nature
biotechnology

viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia

El-ad David Amir[1], Kara L Davis[2,3], Michelle D Tadmor[1,3], Erin F Simonds[2,3], Jacob H Levine[1,3], Sean C Bendall[2,3], Daniel K Shenfeld[1,3], Smita Krishnaswamy[1], Garry P Nolan[2,4] & Dana Pe'er[1,4]

New high-dimensional, single-cell technologies offer unprecedented resolution in the analysis of heterogeneous tissues. However, because these technologies can measure dozens of parameters simultaneously in individual cells, data interpretation can be challenging. Here we present viSNE, a tool that allows one to map high-dimensional cytometry data onto two dimensions, yet conserve the high-dimensional structure of the data. viSNE plots individual cells in a visual similar to a scatter plot, while using all pairwise distances in high dimension to determine each cell's location in the plot. We integrated mass cytometry with viSNE to map healthy and cancerous bone marrow samples. Healthy bone marrow automatically maps into a consistent shape, whereas leukemia samples map into malformed shapes that are distinct from healthy bone marrow and from each other. We also use viSNE and mass cytometry to compare leukemia diagnosis and relapse samples, and to identify a rare leukemia population reminiscent of minimal residual disease. viSNE can be applied to any multi-dimensional single-cell technology.

Emerging single-cell technologies have revealed an extensive degree of heterogeneity between and within tissues[1]. Analysis of single-cell data has shed light on many different cellular processes[2–7] and recent technological advances have enabled the study of a large number of parameters in single cells at unparalleled resolution. For example, mass cytometry[8] can measure up to 45 parameters simultaneously in tens of thousands of individual cells. High-resolution microscopy[9,10] and single-cell RNA quantification[11–14] allow analysis of 100 parameters in dozens and soon hundreds of individual cells. These innovations promise to transform the way we think about development, differentiation, and disease[1,15,16].

However, it is difficult to visualize such high numbers of dimensions in a meaningful manner. Single-cell data are often examined in two dimensions at a time in the form of a scatter plot[17]. Yet, as the number of parameters increases, the number of pairs becomes overwhelming. A typical mass cytometry data set allows several hundred pairwise combinations. In addition, a pairwise viewpoint could miss biologically meaningful multivariate relationships that cannot be discerned in two dimensions. Several computational tools, such as SPADE[18], have been developed to address these problems[19,20]. However, these approaches typically cluster cells and examine the average of each cluster, resulting in the loss of single-cell resolution of the data. Principal component analysis (PCA), another computational tool, has been applied to mass cytometry data sets[21] and can be used to project data into two dimensions while maintaining single-cell resolution. However, PCA is a linear transformation that cannot faithfully capture the nonlinear relationships that are a hallmark of many

single-cell data sets. Therefore, we need new tools to visualize and interpret high-dimensional single-cell data such as those produced by mass cytometry. An ideal tool would enable visualization at single-cell resolution, preserve the geometry and nonlinearity of the data, represent both abundant and rare populations, and provide a robust, interpretable view of the data.

We developed viSNE for this purpose. viSNE allows visualization of high-dimensional single-cell data and is based on the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm[22]. viSNE finds the two dimensional representation of single-cell data that best preserves their local and global geometry. The resulting viSNE map provides a visual representation of the single-cell data that is similar to a biaxial plot, but the positions of cells reflect their proximity in high-dimensional rather than two-dimensional space. We utilize color as a third dimension to interactively visualize features of these cells. Here we apply viSNE to interpret mass cytometry data derived from healthy and leukemic human bone marrow.

RESULTS

Preserving high-dimensional relationships in single-cell data

In viSNE, each cell is represented as a point in high-dimensional space. Each dimension is one parameter (that is, the expression level of one protein). An optimization algorithm searches for a projection of the points from the high-dimensional space into two or three dimensions such that pairwise distances between the points are best conserved between the high- and low-dimensional space (Online Methods). The resulting low-dimensional projection, which we call the viSNE map,

Figure 1 viSNE map of healthy human bone marrow. (**a**) In a synthetic toy example, (1,000 points randomly distributed with normally distributed noise around a polynomial of the third degree) viSNE projects a one-dimensional curve embedded in three dimensions (left) onto two dimensions (right). The color gradient shows that points that are in close proximity in three dimensions remain in close proximity in two dimensions. (**b**) Application of viSNE to a healthy human bone marrow sample, stained with 13 markers and measured with mass cytometry[21], automatically separates cells into spatially distinct subsets based on the combination of markers that they express. Each point in the viSNE map represents an individual cell and its color represents its immune cell subset as designated by independent expert manual gating (manual gates are defined at the bottom). Gray points were not classified by manual gating. The axes are in arbitrary units. (**c**) Biaxial plots represent the same data shown in **b**, and show the gates drawn manually by expert operators. The colors of the squares match the colors in **b**. The actual manual gating used here is more complex and uses a series of biaxial plots to gate each population[21]. Note, that unlike in **b**, no single biaxial plot spatially separates all subsets. (**d**) The same viSNE map shown in **b** is colored according to intensity of CD11b expression. Many of the cells within the dotted line gate were not classified as monocytes by manual gating (gray cells, **b**). (See **Supplementary Fig. 1** for additional analysis of marker expression on these cells.)
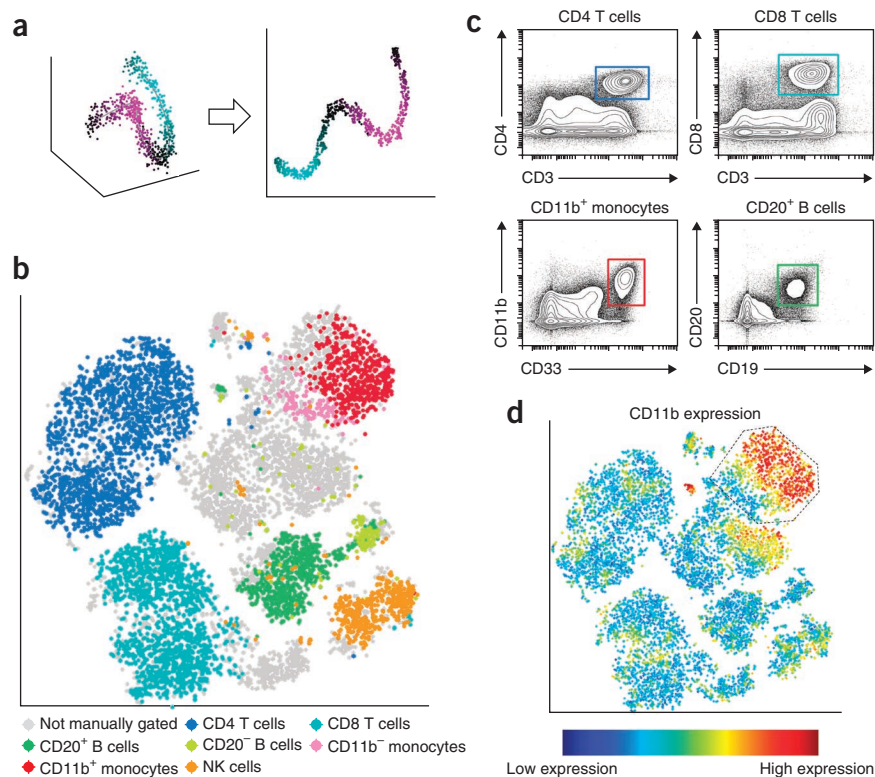
is visualized as a scatter plot, where a cell's location in the plot represents information from all of the original dimensions.

