Published in final edited form as:

Methods. 2012 July ; 57(3): 251–258. doi:10.1016/j.ymeth.2012.02.009.

Cytometry: Today's Technology and Tomorrow's Horizons

Pratip Chattopadhyay and Mario Roederer

ImmunoTechnology Section, VRC, NIAID, NIH, 40 Convent Dr., Room 5509, Bethesda, MD 20817

Abstract

Since the invention of flow cytometry in the 1960's, advances in the technology have come hand-in-hand with advances in the recognition and characterization of new leukocyte subsets. In the early years, with the advent of one- and two-color flow cytometers, major lymphocyte lineages comprising the cellular arm (T-cells) and the humoral arm (B-cells) were identified^{1, 2}. Through the 1980's, the ability to perform three- and four-color flow cytometry experiments enabled the enumeration of cells expressing combinations of CD3, CD4, and CD8 from a single tube; this was a necessity driven by the clinical demands of the emerging HIV epidemic³. The following decade saw continued development in multicolor technology and immunology, with the advent of polychromatic flow cytometry (detection of 5 or more markers simultaneously) enabling identification of naïve and memory T-cell subsets⁴ and detailed functional characterization of antigen-specific lymphocytes (such as measurement of multiple cytokine production from individual cells⁵). Most recently, the new millennium brought 12–18 color technology^{6, 7} and an unprecedented resolution to immune analysis (including the identification of regulatory T-cells⁸, follicular helper T-cells⁹, TH17 cells¹⁰, and the ability to combine functional and phenotypic analyses¹¹; Figure 1). The ongoing development of flow cytometry technology has left its mark on the analysis of hematopoietic development, cell signaling networks, and leukemia/lymphoma diagnoses.

Keywords

Flow cytometry; Single-cell analysis; Data analysis; Polychromatic Flow Cytometry; Fluorescence Reagents

Introduction

Since the invention of flow cytometry in the 1960's, advances in the technology have come hand-in-hand with advances in the recognition and characterization of new leukocyte subsets. In the early years, with the advent of one- and two-color flow cytometers, major lymphocyte lineages comprising the cellular arm (T-cells) and the humoral arm (B-cells) were identified^{1,2}. Through the 1980's, the ability to perform three- and four-color flow cytometry experiments enabled the enumeration of cells expressing combinations of CD3, CD4, and CD8 from a single tube; this was a necessity driven by the clinical demands of the emerging HIV epidemic³. The following decade saw continued development in multicolor technology and immunology, with the advent of polychromatic flow cytometry (detection of 5 or more markers simultaneously) enabling identification of naïve and memory T-cell

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

subsets⁴ and detailed functional characterization of antigen-specific lymphocytes (such as measurement of multiple cytokine production from individual cells⁵). Most recently, the new millennium brought 12–18 color technology^{6,7} and an unprecedented resolution to immune analysis (including the identification of regulatory T-cells⁸, follicular helper T-cells⁹, TH17 cells¹⁰, and the ability to combine functional and phenotypic analyses¹¹; Figure 1). The ongoing development of flow cytometry technology has left its mark on the analysis of hematopoietic development, cell signaling networks, and leukemia/lymphoma diagnoses.

Remarkable advances in hardware technology (e.g., high powered lasers emitting blue, red, green, violet, and yellow light), fluorochrome chemistry (e.g., tandem dyes, nanocrystals), and software tools have driven this co-evolution of flow cytometry and immune analysis¹². However, data analysis strategies are still relatively underdeveloped; that is, the multitude of data available from a polychromatic flow cytometry experiment is rarely analyzed easily or in its entirety. To a lesser extent, improvements are possible in instrument calibration, experimental design, and reagents. In light of this, two questions can be posed: do we really need to measure this many parameters simultaneously, and, if so, what does the future hold in terms of advances in flow cytometry and the development of other single-cell technology platforms? This chapter will address these questions by examining the current and future states of cytometric analyses.

