Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity

Peter van Galen[1,2,3,4,#], **Volker Hovestadt**[1,2,3,4,#], **Marc Wadsworth II**[3,5,8], **Travis Hughes**[3,5,8], **Gabriel K. Griffin**[1,2,3,4,6], **Sofia Battaglia**[1,2,3,4], **Julia A. Verga**[1,2,3,4], **Jason Stephansky**[2,6,7], **Timothy J. Pastika**[2,6,7], **Jennifer Lombardi Story**[4], **Geraldine S. Pinkus**[6], **Olga Pozdnyakova**[6], **Ilene Galinsky**[7], **Richard M. Stone**[7], **Timothy A. Graubert**[4], **Alex K. Shalek**[3,5,8], **Jon C. Aster**[6], **Andrew A. Lane**[7], and **Bradley E. Bernstein**[1,2,3,4,9]

[1.]Department of Pathology, Massachusetts General Hospital, Boston, MA 02114, USA

[2.]Ludwig Center at Harvard Medical School, Boston, MA 02115, USA

[3.]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[4.]Center for Cancer Research, Massachusetts General Hospital, Boston, MA 02114, USA

[5.]Institute for Medical Engineering & Science, Department of Chemistry, and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[6.]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA

[7.]Dana-Farber Cancer Institute, Boston, MA 02215, USA

[8.]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA

[9.]Lead Contact

Summary

Acute myeloid leukemia (AML) is a heterogeneous disease that resides within a complex microenvironment, complicating efforts to understand how different cell types contribute to disease progression. We combined single-cell RNA-sequencing and genotyping to profile 38,410 cells from 40 bone marrow aspirates, including 16 AML patients and 5 healthy donors. We then applied a machine learning classifier to distinguish a spectrum of malignant cell types whose abundances varied between patients, and between subclones in the same tumor. Cell type compositions correlated with prototypic genetic lesions, including an association of FLT3-ITD with abundant progenitor-like cells. Primitive AML cells exhibited dysregulated transcriptional programs with co-expression of stemness and myeloid priming genes and had prognostic

Corresponding author: Bradley E. Bernstein (Bernstein.bradley@mgh.harvard.edu).
[#]These authors contributed equally to the work

significance. Differentiated monocyte-like AML cells expressed diverse immunomodulatory genes and suppressed T-cell activity *in vitro*. In conclusion, we provide single-cell technologies and an atlas of AML cell states, regulators and markers with implications for precision medicine and immune therapies.

eTOC

A combination of transcriptomics and mutational analyses in single cells from acute myeloid leukemia patients reveals the existence of distinct functional subsets and their associated drivers

Graphical Abstract

Introduction

Acute myeloid leukemia (AML) is an aggressive blood cancer characterized by an accumulation of immature cells of the myeloid lineage. Although most patients initially respond to chemotherapy, ~75% relapse and succumb to the disease within 5 years of diagnosis. Efforts to link AML relapse to genetically resistant clones have had limited success, raising interest in non-genetic drivers of functional heterogeneity (Kreso and Dick, 2014; Pollyea and Jordan, 2017).

One source of AML cell diversity is the partial recapitulation of myeloid development. Normal hematopoietic stem cells (HSCs) give rise to mature blood cell types of the myeloid, lymphoid, and erythroid / megakaryocyte lineages. HSC commitment proceeds through a series of increasingly lineage-committed progenitor states (Laurenti and Gottgens, 2018), as characterized in recent single-cell RNA-sequencing (scRNA-seq) studies (Karamitros et al., 2018; Velten et al., 2017; Weinreb et al., 2018). AML tumors also comprise primitive and differentiated cells. Primitive AML cells, commonly referred to as leukemia stem cells (LSCs), sustain the disease and display stem cell properties such as self-renewal, quiescence, and therapy resistance (Pollyea and Jordan, 2017). Differentiated AML cells lack self-

renewal capacity, but could impact tumor biology through pathologic effects on tumor microenvironment or hematopoietic function.

AML progression is influenced by normal cells in the microenvironment. The immune system can limit tumor cell expansion until sub-populations that evade or suppress host immunity emerge. Conversely, intrinsic properties of AML cells, including expression of immunomodulatory factors, and extrinsic microenvironmental changes can lead to an accumulation of suppressive T-regulatory cells (T-reg) and impair cytotoxic T-lymphocyte (CTL) activation (Austin et al., 2016). Enhancing T-cell-mediated AML cell clearance is an attractive therapeutic strategy, but immunotherapy trials have been less successful than in other cancers (Lichtenegger et al., 2017). This highlights a critical need to better understand the cellular components and mechanisms that underlie immunosuppression in the AML microenvironment.

scRNA-seq provides a powerful means to characterize malignant and stromal cell populations in tumors (Giustacchini et al., 2017; Puram et al., 2017; Zheng et al., 2017). In AML, scRNA-seq could potentially address questions related to stemness, developmental hierarchies, and interactions between malignant and immune cells. However, AML presents unique challenges related to its complex differentiation hierarchies and similarities between malignant and normal cells in the ecosystem (Levine et al., 2015). To comprehensively analyze AML heterogeneity, transcriptional data on thousands of cells must be complemented by genotyping data to distinguish malignant from normal cells. Standard plate-based scRNA-seq methods that capture full-length transcripts lack sufficient throughput. Recent droplet- and nanowell-based methods offer higher throughput, but the resulting sequencing data are biased to 3' transcript ends and cannot efficiently detect mutations specific to malignant cells (Giladi and Amit, 2018). These considerations emphasize the need for combined single-cell transcriptional and genetic profiling methods to characterize AML ecosystems.

Here, we adapt nanowell-based technology (Gierahn et al., 2017) to acquire transcriptional and mutational data for thousands of single cells from BM aspirates. We profiled 30,712 cells from 16 AML patients and 7,698 cells from 5 healthy donors by scRNA-seq, and acquired genotyping information for 3,799 cells. We also incorporated long-read nanopore sequencing to phase mutations, detect insertions and fusions, and distinguish subclones. We integrated these data in a machine learning classifier that distinguished malignant from normal cells, and identified six malignant AML cell types that project along the HSC to myeloid differentiation axis. We use this resource to relate developmental hierarchies to genotypes, to evaluate properties and prognostic significance of primitive AML cells, and to identify differentiated AML cells with immunomodulatory properties.

Results

Identification of cell populations in healthy BM samples

To characterize the baseline cellular diversity in healthy BM, we carried out scRNA-seq using a high-throughput nanowell-based protocol, termed Seq-Well (Gierahn et al., 2017). We profiled viably frozen cells from four healthy donors (age 21–56), and enriched

progenitors from a fifth donor (CD34$^+$CD38$^-$ and CD34$^+$) (Figure S1A–B and Table S1). Barcoded sequencing reads were assigned to cells and aligned to the transcriptome, and individual mRNA molecules were counted using unique molecular identifiers (UMIs). We acquired high quality data for 7,698 healthy donor BM cells.

We distinguished cell types by unsupervised clustering using BackSPIN (Figure 1A and S1C–D) (Zeisel et al., 2015). Cell clusters expressed established markers of hematopoietic populations, such as *CD34* for HSC/Prog cells, *CD14* for monocytes and *CD3* for T-cells (Figure 1B). This allowed us to merge 31 clusters into 15 main cell populations. We captured a broad representation of cell types, including HSCs and progenitors, as well as multiple myeloid, erythroid and lymphoid populations. All 15 cell types were identified in at least three donors (Figure 1C), while the sorted CD34$^+$CD38$^-$ and CD34$^+$ cells were highly enriched for HSCs and progenitors.

We next explored the relationships between these cell types by visualizing K-nearest-neighbor (KNN) graphs that connected all single cells in our dataset to their five nearest neighbors in gene expression space (Weinreb et al., 2018). This revealed putative differentiation trajectories, including a continuum of cells from HSCs to monocytes with several intermediate states and gene expression gradients (Figure 1D and S1E–G). Our cell type annotations are consistent with recent scRNA-seq studies and published gene signatures (Figure S1H) (Hay et al., 2018; Karamitros et al., 2018; Laurenti et al., 2013; Novershtern et al., 2011; Velten et al., 2017). Thus, scRNA-seq of normal BM reveals diverse hematopoietic cell types and implies differentiation trajectories consistent with current views of hematopoiesis.

Single-cell profiling of AML tumor ecosystems

To examine cellular diversity in AMLs, we obtained 35 cryopreserved BM aspirates from 16 AML patients at diagnosis and during treatment (Figure 2A). This cohort includes multiple WHO subtypes and is genetically diverse (age 26–74, Table S1). Targeted DNA sequencing of all samples in our cohort showed expected mutation frequencies, including *DNMT3A* (44% of patients), *FLT3* (38%), and *NPM1* (31%, Figure 2B) (Cancer Genome Atlas Research, 2013). We performed scRNA-seq for these 35 samples without enrichment to achieve a broad overview of the AML ecosystem.

We acquired 30,712 high-quality transcriptomes and visualized cells for each patient using t-Distributed Stochastic Neighbor Embedding (t-SNE). This revealed distinct cell types whose proportions changed markedly over the clinical course (Figure 2C–D). In addition to malignant cells, these data revealed presumed normal hematopoietic cell types in the tumor ecosystem expressing lineage-specific genes such as hemoglobin (erythroid cells) and *CD3* (T-cells). Samples collected after induction chemotherapy had a predominance of T / NK cells, consistent with clearance of AML blasts and histological stains showing frequent lymphocytes. Although other cell populations also expressed markers associated with specific hematopoietic cell types, their identity as normal or malignant could not be distinguished *a priori* from their expression programs. We therefore explored additional methods for distinguishing malignant AML cells.

Single-cell genotyping by short-read and nanopore sequencing

Prior scRNA-seq studies of tumors have leveraged gene mutations detected in full-length transcriptomic data and chromosomal copy number variations (CNVs) to distinguish malignant cells (Filbin et al., 2018; Giustacchini et al., 2017). However, high-throughput methods yield 3'-biased transcript coverage, which constrains mutation detection. Moreover, AMLs frequently lack CNVs. We therefore adapted Seq-Well to amplify and sequence portions of transcripts that contain AML mutations (Figure 3A and S2A). We took advantage of an intermediate whole transcriptome amplification (WTA) step that yields full-length cDNAs with cell barcodes (CBs) appended to their 3' ends. We designed 43 primers adjacent to all mutations detected in our cohort by targeted DNA sequencing, and generated amplicons containing mutational sites appended to CBs. Sequencing of these products enabled us to overlay mutational status onto our scRNA-seq data.

We applied mutation-specific single-cell genotyping to each of the 35 AML samples. We successfully detected wild-type and/or mutant transcripts at 27 of the 43 targeted sites (Table S2). We detected transcripts in 14 out of 16 patients, with an average of 355 transcripts mapping to 258 cells per patient. Mutations near 3' transcript ends of highly expressed genes were more efficiently detected (Figure 3B). For example, NPM1 is highly expressed and its W288fs hotspot mutation is 342 bp from the nearest polyA signal, allowing identification of a wild-type or mutant transcript in up to 31% of cells. DNMT3A-R882 mutations are only 161 bp from the nearest polyA signal, but expression is low, such that a wild-type or mutant transcript was identified in up to 6% of cells. Application of the method across our patient cohort identified 3,745 wild-type and 1,230 mutant transcripts (Table S2). Mutations were not detected in healthy donor BM samples and were markedly decreased in AML patients in clinical remission (Figure S2B–D). Furthermore, our detected mutation frequencies strongly correlated with variant allele frequencies (VAFs) obtained with targeted DNA sequencing (r = 0.87, Figure 3C).

To further expand the applicability of our single-cell genotyping protocol, we incorporated nanopore sequencing (van Dijk et al., 2018). We reasoned that the long reads provided by this platform could enhance detection of mutations across transcripts and reveal long insertions, deletions and fusion breakpoints. We amplified representative oncogenes, tumor suppressors and fusions along with CBs from three AML patients, and sequenced the amplicons using Oxford Nanopore Technologies MinION (Figure S2E). We acquired 0.97 million reads mapping to the targeted genes, which we consolidated on the basis of CBs to 255 cells. The nanopore data complemented the Illumina data in several ways. First, in the TP53-Q144P and P152R mutant tumor AML328, nanopore sequencing detected mutant or wild-type transcripts in 97 cells, representing a 3-fold improvement (Figure 3D–E). Transcripts detected only by nanopore were significantly longer than those detected by Illumina. Transcripts captured by both methods (n = 30) yielded identical genotyping results in all cases. Phasing of *TP53* alleles showed that the mutations each affect different transcripts, consistent with bi-allelic inactivation of this tumor suppressor. Second, in the *FLT3* mutant tumor AML328, long reads revealed a 60 bp FLT3 internal tandem duplication (ITD) that was missed by short-read sequencing (Figure 3F). Finally, in the *RUNX1* fusion

tumor AML707B, long reads enabled detection of *RUNX1-RUNX1T1* fusion transcript in 32 cells and revealed the exact sequence of the junction (Figure 3G).

In conclusion, we present methods for amplifying barcoded transcripts of genes that are frequently mutated in AML. Sequencing by Illumina and Oxford Nanopore Technologies enabled detection and phasing of point mutations, insertions, deletions and fusions, thereby genotyping individual cells from AML aspirates (Figure 3H–J).

Machine learning classifier distinguishes malignant from normal cells

We next integrated single-cell mutation calls and transcriptomes for all patients, with the goal to distinguish malignant from normal cells. Since informative genetic calls were acquired for only a subset of cells, we proceeded as follows. First, we selected all AML cells for which single-cell genotyping detected mutations in the assessed genes. We then used the random forest machine learning algorithm to classify these putatively malignant cells according to their similarity to all 15 normal BM cell types (Figure 4A and S3A–D). The vast majority of cells with mutations resembled one of six normal cell types along the HSC to myeloid axis (HSC, progenitor, GMP, promonocyte, monocyte or cDC; Figure 4B–C). We therefore annotated cells with detected mutations that were classified along this axis as HSC-like, progenitor-like, GMP-like, promonocyte-like, monocyte-like or cDC-like malignant cells. These malignant cell types were then incorporated as additional classes in a second classifier that was used to annotate all AML cells in our dataset as malignant or normal (Figure 4A, D–E).

We validated our malignant/normal classifications and cell type annotations by several methods. First, 5-fold cross-validation showed that the second classifier distinguishes malignant cells with >95% sensitivity and >99% specificity (Figure S3E). Second, the transcriptomes of non-malignant cells from AML aspirates closely resembled counterparts from normal BM aspirates (Figure S3F–G). Third, AML707B harbored a Y-chromosome deletion and, consistently, Y-chromosome transcripts were not detected in malignant cells from this tumor (Figure S3H). Fourth, AML328 harbored a chromosome 7 deletion, which we detected as loss-of-heterozygosity of a highly expressed SNP in the 3' UTR of *ACTB* specifically in malignant cells from this tumor (Figure S3I).

Overall, we detected 13,489 malignant AML cells (44% of cells, Figure S4A–C). The fraction of single cells classified as malignant for any given tumor was consistent with clinical blast counts (r = 0.93, Figure 4F). Together, these data support the accuracy of our approach for distinguishing malignant from normal cell types in AML tumors.

Intra-tumoral heterogeneity of malignant AML cells

Intra-tumoral heterogeneity has been extensively studied using cell surface markers (Kreso and Dick, 2014). However, this approach relies on predefined markers that may not accurately represent underlying transcriptional programs and may be expressed by both malignant and normal cells (Levine et al., 2015). We therefore explored the potential of our unbiased transcriptomic classification to provide additional insights. The six malignant cell types identified by our classifier were each represented by at least 1,000 cells in our dataset and identified in at least ten patients (Figure S4). However, their relative abundances varied

markedly between tumors, with some consisting primarily of one or two cell types, and others comprising a spectrum of malignant cell types (Figure 4G–H and S5A). The cell type abundances estimated by our classifier corresponded closely to clinical parameters determined by cell morphology and surface phenotypes (Figure 4I). For example, AML707B had a high proportion of cells classified as GMP-like, consistent with flow cytometry showing low levels of myeloid differentiation markers. In contrast, AML419A had a higher proportion classified as differentiated myeloid cells (60%), consistent with the clinical diagnosis of AML with monocytic differentiation. Despite a strong overall correlation with clinical flow-based estimates of myeloid differentiation (r = 0.87, Figure S5B), the scRNA-seq data revealed more extensive malignant cell diversity than could be appreciated from a limited number of markers. For example, AML921A and AML329 had representation for all six malignant cell types including cDC-like cells (Figure 4G–I). Thus, scRNA-seq data are consistent with clinical parameters, but provide more detailed information on AML cell types and differentiation states.