We also developed *cyt*, an interactive tool for visualization of viSNE maps. *cyt* has multiple features, including plotting the maps, coloring cells by marker expression, sample or subtype, and gating. **Figure 1a** illustrates how viSNE works on a synthetic example; the optimization algorithm identifies the global structure of the data (a one-dimensional line, along with its curvature, embedded in three-dimensional space) and the local structure (pairwise distances between points along the line are conserved).

viSNE map of healthy human bone marrow

First, we evaluated viSNE's ability to map the well-characterized system of human bone-marrow hematopoiesis[23]. We analyzed previously generated data of healthy human bone marrow (sample Marrow1) stained with elemental isotope-conjugated antibodies specific for 13 surface markers[21]. When applied to this data set, viSNE generated a map that clearly separated different cell subsets in space (**Fig. 1b**). To validate and label the map, we used an independently derived classification of the cells based on expert manual gating of a series of biaxial plots (**Fig. 1c** and Online Methods). Although viSNE was not provided with this classification or with any knowledge of immune subsets, it successfully grouped cells of the same subset together and cells of different subsets separately (**Fig. 1b–c**). viSNE accurately distinguished CD4+ and CD8+ T cells, mature and immature B cells, mature and immature monocytes and natural killer (NK) cells. Notably, NK cells formed a distinct subset even though CD56, the canonical marker associated with this lineage, was not included in the antibody staining panel.

To further compare the expert manual gating and the viSNE map, we used the *cyt* feature to gate subsets directly from the viSNE map

(**Supplementary Fig. 1a**). In all cases, the viSNE gate included cells that were not classified by the expert manually gated biaxial plots; these cells are labeled in gray in the viSNE map. Examination of the marker expression of these cells reveals that they are typically just beyond the threshold of one marker, but the viSNE classification is strongly supported based on the expression of all other markers. For example, in **Figure 1d**, wherein cells are colored for CD11b marker expression, the cells in the gated region express the canonical monocyte marker CD33 (**Supplementary Fig. 1b**). However, only 47% of these cells were classified as monocytes by the manual gating (**Fig. 1b**). In addition, the marker intensity distributions between the CD11b– monocytes in the viSNE map monocyte gate and in the manually set monocyte gate (**Supplementary Fig. 1c**) are similar, supporting the notion that the cells gated in the viSNE map are indeed CD11b– monocytes.

Traditional gating relies on hard thresholds to classify cells into subsets. Thus cells whose marker values are slightly below or above the threshold might not be classified correctly, or classified at all (**Supplementary Fig. 1d**). When dealing with the hematopoietic continuum, this may result in the inability to accurately capture transitional cell types. For example, using *cyt* to color cells based on marker intensity revealed that viSNE organized monocytes based on a gradient or smooth increase in expression of CD11b, a marker of monocyte maturity (**Fig. 1d**). This finding highlights the continuous and gradual nature of CD11b expression during monocyte maturation and better represents the continuum of normal differentiation[24]. viSNE takes into account all phenotypic markers concurrently instead of relying on hard thresholds and, as a result, classifies more cells and captures a more accurate view of the variability within each subset when compared to biaxial gating. The single-cell resolution of the viSNE map provides fine detail of each subset, going beyond clustering and enabling investigation of the variation, structure and transitions within each subset.
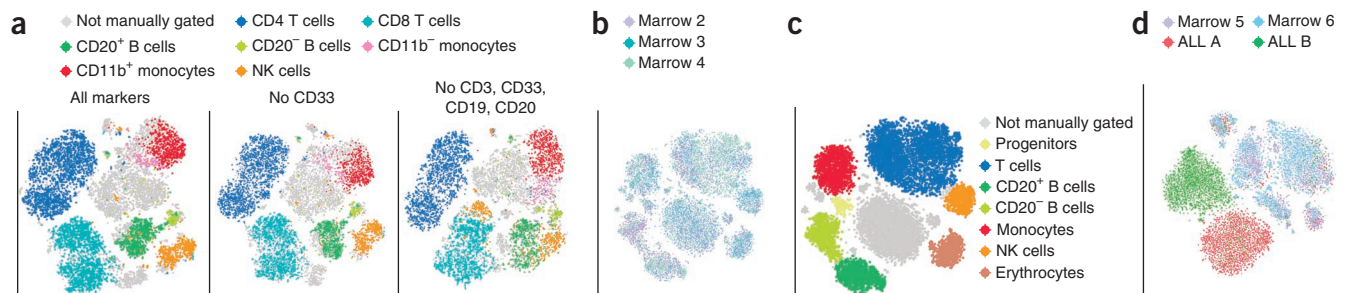
Figure 2 viSNE is robust, consistent, and does not require canonical markers. (**a**) The left map is the same as in **Figure 1b**, and was generated by considering all 13 markers. Middle: viSNE map of the same cells, projected after removing CD33. Right: viSNE map of same cells, projected after removing CD33, CD3, CD19 and CD20. Despite removing four canonical markers, viSNE separates most major subtypes using the remaining nine channels. (See **Supplementary Fig. 3** for analysis after removal of additional markers.) (**b**) Bone marrow samples from three healthy donors (Marrow2–4) were mapped using viSNE. Each point represents a single cell, and different colors represent different samples. (See **Supplementary Fig. 5** for a separate plot for each sample.) (**c**) The map is the same map as in **b**, but is color coded by cell subset as identified by analyzing expression of individual markers as shown in **Supplementary Figure 6**. Subsets are indicated next to the map. (**d**) Bone marrow samples from two healthy donors (Marrow5–6) and two ALL patients (ALL A, B) were mapped using viSNE. Samples are color-coded as indicated in key.

Robust subset classification even without canonical markers

We performed a number of analyses to evaluate the robustness of viSNE. The viSNE map in **Figure 1b** includes 10,000 cells that were subsampled from the complete data set of Marrow1 (ref. 21). We independently subsampled multiple subsets of the data and ran viSNE on each. Reassuringly, these separate analyses resulted in similar viSNE maps that conserved the spatial separation between subsets (**Supplementary Fig. 2**). Thus the viSNE map consistently and reliably represents real structure in the data.

