Cancer Immunol Immunother (2013) 62:955–965
DOI 10.1007/s00262-013-1416-8

FOCUSSED RESEARCH REVIEW

# An introduction to mass cytometry: fundamentals and applications

**Scott D. Tanner · Vladimir I. Baranov · Olga I. Ornatsky · Dmitry R. Bandura · Thaddeus C. George**

Received: 3 October 2012 / Accepted: 11 March 2013 / Published online: 7 April 2013


**Abstract** Mass cytometry addresses the analytical challenges of polychromatic flow cytometry by using metal atoms as tags rather than fluorophores and atomic mass spectrometry as the detector rather than photon optics. The many available enriched stable isotopes of the transition elements can provide up to 100 distinguishable reporting tags, which can be measured simultaneously because of the essential independence of detection provided by the mass spectrometer. We discuss the adaptation of traditional inductively coupled plasma mass spectrometry to cytometry applications. We focus on the generation of cytometry-compatible data and on approaches to unsupervised multivariate clustering analysis. Finally, we provide a high-level review of some recent benchmark reports that highlight the potential for massively multi-parameter mass cytometry.

**Keywords** Mass cytometry · Multivariate cluster analysis · Hematopoiesis · Memory T cells · Multiplexed cell analysis · CIMT 2012

This paper is a Focussed Research Review based on a presentation given at the Tenth Annual Meeting of the Association for Cancer Immunotherapy (CIMT), held in Mainz, Germany, May 23–25, 2012. It is part of a CII series of Focussed Research Reviews and meeting report.

S. D. Tanner (✉) · V. I. Baranov
Department of Chemistry, University of Toronto,
80 St. George Street, Toronto, ON M5S 3H6, Canada
e-mail: sd.tanner@utoronto.ca

S. D. Tanner · V. I. Baranov · O. I. Ornatsky · D. R. Bandura
DVS Sciences Inc., #12-70 Esna Park Drive,
Markham, ON L3R 6E7, Canada

T. C. George
DVS Sciences, Inc., 639 North Pastoria Avenue,
Sunnyvale, CA 94085, USA

## Introduction

Mass cytometry is a novel adaptation of atomic mass spectrometry that enables uniquely high-dimensional flow cytometry applications. It employs stable isotopes of the transition elements as tags in the same manner that fluorophores have been used, and "reads" these with a mass spectrometer that has been developed for high-accuracy elemental analysis and isotope ratio measurements.

The inductively coupled plasma mass spectrometer (ICP-MS) has been the state-of-the-art tool for the determination of the elemental and isotopic composition of matter almost from the time of its implementation [1] and subsequent commercial launch in 1983. It has found ubiquitous use in the determination of environmental contamination by heavy (lead, uranium) or toxic (chromium, aluminum, tin, antimony) metals, the analysis of metals (arsenic, chromium, cadmium) in drinking water, clinical assay of blood (for thallium, lead), the determination of neutron-capture isotopes in nuclear fuel materials, and the detection of fast-effusing elements (lithium, potassium, calcium) in semiconductor production. Importantly, it was realized early that ICP-MS was a powerful means to independently measure the various isotopes of elements, which is of profound significance in geology (rubidium-strontium dating of rocks, prospecting and profiling mineralization through fingerprint rare earth elements) and more recently in provenance determination of rice, wines, juices, and fruits through determination of trace element isotope ratios that are characteristic of local

regions. Until very recently, this "flavor" of mass spectrometry has emphatically not been popularized in biologic applications, simply because its protagonists are interested in atomic composition, whereas the biologic analytical community attends to the molecular composition and interactions that drive the genesis of life and health. However, the ability to simultaneously and exquisitely resolve multiple isotopes, when these are configured to report on biomarkers, is well suited to meet the multiparametric demands common in biologic applications.

For various reasons (including for internal standardization of assays using isotope dilution mass spectrometry), a wide variety of elements can be obtained in the form of enriched isotopes. For example, samarium has six stable isotopes, and neodymium has seven stable isotopes, all of which are available in enriched form. The thirteen lanthanide elements are provided in at least 37 isotopes that have non-redundantly unique masses. Figure 1 provides the mass spectrum, recorded using the CyTOF® mass cytometer, of thirty of these enriched isotopes. The "resolution," frequently defined as the full width at half maximum intensity (FWHM), or at 10 or 1 % maximum, is not really the analytical issue, because the isotopes of the same element differ by a full atomic mass unit (or dalton, essentially the mass of a neutron). Of more importance for the atomic mass spectrometry community is the abundance sensitivity, defined as the residual signal owing to a dominant isotope at its neighbor (M + 1 and M − 1) masses, which in some respects emulates fluorescent overlap into neighboring photomultiplier detector channels. However, because of the importance of isotopic analysis over very large dynamic range, ICP-MS instruments have been developed to have exceedingly high (or low, depending on how you look at it) abundance sensitivity, with conventional quadrupole-based instruments having less than 1 part in $10^6$ overlap (0.0001 %) in the adjacent mass channels. For all intents and purposes, this offers the potential for no equivalent need for fluorescence-style compensation. What remains, in some instances, is a correction for the purity, or rather the amount of impurity isotopes, in the enriched isotope (typically less than 1 % with exceptions for isotopes of very low natural abundance), a fraction that is determined at the time of enrichment and is not affected by the assay and accordingly does not require measurement at the time of analysis, and a typically even lesser contribution for oxide ions (M + 16) that is determined by the ion-oxide bond strength and the temperature of the plasma (ICP).

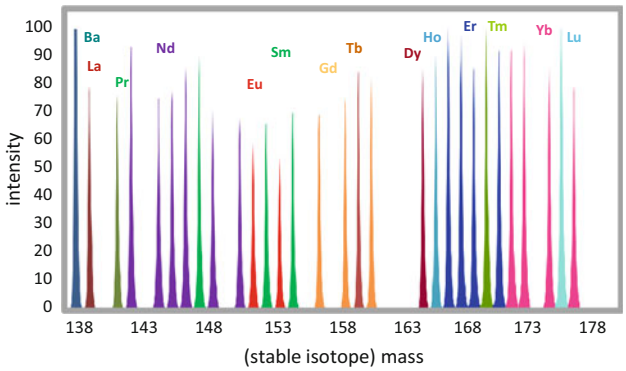


**Fig. 1** Mass spectrum of 30 enriched stable isotopes of the lanthanides, recorded for solution analysis at concentrations of approximately 20 ng/L (20 parts per trillion W/W) for each isotope

The ICP is a high-temperature plasma supported in a flowing argon stream, about 7,000°K at the skin and 5,500°K in the core, and substantially ionized. Because it is a toroidal plasma, sample can be injected along the axis, whereupon it is heated by convection and radiation. Sufficiently small particles (e.g., cells) are vaporized and atomized as they flow through the plasma. These conditions convey some essential and beneficial analytic characteristics. Under local thermodynamic equilibrium conditions, atoms having ionization potentials below about 9 eV (the majority of the periodic table of the elements) are practically 100 % ionized. Accordingly, the ionization efficiency is not significantly affected by the cellular matrix, the sensitivities for all atoms that are similarly efficiently ionized are similar, and the number of ions extracted from the plasma are directly and quantitatively related to the number of atoms introduced into the plasma—these form the basis of quantitative elemental analysis.