The Need for Multiparametric Analysis

A basic advantage of polychromatic experiments (compared to two- or three-color analyses) is better specificity for desired cells, especially with the detection of low frequency populations. Such populations include antigen-specific T- and B-cells, which are commonly identified with fluorescent multimers of peptide+MHC¹³ or antigen¹⁴. These multimeric complexes will often bind non-specifically to various leukocyte populations (particularly monocytes) present in the sample. Similarly, multimeric proteins or the antibody conjugates used for staining may bind dying cells non-specifically. Thus, the use of antibodies to negatively select for undesired populations (such as CD14 to identify and eliminate monocytes) as well as the use of viability dyes (to exclude dead cells, a common source of non-specific interactions) is critical. Even when these negative-selection reagents are combined in a single “dump” channel, simple experiments quickly become polychromatic. For example, basic phenotyping of an antigen-specific T-cell population requires one channel for exclusion markers, three channels to identify T-cell lineages, one channel for identification of the antigen-specific population (e.g., with a peptide-MHC Class I multimer), and ideally at least three channels for determination of T-cell differentiation (memory, effector, etc)³. In total, this phenotyping experiment requires 8-color flow cytometry.

A similar issue arises when trying to interpret the biology of cells with a particular phenotypic profile. This is particularly important when trying to distinguish biologically distinct cell types that share expression of some markers; our interpretation must be validated biologically. The identification of regulatory T-cells (Treg) is a notable example. Initial descriptions^{15–17} of these cells relied solely on the level of CD25 expression (e.g., the top 2% of expressing cells were considered Treg); however, variability between reagents as well as samples led to considerable confusion with such identification. Suboptimal staining of CD25 had important consequences as expression of this marker is also observed on activated cells (a cell whose role is to amplify the immune response, as opposed to down regulate it!). Subsequent studies associated expression of the intracellular marker FoxP3 with regulatory T-cells¹⁸; however, again, stimulated cells also express this transcription factor. Most recently, the prevailing dogma requires multiple markers in combination to detect these cells (e.g., CD25/FoxP3/CD127/CD39) or the analysis of cytokine production

following stimulation (TGF β , IL10)^{19,20}. In either case, at a minimum, to ensure biologically valid identification of regulatory T-cells, six color flow cytometry experiments are needed.

The number of markers expressed by Treg, and the fact that many of these markers are also expressed by other cell types, provides a glimpse into the incredible diversity of memory T-cells. For example, with an 11-color flow cytometry experiment designed to phenotype memory T-cell subsets, expression of six markers (CD45RA, CCR7, CD27, CD127, CD57, CD28) can be investigated. If expression of each marker is considered to be either on or off (and ignoring variation in expression, a maximum number of 64 cell subsets within each of CD4 or CD8 T could theoretically be identified. However, if phenotypic diversity were limited, only a fraction of these T-cell subsets might be represented in the peripheral blood. In fact, this is not the case, as cells expressing each combination of these markers can be detected²¹. This suggests that there are at least 64 different subsets of circulating CD4 and CD8 T-cells—albeit with as-yet largely unknown functional correlates. As the number of markers used to characterize cells increases, the number of different cell populations detected in the peripheral blood rises as well, implying that the true diversity of memory T-cell subsets is even greater. Quantifying this diversity and determining which subsets of cells are functionally similar is fundamental to our understanding of the immune system.

More importantly, the ability to finely define T-cell subsets (based on multiple parameters) improves the likelihood of identifying cells important in disease pathogenesis or vaccination. The identification of critically-relevant cells is important from the standpoint of immune monitoring and optimizing vaccine efficacy. For example, in recent years, a body of literature has emerged that describes the importance of polyfunctional T-cells (cells simultaneously expressing a majority of the measured functional markers, including e.g. IFN γ , IL2, TNF, MIP1 β , and CD107a)²²; the frequency of these cells has been associated with less pathogenic disease (HIV in long-term non-progressors²³, HIV-2²⁴) or protective immunity (Leishmania²², Vaccinia²⁵, Yellow Fever²⁶, SIV²⁷). Notably, the relationship between these functional markers and disease was not apparent in earlier studies, measuring only one or two functional markers at a time, and can only be revealed with 9+ color flow cytometry.

The reason underlying the failure of single parameter approaches is illustrated in Figure 2. In this hypothetical example, the measurement of four functions together defines a putative population of cells (IFN γ + IL2– TNF+ CD107a+) whose frequency has a strong, inverse correlation with disease progression. When TNF is not measured in the experiment, then TNF-negative cells will also fall into the measured subset, thereby introducing noise into the assay and weakening the correlation. When two markers are absent, more noise is introduced into the assay, further reducing the power to predict disease progression. Finally, when only a single marker is measured (IFN γ alone, for example) predictive power could be completely lost. Thus, bulk measurements of only one or two cell types may reduce the power to detect cell populations important in disease. In this way, multiparametric technology (such as polychromatic flow cytometry) provides not only a better theoretical understanding of immunity, but can also have practical benefit.