To test viSNE's reliance on specific markers for classification of immune cell subsets, we generated multiple viSNE maps but excluded some of the markers when generating each map. The viSNE map remained consistent in terms of spatial separation of subsets even after removal of any single marker (**Fig. 2a** and **Supplementary Fig. 3**). Remarkably, even after excluding the canonical markers of B cells, T cells and myeloid cells (CD19/CD20, CD3 and CD33, respectively), the viSNE map remained consistent with the map constructed using all 13 markers (**Fig. 2a**). These findings imply that noncanonical markers, when analyzed together, contain the information needed to separate distinct immune cell subsets. This speaks to a previously unappreciated level of organization where specialized immune subtypes have tightly coordinated surface marker expression beyond their canonical identifiers. The different subtypes reside in distinct well-separated subset shapes in high-dimensional space (**Supplementary Fig. 4** and **Supplementary Data 1**).

Consistent and reproducible healthy bone marrow map

Having demonstrated viSNE's robustness when applied to a single healthy bone marrow sample, we examined its robustness across bone marrow samples from multiple healthy individuals. Three healthy bone marrow samples (Marrow2–4) were assayed by mass cytometry using a panel of 31 phenotypic surface markers. The resulting viSNE map grouped cells into distinct subpopulations, and cells from all three individuals overlapped within each subpopulation (**Fig. 2b** and **Supplementary Fig. 5**). We used the Jenson-Shannon (JS) divergence to quantify the similarity between the viSNE maps of the three individuals. The JS divergence between each pair of healthy individuals was 0.04, confirming that there is almost no divergence between the viSNE maps of healthy samples. Using *cyt* to visualize expression of individual markers, we gated specific immune cell subsets in viSNE maps (**Fig. 2c** and **Supplementary Fig. 6**). The 31 markers used led to the identification of slightly different immune subsets, for example, recognition of erythrocyte and progenitor subsets was aided by the addition of anti-CD61 and anti-CD117

to the panel, whereas the CD4 and CD8 T-cell populations were merged because those markers were omitted (**Supplementary Fig. 7**).

To further evaluate the robustness of viSNE's map of healthy bone marrow, we applied viSNE to an additional bone marrow sample collected using conventional fluorescence-based flow cytometry. The resulting viSNE map is similar to the map generated by mass cytometry (**Supplementary Fig. 8**), demonstrating not only consistency in the map between healthy samples, but also that viSNE is well-suited for the analysis of fluorescence-based cytometry data. The cellular subtypes that make up the human immune system are reproducibly represented by viSNE, and the fidelity of this structure is maintained across multiple cytometry platforms, marker panels, and, most importantly, individuals.

Deformed shapes of leukemic bone marrow maps

Encouraged by the consistency and robustness of viSNE maps of healthy bone marrow samples, we used viSNE to analyze leukemic bone marrow. We stained two bone marrow samples donated from healthy individuals and two from acute lymphoblastic leukemia (ALL) patients with a panel of 29 antibodies optimized for the analysis of ALL (**Supplementary Table 1**).

The maps of the two healthy bone marrow samples (Marrow5–6) overlap (JS divergence 0.04) (**Fig. 2d**). In contrast, the two ALL samples occupy a completely separate region within the viSNE map (JS divergence 0.45), and each forms a distinct population separate from the other ALL sample (JS divergence 0.42). Some cells from the ALL samples (~5%) overlap with cells from the healthy samples. Inspection of these cells revealed marker combinations that correspond to healthy immune cells, supporting their placement with the other healthy cells.

When we applied viSNE separately to each ALL sample, each sample mapped into a large deformed shape (**Fig. 3a**) and several smaller shapes; the latter corresponded to healthy immune cell populations, indicating that the malignant cells were related to each other but sufficiently distinct from healthy cells. We also applied viSNE to two acute myeloid leukemia (AML) patient samples. The viSNE maps of these AML samples also displayed a single, large, deformed shape (**Fig. 3b**), in contrast to the separated and distinct subpopulations of healthy samples. Even when constructing a three-dimensional viSNE map, the AML projects onto a single continuous shape (**Supplementary Data 2–4**). We noted a considerable population structure within each cancer, as discerned by multiple peaks and saddle points in the contour map. Moreover, each cancer sample formed a unique viSNE map, in which healthy subpopulations were consistently separated from the abnormal leukemic subpopulations.
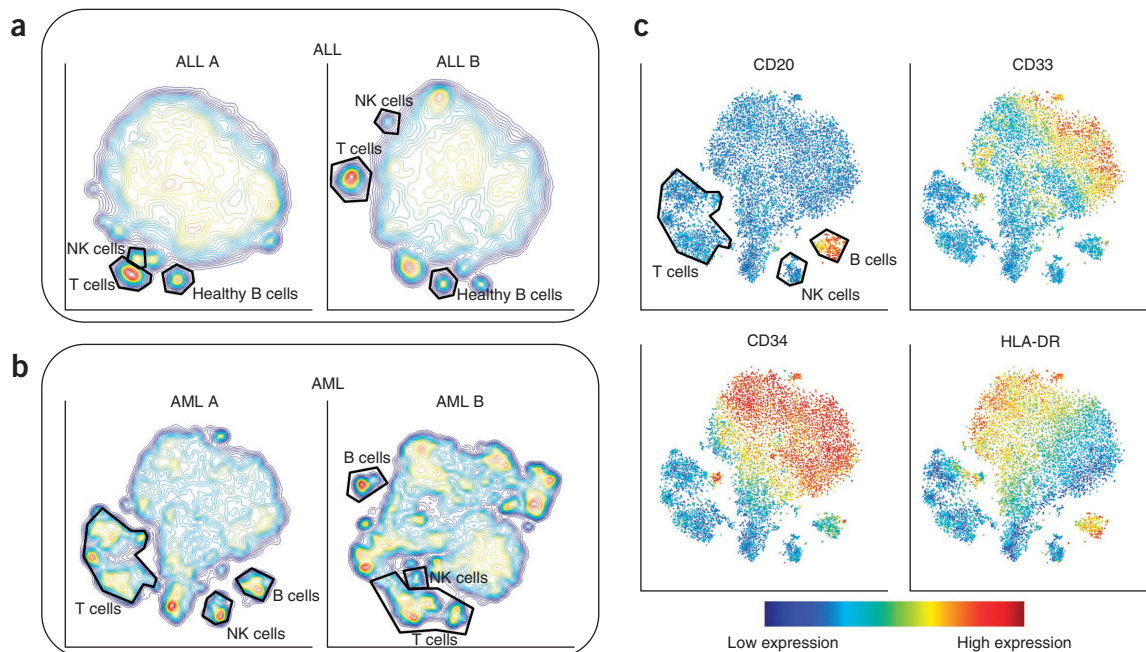
Figure 3 Cancer samples form contiguous but heterogeneous shapes. (**a**,**b**) Contour plots of the viSNE maps of two different ALL (**a**) and AML (**b**) samples. The contours represent cell density in each region of the map. Small gated populations represent healthy immune subtypes, as indicated, as revealed by examination of their marker expression. Because the structure of each tumor dramatically changes between samples, viSNE places the healthy regions in different locations in each map (as each healthy subtype is positioned as close as possible to the most similar cancer cells). (**c**) viSNE map of a diagnosis bone marrow sample from AML patient 1. Cells are colored according to intensity of expression of the indicated markers. CD20 helps identify the healthy B-cell subpopulation.

viSNE can explore cancer heterogeneity

Although healthy samples can be studied by biaxial gating based on known surface phenotypes of individual immune cell subsets, exploring cancer heterogeneity in high dimensions can be a daunting task as cancer samples frequently display abnormal combinations of surface markers and there are hundreds of possible biaxial combinations of surface markers. In current clinical practice, hematopoietic malignancies are analyzed using at most four to eight markers simultaneously. Hematopoietic malignancy immunophenotyping results have typically been displayed using biaxial plots focused on key markers (**Supplementary Fig. 9**). However, by combining mass cytometry with viSNE, we are able to visualize cancer at single-cell resolution in a single map that takes into account ~30 markers; this sort of analysis can reveal additional structure, abnormal marker combinations and subpopulations.

