Published in final edited form as:

Cytometry B Clin Cytom. 2015 ; 88(5): 294–304. doi:10.1002/cyto.b.21252.

Detection of Minimal Residual Disease in B Lymphoblastic Leukemia Using viSNE

Joseph A. DiGiuseppe[1,*], **Michelle D. Tadmor**[2,3], and **Dana Pe'er**[2,3,*]

[1]Department of Pathology & Laboratory Medicine, Hartford Hospital, Hartford, Connecticut

[2]Department of Biological Sciences, Columbia University, New York, New York

[3]Department of Systems Biology, Columbia University, New York, New York

Abstract

Background—Minimal residual disease (MRD) following treatment is a robust prognostic marker in B lymphoblastic leukemia. However, the detection of MRD by flow cytometric immunophenotyping is technically challenging, and an automated method to detect MRD is therefore desirable. viSNE, a recently developed computational tool based on the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm, has been shown to be capable of detecting synthetic "MRD-like" populations of leukemic cells created in vitro, but whether viSNE can facilitate the immunophenotypic detection of MRD in clinical samples has not been evaluated.

Methods—We applied viSNE retrospectively to 8-color flow cytometric immunophenotyping data from normal bone marrow samples, and samples from B lymphoblastic leukemia patients with or without suspected MRD on the basis of conventional manual gating.

Results—In each of 14 bone marrow specimens containing MRD or suspected MRD, viSNE identified a putative MRD population; an abnormal composite immunophenotype was confirmed for the putative MRD in each case. MRD populations were not identified by viSNE in control bone marrow samples from patients with increased normal B-cell precursors, or in post-treatment samples from B lymphoblastic leukemia patients who did not have detectable MRD by manual gating.

Conclusion—viSNE shows promise as an automated method to facilitate immunophenotypic MRD detection in patients treated for B lymphoblastic leukemia.

Key terms

B lymphoblastic leukemia; ALL; minimal residual disease; viSNE; flow cytometry

[*]Correspondence to: Joseph A. DiGiuseppe, Department of Pathology & Laboratory Medicine, Hartford Hospital, 80 Seymour St Hart-ford, CT 06102-5037, USA. joseph.digiuseppe@hhchealth.org or Dana Pe'er, Department of Biological Sciences and Department of Systems Biology, Columbia University, 1212 Amsterdam Ave, New York, NY 10027, USA. dpeer@biology.columbia.edu.

Additional Supporting Information may be found in the online version of this article.

INTRODUCTION

In patients who have acute lymphoblastic leukemia (ALL), “minimal residual disease” (MRD) is the persistence of rare leukemic cells despite the attainment of morphologic remission. MRD is an important adverse prognostic factor [reviewed in Refs. 1,2)] that informs contemporary risk-adapted therapeutic protocols (3,4). Two principal methodologies have been used to detect MRD in ALL: flow cytometric immunophenotyping and PCR-based amplification of antigen-receptor genes (1,2). Of these methodologies, flow cytometry is technically simpler, more rapid, and less expensive (2). Moreover, flow cytometry permits direct quantification of viable leukemic cells (2), and yields quantitative results comparable with those obtained by PCR (5). Until recently, the sensitivity of flow cytometry-based methods (10^{-4}, or 1 leukemic cell in 10,000 total cells) was inferior to that of PCR-based methodologies (10^{-5}, or 1 leukemic cell in 100,000 total cells); however, with the identification of newer immunophenotypic markers, sensitivity of 1 in 100,000 has also been demonstrated by flow cytometry (2,6).

Arguably, the most challenging aspect of MRD detection by flow cytometry is the user-dependent expertise required to identify rare abnormal cells among an overwhelming preponderance of normal cells, many of which display composite immunophenotypes that closely resemble those of the leukemic cells being sought (2,7). Immunophenotypic data are analyzed in a series of 2-dimensional dot plots, each with its own characteristic set of antigenic expression patterns. A 9-color experiment, for instance, yields 55 unique 2-dimensional displays of light scatter and fluorescent parameters (8), any of which might potentially contain discriminatory phenotypic information. As a practical matter, the expertise required for immunophenotypic MRD detection has limited its implementation in multi-center studies to a relatively small number of reference laboratories (3,4,6). Standardization of immunophenotypic MRD detection among different laboratories has also been problematic; although standardization has been achieved to varying extents (9–12), interlaboratory concordance has been lower for MRD levels below 0.1% (9), and much of the interlaboratory discordance has been attributed to discrepancies in data analysis (11).

Automated data analysis represents one potential solution to the challenges posed by immunophenotypic MRD detection. Indeed, computational methods for analysis of multiparametric flow cytometric data have become increasingly important (reviewed in Ref. 13)) as conventional manual gating of large single-cell data sets generated by contemporary fluorescence-based cytometers (14,15), and more recently, time-of-flight mass spectrometry (16), is both impractical and unlikely to extract much of the information contained within the data. Aghaeepour et al. (13) have assessed numerous automated methods for clustering subpopulations within multiparametric data sets, but these methods are insufficiently sensitive to detect MRD populations. Fišer et al. employed a combination of hierarchical clustering analysis and support vector machine (SVM) learning in an effort to automate MRD detection, but distinction of MRD from regenerating normal B-cell precursors (hematogones) was problematic for the SVM classifier (17).

More recently, a computational tool called “viSNE” (after the t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm on which it is based (18)), was shown to be capable

of detecting synthetic "MRD-like" populations of leukemic cells created in vitro by dilution of leukemic bone marrow samples into normal bone marrow samples (19). In viSNE, each cell is represented as a point in high-dimensional space, each coordinate of which reflects a single parameter (e.g., the expression level of one surface protein). An optimization algorithm searches for a projection of the points in high-dimensional space (e.g., 8 dimensions when analyzing the fluorescent parameters of an 8-color experiment) into two dimensions, so that pairwise distances between the points are optimally conserved. The resulting low-dimensional projection, termed the *viSNE map*, is visualized as a scatter plot, in which the location of a cell reflects information from all of its original dimensions. The resulting viSNE map provides a visual representation of the single-cell data that is similar to a biaxial plot, in which the positions of individual cells reflect their proximity in high-dimensional space. Moreover, the multiparameteric characteristics of individual cells in the viSNE map can be visualized interactively by incorporating color as a third dimension. In the recently published synthetic example of MRD detection, viSNE accurately distinguished a leukemic cell population (which comprised 0.25% of all cells) from normal bone marrow cells (19), demonstrating the ability of the algorithm to identify rare populations on the basis of immunophenotype.

