METHODS ARTICLE

Visualizing Intrapopulation Hematopoietic Cell Heterogeneity with Self-Organizing Maps of SIMS Data

Vahid Mirshafiee, PhD,^{1,*} Brendan A. C. Harley, ScD,^{1,2} and Mary L. Kraft, PhD^{1,3,4}

Characterization of the heterogeneity within stem cell populations, which affects their differentiation potential, is necessary for the design of artificial cultures for stem cell expansion. In this study, we assessed whether self-organizing maps (SOMs) of single-cell time-of-flight secondary ion mass spectrometry (TOF-SIMS) data provide insight into the spectral, and thus the related functional heterogeneity between and within three hematopoietic cell populations. SOMs were created of TOF-SIMS data from individual hematopoietic stem and progenitor cells (HSPCs), lineage-committed common lymphoid progenitors (CLPs), and fully differentiated B cells that had been isolated from murine bone marrow via conventional flow cytometry. The positions of these cells on the SOMs and the spectral variation between adjacent map units, shown on the corresponding unified distance matrix (U-matrix), indicated the CLPs exhibited the highest intrapopulation spectral variation, regardless of the age of the donor mice. SOMs of HSPCs, CLPs, and B cells isolated from young and old mice using the same surface antigen profiles revealed the HSPCs exhibited the most age-related spectral variation, whereas B cells exhibited the least. These results demonstrate that SOMs of single-cell spectra enable characterizing the heterogeneity between and within cell populations that lie along distinct differentiation pathways.

Keywords: intrapopulation heterogeneity, single-cell analysis, differentiation, imaging mass spectrometry, self-organizing maps, screening platforms

Introduction

MICROSCALE SCREENING PLATFORMS that contain microarrays of distinct microenvironments, each large enough to accommodate just a few cells, have been developed to minimize the number of rare hematopoietic stem cells (HSCs) that must be used for screening experiments.^{1–4} However, the small scale that makes these platforms so advantageous introduces a new complication: intrapopulation heterogeneity. Cell-to-cell heterogeneity within hematopoietic cell populations, ranging from the most primitive HSCs⁵ to differentiated myeloid and lymphoid cells,⁶ may be required to produce the bioactivities that are attributed to the whole population.⁷

The necessity for intrapopulation heterogeneity has serious implications on the expansion of hematopoietic stem and progenitor cells (HSPCs) in cultures for mechanistic

studies and therapeutic use.⁷ Small stem or progenitor cell screening platforms may not contain enough cells to display the full range of fate decisions exhibited by the native population. Conversely, large-scale HSC expansion may produce a cell population with a high proliferative activity at the expense of less proliferative, but functionally essential HSC subtypes.⁷ Information about intrapopulation heterogeneity is required to design artificial cultures that contain enough HSCs, and thus, cell-to-cell variability to recapitulate hematopoiesis and to interpret the responses of the small number of HSCs in each microenvironment on a microscale screening platform.

To probe the heterogeneity of a cell population, signatures of cell phenotype must be acquired from individual cells, and the variance between cells from the same population must be related to other populations. Recent reports have demonstrated that the compositional information

¹Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois.

²Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois.

³Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois.

⁴Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois.

*Current affiliation: Center for Environmental Implications of Nanotechnology (CEIN), California NanoSystems Institute (CNSI), University of California Los Angeles, Los Angeles, California.

acquired from individual, unlabeled cells with time-of-flight secondary ion mass spectrometry (TOF-SIMS) and Raman spectroscopy can be correlated to cell phenotype with multivariate analysis techniques.^{8–24} We have previously shown that murine bone marrow-derived hematopoietic cells could be accurately classified as HSPCs, common lymphoid progenitors (CLPs), or mature B cells by applying partial least-square discriminant analysis (PLS-DA) models of TOF-SIMS data from hematopoietic cells of known differentiation stage to their spectra.¹³ Although this PLS-DA approach enabled objectively identifying the differentiation stages of individual hematopoietic cells with location specificity, it yielded little insight into cell population heterogeneity.

Spectral analysis with another multivariate method, principal component analysis (PCA), is better suited for visualizing the extent of spectral similarities and differences between and within cell populations.^{9–16} However, PCA cannot capture nonlinear relationships between spectral features, and only a fraction of the spectral variance between individual samples can be visualized on a single two-dimensional (2D) plot.²⁵ Various multivariate analysis techniques can be used to model nonlinear relationships between variables,²⁶ and dendrograms can be used to depict the heterogeneity between samples with respect to all variables.²⁷ However, dendrograms that describe large numbers of samples can be difficult to interpret.

An alternative multivariate analysis technique, called a self-organizing map (SOM), enables visualizing the heterogeneity between samples with respect to linear and nonlinear spectral variance on a single 2D plot that is easy to interpret.^{28–31} However, the feasibility of using SOMs of single-cell spectra to investigate the heterogeneity between and within different hematopoietic cell populations along the B cell differentiation pathway has not been assessed.