In mass cytometry, cells are immunologically stained in the familiar manner, except that the probes are tagged with metal isotopes instead of fluorophores [2]. In order to achieve sensitivities that are comparable to fluorophores, about 100 atoms of a given isotope should be bound to each antibody. This is conveniently achieved by constructing monodisperse polymers containing 30 monomer units each containing a pendant diethylenetriaminepentaacetic acid (DTPA) chelator: DTPA binds 3+ lanthanide ions with a dissociation constant $K_D \sim 10^{-16}$ (recall that streptavidin–biotin has a $K_D \sim 10^{-15}$), so that the tag atoms are essentially covalently bound and do not "cross talk" with differentially tagged polymers. Each polymer terminates in a thiol that reacts with a bismaleimide linker that similarly binds the metal-chelated polymer (MCP) to the reduced disulfides in the Fc region of an antibody. Typically, 4–5 such polymers will be linked to each antibody in order to achieve the desired sensitivity [3].

## Experimental

The present experiments used either straight-chain ("X8") polymers or third-generation dendritic ("DN3") polymers, both available as MAXPAR® reagents (DVS Sciences,

Sunnyvale, CA, USA). Most recent results take advantage of metal-conjugated antibodies that have become available. Though the synthesis protocol has changed, the X8 polymer is approximately similar in structure to that described by Lou et al. [2]; the DN3 polymer is a derivative construct that replaces the methacrylate backbone with a PAMAM-core dendrimer (Dendritic Nanotechnologies, Midland, MI, USA) that has been modified with DTPA chelators and the maleimide linker. Generally, the same tagging construct is used for all antibodies, since the DTPA chelator is equally efficacious for all lanthanide isotopes: the difference in probe is only which isotope is loaded into the polymer. Large dimensional panels are constructed using both cell surface and intracellular antibody probes: generally, an antibody that generates a useful flow cytometry response is also effective in mass cytometry. Because of the essential independence of the detection channels, it is possible to perform a single titration of all probes simultaneously, and "N − 1" assays are not required because of the lack of fluorescence-like compensation.

Almost all experiments reported here used rhodium (Rh) and iridium (Ir) metallo-intercalators to probe DNA (and some RNA) [3–5]. Though the intercalators are not tight binders, their pseudo-equilibrium distribution led them to be originally adopted in order to provide information regarding the amount of DNA [4]. However, today they serve two other important functions in mass cytometry. Because the intercalators are unable to penetrate the membranes of live cells [4], the live sample can be probed with the Rh intercalators, for which only the dead cells will be labeled. Subsequently, the cells are fixed and permeabilized and stained with the Ir-intercalators, which will then label all cells, and the data can be gated on Rh-negative cells to extract results only for live cells [3]. The Ir-intercalator provides the additional benefit of being a universal probe (for DNA-containing cells) that appears at high signal level and provides a convenient trigger for the identification of single cell events (see following section on "Data generation").

A recent publication [6] reports that dead cells bind cisplatin more readily than live cells during short exposure to the drug. It is shown that a pulse treatment (ca. 1 min at 25 μM) allows distinction of dead cells by the presence of Pt. Since the Pt appears to be covalently bound to the cell, it is reported that the cisplatin method is more resilient to stringent washing of samples.

Results reported and reviewed here were obtained at various installations of the CyTOF® mass cytometer (DVS Sciences, Toronto, Canada). The original prototype of the instrument has been described in some detail by Bandura et al. [7]: the fundamental core of the instrument remains similar to that description, one of the more significant changes being to the cell introduction method which has dispensed with the aerosol splitter providing an improvement in cell injection efficiency to approximately 30 %. At its heart, the mass cytometer is an ICP-MS instrument, but it has been adapted specifically to analyze single cell suspension-based samples common to traditional flow cytometry. Once injected, the cells are captured in an argon gas stream, where they are largely stripped of buffer. Subsequently transported into the central core of the plasma, the cells are rapidly and sequentially vaporized, atomized, and ionized. The cell is then represented by a cloud of ions (see "Data generation" section) that is extracted through the vacuum interface into the ion optics section of the mass cytometer. The ion current extracted from the ICP is enormous: approximately 1.5 mA passes through the vacuum interface, and this generates a self-defocusing space-charge-limited field [8, 9]. The majority of these ions are low mass, associated with argon, and since the mass cytometer is required only to determine the ions that are used as cell probes, significant advantage is had by maintaining hard extraction conditions while removing the ions below 100 dalton: the result is a less dense ion beam enriched in the masses that include the 100–200 dalton region and that can be manipulated and focused more efficiently. This pre-selection allows the ICP-MS to be configured with a time of flight (TOF) mass analyzer, which is normally less popular for atomic analysis because of the compromises required to deal with the effects of space charge in the TOF accelerator and flight tube. Removing the space charge throttle earlier in the ion optics allows advantage to be taken of the high spectral scan rate of the TOF, which will be seen to be critical for cell analysis. The transmitted ions are then focused into a sheet beam in the accelerator, whereupon they are pulsed through a fixed potential: under the same acceleration potential, low mass ions fly faster than high mass, and so the ions arrive at the detector in a manner such that their times of flight have a (square-root) dependence on ion mass. Resolution of the mass channels is thus a function of the "purity" of the ion energies and in this configuration can be established to provide abundance sensitivity sufficient to resolve adjacent mass channels by approximately 3 orders of magnitude. All the data are collected on a single detector, digitized at 1 Gs/s, and approximately 3,200 single-byte one-nanosecond samples are recorded for each spectrum. Each mass window is about 25 one-nanosecond windows wide. The flight time of the slowest atomic ions is approximately 13 μs, and so the TOF generates 76,800 spectra per second—and each spectrum is separately processed. Ultimately, as described below, the data are integrated and stored as an *.fcs file (each cell associated with $n$ integrated signals, where $n$ is the number of parameters requested in the experiment) that can be ported into third-party flow cytometry analysis software.

## Data generation

Each cell passes through the plasma (from injector to vacuum interface) in about 0.7 ms [10]. The cell exits the injector at a velocity of about 5 m/s and accelerates as it passes through the hot plasma to about 10 m/s. During that transit, the cell is vaporized, atomized, and ionized. Following ionization, the "cloud of ions" expands in a diffusion-limited manner. At the point of extraction into the vacuum interface, the cloud of ions for a single cell is approximately 2 mm diameter. Importantly, since the plasma is diffusion limited, the size of this cloud of ions is essentially independent of the size of the original (intact) cell: a 10-μm-diameter cell generates a 2 mm cloud of ions as does a 70-μm-diameter cell. At 10 m/s, this (essentially spherical) cloud of ions passes through the plane of the sampling aperture in about 200 μs: assuming that the temporal profile is maintained through the accelerator of the TOF where the transient signal is segmented into 13 μs pulses, this is essentially carried through the detector, and a single cell event is characterized by a transient signal of pseudo-Gaussian shape with a FWHM (full width at half maximum) of some 200 μs. Two implications arise from this. First, with stochastic cell introduction, the maximum cell throughput in order to minimize coincident arrival of ions from separate cell events is of the order of 1,000 cells per second. Second, the temporal evolution of the signal can be used to distinguish single cell events (Gaussian, 200 μs FWHM) from duplex (concomitant) cell events (bimodal, wider arrival time distribution) and cell fragment events (anticipated to be more broadly spread in time, and less regular).