Whether viSNE can facilitate the immunophenotypic detection of MRD in clinical samples, though, has not been evaluated. Although viSNE was able to detect a 0.25% MRD population (19), a sensitivity of at least 10^{-4} (0.01%) would be required to be clinically useful (1,2). Moreover, in the synthetic example described above, a leukemic sample consisting predominantly of blasts was used to simulate an MRD sample (19), whereas clinical samples obtained after treatment not only contain a potential MRD population, but also normal B-cell precursors (20). Since normal B-cell precursors are both immunophenotypically similar to leukemic B-cell lymphoblasts, and particularly numerous in actively regenerating bone marrow following chemotherapy (20), their presence in clinical samples represents a potential impediment to MRD detection by viSNE. An additional requirement for clinical usefulness would be the ability of the algorithm to distinguish the spectrum of abnormal immunophenotypes encountered in clinical practice, while minimizing identification of rare immunophenotypic subsets of normal B-cell precursors. Here we apply viSNE retrospectively to 8-color flow cytometric immunophenotyping data from normal bone marrow samples, and samples from B lymphoblastic leukemia patients with or without suspected MRD following therapy. We demonstrate that viSNE overcomes these challenges, and thus shows promise as an effective aid for distinguishing MRD populations in clinical samples.

MATERIALS AND METHODS

Patient Samples and Flow Cytometric Immunophenotyping

EDTA-anticoagulated bone marrow aspirate samples from patients with or without B lymphoblastic leukemia (Supporting Information Table 1) were received in the Special Hematology Laboratory of Hartford Hospital for flow cytometric immunophenotyping between October, 2011 and March, 2014. The diagnosis of B lymphoblastic leukemia was established in each case using standard WHO criteria (21,22). Bone marrow suspensions

containing 2×10^6 cells were incubated within 12 h of their procurement with each of three validated cocktails containing eight or nine fluorochrome-conjugated antibodies [(23), Supporting Information Table 2] for 15 min at room temperature. Labeled samples were processed for flow cytometric immunophenotyping as described previously (24). Up to 1×10^6 cells per sample were acquired using BD FACSCanto II flow cytometers equipped with violet, blue, and red lasers running BD FACS-Diva software v. 6.1.3 (BD Biosciences, San Jose, CA). Instrument performance was verified daily using the BD Cytometer Setup & Tracking Beads kit (BD Biosciences). CD45-V500 was compensated using cellular controls as described previously (25). Compensation of all other fluorochromes was performed using BD CompBeads (BD Biosciences) according to the manufacturer's instructions. Compensated list-mode data were analyzed using WoodList 2.7.7. Doublets and other cellular aggregates were excluded using a display of FSC peak area vs.FSC peak height, and viable cells were gated using standard displays of FSC versus SSC (7,8). Since all of our samples were procured locally, and processed within 12 h of procurement, a viability stain was not found to be necessary to exclude dead cells. Detection of MRD by conventional manual gating (2,6,7,26) was performed by one of us (JAD) without any knowledge of the results obtained using viSNE. This study was approved by the Institutional Review Board of Hartford Health Care (Project #4214HU).

viSNE Analysis

cyt, an interactive visualization tool written in Matlab (19), is available as a download with viSNE (http://www.c2b2.columbia.edu/danapeerlab/html/cyt-download.html). Matlab (8.1.0.604) was run using a MacBook Pro (Mac OSX version 10.7.5). Viable singlet CD19+ B cells from each case were gated in WoodList, and imported into *cyt* as FCS 2.0 files. Multiple samples are needed to cover the range of behaviors observed with normal bone marrow. Additionally, it is important to include enough cells from each sample, so that rare normal precursor populations will be represented. However, computational burden increases with number of cells. To balance the numbers of samples and cells with computational burden, we used 10 different bone marrow samples from patients with no history of ALL, each of which contributed ~20,000 B cells. To enable direct comparison among cell populations, all samples must be simultaneously mapped by viSNE. For each sample evaluated, viSNE was run on all viable singlet B-cell files from each of the three antibody combinations evaluated (Supporting Information Table 2), together with the 10 control B-cell files from the corresponding antibody combination. Files from the same 10 control samples were used for all analyses. All eight fluorescent parameters were incorporated into the analysis. The addition of viable singlet B cells from each test sample to the normal reference files resulted in a total of ~200,000–250,000 cells to be mapped at a time, an order of magnitude more than previously demonstrated (19). To overcome both the computational and crowding burden imposed by so many cells, we used a newly published algorithm for t-SNE, the Barnes–Hut approximation (27), which is currently included within the *cyt* software (set as default). Typical run times for the generation of a t-SNE map for ~200,000–250,000 cells were ~40 min.

RESULTS

A fundamental principle that enables MRD detection by flow cytometric immunophenotyping is that leukemic cells typically express a combination of antigens (or composite immunophenotype) not observed in their normal counterparts (1–3). In the case of B lymphoblastic leukemia, basic phenotypic patterns of normal B-lymphopoiesis have been well documented (20,28), and with a suitable antibody panel, nearly all cases of B lymphoblastic leukemia are detectable by their abnormal immunophenotype (often referred to as a leukemia-associated immunophenotype) (6). In many studies, post-treatment samples have been analyzed solely for residual populations expressing the specific leukemia-associated immunophenotype documented at the time of diagnosis. More recently, it has become clear that treatment may induce phenotypic changes in leukemic B lymphoblasts (29); for instance, leukemic blasts expressing both CD10 and CD34 at diagnosis may be negative for both antigens after induction chemotherapy. A more robust approach to immunophenotypic MRD detection, which is neither subject to such treatment-induced immunophenotypic modulations, nor requires prior knowledge of the specific diagnostic immunophenotype, is to identify B-cell populations whose composite immunophenotypes differ from those characteristic of normal B-cell precursors (7,28). Since viSNE plots healthy and leukemic marrows to distinct areas of 2-dimensional maps representing high-dimensional immunophenotypic data, and is capable of detecting rare populations with distinct immunophenotypes (19), we hypothesized that viSNE might facilitate identification of MRD populations as phenotypically different from normal B-cell precursors.

In conventional immunophenotypic MRD analysis, post-treatment samples are compared visually with identically stained normal B-cell precursors using a large series of 2-dimensional dot plots (7,24,26,28). To distinguish potential MRD populations from normal B-cell precursors with viSNE, one only needs to generate a single 2-dimensional map for each antibody combination. For each of the three 8-color antibody-staining combinations employed (A–C, Supporting Information Table 2), we used viSNE to create a 2-dimensional map of normal bone marrow B cells. Since a single normal reference sample would be unlikely to encompass all of the potential high-dimensional immunophenotypic diversity seen among individual patients, we created a composite map incorporating normal bone marrow B cells from multiple individuals. In a series of initial experiments, both the number of normal samples used to construct the maps, and the number of normal B cells per sample were varied. Reproducible results were obtained using 10 normal samples (“Control Cases,” Supporting Information Table 1), each with ~20,000 B cells; we therefore established that as our standard.

The normal viSNE map obtained for one of the three antibody combinations (tube A) is shown in Figure 1. Note that the map has a structure, and that each of the 10 control samples conforms to that structure (Fig. 1A). Furthermore, there is extensive mixing of cells, such that each region of the map comprises cells from multiple control samples from different individuals (Fig. 1A). As shown in Figure 1B, the structure of the map reflects the patterns of antigens expressed by the cells in any given region; despite the contribution of multiple samples to each region of the map (Fig. 1A), antigenic expression in each region is homogeneous (Fig. 1B). In aggregate, the control samples faithfully recapitulate expected

patterns of normal B-lymphopoiesis, including the least-mature B-cell precursors (with high expression of CD10 and CD34, and low expression of CD45), mature B cells (with low expression of CD10 and CD38, and high expression of CD22), as well as intermediate forms (Fig. 1B). As expected, individual control samples are non-uniformly represented among different maturational stages; for example, control sample 1 contains relatively few mature B cells, and instead contributes cells predominantly to regions corresponding with the least-mature and intermediate forms (Fig. 1A).