In this study, we explore whether SOMs of our previously studied TOF-SIMS data of individual hematopoietic cells isolated from mouse bone marrow¹³ provide insight into the heterogeneity between and within three distinct hematopoietic cell populations along the B cell differentiation pathway: immature HSPCs, CLPs, and fully differentiated B cells. Because HSCs from young and old mice differ in their ability to recapitulate hematopoiesis despite expressing the same differentiation stage-associated cell surface antigens,^{32,33} we also probed whether age-related spectral variation can be detected within hematopoietic cell populations with SOMs.

Materials and Methods

Cell isolation and fixation

The data sets used in this report were compiled using previously reported TOF-SIMS data from individual primary hematopoietic cells that were isolated from the bone marrow of young (2–4 months old) and old (10 months old) C57BL/6 mice.¹³ The hematopoietic cell populations used were (1) HSPCs (Lin[−]Sca-1⁺c-Kit⁺, LSK), which include long-term HSCs, short-term HSCs, and multipotent progenitors; (2) lineage-committed CLPs (Lin[−]IL-7R α ⁺Sca-1^{med}c-Kit^{med}); and (3) mature B cells (B220⁺IgM⁺). The isolation of these cells using flow cytometry, sample preparation, and TOF-SIMS analysis was previously reported.¹³

Briefly, the following antibody cocktails (eBioscience, San Diego, CA) were used to identify each cell population: (1) phycoerythrin (PE)-conjugated Sca-1, allophycocyanin (APC)-conjugated c-Kit, and a fluorescein isothiocyanate (FITC)-conjugated Lineage (Lin) cocktail (CD5, B220, Mac-1, CD8 α , Gr-1, Ter-119) for HSPCs; (2) PE-conjugated Sca-1, APC-conjugated c-Kit, PE-Cy 7-conjugated IL-7R α , and the aforementioned FITC-conjugated Lin cocktail for CLPs; and (3) FITC-conjugated B220 and eFluor 450-conjugated IgM for mature B cells. Flow cytometry was performed on a BD FACS Aria II flow cytometer. Cells from each population were plated on polylysine-coated silicon substrates (Ted Pella, Redding, CA) and were chemically preserved with 4% glutaraldehyde and 0.4% osmium tetroxide.

Acquisition of TOF-SIMS data sets

Cells were analyzed using a PHI Trift-III TOF-SIMS instrument (Physical Electronics, Inc., Chanhassen, MN). The ¹⁹⁷Au⁺ liquid ion gun was operated at 22 kV, and positive ion mass spectral images were acquired in unbunched mode using a total primary ion dose of 3 × 10¹³ ions/cm². Unit mass binned single-cell spectra were extracted from these images. Peaks below 50 *m/z* and above 300 *m/z* were removed from the unit mass binned single-cell spectra because they are either not chemically specific or not reproducibly detected with our instrumentation, respectively. Peaks associated with surface contaminants (*m/z* 73, 133, 147, 207, 221, and 281) and those that are not produced by common lipid species¹² or amino acids¹⁶ were also removed from the spectra.

Five data sets were compiled of the filtered spectra from (1) hematopoietic cells harvested from young mice; (2) hematopoietic cells harvested from old mice; (3) B cells from old and young mice; (4) CLPs from old and young mice; and (5) HSPCs from old and young mice. The average normalized spectra from all the B cells, CLPs, and HSPCs harvested from young mice and old mice are shown in Supplementary Figure S1 (Supplementary Data are available online at www.liebertpub.com/tec).

SOM construction

SOMs were constructed of the normalized and autoscaled TOF-SIMS data sets using the SOM Toolbox (v.2.0. www.cis.hut.fi/projects/somtoolbox/download) run in MATLAB (v.7.11.0.584 R2010b and v.7.14.0.739 R2012a, MathWorks, Inc., Natick, MA). The variables for map construction, initialization, and the training process were selected and defined automatically by the SOM Toolbox.

Briefly, an SOM consists of an array of hexagons, called map units. In our SOMs, each map unit represented a mass spectrum, where all the mass spectra contain the same mass peaks, but the relative intensities of these peaks were unique to a single map unit. Using the SOM Toolbox, each SOM was developed, or self-organized, by assigning each cell to the map unit that represented the mass set of peak intensities that most closely matched its own spectrum; this map unit is called the best matching unit (BMU). Then, the peak intensities in the spectrum represented by this BMU were adjusted so that they more closely matched the cell spectrum. The peak intensities in the spectrum represented by

FIG. 1. (a) Combined hit histogram shows the BMUs for the B cells (red), CLPs (green), and HSPCs (blue) isolated from young mice. The size of the colored hexagon encodes for the number of hits for each cell type. The map units are labeled with their row number and column letter. Individual hit histograms show the BMUs for the B cells (b), CLPs (c), and HSPCs (d) that were isolated from young mice. The numbers on the colored hexagons specify the number of times the map unit was the BMU for the indicated cell type. Map units that were not a BMU for the indicated cell type are labeled with their coordinates. The color-coded hexagons in the resultant U-matrix (e) represent the dissimilarity between an individual map unit and each of its neighbors, and the average dissimilarity of each map unit to all of its neighbors. Units that represent the average dissimilarity of a map unit to all of its neighbors are labeled with the letter of the column where the corresponding map unit is located in the hit histogram. Units labeled with a line (–, /, or \) represent the dissimilarity between the two map units located at each end of the line. Row numbers that designate the row where the corresponding map unit is located in the hit histogram are given on the left side of the U-matrix; row labels consisting of two row numbers separated by a dash indicate that this row represents the dissimilarity between map units in the two indicated rows (i.e., 1–2 signifies the dissimilarity between map units in row 1 and 2). BMU, best matching unit; CLP, common lymphoid progenitor; HSPC, hematopoietic stem and progenitor cell.