Spectra are recorded at 13-μs intervals. Between cell events, the spectra are largely blank (reflecting only the metal content of the solution between cells, which may contain metals associated with antibodies that have dislodged from the cells—and therefore does not provide a suitable "negative" background). When a cell event arrives at the detector, a series of some 20–35 spectra reflect the pseudo-Gaussian transient and contain all the metal ions associated with the cell, including those metal isotopes that were intentionally attached as probes (of antibodies, DNA and viability indicators). The current version of the software provides a real-time display of the first 3 ms of data acquired each second, as an indicator of satisfactory analytical conditions. An example is provided in Fig. 2, showing a screen capture during the analysis of a PBMC sample stained with 27 metal-tagged antibodies. The display indicates a pixel each time an ion signal is recorded and shows the masses along the *X*-axis and the sequential spectra along the vertical axis. It is clear that the raw data cluster in two dimensions: along the *x*-axis in mass of the isotopes and along the *y*-axis in cell arrival events (e.g., each cell event is marked in about 20–35 sequential spectra of 13 μs each).

Using the Ir-intercalators as an indicator (although any single or combination of mass signals can be used), the two Ir isotopes (191 and 193 dalton) rise and fall in concert, and if their distribution is approximately Gaussian and 200 μs FWHM (e.g., "cell length," being the time for the signal to rise and fall between 3 % full maximum limits, is about 450 μs) then the event is flagged as a single cell event. Accordingly, the data in each mass window are integrated in two dimensions over the ca. 25 one-nanosecond windows of each mass spectrum and over the ca. 20–35 mass spectra corresponding to the cell transient ("cell length"), and a raw text file is written that includes the cell arrival time information and the integrated signals for each isotope indicated as used as a probe. This text file is then translated into an *.fcs file, including any data manipulation that is defined by the user (e.g., randomization of exact zero counts so that the digital data emulate the more diffuse distributions familiar in fluorescence cytometry).

The *.fcs file can be read by any flow cytometry software package. Typically, such software allows interrogation of the data through a set of orthogonal bivariate dot plots that display the correlation of 2 parameters for each cell. A typical four-parameter flow experiment can be represented by 6 such bivariate plots. A 32-parameter mass cytometry data set is represented by 496 bivariate plots, each of which can be gated and expanded in the remaining 30 dimensions. This approach is best described as a supervised assay that is hypothesis testing, since the user selects the gates and essentially asks the software a question that queries the data based on the user's understanding of the biologic system. It is evident that, while generally accepted and quite effective for small dimensional data sets, the approach fails to glean the available from a large dimensional set. Accordingly, it seems that a propitious approach to mass cytometry data will include unsupervised "gating" that can best be regarded as hypothesis generating. Examples that seem effective to date are frequently based on networking algorithms.

## Data analysis with demonstrative applications

The basis of one such approach to unsupervised assay lies in the display of multidimensional data in a polar plot that represents the intensity of each parameter by its distance from the axis of a circular plot. Figure 3 is such a display, presenting the data for 3 distinct samples in 20 parameters. The polar plot can be drawn as an integral for each sample, as here, or integrated over a gated population or a cluster or even for a single cell. The significant point is that in sufficiently multidimensional space, each cell "type" is

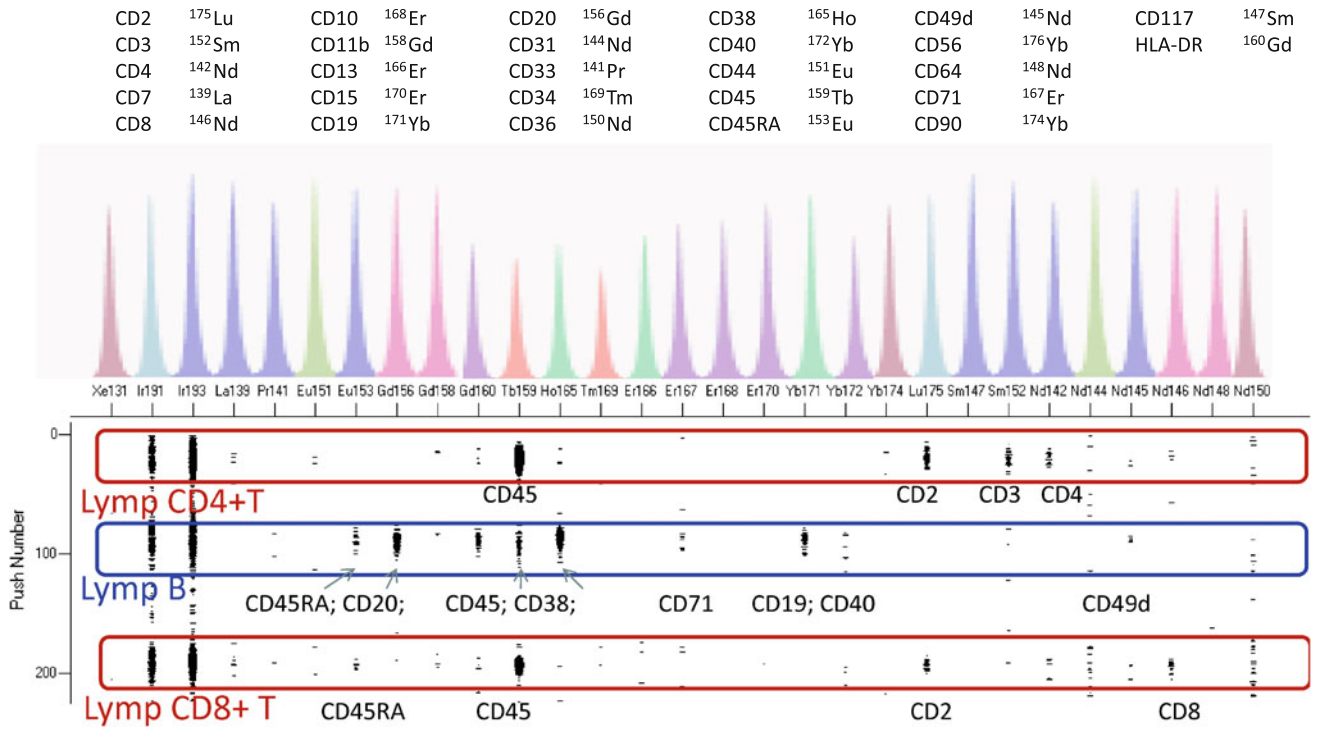


**Fig. 2** Computer screen shot during mass cytometric analysis of adult PBMC. These cells were probed with antibodies against 27 surface antigens. Each antibody was labeled with a different stable isotope (given in the *table* at the *top* of the *figure*: the antigen is indicated, such as CD2, followed by the isotope used to tag the corresponding antibody, $^{175}$Lu). In addition, cellular DNA was labeled with an Ir-intercalator (used as a trigger for cell recognition)

