nature medicine

Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia

Sheng Li[1,18,19], Francine E Garrett-Bakelman[2,19], Stephen S Chung[3], Mathijs A Sanders[4], Todd Hricik[3], Franck Rapaport[3], Jay Patel[3], Richard Dillon[5], Priyanka Vijay[6], Anna L Brown[7–9], Alexander E Perl[10], Joy Cannon[10], Lars Bullinger[11], Selina Luger[10], Michael Becker[12], Ian D Lewis[7,9,13], Luen Bik To[9,13], Ruud Delwel[4], Bob Löwenberg[4], Hartmut Döhner[11], Konstanze Döhner[11], Monica L Guzman[2], Duane C Hassane[2], Gail J Roboz[2], David Grimwade[5], Peter J M Valk[4], Richard J D'Andrea[7–9], Martin Carroll[10], Christopher Y Park[14,15], Donna Neuberg[16], Ross Levine[3], Ari M Melnick[2] & Christopher E Mason[1,17]

Genetic heterogeneity contributes to clinical outcome and progression of most tumors, but little is known about allelic diversity for epigenetic compartments, and almost no data exist for acute myeloid leukemia (AML). We examined epigenetic heterogeneity as assessed by cytosine methylation within defined genomic loci with four CpGs (epialleles), somatic mutations, and transcriptomes of AML patient samples at serial time points. We observed that epigenetic allele burden is linked to inferior outcome and varies considerably during disease progression. Epigenetic and genetic allelic burden and patterning followed different patterns and kinetics during disease progression. We observed a subset of AMLs with high epiallele and low somatic mutation burden at diagnosis, a subset with high somatic mutation and lower epiallele burdens at diagnosis, and a subset with a mixed profile, suggesting distinct modes of tumor heterogeneity. Genes linked to promoter-associated epiallele shifts during tumor progression showed increased single-cell transcriptional variance and differential expression, suggesting functional impact on gene regulation. Thus, genetic and epigenetic heterogeneity can occur with distinct kinetics likely to affect the biological and clinical features of tumors.

AML is a predominantly fatal hematopoietic malignancy[1–3]. Even when leukemia cells appear to have been eradicated from the bone marrow after chemotherapy treatment, most patients relapse eventually. Therefore, it is important to understand how subpopulations of AML cells are resilient to chemotherapy and can give rise to progressively more refractory disease.

Several mechanisms have been hypothesized to endow subpopulations of cells within AML tumors with the capacity to survive exposure to therapeutic agents. These mechanisms include the presence of subsets of quiescent cancer stem cells with inherently greater self-renewal properties and reduced sensitivity to chemotherapy drugs[4]. Somatic mutations that can facilitate the expansion of putative leukemia stem cell populations can also emerge during disease establishment and progression[5–8].

Genetic heterogeneity within tumors can also increase evolutionary fitness[9,10]. Genetic diversity among individual cells in a tumor presumably provides the greatest chance for a subset of cells with particular combinations of mutations (single-nucleotide variants (SNVs), insertions and deletions (indels), copy-number aberrations, and/or translocations) to survive when challenged by cytotoxic drugs. Genetic heterogeneity has been appreciated in AML since early karyotyping studies[11,12], and tumor heterogeneity is relevant in solid tumors and lymphoid malignancies[13–19], where several studies have linked the degree of genetic heterogeneity to clinical outcome[16,20]. However, cells from patients with AML have a paucity of genetic lesions compared to most solid tumors[21,22], suggesting that other factors, such as epigenetic changes, may contribute to the aggressive behavior of AML.

[1]Department of Physiology and Biophysics and the HRH Prince Alwaleed Bin Talal Bin Abdulaziz Al-Saud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, New York, USA. [2]Division of Hematology–Medical Oncology, Department of Medicine, Weill Cornell Medicine, New York, New York, USA. [3]Leukemia Service, Department of Medicine, Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [4]Erasmus University Medical Center, Department of Hematology, Rotterdam, the Netherlands. [5]Department of Medical and Molecular Genetics, King's College London, Faculty of Life Sciences and Medicine, London, UK. [6]Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Medicine, New York, New York, USA. [7]Center for Cancer Biology, SA Pathology and University of South Australia, Adelaide, South Australia, Australia. [8]School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, South Australia, Australia. [9]Department of Hematology, SA Pathology and Royal Adelaide Hospital, Adelaide, South Australia, Australia. [10]Division of Hematology and Oncology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [11]Department of Internal Medicine III, University Hospital of Ulm, Ulm, Germany. [12]University of Rochester Medical Center, Rochester, New York, USA. [13]School of Medicine, University of Adelaide, Adelaide, South Australia, Australia. [14]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [15]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [16]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [17]The Feil Family Brain and Mind Research Institute, New York, New York, USA. [18]Present address: Department of Neurosurgery, Weill Cornell Medicine, New York, New York, USA. [19]These authors contributed equally to this work. Correspondence should be addressed to A.M.M. (amm2014@med.cornell.edu) or C.E.M. (chm2042@med.cornell.edu).

Received 14 March; accepted 11 May; published online 20 June 2016; doi:10.1038/nm.4125

Indeed, aberrant epigenetic patterning is a hallmark of AML[23,24]. Gain or loss of cytosine methylation at specific loci disrupts promoter activity and regulation of gene expression[23,25,26]. Epigenetic marks have great plasticity and may vary over time or through exposure to environmental stimuli. Thus, there is potential for epigenetic diversification to emerge in tumor cell populations and to change during disease progression. Sequence-based profiling can be used to determine quantitative cytosine methylation at base-pair resolution. DNA methylation at sequential cytosine–phosphate–guanine (CpG) dinucleotides can create a phased epigenetic pattern (epiallele) that represents a native 'cellular barcode'[27,28]. Such data can be used to measure epigenetic heterogeneity, which has been associated with worse clinical outcomes in lymphoid malignancies[28–31]. But lymphoid malignancies, in contrast to AML, display a high burden of somatic mutations and can arise from cell types such as B cells that are naturally prone to epigenetic heterogeneity[32].

It is has been suggested that genetic mutations and/or cytogenetic abnormalities and epigenetic modifications diversify along similar lines during disease progression[28,33]. This may reflect a selection process of genetic and epigenetic changes, which enhance tumor fitness through expression of oncogenes and/or repression of tumor suppressors, or a process of genetic diversification and clonal selection[34] in parallel with epigenetic modifications[35] that leads to tumor progression. However, the interplay and/or independence of genetic and epigenetic heterogeneity in AML have not been thoroughly explored. Here we performed a large-scale comparative genomics and epigenomics study in serial specimens obtained from patients with AML. We assembled a cohort of 138 clinically annotated, paired AML patient samples (taken at diagnosis (diagnostic) and at relapse; **Supplementary Tables 1** and **2**). Patient samples were characterized with genome-wide methylome sequencing using a modified version of reduced representation bisulfite sequencing[36,37] called enhanced reduced representation of bisulfite sequencing (ERRBS)[38,39]. In a subset of this cohort, we conducted whole-exome sequencing (WES) in paired patient samples with matched germline control DNA, RNA-seq in paired patient samples, and single-cell RNA-seq (scRNA-seq) for the patient for whom we had samples from the most time points. This enabled assessment of epigenetic heterogeneity and comparison to genetic and transcriptional variance during disease progression. The epigenetic features explored in this study are conceptually different from previously reported epigenetic signatures and uniform patterns of differentially methylated genes in AML[23,24]. Notably, we found that genetic and epigenetic allelic variations follow distinct and often independent kinetics and patterns. Moreover, shifts in epiallele composition are linked to transcriptional deregulation at the global and single-cell levels, and higher epiallele burden is linked to clinically more aggressive disease. Collectively, our data show that epigenetic allelic diversification occurs during AML establishment and also that disease progression and may be independent of the genetic landscape.

RESULTS

Epiallele burden is linked to poor clinical outcome

We calculated loci with epigenetic allelic variance in our patient cohort using the methclone compositional entropy equation[40]. By definition, the epigenetic state of each locus is comprised of cytosine methylation at four consecutive CpG dinucleotides. Each of the possible 16 CpG methylation patterns at these loci is an epiallele (**Supplementary Fig. 1**). Epigenetically shifted loci (eloci) occur when the epiallele proportions at these sites undergo a statistically significant entropy shift (calculated as change in Boltzmann entropy $\Delta S < 90$) in their composition when comparing two specimens[40]. The global metric eloci per million loci (EPM), which normalizes for the variable depth of coverage per specimen and the number of loci measured[40], was used to determine the overall magnitude of epiallele shifting across the genome and epiallele burden. Shifting can include gain and/or loss of epialleles between two specimens. The epiallele and eloci measurements were determined using methylome data from ERRBS (**Supplementary Table 3**) and validated in a subset of specimens using orthogonal methylome sequencing methods on two different platforms (Agilent and Roche; **Supplementary Fig. 2**).

We divided patients into two cohorts on the basis of their EPM values (highest and lowest 50%). There was no difference in the burden of somatic mutations between the groups (**Supplementary Fig. 3**). We then plotted clinical outcome as time to relapse, and we observed that patients with high EPM at diagnosis (compared to normal bone marrow (NBM) controls) had a shorter time to relapse as compared to the low-EPM cohort (P = 0.0396, Mantel–Cox log-rank test; **Fig. 1a**). This association was most significant for EPM assessed on the basis of promoter-associated eloci (P = 0.0077, Mantel–Cox log-rank test; **Fig. 1b**).