each map unit adjacent to the BMU were also adjusted so that they were more similar to the neighboring BMU.^{28,29} This process was reiterated until the mass spectra represented by the map units reflected the distribution of peak intensities in the sample spectra, and map units with similar spectra were positioned closer together in the SOM than those with dissimilar spectra, which were further apart.²⁹

A hit histogram was constructed that shows the number of times each map unit was the BMU for a specific type of cell. To facilitate locating specific regions of the SOM, the map units that were not the BMU for a cell were labeled with their coordinates, which consist of a row number and column letter. A corresponding U-matrix was constructed to show the degree of dissimilarity between the map units, and thus, the cells assigned to them. The U-matrix contains color-coded hexagons that represent the average degree of dissimilarity between each map unit in the hit histogram and all of its neighboring map units, plus additional hexagons that represent an individual map unit's dissimilarity to each of its neighbors.²⁹ Consequently, a hit histogram that has n rows and m columns of map units has a U-matrix that has $2n-1$ rows and $2m-1$ columns of hexagons.

The map units in the U-matrix are labeled with a column letter showing the average dissimilarity between the map unit in the hit histogram that has the same row number and column letter and all of its neighbors. For example, unit B in row 2 of the U-matrix shown in Figure 1e represents the

average dissimilarity between map unit 2B in the hit histogram (Fig. 1a) and of its neighboring map units. Units labeled with a line (–, /, or \) show the dissimilarity between the two map units in the hit histogram that the map units on either side of the line correspond to. For example, in Figure 1e, the leftmost hexagon in row 2–3 represents the spectral dissimilarity between map units 2A and 3A in the hit histogram (Fig. 1a).

For each model, the mean quantitation error (mqe), which is the average distance between each cell spectrum and that of its BMU, was used to assess how well each cell spectrum matched its BMU. Each SOM presented herein had a low topographical error (tge), which is the proportion of the cells whose BMU was not adjacent to its second BMU, indicating the SOM accurately depicts the spectral similarities and differences between the cell spectra.²⁹ The mqe and tge for each SOM are listed in Table 1.

Results

A SOM was constructed of the previously reported TOF-SIMS data from 20 B cells, 20 CLPs, and 14 HSPCs that were harvested from the bone marrow of young (2 to 4 months old) C57BL/6 mice.¹³ Each map unit in the SOM represents a unique set of intensities for every peak in the spectra, and each cell was assigned to the unit that best matched its individual spectrum. The number of times that a

TABLE 1. MEAN QUANTITATION ERROR AND TOPOGRAPHICAL ERROR CALCULATED FOR EACH SELF-ORGANIZING MAP

	Hematopoietic cells, young mice	Hematopoietic cells, old mice	B cells, old and young mice	CLPs, old and young mice	HSPCs, old and young mice
mqe	5.78	6.70	7.13	6.16	6.16
tge	0.019	0.012	0.000	0.000	0.000

CLPs, common lymphoid progenitors; HSPC, hematopoietic stem and progenitor cell; mqe, mean quantitation error; tge, topographical error.

map unit was the BMU for a B cell, CLP, or HSPC is shown in red, green, and blue, respectively, on the multipopulation and individual hit histograms (Fig. 1a–d). The distance between the cells on the hit histogram encodes for their spectral similarities, where cells on neighboring map units have spectra that are more similar than those on units that are far apart.^{28–31} Based on the tge of 0.019 for this SOM (Table 1), only one of the 54 cells was assigned to a BMU that was not adjacent to its second-best unit, which indicates the SOM was an accurate depiction of the similarities and differences in the cell spectra. The degree of dissimilarity between individual map units is shown on the U-matrix (Fig. 1e).

Because cells from the same population have similar surface chemistries, and these cell surface chemistries encode for differentiation stage,¹³ we expected that cells belonging to the same population would be assigned to the same or adjacent map units. The hit histograms shown in Figure 1 are largely consistent with this expectation. The BMUs from each of the three hematopoietic cell populations were largely segregated within nonoverlapping regions of the map space, and multiple cells from the same population were often assigned to the same map unit (Fig. 1a–d). However, one HSPC was assigned to the same map unit as a B cell, (unit 3E, Fig. 1b, d) and this map unit was adjacent to the BMUs of other B cells, but not the clustered HSPCs on map units 3A, 3B, 3C, 4A, 4B, 4C, 5B, and 5C in Figure 1d.