We used viSNE to comprehensively characterize a diagnostic AML bone marrow sample. Although the overall viSNE map shape of cancer is deformed compared to that of healthy bone marrow, some markers (e.g., CD33, CD34 and HLA-DR) show gradients of expression whereas others (e.g., CD79b) show clustered expression (**Fig. 3c** and **Supplementary Fig. 10**). Within the subpopulation of cells that highly express CD34 (a marker of stem/progenitor cells) is a gradient of expression of CD33 (a marker of monocytes; **Fig. 3c**). This marker combination suggests a derailed development program in cancer, because during normal healthy immune cell development, as monocytes mature, expression of CD34 (a marker of immaturity) decreases. Perhaps in this cancer, oncogene activity promoted a progenitor-like CD34+ state, but the cells continued to differentiate aberrantly as indicated by the induction of CD33 expression. The single-cell resolution of viSNE highlights cancer as a continuum of heterogeneous phenotype states, demarcated by gradients of marker expression rather than distinct subpopulations.

Comparing diagnosis and relapse samples

Because the viSNE map might reflect aspects of cancer progression, we used viSNE to analyze two samples from a single AML patient: one sample was taken before chemotherapy and the other was taken after disease relapse. Using viSNE on a merged data set representing both samples, we could clearly visualize a separation between the diagnosis and relapse samples (**Fig. 4a**, **Supplementary Figs. 11** and **12**). viSNE reveals phenotypes unique to the diagnosis sample, which are presumably eliminated by chemotherapy, as well as phenotypes that arise only at relapse. Notably, the viSNE map identifies a region of phenotypes occupied by both samples, but that is considerably rarer at diagnosis. This may suggest enrichment of a drug-resistant clone that maintained a consistent phenotype from diagnosis to relapse. We also note populations of healthy cells that overlap in the diagnosis and relapse sample viSNE maps; these provide an internal technical control for the similarity of staining between samples. Regarding specific markers, FLT3 expression is pervasive in the diagnosis sample, but diminished in the relapse sample. Genetic analysis of the diagnosis sample revealed an internal-tandem duplication of FLT3, a common mutation in AML[25], suggesting relapse derived from a clone lacking this mutation (FLT3 genetic status at relapse was unavailable). The clone that reemerged at relapse had an altogether different and more immature phenotype, with cells expressing both high CD34 and CD33 throughout a large fraction of the sample (**Fig. 4b** and **Supplementary Fig. 12**). The relapse sample was highly heterogeneous, as distinct regions expressed different markers from the myeloid lineage (CD64 and CD15) and lymphoid lineage (CD7) (**Fig. 4b**).

To allow further dissection of the heterogeneity of the AML sample using experimental tools such as DNA and RNA sequencing, we used the viSNE map to devise a gating scheme that is compatible with fluorescence-assisted cell sorting (FACS). We divided the AML sample

Figure 4 viSNE reveals the progression of cancer from diagnosis to relapse. (**a**) Contour plots of the viSNE maps of diagnosis and relapse AML samples in patient AML B. The contours represent cell density in each region in the map. The map is the same in each sample. Each point represents a cell from the diagnosis (top, purple) or relapse (bottom, red) sample. (**b**) Cells from both diagnosis and relapse samples are shown in each map, and the map is the same as in **a**. Cells are colored according to intensity of expression of the indicated markers, enabling the comparison of expression patterns before and after relapse. For example, Flt3 is expressed primarily in the diagnosis sample. CD34 emerges in the relapse sample, as do CD64 and CD7. There is a CD33 gradient in both samples. The overlapping region has cells that express high levels of CD49d.
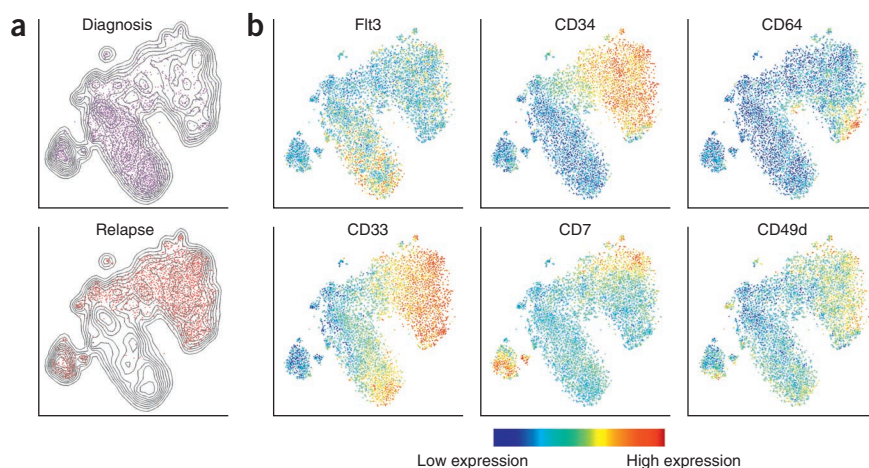
into subpopulations based on expression of CD33, CD34, CD7 and CD64. We classified each marker as "on" (positive) or "off" (negative) according to a threshold that was chosen using the map (**Fig. 5a**, black lines). We intersected CD34 and CD33 gates, selected the CD34+CD34+ population, and applied the CD64 and CD7 gates to it (**Fig. 5b**). When examining the intersection of all eight groups (two for each marker), we identified six distinct subpopulations having at least 20 cells each that spread across the viSNE map (**Fig. 5c**). The next step would be to physically separate the relapse sample into these subpopulations using FACS and characterize them by means of downstream experiments.

viSNE detects minimal residual disease

The ability to detect, by flow cytometry, small numbers of cancerous cells displaying an aberrant phenotypic "fingerprint" is used to stratify patients by risk of relapse and direct treatment decisions. The presence of such minimal residual disease (MRD) can be associated with risk of relapse[26,27]. The detection of MRD indicates a need for intensified therapy that unfortunately carries an increased risk of toxicity. Consequently, accurate detection of rare malignant populations is paramount in correctly assigning risk to an individual patient.