We also performed a multivariate analysis using linear regression to determine whether EPM retained its association with clinical outcome independently of other clinical variables. We found that EPM maintained a significant association with clinical outcome (P = 0.024, Cox proportional-hazards regression model) when age, white blood cell (WBC) count at the time of diagnosis, and sex were taken into account. None of these criteria were independently linked to time to relapse in this cohort (**Table 1**). In contrast to the results with EPM, the burden of somatic mutations as determined by WES (n = 48; **Supplementary Table 4**) was not significantly associated with time to relapse (P = 0.272, Mantel–Cox log-rank test; **Fig. 1c**).

Relapsed AMLs show variable changes in epiallele burden

To further understand the nature of epiallele shifting in AML, we compared diagnostic and relapsed AML specimen methylomes to NBM controls (**Fig. 2a**). All AML patients showed substantial epiallele shifting at diagnosis and at relapse compared to NBM controls (**Fig. 2b**,**c**). Although the degree of EPM increase varied, it occurred across all patients and disease stages (**Fig. 2d**). We next compared the

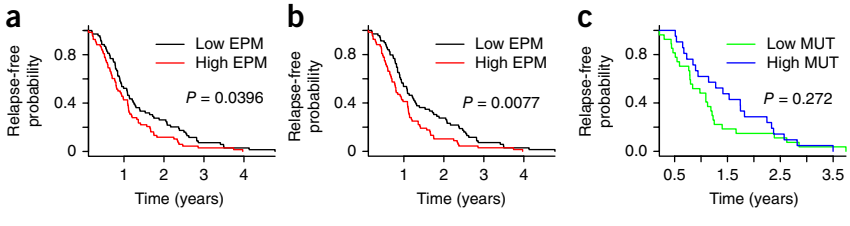

Figure 1 EPM levels at diagnosis compared to normal bone marrow controls segregate patients into two groups with distinct clinical outcomes. (**a**) Time-to-relapse analysis for patients (n = 137) with high (red) or low (black) EPM values at diagnosis compared to normal bone marrow. (**b**) Time-to-relapse analysis for patients (n = 137) with high (red) or low (black) EPM values assessed from promoter-annotated eloci (loci in promoters that were shared by at least 75% of patients were included). (**c**) Time-to-relapse analysis for patients with high (blue) or low (green) somatic mutation burden in diagnosis samples (n = 48). Mantel–Cox log rank test was used for the survival analysis (**a**–**c**).

Table 1 Multivariate analysis of EPM association with time to relapse

Variable	*P* value	Hazard ratio
EPM	0.024	1.559
Age	0.930	0.994
Sex	0.303	1.223
WBC count	0.339	0.999

Multivariate Cox proportional-hazards regression model using relapse time as response variable to test EPM and clinical parameters as variables in the entire cohort ($n = 127$).

EPM values of individual patients at relapse and at diagnosis (**Fig. 2a**). We observed substantial intra-patient variation (**Fig. 2d**,**e**). Although most patients showed a considerable degree of epiallele shifting during disease progression (median $\log_{10}$(EPM) = 2.29), the magnitude of this difference was highly variable. Indeed, a few patients (8%) showed no epigenetic changes (EPM = 0; **Fig. 2e**) between diagnosis and relapse. Overall, these data suggest that global epigenetic allele shifting is a universal feature of AML relative to NBM controls, but epiallele shifting is highly variable within individual patients during disease progression. Notably, the changes in epiallele burden were independent of patients' age, percentage blast purity of specimens, disease French–American–British (FAB) subtype, and abundance of somatic mutations (**Supplementary Fig. 4a**,**h**).

We also noted that eloci were significantly overrepresented at CpG islands and promoters ($P = 2.9 \times 10^{-3}$ and $P = 0.027$, respectively, compared to NBM controls; Wilcoxon signed-rank test) in AML patients at diagnosis (**Supplementary Fig. 5a**,**f**). Eloci were also significantly enriched at enhancers at diagnosis (active enhancers, $P = 1.4 \times 10^{-4}$; poised enhancers, $P = 7.5 \times 10^{-3}$; Wilcoxon signed-rank test; **Supplementary Fig. 5j**,**k**). At relapse enrichment for eloci was more significant at intronic and intergenic regions ($P = 5.3 \times 10^{-3}$ and $P = 7.6 \times 10^{-4}$, respectively; Wilcoxon signed-rank test; **Supplementary Fig. 5h**,**i**). These differences point to substantial disruption of regions with potential for gene-regulatory effects during AML progression.

Independent patterns of epigenetic and genetic diversity

The EPM measurement provides assessment of epigenetic allele shifting at the global level but does not provide information on how specific eloci vary during leukemia progression. The latter information would be similar to that from evaluation of individual genetic mutations in tumors. From the genetic perspective, AML progression has been examined in patients by next-generation sequencing and other techniques[5,41–44]. These studies showed that although most genetic lesions did not change during disease progression, some were gained or lost at relapse. A subset of patients showed no new somatic mutations at relapse.

Our current data set enabled simultaneous comparison of epigenetic and genetic allele burdens to discern whether they are correlated and how their composition might vary during leukemia progression. To address each of these points, we called mutations at diagnosis and relapse for each patient sample with WES data, then defined three categories of eloci: (i) eloci unique to diagnosis (diagnosis-specific), (ii) eloci unique to relapse (relapse-specific), and (iii) eloci present both at diagnosis and at relapse (shared).

We examined the eloci in several ways. First, we performed *K*-means clustering based on the proportion of diagnosis-specific, relapse-specific, or shared eloci present in each patient (**Supplementary Fig. 6a** and **Fig. 3a**). This analysis segregated AML patients into three clusters: (i) those having a significant predominance of diagnosis-specific eloci (cluster 1; $n = 57$, $P = 5.3 \times 10^{-11}$, Wilcoxon signed-rank test); (ii) those with mostly shared eloci and no predominance of diagnosis- or relapse-specific eloci (cluster 2; $n = 39$); and (iii) those with a significant predominance of relapse-specific eloci (cluster 3; $n = 42$, $P = 4.5 \times 10^{-13}$, Wilcoxon signed-rank test; **Fig. 3a**). None of these clusters was significantly linked with age, WBC count, or FAB classification (**Supplementary Fig. 6b–d**).

We next determined the abundance of somatic genetic mutated alleles (SNVs and indels) for patients for whom WES was available ($n = 48$). We examined the abundance of somatic mutations at diagnosis and relapse in each epigenetic cluster. We annotated somatic mutations as diagnosis-specific, relapse-specific, or shared between the two time points, similarly to the classification used above. The proportions of mutations in these categories did not follow the distributions observed for eloci in clusters 1 and 3 (cluster 1, $P = 0.89$, cluster 3, $P = 0.40$; Wilcoxon signed-rank tests; **Fig. 3b**). For example, samples in epigenetic cluster 1 showed significantly lower frequencies of somatic mutation at diagnosis and relapse than samples in cluster 3 (**Fig. 3c**, $P = 0.048$; **Fig. 3d**, $P = 0.008$, Wilcoxon rank-sum tests; **Supplementary Table 5**). Concordantly, a reciprocal analysis in which we used the median mutation burden per subject to divide patients revealed enrichment for higher mutation rates in cluster 3 (**Supplementary Fig. 6e**). These data indicate distinct combinations of genetic and epigenetic subtypes among epigenetic clusters.

We next examined whether specific individual somatic mutations might be linked to the tendency to develop eloci. Such considerations are of interest given that many recurrent somatic mutations in AML occur in genes encoding epigenetic modifiers[11,45]. However, there was no significant association between specific mutations in

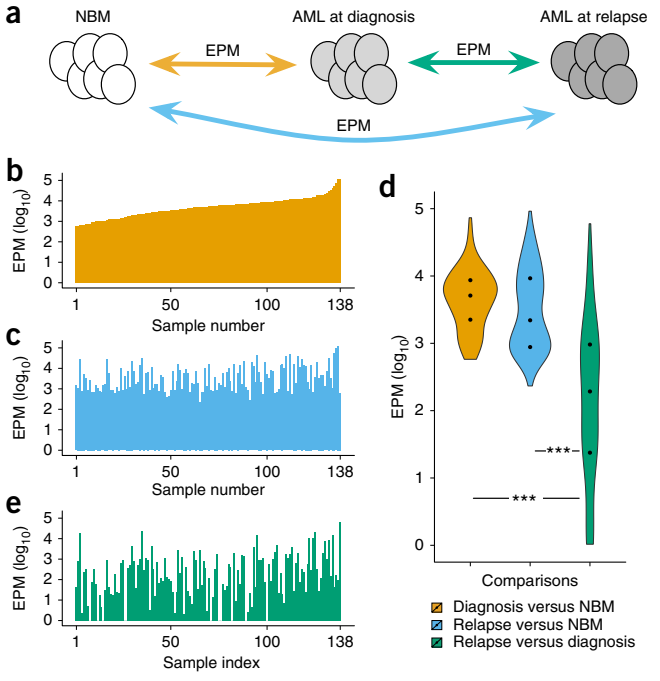

Figure 2 AML is characterized by high epiallele shift and variance. (**a**) Schematic of DNA methylation patterns compared among NBM controls, AML diagnostic samples, and AML relapse samples. (**b**,**c**) $\log_{10}$(EPM) values of diagnostic ($n = 138$) (**b**) and relapse ($n = 138$) (**c**) samples versus NBM controls ($n = 14$). (**d**) Violin plot of EPM values between AML patient samples and NBM controls and between relapse and diagnosis for individual patients. ***$P < 0.001$, Wilcoxon rank-sum tests. (**e**) $\log_{10}$(EPM) values between AML diagnosis and relapse samples.