Pairwise immunophenotypic relationships that exist among individual cells in high-dimensional space are conserved in 2-dimensional viSNE maps (19), and as noted above, leukemic B-lymphoblasts typically differ phenotypically from normal B cells. We therefore tested whether leukemic B-lymphoblast populations would occupy a unique space in the resulting viSNE map when analyzed in conjunction with the 10 control samples. As illustrated in Figure 2A using four representative patients' diagnostic bone marrow samples, leukemic B-lymphoblast populations occupy a unique space in their respective viSNE maps, which is distinct from the space occupied by normal B cells. These results are in keeping with those reported previously using whole bone marrow samples (19). However, we demonstrate here that eight surface antigens are sufficient to achieve distinct separation, whereas such separation was previously demonstrated with 31 markers (19). In Figure 2B, the viSNE map incorporating one of the leukemic samples (ALL 1, circled) is colored for expression of several antigens. In this example, the leukemic population expresses CD34, CD49f, and CD58 at higher levels than the normal B cells (Fig. 2B). Having used viSNE to facilitate the characterization of abnormal antigen expression in the leukemic population, we can now visualize these abnormally expressed antigens in a conventional 3-dimensional plot; indeed, abnormally increased expression of these three antigens by the leukemic cells relative to normal B cells is also apparent in the 3-dimensional dot plot (Fig. 2C).

In some instances, the phenotype of leukemic B-lymphoblasts may differ substantially from that of their normal B-cell counterparts, yet display partial immunophenotypic overlap with normals. Such a relationship is reflected in the viSNE map illustrated in Figure 3A. In this example, the leukemic cells occupy a space in the map that is distinct from a preponderance of the normal B cells, but rare normal B cells (enlarged) nonetheless map to the leukemic rather than normal B-cell space. Since the viSNE map reflects high-dimensional pairwise distances among all cells included in the analysis (both normal and abnormal), the inclusion of leukemic cells in the analysis should affect the two-dimensional location of individual normal cells in the resulting map. Specifically, the presence of a prominent abnormal (i.e., leukemic) population would be expected to increase the number of repelling forces pushing normal cells away from that region of the map, whereas a smaller number of cells (i.e., ~1–3) may fail to create such separation. This finding recapitulates an observation made initially with a synthetic MRD sample, in which the suspect region in the viSNE map containing ALL cells included rare normal cells (19). The basis for this overlap of rare normal cells with leukemic cells in the viSNE map is illustrated in the dot plot in Figure 3B, in which it is apparent that these rare normal B cells are more phenotypically similar to a subset of leukemic cells than to any of the remaining normal B cells. This observation has implications for MRD detection using viSNE. Specifically, the presence of rare normal cells within a suspect region in the viSNE map does not preclude identification of leukemic cells

in that region, provided that the region contains a distinctly identifiable population of cells with an abnormal immunophenotype.

Having established viSNE maps for normal bone marrow B cells and demonstrated that leukemic B-cell populations with abnormal composite immunophenotypes occupy spaces distinct from normal B cells in viSNE maps, we sought to test whether rare MRD populations would also be distinguished from normal B cells by viSNE. For each of 14 samples in which a suspect MRD population had been identified by manual gating, we generated viSNE maps for three different 8-color antibody combinations in conjunction with their corresponding identically stained control samples. Representative examples of these analyses are illustrated in Figure 4. In the cases illustrated (Fig. 4), most of the B cells from the suspect files localize to regions of the viSNE maps that are also populated by normal B cells from the control files. In addition, though, each suspect file contains a population that does not localize to an area of the viSNE map populated by control B cells (Fig. 4). In conventional 3D-dot plots, these outlier populations represent clusters of cellular events, whose composite immunophenotypes differ from those of the normal B cells in patterns typical of leukemic B-lymphoblasts (Fig. 4). For example, abnormally increased expression of CD10 and CD86, and diminished expression of CD38, as seen in Figure 4A, are common phenotypic aberrancies in B lymphoblastic leukemia (6,7).

Of the 42 viSNE maps generated from the 14 MRD-containing samples stained with three different antibody combinations, a putative MRD population residing outside of regions populated by normal B cells was present in 34 (Table 1); in each instance, an abnormal immunophenotype consistent with MRD was verified by visual inspection of dot plots, as illustrated in Figure 4. (The *cyt* software enables one to identify uniparametric differences between two populations (e.g., the putative MRD highlighted by viSNE and normal), thereby highlighting the distinguishing features of such populations, and facilitating rapid generation of the most relevant conventional dot plots.) Note also that multiple cells that are phenotypically similar in high-dimensional space may be represented in the 2-dimensional viSNE map as very close together (Fig. 4A) or overlapping (Figs. 4B and 4C). Importantly, in each of the 14 MRD-containing samples, viSNE distinguished a putative MRD population with at least one of the three antibody combinations (Table 1). Whereas the manual identification of candidate MRD populations is an iterative process, requiring the examination of multiple 2-dimensional dot plots by a trained analyst to recognize potential MRD populations with abnormal phenotypic features, a single 2-dimensional viSNE plot was capable of segregating each of these abnormal immunophenotypes, demonstrating the ability of the algorithm to detect “different from normal” across a wide range of phenotypes.

Although viSNE distinguished a putative MRD population in each of the 14 MRD-containing samples with at least one of the three antibody combinations, in some instances, the algorithm failed to detect an abnormal population in individual tubes in which MRD was detectable by manual gating. In some of these instances, the most discriminatory phenotypic abnormalities were not evaluated in the tube in question, and MRD was only recognizable by manual gating through inference. For example, in Case 9, aberrant phenotypes sufficiently distinct from normal B-cell precursors to permit definitive characterization as MRD were only demonstrable in tubes B and C. In this case, overexpression of CD49f (in

tube B) and CD86 (in tube C) clearly distinguished an abnormal B-cell population, which was further characterized by CD10 expression and dim CD45 expression. Although no specific phenotypic aberrancies diagnostic of MRD were revealed by the antibodies included in tube A, the 694 cells identified by manual gating (Table 1) were nonetheless recognizable in tube A inferentially, by virtue of their shared phenotypic features [i.e., CD10+/CD45(dim)+] with the corresponding population in the other tubes, in which definitive aberrant phenotypes [i.e., overexpression of CD49f and CD86 in tubes B and C, respectively] were demonstrable. In this case, therefore, it is not surprising that viSNE detected an abnormal population in tubes B and C, but not in tube A. In other cases, the precise reason for the failure of viSNE to detect the abnormal population was unclear. For example, in case 14, viSNE failed to detect an abnormal population in tube C, in which overexpression of CD86 readily permitted direct manual identification of a population of 103 cells.