There are two competing manual methods for assessing MRD by flow cytometry. The first involves identifying aberrant antigen expression (leukemia-associated immunophenotype) on the leukemia cells at diagnosis, and then looking for that same phenotype on samples taken after chemotherapy[28]. The second method involves identifying leukemic cells based on a 'different-from-normal' phenotype by comparing them to a historical bank of healthy bone marrow samples[29]. Although the prognostic value of MRD measurement has been validated in several
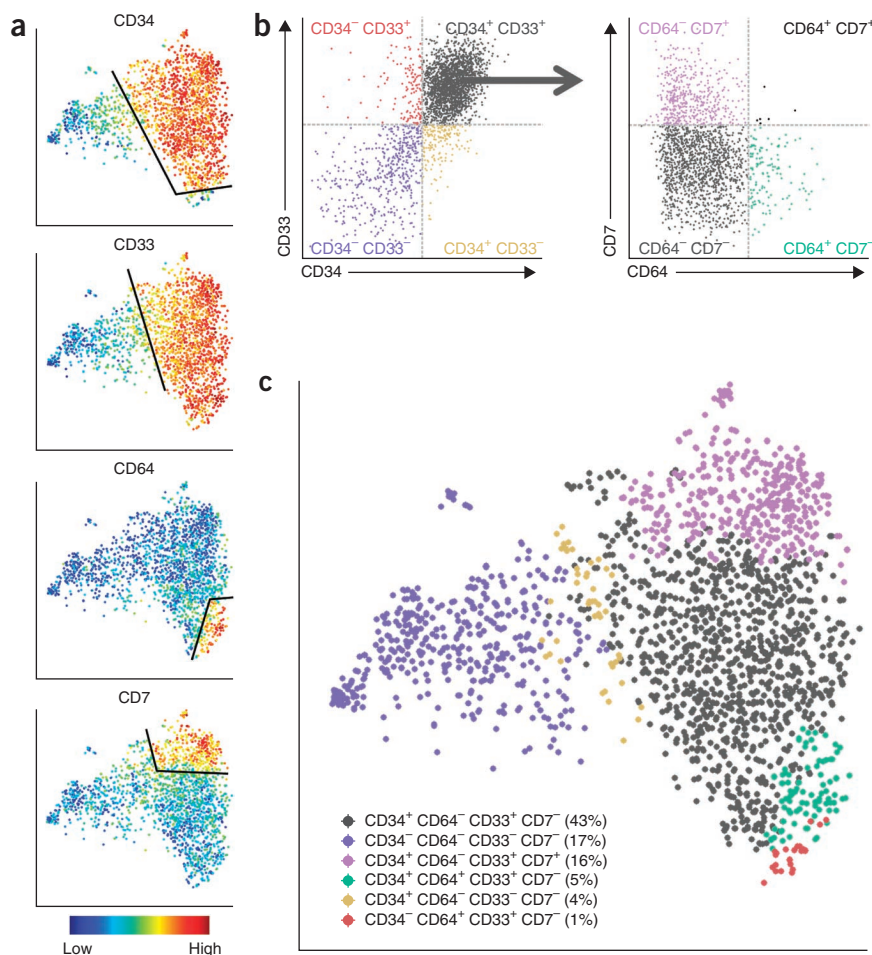
Figure 5 A gating scheme for FACS of an AML relapse sample in patient B based on the viSNE map. (**a**) The viSNE map, colored by intensity of expression of (from top to bottom) CD34, CD33, CD64 and CD7. For each marker, cells were separated into two subpopulations: "on" (positive) and "off" (negative), based on an expression threshold (black lines). (**b**) Left: biaxial plot of CD34 versus CD33. Right: biaxial plot of CD64 versus CD7 applied to the CD34+ CD33+ subpopulation from the upper right quadrant of the left plot. In all cases, cells are colored and labeled by the quadrants. (**c**) Six subpopulations (the only six populations having more than 20 cells each), revealed by comparisons such as that in **b**, were projected onto the viSNE map. Cells are colored by their respective subpopulation from **b**. The relapse sample can now be sorted into these subpopulations by FACS and further studied through downstream experiments such as DNA and RNA sequencing.
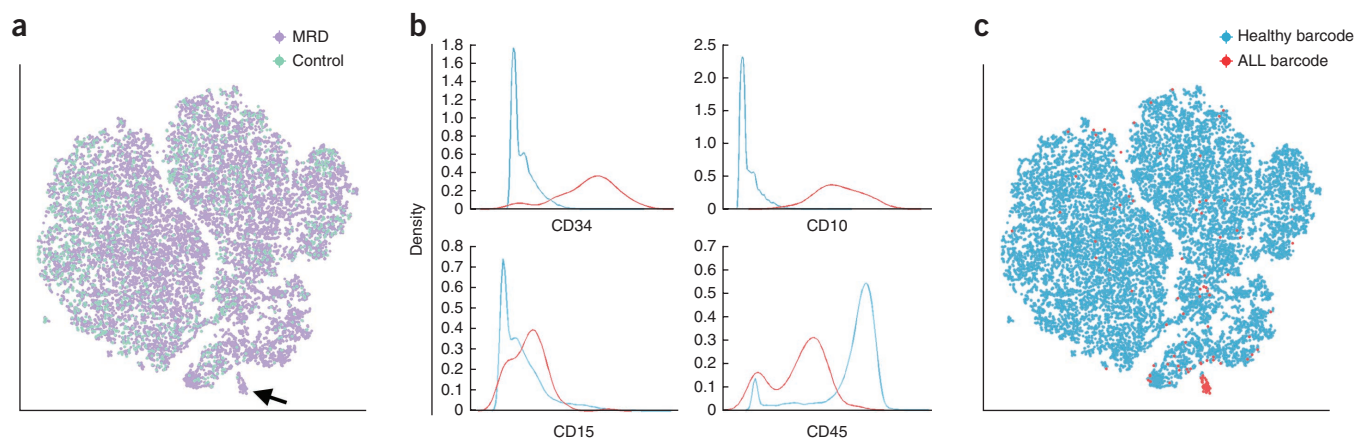
Figure 6 Using viSNE to identify synthetic MRD. (**a**) A synthetic MRD sample was created by spiking a healthy bone marrow sample with metal-barcoded ALL cells. This synthetic MRD sample was compared to an unmanipulated healthy bone marrow sample (the viSNE algorithm was blinded to the metal-barcode channel). The viSNE map of the synthetic MRD sample (purple) and a healthy control sample (cyan) includes a suspect region (marked by an arrow) composed almost entirely of cells from the synthetic MRD sample (purple). (**b**) Expression of indicated markers on cells in the suspect region (red) and the nonsuspect region (cyan). *X*-axis represents marker expression level and *y*-axis represents density of cells. (**c**) The viSNE map from **a**, after removing the blinding of the metal-barcoding channel. The map is colored according to expression of the metal barcode in the spiked-in ALL cells (red). The suspect region is indeed composed almost entirely of ALL cells. ALL cells outside of the suspect region have marker expression levels conforming to healthy cells. (See **Supplementary Fig. 13** for an additional example and **Supplementary Fig. 14** for a control example.)

clinical trials, both of these approaches require an expert pathologist to manually inspect biaxial plots, and both approaches have shortcomings. It can be difficult to identify abnormal cells that are sufficiently phenotypically distinct from normal bone marrow. If one relies on and searches only for cells with the phenotype of the diagnostic sample, one may fail to detect other malignant populations displaying distinct yet abnormal phenotypes. Thus, a tool that automatically identifies abnormal cellular phenotypes would allow clearer identification and evaluation of remaining cancer cells.