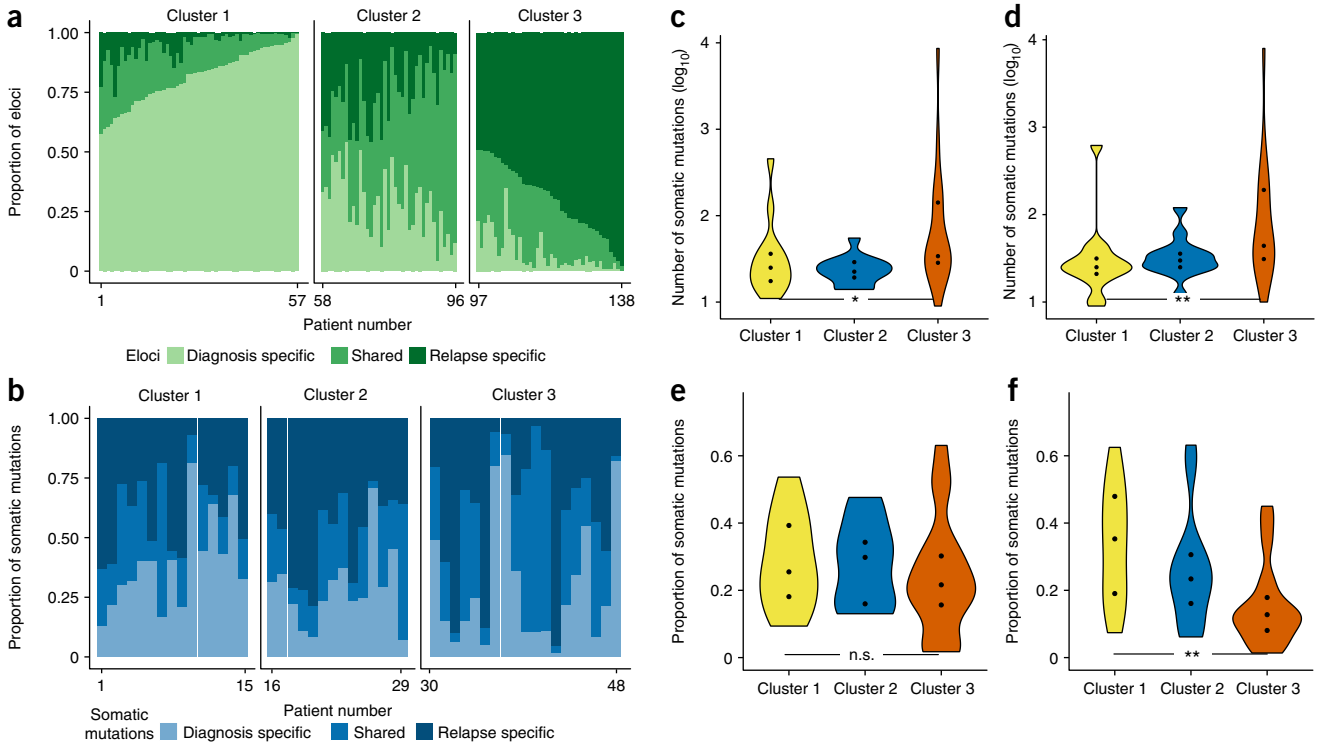

Figure 3 Disease-stage-specific epiallele patterns define unique subsets of AML patients. (**a**) Proportions of eloci that are diagnosis specific (light green), shared (green), or relapse specific (dark green) for each cluster, defined using *K*-means clustering. (**b**) Proportions of somatic mutations that were diagnosis specific (light blue), shared (blue), or relapse specific (dark blue) in patients for whom WES data were available within each epigenetic cluster described in **a**. (**c**,**d**) Number of somatic mutations ($\log_{10}$) for each epigenetic cluster at diagnosis (**c**; $P = 0.048$) or relapse (**d**; $P = 0.008$). (**e**,**f**) Proportion of somatic mutations whose VAFs increased (**e**; $P = 0.367$) or decreased (**f**; $P = 0.0012$) by ≥10% at relapse compared to diagnosis. $^*P < 0.05$; $^{**}P < 0.01$, Wilcoxon rank-sum test; n.s., not significant.

genes reported to be recurrent in AML[11,45] and the abundance of eloci (all genes $P > 0.05$, chi-squared test with Monte Carlo simulation; **Supplementary Fig. 6f** and **Supplementary Table 6**). As in many tumors, somatic mutations in AML occur in a heterogeneous pattern, representing the presence of genetically distinct subclones. Thus, it is possible that the genetic clonal complexity of AML might mirror the abundance of eloci. Using sciClone[46] we determined that 32.5% of patients showed increasing clonal complexity, 20% showed decreasing clonal complexity, and 47.5% showed no change in clonal complexity between diagnosis and relapse (**Supplementary Fig. 7**). Yet none of these clonal evolution patterns were enriched in epigenetically defined clusters ($P = 0.915$, chi-squared test; **Supplementary Fig. 8**), further indicating independent kinetics of genetic and epigenetic changes during AML evolution.

We also evaluated whether changes in variant allele frequency (VAF) of somatic mutations between diagnosis and relapse were linked to the pattern of eloci occurrence. Each patient exhibited variable numbers of alleles that decreased, increased, or remained stable in frequency between diagnosis and relapse samples. We focused our assessment on the patients in epigenetic clusters 1 and 3, which showed the greatest differences in eloci distributions between diagnosis and relapse samples. We assessed somatic mutations in diagnostic and relapse samples (compared to matched germline controls) for changes in VAFs during disease progression. An increase in VAF was defined as a minimum gain of 10% (**Fig. 3e**), and a decrease in VAF was defined by a minimum loss of 10% (**Fig. 3f**). Among patients in epigenetic clusters 1 and 3, there was no significant difference in the tendency to gain somatic mutation VAF at relapse (**Fig. 3e** and **Supplementary Table 5**). However, epigenetic cluster 3 patients showed a significantly lower proportion of somatic mutations with reduced VAF at relapse ($P = 0.0012$, Wilcoxon rank-sum test; **Fig. 3f** and **Supplementary Table 5**). These data suggest for the first time that disease progression in AML can be classified on the basis of patterns of DNA-methylation heterogeneity. Furthermore, epigenetic heterogeneity and genetic diversification do not necessarily follow the same kinetics between diagnosis and relapse.

We next investigated whether the differences in epiallele shift patterns between clusters 1 and 3 were associated with perturbations in the transcriptional landscape. We conducted RNA-seq in paired diagnostic and relapse samples representing patients from each cluster ($n = 19$; **Supplementary Table 7**). A supervised analysis (Online Methods) revealed 114 genes differentially expressed between diagnostic samples in clusters 1 and 3 (**Supplementary Table 8**). These included genes encoding the cell cycle inhibitor CDKN1A and the transcription factor SPI-B, which were expressed more highly in cluster 3, and the gene encoding the transcription factor FOXC1, which was expressed more highly in cluster 1. Overall, we detected upregulation of genes encoding protein kinases and other signaling molecules in cluster 1, and enrichment of genes linked to inflammation and immune response in cluster 3 (**Supplementary Table 9**). Differential expression of transcription factors inflammatory and signal-transduction gene sets between the two groups suggested a potential for alternative mechanisms of oncogenesis between the patients in epigenetic clusters 1 and 3.