The suspect MRD populations in these 14 samples comprised between 0.001% and 0.7% of viable single cells in the original samples. Since the total number of viable bone marrow cells obtained from different patients' samples also varies, the actual number of MRD cells varies widely among the different samples evaluated (Table 1). As shown in Figure 5, the number of putative MRD cells identified by viSNE correlated well with the number of cells that had been previously identified as MRD by manual gating ($r^2 = 0.99$; two-tailed $P < 0.0001$). Moreover, in two different samples, MRD populations comprising as few as six cells were detected by viSNE (Table 1). For comparison, visual identification of cellular populations comprising fewer than 10–20 cells is technically difficult (7,30).

To evaluate the specificity of abnormal population identification by viSNE, we studied seven bone marrow samples from patients with no history of B lymphoblastic leukemia ("Negative Controls," Supporting Information Table 1), eight samples from patients with B lymphoblastic leukemia in remission who had no detectable MRD by conventional manual gating ("No detectable disease in bone marrow," Supporting Information Table 1), and three samples from a patient with B lymphoblastic lymphoma who had no detectable leukemic blasts in the bone marrow by conventional manual gating ("No detectable disease in bone marrow," Supporting Information Table 1). Each sample was stained with the identical 8-color antibody combinations used for MRD detection, and viSNE maps were generated in conjunction with the 10 control samples, as with the MRD-containing samples. No outlier populations were identified in 17 of the 21 viSNE maps generated from patients with no history of B lymphoblastic leukemia; a representative map is illustrated in Figure 6A (left panel). Similarly, no outlier populations were identified in 22 of the 24 viSNE maps generated from patients with B lymphoblastic leukemia in remission, or in eight of the nine maps generated from a patient with B lymphoblastic lymphoma without detectable marrow involvement; a representative map is illustrated in Figure 6A (right panel).

Outlier populations were identified in 7 of the 54 maps generated. In three of these samples, the outlier population comprised a single cell, and therefore could not be confirmed as a distinct cell population; as noted above, a population consisting of at least 10 cells with an abnormal immunophenotype is generally required for reliable identification (30). In the remaining four samples, no outlier population comprised [mt]4 cells. The map containing the largest outlier population (4 cells) is shown in Figure 6B (left panel). These cells express

a mature composite immunophenotype, with bright expression of CD20, CD22, and CD45, and absence of CD10 and CD34; in conventional dot plots, the outlier population cannot be distinguished from normal B cells (Fig. 6B, right panel). Likewise, in no instance did any of these small outlier populations display an abnormal composite immunophenotype consistent with MRD. These data indicate that outlier populations segregated by viSNE as "different from normal" must still be con-firmed to express an abnormal composite immunophenotype consistent with MRD before being classified definitively as such. viSNE comes accompanied with *cyt* and an interactive tool that automatically identifies differences in the expression of individual antigens between populations of interest in the viSNE map (e.g., between an outlier population and normal controls), thereby facilitating the process of confirmation of a putative outlier population as MRD.

DISCUSSION

We demonstrate here for the first time in clinical samples facilitated immunophenotypic detection of MRD in B lymphoblastic leukemia using viSNE. It is well established that the presence of MRD following treatment is among the strongest prognostic factors in patients with B lymphoblastic leukemia (1,2), and contemporary protocols employ MRD detection to stratify patients more precisely for risk-adapted therapy (3,4). However, the expertise and experience required for accurate immunophenotypic MRD detection have limited its implementation in multicenter studies to a relatively small number of reference laboratories (3,4,6). Moreover, despite efforts to standardize immunophenotypic MRD detection (9–12), interlaboratory concordance for MRD levels below 0.1% has been problematic (9), and much of the interlaboratory discordance has been attributed to discrepancies in data analysis (11).

Automated data analysis represents one potential solution to the challenges posed by immunophenotypic MRD detection. Numerous automated methods for clustering subpopulations within multiparametric data sets have been developed [reviewed in Ref. 13)], but these methods are insufficiently sensitive to detect MRD populations. In one study, a combination of hierarchical clustering analysis and SVM learning was used in an attempt to automate MRD detection. However, distinction of MRD from regenerating normal B-cell precursors (hematogones) was problematic for the SVM classifier (17). More recently, viSNE, a computational tool based on the t-SNE algorithm (18), was shown to be capable of detecting a synthetic "MRD-like" population of leukemic cells created in vitro by dilution of a leukemic bone marrow sample into a normal bone marrow sample (19).

Although viSNE has been shown to be capable of detecting a synthetic "MRD-like" population of leukemic cells (19), this synthetic example fails to take into account several of the potential challenges inherent in MRD detection in clinical samples. For example, whereas the target population represented 0.25% of all cells in the synthetic example (19), a sensitivity of at least 0.01% (and up to 0.001% with newer markers) is standard in clinical MRD detection (1,2,6). The incorporation of a newly-published algorithm for t-SNE [the Barnes–Hut approximation (27), which is currently included within the *cyt* software] has overcome both the computational and crowding burden imposed by the large numbers of cells required for MRD detection. Indeed, we have shown here that viSNE is capable of

accurately identifying MRD populations comprising as few as six cells, and representing as little as 0.001% of viable cells in clinical samples. In this regard, it should be noted that the visual identification of cellular populations comprising fewer than 10–20 cells is technically difficult (7,30).

Since viSNE takes into account differences in all dimensions simultaneously, it may potentially be more sensitive than manual gating when identifying small (i.e., <10 cells) abnormal populations. Determination of the minimum number of events required for reliable detection of an abnormal population by viSNE, though, would require comparison with an alternate method, such as next-generation sequencing (31). Moreover, the minimum number of events required for reliable detection of an abnormal population by viSNE would likely be affected by the degree of phenotypic differences between the abnormal population and the control cells within the context of each antibody combination evaluated; determination of this detection limit would therefore require extensive testing, which is beyond the scope of this study. Nonetheless, in the example illustrated in Figure 4C (Case 3, tube A), we have performed preliminary experiments to address the detection limit of viSNE. We randomly subsampled the MRD population, which comprised eight cells, to represent four cells, two cells, or one cell. We then subjected each sub-sample to analysis by viSNE in conjunction with the 10 standard control samples. We found that the subsample containing four putative MRD cells still appeared as an outlier in the resulting viSNE map, whereas neither the 2-cell nor 1-cell subsamples was distinguished from normal in the corresponding viSNE maps.

Other potential challenges inherent in MRD detection in clinical samples not reflected in the synthetic example include the presence within regenerating bone marrow of normal B-cell precursors (20), the wide spectrum of abnormal immunophenotypes encountered in clinical practice, and day-to-day variability in instrument and reagent performance. In spite of these potential pitfalls, our data demonstrate the ability of viSNE accurately to identify MRD populations in clinical samples. Normal B-cell precursors in samples from patients with ALL, and individuals with no history of leukemia were appropriately classified by the algorithm as being similar to their counterparts in the normal viSNE map. Furthermore, MRD populations were correctly identified in all 14 samples from 10 different patients, each with a different abnormal composite immunophenotype. Finally, in spite of efforts to minimize variability in the day-to-day performance of instrumentation and reagents (8), it remained possible that the algorithm might be sensitive to slight technical perturbations that are not, in fact, indicative of an abnormal immunophenotype. That our data were acquired over a time frame in excess of two years, yet no such technical artifacts were detected, indicates that the sensitivity of viSNE to differences in phenotype does not preclude its use in the setting of a clinical laboratory employing routine quality-assurance measures.