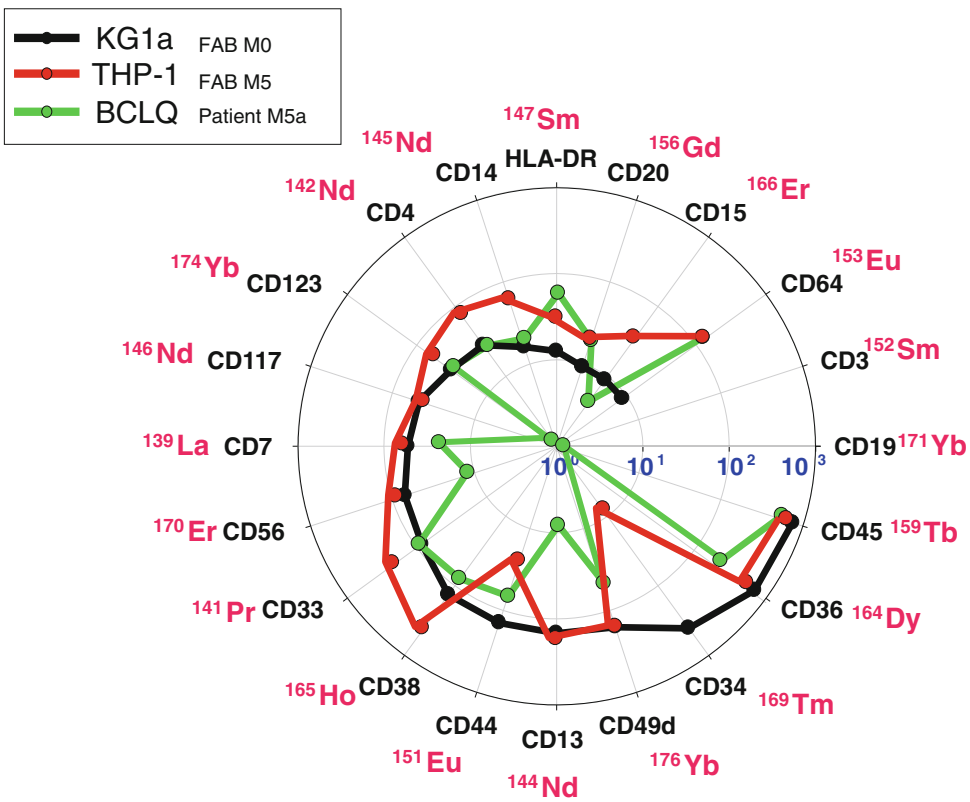


**Fig. 3** Radial polar plot display of 20-parameter data for two cell lines (KG1a and THP-1) and a patient bone marrow sample (BCLQ), representative of different acute myeloid leukemia subtypes. Intensity is displayed in logarithmic form as the distance from the axis of the plot. In this instance, the *plots* provide the marker signals averaged over a cell population: similar displays can be generated for gated populations or even for individual cells

distinguished from each other by its biomarker fingerprint: the more parameters that are used (regardless of whether they are positive, dim, or negative) the more confidently the fingerprint points to a specific phenotype or state.

The ability to distinguish cells by their fingerprints commends the application of bioinformatics for multi-parametric cytometry data sets, and specifically data clustering to identify the cell subpopulations in a sample [11]. Clustering (or computer-assisted gating), particularly in the instance that compensation is not required, should allow populations to be more readily distinguished (better defined, less biased, boundaries), and the ability to determine many parameters simultaneously suggests that unsuspected or rare populations might be more readily recognized. The latter speaks to the advantage of unsupervised assay. One such approach that has found popularity in other multivariate applications is the unsupervised neural network (UNN) [12, for example]. A discussion of the various algorithmic approaches to UNN is beyond the scope of this work, but in general clusters are identified by their proximity in *n*-dimensional space to their most similar characteristics (boundary of the cluster) and by their distance from other clusters of dissimilar character using a variety of distance or similarity measures. The input to the UNN (or training) includes the raw data, and the number of distinguishable clusters that are desired, dimensionality, and type of similarity measure; biologic presumption, and even information regarding the sample, is not required. A trained UNN includes the boundary conditions that split the multi-parametric space into trained clusters spaces. The trained UNN recognizes this underlying structure in other samples belonging to the same set.

Unsupervised neural networks have two principal applications (1) to recognize a cell population that is included in a training set and (2) to distinguish *n* clusters in a sample, such clusters being subsequently subject to