AML419A harbors subclones with distinct cell type compositions

We next considered the underlying causes of variability in malignant cell type abundances. While most of our tumors were predominated by a few proximate cell types or a spectrum along the myeloid axis, AML419A contained two malignant cell types at opposite ends of the developmental axis (Figure 4G–H). We hypothesized that the corresponding populations of HSC/Prog-like cells and differentiated monocyte-like cells reflected different genetic subclones. Genotyping of AML419A revealed three activating *FLT3* mutations: FLT3-ITD, FLT3-A680V and FLT3-N841K (Table S1). Analysis of nanopore sequencing reads allowed each mutation to be assigned to a different allele, while a fourth allele was wild-type (Figure 5A). Given that this AML was cytogenetically normal and that targeted DNA sequencing failed to detect CNVs, the four *FLT3* alleles implied the existence of multiple subclones. Consistently, although we detected FLT3-ITD and FLT3-A680V transcripts in the same cells, the FLT3-N841K mutation never co-occurred with other mutant alleles in the same cell (Figure 5B). Integration of these data with VAFs from bulk DNA sequencing enabled us to infer a putative phylogeny of AML419A: that it evolved one subclone 'A' with a FLT3-A680V mutation, a second subclone 'B' with an additional FLT3-ITD mutation on the opposite allele, and an independent third subclone 'C' with a FLT3-N841K mutation only (Figure 5C).

As these mutations confer FLT3 gain-of-function by different mechanisms (Figure 5D) (Leick and Levis, 2017), we investigated whether the corresponding subclones have different phenotypes. We could confidently assign 13 cells to subclone C, based on detection of FTL3-N841K transcripts, and another 10 cells to subclone B, based on detection of FLT3-ITD transcripts (Figure 5E). We could not definitively assign the 17 cells for which we only detected a FLT3-A680V transcript, and therefore annotated them as either clone A or B. These assignments enabled us to compare the cell type compositions of the different subclones by evaluating the expression of genes that are specific to each of the six malignant cell types (Table S3). A majority of cells in subclones A/B expressed signature genes associated with progenitor-like cells (19/27 cells, Figure 5E). In contrast, nearly all subclone C cells expressed genes associated with differentiated monocyte-like or cDC-like cells

(12/13 cells). Thus, combined genetic and transcriptional analysis suggest that AML419A contains a FLT3-A680V/ITD subclone consisting mostly of primitive cells, and a separate FLT3-N841K subclone consisting mostly of differentiated cells. The two subclones converge on cells with similar transcriptional states, but harbor them in different proportions. These results suggest that alternate *FLT3* genotypes can profoundly influence cellular hierarchies of AML subclones within a single tumor.

AML cellular hierarchies correlate with underlying genetic alterations

These *FLT3* associations prompted us to examine the relationship between genotype and cellular hierarchy across a larger AML cohort. To this effect, we used the scRNA-seq data to derive gene signatures for each of the six malignant cell types (Table S3). These signatures were designed to equally weight each malignant cell type and to exclude genes that are expressed in normal cell types that can be prevalent in AML tumors, thus distinguishing our approach from prior studies that have stratified AMLs by variable genes or signatures of sorted populations (Cancer Genome Atlas Research, 2013; Ng et al., 2016). We used our signatures to score bulk expression profiles of 179 diagnostic AML aspirates from the Cancer Genome Atlas (TCGA), and thereby infer their cell type compositions.

Hierarchical clustering of the TCGA AMLs by these signatures revealed seven clusters of tumors with distinct malignant cell type compositions (Figure 5F and S5C). Several clusters included tumors with high abundances of specific cell types, such as GMP-like (clusters A-B), progenitor-like (cluster D) or monocyte-like cells (cluster E). Others comprised tumors that contain a spectrum of malignant cell types along the HSC to myeloid axis (cluster G). These inferences indicate marked variability in cell type compositions and developmental hierarchies.

We next examined the relationship between these inferred hierarchies and underlying genotypes. Remarkably, the clusters derived solely from cell type abundances corresponded closely to the genetics of the AMLs (Figure 5G). For example, TCGA tumors with uniquely high GMP-like scores (cluster B) perfectly overlapped with *RUNX1-RUNX1T1* fusions. In line with this observation, the one AML in our scRNA-seq dataset harboring this genetic alteration (AML707B) consisted almost entirely of GMP-like cells (Figure 4G–H). A second cluster of tumors with a spectrum of cell types and relatively high monocyte-like and cDC-like scores (cluster F) overlapped almost perfectly with *CBFB-MYH11* fusions ($P < 0.001$). Consistently, the one AML in our scRNA-seq dataset harboring this genetic alteration (AML1012) showed similar cell type abundances (Figure S5A). A third cluster with high GMP-like scores (cluster A) perfectly overlapped with acute promyelocytic leukemias (APL) with *PML-RARA* fusions. Two other clusters were enriched for cytogenetically complex tumors and those harboring *CEBPA*, *RUNX1*, and *TP53* mutations (clusters C, G). These clusters have distinct malignant cell type compositions, with cluster C representing the most undifferentiated group of AMLs (enriched for FAB M0 subtype) and cluster G recapitulating a spectrum of differentiation.

Taken together, our analyses reveal striking variability in the abundances of malignant cell types across AMLs, and suggest a prominent role for genotype in determining the cell type composition and hierarchy of a given tumor.

Differential effects of FLT3 genotypes on AML differentiation

The remaining two TCGA clusters (D, E) both contained *NPM1* mutant tumors, but differed markedly in their cell type compositions (Figure 5F–G). Cluster D was enriched for undifferentiated HSC/Prog-like cell signatures and harbored multiple FLT3-ITD mutant tumors. Cluster E was enriched for monocyte-like and cDC-like cell signatures and harbored FLT3-TKD mutant tumors. These associations are reminiscent of the alternate *FLT3* subclones in AML419A (Figure 5E), and suggest that FLT3-ITD confers a strong differentiation block.

To test this hypothesis, we expressed FLT3-WT (wild-type), FLT3-TKD or FLT3-ITD in the MUTZ-3 cell line and examined the resulting cellular phenotypes by flow cytometry. We found that FLT3 expression increased the percent of primitive CD34$^+$ MUTZ-3 cells (Figure 5H–I and S5D–F). This effect was most pronounced with the FLT3-ITD construct, consistent with the potent signaling activity of the corresponding protein variant and with the primitive composition of FLT3-ITD AMLs. Although FLT3 mutations have been primarily described as enhancers of proliferation, our findings point to additional effects on cell differentiation that may help explain why FLT3-ITD AMLs have worse outcomes than FLT3-TKD mutant tumors (Leick and Levis, 2017).

Dysregulated transcriptional programs in primitive AML cells

We next turned our focus to primitive AML cell types, which fuel tumor growth. We found that primitive AML cells upregulate genes involved in stress response and redox signaling (*XBP1*, *GPX1*), proliferation (*FLT3*, *PIM1*, *MYC*) and self-renewal (*HOXA9*, *BMI1*), relative to their normal counterparts (Figure 6A, S6A–E, Table S4). We also evaluated preferentially-expressed surface markers as these provide opportunities for targeted therapies (Pollyea and Jordan, 2017). This highlighted established LSC markers such as *CD96*, *CD47* and *IL1RAP* (Figure 6B, S6C), as well as additional candidates such as *CD36* and *CD74* (MHCII invariant chain).

To further relate the differentiation states of primitive malignant and normal cells, we generated three gene signatures that represent successive stages of normal hematopoietic development: HSC/Prog (including *MEIS1*, *NRIP1*, *MSI2*), GMP (including *MPO*, *ELANE*, *AZU1*) and differentiated myeloid (including *LYZ*, *MNDA*, *CD14*) (Figure S6F and Table S3). As expected, application of these signatures to single cells from normal BMs clearly distinguished major cellular subsets of HSC/Prog, GMP and differentiated myeloid cells (Figure 6C–D). However, a distinct pattern emerged when we applied these signatures to malignant AML cells. HSC/Prog signature genes and GMP signature genes were frequently co-expressed in the same malignant cells, contrasting markedly with their exclusivity in normal hematopoiesis. Malignant HSC/Prog-like cells also expressed myeloid factors, such as *MPO* and *ELANE*, that are absent in normal HSC/Prog cells, consistent with prior reports of lineage priming in LSCs (Goardon et al., 2011; Krivtsov et al., 2006).

Finally, we considered the clinical implications of the primitive AML populations. The relative proportions of HSC/Prog-like and GMP-like cells varied markedly among the tumors in our cohort. We therefore used our signatures for these malignant cell types to

score the 179 TCGA AMLs. The signatures were anti-correlated across the bulk expression profiles (r = −0.24, Figure 6E), supporting the observation that the developmental states of primitive AML cells vary between tumors. We partitioned the AMLs into a group with relatively higher expression of HSC/Prog-like genes (n = 98), and a group with higher expression of GMP-like genes (n = 81). We found that patients with higher HSC/Prog-like signals, whose tumors presumably contain more primitive LSCs, had significantly worse outcomes ($P < 0.0001$, log-rank test; Figure 6F). This survival difference was more pronounced than those of the individual signatures (Figure S6G) and maintained when we excluded APL cases ($P = 0.0013$). Although prior studies have correlated stem cell signatures to AML outcome (Ng et al., 2016), our single-cell data nominate specific HSC/Prog-like cell states and transcriptional programs that may underlie these associations and bear further study.

T-cell signatures are suppressed in AML patients

T-cells can in principle eliminate AML cells, as demonstrated by the ability of graft-versus-leukemia to yield durable cures following stem cell transplantation, but may be compromised in AML (Austin et al., 2016). We therefore examined the T-cells in our single-cell data. In normal BM, we identified two T-cell subsets, naïve T-cells (*IL7R*, *CCR7*) and CTLs (*CD8A*, *GZMK*), and a related population of NK cells (*NCAM1/CD56*, *KLRD1*) (Figure 1). We recovered the same three populations when we performed unsupervised clustering of all T- and NK cells from tumor and normal samples (Figure 7A). Supervised analysis also identified a subset of cells expressing T-reg markers, but their limited numbers precluded further analysis.

AML aspirates tended to have proportionally fewer T-cells and CTLs than normal controls (Figure 7B–C). To corroborate this finding, we used immunohistochemistry (IHC) to quantify CD3$^+$ T-cells, CD8$^+$ CTLs, and CD25$^+$FOXP3$^+$ T-regs in an additional cohort of 15 diagnostic AMLs and 15 normal BMs. We again found that AMLs contained significantly fewer T-cells and CTLs and had a reduced CTL:T-cell ratio (Figure 7D–F). Conversely, the tumors had relatively greater numbers of T-regs, consistent with prior reports that this suppressive subset is increased in AML (Ustun et al., 2011). Thus, scRNA-seq and IHC reveal consistent changes in T-cell numbers and composition, indicative of an immunosuppressive tumor environment.

Differentiated AML cells suppress T-cell activation *in vitro*

The altered immune microenvironment in AML could potentially reflect activities of specific malignant cell types, such as differentiated myeloid cells (Figure 7G). We therefore tested whether AML cells suppress T-cell activation *in vitro*. We used a bioassay based on a CD4$^+$ T-cell line with a Nuclear Factor of Activated T-cells (NFAT) reporter. We stimulated these T-cells with CD3/CD28 beads and measured NFAT activation in the presence of the AML cell lines MUTZ-3 and OCI-AML3. The MUTZ-3 cells had a strong inhibitory effect, reducing T-cell activation up to 5-fold in a dose-dependent manner (Figure 7H and S7A). We validated this result by showing that MUTZ-3 cells also inhibited activation of primary T-cells from healthy donors, as assessed by flow cytometry for the T-cell activation marker CD69 (Figure 7I).

We next investigated whether the immunomodulatory properties of MUTZ-3 are mediated by specific sub-populations. We performed co-culture assays with sorted HSC/Prog-like (CD34$^+$) or monocyte-like (CD14$^+$) MUTZ-3 cells. The CD14$^+$ cells reduced T-cell activation by 10-fold ($P < 0.0001$, Figure 7J), while the CD34$^+$ cells had little effect. This prompted us to examine the immunomodulatory functions of monocyte-like cells from primary AMLs. We isolated CD14$^+$ and CD14$^-$ cells from five AML tumors and six normal BMs. The leukemic origin of the CD14$^+$ AML cells was verified by targeted DNA sequencing of the sorted populations (Figure S7B). We found that CD14$^+$ cells from three out of five AMLs strongly inhibited T-cell activation (up to 5.3-fold), whereas CD14$^-$ cells had little or no effect (Figure 7K). Notably, CD14$^+$ cells from normal BM had only a subtle effect in this assay (1.4-fold). These results suggest that a subset of AMLs give rise to CD14$^+$ monocyte-like that can suppress T-cell activation.

Evidence for immunomodulatory functions led us to examine the expression states of monocyte-like AML cells across our tumors. Recent studies have classified normal monocytes into four subtypes: classical, non-classical, and two intermediate states with either trafficking or cytotoxic features (Villani et al., 2017). Analysis of the corresponding signatures revealed that monocyte-like AML cells exhibit features of classical and non-classical monocytes, but lack cytotoxic signature genes (Figure 7G and S7C). Despite their malignant origin, monocyte-like cells were similar to normal monocytes with respect to these subtype signatures.

Nonetheless, an unbiased comparison of malignant and normal cells revealed 296 genes that were preferentially expressed in malignant monocyte-like cells from one or more tumors (Figure S7D and Table S4). This set contained many genes with annotated functions in immune regulation, whose expression varied markedly between patients (Figure 7L). For example, monocyte-like cells from specific tumors over-expressed TNF pathway genes (*TRAIL*/*TNFSF10*, *TNFAIP2*), IL-10 pathway genes (*STAT1, HMOX1*), or regulators of reactive oxygen species (*TXNIP*), all of which have been associated with myeloid-derived suppressor cells (Hartwig et al., 2017; Veglia et al., 2018). Furthermore, several tumors strongly expressed two surface markers associated with immunosuppressive myeloid cells: *CD206*/*MRC1* and *CD163* (Biswas and Mantovani, 2010). We found that expression of these markers was associated with poor outcome in the TCGA cohort (Figure 7M). Our data suggest that differentiated monocyte-like AML cells have immunomodulatory functions that contribute to the pathogenesis of this disease.

Discussion

Intratumoral heterogeneity in AML has been appreciated since the 1960s, but it has only recently become possible to study the complexity of tumors using high-dimensional single-cell analyses (Giustacchini et al., 2017; Levine et al., 2015; Zheng et al., 2017). Here we combined scRNA-seq and genotyping to characterize AML tumor ecosystems, distinguish malignant from normal cells, and elucidate subclones. We identified six malignant cell types along the HSC to myeloid differentiation axis whose abundances vary widely between AMLs with different genotypes. We further investigated cells at opposite ends of this axis, documenting dysregulated transcriptional programs in HSC-like AML cells and

immunomodulatory properties of monocyte-like AML cells. Our study provides powerful single-cell technologies, a rich resource of single-cell transcriptomes, and insights into AML hierarchies, LSC programs and tumor-immune interactions.

To address unique challenges posed by the AML ecosystem, we established methods that combine high-throughput scRNA-seq with single-cell genotyping of recurrently mutated AML genes. We used short-read and nanopore sequencing to detect and phase point mutations, insertions, deletions and fusions in individual cells. We then integrated transcriptional and genetic data for AMLs and normal BMs in a machine learning classifier, which identified six malignant AML cell types that shared features with normal hematopoietic cells. Although we primarily deployed these technologies to distinguish malignant cells, our identification of genetic subclones in a *FLT3* mutant AML suggest their potential, with further innovations, to characterize pre-malignant clones and LSC evolution.

The relative abundances of malignant cell types varied markedly between the 16 AMLs that we profiled at the single cell level, as well as across 179 bulk AML profiles queried with cell type-specific gene signatures. Single-cell data were instrumental for the latter analysis as they enabled generation of precise malignant cell signatures that were robust to confounding signals from non-malignant cells in tumors. Unsupervised clustering of TCGA bulk expression profiles by these gene signatures yielded seven groups of AMLs with distinct cell type compositions, indicative of shared differentiation programs or cellular hierarchies. Remarkably, each AML group was strongly enriched for characteristic genetic lesions. Hence, the genotypes that define molecular subtypes used for patient risk stratification are also associated with characteristic cellular hierarchies.

Our results provide particular insight into the functional significance of *FLT3* genotypes. TCGA tumors with FLT3-TKD mutations are enriched for differentiated AML cells, while those with FLT3-ITD insertions have higher abundances of primitive cells. A similar conclusion emerged from an in-depth analysis of genetic subclones in a single AML tumor. Nanopore sequencing enabled us to phase four distinct FLT3 alleles in this tumor and assign them to distinct subclones. We found that a FLT3-ITD subclone primarily contained primitive cells, while a FLT3-TKD subclone in the same tumor primarily contained differentiated cells. We also demonstrated that FLT3-ITD expression suppresses differentiation of AML cells *in vitro*. These inter-tumoral, intra-tumoral and *in vitro* findings suggest that FLT3 variants differentially affect AML differentiation, and may explain the association of FLT3-ITD with poor patient outcomes (Leick and Levis, 2017).