Epiallele and mutation shifts throughout AML progression

To create a higher-resolution view of the progression kinetics of genetic and epigenetic heterogeneity in AML, we performed whole-genome

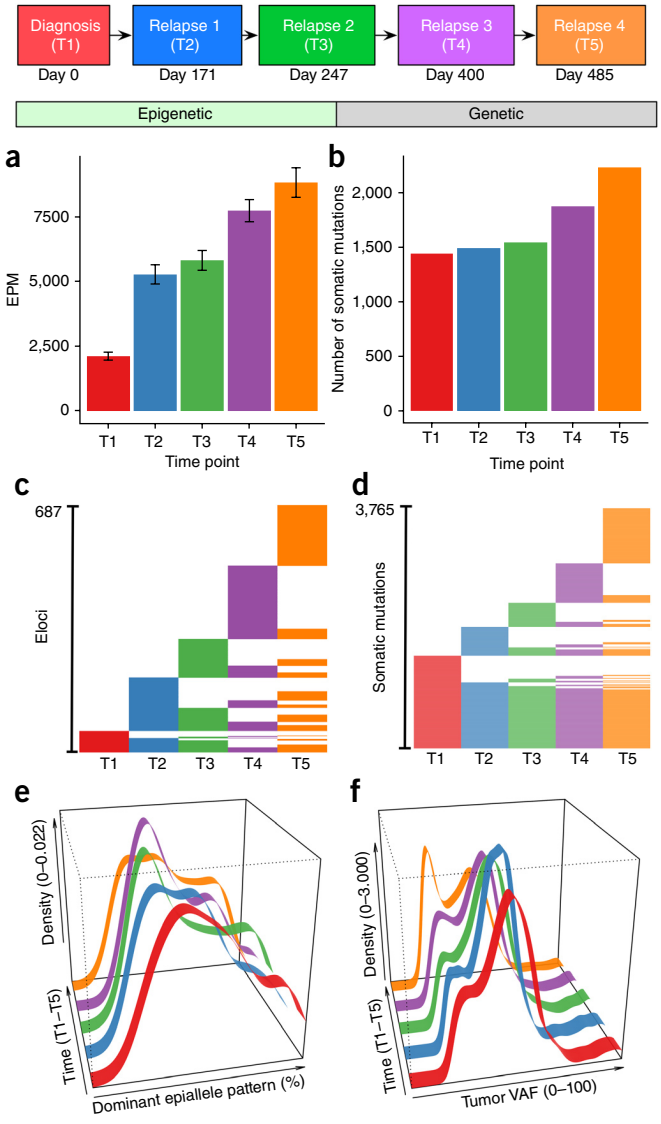

Figure 4 Assessment of epiallele shifts and genetic changes in serial samples from patient AML_130. (**a**) EPM in AML samples compared to NBM controls ($n = 14$) at time points T1–T5 (error bars, mean ± s.e.m.). (**b**) Somatic mutation burden at each time point. (**c**,**d**) Numbers of shared and unique eloci (**c**) and somatic mutations (**d**) among time points. (**e**) Density plot of the dominant epiallele frequency detected at eloci across all time points. (**f**) Density plot of the tumor VAFs detected at each time point.

sequencing (WGS) and ERRBS on leukemia cells from a patient (AML_130) from whom five serial specimens were available (T1 = diagnosis; T2–T5 = serial relapse time points; **Supplementary Tables 2–4**). We first evaluated the global epigenetic allele burden (EPM) for each time point compared to NBM controls (**Fig. 4a**). We observed a marked increase (150%) in epiallele burden occurring at T2. In contrast, the T2 sample showed little increase (2.7%) in the number of somatic mutations compared to germline DNA from the same patient (**Fig. 4b**). Increases in EPM at time points T3–T5 were relatively smaller; however, at T4 there was a larger increase (29%) in the abundance of somatic mutations than at other time points (5.4%), indicating that the initial epigenetic shift at T2 was antecedent to the later genetic evolution in AML cells from this patient at T4. Furthermore, genomic analysis did not yield evidence of acquisition or loss of a somatic mutation linked to known epigenetic modifier genes at T2 that might explain the jump in epiallele diversity at T2 (**Supplementary Table 10**).

We next examined the unique occurrences of individual eloci and somatic mutations at each time point. Analysis of eloci revealed a tendency of the epigenome to evolve continuously over time (**Fig. 4c**). A majority of eloci (65.79%) were unique to a single time point; far fewer eloci were shared among two, three, or four time points (18.78%, 10.04%, and 3.64%, respectively), and only 1.75% of eloci were shared across all five time points (**Supplementary Fig. 9a**). In sharp contrast, genetic alleles were much more stable; 24.6% of somatic alleles were detectable over all five time points (**Fig. 4d**, **Supplementary Fig. 9b**, and **Supplementary Table 11**). Hence, genetically defined, mutated alleles showed significantly greater stability over time than eloci ($P < 2.2 \times 10^{-16}$, chi-squared test).

To better discern the relationship between epigenetic and genetic heterogeneity, we examined which of the 16 possible epialleles was dominant within each elocus at each time point. The proportion of these specific epialleles was used as a measure of epigenetic clonality. Reciprocally, to evaluate genetic clonal complexity, we examined the VAF of somatic mutations at each time point. Notably, we observed a shift in dominant epiallele frequency over time. At early time points, the dominant epialleles tended to occur at a frequency of >50% (**Fig. 4e**; median = 53%). A marked shift in epiallele frequency occurred at T3, when the dominant pattern shifted to a lower frequency (25%). By T5 the leukemia cells were again enriched for eloci with higher dominant epiallele frequency. However, the pattern was different when considering genetic VAF. At early time points (T1–T3), the median VAF was 44%. A significant shift to low-frequency genetic mutations (VAFs) occurred (**Fig. 4f**) only at T5. Genetic alleles were also distinct from dominant epialleles in that they exhibited a tighter allele frequency distribution ($P = 0.008$, Wilcoxon rank-sum test; **Fig. 4e**,**f**), perhaps because the epigenome is more plastic than the genome. Collectively, these data support the notion that epigenetic and genetic heterogeneity in AML are not necessarily linked and may follow distinct patterns and distributions during disease progression.

Promoter epiallele shifts linked to transcription variance

Given that transcription is regulated through epigenetic marks, we next determined whether the presence of eloci was linked to alterations in gene expression. We focused on genes containing eloci within promoters, derived from the comparison of diagnostic to relapse specimens ($n = 19$ pairs; **Supplementary Table 7**). Genes containing shifts in epiallele composition at promoters (promoter eloci) showed significantly greater variance in transcript abundance between relapse and diagnosis than genes without such epiallele shifts in their promoters ($P < 0.001$; Wilcoxon rank-sum test; **Fig. 5a**,**b**). We next focused our analysis on genes that were significantly differentially expressed in relapse versus diagnostic AML samples. Here again, a higher proportion of differentially expressed transcripts were significantly associated with promoters harboring eloci than with promoters not harboring eloci ($P < 0.001$; Wilcoxon signed-rank test; **Fig. 5c**).

Finally, we used an additional approach to examine potential links between promoter epigenetic heterogeneity and transcription. We performed scRNA-seq in 96 cells from the first relapse sample (T2) from patient AML_130. We found that the genes with higher epiallele heterogeneity within their promoters also had significantly higher levels of cell-to-cell transcriptional heterogeneity as measured by the coefficient of variation ($P < 2.2 \times 10^{-16}$, ANOVA, **Fig. 5d**). Hence the presence of eloci at gene promoters resulted in greater tendency of associated genes to show deregulated expression.

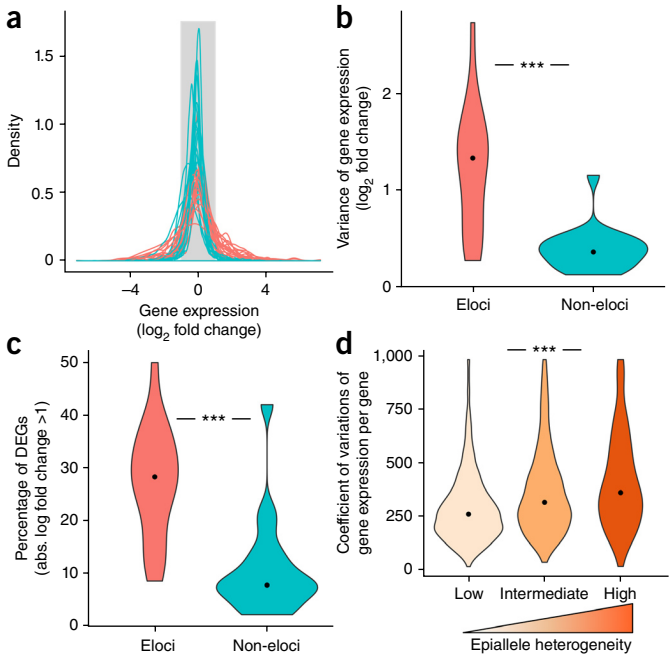

Figure 5 Transcriptional variance is associated with high epiallele shift at promoters. (**a**) Density plot of $\log_2$ fold change in transcript levels of genes with (red) and without (blue) eloci in their promoters, as measured from bulk cell populations ($n = 19$ paired patient samples). (**b**) Violin plot of variance in transcript expression from genes with or without eloci in their promoters in bulk cell populations (Wilcoxon signed-rank test; $P = 3.82 \times 10^{-6}$). (**c**) Violin plot of the percentage of genes that are differentially expressed (DEGs; absolute $\log_2$ fold change >1; Wilcoxon signed-rank test: $P = 3.82 \times 10^{-6}$) with or without eloci in their promoters in bulk cell populations. (**d**) Violin plots of variance in transcript expression level as measured by scRNA-seq (AML_130 relapse sample T2; $n = 96$ cells) and association (ANOVA test, $P < 2.2 \times 10^{-16}$) with low (0.05), intermediate (0.05 0.2) and high (≥0.2) epipolymorphism within their respective gene promoters. ***$P < 0.001$, Wilcoxon signed-rank and ANOVA tests.