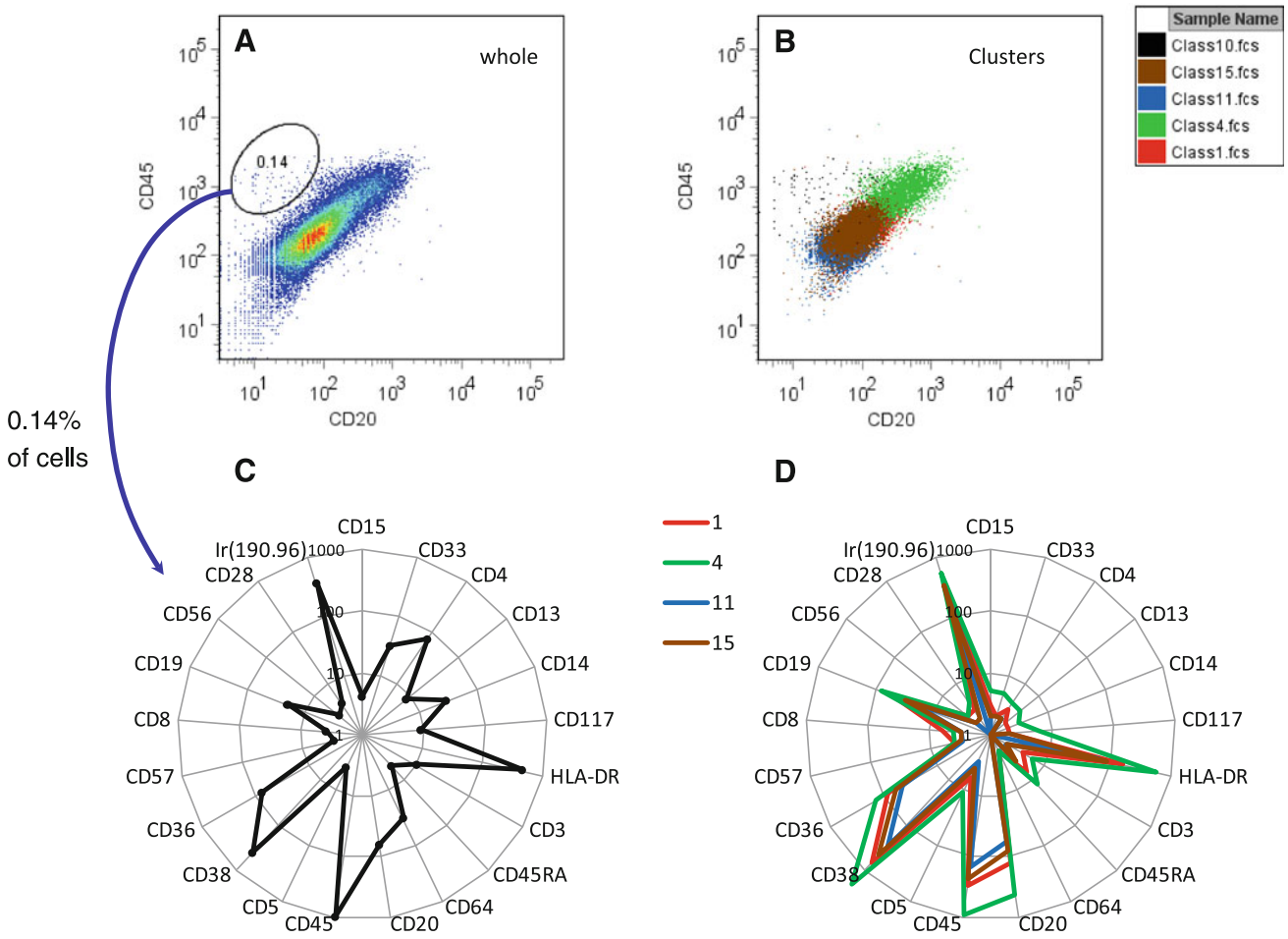


**Fig. 4** Unsupervised neural network analysis of Ramos B cells contaminated with 0.1 % PBMC from a leukemia patient. **a** Two-dimensional dot plot (CD45 vs. CD20) for the entire sample. These data were subjected to a UNN requesting identification of 15 distinguishable clusters of cell, 5 of which are shown in **b**. Radial polar plots are shown for these 5 clusters in **d**: 14 of the 15 clusters were similar and were distinguished as "different clusters" principally on the basis of intensity. A 15th cluster corresponding to the gated region in **a** was identified on the basis of its polar plot shown in **c** as the cells corresponding to the leukemia patient

analyst identification. One example of use of a UNN to identify an unanticipated cell population is shown in Fig. 4. In this instance, a Ramos B cell sample was contaminated with 0.1 % of a leukemia patient's blast cells. The 20-parameter (plus intercalators) data were submitted to UNN for analysis without provision of any sample information other than the raw per cell data. The UNN was asked to identify 15 clusters or which 14 were characteristic of the B cell fingerprint: because the fingerprint signatures were so similar, the clusters in this case were differentiated on their overall intensities. However, a cluster defined by a distinct fingerprint was distinguished and associated with 0.1 % of the cell population, and furthermore, the fingerprint associates that cluster with the patient whose cells were spiked into the sample. The example points to a potential clinical application of determining rare cell populations in a complex matrix, and identifying that population once the UNN is trained on multiple representative samples.

There are a variety of other clustering approaches that have been applied to flow cytometry: for example, SamSpectral [13], Flowclust [14], and FLAME [15]. Very recently, this suite of algorithms has been complemented to discern cellular hierarchy via RchyOptimyx [16] and applied to mass cytometric data. It is shown that this latest approach, based on graph theory and related optimization algorithms, is able to generate a minimal but sufficient top-down relationship of cell populations that are connected using parent–child relationships.

We close out this discussion of data analysis with three quite distinct application-specific analytical approaches that have been reported in the recent literature.

In their groundbreaking paper using mass cytometry to characterize signaling responses, Bendall et al. [17] simultaneously combined 31 markers to elucidate the functional response of the entire human hematopoietic system to immune modulators and small molecule drug inhibitors. In