Approaches to Polychromatic Experiments: Limitations and Future Directions

Today, commercial flow cytometers capable of measuring 10 parameters are common, and those capable of 12–20 parameters are widespread. This rapid dissemination of hardware technology has revealed pressing needs in other areas: 1) practical and efficient methods for instrument calibration and quality control (QC), 2) availability of a wider variety of

fluorochromes and antibody conjugates, 3) assistance with experimental design, and 4) better strategies for data analysis.

Instrument Calibration and QC

The success of polychromatic flow cytometry depends critically on instrument calibration; however, the old paradigm of analyzing unstained cells, and adjusting the gain (voltage) for all detectors until the signal falls under the second log decade persists. This method is problematic, since it does not consider the precision of fluorescence measurements in choosing voltages, nor does it ensure that the signal from a particular fluorochrome is detected maximally in the channel designated for it. The former can impact resolution of dim populations, by increasing the “spillover spreading” of data²⁸, while the latter complicates compensation for spectral overlap. Most critically, the method puts enormous importance on the ability to quantify “negative” signals – unstained cells – where the relative error in measurement is huge. The consequence can be poor performance of staining panels, and an inability to interpret experimental data.

In recent years, two systems for instrument calibration have emerged. Becton Dickinson (BD) has developed a proprietary system, known as Cytometer Setup and Tracking, or CS&T, for integration with their instruments and analysis software. The cornerstone of this system are beads that yield varying amounts of fluorescent signals, mimicking positive, dim, and negative staining. Analysis of these beads (which is performed automatically within the instrument software) checks laser and detector performance, linearity, and reports the gain settings that minimize the CV of dim populations. The method is fast and automatically provides a wealth of information about instrument performance, which can be tracked over time as a QC tool. However, there are important caveats beyond the restricted availability on instruments from solely this manufacturer. First, the signal of the brightest peak in each channel must be calibrated against a sample stained with the antibodies used in the experiment. Thus, the calibration routine requires use of a stained sample, and must be repeated every time a new staining panel is employed. Second, the algorithm requires that the positive signal be limited to ~10X the background signal. This setting is arbitrary, and may result in suboptimal quantification of extreme signals (very bright markers). Third, the method is currently incompatible with high power lasers, which perform so well that there is no measurable difference between the CVs of the lower and upper peaks. This is an important limitation, because it excludes green laser systems, which have been shown to provide much improved resolution of signals from PE and PE-based tandems compared to blue lasers. It also excludes high power red lasers, where the increased generation of photons may help resolve far red fluorochromes. Finally, the beads do not employ dyes commonly used in flow cytometry and vary from lot-to-lot. As such, calculations must be performed by downloading conversion factors, both for flow cytometry dyes (some of which may not be supported by BD) and for the lot identities.

A second system for instrument calibration was developed in our lab, and has been described previously²⁹. Briefly, this method employs antibody capture beads, which can be incubated with the fluorochrome-conjugated antibodies commonly used in experiments. In this way, the optimal gain values are calculated using all the fluorochromes the investigator might use. These beads have been manufactured with varying amounts of capture antibody, so that the beads provide a mix of five fluorescence levels (not simply bright, dim, and negative), thereby providing more information about the resolution of dim markers. When the beads are analyzed, the CV of signals are determined across a range of gain settings, and the voltage that provides the lowest CV, but still gives the brightest signal is chosen. Singly-stained compensation beads, and cell samples stained with a prototypic panel, are analyzed to verify that gain settings are optimal. These aspects of the method have important advantages over the CS&T system: first, it is applicable to all experiments, regardless of the

fluorochromes chosen. Second, this calibration need not be performed daily (in contrast to the CS&T system). Third, it allows more consistent instrument monitoring, using an independent bead system (8-peak rainbow beads).

The next generation of cytometry will address the major limitations of these two systems. First, automated analysis of calibration data – which is possible to some extent with the CS&T system, but not at all with our system – will likely become a reality soon, as flow cytometry software vendors offer calibration platforms. Second, these systems will eventually employ better measures of instrument performance, beyond changes in gain settings over time or CVs of stained populations. These numbers can vary dramatically from detector to detector, with no consequence to cell staining, or they can be consistent even in the face of dramatic changes. The reason for this is the relatively large variance in the performance of detectors. Therefore, new measures that are applicable across detectors and from instrument to instrument are needed. Two relevant measures, the efficiency of a detector (Q) in measuring a particular fluorochrome and the optical background (B) or electronic noise, should be considered in this regard. With such measures, the ability to standardize performance across instruments may be possible in the future. This would provide a great advantage over current systems, which can only standardize the behavior of a single instrument over time.