Although this study demonstrates in principle that viSNE may facilitate MRD detection in clinical samples, it will be important to confirm these findings in a separate cohort. Since *cyt* (which includes viSNE) is freely available as a download (http://www.c2b2.columbia.edu/danapeerlab/html/cyt-download.html), this approach may easily be applied to any existing large data sets from cooperative groups or other large collaborative efforts. Moreover, since the ability of the algorithm to distinguish abnormal from normal depends upon the precise antibody combination used, it will be of interest to evaluate this approach using different

antibody panels. Newer instruments approved for clinical use are capable of 10-color analysis, enabling the incorporation of additional discriminatory markers per tube in addition to those required for standard gating (2); the number of available parameters is expected to increase further as the technology advances. A key advantage to viSNE is that it is designed to deal with higher dimensionality, and can easily accommodate 50 markers (we note that 50 markers result in 1,225 biaxial plots, which would be daunting for any expert). The incorporation of additional discriminatory markers is likely to improve the specificity of candidate MRD population identification by viSNE.

In this context, it is important to note that viSNE does not entirely replace the requirement for an analyst in immunophenotypic MRD detection. Detection of MRD by manual gating fundamentally includes two distinct steps: identification of candidate populations, and confirmation that a given candidate population in fact represents a population of leukemic blasts rather than a technical artifact (e.g., non-specific staining). The first step, identification of candidate MRD populations, is time-consuming, requiring visual inspection of many 2-dimensional dot plots for potential outlier populations, and is subject to false-negative errors. It is this first step that is facilitated by viSNE. As with manual gating, though, the candidate outlier populations identified by viSNE proved not to represent actual MRD in a minority of instances. This observation indicates that, at least with the antibody combinations and control samples employed in this study, confirmation of candidate populations detected by viSNE as actual MRD is required. It is possible that the specificity of detection might increase with the inclusion of additional control samples from patients of widely varying ages. Nonetheless, the availability of a sensitive automated screening method for candidate MRD populations should facilitate more widespread implementation of immunophenotypic MRD detection in clinical flow cytometry laboratories.

Finally, although this study has focused on facilitated immunophenotypic MRD detection in B lymphoblastic leukemia, flow cytometry is also used to detect MRD in other hematolymphoid malignancies, including acute myeloid leukemia (32), chronic lymphocytic leukemia (33), and plasma cell myeloma (34). As in ALL, immunophenotypic MRD detection in these other malignancies relies upon the identification of abnormal composite immunophenotypes, which differ from those seen in their respective normal cellular counterparts. The general approach used in this study, including the generation of a reference viSNE map with a series of normal samples, may also serve as a model for analysis in these different systems. Further study is warranted to test the applicability of viSNE in these and other settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant sponsor: National Science Foundation graduate fellowship.

Grant sponsor: National Institutes of Health Roadmap Initiative; Grant number: 1-DP2-OD002414-01, DP1 HD084071-01.

Grant sponsor: National Institutes of Health; Grant number: R01CA164729.

The authors gratefully acknowledge Dr. Brent Wood, University of Washington, Seattle, WA, for his generous gift of WoodList, and the members of the Special Hematology Laboratory at Hartford Hospital, Jolene Cardinali, Beth Constant, Deb Haas, Len Iorio, Stephanie Lacroix, Kathy Pisk, and Sandy Vicki for their dedication and outstanding technical support. JAD gratefully acknowledges Kristina R. (Gigi) DiGiuseppe for support with software.

LITERATURE CITED

1. Campana D. Minimal residual disease in acute lymphoblastic leukemia. ASH Educ Program Book. 2010; 2010:7–12.
2. Gaipa G, Basso G, Biondi A, Campana D. Detection of minimal residual disease in pediatric acute lymphoblastic leukemia. Cytometry Part B Clin Cytom. 2013; 84B:359–369.
3. Borowitz MJ, Devidas M, Hunger SP, Bowman WP, Carroll AJ, Carroll WL, Linda S, Martin PL, Pullen DJ, Viswanatha D, 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. 2008; 111:5477–5485. [PubMed: 18388178]
4. Basso G, Veltroni M, Valsecchi MG, Dworzak MN, Ratei R, Silvestri D, Benetello A, Buldini B, Maglia O, Masera G, et al. Risk of relapse of childhood acute lymphoblastic leukemia is predicted by flow cytometric measurement of residual disease on day 15 bone marrow. J Clin Oncol. 2009; 27:5168–5174. [PubMed: 19805690]
5. Coustan-Smith E, Campana D. Immunologic minimal residual disease detection in acute lymphoblastic leukemia: A comparative approach to molecular testing. Best Pract Res Clin Haematol. 2010; 23:347–358. [PubMed: 21112034]
6. Coustan-Smith E, Song G, Clark C, Key L, Liu P, Mehrpooya M, Stow P, Su X, Shurtleff S, Pui C-H, et al. New markers for minimal residual disease detection in acute lymphoblastic leukemia. Blood. 2011; 117:6267–6276. [PubMed: 21487112]
7. DiGiuseppe, JA. Acute lymphoblastic leukemia/lymphoma: Diagnosis and minimal residual disease detection by flow cytometric immunophenotyping. In: Detrick, B.Hamilton, R., Schmitz, J., editors. Manual of Molecular and Clinical Laboratory Immunology. 8. Washington, D.C: ASM Press; in press
8. Wood B. 9color and 10-color flow cytometry in the clinical laboratory. Arch Pathol Lab Med. 2006; 130:680–690. [PubMed: 16683886]
9. Dworzak MN, Gaipa G, Ratei R, Veltroni M, Schumich A, Maglia O, Karawajew L, Benetello A, Poetschger U, Husak Z, et al. Standardization of flow cytometric minimal residual disease evaluation in acute lymphoblastic leukemia: Multicentric assessment is feasible. Cytometry Part B Clin Cytom. 2008; 74:331–340.
10. Irving J, Jesson J, Virgo P, Case M, Minto L, Eyre L, Noel N, Johansson U, Macey M, Knotts L, et al. Establishment and validation of a standard protocol for the detection of minimal residual disease in B lineage childhood acute lymphoblastic leukemia by flow cytometry in a multi-center setting. Haematologica. 2009; 94:870–874. [PubMed: 19377076]
11. Luria D, Rosenthal E, Steinberg D, Kodman Y, Safanaiev M, Amariglio N, Avigad S, Stark B, Izraeli S. Prospective comparison of two flow cytometry methodologies for monitoring minimal residual disease in a multicenter treatment protocol of childhood acute lymphoblastic leukemia. Cytometry Part B Clin Cytom. 2010; 78:365–371.
12. Mejstříková E, Froňková E, Kalina T, Omelka M, Batinić D, Dubravčić K, Pospisilova K, Vaskova M, Luria D, Cheng SH, et al. Detection of residual B precursor lymphoblastic leukemia by uniform gating flow cytometry. Pediatr Blood Cancer. 2010; 54:62–70. [PubMed: 19760767]
13. Aghaeepour N, Finak G, Hoos H, Mosmann TR, Brinkman R, Gottardo R, Scheuermann RH. Critical assessment of automated flow cytometry data analysis techniques. Nat Methods. 2013; 10:228–238. [PubMed: 23396282]
14. Chattopadhyay PK, Hogerkorp CM, Roederer M. A chromatic explosion: The development and future of multiparameter flow cytometry. Immunology. 2008; 125:441–449. [PubMed: 19137647]