The distance between the one HSPC that had the same BMU as a B cell (unit 3E, Fig. 1) from the other HSPCs, and its proximity to other B cells suggested that this HSPC had spectral features and surface chemistries that were more similar to B cells than to the other HSPCs. Furthermore, the separation between the clustered HSPCs and the B cells on the hit histogram, and the orange and red hexagons in row 2-3 of the U-matrix (Fig. 1e), which signify high spectral dissimilarity between the B cells and clustered HSPCs, indicate that no other HSPCs were spectrally similar to B cells. To explore the spectral variation between the HSPC with the same BMU as a B cell and the nearby B cells and HSPCs in the SOM, we compared the spectra represented by unit 3E to those of the closest map units that were the BMUs for multiple HSPCs (unit 4C, Fig. 1) and multiple B cells (unit 2E,

Fig. 1). Figure 2 shows the mass spectra represented by the BMU for both HSPC and B cell (unit 3E, Fig. 1), the neighboring unit that was the BMU for two B cells (unit 2E, Fig. 1), and the unit that was the BMU for four HSPCs (unit 4C). Table 2 lists the biomolecules that are known to produce the mass fragments whose intensities seemed to vary the most between the map units.^{12,16}

Peaks m/z 55, 57, 69, 77, 81, 91, 93, 95, 107, 117, 128, 143, 159, 206, and 219 were more intense in the spectrum for unit 4C, the BMU for four HSPCs, whereas peaks m/z 58, 59, 60, 70, 72, 166, 184, and 224 were more intense in the spectrum for unit 2E, the BMU for two B cells. Most of the aforementioned peaks may be produced by multiple biomolecular building blocks during TOF-SIMS analysis.^{12,16} However, peaks m/z 77, 107, 128, 143, and 159, which were more intense in the spectra for unit 4C, are each reported to be associated with only a single amino acid, namely phenylalanine, tyrosine, serine, and tryptophan.^{12,16} This suggests these amino acids were more abundant on the HSPCs on unit 4C than the B cells on unit 2E. All of the peaks that were more intense in the spectra represented by unit 2E were associated with lipid fragments,¹² which may suggest that the lipid to protein ratios on the B cells assigned to unit 2E were higher than on the HSPCs assigned to unit 8E. The spectrum of unit 3E, the BMU for both a B cell and an HSPC, had peak intensities that were often closer to those in the spectrum for unit 2E, the BMU for the two B cells. This suggests the cells on unit 3E had lipid to protein ratios that were more similar to the B cells on unit 2E than the HSPCs on unit 4C.

The U-matrix sheds light on the relative magnitude of the variation in spectral features, between and within the distinct hematopoietic cell populations. The dark to medium blue colored hexagons in rows 1, 1-2, and 2 in the U-matrix denote low spectral dissimilarity (high similarity) between the BMUs of the B cells (Fig. 1e). Therefore, the surface chemistries of the individual B cells that were isolated from the young mice were relatively similar to each other. In comparison, the blue to green colors of the hexagons in the U-matrix that encode for the dissimilarity between the clustered HSPCs denote low to intermediate spectral dissimilarity. This combined with the assignment of one HSPC to a map unit that was separated from the

FIG. 2. Normalized and filtered mass spectra for specific map units in the SOM of the cells harvested from young mice. The spectrum for unit 3E, which was the BMU for one HSPC and one B cell, is shown in purple. The red spectrum corresponds to adjacent map unit 2E, which was the BMU for two B cells. The blue spectrum is for unit 4C, which was the BMU for four HSPCs. SOM, self-organizing map.

TABLE 2. LIST OF POSITIVE-ION MASS PEAKS WITH HIGH INTENSITIES OR VARIATIONS IN THE SPECTRA OF MAP UNITS IN THE SELF-ORGANIZING MAPS, AND THE BIOMOLECULES KNOWN TO PRODUCE THESE FRAGMENT IONS

Mass (m/z)	Related biomolecule
55	Lipid fragment ($C_4H_7^+$), arginine, lysine, cysteine, valine, leucine, histidine
57	Lipid fragment ($C_4H_9^+$), valine, serine, alanine, arginine, cysteine, aspartic acid, leucine, threonine, glycine
58	Lipid fragment ($C_3H_8N^+$), isoleucine
59	Lipid fragment ($^{13}C^{12}C_2H_8N^+$), arginine, valine
60	Lipid fragment ($C_3H_{10}N^+$), serine
67	Lipid fragment ($C_5H_7^+$)
69	Lipid fragment ($C_5H_9^+$), isoleucine, histidine, lysine
70	Lipid fragment ($C_4H_8N^+$), proline, leucine, glutamic acid, asparagine, arginine
72	Lipid fragment ($C_4H_{10}N^+$), valine
77	Phenylalanine
81	Lipid fragment ($C_6H_9^+$), alanine
83	Lipid fragment ($C_6H_{11}^+$)
86	Lipid fragment ($C_5H_{12}N^+$), isoleucine, leucine
91	Lipid fragment ($C_7H_7^+$), serine, phenylalanine, methionine
93	Lipid fragment ($C_7H_9^+$)
95	Lipid fragment ($C_7H_{11}^+$), histidine
104	Lipid fragment ($C_5H_{14}NO^+$), lysine
107	Tyrosine
117	Tryptophan, proline
128	Serine
143	Tryptophan
148	Lipid fragment ($C_5H_{11}NPO_2^+$), glutamic acid
159	Tryptophan
166	Lipid fragment ($C_5H_{13}NPO_3^+$), methionine, phenylalanine
184	Lipid fragment ($C_5H_{15}NPO_4^+$)
205	Tryptophan
206	Lipid fragment ($C_5H_{14}NPO_4Na^+$)
219	Aspartic acid, cysteine
224	Lipid fragment ($C_7H_{15}NPO_5^+$)