A Wider Variety of Fluorochromes and Antibody Conjugates

To perform (up to) 18-color flow cytometry analyses, a wide range of reagents is required. At a minimum, 18 different fluorescent dyes must be available, and a library of multiple antibodies (both in terms of specificity and clone) must be conjugated to each dye. In practice, many more fluorochromes are needed, since some fluorochrome conjugates of the same antibody perform better than others.

For many years, the variety of antibody-fluorochrome conjugates available was limited. This was especially true for UV- and violet-laser excited fluorochromes (Figure 3). With the introduction of quantum dots in 2004, seven or eight additional parameters could be measured in flow cytometry experiments, with excellent sensitivity. However, despite their utility, the availability of these materials has been limited. Moreover, many conjugates have not been suitable for intracellular staining (because of unpurified free dye)³⁰. Finally, quantum dot fluorescence can be fatally compromised by exposure to low level heavy metal contaminants in buffers³¹.

Recently, the development of a new class of dyes has been reported. Brilliant violet (BV) dyes were developed as a consequence of the Nobel Prize-winning finding that some organic polymers could conduct electrons as well as inorganic materials. This led to the development of π -conjugated polymers, in which each monomer harvests laser light, and transfers electrons from this light along the polymer chain. The collected light is cooperatively emitted as fluorescence, providing a highly sensitive probe (Figure 4). In fact, the first of these dyes, emitting light at 421nm (BV421) has a staining index rivaling PE. Because the design of these polymers is well-defined, conjugation to antibodies occurs at precise sites (limiting aggregation or multimeric complexes) and the materials are very stable. These dyes are suitable for high-sensitivity applications, such as the identification of antigen-specific T-cells (by peptide-MHC Class I multimer or CD154 staining) and can easily be substituted for Pacific Blue in multicolor staining panels. Notably, BV421 can be conjugated to other fluorochromes to make tandem dyes that emit light at a wide variety of wavelengths. These tandem dyes can be used intracellularly, providing a new sensitive option for intracellular detection from the violet laser.

The Challenges Associated with Experimental Design

A major obstacle in polychromatic flow cytometry is the development of optimized staining panels. Without carefully constructed panels, artefacts in staining patterns may be observed (caused by incorrect compensation or incompatible antibody combinations) and reproducibility may suffer. The importance of optimized panel design in experimental design cannot be overemphasized. However, panel design is a lengthy and labor-intensive process. The process is reviewed elsewhere³², but briefly, it consists of the following steps: 1) stratification of the markers into those required for the study, those that would be useful, and those that could be interesting but are not necessary, 2) purchase and titration of a wide variety of antibody reagents specific for the markers considered, 3) ranking of the best antibody reagents for each marker, 4) characterization of expression levels (does the marker exhibit on/off expression, a continuum of staining, or is it dim?), 5) construction of putative panels by placing the dim markers on the brightest channels (PE, APC, and now BV421) and markers with on/off expression on the dimmest channels (QD565) or channels that are subject to significant spreading (Cy7APC), and 6) testing of the panels by staining tubes with increasing numbers of reagents, and examining the staining patterns and spreading of each panel upon the addition of a new antibody. Repeated iteration of the last two steps with alternate reagent combinations is needed to optimize the panel.

As with instrument calibration, there are opportunities to improve this process using automated tools. In fact, such tools could be integrated into instrument calibration software by using calculated Q and B values to determine the minimal signal that can be separated from background for each detector. This information could be used to determine the best channels for markers with on/off, continuum, or dim expression, and to build putative panels. Using theoretical compensation matrices for each putative panel, built from the titration data of each reagent, the spreading error in each detector could be calculated to predict experimental results. Dim markers that overlap with the negative population, after adjustment of spreading, could be flagged, or excluded, so that a final list of optimal staining combinations would be reported to the investigator.

Better Strategies for Data Analysis

Once instrument calibration, reagent choice, and panel design are complete, investigators face an enormous hurdle of data analysis. A fundamental advantage of polychromatic flow cytometry is the ability to examine finely-defined subsets of cells. A few such subsets can be identified *a priori* by the investigator; however, in typical experiments, there is an interest in exploring hundreds of phenotypic combinations (including all the markers). A number of tools and approaches are currently being evaluated for use; these include multidimensional visualization (exploration) tools, gating tools, and post-analysis data aggregation tools.