While our data emphasize prominent roles for genetics in shaping AML hierarchies, they do not exclude additional effects of genotype on the intrinsic transcriptional states of specific cell types. Indeed, we find such examples in supervised analyses of malignant progenitors in *NPM1* mutant and *RUNX1-RUNX1T1* fusion tumors (Figure S6D–E). These genotype-specific alterations are superimposed upon already deranged transcriptional states of malignant progenitors (Figure 6C–D), which conflate stemness and myeloid programs. Our definition of the transcriptional states of malignant AML cell types, their inter-tumoral variability, and their close association to tumor genetics is an important milestone with

implications for treatment response, relapse and the development of genotype-specific precision therapies.

Lastly, while AML studies often focus on primitive cell types with self-renewal capacity, we find evidence that differentiated malignant cells also contribute to AML biology. Monocyte-like AML cells, which are present at variable abundances in a majority of our tumors, potently inhibit T-cell activation *in vitro*. These cells could contribute to altered T-cell phenotypes and an immunosuppressive AML microenvironment (Austin et al., 2016). Monocyte-like cells express a range of immunomodulatory genes, including TNF and IL-10 pathway genes, antigen presentation components and leukocyte immunoglobulin-like receptors (Deng et al., 2018; Hartwig et al., 2017). However, the expression of these genes varies markedly between tumors, which may confound efforts to characterize and modulate their activities. Nevertheless, our data can guide future efforts to define the functions and mechanisms of immunomodulatory AML populations, to evaluate their relationship to myeloid-derived suppressor cells, and to modify their activities for therapeutic benefit (Lichtenegger et al., 2017; Pyzer et al., 2017; Veglia et al., 2018).

In summary, we leveraged innovations in single-cell transcriptomics and genotyping to parse heterogeneous AML ecosystems. Our results provide insight into the aberrant regulatory programs of primitive AML cells, reveal a striking correspondence between developmental hierarchies and tumor genetics, and identify differentiated AML cells with immunosuppressive properties. Our data and findings can guide therapeutic strategies to target critical and specific components of the malignancy.

STAR * METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Bradley E. Bernstein (bernstein.bradley@mgh.harvard.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human tumor specimens—All acute myeloid leukemia (AML) patients consented to an excess sample banking and sequencing protocol that covered all study procedures and was approved by the Institutional Review Board (IRB) of the Dana-Farber Cancer Institute. Normal bone marrow (BM) and CD4$^+$ T-cell donors consented to the same protocol or an IRB-approved protocol from Lonza that covers all study procedures. Demographic and clinical details are provided in Table S1.

Cell lines—MUTZ-3 cells were purchased from DSMZ (ACC-295), 5637 cells were purchased from ATCC (HTB-9) and OCI-AML3 cells were received from Dr. Mark Minden (University of Toronto). Cell line verification by Short Tandem Repeat profiling was performed upon receipt and every six months (ATCC 135-XV). OCI-AML3 and 5637 cells were cultured at 37°C in RPMI-1640 with Glutamax (Thermo 61870–036) with 10% heat-inactivated FBS (Peak Serum PS-FB1) and P/S (RPMI+). MUTZ-3 cells were cultured at 37°C in MEM-alpha (Thermo 12571–063) with 20% heat-inactivated FBS (Peak Serum PS-FB1), P/S, and 10% 5637-conditioned medium. We used the heterogeneous MUTZ-3 cells

to confirm that cell type annotations correlated with phenotypic and functional characteristics (Figure S5D–F). Specifically, culture initiation and proliferation were restricted to CD34[+] HCS/Prog-like MUTZ-3 cells, whereas CD14[+] monocyte-like cells lacked these properties.

METHOD DETAILS

Cell preparation—Aspirates from the iliac crest of normal BM donors and AML patients were processed using density gradient centrifugation to isolate mononuclear cells, viably frozen with 10% DMSO and stored in liquid nitrogen (only BM5 was not frozen). Frozen cells were thawed using standard procedures, and viable cells were enriched using magnetic removal of dead cells (Miltenyi Biotec 130–090-101) or flow cytometry to sort propidium iodide-negative cells.

To sort primitive cells from a fresh BM aspirate (BM5), 200 million cells were enriched for CD34 using Miltenyi Biotec magnetic enrichment microbeads (130–046-702) and sorted by flow cytometry. Briefly, cells were resuspended in 600 μl PBS with 2% FBS (Peak Serum PS-FB1) and incubated with 200 μl FcR blocking reagent and 200 μl CD34 antibody-conjugated magnetic beads (epitope QBEND/10). Cells were applied to an MS column (Miltenyi Biotec 130–042-201) on a magnet followed by collection of the CD34[+] fraction. Next, CD34[+] cells were resuspended in 1 ml PBS with 2% FBS, 10 μl PE-conjugated CD38 antibody (BD 347687) and 10 μl APC-conjugated CD34 antibody (BD 340441, clone 8G12 which is not inhibited by QBEND/10 antibody binding). Cells were stained for 15 minutes on ice, washed with PBS 2% FBS, resuspended in PBS 2% FBS with DAPI, and sorted on a Sony SH800 sorter (Figure S1B). Post-sorting analysis showed 95–96% purity. Then, 10,000 CD34[+] cells and 10,000 CD34[+]CD38[−] cells were used for Seq-Well scRNA-seq.

Single-cell transcriptome profiling—Seq-Well was performed as described (Gierahn et al., 2017), with the following changes: we performed 18 PCR cycles for WTA, and we used a template switching oligo with an LNA-modification of the last guanine (Table S2). Briefly, an array with ~90,000 nanowells is loaded with mRNA capture beads that are bound to oligos with a primer binding sequence, cell barcode (CB), unique molecular identifier (UMI) and polyT sequence (Chemgenes NC0927472). The size of the beads relative to the wells of the array ensures that only one bead will occupy each well. Then, a cell suspension of 200 μl contained 10,000 cells is loaded and cells enter nanowells by gravity in approximately 20 minutes. The cell : nanowell ratio ensures that nanowells only contain a single cell. A partially permeable polycarbonate membrane (Sterlitech Custom Order) is used to seal the surface of the array, which allows buffers to pass through but traps the bead and the cell. Cells are lysed with a lysis buffer and mRNA binds to the polyT sequence on the oligo that is linked to the bead, which is contained in the same well. Following a bead removal process and pooling of all the beads, the bead-bound mRNA is reverse transcribed to produce cDNA which is then used for WTA PCR. Sequencing libraries are prepared using Nextera reagents (Illumina FC-131–1096). Libraries from 2–3 nanowell arrays were sequenced per run, yielding 350–550 million reads on an Illumina NextSeq 500 instrument. Sequencing was performed according to manufacturer's instructions, with the following adjustments: (1) Libraries were loaded at 2.5 pM, (2) a Custom Read 1 Primer (CR1P, Table S2) was used by

diluting 6.6 μl of CR1P (100 μM) to 2.2 ml with HT1 buffer (provided by Illumina), (3) we did not use PhiX because it would be incompatible with CR1P. Read length was 20 cycles for Read 1, 8 cycles for the library index, and 50 or 64 cycles for Read 2 (64 cycles were used for single-cell genotyping; all single-cell Seq-Well reads were shortened to 50bp for comparability).

Reproducibility of the Seq-Well protocol was supported by the following measures: (1) BM samples from healthy donors were processed months apart (BM1 processed on April 11, 2017, BM2 processed on April 24, 2017, BM4 processed on June 10, 2017, BM3 processed on July 24, 2017 and BM5 processed on November 15, 2017). However, we did not observe batch dependent clustering (Figure S1E) and the variability in cell type frequencies between individuals was within the expected range (Figure 1C). (2) We verified that non-malignant cells from different AML patients cluster together (Figure S3G). (3) AML921A was processed in two technical replicates, resulting in highly similar data. (4) Different samples from the same patient were always thawed and processed simultaneously.

Targeted DNA sequencing—Targeted sequencing of genetic mutations was performed using the Rapid Heme Panel (RHP) platform, which is a service by the Center for Advanced Molecular Diagnostics of Brigham and Women's Hospital (Kluk et al., 2016). Briefly, a set of 95 genes that are recurrently mutated in hematological malignancies are amplified and sequenced at 1500× coverage on average. Single nucleotide variants and small insertions/ deletions are detected at allele frequencies of 5%. This platform was used for every AML patient at diagnosis, some patients at later time points (Table S1), and sorted CD14[+] cells (Figure S7B). The frequency of mutations identified in our cohort (e.g. *DNMT3A* mutations in 44% of patients, *FLT3* alterations in 38% of patients, and *NPM1* alterations in 31% of patients; Figure 2B) was comparable with larger AML genome sequencing cohorts (Cancer Genome Atlas Research, 2013).

Single-cell genotyping by short-read sequencing—We designed an adaptation of the Seq-Well method for targeted amplification of known mutations from the WTA product (Figure S2A). The starting material for this single-cell genotyping method is the product of the Seq-Well WTA reaction (only a fraction of which is used for scRNA-seq). The method consists of two PCR reactions with a streptavidin bead enrichment in between. The first PCR reaction serves to add a biotin tag and Nextera adapter (NEXT) to the mutation of interest while retaining the UMI and CB of the transcripts. We designed biotinylated primers to detect specific mutations (Table S2), that were known because every patient underwent targeted DNA sequencing (see above). For every AML sample, a 10× primer mix is created containing the SMART-AC primer at 3 μM, which is common to all initial reactions, and 1– 6 mutation-specific primers (such as Next_NPM1_833) at a combined concentration of 3 μM.

To prepare the template for the single-cell genotyping reaction, WTA products from an AML sample are pooled and diluted to be used at 10 ng in a total volume of 10 μL (every AML sample is split into several WTA reactions during the Seq-Well protocol). Next, 2.5 μL of the 10× primer mix and 12.5 μL of KAPA HiFi Hotstart ReadyMix (Fisher Scientific KK2602) are added to the template and PCR is performed using the following conditions:

initial denaturation at 95°C for 3 minutes, followed by 12 cycles of 98°C for 20 seconds, 65°C for 15 seconds, and 72°C for 3 minutes, and final extension at 72°C for 5 minutes. Following amplification, the PCR product is purified with 0.7× AMPure XP beads to remove primers (Beckman Coulter A63881). Since the SMART-AC primer is nearly complementary to both ends of the WTA product, this first PCR yields many unintended full-length fragments. Using Streptavidin-coupled Dynabeads, only biotinylated fragments containing the mutation of interest are captured (following manufacturer's instructions, ThermoFisher 60101). Dynabeads/DNA-complex is eluted in 23 μL H$_2$O.

To add Illumina adapters (P5, P7), an index barcode to identify the library (Index_BC), and a CR1P binding sequence to the fragments, a second PCR is performed using 23 μL of streptavidin-bound product as template, with 2 μL of 5 μM primer mix (P5_SMART_Hybrid and N70_BCXX, Table S2) and 25 μL PFU Ultra II HS 2× Master Mix (ThermoFisher Q32854). The parameters used for PCR2 are an initial denaturation at 95°C for 2 minutes, then 4 cycles of 95°C for 20 seconds, 65°C for 20 seconds, and 72°C for 2 minutes, followed by 10 cycles of 95°C for 20 seconds and 72°C for 2 minutes and 20 seconds, and then a final extension at 72°C for 5 minutes.

After the second PCR, the streptavidin beads are magnetized and the supernatant is saved and then purified with 0.7× AMPure XP beads. After eluting in 20 μL of TE, the AMPure XP beads are magnetized and the supernatant is saved for sequencing. The resulting libraries are similar to Seq-Well scRNA-seq libraries but with targeted integration of the NEXT sequencing primer binding site adjacent to the mutation of interest. The libraries were generally 0.5–30 ng/μl and 200–800 bp in size. Single-cell genotyping libraries were sequenced together with Seq-Well scRNA-seq libraries on an Illumina NextSeq500 instrument. Genotyping and scRNA-seq libraries from the same AML sample were not sequenced in the same run to prevent cross-contamination of libraries. The computational analysis of single-cell genotyping data is described in detail below.

Single-cell genotyping by nanopore sequencing—In addition to short-read Illumina sequencing, we performed long-read nanopore sequencing for three AML samples. We reasoned that nanopore sequencing would improve detection of long amplicons and would enable detection of large tandem duplications and gene fusions for which the exact fusion junction was unknown. WTA products were amplified using primers designed for *TP53*, *FLT3*, and the *RUNX1-RUNX1T1* (Table S2). Amplicons containing the mutation site, UMI and CB were amplified as described above for the single-cell genotyping procedure. PCR products were further amplified with P5 and P7 primers for an additional 15 cycles to obtain enough material for nanopore library preparation. Amplicons were purified using a 0.7× AMPure XP bead cleanup, and library construction was performed using the SQKLSK108 (1D) and SQK-LSK308 (1D^2) Ligation Sequencing Kit (Oxford Nanopore Technologies, ONT) according to manufacturer's instructions, with some modifications. Briefly, 1 μg purified DNA was subjected to end repair and dA-tailing using the NEBNext Ultra II End-Repair/dA-tailing Module. Next, a 1X volume AMPure XP bead cleanup was performed and nanopore sequencing adapters were ligated to the eluted DNA using the Blunt/TA Master Mix (NEB). In order to capture shorter amplicons, the final clean-up of the adapter-ligated DNA was modified and performed with 0.7× AMPure XP beads. The purified-ligated DNA

was quantified by fluorometry (Qubit) and 300–500 ng was DNA mixed with RBF (Running Buffer with Fuel mix, ONT) and LLB (Library Loading Beads, ONT) before loading on R9.4/R9.5 flow cells (FLOMIN106/FLO-MIN107, ONT). We produced data from four different flow cells that were run for 10–48 h on a MinION sequencing device as per manufacturer's guidelines and controlled using the MinKNOW 2.2 software. Reads were base-called using the albacore software (version 2.3.3). Individual (unpaired) reads were used for samples processed using the $1D^2$ sequencing kit. The computational analysis of long-read single-cell genotyping data is described in detail below.

Lentiviral expression of FLT3 mutants—FLT3-WT and FLT3-ITD containing plasmids were provided by Dr. Andrew Lane. The coding sequences were amplified using the primers (for) 5'-CACCATGCCGGCGTTGGCG-3' and (rev) 5'-CTACGAATCTTCGACCTGAGC-3' and cloned into the pENTR/D-TOPO vector using manufacturer's instructions (Thermo Fisher K240020). pENTR-FLT3-D835Y was generated from pENTR-FLT3-WT using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent 210518) with the primers 5'-GTTGGAATCACTCATGATATATCGAGCCAATCCAAAGTCAC-3' and 5'-GTGACTTTGGATTGGCTCGATATATCATGAGTGATTCCAAC-3'. Next, pENTR-FLT3-WT, pENTR-FLT3-D835Y and pENTR-FLT3-ITD were recombined with pMAL to generate the expression vectors FLT3-WT, FLT3-D835Y and FLT3-ITD marked by eGFP from a bidirectional minimal hCMV-hPGK promoter. As a control (CTRL), we used pMAL-LUC (expressing humanized *Renilla* luciferase instead of FLT3). Vector sequences were verified by Sanger sequencing. CTRL, FLT3-WT, FLT3-D835Y and FLT3-ITD lentiviral particles were packaged in 293T cells by co-transfecting VSV.G (Addgene Plasmid 14888), dRT-pMDLg/pRRE (Addgene Plasmid 60488) and pRSV-Rev (Addgene Plasmid 12253) using Fugene (Promega E2311). Lentiviral particles were concentrated 50X using LentiX Concentrator (Takara 631232), resuspended in RPMI+ and stored at −80°C.

MUTZ-3 cells were transduced with 1–3 μl virus to achieve 20–40% GFP+ cells. Cells were cultured in the presence of FLT3 ligand blocking antibody (R&D Systems MAB308–100) to minimize the impact of endogenous (wild-type) FLT3 signaling. After four days, MUTZ-3 cell differentiation was read out using flow cytometry for GFP (to gate on transduced cells), CD34 (APC, BD 340441) and CD14 (PE-Cy7, Coulter A22331).

Flow cytometry—Flow cytometry for surface marker analysis was performed using similar procedures as cell sorting (see "Cell preparation") and analyzed on a BD Cytoflex or BD LSRII. Fluorochromeconjugated antibodies are listed in the Key Resources Table. Cell cycle analysis was performed as follows: (1) stain MUTZ-3 cells with CD14-PE-Cy7 (Coulter A22331) and CD34-FITC (BD Pharmingen 348053) antibodies; (2) spin and permeabilize using Phosphoflow Perm 2 solution (BD 347692 diluted 10× in H_2O); (3) wash with PBS 2% FBS; (4) stain with Ki67 Alexa Fluor 647 (BD Pharmingen 561126); (4) wash; (5) resuspend in Cytofix buffer (BD 554655 diluted 4× in PBS) with DAPI at 1 μg / ml; (5) analysis on an BD LSRII analyzer within 30 minutes. Data was analyzed using FlowJo software (Tree Star, Inc.).