other HSPCs indicates the HSPCs had more heterogeneous surface chemistries than the B cells. Finally, the CLP's high spread on the map space and the high spectral dissimilarity between the BMUs of many CLPs denoted by the yellow, orange, and red hexagons at the bottom right side of the U-matrix (Fig. 1e) indicate the surface chemistries of the CLPs varied more than the HSPCs. Therefore, the intrapopulation heterogeneity in cell surface chemistry was lowest for the B cells and highest for the CLPs isolated from young mice.

The SOMs of the TOF-SIMS data from 30 B cells, 25 CLPs, and 29 HSPCs isolated from five relatively old (10 months old) mice are shown in Figure 3. The tge of 0.012 for this SOM (Table 1) indicates that only one of the 84 cells on the SOM was assigned to a BMU that was not adjacent to its second-best unit. Thus, this SOM accurately depicted the similarities and differences in the cell spectra.

The trends in the relative heterogeneity within the B cell, CLP, and HSPC populations from old mice were similar to

those observed for the young mice. First, the B cells harvested from old mice were the least dispersed population on the hit histograms, and the U-matrix indicated low to intermediate dissimilarity between their BMUs (Fig. 3e). Second, the HSPCs harvested from old mice were more dispersed on the hit histogram than the B cells (Fig. 3a), and the U-matrix indicated the BMUs for the HSPCs had higher spectral dissimilarity than those of the B cells (Fig. 3e). In addition, two HSPCs were assigned to the same map unit as a B cell (unit 5D, Fig. 3a) in the SOM for the old mice. Comparison of the mass spectra represented by the units that were the most common BMUs for only B cells (unit 1E, Fig. 3) and HSPCs (unit 8E, Fig. 3) revealed similar trends in the peak intensities that were observed for the young mice (Supplementary Fig. S2). This suggests that like the cells from young mice, the SOM for the cells from the old mice indicated the ratio of lipids to proteins was higher on the surfaces of the B cells on unit 1E than the HSPCs on unit 8E. However, the one B cell and two HSPCs with the same BMU (unit 5D, Fig. 3) had peak intensities, and thus surface chemistries, that were between those of the B cells on unit 1E and the HSPCs on unit 8E (Supplementary Fig. S3). Finally, the CLPs were the most dispersed population on the hit histograms of cells from old mice (Fig. 3c), and the U-matrix indicated higher dissimilarity between the BMUs of the CLPs relative to that between the BMUs of the B cells or HSPCs harvested from old mice (Fig. 3e).

Unlike the SOM of the cells from young mice, multiple map units in the SOM of the cells from old mice were the BMUs for both CLPs and HSPCs (unit 1A, 8C, 10A, 10C, and 10E in Figs. 3c and 2d), and more CLPs had BMUs that were adjacent to those of HSPCs. The higher number of HSPCs and CLPs with the same map units in the SOM of cells harvested from relatively old mice may suggest an age-related increase in the fraction of HSPCs and CLPs at intermediate stages of differentiation, or the surface chemistries of these two populations become more similar as the mice age. Plots of the mass spectra represented by the BMUs for both HSPCs and B cells, and those of the units that were the BMUs for only CLPs and only HSPCs (units 5A and 8E in Fig. 3, respectively) provided insight into the spectral features and surface chemistries that varied between these cells (see Supplementary Fig. S3 for details).

HSCs exhibit age-related differences in functional capacity despite an absence of changes in surface antigen expression.^{32,33} To explore whether SOMs could be used to detect age-related differences in surface chemistries within each population, we used the spectra from cells isolated from young and old mice using the same surface antigen profiles to create an SOM for each hematopoietic cell population (Fig. 4). All three SOMs had a tge of zero (Table 1), indicating they accurately captured the similarities and differences in the cell spectra.

Of these three cell populations, the B cells from the young and old mice were the most interdispersed on the map space. Four map units were the BMUs for B cells from both young and old mice (units 1C, 1E, 5B, and 7D, Fig. 4a). However, the BMUs of the B cells from the young mice were still more abundant on the right side of the hit histogram (Fig. 4b). The U-matrix also occasionally indicated higher spectral dissimilarity between the BMUs for B cells from young and old mice (Fig. 4d).

