University of Alabama Birmingham, Birmingham, Alabama, USA.

ADR’s present address is: Department of Gastroenterology, Hepatology and Nutrition, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.