DISCUSSION

AML is well suited for examining questions about epigenetic heterogeneity, as this type of tumor generally manifests a relative paucity of genetic lesions[21,22] but still shows genetic clonal complexity and clonal evolution during progression. Mazor *et al.*[33] recently examined DNA methylation at individual CpG sites and somatic mutations during tumor progression in gliomas, where a phylogenetic analysis showed a co-dependency of change for the two features examined. However, an examination of overall DNA methylation at CpG sites with array-based measurements cannot fully address epigenetic heterogeneity as measured by epiallele diversification; thus, it has remained unclear whether epigenetic alleles and genetic alleles follow similar courses during disease evolution. Although others have documented epigenetic heterogeneity in B cell neoplasms[28,29,32], these tumors arise mostly from cells that display intrinsic epigenetic heterogeneity, and the studies did not explore the link between genetic and epigenetic allelic diversity during tumor progression. This is the first report of dynamic epiallele shifting in a human tumor. Previous reports have used DNA methylation heterogeneity metrics such as epipolymorphism approaches, which calculate diversity within individual specimens, rather than methclone compositional entropy, which measures shifting between samples[28,29] and, hence, different properties of the epigenome.

Here we establish that tumor genetic and epigenetic heterogeneity in AML may arise as independent, biologically distinct phenomena, each presumably having a unique functional significance. The degree of leukemic epigenetic allelic burden was independent of age and other clinical parameters. There was no apparent association between the degree of epigenetic heterogeneity and the presence of somatic mutations affecting epigenetic modifier genes such as *DNMT3A*, *TET2*, and *IDH1* or *IDH2*. We also explored whether dominant epigenetic alleles behaved in a similar or distinct manner from that of genetic alleles during clonal evolution. We found no association between the kinetics and pattern of genetic and epigenetic alleles during leukemic progression.

It is notable that AML patients can be classified on the basis of epiallele pattern kinetics and somatic mutation burdens during disease progression. Our results suggest that in at least some cases, AML patients may be divided at diagnosis as having high epiallele diversity and low somatic mutation burden (cluster 1, epigenetically driven) or as having lower epiallele diversity and higher mutation burden (cluster 3, genetically-driven) (**Fig. 3a**,**c**). The latter develop increasing epigenetic diversity during progression (**Fig. 3a**). In both cases, genetic clonal composition remains predominantly stable (**Supplementary Table 5**). Most notably, epigenetic instability was not necessarily linked to genetic instability or specific somatic mutations. This observation may point to alternative modes of dominant heterogeneity in newly diagnosed patients: one genetic and one epigenetic. Tumor heterogeneity in AML can thus not be determined strictly by genetic or epigenetic analysis alone, and relying only on genetic mapping may underestimate the true tumor diversity. We speculate that *de novo* acquisition of eloci might represent the response of hematopoietic cells to environmental stresses at various points during the development and/or progression of disease. Conversely, exposure to chemotherapy may induce epigenetic plasticity, contributing to the increased number of eloci seen at relapse in some patients. Indeed, longitudinal profiling of a patient treated with four different regimens showed distinct patterns of epiallele shifting after each treatment. It is also possible that unrecognized mutations or alterations in coding or noncoding regions could destabilize the epigenome at specific loci and contribute to epigenetic heterogeneity. One example is the protein CTCF, which can regulate cytosine methylation patterning and boundaries[47]. A CTCF haploinsufficiency mouse model results in the acquisition of epigenetic hypervariability hot spots and development of tumors[48]. In the case of B cells, epigenetic heterogeneity may be linked to the actions of activation-induced cytosine deaminase[48]. In the case of AML, biological differences among patients with different modes of eloci progression patterns may be mediated through differential expression of hematopoietic transcription factors and functional gene sets.

The acquisition of eloci can also lead to functional consequences, as we found that genes containing eloci at their promoters showed more pronounced transcriptional variability and differential expression during disease progression. A recent scRNA-seq study on newly diagnosed CLL patients also showed that sites of epigenetic variability had greater cell-to-cell transcriptional plasticity[28]. A second indication of biological relevance is the inferior outcome of patients with higher EPM burdens in both univariate and multivariate analyses. The fact that the link to inferior outcome is even stronger for promoter epialleles suggests that eloci that perturb transcriptional regulation are particularly important in contributing to disease phenotype.

Using an epipolymorphism-based approach, we observed decreasing DNA methylation heterogeneity upon AML progression, which matches a previous report in diffuse large B cell lymphoma (DLBCL)[29] (**Supplementary Fig. 10a**). However, unlike in DLBCL, the abundance

of epipolymorphisms in AML did not segregate patients into distinct clinical groups on the basis of progression-free survival (**Supplementary Fig. 10b**); only the EPM separated the cohorts. Epiallele shifting as detected by methclone is thus a candidate biomarker in AML and may be indicative of more aggressive disease, perhaps reflecting greater epigenetic plasticity and adaptability in certain tumors. Given the heterogeneous nature of AML, our study was underpowered to detect specific eloci that could be used as outcome biomarkers, although there was a trend for a set of 21 promoter-annotated eloci to associate with shorter time to relapse (**Supplementary Table 12**).

To our knowledge, this study is the first to examine the longitudinal interplay of genetic and epigenetic heterogeneity in AML. It is possible that genetic and epigenetic heterogeneity are linked by other means in AML or in different manners in other types of tumors. Studies in larger cohorts using improved sequencing technologies are likely to yield additional insights. Collectively, our data suggest that epigenetic heterogeneity has important clinical and functional impacts in AML and does not fully follow the kinetics and patterns in the genetic compartment, thus creating additional means by which a tumor can evolve.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Database of Genotypes and Phenotypes (dbGaP): data from this study are available under the accession number phs001027.v1.p1.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank J. Phillips, J. Ishii, L. Wang, J. Busuttil, T. Lee, P. Zumbo, J. Gandara, and A. Zeilemaker for technical support; C. Sheridan for technical support, assistance with organization, and maintenance of sample database and banking; M. Perugini, D. Iarossi, and I.S. Tiong for assistance with clinical database management; and Y. Neelamraju, Z. Li, J. Glass, and M.R. De Massy for data annotation and management. Next-generation sequencing protocols and sequencing were performed by the Weill Cornell Medicine Epigenomics Core and the New York Genome Center. We thank A. Viale from the Integrated Genomics Operation and N. Socci from the bioinformatics core at Memorial Sloan Kettering Cancer Center for sequencing services. We thank the South Australian Cancer Research Biobank for access to clinical samples. We thank F. Michor for recommendations regarding data analyses. This work was supported by Starr Cancer Consortium grant I4-A442 (A.M.M., R.L., and C.E.M.), STARR Cancer Consortium grant I7-A765 and I9-A9-071 (C.E.M.), the Irma T. Hirschl and Monique Weill-Caulier Charitable Trusts, Bert L. and N. Kuggie Vallee Foundation and the WorldQuant Foundation, Pershing Square Sohn Cancer Research Alliance, and NASA (NNX14AH50G) (C.E.M.); LLS SCOR 7006-13 (A.M.M.); NCI K08CA169055 (F.E.G.-B.); an American Society of Hematology (ASHAMFDP-20121) award under the ASH-AMFDP partnership with the Robert Wood Johnson Foundation and ASH/EHA TRTH (F.E.G.-B.); a Doris Duke Medical Foundation, Leukemia and Lymphoma Society Translational Research Program, and Geoffrey Beene Cancer Center (C.Y.P.); a Leukaemia and Lymphoma Research award (D.G. and R. Dillon); German Research Foundation (DFG) grant SFB 1074 (project B3; K.D. and L.B.); DFG Heisenberg-Stipendium BU 1339/3-1 (L.B.); an Australian National Health and Medical Research Council and the Royal Adelaide Hospital Contributing Haematologists Fund financial support (R.J.D., A.L.B., and I.D.L.); US National Institutes of Health R01CA102031 (G.J.R. and M.L.G.) and R01NS076465 (C.E.M. and A.M.M.); and Leukemia Fighters funding (G.J.R., M.L.G., and D.C.H.).