FIG. 3. (a) Combined hit histogram for all three hematopoietic cell populations isolated from old mice. The sizes of the red, green, and blue hexagons indicate the number of times the map unit was the BMU for a B cell, CLP, and HSPC, respectively. The map units that are not a BMU are labeled with their row number and column letter. Individual hit histograms show the BMUs for the B cells (b), CLPs (c), and HSPCs (d) that were isolated from old mice. The colored hexagons are labeled with the number of times the map unit was the BMU for the indicated cell type. Empty map units are labeled with their coordinates. (e) The magnitude of dissimilarity between neighboring map units in the hit histogram is encoded by hexagon color in the U-matrix. See the Figure 1 caption for an explanation of markings on the U-matrix.

FIG. 4. Combined hit histograms show the BMUs for the (a) B cells, (e) CLPs, and (i) HSPCs isolated from young and old mice, where the cells from the old mice are shown in a primary color and the cells from the young mice are represented by the darker shade. The individual hit histograms show the BMUs for the B cells from (b) young and (c) old mice, CLPs from (f) young and (g) old mice, and HSPCs from (j) young and (k) old mice. The magnitude of dissimilarity between neighboring map units in the hit histogram for the (d) B cells, (h) CLPs, and (l) HSPCs is encoded by hexagon color in the corresponding U-matrix. See the caption for Figure 1 for an explanation of markings on the U-matrix.

The hit histograms for the SOMs of the CLPs and HSPCs showed more distinct separation between the cells harvested from young and old mice (Fig. 4e–l). Only two map units were the BMUs for CLPs from both young and old mice (units 1E and 7A, Fig. 4e). The BMUs of only a few CLPs from old mice were directly adjacent to the BMUs for CLPs from young mice, and the U-matrix showed that these neighboring BMUs usually had higher spectral dissimilarity (Fig. 4h). No map units in the hit histogram for the HSPCs were the BMUs for HSPCs from both young and old mice. The BMUs of only two HSPCs from young mice (units 6B and 7D, Fig. 4j) were adjacent to the BMUs of HSPCs from old mice (units 5C and 7E, Fig. 4k), and the U-matrix indicated high spectral variation between these BMUs (Fig. 4l).

Altogether, these findings suggest that the surface chemistries of the cells isolated from old and young mice differed the least for the B cells and the most for the HSPCs.

Discussion

Knowledge of the cell-to-cell heterogeneity within hematopoietic cell populations is necessary to design microscale screening platforms that contain enough HSCs to exhibit the full range of cell fate decisions that occur in the body.⁷ Methods to accurately identify the differentiation stages of individual cells in microscale screening platforms have been developed, but these techniques yield little insight into the extent of heterogeneity between and within different cell populations. In this study, we assessed whether SOMs of single-cell TOF-SIMS data provide information about the spectral and chemical heterogeneity that ultimately contribute to the functional heterogeneity between and within three hematopoietic cell populations. By creating SOMs of single-cell spectra from B cells, CLPs, and HSPCs, we were able to visualize the extent of the spectral variation between individual cells and identify the spectral features that were responsible for this heterogeneity. Databases of the mass fragments produced by specific biomolecules during TOF-SIMS analysis^{12,16} enabled translating these spectral differences into information about the differences in surface chemistries that may have produced them.

Although SOMs provide insight into the spectral features, and thus surface chemistries that vary between cells of interest, the ease at which the extent of the spectral heterogeneity between and within hematopoietic cell populations can be visualized is a major advantage of this approach. The SOM models presented herein revealed the intrapopulation spectral heterogeneity was the lowest for B cells and the highest for CLPs, regardless of the age of the donor mice. The cell markers used for isolation produce a CLP population consisting of six CLP subtypes with differing lineage potentials, gene expression profiles, and surface chemistries.^{34,35} In comparison, the cell markers used to isolate HSPCs for this study produce a mixture of long-term stem cells, short-term stem cells, and multipotent progenitors,^{36,37} and these subpopulations may consist of additional cell subtypes.^{38,39}

Based on this diversity within the HSPC populations isolated for this study, one might have expected that the HSPCs would have exhibited higher spectral heterogeneity than the CLPs, which is contrary to our findings. One possibility is that the HSPC data sets used to create each SOM were not large enough to contain the rarest HSPC subtypes, resulting in lower intrapopulation spectral variation than that

exhibited by the entire native HSPC population. To explore this possibility, cell surface markers and sorting conditions could be used to isolate cells from each HSPC subtype identified to date, producing an HSPC population with spectral variation characteristic of the native HSPC population. An SOM could then be used to compare the extent of spectral heterogeneity within the complete HSPC population and the smaller HSPC or CLP data sets used herein as a proof of concept. Alternatively, the number of HSPC cells that need to be isolated to capture spectral heterogeneity within the native HSPC population might be identified by systematically increasing the number of HSPCs that are used to develop the SOM until no further increases in spectral heterogeneity, as assessed by the proximities of their BMUs on the map, are observed. This approach could also be used to identify the minimum number of HSPCs that must be present in a microscale screening platform to recapitulate spectral variation, and thus, related functional capabilities that are required to recapitulate hematopoiesis.