Immunohistochemistry—Slides were run on the Bond III Imunostainer (deparaffinized on the machine) by Leica (Buffalo Grove, IL). The machine uses several retrievals including ER1 which is low pH (citrate based) and ER2 which is high pH (EDTA based) and heats the slides during the retrieval. The following antibodies were used: CD3 clone LN-10 (Leica), 1:300, 1 hour incubation with primary, retrieval ER2 20 minutes, detection with Bond Polymer Refine (DAB); CD8 clone C8/144R (Dako), 1:200, 1 hour incubation with primary, retrieval ER2 30 minutes, detection with Bond Polymer Refine; FOXP3/CD25 double stain: FOXP3 clone 206D (Biolegend), 1:50, retrieval ER2 40 minutes, detection with Bond Polymer Refine (DAB), CD25 clone 4C9 (Lifespan), 1:50, 1 hour incubation, detection with Bond Polymer Refine Red Detection.

T-cell activation bioassay—The T-cell activation bioassay was purchased from Promega (J1621) and carried out according to manufacturer's instructions. Briefly, 25 μl RPMI+ containing 100,000 Human T-Activator CD3/CD28 beads (Thermo Fisher 11131D) was combined with 25 μl RPMI+ containing 100,000 BM or AML cells and 25 μl RPMI+ containing 100,000 TCR/CD3 Effector Cells (total 75 μl / well). The TCR/CD3 Effector Cells are Jurkat cells with endogenous TCR, CD3, CD4 and CD28 expression and a luciferase reporter driven by a Nuclear Factor of Activated T-cells Response Element (NFAT-RE). Engagement of TCR/CD3 with an appropriate ligand, such as CD3/CD28 beads, results in NFAT-RE mediated luminescence. The beads and cells were incubated at 37°C for 6 hours followed by reading out luciferase using Bio-Glo (Promega G7941) on a BioTek SYNERGY HT machine. Positive control wells contained Human T-Activator CD3/CD28 beads and TCR/CD3 Effector cells (no BM / AML cells, 100% luminescence). Background control wells contained 75 μl RPMI+, and never exceeded 1% of positive controls. Negative controls wells contained TCR/CD3 Effector cells ± BM / AML cells (no beads), and never exceeded background levels. Luminescence was calculated by subtracting background and shown as a percentage of positive control wells.

Activation of primary CD4+ T-cells was tested by adding together 25 μl RPMI+ containing 100,000 Human T-Activator CD3/CD28 beads and 25 μl RPMI+ containing 100,000 MUTZ-3 cells. Primary CD4+ T-cells (Lonza 2W-200) were thawed and 100,000 cells were added per well in 25 μl RPMI+ (total 75 μl / well). The beads and cells were incubated at 37°C for 6 hours followed by flow cytometry for CD4-FITC (Biolegend 357405), CD33-PE-Cy7 (BD 333946) and CD69-APC (Biolegend 310909). After gating on (P1) lymphocytes and (P2) CD4+CD33− cells, the mean fluorescence of CD69 was used as a measure of activation.

To specifically assess the effect of HSC/Prog-like (CD34+) and monocyte-like (CD14+) AML cells in the T-cell bioactivation assay, cells were sorted using Miltenyi Biotec magnetic enrichment microbeads (130–046-702 or 130–050-201) according to manufacturer's instructions. Briefly, cells were incubated with antibody-conjugated magnetic beads and applied to an MS column on a magnet followed by collection of negative (flow-through) and positive fractions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Cell barcode processing—All sequencing data was first assessed by looking at general quality metrics such as cluster density, total yield, and per-cycle base quality. Sequencing libraries were then split by library barcodes using bcl2fastq version 2.15.0.4 and default settings, except for allowing for 2 mismatches to library barcode sequences when appropriate. Read 1, containing a 12 bp CB and an 8 bp UMI, yielded 20 bp reads. Read 2, containing part of the transcript, yielded 50 bp reads. For some of the sequencing runs Read 2 was sequenced for up to 64 cycles. The extra bases were used only for single-cell genotyping analysis. All downstream analyses were performed using the R programming language (version 3.4), unless otherwise noted. We made extensive use of the data.table and Rsamtools packages.

To analyze our single-cell sequencing data, we employed an approach to annotate sequencing reads by CB before sequence alignment and quantification. First, we counted all unique 12 bp CBs for each library. We excluded CBs occurring less than 100 times, and filtered barcodes containing stretches of eight identical nucleotides. Next, we excluded CBs that were associated with non-random UMIs. For all reads associated with a given CB, we checked that the frequency of any nucleotide did not exceed 90% at each base of the UMI. The majority of reads filtered this way contained part of the Tn5 binding sequence, i.e. reflected events in which the transposase integrated within the CB/UMI, yielding very short fragments with invariable (non-random) UMI sequences.

We noticed that a number of CBs (5–20%, depending on the batch of barcoded beads) were associated with UMIs that contained a Thymine as the last nucleotide. These sequences often represent CBs in which a single nucleotide is missing due to errors in the split-pool synthesis. In this case reads start with a 11 bp CB, followed by the 8 bp UMI and the first base of the poly-T sequence that hybridizes to the poly-A tail of captured mRNAs. If not corrected, this causes a single cell to produce four different single-cell transcriptomes. We corrected these barcodes if in fact four different CBs were detected with a similar number of total reads that were variable in their last base. The UMIs were also corrected accordingly.

To filter out CBs that likely resulted from sequencing errors, we ranked all CBs according to their number of reads (requiring at least 1,000 reads). We filtered out all CBs that had a higher ranked CB that was different in only one position (hamming distance 1).

This final list of CBs was then used to generate a fastq file containing the Read 2 sequences of the remaining cells. The library barcode, the CB, and the UMI were appended to the read identifier. For some of the sequencing runs we noticed a higher number of reads that were excluded because the library barcode was not detected accurately. We rescued these reads if their CB matched uniquely to one of the libraries that were sequenced together in the respective run. Resulting fastq files for each sample were deposited in GEO (GSE116256).

Sequence alignment and gene quantification—Sequencing reads were aligned to the human genome (hg38) using STAR version 2.5.2b and default parameters. Alignments were guided by using RefSeq gene annotations. Transcripts were quantified using the "--quantMode TranscriptomeSAM" option. This resulted in two alignment files, one in which

reads were aligned to the genome, and one which contained pseudo-alignments to the transcriptome.

The transcriptome alignments were used to quantify gene expression. For every read all the unique gene names of the transcripts the read aligned to were recorded. Some reads aligned to multiple genes, which often reflected a primary gene and one or more pseudo-, antisense-, or readthrough-genes. We checked if any of the gene names was contained in all the other gene names, with a "-" before or after (antisense- and readthrough-genes), or followed by "P" and a digit (pseudo-genes). If this was the case, we only kept the primary gene. Reads that still mapped to multiple genes were filtered. In a second step, all reads that mapped to the same gene and had an identical UMI sequence were collapsed. This yielded a digital expression matrix consisting of UMI counts for each cell and gene.

For all downstream analysis we required cells to have at least 1,000 UMIs (gene counts, indicative of the number of captured transcripts) mapping to at least 500 unique genes. We additionally excluded cells for which more than 20% of the gene counts reflected either mitochondrial genes or ribosomal RNAs, as these likely reflected poor quality cells. Resulting digital expression matrices for each sample were deposited in GEO. For downstream analyses, we normalized gene counts to a total of 10,000 for each cell.

BackSPIN clustering—Initial QC yielded 7,698 cells from normal BM donors. We randomly filtered 783 cells of 1,590 BM5 CD34$^+$CD38$^-$ cells to reduce representation of this population. The remaining 6,915 normal BM cells were then clustered into cell types using BackSPIN (Zeisel et al., 2015). BackSPIN employs a bi-clustering algorithm which iteratively splits both cells and genes, until a predetermined number of splits is reached. We selected the BackSPIN algorithm because it performs well when dealing with a relatively large number of cell populations. This is especially true for datasets in which some clusters are demarcated by a large number of differentially expressed genes (e.g. between the myeloid and lymphoid lineages), and others by relatively few genes (e.g. different populations within the myeloid lineage).

For clustering, we first determined the most variably expressed genes in the dataset. We performed a linear fit of the log-transformed average expression values and the log-transformed coefficients of variation (standard deviation divided by the average expression). Variably expressed genes were determined as genes associated with a residual larger than two times the standard deviation of all residuals. From these genes we excluded a set of genes that were associated with cell cycle (*ASPM, CENPE, CENPF, DLGAP5, MKI67, NUSAP1, PCLAF, STMN1, TOP2A, TUBB*). This yielded in the order of 1,000 to 2,000 variably expressed genes depending on the set of cells (Figure S1C, S4A). Expression values were log-transformed (after addition of 1) before performing BackSPIN clustering. We used default settings and a maximum splitting depth of 5. In the healthy BM data this yielded a final set of 31 clusters.

In a first post-processing step we calculated the average expression level of each gene for each cluster. If gene expression of a single cell correlated higher to the average gene expression of another cluster than the cluster it was part of, we reassigned the cell to the

cluster it was most highly correlated to. For the healthy BM data, we merged clusters if their average gene expression profiles were highly correlated and if they were characterized by similar cell type-specific marker genes. This yielded 15 cell types across the undifferentiated compartment and the three main lineages (erythroid, lymphoid, and myeloid, Figure 1A).

We independently clustered normal BM cells using SC3, a different clustering algorithm that is also designed for single cell analysis. We used a two-step strategy that first separates the main lineages (Undifferentiated, Myeloid, Erythroid, and Lymphoid), and then clustered again within each lineage. The results were concordant with our BackSPIN clustering results (data not shown). We conclude that the BackSPIN algorithm is an appropriate choice for clustering cell types in our scRNA-seq data.

KNN and t-SNE visualization—We employed two different methods for visualizing similarities between cells in two-dimensional space: visualization of k-nearest-neighbor (KNN) graphs and t-distributed stochastic neighbor embedding (t-SNE). For both methods we started with the same set of variable genes as for the BackSPIN clustering. For KNN visualization we calculated pairwise correlation coefficients between single cells. Then we constructed a graph by connected each cell to its five most highly correlated neighbors. This graph was visualized using SPRING, an interactive tool that uses force-directed graph drawing. For t-SNE visualization we used the Rtsne implementation in R and default parameters, except setting the maximum number of iterations to 2,000 (5,000 for the healthy BM data). Throughout the study, we show only two different KNN visualizations (healthy BM and T / NK cells, Figure 1D and 7A, respectively) and two different t-SNE visualizations (AML556 and AML707B, Figure 2C–D). These visualizations are reused in other figures to highlight additional cell parameters, such as sample-of-origin, mutation status, and gene expression levels.

Short-read single-cell genotyping analysis—Short reads from libraries of single-cell genotyping by Illumina sequencing were processed using the CBs detected from the regular Seq-Well protocol. This ensured detection of fragments even if there were only few reads for a given CB. Resulting fastq files for all samples were deposited in GEO. All genotyping reads were aligned to a short reference index consisting only of the expected transcripts using BWA-MEM and default mapping parameters.

For each mutational site and AML sample, we then determined the expected read sequence for both the wildtype and the mutant allele. These were identical to the most frequently detected read sequences for most of the sites. For some primers we observed unspecific amplification of other transcripts. This however did not affect our interpretation of the targeted site. We only retained reads that contained the exact mutant or wildtype sequence at the expected position. In case of short insertions and deletions (e.g. NPM1 insertion), we required the exact mutant sequence to be detected. We allowed for one mismatch to the reference transcript in the remaining read sequence.

We then counted the number of reads supporting the mutant or wild-type transcript for each CB and UMI. Since most transcripts were detected hundreds of times, we required at least 10 sequencing reads per CB and UMI. For each mutant transcript we frequently also

detected a wild-type transcript with the same CB and UMI, albeit at a much lower frequency (0.1–1%). The same was observed for wild-type transcripts, for which we also detected the mutant transcript at much lower frequency. This is consistent with a low background sequencing error rate. These erroneous transcripts were filtered out. Similarly, we filtered transcripts that likely resulted from sequencing errors in the UMI if for the same CB there was a similar UMI that was different in only one base and detected with more reads. For each cell and mutational site, we then summarized the detected mutant and wild-type transcripts and used these annotations throughout the study. A detailed overview of this data is presented in Table S2. Genotyping results for each single cell from every sample are provided within annotation files that were deposited in GEO.

Long-read single-cell genotyping analysis—Long reads from libraries of single-cell genotyping by nanopore sequencing were aligned to the human genome using minimap2 using the following non-standard parameters: -x splice -u b -k 14. We proceeded with 0.97 million reads aligned to the targeted genes (*TP53* and *FLT3* in AML328, *FLT3* in AML419A, and *RUNX1/RUNX1T1* in AML707B). Fastq files containing these reads were deposited in GEO.

For each gene and sample, we then matched every CB detected from the regular Seq-Well data to the last 250 bp of the 3' end (as determined by mapping orientation) of all aligned reads. We allowed for one mismatch (including 1 bp insertions and deletions). We included the last 3 bases of the invariable adapter sequence (TAC) before the CB (12 bp) to improve specificity. We found that most CBs mapped to defined positions within the nanopore reads, consistent with their expected position 3' of the poly-A tail and 5' of the SMART sequence, Illumina P5 adapter and Nanopore adapter (illustrated in Figure S2A). To ensure accurate barcode matching, we performed three additional filtering steps: First, we narrowed the window of expected starting positions of the CBs to 70–130 for reads aligning to the gene in forward orientation, and 40–110 for reads aligning in reverse orientation. For data generated using the $1D^2$ kit we used windows from 80–180 and 40–160. We proceeded with CBs that mapped to the respective windows in   90% of reads in which it was detected. Second, we checked for the location of the poly-A tail relative to the mapped CB. For all reads associated with a given CB, we required an average number of at least 5 A's starting within 7 to 11 bp from the last base of the CB (the 8 bp UMI lies in between the CB and poly-A) and removed CBs which did not fulfill these criteria. Lastly, we filtered a small number of reads associated with multiple CBs. For all remaining CBs, we sampled a maximum number of 1,000 reads for downstream genotyping.

To perform single nucleotide variant calling for *TP53* and *FLT3*, we separately aligned all reads for each CB to a transcript reference that consisted only of the targeted gene using minimap2. We separately aligned reads to references corresponding to every mutant allele we detected. We then used nanopolish (version 0.10.2) to first create indices for all fast5 files associated with each CB, and then performed variant calling from the raw nanopore signal at the exact positions of the variants. The following non-default parameters were used: --ploidy=2 --min-flanking-sequence=10 --calculate-all-support --min-candidate-frequency=0.1. As nanopolish only reports variants but not wild-type calls, we repeated the analysis for each mutant allele reference in which wild-type transcripts appear as mutant,

and subsequently merged all variant calls. For the final genotyping, we required a coverage of at least 10 reads, and a supporting base fraction of at least 0.5, as reported by nanopolish.

To call *FLT3* internal tandem duplications (ITDs) in AML328 and AML419A, we aligned all reads for each CB to a transcript reference that consisted of the FLT3 transcript containing the ITD using minimap2. Using this reference, alignments of reads from wild-type transcripts appear to have a deletion, whereas transcripts containing the ITD do not. For each read, we calculated the average coverage in two 50 bp windows on either side of the duplicated sequence, and only considered reads with at least 80% coverage in both windows. We then calculated the average coverage of the duplicated sequence, which showed a bimodal pattern corresponding to the ITD (80% coverage) and wildtype allele. For the final genotyping we required a coverage of at least 10 reads. Cells associated with 80% of reads corresponding to the ITD were called as mutant, and cells associated with 20% of reads corresponding to the ITD were called as wild-type.

To call *RUNX1-RUNX1T1* fusion transcripts in AML707B, we aligned all reads for each CB to a transcript reference that consisted only of the *RUNX1* gene, and independently to a reference containing only the *RUNX1T1* gene. For each read, we then calculated the covered bases in either alignment. For each cell, we calculated the fraction of reads aligning for more than 100 bp to either reference and required at least 10 aligned reads. For the final genotyping we only considered cells that were associated with 30% of reads aligning to *RUNX1*. Cells associated with 10% of reads aligning to *RUNX1T1* were called as mutant, and cells associated with 1% of reads aligning to *RUNX1T1* were called as wild-type.

For visualization of the genotyping analysis by nanopore sequencing (Figure 3D, 3F–G, 5A) we selected representative transcripts for each mutant and wild-type allele, and then selected 100 supporting reads for display. Genomic alignments were visualized using the Integrated Genomics Viewer. We further processed IGV images to only show variants at the mutated sites, which were broadened to the width of the entire exon for clarity. Nanopore genotyping results for each single cell were deposited in GEO. Additionally we deposited raw nanopore signal (fast5) files for all reads associated with mutant or wild-type transcripts.