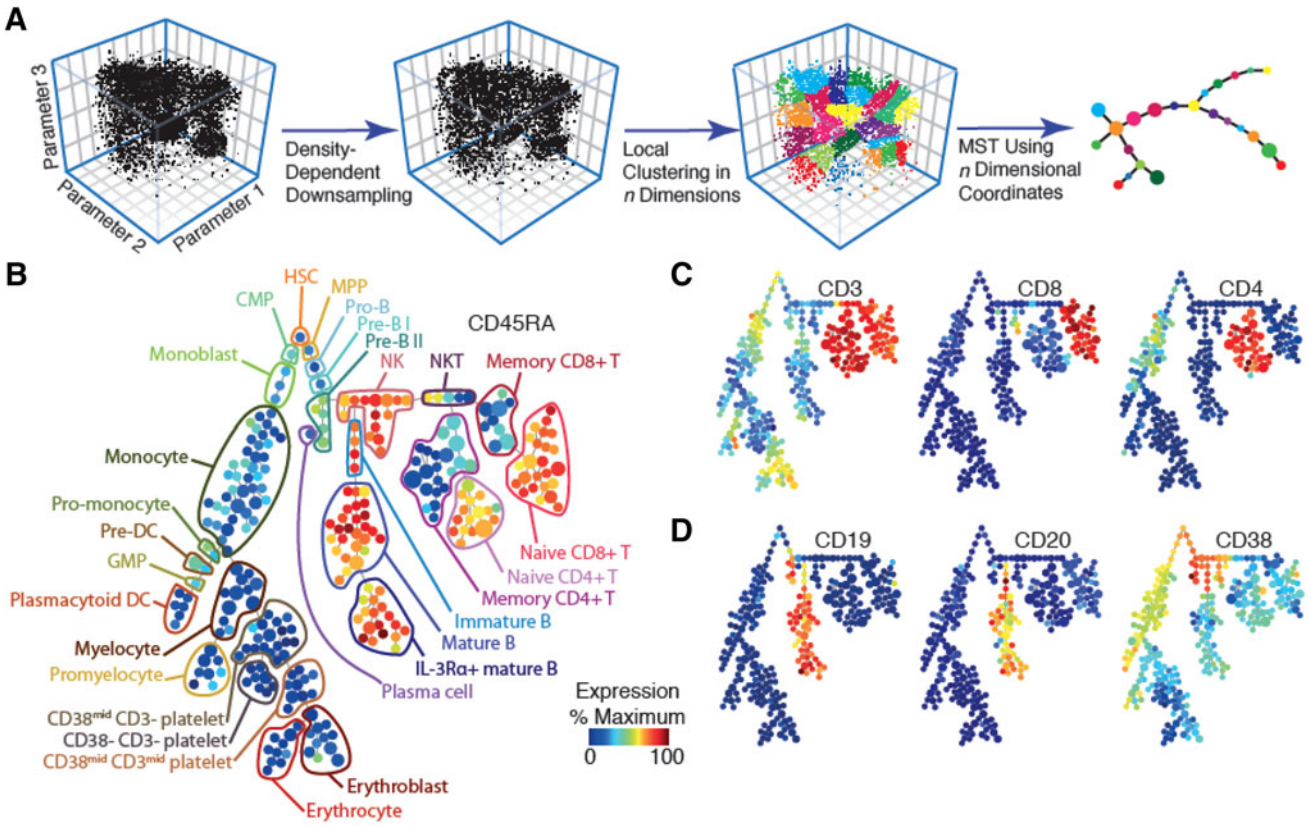


**Fig. 5** Bone marrow samples were analyzed for 31 proteins, of which 13 were cell surface markers used for subpopulation analysis, and 18 were intracellular signaling molecules used to measure functional response to stimulation and inhibition. The data presented here focus on the 13 cell differentiation markers. The data were subjected to unsupervised cluster analysis (SPADE), which identifies distinct phenotype populations and determines the relationships based on nearest neighbor populations (see [17, 18] for detailed information). **a** Highlights the major steps in defining the SPADE hierarchical tree. The outcome of the analysis is the "tree" shown in **b** which is interpreted to reflect immunological populations based on the marker distributions, a subset of which are shown in **c** and **d**. The "tree" that is formed is reminiscent of the hierarchical progression that is consistent with models for hematopoiesis. From Bendall et al. [17]. Reprinted with permission from AAAS

their model system, human bone marrow samples were incubated with the various stimuli in the presence or absence of small molecule signaling inhibitors. Each sample was then simultaneously probed with 13 immunophenotyping and either 18 additional subset-specific surface markers or 18 functional markers (including 16 phospho-specific signaling markers) and analyzed on the CyTOF® instrument. The surface markers were used to generate an immunophenotyping tree with the SPADE algorithm [18] (as implemented in Cytobank software), which groups cells of identical phenotype into clusters, connected to each other based on similarity in a tree structure (Fig. 5). Within the tree, all the traditional bone marrow cell types were identified by backgating known subset immunophenotypes, which allowed grouping of the clusters as indicated. This analysis greatly increases the understanding of cell types and their relationship to each other within normal bone marrow, and enables presentation of the bone marrow as a continuum of phenotypes. The paper goes on to show the impact of stimuli and small molecule inhibitors on signaling pathways for each of the bone marrow subsets. For example, BCR activates PLCγ2 specifically in B cells, while pervanadate induces PLCγ2 phosphorylation in most subsets, and activation by both immunomodulators is abolished with dasatinib. Alternatively, IL7 specifically induces STAT5 activation in T cells, while pervanadate acts as a pan-subset STAT5 activator. Dasatinib abolishes PVO4-induced STAT5 activation in all subsets except plasmacytoid dendritic cells, but has no impact on IL7 stimulation. These data

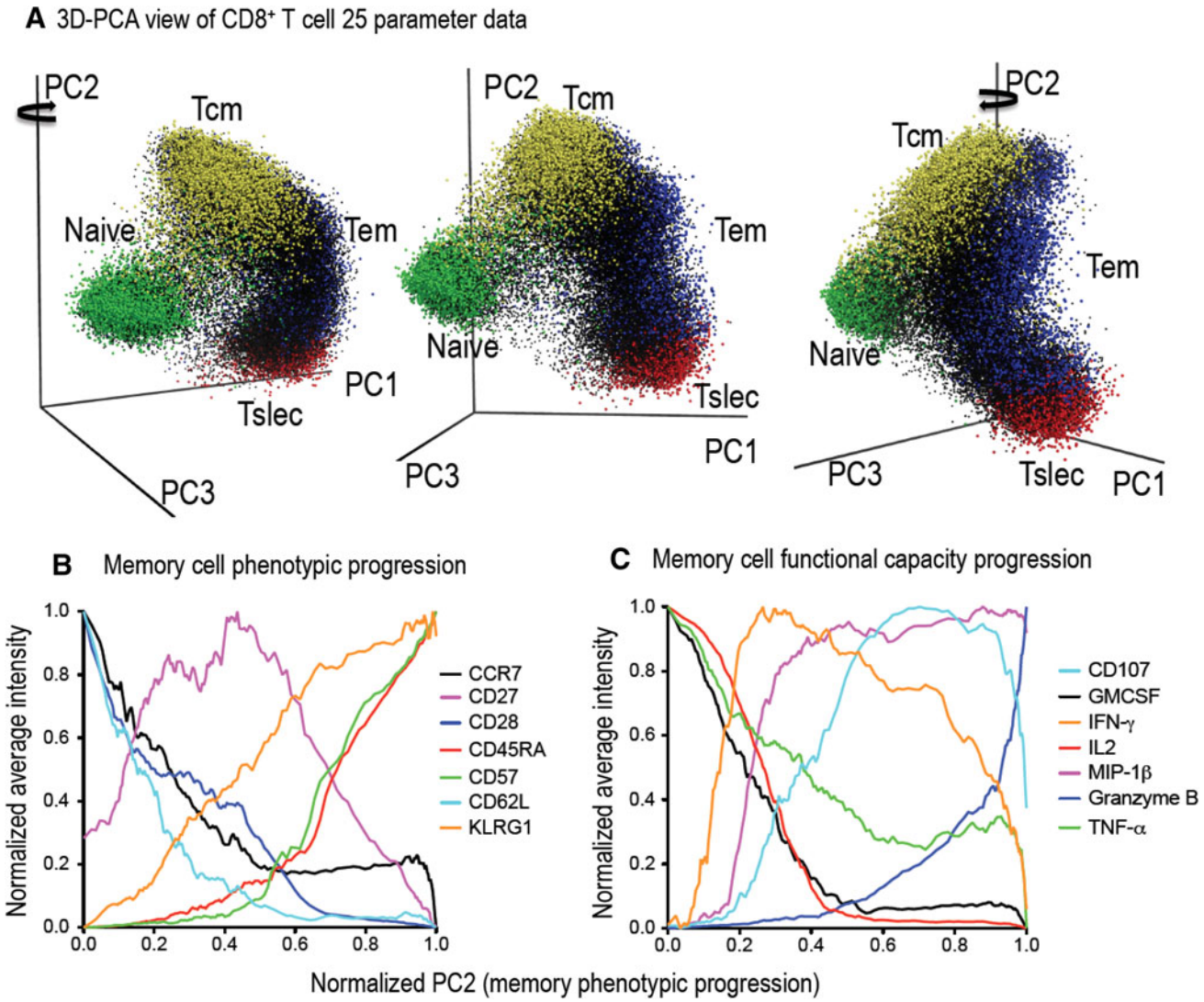


**Fig. 6** Major T cell clusters were determined using metal-encoded tetramer probes and examined using 25 metal-labeled antibodies against cell differentiation and signaling molecules. The data are displayed in **a** as 3D PCA plots constructed from the 25 antibody response signals, principally distinguishing differentiation state, memory segregation, and memory status (see [19] for details). Phenotypic and functional capacity progression are displayed in **b** and **c** by expansion along the PC2 (memory progression) axis. Reprinted from Newell et al. [19] with permission from Elsevier

demonstrate that the effects of inhibitors are both cell type-specific and stimulation-specific. This information is critically important for understanding the expected drug efficacy in different disease states and/or treatment regimens, and CyTOF® instrument provides the multi-parametric dimensionality to attain this detailed system information.