SOMs of single-cell TOF-SIMS data also enabled assessing the relative magnitude of the age-related spectral variance within the B cell, CLP, and HSPC populations. Comparison of the SOMs of each cell population revealed that surface chemistries of B cells harvested from young and old mice differed the least, whereas surface chemistries of HSPCs harvested from young and old mice differed the most. Previous investigations demonstrate that murine HSCs experience an age-related decline in differentiation potential without appreciable changes in known differentiation-related cell surface antigens.^{18,19} The high variation in the cell surface chemistry detected between HSPCs harvested from young and old mice may be indicative of these age-related functional changes. Although complementary metabolomic studies are required to test this hypothesis, altogether, this work demonstrates that SOMs of single-cell spectra enable assessing the heterogeneity within cell populations that lie along a differentiation pathway.

Conclusion

SOMs of single-cell TOF-SIMS data provide insight into the spectral, compositional, and thus, functional heterogeneity between and within cell populations. By using SOMs of single-cell spectra to characterize the variability between and within stem and progenitor cell populations, a better understanding of the functional significance of intrapopulation heterogeneity might be obtained.

Acknowledgments

This work was supported by the National Institute of Biomedical Imaging and Bioengineering under Award Number R21 EB018481. The authors acknowledge additional funding provided by the Dept. of Chemical and Biomolecular Engineering at the University of Illinois.

Disclosure Statement

No competing financial interests exist.

References

1. Ranga, A., and Lutolf, M.P. High-throughput approaches for the analysis of extrinsic regulators of stem cell fate. *Curr Opin Cell Biol* **24**, 236, 2012.