15. Chattopadhyay PK, Roederer M. Cytometry: Today's technology and tomorrow's horizons. Methods. 2012; 57:251–258. [PubMed: 22391486]
16. Bendall SC, Simonds EF, Qiu P, El-ad DA, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science. 2011; 332:687–696. [PubMed: 21551058]
17. Fišer K, Sieger T, Schumich A, Wood B, Irving J, Mejstříková E, Dworzak MN. Detection and monitoring of normal and leukemic cell populations with hierarchical clustering of flow cytometry data. Cytometry Part A. 2012; 81:25–34.
18. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008; 9:85.
19. Amir, E-Ad, Davis, KL., Tadmor, MD., Simonds, EF., Levine, JH., Bendall, SC., Shenfeld, DK., Krishnaswamy, S., Nolan, GP., Pe'er, D. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol. 2013; 31:545–552. [PubMed: 23685480]
20. McKenna RW, Washington LT, Aquino DB, Picker LJ, Kroft SH. Immunophenotypic analysis of hematogones (B-lymphocyte precursors) in 662 consecutive bone marrow specimens by 4-color flow cytometry. Blood. 2001; 98:2498–2507. [PubMed: 11588048]
21. Borowitz, MJ., Chan, JKC. B lymphoblastic leukaemia/lymphoma, not otherwise specified. In: Swerdlow, SH.Campo, E.Harris, NL.Jaffe, ES.Pileri, SA.Stein, H.Thiele, J., Vardiman, JW., editors. WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues. 4. Lyon: International Agency for Research on Cancer (IARC); 2008. p. 168-170.
22. Borowitz, MJ., Chan, JKC. B lymphoblastic leukaemia/lymphoma, with recurrent genetic abnormalities. In: Swerdlow, SH.Campo, E.Harris, NL.Jaffe, ES.Pileri, SA.Stein, H., et al., editors. WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues. 4. Lyon: International Agency for Research on Cancer (IARC); 2008. p. 171-175.
23. Vrotsos E, Gorgan M, DiGiuseppe JA. Detection of small abnormal B-lymphoblast populations at diagnosis of chronic myelogenous leukemia, BCR-abl1+: Incidence, phenotypic features, and clinical implications. Cytometry Part B Clin Cytom. in press.
24. DiGiuseppe JA, Fuller SG, Borowitz MJ. Overexpression of CD49f in precursor B-cell acute lymphoblastic leukemia: Potential usefulness in minimal residual disease detection. Cytometry Part B Clin Cytom. 2009; 76:150–155.
25. DiGiuseppe JA, Cardinali J. Improved compensation of the fluorochrome AmCyan using cellular controls. Cytometry Part B Clin Cytom. 2011; 80:191–194.
26. DiGiuseppe JA. Acute lymphoblastic leukemia: Diagnosis and detection of minimal residual disease following therapy. Clin Lab Med. 2007; 27:533–549. [PubMed: 17658406]
27. Van der Maaten L. Barnes-Hut-SNE. arXiv 2013;1301.3342:1–11.
28. Wood, B. Multicolor immunophenotyping: Human immune system hematopoiesis. In: Darzynkiewicz, Z.Roederer, M., Tanke, HJ., editors. Cytometry, 4th ed.: New Developments. Methods in Cell Biology. Vol. 75. San Diego, CA: Elsevier; 2004. p. 559-576.
29. Gaipa G, Basso G, Aliprandi S, Migliavacca M, Vallinoto C, Maglia O, Faini A, Veltroni M, Husak D, Schumich A, et al. Prednisone induces immunophenotypic modulation of CD10 and CD34 in nonapoptotic B-cell precursor acute lymphoblastic leukemia cells. Cytometry Part B Clin Cytom. 2008; 74:150–155.
30. Soma L, Wood B. Minimal residual disease testing in acute leukemia. Int J Hematol Oncol. 2013; 2:467–485.
31. Wood BL. Principles of minimal residual disease detection for hematopoietic neoplasms by flow cytometry. Cytometry Part B Clin Cytom. in press.
32. Walter R, Gooley T, Wood B, Milano F, Fang M, Sorror M, Estey EH, Salter AI, Lansverk E, Chien JW, et al. Impact of pretransplantation minimal residual disease, as detected by multiparametric flow cytometry, on outcome of myeloablative hematopoietic cell transplantation for acute myeloid leukemia. J Clin Oncol. 2011; 28:1–8.
33. Rawstron A, Villamor N, Ritgen M, Böttcher M, Ghia P, Zehnder J, Lozanski G, Colomer D, Moreno C, Geuna M, et al. International standardized approach for flow cytometric residual disease monitoring in chronic lymphocytic leukaemia. Leukemia. 2007; 21:956–964. [PubMed: 17361231]

34. Rawstron A, Child J, de Tute R, Davies F, Gregory W, Bell S, Szubert AJ, Navarro-Coy N, Drayson MT, Feyler S, et al. Minimal residual disease assessed by multiparameter flow cytometry in multiple myeloma: Impact on outcome in the medical research council myeloma IX study. J Clin Oncol. 2013; 31:2540–2547. [PubMed: 23733781]

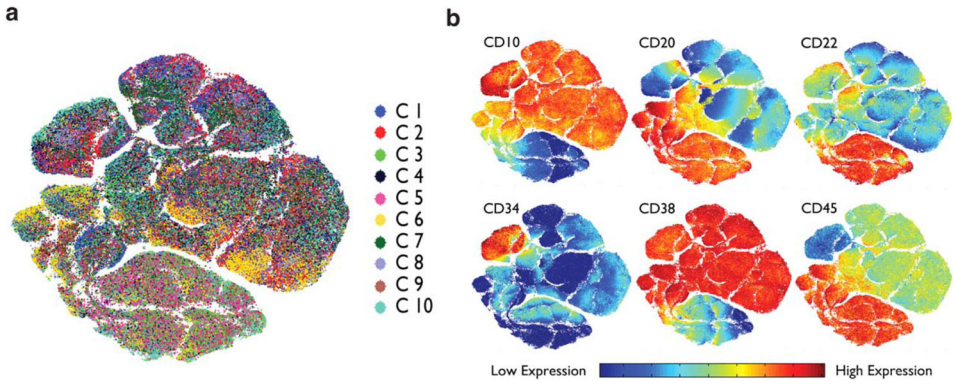

Fig. 1.

viSNE map of normal bone marrow B cells labeled with a single 8-color antibody combination (tube A). Panel **A**: Each of the 10 control bone marrow samples used to construct the map (C1 through C10) is identified individually with a unique color. Each point in the viSNE map represents an individual cell colored by sample identity. Panel **B**: viSNE map shown in Panel (A), in which each cell is colored to reflect intensity of antigen expression for six different markers in tube A. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