Using a different analytical approach, Newell et al. [19] used 32 parameter mass cytometry data to uniquely characterize the functional and phenotypic diversity virus-specific CD8+ memory cells within human peripheral blood. Sixteen immunophenotyping markers and 9 functional markers (including chemokines, cytokines and cytotoxic granule markers) were combined using principal component analysis (PCA) to cluster cells based on phenotypic and functional characteristics (Fig. 6). The composition of each major cluster was determined by displaying naïve (green) and memory (central (Tcm), yellow; effector (Tem), blue; short-lived effector (Tslec), red) subpopulations gated by known surface marker phenotype. Doing so revealed that PC1 primarily discriminates naïve from memory cells, while PC2 differentiates memory status. Following the expression of memory phenotypic and functional markers along the PC2 axis shows progressive gain in markers associated with effector memory cells (granzyme B and CD107) and loss in markers (CD62L, CCR7, and IL2) associated with central memory cells (Fig. 6b, c). Further studies in the paper showed that hundreds of functionally distinct subpopulations, as defined by the array of expressed cytokines, chemokines, and granule components, could be discriminated upon stimulation, revealing a previously unappreciated heterogeneity of memory cell diversity. Additional studies utilized metal-tagged peptide-MHC complexes to identify and functionally characterize the immune response of virus-specific memory cells (data not shown, see paper). This approach revealed that CMV memory displays a short-lived effector memory bias, consistent with the chronic and highly immunogenic nature of CMV infection, while Flu memory displays a more central memory bias, consistent with the episodic nature of Flu infection and the fact that none of the donors were ill at the time of donation. Each virus-specific niche favored different combinations of functionally distinct subpopulations. In all, the highly multi-parametric nature of mass cytometry data display a continuous

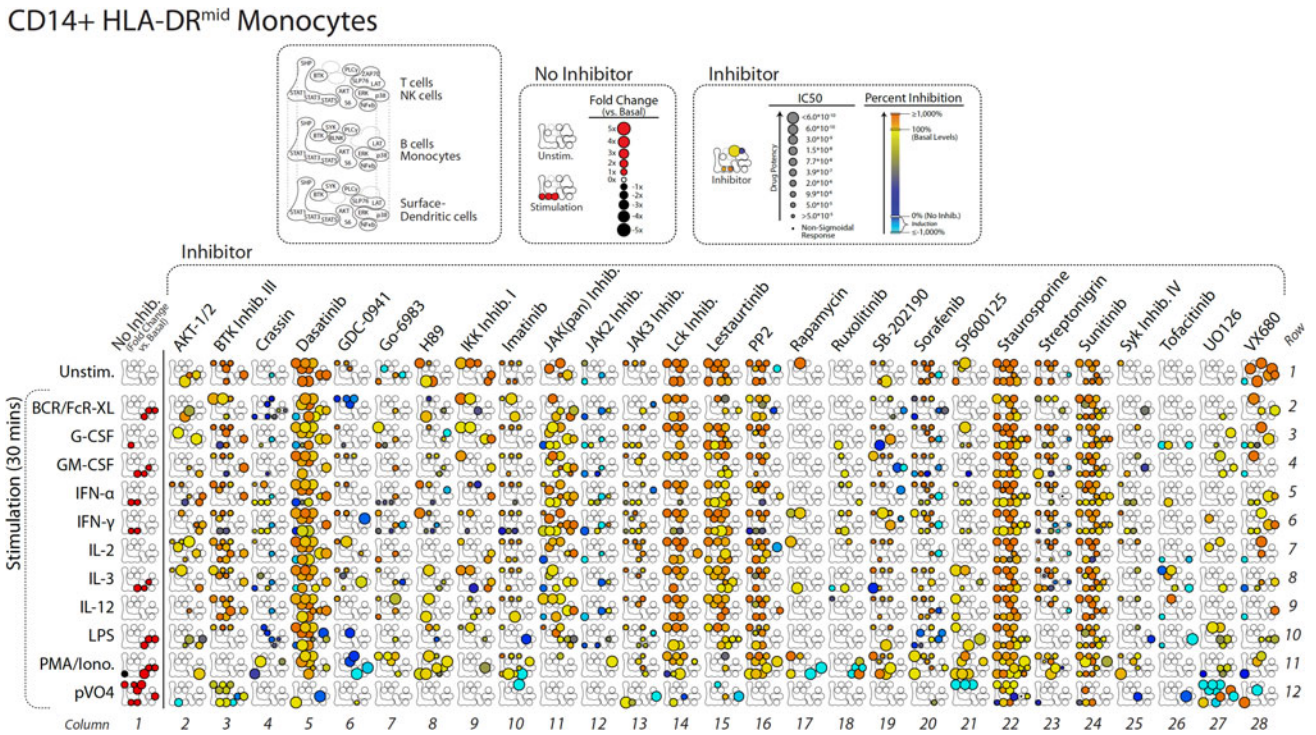


**Fig. 7** Exemplary data from an assay of PBMC using 9 probes to distinguish 14 immunological populations, 14 probes against 14 intracellular signaling molecules, and 7 bar-coding elements to multiplex 96 samples that were probed with 8 concentrations each of 12 stimulants (96 samples) and inhibited with each of 27 inhibitors. In this instance, the data are presented for the population identified as $CD14^{+}$ $HLA\text{-}DR^{mid}$ showing the signaling responses for the 14 signaling molecules (arrayed as in the *upper left figure*) to the 12 stimulants indicated at the *left* (*one per row*) and inhibitors indicated in each column. The *size* and *color* of each *spot* indicates the IC50 and percent inhibition observed in each instance (see [20] for details). For example, in this instance, in the presence of ruxolitinib, inhibition of phosphorylation of STAT1 ($IC_{50}$ = 23 nM, 93 % inhibition) and STAT3 ($IC_{50}$ = 4 nM, 147 % inhibition) was observed. Reprinted by permission from Macmillan Publishers Ltd.: Bodenmiller et al. [20]

nature of CD8+ T cell differentiation with a high degree of functional diversity, revealing a remarkable degree of flexibility in the immune response to viral infection.

