Supplementary Methods

AML subtype classification and comparison groups

Samples were classified first by karyotype and then by FLT3-ITD status. All AML patients were tested for insertional mutations in NPM and FLT3-ITD, FLT3 tyrosine kinase domain mutations (D835), and CEBPα mutations as part of standard diagnostic testing. Additionally, testing for exon 8 and 17 mutations in *c-Kit* was performed in 23 patients (at the discretion of the treating physician) and expanded mutational analysis (for common mutations in *APC*, *BRAF*, *CTNNB1*, *DNMT3a*, *EGFR*, *ERBB2/HER2*, *IDH1*, *IDH2*, *KRAS*, *NRAS*, *Notch1*, *PIK3Ca*, *PTEN*, and *p53*) became available in 2011 and was performed on samples from 24 of the patients. Two patients (AML#36 and AML#42) who underwent biopsy for suspected AML were found to have blast counts of less than 20%, both of these patients developed frank AML within 60 days of biopsy and were analyzed with the other AML samples (both were normal karyotype, FLT3wt), as both could have potentially been eligible for AML induction treatment. Two patient samples could not be used for all analyses: IdU was not added to the sample from AML#33 and could not be analyzed for cell cycle. The sample from APL#1 was left at room temperature for 3 hours prior to processing and was only used for immunophenotypic analysis. One sample, AML#18, came from a patient originally found to have NK-AML with a FLT3-ITD mutation, but at the time of biopsy had developed an adverse-risk karyotype (inv(1), t(11:15), -18) with a persistent FTL3-ITD mutation. This sample was classified as an adverse-risk karyotype (ARK) for most analyses, and was ignored for comparisons of FLT3-ITD+ vs. FLT3wt (though all differences between FLT3-ITD+

AML and FLT3wt AML samples remain statistically significant regardless of whether or not sample AML18 is included).

For cell cycle analyses, APL samples were excluded from subtype comparisons as this disease is currently treated with novel therapeutic approaches that are much less dependent on cell cycle state. Similarly, as the clinical significance of FLT3 tyrosine kinase domain mutations remains unclear, these samples were not included in these comparisons as the goal was to correlate cell cycle properties to known prognostic markers. Given the small number of samples from patients in complete remission (CR) and the diverse genetic and karyotypic properties of AML cells from these patients, these samples were not used for the majority of analyses; however, the one patient who achieved a sustained CR (>2 years; CR#2) was similar to the normal samples both immunophenotypically and with regard to cell cycle and intracellular signaling properties.

For analyses of particular aberrant markers, we focused on just the $Lin^-CD34^+CD38^{low}$ compartment. This was done for several reasons. First, we wanted to have the ability to compare populations across all markers without creating a bias by subgating populations by gating with other makers. Second, the $Lin^-CD34^+CD38^{low}$ gate was both one of the widest gates (allowing for the greatest sensitivity to detect variation) and the most consistent across different normal and patient samples. Third, we felt that this gate would be the easiest for other researchers (using either mass cytometry or high dimensional flow cytometry) to replicate. Fourth, as there is already a great deal of literature focused on Lin^-

$CD34^{+}CD38^{low}$ cell populations, we wanted to report results that could be compared to these studies. Medians for each marker in each of the individual gated subpopulations of HSPCs can be found in Supplementary Table 4 and demonstrate similar trends as the larger $CD34^{+}CD38^{\mathbf{low}}$ population.

Barcoding

Mass-tag barcoding was performed in groups of 20 samples using a transient partial permeabilization protocol(1). The unique pattern of three of the six stable Pd isotopes enables removal of doublet events (2). Each sample was split in half into the same position of the two 20-well barcoding plates (one for each of the two antibody staining panels). Each barcoding plate included at least three sample aliquots from one of the five healthy donors. An aliquot of the sample from donor #6 (Nl-6) was included in each barcoding plate and staining reaction as an internal reference standard to ensure that staining and antibody detection was consistent. Barcoding was performed on approximately 2 million fixed cells per sample placed into racked, 1.1-mL microtubes (BioExpress, Kaysville, UT, USA) using a multichannel pipette and a multichannel aspirator. Fixed cells were washed once in CSM and then washed once in PBS, followed by a second wash in PBS plus 0.02% saponin (Sigma-Aldrich, St. Louis, MO, USA). All pre-barcoding saponin washes were performed at 4 °C. Each wash step (re-suspension of cells, centrifugation of cells for 5 minutes at 600 X g, and aspiration of supernatant) was completed in approximately 10 minutes. After these washes, cells were resuspended in ~60 μL PBS plus 0.02% saponin and maintained at 4 °C prior to the application of mass-tag

barcoding reagent. A 100x DMSO stock of the mass tag barcoding reagent was then rapidly (<20 seconds) diluted into 1 mL ice-cold PBS plus 0.02% saponin and then quickly (<20 seconds) applied to the resuspended cells. Cells were incubated for 15 minutes (at room temperature) to allow covalent reaction of the barcode mass tags with the cells. After barcoding, cells were washed twice with CSM and then combined in a single tube. Cells were not re-exposed to saponin in subsequent manipulations or antibody staining steps. All antibody staining, methanol permeabilization, and sample measurement was performed with all cells (~40 million total) simultaneously in the same tube.