2. Martin, T.A., Caliri, S.R., Williford, P.D., Harley, B.A., and Bailey, R.C. The generation of biomolecular patterns in highly porous collagen-GAG scaffolds using direct photolithography. *Biomaterials* **32**, 3949, 2011.
3. Mei, Y., Saha, K., Bogatyrev, S.R., *et al.* Combinatorial development of biomaterials for clonal growth of human pluripotent stem cells. *Nat Mater* **9**, 768, 2010.
4. Luo, W., and Yousaf, M.N. Tissue morphing control on dynamic gradient surfaces. *J Am Chem Soc* **133**, 10780, 2011.
5. Schroeder, T. Hematopoietic stem cell heterogeneity: subtypes, not unpredictable behavior. *Cell Stem Cell* **6**, 203, 2010.
6. Galli, S.J., Borregaard, N., and Wynn, T.A. Phenotypic and functional plasticity of cells of innate immunity: macrophages, mast cells and neutrophils. *Nat Immunol* **12**, 1035, 2011.
7. Phinney, D.G. Functional heterogeneity of mesenchymal stem cells: implications for cell therapy. *J Cell Biochem* **113**, 2806, 2012.
8. Ilin, Y., and Kraft, M.L. Secondary ion mass spectrometry and Raman spectroscopy for tissue engineering applications. *Curr Opin Biotechnol* **31**, 108, 2015.
9. Barnes, C.A., Brison, J., Robinson, M., Graham, D.J., Castner, D.G., and Ratner, B.D. Identifying individual cell types in heterogeneous cultures using secondary ion mass spectrometry imaging with C60 etching and multivariate analysis. *Anal Chem* **84**, 893, 2012.
10. Baker, M.J., Brown, M.D., Gazi, E., Clarke, N.W., Vickerman, J.C., and Lockyer, N.P. Discrimination of prostate cancer cells and non-malignant cells using secondary ion mass spectrometry. *Analyst* **113**, 175, 2008.
11. Thompson, C.E., Fletcher, J.S., Goodacre, R., Henderson, A., Lockyer, N.P., and Vickerman, J.C. ToF-SIMS studies of *Bacillus* using multivariate analysis with possible identification and taxonomic applications. *Appl Surf Sci* **252**, 6719, 2006.
12. Anderton, C.R., Vaezian, B., Lou, K., Frisz, J.F., and Kraft, M.L. Identification of a lipid-related peak set to enhance the interpretation of TOF-SIMS data from model and cellular membranes. *Surf Interface Anal* **44**, 322, 2012.
13. Frisz, J.F., Choi, J.S., Wilson, R.L., Harley, B.A., and Kraft, M.L. Identifying differentiation stage of individual primary hematopoietic cells from mouse bone marrow by multivariate analysis of TOF-secondary ion mass spectrometry data. *Anal Chem* **84**, 4307, 2012.
14. Ilin, Y., Choi, J.S., Harley, B.A., and Kraft, M.L. Identifying states along the hematopoietic stem cell differentiation hierarchy with single cell specificity via Raman spectroscopy. *Anal Chem* **87**, 11317, 2015.
15. Ilin, Y., and Kraft, M.L. Identifying the lineages of individual cells in cocultures by multivariate analysis of Raman spectra. *Analyst* **139**, 2177, 2014.
16. Kulp, K.S., Berman, E.S., Kinze, M.G., *et al.* Chemical and biological differentiation of three human breast cancer cell types using time-of-flight secondary ion mass spectrometry. *Anal Chem* **78**, 3651, 2006.
17. Pascut, F.C., Kalra, S., George, V., Welch, N., Denning, C., and Nottingher, I. Non-invasive label-free monitoring the cardiac differentiation of human embryonic stem cells in vitro by Raman spectroscopy. *Biochim Biophys Acta* **1830**, 3517, 2013.
18. Chan, J.W., Lieu, D.K., Huser, T., and Li, R.A. Label-free separation of human embryonic stem cells and their cardiac derivatives using Raman spectroscopy. *Anal Chem* **81**, 1324, 2009.
19. Ghita, A., Pascut, F.C., Mather, M., Sottile, V., and Nottingher, I. Cytoplasmic RNA in undifferentiated neural stem cells: a potential label-free Raman spectral marker for assessing the undifferentiated status. *Anal Chem* **84**, 3155, 2012.
20. Jell, G., Nottingher, I., Tsigkou, O., *et al.* Bioactive glass-induced osteoblast differentiation: a noninvasive spectroscopic study. *J Biomed Mat Res Part A* **86A**, 31, 2008.
21. Tsai, T.-H., Short, M.A., McLean, D.I., Zeng, H., McElwee, K., and Lui, H. Label-free identification and characterization of murine hair follicle stem cells located in thin tissue sections with Raman micro-spectroscopy. *Analyst* **139**, 2799, 2014.
22. Kim, B.S., Lee, C.I., Christensen, J.E., Huser, T.R., Chan, J.W., and Tarantal, A.F. Growth, differentiation, and biochemical signatures of rhesus monkey mesenchymal stem cells. *Stem Cells Dev* **17**, 185, 2008.
23. Swain, R.J., Jell, G., and Stevens, M.M. Non-invasive analysis of cell cycle dynamics in single living cells with Raman micro-spectroscopy. *J Cell Biochem* **104**, 1427, 2008.
24. Swain, R.J., Kemp, S.J., Goldstraw, P., Tetley, T.D., and Stevens, M.M. Spectral monitoring of surfactant clearance during alveolar epithelial type II cell differentiation. *Biophys J* **95**, 5978, 2008.
25. Yin, H. Learning nonlinear principal manifolds by self-organising maps. In: Gorban, A.N., Kégl, B., Wunsch, D.C., and Zinovyev, A.Y., eds. *Principal manifolds for data visualization and dimension reduction*. Heidelberg, Berlin: Springer, 2008, pp. 68–95.
26. Martinez, W.L., and Martinez, A.R. Dimensionality reduction - nonlinear methods. In: Lafferty, J., Madigan, D., Murtagh, F., Smyth, P., eds. *Exploratory Data Analysis with MATLAB: Series in Computer Science and Data Analysis*. Boca Raton: CRC Press, 2005, Ch. 3, pp. 61–101.
27. Martinez, W.L., and Martinez, A.R. Visualizing clusters. In: Lafferty, J., Madigan, D., Murtagh, F., Smyth, P., eds. *Exploratory Data Analysis with MATLAB: Series in Computer Science and Data Analysis*. Boca Raton: CRC Press, 2005, Ch. 8, 233–257.
28. Kohonen, T. The self-organizing map. *Proc IEEE* **78**, 1464, 1990.
29. Cuss, C.W., and Gueguen, C. Analysis of dissolved organic matter fluorescence using self-organizing maps: mini-review and tutorial. *Anal Methods* **8**, 716, 2016.
30. Lloyd, G.R., Brereton, R.G., and Duncan, J.C. Self Organising Maps for distinguishing polymer groups using thermal response curves obtained by dynamic mechanical analysis. *Analyst* **133**, 1046, 2008.
31. Brereton, R.G. *Chemometrics for pattern recognition*. Chichester, UK: John Wiley & Sons, Ltd., 2009, Ch. 3, pp. 47–105.
32. Chambers, S.M., Shaw, C.A., Gatz, C., Fisk, C.J., Donehower, L.A., and Goodell, M.A. Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. *PLoS Biol* **5**, e201, 2007.
33. Warren, L.A., and Rossi, D.J. Stem cells and aging in the hematopoietic system. *Mech Ageing Dev* **130**, 46, 2009.
34. Mansson, R., Zandi, S., Welinder, E., *et al.* Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity. *Blood* **115**, 2601, 2010.

35. Zhang, Q., Esplin, B.L., Iida, R., *et al.* RAG-1 and Ly6D independently reflect progression in the B lymphoid lineage. *PLoS One* **8**, e72397, 2013.
36. Adolfsson, J., Borge, O.J., Bryder, D., *et al.* Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* **15**, 659, 2001.
37. Adolfsson, J., Mansson, R., Buza-Vidas, N., *et al.* Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295, 2005.
38. Challen, G.A., Boles, N.C., Chambers, S.M., and Goodell, M.A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF- β 1. *Cell Stem Cell* **6**, 265, 2010.
39. Beerman, I., Bhattacharya, D., Zandi, S., *et al.* Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc Natl Acad Sci U S A* **107**, 5465, 2010.

Address correspondence to:

Mary L. Kraft, PhD

Department of Chemical and Biomolecular Engineering

University of Illinois at Urbana-Champaign

208 Roger Adams Laboratory MC 712

600 S. Mathews Avenue

Urbana, IL 61801

E-mail: mlkraft@illinois.edu

Received: September 5, 2017

Accepted: March 15, 2018

Online Publication Date: May 7, 2018