A recent paper from Bodenmiller et al. [20] that comprehensively map the signaling responses of human PBMC takes multiplexing of multi-parameter mass cytometry data to a new plane. In this report, the impact of 27 different drugs on 14 signaling pathways in 14 PBMC subpopulations was measured with mass cytometry. Each inhibitor profile consisted of an 8-point dose response of inhibitor for each of 12 stimuli and thus occupied an entire 96-well plate. Because of the scale of this experiment, mass-tag cellular barcoding (MCB) was employed that allowed combination of all 96 wells of an inhibitor plate into a single tube for subsequent sample processing, data collection, and decoding for individual sample analysis. MCB involves labeling each sample with a unique binary (present/not present) array of isotope tags (linked to thiol-reactive moieties that link to surface-exposed cysteines) prior to combining samples for processing and analysis. Seven isotope tags allow unique barcoding of $2^7$ samples and were thus used in this experiment to encode each inhibitor plate. MCB not only dramatically increases sample throughput, but also ensures uniform sample labeling and data collection, and reduces experimental cost by requiring less staining reagents. Following MCB, samples were simultaneously probed with 9 immunophenotyping markers to identify 14 PBMC subpopulations and 14 phospho-specific markers to measure activation state of signaling pathways. This experiment generated an enormous database of greater than 60,000 $IC_{50}$ values (27 inhibitors on 12 stimuli for each of 14 signaling pathways in 14 subpopulations) that enables context-specific classification of inhibitor function and cellular response. One can visualize and query this database using a signaling response map (Fig. 7). For each signaling pathway (indicated by spatial address), the $IC_{50}$ (indicated by size of bubble) and percent inhibition (heat map color) for each of the 27 inhibitors (columns 2–28) on the 12 stimuli (rows 1–12) are shown for the indicated monocyte subpopulation. Looking at all the data, it was clear that inhibitor activity was cell type and context specific. Furthermore, PCA-based compound classification suggested novel inhibitor mechanisms of action. The large database of cellular signaling states could be queried to influence preclinical development, potentially enabling predictive selection of drugs for defined disease states.

## Conclusions

Mass cytometry is a specific implementation of inductively coupled plasma mass spectrometry (ICP-MS, atomic mass spectrometry) that enables an unprecedented degree of multi-parametric single cell analysis and provides an opportunity for absolute quantification. Though fundamentally a different technology, mass cytometry has adapted key attributes common to conventional flow cytometry that make it attractive for subpopulation and rare event analysis. Mass cytometry data are compatible with conventional cytometry analysis approaches, which can largely be categorized as supervised and hypothesis-testing, and which are most appropriate for low-dimensional data. However, high dimensionality data analysis benefits from unsupervised approaches and naturally lends itself to hypothesis-generating interrogation. UNNs, PCA, and spanning tree progression analysis (SPADE) have been successfully applied to 30+ dimensional mass cytometry data and serve as benchmarks for further developments that are evidently in process. The dimensionality, not to mention sample throughput, is dramatically increased with the introduction of metal encoding, which enables multiplexed multi-parameter cellular analysis.

**Acknowledgments** The University of Toronto investigators gratefully acknowledge receipt of financial support from the Ontario government through the Ontario Research Fund—Global Leadership in Genomics and Life Sciences (ORF-GL2-01-003). Scott Tanner further wishes to acknowledge previous enabling research funding from Genome Canada (Applied Human Health, and Technology Development) and ongoing support from the NIH-Office of AIDS Research.

**Conflict of interest** The authors are employees of, and receive remuneration from, DVS Sciences Inc. Scott Tanner, Vladimir Baranov, Olga Ornatsky, and Dmitry Bandura are co-founders of and equity shareholders in DVS Sciences Inc. Scott Tanner is a member of the Board of Directors of DVS Sciences, Inc.

## References

1. Houk RS, Fassel VA, Flesch GD, Svec HJ, Gray AL, Taylor CE (1980) Inductively coupled argon plasma as an ion source for mass spectrometric determination of trace elements. Anal Chem 52:2283–2289
2. Lou XD, Zhang G, Herra I, Kinach R, Ornatsky O, Baranov V, Nitz M, Winnik MA (2007) Polymer-based elemental tags for sensitive bioassays. Angew Chem Int Ed Engl 46:6111–6114
3. Ornatsky O, Bandura D, Baranov V, Nitz M, Winnik MA, Tanner S (2010) Highly multiparametric analysis by mass cytometry. J Immunol Methods 361:1–20
4. Ornatsky OI, Lou X, Nitz M, Schafer S, Sheldrick WS, Baranov VI, Bandura DR, Tanner SD (2008) Study of cell antigens and intracellular DNA by identification of element-containing labels and metallointercalators using inductively coupled plasma mass spectrometry. Anal Chem 80:2539–2547
5. Majonis D, Herrera I, Ornatsky O, Schulze M, Lou X, Soleimani M, Nitz M, Winnik MA (2010) Synthesis of a functional metal-chelating polymer and steps toward quantitative mass cytometry bioassays. Anal Chem 82:8961–8969
6. Fienberg H, Simonds EF, Fantl WJ, Nolan GP, Bodenmiller B (2012) Platinum-based covalent viability reagent for single cell mass cytometry. Cytometry A 81A:467–475

7. Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick JE, Tanner SD (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal Chem 81:6813–6822
8. Gillson GR, Douglas DJ, Fulford JE, Halligan KW, Tanner SD (1988) Non-spectroscopic interelement interferences in inductively coupled plasma mass spectrometry (ICP-MS). Anal Chem 60:1472–1474
9. Tanner SD (1992) Space charge in ICP-MS: calculation and implications. Spectrochimica Acta Part B 47B:809–823
10. Olesik JW, Gray PJ (2012) Considerations for measurement of individual nanoparticles or microparticles by ICP-MS: determination of the number of particles and the analyte mass in each particle. J Anal At Spectrom 27:1143–1155
11. Lugli E, Roederer M, Cossarizza A (2010) Data analysis in flow cytometry: the future just started. Cytometry A 77A:705–713
12. Balfoort HW, Snoek J, Smiths JRM, Breedveld LW, Hofstraat JW, Ringelberg J (1992) Automatic identification of algae: neural network analysis of flow cytometric data. J Plankton Res 14:575–589
13. Zare H, Shooshtari P, Gupta A, Brinkman RR (2010) Data reduction for spectral clustering to analyze high throughput flow cytometry data. BMC Bioinform 11:403
14. Pyne S, Hu X, Wang K, Rossin E, Lin TI, Maier LM, Baecher-Allan C, Mclachlan GJ, Tamayo P, Hafler DA, De Jager PL, Mesirov JP (2009) Automated high-dimensional flow cytometric data analysis. Proc Natl Acad Sci 106(21):8519–8524
15. Lo K, Hahne F, Brinkman R, Gottardo R (2009) flowClust: a Bioconductor package for automated gating of flow cytometry data. BMC Bioinform 10:1–145
16. Nima Aghaeepour N, Jalali A, O'Neill K, Chattopadhyay PK, Roederer M, Hoos HH, Brinkman RR (2012) RchyOptimyx: cellular hierarchy optimization for flow cytometry. Cytometry A. Published online 8 Oct 2012. doi:10.1002/cyto.a.22209
17. Bendall SC, Simonds EF, Qiu P, Amir ED, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sach K, Pe'er D, Tanner SD, Nolan GP (2011) Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 332:687–696
18. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Bruggner RV, Linderman MD, Sachs K, Nolan GP, Plevritis SK (2011) Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol 29:886–891
19. Newell EW, Sigal N, Bendall SC, Nolan GP, Davis MM (2012) Cytometry by Time-of-Flight Shows Combinatorial Cytokine Expression and Virus-Specific Cell Niches within a Continuum of CD8+ T Cell Phenotypes. Immunity 36:142–152
20. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, Simonds EF, Bendall SC, Sachs K, Krutzik PO, Nolan GP (2012) Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat Biotechnol 30:858–867