In total, our study acquired transcriptomes for 38,410 cells and genotyping information for 3,799 of these cells using both short-read Illumina sequencing and long-read nanopore sequencing. It is the first to combine single-cell transcriptomics and genotyping in a high-throughput format (droplet or nanowell). For comparison, Smart-seq2 studies typically acquire full-length transcriptomes for a few thousand cells, and has been adapted for targeted genotyping of ~1000 cells (Giustacchini et al., 2017). However, its plate-based format cannot efficiently scale to the numbers of cells that we process here. Other published methods assess genomic DNA and transcriptomes from the same cell, allowing detecting of mutations in lowly-expressed genes and non-transcribed regions. However, current iterations are limited in throughput to a few hundred cells (Macaulay et al., 2017). Recent online reports combine scRNA-seq and mutation detection to analyze hematologic malignancies (Nam et al., bioRxiv: 444687, Petti et al., bioRxiv: 434746, Rodriguez-Meira et al., bioRxiv: 474734, Velten et al., bioRxiv: 500108). Each approach has its specific advantages.

Generation of the Random forest classifier—The Random forest algorithm is a machine learning approach that uses a large number of binary decision trees that are learned from random subsets of a training set. These trees (the forest) can then be applied to a given sample to generate a class probability that reflects its similarity to a given class of the training set. If a single class prediction is required, the class with the highest probability score is used (majority vote). Random forest classifiers are particularly well suited if the dataset contains many different classes, many samples and many features. In our case samples represent single-cell expression profiles, features represent genes, and classes represent different cell types. For our analysis we used the randomForest R package version 4.6–14.

We used Random forest-based classification for two different purposes: To predict similarity of single cells to the 15 different cell types detected in healthy BM (classifier 1), and to predict if a single cell from a tumor sample is malignant or normal (classifier 2). To train the first classifier, we first performed a feature selection step to select the most informative genes from all 14,554 expressed genes in the dataset (average expression > 0.01). Feature selection was performed by training an "outer" random forest classifier on all expressed genes. We trained 1,000 trees, using a random subset of 50 cells from each cell type for each tree. Based on the reported overall gene importance in the "outer" classifier, we then selected only the 1,000 most informative genes for training of the "inner" classifier. The reported out-of-bag (OOB) error (i.e. misclassification error of cells that were not used for learning of a given tree) was 20% lower for the "inner" than for the "outer" classifier, justifying the use of an initial feature selection step. The "inner" classifier was further evaluated by performing 5-fold cross-validation by splitting the training dataset into five equally sized parts. In each iteration of the cross-validation, four of these parts were used to generate a classifier that was then used for predicting class probabilities of the remaining part. Results of the cross-validation are provided in Figure S3A.

The second classifier is used for determining if a cell for which we did not detect a mutant transcript is malignant or normal, based on its similarity to normal and malignant cells (i.e. cells from healthy BM and HSC to myeloid-like cells from tumor samples for which we detected mutant transcripts). We first attempted to use a classifier that distinguishes between just these two classes. However, we achieved much better results by using all 15 normal and six malignant cell types in a combined training set (21 classes), presumably because a malignant monocyte-like cell is more similar to a normal monocyte than to a malignant HSC-like cell. For malignant cells we used cell type annotations as predicted by the first classifier, with the following exceptions: to have at least 65 HSC-like cells for each malignant class (required for having >50 cells for 5-fold cross-validation), we reclassified 23 cells initially classified as progenitor-like with highest prediction scores for the HSC cell type as HSC-like cells. We also reclassified 29 cells that were initially classified as early Erythroid progenitors as progenitor-like cells, if their prediction score for the Progenitor cell type was higher than their prediction score for the late Erythroid cell type. The second classifier was then generated using the combined training set of 21 classes and the same parameters as for the first classifier. The second classifier reached 95.2% sensitivity and

99.7% specificity in distinguishing malignant from normal cells, as measured by 5-fold cross-validation. Results of the 5-fold cross-validation are provided in Figure S3E.

To exclude the possibility that the high frequency of cells with detected NPM1 mutations affected the classifier, we generated a separate classifier that does not consider NPM1 mutant calls. This separate classifier had equally high specificity (99.8% of normal cells correctly called normal), and sensitivity (93% of malignant cells correctly called malignant) in 5-fold cross-validation. It is also consistent with the original classifier: 97% of cells originally classified as normal were classified as normal; 91% of cells originally classified as malignant were classified as malignant. These results indicate that the classifier is robust to the frequency of NPM1 mutations in the training set.

To independently assess whether the Random Forest classifier was an appropriate choice for classifying cell types, we compared the performance of our first random forest (RF) classifier to an independent Support vector machine (SVM) classifier. We used the e1071 R package and default parameters, except for assigning class weights inverse to the class size to account for differences in cell numbers per cell population. While the SVM classifier generated reasonable results, it did not perform as well as our random forest classifier in cross-validation. For example, a larger number of cells are misclassified to a different lineage (9.2% vs 3.8%). We conclude that the random forest algorithm is an appropriate choice for classifying cell types in our scRNA-seq data.

Random-forest based classification—When applying both classifiers to single cells from tumor samples, we first determined from the second classifier if the prediction score was highest for a malignant or normal cell type. If a cell was classified as malignant, we then used the highest prediction score of the HSC to myeloid cell types (HSC, progenitor, GMP, promonocyte, monocyte, cDC) from the first classifier for cell type assignment. If a cell was classified as normal, we used the highest prediction score from the first classifier.

We evaluated normal and malignant cell predictions by performing unsupervised BackSPIN clustering of all cells that were predicted as one of six HSC to myeloid cell types. This analysis was performed for each patient separately. We included 250 normal cells of each cell type from healthy BM samples in this clustering. For some samples we identified cells for which we could make a better judgement by considering the additional evidence at hand (e.g. mutated transcripts, targeted DNA sequencing results). We then refined these cells as malignant or normal. In total 578 cells were refined as malignant (1.9% of cells), and 573 cells were refined as normal (1.9%). An example of this evaluation is shown in Figure S4A. We also identified eight samples from four different patients for which we were not confident about the classification results (AML314, AML371, AML722B and AML997, 3.7% of cells). These samples were of poorer quality and had fewer detected cells, and were excluded from downstream analyses of malignant cells. Final classification results for each single cell from every sample are provided within annotation files that were deposited in GEO.

We used prediction scores for projecting single cells onto the KNN visualization of normal BM cells (shown in Figure 4B, 4H, S5A, S5D). For this purpose, we placed a grid of 50×50

equally sized bins onto the two-dimensional visualization of normal BM cells. We then identified the 20 most highly correlated normal BM cells for every cell to be projected, and recorded the density of these 20 most correlated BM cells for each bin of the grid.

Generation of gene signatures—We generated cell type-specific gene signatures by correlating log-transformed gene expression values to cell type prediction scores from the first Random forest classifier, and then considering the most highly correlated genes. This analysis was performed for each cell type along the HSC to myeloid differentiation axis (6 out of 15 detected cell types) across all malignant cells from AML patient samples at diagnosis (11,641 cells). For each gene, the second-highest correlation coefficient was subtracted from the highest correlation coefficient, to ensure that a gene is specific to a certain cell type (and not also highly correlated to another cell type). This correction included correlation coefficients of normal cells from healthy BM (4,430 cells) to the remaining nine cell types. This prevents genes that are more highly correlated to the erythroid and lymphoid cell types to be part of the HSC to myeloid signatures. We also included correlation coefficients of gene expression values to cell cycle signature scores (described below), which prevents genes that are highly expressed in cycling cells to be associated with a certain cell type. After this correction, the 30 most highly correlated genes to each cell type defined six tumor-derived gene signatures (shown in Figure 5E–F, S5C). In addition, we generated gene signatures for combined HSC and progenitor prediction scores, and for combined promonocyte, monocyte and cDC prediction scores, as a number of genes were correlated to each of these classes (shown in Figure 6E, S6G). We also generated three normal-derived combined signatures by correlating expression values to cell type prediction scores across normal cells from healthy BM (shown in Figure 6C–D, S6F). All gene signatures are provided in Table S3.

Scoring cells for gene signatures—We calculated cell cycle-gene expression scores in single cell profiles by using a minimal gene signature of ten genes that are highly expressed in cycling cells (*ASPM, CENPE, CENPF, DLGAP5, MKI67, NUSAP1, PCLAF, STMN1, TOP2A, TUBB*). For each of these genes, we selected the 100 genes with the smallest difference in average expression level as a background gene set. The average expression of the background gene set was then subtracted from the respective signature gene, and the average of the resulting values of all signature genes was kept as the cell cycle-gene expression score. A similar strategy for scoring gene signatures from single-cell expression data has been described previously (Puram et al., 2017). For most illustrations, the signature scores were binarized (e.g. Figure 4G, 6C). Cells were classified as cycling if the signature score was larger than a threshold value. The threshold represented the median score plus 1.5× the median absolute deviation, as calculated from the normal BM data.

To compare our normal BM single cell data with prior publications, we scored published gene expression signatures in the same way as the cell cycle signature described above. In Figure S1H, we show scores of our normal BM cells for microarray-derived expression signatures of HSCs and NK cells (Laurenti et al., 2013; Novershtern et al., 2011), and scRNA-seq derived signatures of CD34+ MultiLin progenitors (Hay et al., 2018), GMPs (Karamitros et al., 2018) and megakaryocyte erythrocyte progenitors (Velten et al., 2017). In

Figure S7C, we show scores of our monocyte and monocyte-like cells for scRNA-seq derived signatures of Mono1 (classical), Mono2 (non-classical), Mono3 (intermediate/trafficking) and Mono4 (intermediate/cytotoxic) (Villani et al., 2017). To visualize these signature scores in Figure 7G, monocytes and monocyte-like cells were placed (if Mono1 > Mono2) at x = Mono1 or (if Mono1 < Mono2) at x = -Mono2, and (if Mono 3 > Mono4) at y = Mono3 or (if Mono3 < Mono4) at y = -Mono4.

Bulk expression analysis—Bulk RNA-seq expression levels from the TCGA-LAML study were downloaded from the companion website of the original publication (Cancer Genome Atlas Research, 2013) (https://tcga-data.nci.nih.gov/docs/publications/laml_2012). We downloaded processed RPKM expression levels of 179 samples (laml.rnaseq.179_v1.0_gaf2.0_rpkm_matrix.txt.tcgaID.txt.gz). Information on cytogenetic alterations, genetic mutations and FAB classification was gathered from the updated supplementary table (SuppTable01.update.2013.05.13.xlsx). The most recent survival data was downloaded from the cBioPortal.

For unsupervised clustering of TCGA samples according to six tumor-derived gene signatures (Figure 5F) we first log-transformed the RPKM expression levels and calculated Z-scores. We then calculated pairwise Euclidean distances between samples and performed hierarchical clustering using Ward's method. Seven clusters were identified based on the resulting dendrogram. Association of clusters with genetic alterations and histological variants was tested using Fisher's exact test between all seven clusters (Figure 5G). The same information is shown in Figure S5C, but samples are clustered according to the analysis performed in TCGA (Cancer Genome Atlas Research, 2013).

In addition to the unsupervised clustering analysis, we also calculated expression scores for cell type-specific signatures in bulk profiles using a similar approach as described above for the scoring of gene signatures in single cells: For each gene in a given signature, we selected the 100 genes with the smallest difference in average expression level as a background gene set. The average expression of the background gene set was then subtracted from the respective signature gene, and the average of the resulting values of all signature genes was kept as the signature score. A similar approach for scoring bulk expression samples has been described previously (Puram et al., 2017). These scores were used to stratify patients into two groups, followed by Kaplan-Meier survival analysis (shown in Figure 6E–F and S6G).

Statistical testing—We used unpaired Student's *t* tests for *in vitro* assays of myeloid differentiation (Figure 5I), T-cell type frequencies (Figure 7E–F) and T-cell activation (Figure 7H–K, Figure S7A). We used a paired Wilcoxon test to compare correlations of gene expression to prediction scores (Figure 6D). We used Fisher's exact test for analysis of clinical parameters of TCGA samples (Figure 5G, S5C). We used the Kruskal-Wallis rank sum test for analyzing gene expression differences (Figure 6B, Figure S6D–E, Figure S7D). We used a log-rank test for analyzing survival analysis (Figure 6F, Figure 7M, Figure S6G). We used ELDA software to assess culture-initiating cell frequencies (Figure S5E). Statistical analyses were performed using Microsoft Excel (Student's *t* test), the ELDA website (http://bioinf.wehi.edu.au/software/elda) and using the R language for Statistical Computing (all others). Parameters such as sample size, number of replicates, number of independent

experiments, measures of center, dispersion, and precision (mean ± SD) and statistical significance are reported in Figures and Figure Legends.

DATA AND SOFTWARE AVAILABILITY

The raw data, gene expression matrices, genotyping information and cell annotations have been deposited in the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE116256. R scripts written for performing gene quantification, unsupervised clustering, training and application of the random forest-based classifier and other analyses are shared via GitHub (https://github.com/BernsteinLab/aml2019).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Declaration of Interests

B.E.B. is an advisor and equity holder for Fulcrum Therapeutics, 1CellBio, HiFiBio and Arsenal Biosciences, is an advisor for Cell Signaling Technologies, and has equity in Nohla Therapeutics. A.A.L. receives research support from Stemline Therapeutics and is a consultant for N-of-One. J.C.A. is an advisor for Cellestia, Ayala and Epizyme. A patent application covering single-cell genotyping has been filed by the Broad Institute.

References

Austin R, Smyth MJ, and Lane SW (2016). Harnessing the immune system in acute myeloid leukaemia. Crit Rev Oncol Hematol 103, 62–77. [PubMed: 27247119]

Biswas SK, and Mantovani A (2010). Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. Nat Immunol 11, 889–896. [PubMed: 20856220]

Brunetti L, Gundry MC, Sorcini D, Guzman AG, Huang YH, Ramabadran R, Gionfriddo I, Mezzasoma F, Milano F, Nabet B, et al. (2018). Mutant NPM1 Maintains the Leukemic State through hOx Expression. Cancer Cell 34, 499–512 e499. [PubMed: 30205049]

Cancer Genome Atlas Research (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368, 2059–2074. [PubMed: 23634996]

Deng M, Gui X, Kim J, Xie L, Chen W, Li Z, He L, Chen Y, Chen H, Luo W, et al. (2018). LILRB4 signalling in leukaemia cells mediates T cell suppression and tumour infiltration. Nature 562, 605–609. [PubMed: 30333625]

Filbin MG, Tirosh I, Hovestadt V, Shaw ML, Escalante LE, Mathewson ND, Neftel C, Frank N, Pelton K, Hebert CM, et al. (2018). Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. Science 360, 331–335. [PubMed: 29674595]

Gierahn TM, Wadsworth MH 2nd, Hughes TK, Bryson BD, Butler A, Satija R, Fortune S, Love JC, and Shalek AK (2017). Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat Methods 14, 395–398. [PubMed: 28192419]

Giladi A, and Amit I (2018). Single-Cell Genomics: A Stepping Stone for Future Immunology Discoveries. Cell 172, 14–21. [PubMed: 29328909]

Giustacchini A, Thongjuea S, Barkas N, Woll PS, Povinelli BJ, Booth CAG, Sopp P, Norfo R, Rodriguez-Meira A, Ashley N, et al. (2017). Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. Nat Med 23, 692–702. [PubMed: 28504724]

Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, Woll P, Mead A, Alford KA, Rout R, et al. (2011). Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell 19, 138–152. [PubMed: 21251617]

Hartwig T, Montinaro A, von Karstedt S, Sevko A, Surinova S, Chakravarthy A, Taraborrelli L, Draber P, Lafont E, Arce Vargas F, et al. (2017). The TRAIL-Induced Cancer Secretome Promotes a Tumor-Supportive Immune Microenvironment via CCR2. Mol Cell 65, 730–742 e735. [PubMed: 28212753]

Hay S, Ferchen K, Chetal K, Grimes HL, and Salomonis N (2018). The Human Cell Atlas Bone Marrow Single-Cell Interactive Web Portal. Exp Hematol.