Mass-tagged barcoding reagents were prepared as described (2). Briefly, barcoding was performed with a pattern of three of the six stable Pd isotopes (102, 104, 105, 106, 108, 110) for each sample using isotopically purified palladium nitrate (Trace Sciences International, Richmond Hill, Ontario, Canada) and isothiocyanobenzyl-EDTA (Dojindo Molecular Technologies, Rockville, MD, USA) as the chelator. These were mixed in aqueous solution then immediately frozen and lyophilized. The resultant solid mass tag barcoding reagent was diluted in DMSO into 100x concentrated stock 96-well PCR plates (with each of the 20 wells containing a unique metal pattern) and frozen at -80 °C for up to 12 months. Plates were thawed immediately prior to use. Mass tag barcoding was performed at a final metal concentration of 300 nM; staining was equivalent for all Pd isotopes. The barcode signal from each cell was de-convoluted back into individual samples using a Matlab software application, which also allowed removal of doublet events.

Data analysis and gating

All mass cytometry data are displayed with an arcsinh transformation and a scale argument of five (except for linear scales used for Ir intercalator and cell length parameters). During data acquisition the cell subtraction value was set to -100 (thereby adding 100 counts to each channel). After acquisition, the effect of the cell subtraction setting was negated by subtracting a value of 100 from every channel of each FCS file using the flowCore package for R (10). These manipulations were performed to better estimate the effect of background subtraction and experimental noise for cells with low signal by allowing negative values to be displayed (3). Immunophenotypic assignments were based on previous studies from our laboratory (3, 4) and others (5). All gating and extraction of median expression levels was performed using Cytobank (www.cytobank.org). SPADE analysis was performed as previously described (6), clustering markers are indicated in Supplementary Table 2. Clusters were manually grouped and annotated into immunophenotypic populations based on examination of relevant biaxial plots (e.g., CD3 vs. CD45) of the cell events in each cluster by utilizing information from previous reports (3, 4), (5).

ViSNE analysis was performed in two tSNE dimensions using the CYT software tool and the viSNE analysis tool in Cytobank as described previously. (7) Data files were down-sampled to ≤5,000 events each, and the surface markers used for the analysis are shown in Supplementary Table 2. Statistical analysis of the viSNE results was performed by comparison of the calculated median tSNE1 and tSNE2 values by Mann-Whitney U testing. NK-AML FLT3-ITD$^+$ samples were significantly different from normal in both the tSNE1 and tSNE2 dimensions

(p=4.5x10^{-7} and 0.00023, respectively) as were the APL samples (p=0.012 and 0.0013, respectively). The FLT3wt NK-AML and CBF-AML samples were different from normal in the tSNE1 dimension (p=0.0023 and 0.00017, respectively). The FLT3-ITD^{+} NK-AML samples were also significantly different from the FLT3wt NK-AML samples in both dimensions (p=0.0049, and 0.0019) and different from the CBF-AML in the tSNE1 dimension (p=0.00046).

Immunophenotypic gates used in manual gating were defined based on the normal donor cell samples (Supplementary Figure 9), and the same gates were applied to all samples since staining of the normal and AML samples was performed in the same tube simultaneously (with the exception of minor adjustments to the CD7 vs. CD45 gate in samples with aberrant CD7 expression) (1). For some populations, CD33 appeared to better discriminate immunophenotypic HSC and multipotent progenitor (MPP) populations from more mature populations and was thus used instead of CD45RA, which did not stain as brightly by mass cytometry; the resulting CD33-negative HSC and MPP populations were all negative for CD45RA as well. Because of this, we feel that the gated boundaries of the CMP population may not be perfectly precise (it may include a small fraction of MPP or GMP cells).

As not all immunophenotypic makers were included in the intracellular signaling panel (staining panel B; Supplementary Table 2), distinct HSC and MPP populations could not be determined due to the lack of CD90. As a result, for the intracellular signaling analysis, an HSC/MPP gate (lin^{-}CD34^{+}CD38^{low}CD33^{low}), an MPP/common myeloid progenitor (CMP) gate (lin^{-}CD34^{+}CD38^{low}CD33$^{low\text{-}mid}$), and a CMP/granulocyte-macrophage progenitor (GMP) gate (lin^{-}CD34^{+}CD33^{mid}CD38$^{mid\text{-}}$

high) were used. Comparison of the SPADE analysis (in which cell samples stained with each of the panels were analyzed simultaneously) demonstrated that cells gated with these different methods clustered in the similar SPADE populations. Statistical comparisons of median marker expression levels and population frequencies between sample groups were performed with the Mann-Whitney U test. Comparisons of S-phase fractions across the immunophenotypic populations were the primary objective of the study, analysis of phosphorylated STAT5 was also pre-planned; all other comparisons were post-hoc analyses.

1. Behbehani GK, Thom C, Zunder ER, Finck R, Gaudilliere B, Fragiadakis GK, et al. Transient partial permeabilization with saponin enables cellular barcoding prior to surface marker staining. Cytometry A. 2014.
2. Zunder E.R. FR, Behbehani G.K., Amir E.D., Krishnaswamy S., Gonzalez V.D., Lorang C.G., Bjornson Z., Spitzer M.H., Bodenmiller B, Fantl W.J., Pe'er D., Nolan G.P. Palladium-based Mass-Tag Cell Barcoding with a Doublet-Filtering Scheme and Single Cell Deconvolution Algorithm. In Press. 2015.
3. Bendall SC, Simonds EF, Qiu P, Amir el AD, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science. 2011;332:687-96.
4. Behbehani GK, Bendall SC, Clutter MR, Fantl WJ, Nolan GP. Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry A. 2012;81:552-66.
5. van Lochem EG, van der Velden VHJ, Wind HK, te Marvelde JG, Westerdaal NAC, van Dongen JJM. Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: Reference patterns for age-related changes and disease-induced shifts. Cytometry Part B: Clinical Cytometry. 2004;60B:1-13.
6. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotech. 2011;29:886-91.
7. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol. 2013;31:545-52.