AUTHOR CONTRIBUTIONS

A.M.M. and C.E.M. conceived of the studies, designed analytical approaches, analyzed results, wrote the manuscript, and designed experiments with F.E.G.-B. S. Li conceived of computational analyses, wrote code and performed computational analysis, wrote the manuscript, and generated figures. F.E.G.-B. performed, coordinated and/or supervised all patient sample experimental procedures, performed computational and bench experimental data management and analyses, and wrote manuscript. S.S.C. and R. Dillon performed experiments (flow cytometry and sorting of subject samples) and associated data analysis. T.H., F.R., and J.P. performed computational analyses. M.A.S. and P.J.M.V. performed experiments (exome capture) and associated computational analysis. A.L.B., A.E.P., J.C., L.B., S. Luger, M.B., I.D.L., L.B.T., B.L., H.D., K.D., P.J.M.V., R.J.D., and M.C. coordinated patient sample collection and analyzed clinical data. D.N. assisted with statistical analyses. P.V. performed single-cell RNA-seq library preparation. R. Delwel, M.L.G., D.C.H., G.J.R., D.G., C.Y.P., and R.L. helped with sample collection, writing, analysis, and patient annotation. All authors read, edited, and approved the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Roboz, G.J. Current treatment of acute myeloid leukemia. *Curr. Opin. Oncol.* **24**, 711–719 (2012).
2. Grimwade, D. *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5,876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* **116**, 354–365 (2010).
3. Döhner, H. *et al.* Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* **115**, 453–474 (2010).
4. Ishikawa, F. *et al.* Chemotherapy-resistant human AML stem cells home to and engraft within the bone marrow endosteal region. *Nat. Biotechnol.* **25**, 1315–1321 (2007).
5. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
6. McKerrell, T. *et al.* Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep.* **10**, 1239–1245 (2015).
7. Moran-Crusio, K. *et al. Tet2* loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
8. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
9. Landau, D.A., Carter, S.L., Getz, G. & Wu, C.J. Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia* **28**, 34–43 (2014).
10. Klco, J.M. *et al.* Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell* **25**, 379–392 (2014).
11. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
12. Testa, J.R., Mintz, U., Rowley, J.D., Vardiman, J.W. & Golomb, H.M. Evolution of karyotypes in acute nonlymphocytic leukemia. *Cancer Res.* **39**, 3619–3627 (1979).
13. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
14. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumors. *Nature* **490**, 61–70 (2012).
15. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015).
16. Landau, D.A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
17. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl. Acad. Sci. USA* **110**, 4009–4014 (2013).
18. Zhang, J. *et al.* Genetic heterogeneity of diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. USA* **110**, 1398–1403 (2013).
19. Landau, D.A. *et al.* Mutations driving CLL and their evolution in progression and relapse. *Nature* **526**, 525–530 (2015).
20. Mroz, E.A., Tward, A.D., Hammon, R.J., Ren, Y. & Rocco, J.W. Intratumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. *PLoS Med.* **12**, e1001786 (2015).
21. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
22. Lawrence, M.S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumor types. *Nature* **505**, 495–501 (2014).
23. Figueroa, M.E. *et al.* DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* **17**, 13–27 (2010).
24. Figueroa, M.E. *et al.* Leukemic *IDH1* and *IDH2* mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553–567 (2010).
25. Rampal, R. *et al.* DNA hydroxymethylation profiling reveals that *WT1* mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep.* **9**, 1841–1855 (2014).

26. Shih, A.H. *et al.* Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell* **27**, 502–515 (2015).
27. Landan, G. *et al.* Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues. *Nat. Genet.* **44**, 1207–1214 (2012).
28. Landau, D.A. *et al.* Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813–825 (2014).
29. Pan, H. *et al.* Epigenomic evolution in diffuse large B cell lymphomas. *Nat. Commun.* **6**, 6921 (2015).
30. Chambwe, N. *et al.* Variability in DNA methylation defines novel epigenetic subgroups of DLBCL associated with different clinical outcomes. *Blood* **123**, 1699–1708 (2014).
31. De, S. *et al.* Aberration in DNA methylation in B cell lymphomas has a complex origin and increases with disease severity. *PLoS Genet.* **9**, e1003137 (2013).
32. Shaknovich, R. *et al.* DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B cell differentiation. *Blood* **118**, 3559–3569 (2011).
33. Mazor, T. *et al.* DNA methylation and somatic mutations converge on the cell cycle and define similar evolutionary histories in brain tumors. *Cancer Cell* **28**, 307–317 (2015).
34. Greaves, M. & Maley, C.C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
35. Feinberg, A.P., Koldobskiy, M.A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer etiology and progression. *Nat. Rev. Genet.* **17**, 284–299 (2016).
36. Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868–5877 (2005).
37. Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468–481 (2011).
38. Akalin, A. *et al.* Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet.* **8**, e1002781 (2012).
39. Garrett-Bakelman, F.E. *et al.* Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base-pair resolution. *J. Vis. Exp.* **96**, e52246 (2015).
40. Li, S. *et al.* Dynamic evolution of clonal epi-alleles revealed by methclone. *Genome Biol.* **15**, 472 (2014).
41. Parkin, B. *et al.* Clonal evolution and devolution after chemotherapy in adult acute myelogenous leukemia. *Blood* **121**, 369–377 (2013).
42. Krönke, J. *et al.* Clonal evolution in relapsed *NPM1*-mutated acute myeloid leukemia. *Blood* **122**, 100–108 (2013).
43. Tawana, K. *et al.* Disease evolution and outcomes in familial AML with germline *CEBPA* mutations. *Blood* **126**, 1214–1223 (2015).
44. Chou, W.C. *et al.* The prognostic impact and stability of isocitrate dehydrogenase 2 mutation in adult patients with acute myeloid leukemia. *Leukemia* **25**, 246–253 (2011).
45. Patel, J.P. *et al.* Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1079–1089 (2012).
46. Miller, C.A. *et al.* SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput. Biol.* **10**, e1003665 (2014).
47. Ong, C.T. & Corces, V.G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **15**, 234–246 (2014).
48. Kemp, C.J. *et al.* CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* **7**, 1020–1029 (2014).

ONLINE METHODS

Patient characteristics. 138 clinically annotated, paired samples (diagnosis and relapse) from AML patients (86 males and 52 females) seen at medical centers in Australia, Germany, the Netherlands, and the United States were collected. Patients with acute promyelocytic leukemia were excluded. All patients were treated according to the protocols of the corresponding institutes and hospitals. The clinical and molecular characteristics of these patients are summarized in **Supplementary Table 1**, and detailed descriptions are provided in **Supplementary Table 2**. All patients were treated with combination chemotherapy (cytarabine arabinoside and an anthracycline) during induction phase followed by consolidation chemotherapy treatment with or without stem cell transplantation in first remission, per clinical center standards. Samples from further serial time points were available for patient AML_130, a male with previous high-dose radiation exposure who developed AML at age 57. Samples were available from the diagnostic time point (T1, leukapheresis sample) and during disease progression (T2T5) as follows: he underwent standard induction treatment (cytarabine + idarubicin) and achieved complete remission but developed relapse before initiation of consolidation treatment. The patient was then treated with high-dose cytarabine without response (sample T2; peripheral blood sample), combination chemotherapy (sirolimus, mitoxantrone, etoposide, and cytarabine) without response (sample T3; peripheral blood sample), combination chemotherapy without response (cyclophosphamide followed by clofarabine), hydroxyurea and external beam radiation to tonsillar site of extramedullary hematopoiesis (sample T4; bone marrow aspirate), and investigational FLT3 tyrosine kinase inhibitor (KW-2449) followed by an allogeneic stem cell transplant (sample T5; leukapharesis sample). No treatments beyond the chemotherapy administered immediately after diagnosis in AML_130 induced a complete remission.

Sample collection and processing. Donors (AML patients and individuals without known hematological malignancies) gave informed consent according to the Declaration of Helsinki for collection and use of sample materials in research protocols at the following clinical centers: Erasmus Medical Center (protocol number MEC-2015-155), Royal Adelaide Hospital and SA Pathology (Adelaide, South Australia; 1998–present), University of Pennsylvania (protocol number 703185), University of Rochester Medical Center (protocol number URCC ULEU07047), and the University Hospital of Ulm (protocol number 148/10). Study protocols were approved by the institutional review boards of corresponding institutes and hospitals (protocols noted above), and at Weill Cornell Medicine (WCM; protocol number 0805009783). For cryopreserved specimens from AML patients at the Royal Adelaide Hospital and SA Pathology (Adelaide, South Australia) collected and stored before 1998, the requirement for informed consent was waived by the Royal Adelaide Hospital Human Research Ethics Committee (RAH HREC, protocol number 110304b), and use of the samples obtained in this specific research study was approved by the RAH HREC on 10 September 2010. Of 140 paired patient samples collected, 138 were successfully processed. Patient samples were collected at the time of diagnosis and within 3 months of clinically determined relapse. Samples were subjected to Ficoll separation on the day of collection and viably frozen or immediately subjected to nucleic acid extractions using standard techniques. Deidentified samples were then provided to WCM. Viably frozen cells were available for further processing from patients AML_102 through AML_140, including serial samples from AML_130 (T1–T5). These samples were thawed and depleted of $CD3^+$ and $CD19^+$ cells using magnetic beads per the manufacturer's recommendations (Miltenyi Biotec), yielding a blast percentage enrichment of 89. ± 9.5% as confirmed by post-separation flow cytometry. Lymphocytes were then isolated for germline controls using flow sorting for $CD3^+$ and $CD19^+$ cells on a FACSAria II cell sorter (BD Biosciences), yielding >85% purity as confirmed by post-sort flow cytometry analysis. Germline DNA was isolated from *ex vivo*–expanded lymphocytes for patient samples AML_074 through AML_101. Briefly, the procedure for *ex vivo* expansion was as follows: 1 million bone marrow cells were cultured in RPMI 1640 with L-glutamine and 10% FCS. T cell expansion was induced with T cell expander Dynabeads CD3/CD28 (Thermo Fisher). After 1–3 d rIL2/ml was added. Cells were subsequently purified to > 98% with CD3 microBeads (MACS Miltenyi Biotec). Flow cytometry analysis and sorting was performed using CD19 PerCP–Cy5.5 (HIB19) and CD3 PerCP–Cy5.5 (HIT3a) from Biolegend (San Diego) and human CD45 APC–A780 (2D1) from eBioScience. Validation for each antibody is provided on the manufacturer's website (http://www.biolegend.com/) and at Antibodypedia (http://www.antibodypedia.com/gene/4237/CD19/antibody/539284/302214; http://www.antibodypedia.com/gene/4581/CD3D/antibody/539101/300437) and 1DegreeBio (1DB 1DB-001-0001093758). Lymphocytes were defined as $CD45^{high}SSC^{low}$ $CD19^+$ or $CD3^+$ cells, and leukemic blasts were defined as $CD19^-CD3^-$ $CD45^{low}SSC^{low}$ cells[49]. DNA and RNA were extracted from the isolated cells using standard techniques. For patient samples AML_001 through AML_101, DNA and RNA were extracted from mononuclear cell layers using standard techniques. Normal bone marrow controls (NBM controls: $CD34^+$ mononuclear cells; 7 males and 7 females) were purchased from AllCells ($n = 5$) or isolated using magnetic bead positive selection for $CD34^+$ cells per the manufacturer's recommendations (Miltenyi Biotec) from freshly collected bone marrow samples from volunteers without known hematological malignancies ($n = 9$). Purity was verified using $CD34^+$ PE/CF594 (563) from BD Biosciences (validation provided on manufacturer's website, http://www.bdbiosciences.com/in/index.jsp) in flow cytometry post-separations to greater than 90%. All flow cytometry data were analyzed using FlowJo (TreeStar).