When multiple markers and hundreds of phenotypic combinations are available for exploration, the ability to visualize data in multiple dimensions becomes important. To this end, polychromatic plots³³ have been developed. These are similar to standard dot plots, with the important exception that the color of each dot varies according to the expression of three other markers. Thus, every event is encoded with a shade of red, green, and blue to reflect the expression levels in three additional dimensions. The function that encodes color mapping can be altered, as can the priority of colors/markers, so that various populations can be emphasized. In this way, a two-dimensional dot plot can be translated into a five-dimensional visualization. Although great care must be taken in analyzing and interpreting data generated this way, the method is powerful and accessible (since it preserves the dot plot format we are used to seeing).

When gates defining each phenotypic combination are required for downstream analysis, Boolean algorithms are helpful. Such algorithms require simply a single gate identifying positive cells for each marker, from which negative gates are imputed, and Boolean combinations are constructed consisting of every possible combination. For example, if CD45RA, CCR7, and CD27 are put into the algorithm, the gates defining the following cell types are constructed: CD45RA+ CCR7+ CD27+, CD45RA+ CCR7+ CD27–, CD45RA+ CCR7– CD27+, CD45RA+ CCR7– CD27–, etc. This allows rapid enumeration of cells expressing these combinations of markers, through automated construction of the series of gates necessary for identification. A disadvantage of this tool is that it assumes that all subsets can be discriminated with equal sensitivity; this may not always be the case.

In addition to the tools available for visualization of staining patterns, gating, and phenotyping, specialized software can aggregate the frequency of every cell type across multiple specimens. For example, SPICE software³⁴ performs this function, and joins categorical data (time point, disease condition), allowing rapid statistical comparison of cell frequencies across multiple different conditions. In addition, the complete dataset can be visualized as scatter plots, bar graphs, or pie charts, and overlaid with categorical variables. Finally, data can be normalized for background biological controls, as is required for intracellular cytokine assays (where data from mock-stimulated control samples is subtracted from each condition).

There are a number of considerations for employing these and similar approaches. First, when performing hypothesis-driven research, a single subset (or a couple of related subsets) is identified to test against a biological or disease outcome. However, this ignores the bulk of the data generated in the experiment, and limits the ability to detect new relationships. Using Boolean gating and SPICE, the majority of the data can be examined; however, many users consider testing only the terminal phenotypes defined by the Boolean gates (for example, CD45RA– CCR7+ CD127+ CD27+ CCR5– CD57– CD28– cells) against the outcome. What if the cell type relevant to disease is defined by some lesser combination of markers (a parent population, such as CD45RA– CD127+ CD57– cells)? Such tests are simple to perform in SPICE, but the software cannot perform multiple comparisons adjustments; the need for which is compounded by the additional number of populations tested. To compensate for this, the p-value for which differences are considered significant must be reduced in SPICE.

The next generation of cytometry tools may automate many of these processes. Using tools like FlowMeans³⁵, gates can be identified automatically (with a user's input if needed) through computational formulas that examine the pattern of staining, and all Boolean combinations of every combination of markers can be constructed. Thus, the subjectivity and labor of gating are partially eliminated, and a greater number of subsets (e.g., “parent” populations) are included in the analysis.

A related tool, FlowType, takes these phenotypes and tests their relationship to a biologic measurement such as clinical outcome. For example, the tool can identify the cell types for which the association between subset frequency and patient survival time is statistically significant. Since there can be tens of thousands of phenotypes tested in such a scheme, adjustment of statistical significance for multiple comparisons is done. A subsequent analysis can test the correlation between all significant phenotypes, and define clusters of phenotypes that are highly related. From these, a representative phenotype is chosen, and each of the markers defining that phenotype is tested for the necessity of including it, in terms of predicting the biological outcome. Markers with less impact can be dropped from the phenotype, thereby reducing a dataset with tens of thousands of cell populations to only the simplest phenotypes with the strongest association to biological outcome. By defining

the simplest phenotypes with biological relevance, it informs future experiments (which can eliminate the less relevant markers) and allows researchers in resource-poor or strictly regulated clinical settings to design the simplest, most meaningful immunophenotyping panels possible.