Karamitros D, Stoilova B, Aboukhalil Z, Hamey F, Reinisch A, Samitsch M, Quek L, Otto G, Repapi E, Doondeea J, et al. (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. Nat Immunol 19, 85–97. [PubMed: 29167569]

Kluk MJ, Lindsley RC, Aster JC, Lindeman NI, Szeto D, Hall D, and Kuo FC (2016). Validation and Implementation of a Custom Next-Generation Sequencing Clinical Assay for Hematologic Malignancies. J Mol Diagn 18, 507–515. [PubMed: 27339098]

Kreso A, and Dick JE (2014). Evolution of the cancer stem cell model. Cell Stem Cell 14, 275–291. [PubMed: 24607403]

Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, Levine JE, Wang J, Hahn WC, Gilliland DG, et al. (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature 442, 818–822. [PubMed: 16862118]

Laurenti E, Doulatov S, Zandi S, Plumb I, Chen J, April C, Fan JB, and Dick JE (2013). The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat Immunol 14, 756–763. [PubMed: 23708252]

Laurenti E, and Gottgens B (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature 553, 418–426. [PubMed: 29364285]

Leick MB, and Levis MJ (2017). The Future of Targeting FLT3 Activation in AML. Curr Hematol Malig Rep 12, 153–167. [PubMed: 28421420]

Levine JH, Simonds EF, Bendall SC, Davis KL, Amir el AD, Tadmor MD, Litvin O, Fienberg HG, Jager A, Zunder ER, et al. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. Cell 162, 184–197. [PubMed: 26095251]

Lichtenegger FS, Krupka C, Haubner S, Kohnke T, and Subklewe M (2017). Recent developments in immunotherapy of acute myeloid leukemia. J Hematol Oncol 10, 142. [PubMed: 28743264]

Macaulay IC, Ponting CP, and Voet T (2017). Single-Cell Multiomics: Multiple Measurements from Single Cells. Trends Genet 33, 155–168. [PubMed: 28089370]

Ng SW, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, Arruda A, Popescu A, Gupta V, Schimmer AD, et al. (2016). A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature 540, 433–437. [PubMed: 27926740]

Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME, Habib N, Yosef N, Chang CY, Shay T, et al. (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell 144, 296–309. [PubMed: 21241896]

Pabst T, Mueller BU, Harakawa N, Schoch C, Haferlach T, Behre G, Hiddemann W, Zhang DE, and Tenen DG (2001). AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. Nat Med 7, 444–451. [PubMed: 11283671]

Pollyea DA, and Jordan CT (2017). Therapeutic targeting of acute myeloid leukemia stem cells. Blood 129, 1627–1635. [PubMed: 28159738]

Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, Rodman C, Luo CL, Mroz EA, Emerick KS, et al. (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. Cell 171, 1611–1624 e1624. [PubMed: 29198524]

Pyzer AR, Stroopinsky D, Rajabi H, Washington A, Tagde A, Coll M, Fung J, Bryant MP, Cole L, Palmer K, et al. (2017). MUC1-mediated induction of myeloid-derived suppressor cells in patients with acute myeloid leukemia. Blood 129, 1791–1801. [PubMed: 28126925]

Ustun C, Miller JS, Munn DH, Weisdorf DJ, and Blazar BR (2011). Regulatory T cells in acute myelogenous leukemia: is it time for immunomodulation? Blood 118, 5084–5095. [PubMed: 21881045]

van Dijk EL, Jaszczyszyn Y, Naquin D, and Thermes C (2018). The Third Revolution in Sequencing Technology. Trends Genet 34, 666–681. [PubMed: 29941292]

Veglia F, Perego M, and Gabrilovich D (2018). Myeloid-derived suppressor cells coming of age. Nat Immunol 19, 108–119. [PubMed: 29348500]

Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, Hirche C, Lutz C, Buss EC, Nowak D, et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol 19, 271–281. [PubMed: 28319093]

Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S, et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356.

Weinreb C, Wolock S, Tusi BK, Socolovsky M, and Klein AM (2018). Fundamental limits on dynamic inference from single-cell snapshots. Proc Natl Acad Sci U S A 115, E2467–E2476. [PubMed: 29463712]

Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347, 1138–1142. [PubMed: 25700174]

Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat Commun 8, 14049. [PubMed: 28091601]
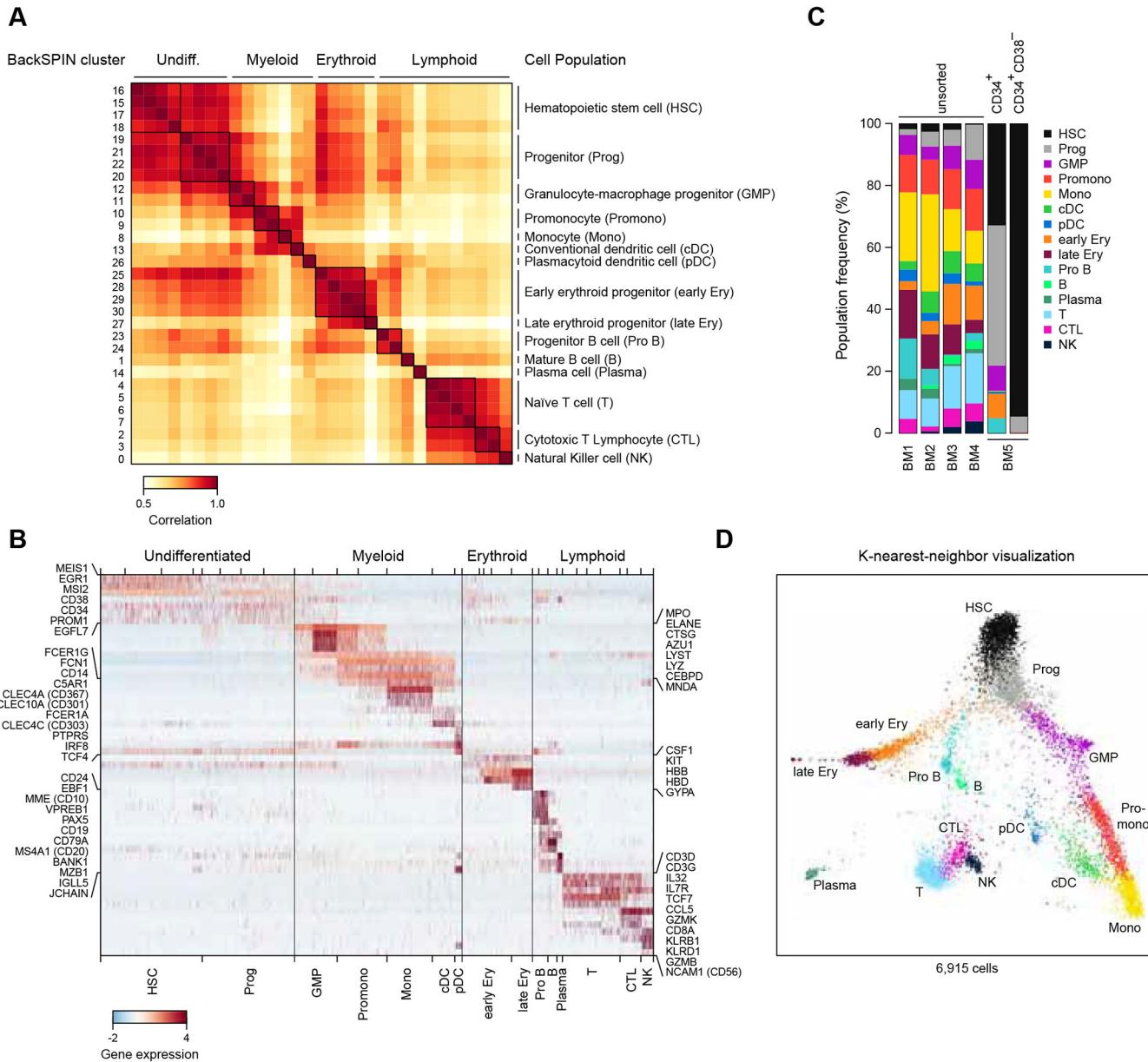
Highlights

- Technology for high-throughput single-cell RNA-sequencing and genotyping

- Variable cell type composition of AML correlates to genetics and outcome

- Primitive AML cells aberrantly co-express stemness and myeloid priming genes

- Differentiated AML cells express immunomodulatory factors and suppress T-cells
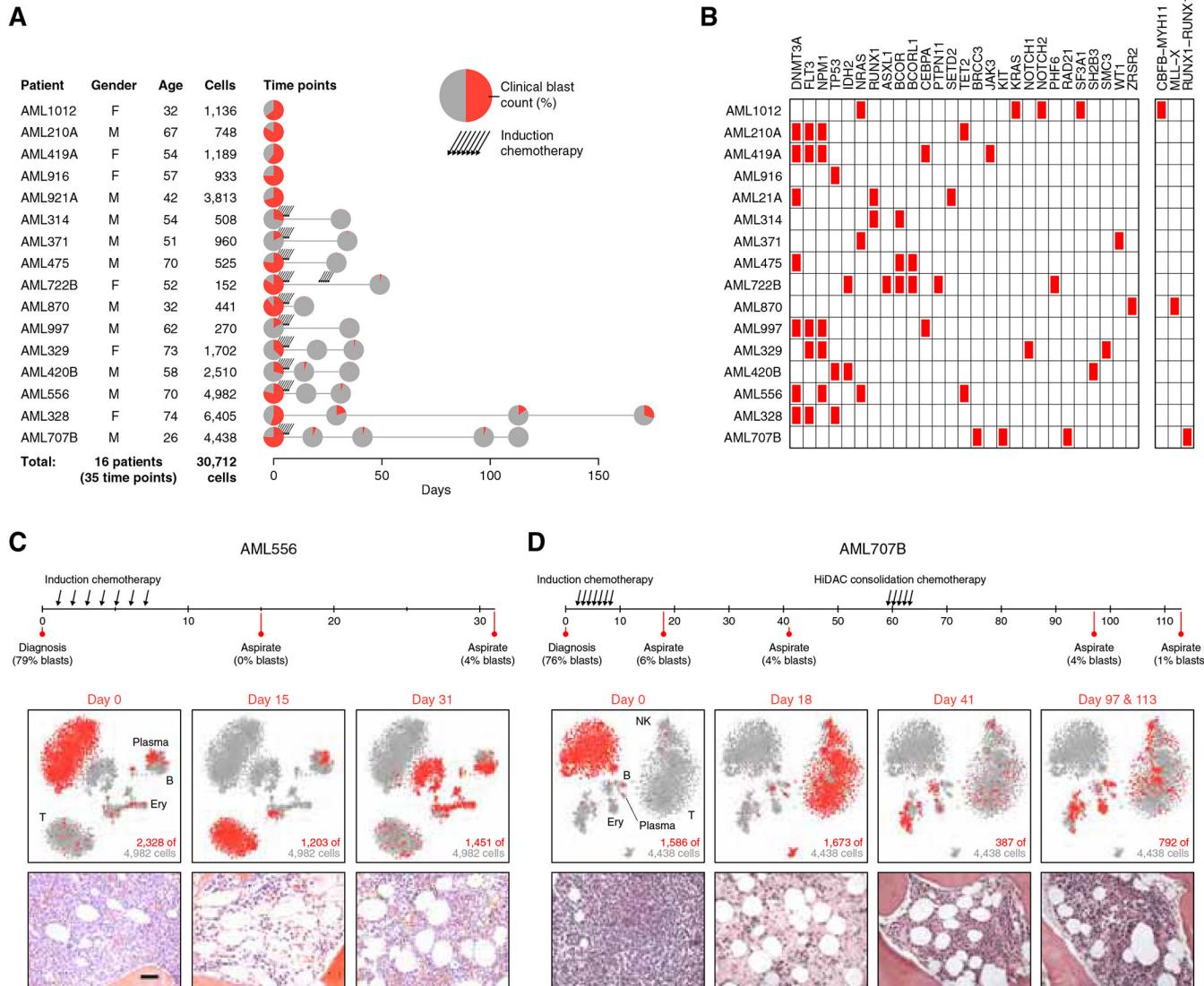
Figure 1. Identification of cell populations in healthy BM samples

A. BackSPIN clustering of scRNA-seq data for 6,915 hematopoietic cells from normal BM identified 31 clusters of cells with similar transcriptional states. Heatmap shows the pairwise correlation between the average expression profiles of these clusters (rows and columns). Clusters were merged into 15 cell populations based on marker gene expression (right).

B. Heatmap shows the expression of 55 selected cell type-specific genes (rows) across 6,195 single cells ordered by the BackSPIN-defined clusters (columns).

C. Stacked barplots show the frequencies of BackSPIN-defined cell types in five normal BMs.

D. KNN visualization of 6,915 single-cell transcriptomes (points), with similar cells positioned closer together. Points are color-coded by cell type annotations as in C.

See also Figure S1.
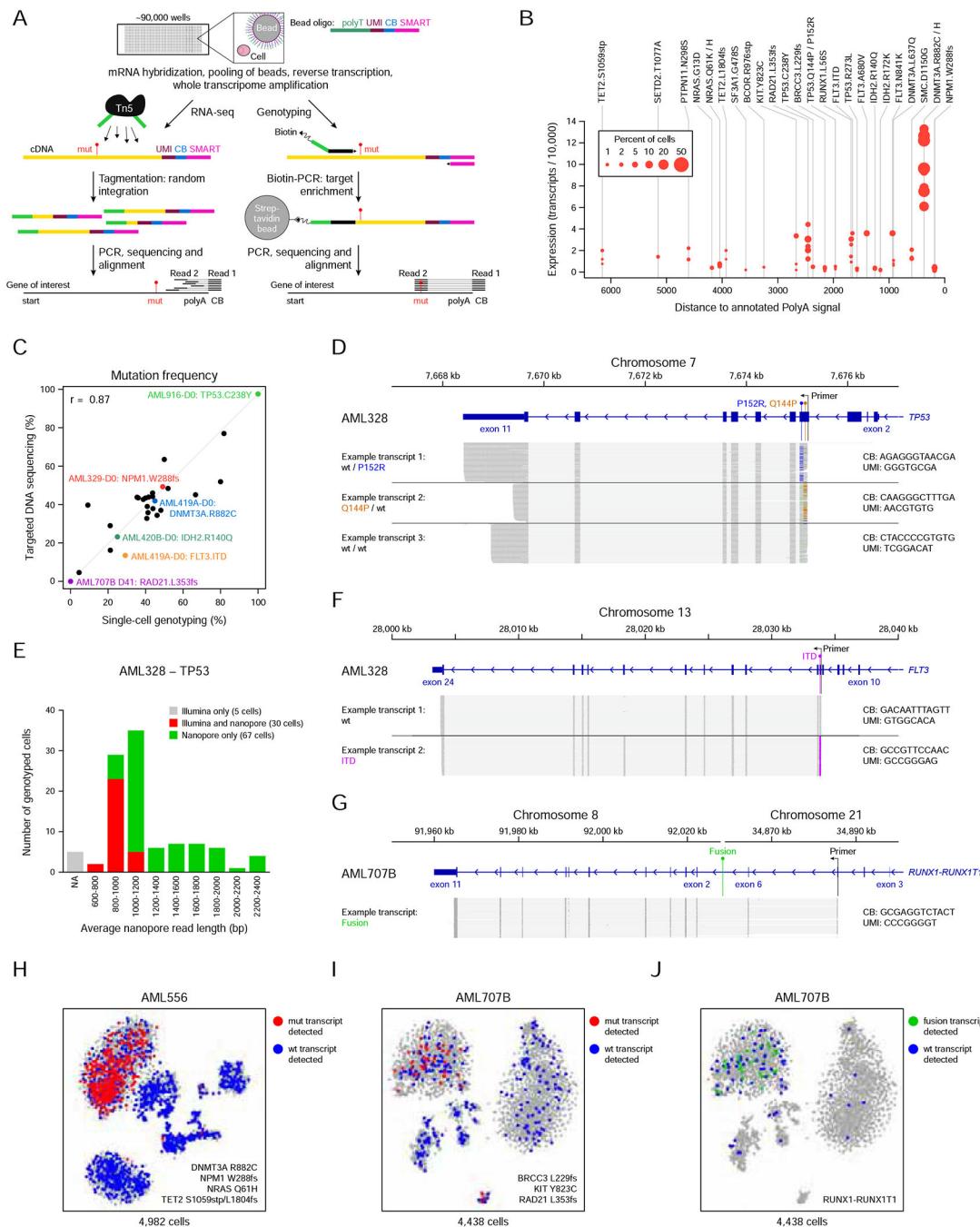
Figure 2. Single-cell profiling of AML tumor ecosystems

A. Overview of AML patients and BM aspirate collections. Cell numbers reflect single-cell transcriptomes that passed quality thresholds. For each patient, pie charts indicate time of sample collections, relative to diagnosis, and clinical blast count.