Enhanced reduced representation bisulfite sequencing. ERRBS is a modified version of RRBS[36,37]. Libraries were prepared by MspI restriction enzyme digestion of high-molecular-weight genomic DNA, followed by end repair, size selection, bisulfite conversion and library amplification as previously described[38,39]. Libraries were sequenced on a HiSeq 2000 Illumina machine using 75-bp single-end reads. Data alignment was performed to human genome hg19 as previously described[38,39]. Average reads per sample were 138,596,488 with a mean alignment rate of unique reads of 63.7%, covering on average 4,399,235 CpGs per sample at a minimum threshold coverage of 10×. Average coverage depth per CpG was 72.3, and average bisulfite conversion was 99.86%, determined as previously described[38,39]. A subset of patient samples were processed using SeqCap Epi 4M CpGiant Enrichment kit (NimbleGen-Roche) and SureSelect^{XT} Human Methyl-Seq (Agilent Technologies) per the manufacturer's recommendations. These were sequenced using a paired-end 100-bp approach on a HiSeq 2000 (Illumina). See **Supplementary Table 3** for detailed sequencing statistics.

Data analysis. The programming framework R[50] version 3.2.1 was used for data analyses. Specific R packages and other tools and software used for analyses are described below and in **Supplementary Methods**.

Code availability. Analysis scripts used in this manuscript are included in **Supplementary Methods** and have been deposited at https://github.com/ShengLi/relapsed_AML.

ERRBS data analysis. We performed bisulfite-treated read alignment to hg19 genome and methylation calls as previously described[38,39]. Briefly, the adaptor sequences were removed by FAR software[51]. Preprocessed reads were then aligned to human genome reference hg19 using the bismark alignment software.

Epiallele shift analysis was performed using methclone[40]. Briefly, the epiallele patterns of compositional changes between different samples were evaluated using methclone to calculate the combinatorial entropy (ΔS) change of epialleles at each locus. This analysis outputs the loci with a ranked list of epiallele changes defined by the entropy change. We first calculated the foreground combinatorial entropy using the epiallele composition detected, which is the sum of the entropy of each locus for two samples considered. We then calculated the background combinatorial entropy, using the epiallele composition after uniformly mixing all patterns of epialleles between two samples. The sum entropy was determined, and the entropy changes were defined by the difference between foreground and background combinatorial entropy values. Epiallele shifts per million loci (EPM) is a normalized measure of the global epiallele changes with $\Delta S < -90$ as the cut-off (eloci). Comparisons of AML specimens to NBM controls report the average EPM determined from assessing each AML case against each NBM ($n = 14$). The diagnosis-specific eloci are eloci that are detectable only at diagnosis stage when compared to NBM controls. The relapse-specific eloci are eloci that are detectable only at relapse stage when compared to NBM controls. The shared

doi:10.1038/nm.4125

eloci are eloci that are detectable at both stages. The epigenetic clusters were determined using *K*-means clustering ($K = 3$, gap statistic for estimating the number of clusters) based on the proportion of eloci that are diagnosis specific, shared, or relapse specific. Within clusters 1 and 3, respectively, the proportion of eloci for diagnosis and relapse were compared using Wilcoxon signed-rank tests within each cluster to measure the significance of dominance. Finally, we assessed the distribution of patients from each clinical center ($n = 5$) in the epigenetically defined clusters. Using multinomial logistic regression analysis to model the clusters with clinical and demographic features, we determined that the center identity was not equally distributed among the three clusters. However, we confirmed no known technical or clinical etiology was related to this distribution for each center (age, WBC, gender, FAB classification, mutations, cytogenetics), and the sample sizes for the subcohorts were not powered to detect subclusters within each cohort.

Intratumor global methylation heterogeneity (MH) was assessed as previously described[29]. Briefly, epipolymorphisms were determined using proportions of DNA methylation patterns at four adjacent CpGs ($1 - \Sigma_{i=1}^{16} p_i^2$, where p_i is the fraction of each DNA methylation pattern i among the cell population)[27]. Epipolymorphism loci shared by at least 75% of the patients ($n = 104$) were considered for further analysis. Because epipolymorphism is dependent on DNA methylation levels, each locus was binned by the average DNA methylation levels from 0–100% (21 bins in total). Therefore, for each sample, all the loci were assigned to one of the 21 different groups on the basis of their mean DNA methylation levels. For each bin, the width was 5% (DNA methylation), except for the first and last bin (2.5% DNA methylation). The median epipolymorphism across all the loci within each bin were calculated. Then the overall epipolymorphism landscape was defined using the median epipolymorphism across the 21 bins spanning the DNA methylation levels from 0–100%. The intratumor overall MH was defined as the area under the median epipolymorphism across the spectrum of methylation percentages in 21 bins. The range of MH is 0–100. Higher MH represents higher overall intratumor heterogeneity. Each locus was covered by at least 60 sequencing reads.

Genomic annotation. Genomic annotation reference files (CpG islands and RefSeq genes) for eloci distribution analyses were obtained from UCSC (https://genome.ucsc.edu/) using the February 2009 (GRCh37/hg19) assembly[52,53]. Promoters were defined as transcription start sites ± 1 kb. CpG shores were defined as 2 kb flanking CpG islands, subtracted by any regions overlapping with nearby CpG islands. CpG shelves were defined by 2 kb flanking CpG shores, subtracted by any regions overlapping with nearby CpG islands and shores. Enhancers were defined on the basis of the US National Institutes of Health Roadmap Epigenomics Project[54] CD34-mobilized primary cell data. Active enhancers were defined as histone H3 monomethylated at K4 (H3K4me1) and histone H3 acetylated at K27 (H3K27ac) peaks without H3K4me3 marks, and poised enhancers were defined as H3K4me1 without H3K27ac or H3K4me3 marks. Wilcoxon signed-rank tests were used to compare the proportions of eloci falling into each of the genomic annotation regions.

RNA sequencing. RNA-seq libraries were prepared in two batches (**Supplementary Table 7**) using TruSeq RNA-Seq by poly(A) enrichment (Illumina) and sequenced on HiSeq2000 (Illumina) using a 50-bp paired-end approach per the manufacturer's recommendations. All paired samples were prepared within the same batch. Alignment was performed using the STAR aligner (version 2.3.0e)[55] and human genome hg19 as reference. Aligned results were annotated using the RefSeq gene model and HTSeq union mode. Gene expression data were normalized using RPKM. Blast-enriched patient samples from AML_130 (lymphocyte depleted as described above) were also used for single-cell isolations. Single cells were captured and mRNA isolated using Clontech's SMARTer chemistry (v2) on the Fluidigm C1 Single Cell Auto Prep system. Illumina's Nextera XT kit was used for library preparation before 100-bp paired-end sequencing on the HiSeq 2500 (Illumina) platform. See **Supplementary Table 8** for sequencing statistics. Data analysis was performed using r-make (http://physiology.med.cornell.edu/faculty/mason/lab/r-make/) for quality control, alignment, and gene expression quantification.

DESeq2 (ref. 56) was used for differential gene expression analysis. The significance of differentially expressed genes (DEGs) was determined using Wald significance test to compare cluster 1 compared to cluster 3. The design matrix includes batch information to control for any possible batch differences. For multiple-hypothesis testing, the significance cutoff used for optimizing the independent filtering was 0.05 (Benjamini–Hochberg). A $\log_2$ fold change greater than 1.2 or less than −1.2 was used for significance.

Transcriptional heterogeneity was assessed using single-cell RNA-seq data. Specifically, the coefficient of variations and transcriptional abundance per gene across cells was determined using $\log_2$ RPKM. Genes were included if the average $\log_2$ RPKM in cells was higher than 1.

Gene Ontology term enrichment analysis was performed using Gene Set Analysis Toolkit[57]. A threshold of two overlapping genes was considered. Hypergeometric tests were used for significance determination, and Benjamini–Hochberg[58] was used for multiple testing correction.