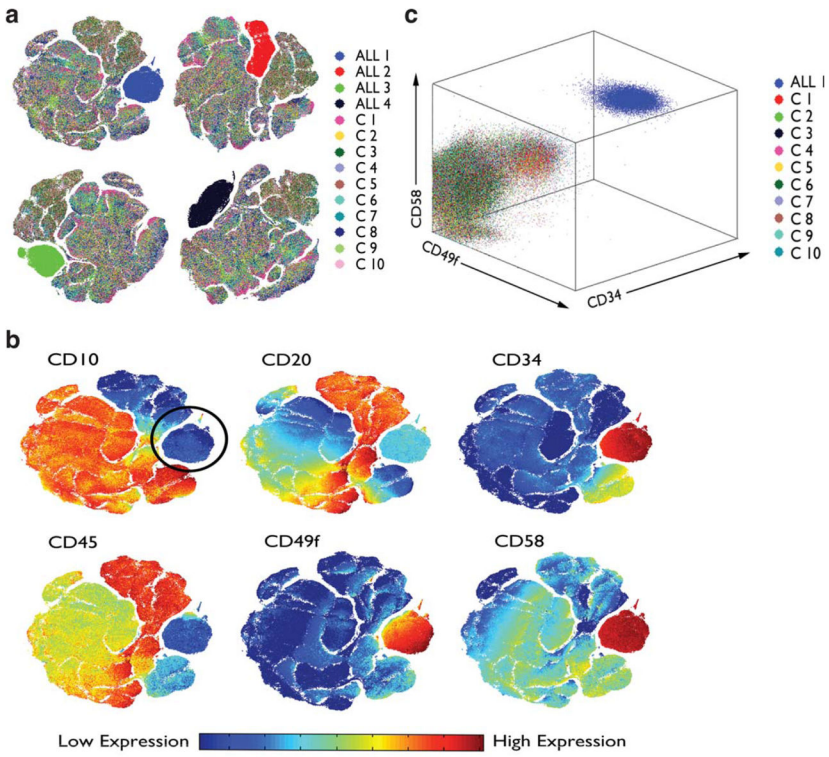

Fig. 2.

Leukemic B-lymphoblast populations occupy different regions of the viSNE map from normal bone marrow B cells. Panel **A**: viSNE maps for leukemic B-lymphoblast populations from four different patients (ALL 1 through ALL 4) run concurrently with identically stained B cells from 10 control bone marrow samples (C1 through C10). As in Figure 1, each cell is colored to reflect sample identity. Panel **B**: viSNE map of leukemic B-lymphoblasts from ALL 1 and B cells from 10 control bone marrow samples colored for antigen expression (tube B). The location of the leukemic B cells is shown by the circle. Panel **C**: Data from Panel (B) are shown in a conventional 3-dimensional dot plot, in which the axes represent expression of the three most discriminatory antigens. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

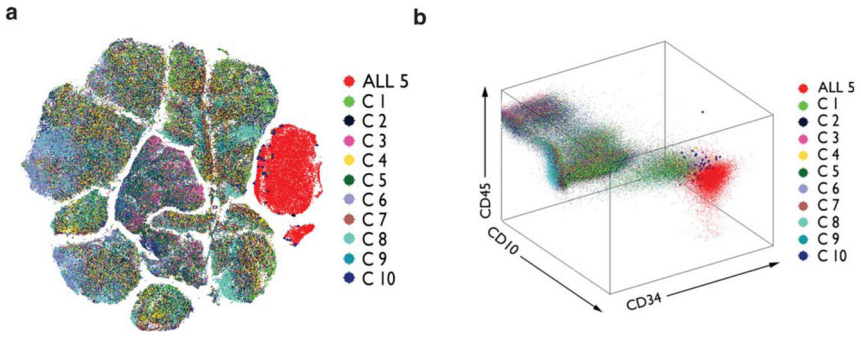

Fig. 3.

Normal bone marrow B cells may display partial immunophenotypic overlap with leukemic B-lymphoblast populations. Panel **A**: viSNE map showing co-localization of rare normal B cells from several different control samples with a single leukemic B-lymphoblast population. Panel **B**: Data from Panel (A) are shown in a conventional 3-dimensional dot plot, illustrating the phenotypic similarity between the rare normal B cells that co-localize with the leukemic B-lymphoblasts in the viSNE map. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

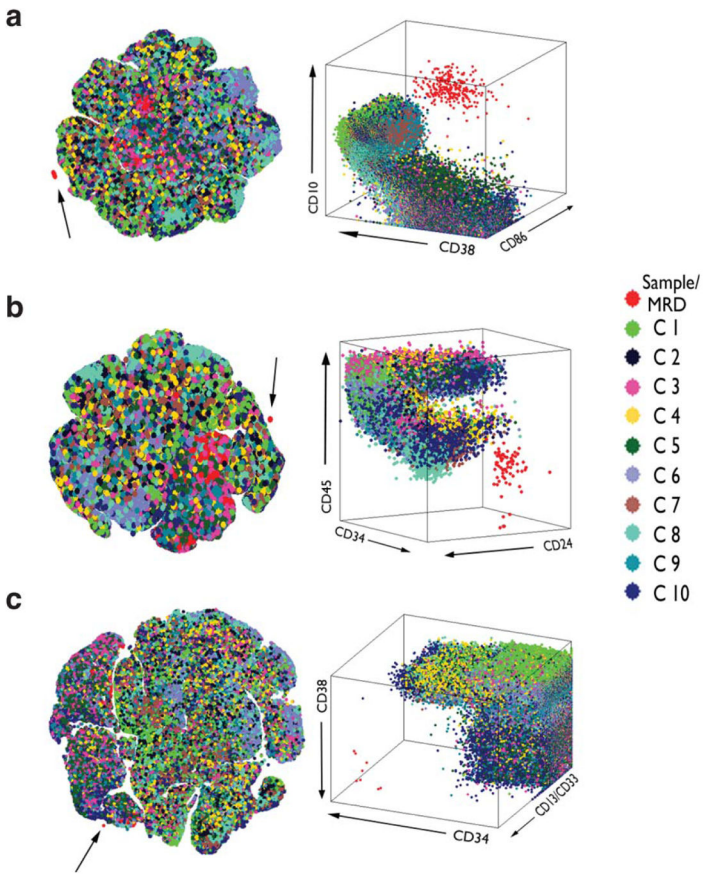

Fig. 4.

viSNE detects rare populations with abnormal immunophenotypes consistent with MRD. Left: viSNE maps from three different patients' samples in which MRD had been detected by manual gating (Panel (**A**) MRD case 10, tube C; Panel (**B**) MRD case 4, tube C; Panel (**C**) MRD case 3, tube A). In each panel, the post-treatment sample containing the candidate MRD population is identified in red. In each case, most of the B cells from the post-treatment sample (red) co-localize with normal B cells from the 10 control samples (each identified with a unique color as in the legend). In addition, an outlier population is detected in each of the post-treatment samples (arrows), which occupy a region of the map different from the normal B cells. Right: The outlier population from each patient's post-treatment sample (red) is plotted against normal B cells from the 10 control samples (again identified with a unique color as in the legend) in conventional 3-dimensional dot plots, in which the axes represent expression of the three most discriminatory antigens. In each case, the outlier population detected by viSNE displays an abnormal composite immunophenotype consistent with MRD. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