B. Chart shows genetic alterations (red) detected in our cohort by targeted DNA sequencing and cytogenetics.

C-D. t-SNE plots show single cells from AML556 (**C**) or AML707B (**D**) at successive collections. Each plot shows cells from the indicated time point (red) and other time points (gray). t-SNE plots and corresponding H&E stains depict marrows dominated by AML cells at presentation (Day 0), hypocellular marrows with T-cells after chemotherapy (Day 15–18), or repopulating hematopoiesis (Day 31–41). Scale bar 50 μm.
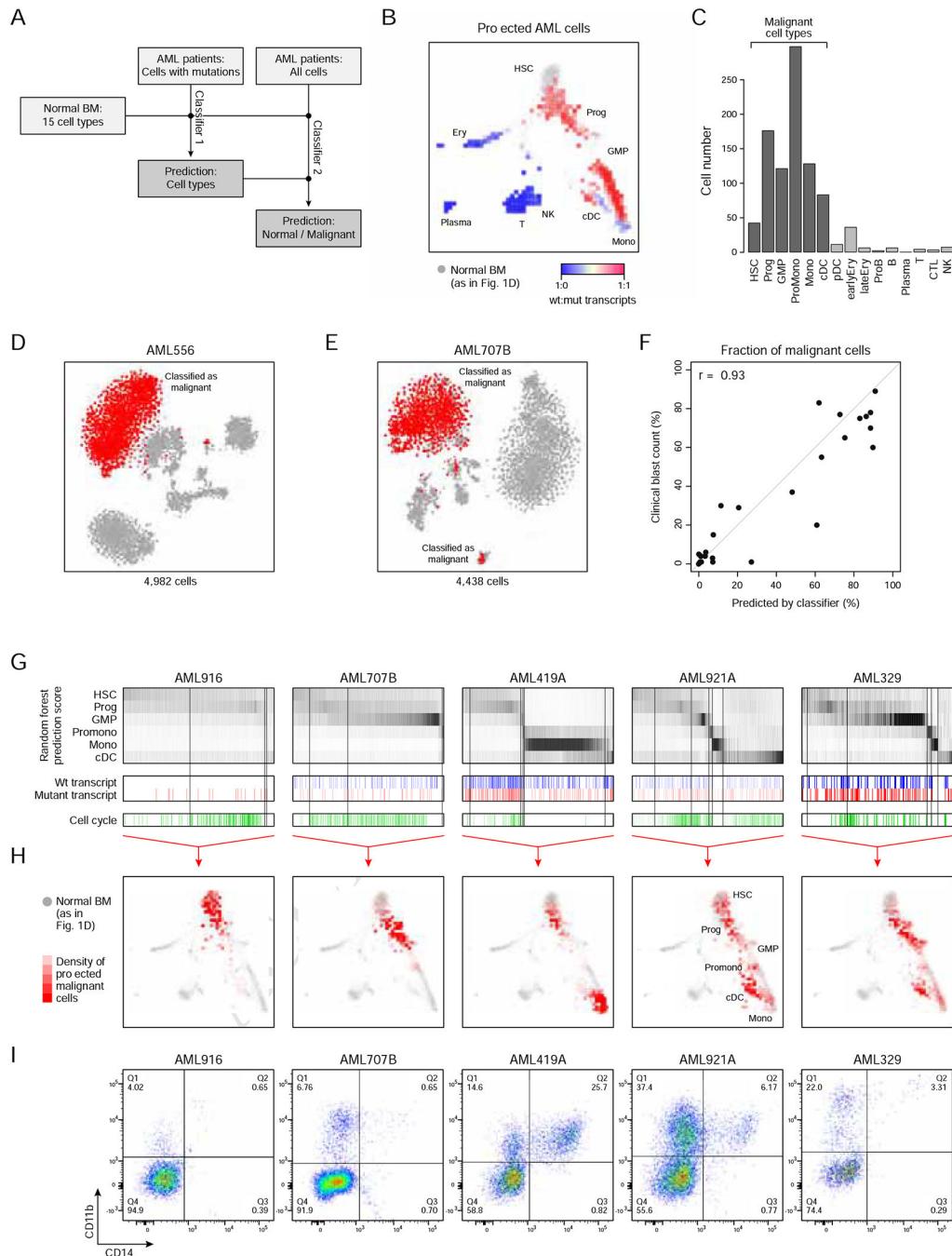
See also Table S1.

Figure 3. Single-cell genotyping by short-read and nanopore sequencing

A. Illustration depicts procedures for acquiring transcriptional and genotypic information from single cells. Nanowell plates and beads generate WTA product wherein each cDNA is appended to a UMI, a cell-specific barcode (CB), and a priming site (SMART). Product is split and used as input for Tn5-mediated scRNA-seq library generation (left) and single-cell genotyping (right). The single-cell genotyping reaction utilizes biotinylated primers to amplify mutational sites in target genes along with corresponding UMI and CB for sequencing.

B. Bubble plot depicts the frequency of mutation detection by single-cell genotyping with short-read Illumina sequencing. Detection is more efficient for mutations in highly expressed genes and near the 3' polyA signal.

C. Scatter plot compares mutation frequencies from DNA sequencing (y-axis) or single-cell genotyping (x-axis). Each point corresponds to a mutational site in a specific AML sample. Six examples are highlighted.

D. Genome plot illustrates long nanopore reads for three selected *TP53* transcripts from AML328. For each transcript, 100 reads are shown (reads were matched by CB and UMI, indicating they came from the same transcript). Black arrow indicates the location of the primer used for amplification. Base mismatches encoding Q144P (orange) or P152R (blue) mutations are indicated.

E. Stacked barplot shows the number of cells in which wild-type or mutant *TP53* transcripts of indicated lengths were detected by short-read (gray) or nanopore sequencing (green), or both (red). Fragment length was determined from the nanopore data.

F. Genome plot illustrates nanopore reads for two selected *FLT3* transcripts from AML328. For each transcript, 100 reads corresponding to the ITD or wild-type allele are shown. Black arrow indicates the location of the primer used for amplification. Base insertions representing a newly detected 60 bp ITD are indicated in exon 14 (pink).

G. Genome plot illustrates nanopore reads of a fusion transcript from AML707B aligning to the *RUNX1T1* (left) and *RUNX1* (right) loci. One hundred reads are shown. Nanopore sequencing enabled detection of the fusion without prior knowledge of the junction.

H-I. t-SNE plots for AML556 (**H**) and AML707B (**I**) as in Figure 2C–D show cells for which wild-type (blue) or mutant (red) transcripts were detected by single-cell genotyping with Illumina sequencing.

J. t-SNE plot for AML707B shows cells for which *RUNX1-RUNX1T1* fusion transcripts (green) were detected by nanopore sequencing.

See also Figure S2 and Table S2.

Figure 4. Machine learning classifier distinguishes cell types in the AML ecosystem
A. Schematic of machine learning classifiers used to predict AML cell types (Classifier 1) and to distinguish malignant from normal cells in AML tumors (Classifier 2).
B. KNN visualization shows single-cell transcriptomes from normal BM (gray; as in Figure 1D). Cells from AML samples for which wild-type or mutant transcripts were detected were projected onto this graph according to their similarity to the normal cells. Boxes depicting the density of projected cells are colored according to the ratio between wild-type and

mutant transcripts. Cells with mutant transcripts (red) project along the HSC to myeloid differentiation axis.

C. Barplot shows classification of AML cells with mutant transcripts by the first Random forest classifier. The majority are classified as one of six cell types along the HSC to myeloid axis, thereby defining six malignant cell types.

D-E. t-SNE plots of AML556 (**D**) and AML707B (**E**) as in Figure 2C–D with cells colored by their classification as malignant (red) or normal (grey).
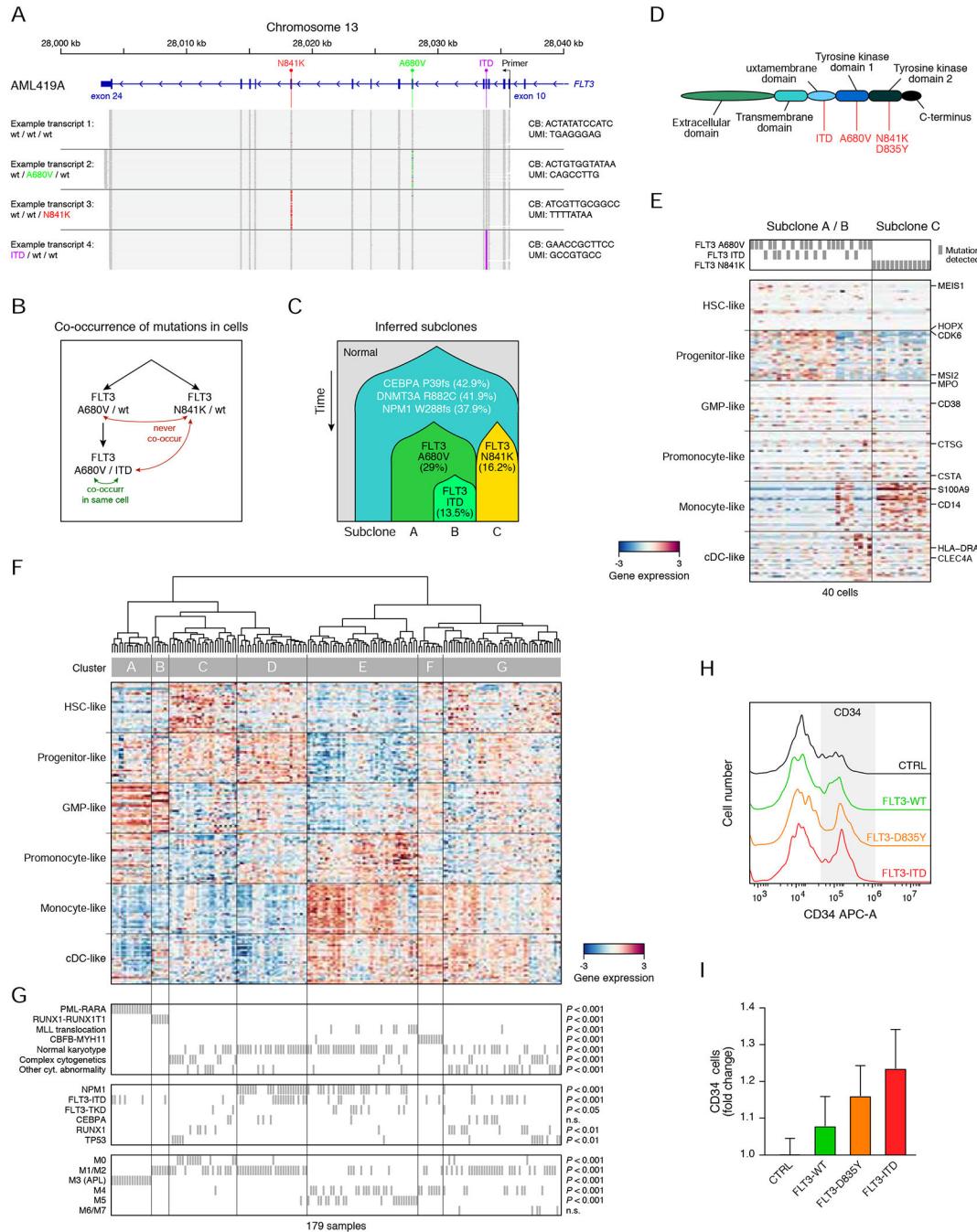
F. Scatter plot compares clinical blast counts (y-axis) to the fraction of cells classified as malignant by the machine learning classifier (x-axis). Each point corresponds to a specific AML BM aspirate (n = 27).

G. Heatmaps show cell type prediction scores (rows) for all malignant cells (columns) from five representative tumors. Cells in which wild-type and/or mutant transcripts were detected, or that express cell cycle signature genes are indicated below.

H. KNN visualizations show single-cell transcriptomes of normal BM cells (gray; as in Figure 1D). Malignant cells from the respective AMLs were projected onto this graph according to their similarity to the normal cells. The density of projected cells (red) conveys the distinct cell type compositions of these tumors.

I. Flow cytometry plots show expression of myeloid differentiation markers by the AML samples.

See also Figure S3–5.

Figure 5. AML cellular hierarchies correlate with underlying genetic alterations

A. Genome plot illustrates nanopore reads for four selected *FLT3* transcripts from AML419A. For each transcript, 100 reads are shown. Black arrow indicates the location of the primer used for amplification (exon 11). Base mismatches encoding A680V (exon 16; green) or N841K (exon 20; red) mutations are indicated. Base insertions representing a 24 bp ITD are indicated in exon 14 (pink). The mutations do not co-occur on the same transcripts.

B-C. Diagrams show AML419 evolution inferred from co-occurrence of mutations in single cells (**B**) and VAFs from bulk DNA sequencing (**C**). The most likely model yields one subclone with an A680V mutation, a second subclone with an ITD, and a third subclone that exclusively harbors an N841K mutation.

D. Diagram shows FLT3 protein domains and location of mutations.

E. Heatmap shows expression of 180 signature genes for the six malignant cell types (rows) in 40 single cells from AML419A (columns). Cells were assigned to subclone A or B, or subclone C on the basis of *FLT3* genotypes.
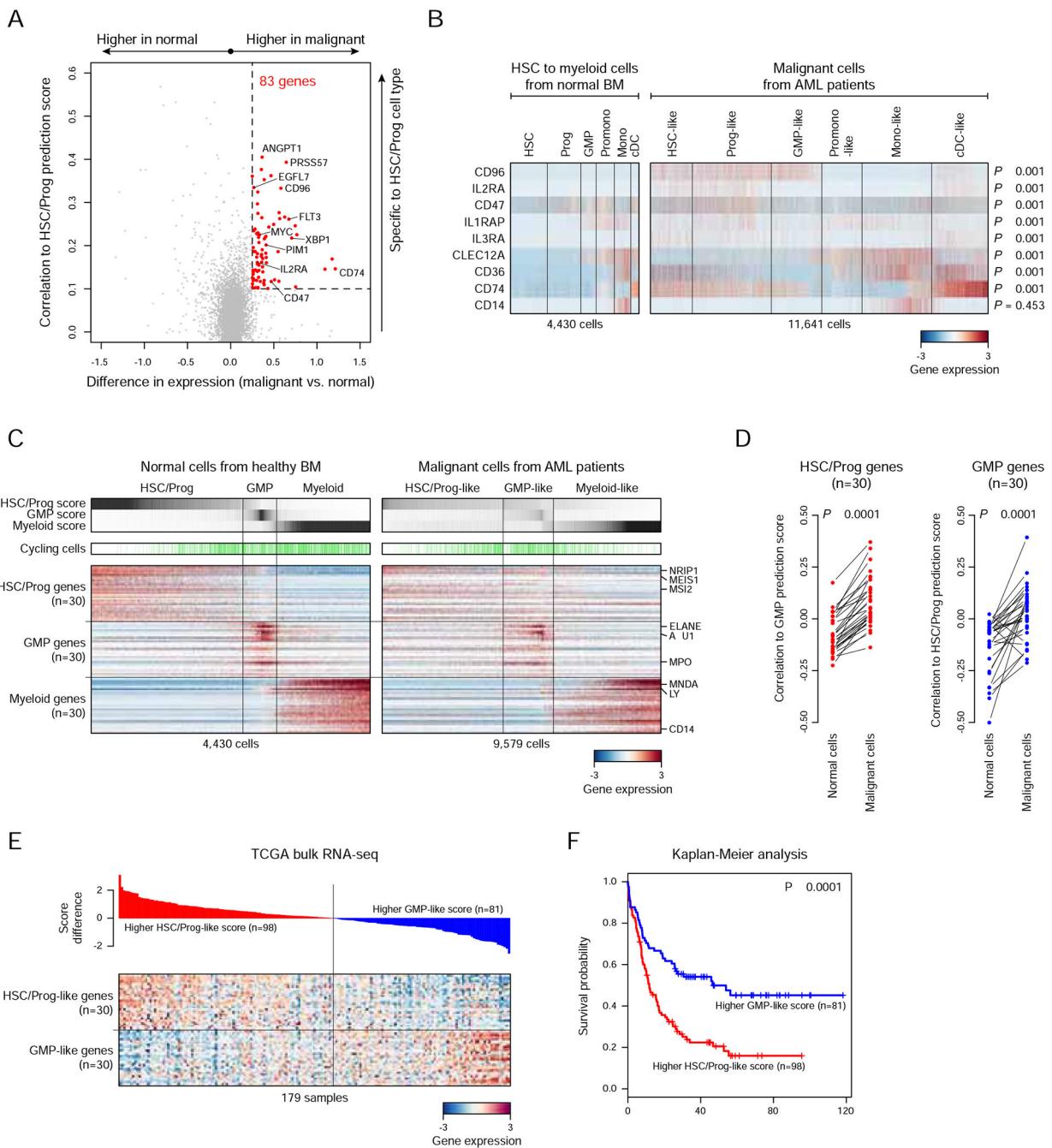
F. Heatmap shows expression of 180 signature genes for the six malignant cell types (rows) in 179 AMLs profiled by bulk RNA-seq (columns). Unsupervised clustering revealed seven subsets of patients with different inferred cell type abundances (clusters A-G).

G. Charts indicate chromosomal aberrations (top), mutations (middle) and FAB classifications (bottom) for AMLs in **F**. Correspondence between cell type compositions and genetics is evident. *P*-values indicate non-random distribution of events between clusters (Fisher's exact test). n.s., not significant.

H. Flow cytometry histograms show expression of the primitive cell marker CD34 in MUTZ-3 cells, four days after transduction with FLT3-WT, FLT3-D835Y, FLT3-ITD or a control gene (*luciferase*).