Because viSNE revealed such a clear contrast between leukemic and healthy bone marrow, we tested whether it could aid manual MRD detection. We spiked metal-barcoded[30] cells from an ALL patient sample into a healthy bone marrow sample, thereby creating a synthetic sample with 0.25% of a MRD-like population. A single healthy bone marrow sample served as a guide for interpretation (similar to the different-from-normal manual approach). We used a biased subsampling method to enrich for unique noncontrol subpopulations and generated a viSNE map using eight markers (CD3, CD7, CD10, CD15, CD20, CD34, CD38, CD45) to emulate an MRD scenario using fluorescence-based flow cytometry. The algorithm was blinded to the metal-barcoded channel.

A good MRD candidate region would be a distinct region of the viSNE map that contains cells from the MRD sample, but no cells from the healthy control sample. Cells from both samples were well-mixed across most of the viSNE map, except for one suspect region that was composed almost entirely of cells from the synthetic MRD sample (**Fig. 6a**). We compared marker expression levels in the suspect region to the rest of the sample (**Fig. 6b**) and found that the suspect cells strongly expressed CD10 and CD34, exhibited below-average expression of CD45 and also expressed CD15, a phenotypic combination often seen in precursor B-cell ALL. Taken together, the combination of these surface markers and the absence of similar cells in the healthy control suggest that these were leukemic cells. Removal of the blinding of the metal-barcoding channel revealed that these cells were positive for the metal barcode and therefore were indeed the spiked ALL cells (**Fig. 6c**). We repeated this analysis with a different set of markers and achieved similar results (**Supplementary Fig. 13**). As a control, we repeated the same procedure with a sample that included

only healthy cells. The two healthy samples were well mixed across the entire viSNE map; there was no region that contained only MRD cells (**Supplementary Fig. 14**), demonstrating that the subsampling method and viSNE do not spuriously create suspect regions. Although only a synthetic example, this demonstrates viSNE's success in identifying a minuscule cancer subpopulation in the data, suggesting that viSNE can be effectively used for MRD detection.

DISCUSSION

viSNE allows visualization of high-dimensional single-cell data on a two-dimensional map. This mapping takes advantage of the inherent structure of the data; for example, different immune subsets reside in separate regions in high-dimensional space. Conventional analysis of cytometry data, which views only two dimensions at a time, ignores the higher-order structure and complex relationships between markers in the data. Whereas viSNE plots resemble conventional biaxial plots, their utility comes from combining and representing information from all dimensions simultaneously.

We found that the viSNE map is consistent across multiple healthy individuals, whereas cancer samples occupy regions distinct from healthy cells and from each other. We illustrated how viSNE can be used to characterize heterogeneity within cancer samples, mark disease progression from diagnosis to relapse, and identify rare cancer populations lurking among predominantly healthy cells.

viSNE has a number of advantageous features for the analysis of single-cell data. It is sensitive to small subsets (**Supplementary Fig. 15**), which allows us to subsample uniformly, preserving the original frequencies of cell populations. It takes into account similarities between all pairs of cells, providing information about both nearby and distant cells. And it plots each individual cell, showing the diversity within populations and maintaining single-cell resolution. This single-cell resolution is one feature that distinguishes viSNE from SPADE[18], and may help explain why viSNE appears to be more robust than SPADE (**Supplementary Fig. 16**). Single-cell resolution may also be critical for viSNE's success in the MRD scenario. For example, when we applied SPADE to the same synthetic MRD sample, the ALL cells were indistinguishable from healthy cells in the

resulting SPADE tree (**Supplementary Fig. 17**), rendering SPADE inappropriate for MRD.

viSNE belongs to the class of nonlinear dimensionality reduction (NLDR) algorithms which, unlike PCA, do not assume linear relationships between parameters. Immune subsets are nonlinear and hence PCA, unlike viSNE, fails to distinguish between them (**Supplementary Figs. 18** and **19**). We evaluated several other NLDR algorithms[31] and among these only viSNE resulted in consistent maps that separated immune subtypes across multiple subsamples of the same data (**Supplementary Fig. 19**). The other methods might be confounded by the noise inherent in biological systems and measurement technologies or by the complex geometry in high-dimensional hematopoietic space.

Despite its extensive utility, as with all dimensionality-reduction tools, viSNE is inherently limited: low-dimensional mapping cannot represent all of the information in high-dimensional space. Therefore, viSNE captures only the most dominant structures. One way to gain more detail is to run viSNE on a well-defined subset of the data. For example, instead of analyzing several cancer samples together (**Fig. 2d**), one can run viSNE on each sample separately (**Fig. 3a,b**). Alternatively, it is possible to limit the mapping to only a subset of the parameters of interest. Another consequence of dimensionality reduction is the 'crowding problem', which typically limits the number of cells we can map to 30,000. This limits applications such as gating, because a subsampling of the cells is required, meaning only cells in the subsample can be classified. An effective solution for gating is combining viSNE with a clustering algorithm. We clustered the cells using FLOCK[19], a state-of-the-art clustering tool for cytometry data, and labeled the viSNE map according to this clustering (**Supplementary Fig. 20**). Although FLOCK separated the immune subtypes, it split each subtype into multiple clusters. The viSNE map helps interpret these clusters and their relationships.

viSNE is an unsupervised algorithm and does not require prior knowledge of the system. It is thus suitable for navigating less explored systems such as cancer. Although structure in healthy samples is formed through an orderly program of development, cancer's derailed developmental program leads to loss of normal order and structure. viSNE helps characterize the plethora of abnormal phenotypes unique to each cancer by exploiting its ability to take all markers into account simultaneously, rather than scanning through hundreds of biaxial plots, two markers at a time.

A characteristic feature that repeated across multiple cancer maps was the emergence of distinct gradients of marker expression levels that resemble developmental progression in healthy cells (**Fig. 1d**). Comparing gradients in AML diagnosis and relapse samples (**Fig. 4b**) supports the notion that the cells first gain CD34 and subsequently acquire highly diverse abnormal combinations of lineage-specific markers without attenuation of CD34. After identifying unexpected cancer populations using viSNE, one can design a sorting strategy for physical isolation and downstream characterization of these populations.

In the future, we expect viSNE to be instrumental in the analysis of mass cytometry data integrating the surface marker panel with a panel of functional markers that probe signaling, cell cycle and metabolism, under many experimental perturbations (such as cytokines and drugs)[3]. In this scenario, viSNE's ability to distinguish rare subsets that make up only a tiny fraction of the population (**Fig. 6** and **Supplementary Fig. 15**) could be advantageous in the identification and characterization of rare drug-resistant subpopulations.