Whole-exome sequencing. Whole-exome capture was performed on DNA isolated from 48 paired diagnosis and relapse samples and patient-matched germline samples. Germline DNA (lymphocytes) was subjected to whole-genome amplification (repli-g kit; Qiagen) in 25 patients owing to limited materials (see **Supplementary Table 2**). To obtain sufficient quantities of DNA from the remaining samples, T cells were expanded *ex vivo* as described above. DNA was extracted using standard techniques. Libraries were prepared per the manufacturer's recommendation using NimbleGen SeqCap EZ Human Exome Library v3.0, Agilent Human Exon V3 (exon 50 Mb), or Agilent SureSelect Human All Exon V4 (51 MB; see **Supplementary Table 2** for specification of kit use per patient tumor and germline samples) and sequenced at a minimum of 50-bp single-read sequencing on a HiSeq 2000 (Illumina) to a mean coverage per base of 73-fold. See **Supplementary Table 4** for sequencing statistics.

Whole-genome sequencing. Illumina TruSeq Nano DNA Library Prep Kit (Illumina) was performed on DNA from time points T1–T5 and germline ($CD19^+CD3^+$ cells) isolated from the first relapse sample from patient AML_130 (T2). Libraries were prepared per the manufacturer's recommendation and sequenced using a 101-bp paired-end sequencing approach on the HiSeq 2000 platform per the manufacturer's recommendations to a mean coverage per base of 43-fold. See **Supplementary Table 4** for sequencing statistics.

Next-generation sequencing data analysis. DNA sequencing data were analyzed to determine somatic mutations, copy-number aberrations, and clonal evolution patterning using publicly available tools (BWA[59], GATK[60–62], Mutect[63], Varscan[64], somaticSniper[65], SnpEFF[66], XHMM[67], DNAcopy library[68], and sciClone[46]). Tool versions; inclusion, exclusion, and significance parameters; statistical tests; and specific commands used and implemented are described in **Supplementary Methods** and have been deposited into GitHub (https://github.com/ShengLi/relapsed_AML).

Integrative analysis. The association between DEGs (derived from relapse versus diagnosis paired AML patient samples) and promoter-associated eloci was determined as follows. Eloci (maximum $\Delta S = -90$) annotated to gene promoters (transcription start sites ± 1 kb) were considered for the integrative analyses. Patients with a minimum of 30 genes with promoter-associated eloci were included ($n = 19$). Epiallele loci within gene promoters that exhibited a minimum ΔS greater than −2 were designated non-eloci. RPKM $\log_2$ fold change was used to measure the transcriptome level dynamics. The variance of gene expression (RPKM $\log_2$ fold change) was calculated for genes with or without eloci within their proximal promoters and compared across all patients using a Wilcoxon signed-rank test. Genes with $\log_2$ fold change greater than 1 were defined as DEGs. The number of DEGs with or without eloci were compared by Wilcoxon signed-rank test.

Assessment for association between epigenetic clusters and mutations in genes recurrently affected in AML was performed. The variant calling pipeline (in-house-developed algorithm)[69] used BAM files generated from the sequencing data of all paired diagnostic, relapse and germline samples to identify SNVs and small indels in driver genes associated with AML[11,45]. For the identification of these somatic genetic aberrations we used a multi-variant-calling approach integrating the output of six different somatic mutation detection algorithms: MuTect[63] (version 1.17), Indelocator, GATK UnifiedGenotyper[61]

(version v3.2-2-gec30cee), SAMtools[70] (version 0.1.19-44428cd), VarScan 2 (version 2.3.6) (ref. 64) and Pindel[71] (version 0.2.5a7). Briefly, MuTect allowed a VAF up to 10% or six reads containing the variant allele in the germline control to prevent the filtering of driver mutations detectable in lymphocyte control material[72]. Indelocator was allowed to detect indels with a VAF of 2% or greater. SAMtools, GATK UnifiedGenotyper and VarScan2 were run with default parameters. Pindel was used to detect the *FLT3*-internal tandem duplication (*FLT3*-ITD) aberration specifically within exons 13, 14, and 15 of the *FLT3* gene by focusing on short insertions or tandem duplications. The detected variants were aggregated in a unique variant call format (VCF) file and subsequently annotated with ANNOVAR[73]. Annotations included dbSNP[74], COSMIC[75], and population-based sequencing efforts, such as the 1000 Genomes Project. The detected variants were further characterized by multiple fragment and regional characteristics by an in-house-developed algorithm previously described[69]. In brief the algorithm determines the number of high-quality (alignment score ≥40) and total number of fragments (irrespective of alignment score) for each detected variant. This algorithm determines the VAF from high-quality and concurrently all fragments on the basis of the detected alternative allele. In addition, the algorithm determines the strand bias of the reads. For each potential somatic variant, the same set of fragment and regional statistics was determined for the germline sample. Somatic mutations were detected by comparing the characteristics from the diagnostic or relapse sample to the germline sample. The *FLT3*-ITD aberration was detected by Pindel (available at https://github.com/ShengLi/relapsed_AML); however, following this approach the VAF was not accurately estimated, and the VAF was therefore considered not determined (**Supplementary Table 8**). In a subset of patients, *NPM1* mutations were not detectable owing to a lack of coverage in the data generated, as indicated in **Supplementary Table 7**. Mutations were considered diagnosis specific if alternative allele frequency at diagnosis was greater than 3% and was not detected at relapse. Mutations were considered relapse specific if alternative allele frequency at relapse was greater than 3% and was not detected at diagnosis. Mutations were considered shared between diagnosis and relapse if alternative allele frequency at diagnosis and at relapse were both greater than 3%. Association between the frequency of each mutation and the epigenetic clusters was assessed using a chi-squared test with Monte Carlo simulation (number of replicates = 1×10^6).

Clinical correlation analysis. The correlation of overall epiallele shift ($\log_{10}$ EPM) with clinical parameters was evaluated. The clinical correlation between EPM ($\log_{10}$) and FAB classes was performed using ANOVA test. FAB classes with a minimum of seven patients were assessed. We used the Pearson correlation (r) between overall epiallele shift and age or blast purity post separation and Hoeffding's D statistics for dependency test[76].

To determine which gene-associated epiallele shift loci (eloci) associated with clinical outcome, we performed the following analysis: (i) we determined eloci between diagnosis samples and NBM controls (we required that the epiallele region considered be covered by a minimum of 5 NBM controls); (ii) we annotated eloci to gene promoters (transcription start site ± 1 kb) and excluded epiallele loci which were covered in less than ten patients; (iii) we divided the patients into groups of longer ($n = 69$) or shorter ($n = 68$) relapse-free survival on the basis of the median value of time to relapse; (iv) we assessed the frequency of each elocus in the patient groups; (v) we used odds ratio to determine the association between early and late relapse group for each locus; (vi) we determined significance using Fisher's exact test for each locus; (vii) we applied the Benjamini–Hochberg correction to *P* values.

Survival analysis. Log-rank (Mantel–Cox) test was used for survival analysis. For relapse-free survival analysis, survival endpoints in this study were time from diagnosis until AML relapse. The patients were divided by the median EPM (low EPM versus high EPM), median number of somatic mutations (low MUT versus high MUT), or median MH (low MH versus high MH) for respective comparisons. Among 138 AML patients in the cohort, time to relapse was available for 137 patients, and WBC count was available for 127 patients.

Multivariate Cox proportional hazards regression model used relapse time as response variable, and included $\log_{10}$ EPM, age, gender, and WBC as variables to be tested ($n = 127$). *P* value of each predictor was used to assess whether this clinical parameter was significantly associated with relapse time.

49. Nicholson, J.K., Hubbard, M. & Jones, B.M. Use of CD45 fluorescence and side-scatter characteristics for gating lymphocytes when using the whole-blood lysis procedure and flow cytometry. *Cytometry* **26**, 16–21 (1996).
50. Team, R.C. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2012).
51. Matthias, D., Roehr, J.T., Ahmed, R. & Dieterich, C. Flexbar—flexible barcode and adapter processing for next-generation sequencing platforms. *Biology* **1**, 895–905 (2012).
52. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
53. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496 (2004).
54. Kundaje, A. *et al.* Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
55. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
56. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
57. Wang, J., Duncan, D., Shi, Z. & Zhang, B. WEB-based gene set analysis toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* **41**, W77–W83 (2013).
58. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995).
59. Li, H. & Durbin, R. Fast and accurate short-read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
60. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
61. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
62. Van der Auwera, G.A. *et al.* From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **11**, 10.1 (2013).
63. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
64. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy-number-alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
65. Larson, D.E. *et al.* SomaticSniper: identification of somatic point mutations in whole-genome sequencing data. *Bioinformatics* **28**, 311–317 (2012).
66. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain *w1118;iso-2;iso-3. Fly (Austin)* **6**, 80–92 (2012).
67. Fromer, M. *et al.* Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am. J. Hum. Genet.* **91**, 597–607 (2012).
68. Venkatraman, E.S. & Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).
69. Gröschel, S. *et al.* Mutational spectrum of myeloid malignancies with inv(3)/t(3;3) reveals a predominant involvement of RAS/RTK signaling pathways. *Blood* **125**, 133–139 (2015).
70. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical-parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
71. Ye, K., Schulz, M.H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium-sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).
72. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
73. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
74. Sherry, S.T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
75. Forbes, S.A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805–D811 (2015).
76. Hoeffding, W. A nonparametric test of independence. *Ann. Math. Statist.* **19**, 546–557 (1948).