I. Plot shows change in the percent of CD34$^+$ cells following transduction of FLT3 variants as in **H**. *P*-values were calculated using Student's *t*-test compared to CTRL (mean + SD of n = 6 transductions). * *P* < 0.05, ** *P* < 0.01, **** *P* < 0.0001.

See also Figure S5 and Table S3.

Figure 6. Dysregulated transcriptional programs in malignant progenitors

A. Scatterplot positions genes (dots) by their preferential expression in malignant HSC/Prog-like cells relative to normal counterparts (x-axis), and by their correlation to HSC/Prog prediction scores across malignant cells (y-axis). Genes in the top right are preferentially expressed in malignant HSC/Prog-like cells, relative to normal progenitors and other malignant cell types (red).

B. Heatmap shows expression of surface markers (rows) in normal BM cells (left, columns) or malignant cells from diagnostic AML aspirates (right, columns). *CD14* is shown for

comparison. *P*-values between HSC-like cells and normal HSCs are calculated by FDR-adjusted Kruskal test.

C. Heatmaps show expression of normal BM-derived signature genes for HSC/Prog, GMP or differentiated myeloid cells (n = 90; rows) in normal BM (left, columns) or malignant AML cells (right, columns). Cells are ordered by their classifier prediction scores (shown on top). Cells that express cell cycle genes are indicated. Primitive AML cells co-express HSC/Prog and GMP programs.
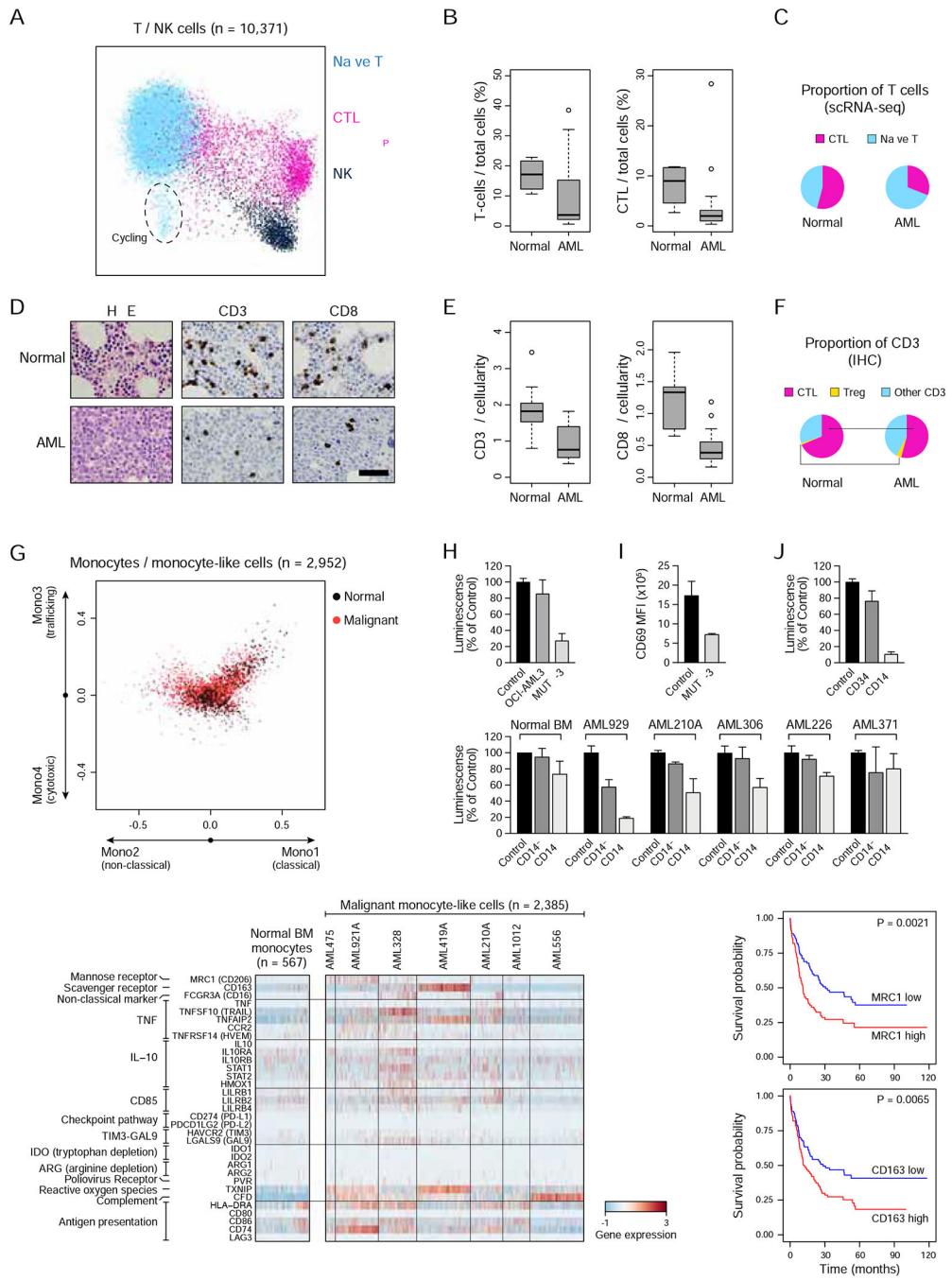
D. Plot shows correlation of 30 normal BM-derived HSC/Prog signature genes (red dots) with GMP prediction score across normal or malignant cells. Right: Plot shows correlation of 30 normal BM-derived GMP signature genes (blue dots) with HSC/Prog prediction score across normal or malignant cells. HSC/Prog genes and GMP genes are aberrantly correlated in malignant cells. *P*-values were calculated by paired Wilcoxon test.

E. HSC/Prog-like and GMP-like signatures were applied to TCGA RNA-seq profiles. Heatmap shows expression of 60 signature genes (rows) across 179 bulk AML profiles (columns).

F. Kaplan-Meier curves show the survival of 179 AML patients stratified by signature scores in **E**. Patients with higher HSC/Prog-like scores have worse outcomes. *P*-value was calculated by log-rank test.

See also Figure S6 and Table S3–4.

Figure 7. AML-derived monocyte-like cells have immunomodulatory properties

A. KNN visualization shows all T- and NK cells identified in normal BM and AML samples. BackSPIN analysis distinguished three clusters of cells that express markers of naïve T-cells, CTLs or NK cells.

B. Boxplots show the relative numbers of cells annotated as T-cells or CTLs by scRNA-seq (median ± quartiles for 4 normal BMs and 16 diagnostic AMLs).

C. Pie charts show relative numbers of cells annotated as CTLs or naïve T-cells by scRNA-seq (mean for two normal BM donors and six diagnostic AMLs with >50 T / NK cells).

D. Representative IHC stains for T-cells (CD3) and CTLs (CD8) in normal BM and AML. H & E stains are also shown. Scale bar 50 μm.

E. Boxplots show relative numbers of T-cells or CTLs identified in IHC stains (median ± quartiles for 15 normal BMs and 15 diagnostic AMLs).

F. Pie charts show relative numbers of CTLs (CD8$^+$), T-regs (CD25$^+$FOXP3$^+$) and other T-cells, per IHC stains (mean for 15 normal donors and 15 AMLs). AMLs have fewer T-cells and CTLs, but greater numbers of T-regs, consistent with an immunosuppressive microenvironment.

G. Scatterplot shows 2,385 malignant monocyte-like cells from AMLs (red dots) and 567 monocytes from normal BMs (black dots). Cells are placed according to their signature scores for Mono1 (right), Mono2 (left), Mono3 (up) and Mono4 (down) (Villani et al., 2017).

H. Barplot shows activation of a CD4$^+$ T-cell line after stimulation with CD3/CD28 beads *in vitro*. T-cell activation was read out by an NFAT reporter. The assay was performed in the absence (Control) or presence of OCI-AML3 or MUTZ-3 cells (mean ± SD of n   3 experiments).

I. Barplot shows activation of primary CD4$^+$ T-cells after stimulation with CD3/CD28 beads *in vitro*. T-cell activation was read out by flow cytometry for CD69. The assay was performed in the absence or presence of MUTZ-3 cells (mean ± SD of n = 6 replicates).

J. Barplots show activation of a CD4$^+$ T-cell line as in **H**. The assay was performed in the presence of 100,000 sorted CD34$^+$ or CD14$^+$ MUTZ-3 cells (n = 3 experiments).

K. Barplots show activation of a CD4$^+$ T-cell line as in **H**, **J**. The assay was performed in the presence of 100,000 sorted CD14$^-$ or CD14$^+$ primary cells from normal BMs (n = 6 donors) or AML aspirates (n   3 technical replicates each). Significance is only indicated when T-cell activation was reduced >1.5-fold compared to Control.

L. Heatmap shows expression of curated immunomodulatory genes (rows) in monocytes from normal BM (left, columns) and monocyte-like cells from AMLs (right, columns). Only AMLs with >50 monocyte-like cells are shown.

M. Kaplan-Meier curves show the survival of 179 AML patients stratified by expression of *MRC1/CD206* or *CD163*. *P*-values in **M** were calculated by log-rank test. All other *P*-values were calculated by Student's *t* test. * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001, **** *P* < 0.0001. See also Figure S7 and Table S4.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse CD15-V450, clone MMA	BD Biosciences	Cat# 642917; RRID: AB_1645751
Monoclonal mouse CD34-FITC, clone 8G12	BD Biosciences	Cat# 348053; RRID: AB_2228982
Monoclonal mouse CD38-PE, clone HB7	BD Biosciences	Cat# 347687; RRID: AB_400341
Mouse CD14-APC, clone RMO52	Beckman Coulter	Cat# IM2580U
Monoclonal mouse CD11b-APC-Cy7, clone ICRF44	BD Biosciences	Cat# 557754; RRID: AB_396860
Monoclonal mouse FLT-3, clone 40406	R & D Systems	Cat# MAB308; RRID: AB_2104978
Mouse CD34-APC, clone 8G12	BD Biosciences	Cat# 340441; RRID: AB_400514
Anti-Human CD14-PC7 antibody	Beckman Coulter	Cat# A22331; RRID: AB_10639528
Monoclonal rat CD4-FITC, clone A161A1	BioLegend	Cat# 357405; RRID: AB_2562356
Mouse CD33-PE-Cy7 , clone P67.6	BD Biosciences	Cat# 333946; RRID: AB_399961
Monoclonal mouse CD69-APC, clone FN50	BioLegend	Cat# 310909; RRID: AB_314844
Monoclonal mouse Ki67-Alexa Fluor® 647, clone B56	BD Biosciences	Cat# 561126; RRID: AB_10611874
Biological Samples		
See Table S1 for a list of patients included in the study.		
Chemicals, Peptides, and Recombinant Proteins		
(3-Aminopropyl)triethoxysilane (APTES)	Sigma	Cat# A3648–100ML
p-Phenylene diisothiocyanate (PDITC)	Sigma	Cat# 258555–5G
Pyridine	Sigma	Cat# 270970–1L
N,N-Dimethylformamide (DMF)	Sigma	Cat# 227056–2L
Chitosan	Sigma	Cat# C3646–100G
Poly(L-glutamic) acid sodium solution	Sigma	Cat# P4761–100MG
Sodium Carbonate ReagentPlus	Sigma	Cat# S2127–500G
Guanidine Thiocyanate (GITC)	Sigma	Cat# G9277–500g
Sarkosyl (10%, 500 ml)	Fisher Scientific	Cat# 50–843-132
Maxima H Minus Reverse Transcriptase	Thermo Fisher	Cat# EP0753
20% Ficoll PM-400	Sigma	Cat# F5415–50mL
Betaine	Sigma	Cat# B0300–5VL
1 M MgCl2	Sigma	Cat# 63069–100ML
1 M Tris-HCl pH 8.0	Boston BioProducts	Cat# BBT-80
10 mM dNTPs	New England BioLabs	Cat# N0447L
RNAse Inhibitor	Thermo Fisher	Cat# AM2696
Exonuclease I	New England Biolabs	Cat# M0293S
Poly(ethylene glycol) (PEG) Mn 400	Sigma	Cat# 202398–250G

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Poly(ethylene glycol) (PEG) BioUltra 8,000	Sigma	Cat# 89510–250G-F
Acetone	Avantor	Cat# 2440–10
BSA	Sigma	Cat# A9418–100G
2-Mercaptoethanol	Fisher Scientific	Cat# NC0753648
Tween-20	Fisher Scientific	Cat# 65–520-4100ML
EDTA (0.5M, pH 8.0)	Boston Bioproducts	Cat# BM-150
Sodium Chloride	Fisher Chemical	Cat# S671–3
UltraPure Distilled Water	Thermo Fisher	Cat# 10977023
Sodium hydroxide	Sigma	Cat# S8045–500G
AMPure XP (SPRI) beads	Beckman Coulter	Cat# A63881
Critical Commercial Assays		
KAPA HiFi Hotstart Readymix PCR Kit	Kapa Biosystems	Cat# KK2602
Nextera XT DNA Library Preparation Kit	Illumina	Cat# FC-131–1096
Hybridization Chamber Kit - SureHyb enabled	Agilent	Cat# G2534A
MACOSKO-2011–10 mRNA Capture Beads	Chemgenes	Cat# NC0927472
Dynabeads™ kilobaseBINDER™ Kit	ThermoFisher	Cat# 60101
PfuUltra II Hotstart PCR Master Mix	Agilent	Cat# 600852
Qubit dsDNA HS Assay Kit	ThermoFisher	Cat# Q32854
BioA High Sensitivity DNA Kit	Agilent	Cat# 5067–4626
High Sensitivity D5000 ScreenTape	Agilent	Cat# 5067–5592
Bio-Glo™ Luciferase Assay System	Promega	Cat# G7941
Jurkat NFAT reporter cells	Promega	Cat# J1621
Dynabeads® Human T-Activator CD3/CD28	Gibco	Cat# 11131D
Dead Cell Removal Kit	Miltenyi Biotec	Cat# 130–090-101
CD34 MicroBead Kit	Miltenyi Biotec	Cat# 130–046-702
CD14 MicroBeads	Miltenyi Biotec	Cat# 130–050-201
MS Columns	Miltenyi Biotec	Cat# 130–042-201
FuGENE® HD Transfection Reagent	Promega	Cat# E2311
Lenti-X™ Concentrator	Takara	Cat# 631232
Ligation Sequencing Kit	Oxford Nanopore Technologies	Cat# SQK-LSK109
NEBNext Ultra II End-Repair/dA-tailing Module	New England BioLabs	Cat# E7645
Blunt/TA Ligase Master Mix	New England BioLabs	Cat# M0367
QuikChange Lightning Site-Directed Mutagenesis Kit	Agilent	Cat# 210518
pENTR™/D-TOPO™ Cloning Kit	ThermoFisher	Cat# K240020
Permeabilizing Solution 2	BD Biosciences	Cat# 347692
BD Cytofix™ Fixation Buffer	BD Biosciences	Cat# 554655

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Raw data	GEO	Accession number GSE116256
Processed data	GEO	Accession number GSE116256
Experimental Models: Cell Lines		
MUTZ-3	DSMZ	Cat# ACC-295; RRID: CVCL_1433
OCI-AML3	University of Toronto, Minden Lab	N/A
5637	ATCC	Cat# HTB-9; RRID: CVCL_0126
THP-1	Broad Institute Genetic Perturbation Platform	N/A
Oligonucleotides		
See Table S2 for a list of oligonucleotide sequences.		
Recombinant DNA		
VSV.G	Addgene	Cat# 14888
dRT-pMDLg/pRRE	Addgene	Cat# 60488
pRSV-Rev	Addgene	Cat# 12253
pMAL	University of Toronto, Dick lab	N/A
Software and Algorithms		
R version 3.4	R Core Team	https://www.r-project.org
R package - data.table	CRAN	https://cran.r-project.org/web/packages/data.table/index.html
R package - Rsamtools	Bioconductor	http://bioconductor.org/packages/release/bioc/html/Rsamtools.html
R package - Rtsne	Github	https://github.com/jkrijthe/Rtsne
R package - randomForest	CRAN	https://cran.r-project.org/web/packages/randomForest/index.html
STAR version 2.5.2b	Github	https://github.com/alexdobin/STAR
SPRING	Kleintools (Weinreb et al., 2018)	https://kleintools.hms.harvard.edu/tools/spring.html
BackSPIN	Github (Zeisel et al., 2015)	https://github.com/linnarsson-lab/BackSPIN
BWA-MEM	Cornell University	https://arxiv.org/abs/1303.3997
FlowJo version 10.4.2	TreeStar	https://www.flowjo.com
Prism 7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Integrative Genomics Viewer (IGV version 2.4.8)	Broad Institute	http://software.broadinstitute.org/software/igv/ download
Albacore version 2.3.3	Github	https://github.com/dvera/albacore
SC3	Bioconductor	http://bioconductor.org/packages/release/bioc/html/SC3.html