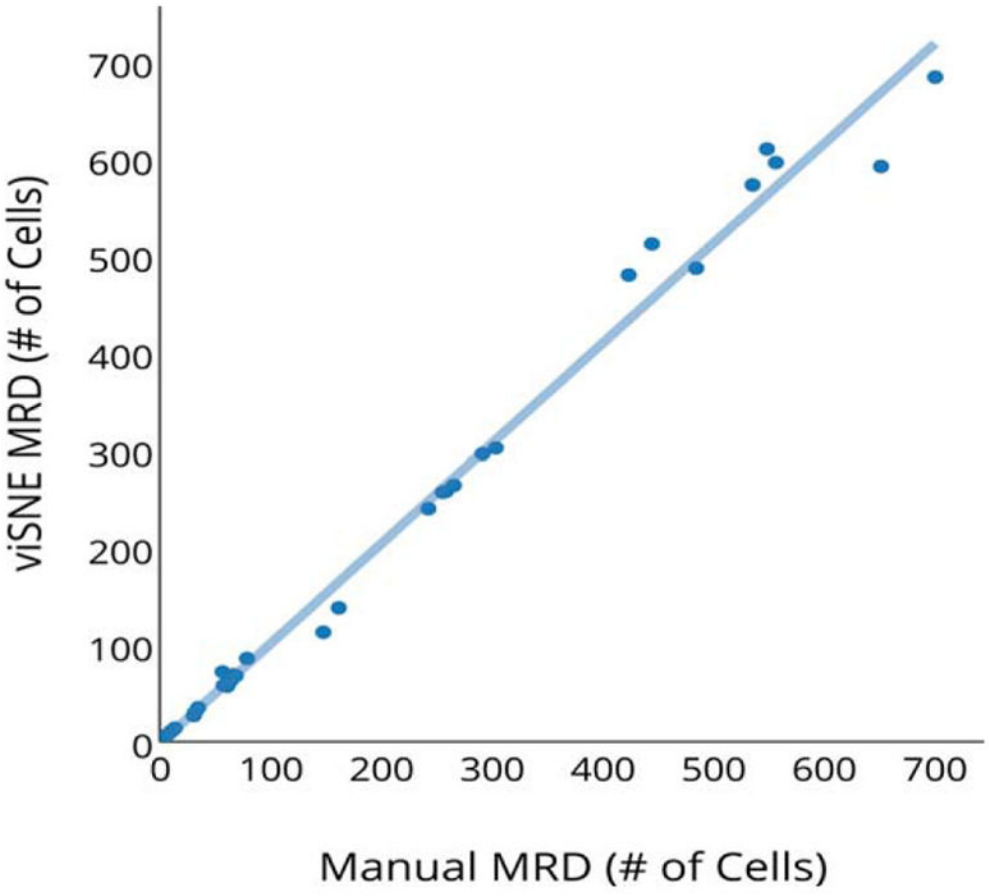

Fig. 5.

Correlation between the number of putative MRD cells identified by viSNE and conventional manual gating ($r^2 = 0.99$; two-tailed $P < 0.0001$). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

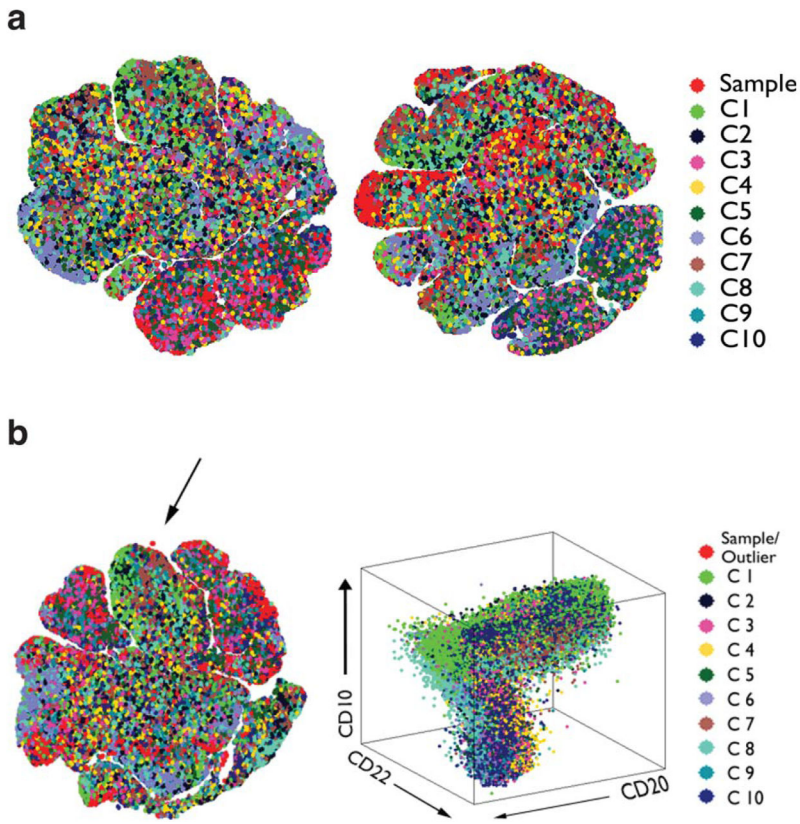

Fig. 6.

Specificity of abnormal population identification by viSNE. Panel **A**: Representative examples of viSNE maps from a patient with no history of B lymphoblastic leukemia (left, negative control 2, tube A) and a patient with B lymphoblastic leukemia after treatment in whom no MRD was detectable by manual gating (right, no detectable disease in bone marrow 7, tube A). No candidate MRD populations are identified. Panel **B**: Left: A candidate outlier population in the viSNE map from a patient with B lymphoblastic lymphoma but no detectable leukemic B-lymphoblast population in bone marrow by manual gating (no detectable disease in bone marrow 9, tube A). Right: In a conventional 3-dimensional dot plot, the candidate population (red) is indistinguishable from normal bone marrow B cells, and cannot be confirmed as MRD. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 1

Number of Cells in MRD Populations as Identified by Manual Gating and viSNE

MRD sample	Tube A manual	Tube A viSNE	Tube B manual	Tube B viSNE	Tube C manual	Tube C viSNE
1	557	596	536	573	485	487
2	4	ND	6	6	7	7
3	8	8	6	6	4	ND
4	61	57	61	61	69	68
5	22	ND	41	ND	14	14
6	549	610	445	512	424	480
7	256	257	292	296	304	302
8	63	61	58	58	67	69
9	694	ND	701	684	652	592
10	243	240	259	258	266	264
11	8	ND	19	ND	11	11
12	79	86	65	64	57	72
13	31	27	32	30	35	35
14	148	113	162	138	103	ND

ND: not detected.