We demonstrated viSNE's capability to analyze mass cytometry and flow cytometry data. Biological research is trending toward dozens of dimensions in tens of thousands of cells. Making sense of these data is a daunting challenge that requires powerful computational

approaches. The utility of viSNE will increase with the number of dimensions capable of being analyzed by mass cytometry and other technologies.

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

Note: Supplementary information is available in the online version of the paper.

AUTHOR CONTRIBUTIONS
E.D.A., G.P.N. and D.P. conceived the study. E.D.A. and D.P. developed the methods. D.K.S. and M.D.T. implemented parallel t-SNE and *cyt*, respectively. E.F.S., S.C.B., K.L.D. and G.P.N. designed and performed mass and flow cytometry experiments. E.D.A., J.H.L., E.F.S., S.C.B., K.L.D., S.K. and D.P. performed the biological analysis and interpretation. E.D.A. and M.D.T. performed robustness analysis of the method. E.D.A., J.H.L., K.L.D., E.F.S. and D.P. wrote the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Bendall, S.C. *et al.* A deep profiler's guide to cytometry. *Trends Immunol.* **33**, 323–332 (2012).
2. Petilla Interneuron Nomenclature Group. *et al.* Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* **9**, 557–568 (2008).
3. Irish, J.M. *et al.* Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217–228 (2004).
4. Sachs, K. *et al.* Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308**, 523–529 (2005).
5. Majeti, R., Park, C.Y. & Weissman, I.L. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. *Cell Stem Cell* **1**, 635–645 (2007).
6. Tarnok, A., Ulrich, H. & Bocsi, J. Phenotypes of stem cells from diverse origin. *Cytometry A* **77**, 6–10 (2010).
7. O'Brien, C.A., Kreso, A. & Dick, J.E. Cancer stem cells in solid tumors: an overview. *Semin. Radiat. Oncol.* **19**, 71–77 (2009).
8. Bandura, D.R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal. Chem.* **81**, 6813–6822 (2009).
9. Cornett, D.S. *et al.* MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat. Methods* **4**, 828–833 (2007).
10. Mercer, J. *et al.* RNAi screening reveals proteasome- and Cullin3-dependent stages in vaccinia virus infection. *Cell Rep.* **2**, 1036–1047 (2012).
11. Ozsolak, F. & Milos, P.M. RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* **12**, 87–98 (2011).
12. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
13. Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
14. Lubeck, E. & Cai, L. Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat. Methods* **9**, 743–748 (2012).

15. Benoist, C. & Hacohen, N. Immunology. Flow cytometry, amped up. *Science* **332**, 677–678 (2011).
16. Bendall, S.C. & Nolan, G.P. From single cells to deep phenotypes in cancer. *Nat. Biotechnol.* **30**, 639–647 (2012).
17. Herzenberg, L.A. *et al.* Interpreting flow cytometry data: a guide for the perplexed. *Nat. Immunol.* **7**, 681–685 (2006).
18. Qiu, P. *et al.* Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat. Biotechnol.* **29**, 886–891 (2011).
19. Qian, Y. *et al.* Elucidation of seventeen human peripheral blood B-cell subsets and quantification of the tetanus response using a density-based method for the automated identification of cell populations in multidimensional flow cytometry data. *Cytometry B Clin. Cytom.* **78** (suppl. 1), S69–S82 (2010).
20. Pyne, S. *et al.* Automated high-dimensional flow cytometric data analysis. *Proc. Natl. Acad. Sci. USA* **106**, 8519–8524 (2009).
21. Bendall, S.C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
22. Van der Maaten, L. & Hinton, G. *Visualizing* (2579–2605): *data using t-SNE*. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
23. Kondo, M. *et al.* Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annu. Rev. Immunol.* **21**, 759–806 (2003).
24. van Lochem, E.G. *et al.* Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow: reference patterns for age-related changes and disease-induced shifts. *Cytometry B Clin. Cytom.* **60**, 1–13 (2004).
25. Wakita, S. *et al.* Mutations of the epigenetics modifying gene (DNMT3a, TET2, IDH1/2) at diagnosis may induce FLT3-ITD at relapse in de novo acute myeloid leukemia. *Leukemia* published online doi: 10.1038/leu.2012.317 (8 November 2012).
26. Campana, D. Status of minimal residual disease testing in childhood haematological malignancies. *Br. J. Haematol.* **143**, 481–489 (2008).
27. Borowitz, M.J. *et al.* Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia and its relationship to other prognostic factors: a Children's Oncology Group study. *Blood* **111**, 5477–5485 (2008).
28. Ossenkoppele, G.J., van de Loosdrecht, A.A. & Schuurhuis, G.J. Review of the relevance of aberrant antigen expression by flow cytometry in myeloid neoplasms. *Br. J. Haematol.* **153**, 421–436 (2011).
29. Loken, M.R. *et al.* Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. *Blood* **120**, 1581–1588 (2012).
30. Bodenmiller, B. *et al.* Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867 (2012).
31. Van der Maaten, L., Postma, E. & Van Den Herik, H. Dimensionality reduction: A comparative review. *J. Mach. Learn. Res.* **10**, 1–41 (2009).

ONLINE METHODS

The t-SNE algorithm. The t-SNE algorithm maps points from high-dimensional space to low-dimensional space by minimizing the difference in all pairwise similarities between points in high- and low-dimensional spaces[22]. The axes of the low-dimensional spaces are given in arbitrary units. The algorithm proceeds as follows. First, the pairwise distance matrix is calculated in high-dimensional space. Next, the distance matrix is transformed to a similarity matrix using a varying Gaussian kernel, so that the similarity between points X_i and X_j represents the joint probability that X_i will choose X_j as its neighbor or vice versa (based on their Euclidean distance and local density). Then, a random low-dimensional mapping is rendered and pairwise similarities are computed for points in the low-dimensional map. However, the low-dimensional similarities are computed using Student's t-distribution rather than a Gaussian distribution. Finally, gradient descent is used to minimize the Kullback-Leibler divergence between the two probability distributions, leading to the final low-dimensional map.

The optimization step may be interpreted as a set of springs. Each pair of points Y_i and Y_j is connected by a spring, which repels or attracts the points from each other depending on whether the similarity between the points in the projection is lower or greater than the similarity in the high-dimensional space. The gradient reduces each point's springs into a single force. The heavy tailed t-distribution helps alleviate the "crowding problem" by exerting more force when pushing distant points further apart. See **Supplementary Methods** and ref. 22 for full technical details.

The viSNE implementation. viSNE is a fast, distributed implementation of the t-SNE algorithm, improved and tailored for the analysis of single-cell data. In viSNE, each cell is represented by a point in high-dimensional space, each coordinate representing one measured parameter (e.g., the protein expression level). Our implementation then maps these points into two or three dimensions. viSNE also includes *cyt*, a tool for the interactive visualization of the resulting maps (**Supplementary Methods**).

The robustness and accuracy of t-SNE derives from the computation of all pairwise similarities. But, the similarity matrix comes at a heavy computational price, limiting the original implementation to 10,000 points. Our distributed implementation relies on the fact that each of t-SNE's computations are local and do not require the entire matrix. The technical computational limit of viSNE is 100,000 points. However, beyond 30,000 the limit is not computational, but rather the "crowding problem"[22]: the volume in high-dimensional space grows polynomial with the number of dimensions, and as a result a two-dimensional map cannot accommodate a large number of points while conserving the high-dimensional distances between them. Instead, distant points collapse onto nearby areas of the map, creating one large, dense region, with no separation between populations. To solve this, viSNE subsamples cells uniformly at random and maps the sampled population. The algorithm is robust to such subsampling (**Supplementary Fig. 2**) and even after subsampling, we still detect rare subpopulations that constitute a mere 0.2% of the population (**Supplementary Fig. 14**). viSNE is publicly available for download at http://www.c2b2.columbia.edu/danapeerlab/html/cyt.html.

Mass cytometry data. Fresh, Ficoll-enriched human bone marrow was obtained from healthy donors from AllCells, Inc. (Emeryville, CA). Samples were obtained with informed consent in accordance with the Declaration of Helsinki and with accordance with Stanford University's review board. Leukemia bone marrow samples were obtained under IRB-approved protocols (protocol number 17552 under Stanford University's IRB) at St. Jude Children's Research Hospital, Memphis, TN (pediatric acute myeloid and lymphoblastic leukemia) or at Princess Margaret Hospital, Toronto, ON (adult acute myeloid leukemia). All samples were deidentified. The age and the sex of the donor, or any additional clinical information, were unknown at the time of the study.

Samples were processed as described[21]. Briefly, cells were used fresh before mass cytometry experiments, or frozen in FCS with 10% DMSO, thawed and re-suspended in 90% RPMI with 10% FCS (supplemented with 20 U/mL sodium heparin (Sigma) and 0.025 U/mL benzonase (Sigma) in the case of frozen samples), 1× L-glutamine and 1× penicillin/streptomycin (Invitrogen).

Cells were fixed with formaldehyde (PFA; Electron Microscopy Sciences, Hatfield, PA) added directly to growth media at a final concentration of 1.6% for 10 min at room temperature. Cells were then centrifuged at 500*g* for 5 min and washed once with staining media (PBS with 0.5% BSA, 0.02% sodium azide) to remove residual PFA, and blocked with Purified Human Fc Receptor Binding Inhibitor (eBioscience Inc., San Diego, CA) following manufacturer's instructions. Surface marker antibodies were added yielding 50 or 100 μl final reaction volumes and stained at room temperature for 30 min (**Supplementary Table 2**). Following staining, cells were washed 2 more times with cell staining media, permeabilized with 4 °C methanol for at 10 min at 4 °C, and then optionally stored at −80 °C for later use. Cells were then washed twice in cell staining media to remove remaining methanol, and stained with surface and phospho-specific antibodies in 50 or 100 μl for 30 min at room temperature. Cells were washed once in cell staining media, then stained with 1 ml of 1:5000 191/193Ir DNA intercalator(2) (www.dvssciences.com; DVS Sciences, Richmond Hill, Ontario, Canada) diluted in PBS with 1.6% PFA for 20 min at room temperature. Cells were then washed once with cell staining media and then finally with water alone before running on the CyTOF.

Processing of mass cytometry data. Data were transformed using hyperbolic arcsin with a cofactor of five. Single cells were gated based on cell length and DNA content (to avoid debris and doublets) as described[21]. The expert manual classification of Marrow1 was taken from reference[21], where the complete biaxial plot gating strategy can be found.

viSNE analysis. Generating viSNE maps included the following steps (exact details can be found in **Supplementary Table 1**). First, between 6,000 and 12,000 cells were uniformly subsampled from the data. After subsampling, viSNE was run for 500 iterations to project the data into 2D. Unlike t-SNE, PCA was not used as a preprocessing step. All runs used an identical random seed and the default t-SNE parameters (perplexity = 30, momentum = 0.5 for initial 250 iterations, momentum = 0.8 for remaining iterations, epsilon = 500, lie factor = 4 for initial 100 iterations, lie factor = 1 for remaining iterations). viSNE maps were visualized using *cyt*, which was also used to generate figures (color coding by immune cell subset (as in **Fig. 1b**), by marker expression levels (as in **Fig. 1d**) and in plotting expression level densities (as in **Fig. 2b**). The raw data and viSNE maps can be downloaded at: http://www.c2b2.columbia.edu/danapeerlab/html/cyt.html.

Quantifying similarity between viSNE maps. We use the Jensen-Shannon (JS) divergence (**Supplementary Methods**) to quantify the similarity between viSNE maps. Each map is converted into a probability distribution. We define the similarity between two maps as the JS divergence between their respective distributions.

A gating scheme for FACS. The viSNE map was used to devise a gating scheme for FACS of the AML relapse sample (**Fig. 5**). Due to the limits of flow cytometry, the gating scheme can employ only a limited number of channels and use hard thresholds. Through manual inspection of the viSNE map we identified four markers that lead to distinct subpopulations that could be of interest for downstream analysis: CD34, CD64, CD33 and CD7 (**Fig. 4**). For each marker we defined a threshold for a binary negative/positive gate. The four binary gates were combined to create a total of sixteen composite gates covering all negative/positive combinations. Only six composite gates had more than 20 cells. The cells residing in each of these six composite gates are color coded on the viSNE map of **Figure 5**.

Subsampling of synthetic MRD sample. We used two samples: the synthetic MRD sample (composed of 99.5% healthy bone marrow cells and 0.5% cells from a metal-barcoded ALL sample) and the control sample (100% healthy bone marrow cells taken from a different donor). To capture a higher proportion of ALL cells for the viSNE map, we devised the following subsampling procedure. The cells from the synthetic MRD sample and from the control sample were combined computationally and clustered using the Louvain algorithm[32]. Next, the clusters were weighted by the proportion of synthetic MRD sample cells in them; the higher the proportion of synthetic MRD sample cells, the higher the weight. Finally, 10,000 cells were chosen one at a time in a two-step process: one of the clusters was chosen randomly (biased by cluster weight) and

a cell was uniformly chosen from that cluster. Note, the subsampling procedure is blind to the metal barcode; it can only access the mass cytometry measurement and the identity of the sample (synthetic MRD or control). Following the subsampling, viSNE was run as described above.

Additional algorithms. Isomap, LLE, KernelPCA and LLTSA were run using the Matlab Toolbox for Dimensionality Reduction[31]. FLOCK was compiled from the code available in the ImmPort FLOCK SourceForge page (http://sourceforge.net/projects/immportflock/). SPADE was run using the implementation available in Cytobank[33].

32. Blondel, V.D. *et al.* Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
33. Kotecha, N., Krutzik, P.O. & Irish, J.M. *Web-based analysis and publication of flow cytometry experiments. Curr. Protoc. Cytom.* **10**, 10.17 (2010).