It is important to realize that these approaches rely on clusters of cells (cell populations) that are tested for biological importance. However, what if the cells of interest do not fall into a distinct cluster defined by a particular phenotype, but are instead enriched within a region of multidimensional space that cuts across defined gates? Two rarely-employed methods, known as probability binning and frequency difference gating^{36,37}, are useful in this regard. The process can be described as follows: First, data from patients within the same disease group is concatenated into a single file. This has the important advantage of increasing the number of events examined, particularly for rare, antigen-specific responses. Data from one patient group is then divided into multi-dimensional “bins” (spanning the measurement space) which contain roughly equal numbers of events; these bins are also then applied to the concatenated data from other disease groups. For each bin, the number of events in the disease groups are compared, and a test statistic (describing the degree of difference) is generated. The bins are then ranked by this test statistic (i.e., from the most similar to the least similar). The bins with most significant differences in frequency represent regions of multidimensional measurement space that identify cells that are present at different frequencies between disease states – importantly, not just present vs. absent, but different in representation. Furthermore, the bins themselves can be used as a “gate” for further analysis and enumeration of those cells.

Recently, two unique approaches for data analysis have been introduced; both tools can examine the relationship between markers or cell populations more directly than the approaches described above. The first is called spanning-tree progression analysis of density-normalized events (SPADE), and consists of four steps³⁸. The initial step takes the dataset, consisting of hundreds of thousands of events, and downsamples it in a density-dependent manner so that it is computationally manageable (but still reflects the original population frequencies). The second step clusters the data, while the third step links the clusters using minimum spanning trees. Finally, the data is upsampled to restore all the cells from the dataset. Results appear in tree-like graphics, where branches represent the relationships/hierarchy underlying the dataset, and the cell populations are depicted as nodes along these branches (colored, as is done for heat maps, to reflect frequency). The algorithm has particular utility in complex, multiparameter datasets generated to explore the relationships between cell types or the response to stimuli; however, it is less suited for comparison of samples across multiple individuals or disease conditions (where statistics are required to confirm differences). In some respects, it is similar to a second recently introduced approach, which is employed in Gemstone software (Verity House). This system uses probability state modeling to reveal relationships between many markers by examining them in the context of one, or a few, markers that describe the progression of a cell population³⁹. To do this, an initial modeling step is necessary, after which the variations in markers from the complete dataset can be visualized in the context of the modeled markers using ribbon pots.

In theory, the methods described above have the power to detect biologically important cells when the precise phenotypic definition of these cells is not known. That is, a biologically important cell type may uniquely express a marker that has not been measured in the polychromatic staining panel. However, this cell type is likely to share expression of other markers with closely related, but less biologically relevant cells. In this scenario, the probability binning/frequency difference gating may be able to detect these cells within a slice (or bin) of multidimensional space, or SPADE and Gemstone may be used to reveal

other candidate markers to describe the key population. Notably, the concept that our assays may not identify a uniquely important marker (because we did not or could not measure it) suggests we need more multiparametric technology than available in current, state-of-the-art flow cytometers.

New Technology Platforms and Their Interface with Flow Cytometry

In recent years, the need for more multiparametric technologies has given rise to new platforms for the characterization of single cells. These technologies include the FluidigmBioMark system, which measures 96 gene transcripts simultaneously from a single cell, and Cytometric Time of Flight (CyTOF), which adapts inductively coupled mass spectrometry for analysis of cellular protein expression. Both systems offer distinct advantages over polychromatic flow cytometry, not only in terms of the number of parameters measured, but for other reasons as well (discussed below). However, both technologies still require flow cytometry to maximize their utility.

The Fluidigm system can quantify mRNA levels for 96 different gene transcripts from a single cell using highly-multiplexed TaqMan real-time PCR. The measurement of mRNA (rather than protein) has advantages and disadvantages. Detection of mRNA is far more sensitive than protein, because the cDNA resulting from reverse transcription can be amplified to detect even a single copy. Moreover, the primers and probes used in gene expression assays are far easier to find, produce, and use than antibody probes. However, important caveats include artefacts arising from low levels of contamination, the inherent noise of gene expression measurements in exponential amplification systems, and the reality that mRNA expression does not always reflect protein expression. Still, the highly multiplexed nature of this technology allows rapid and economical screening of a great many markers that might be related to a disease process or immune response, from a small cell sample. Markers with an interesting relationship to biological outcome could then be studied further by flow cytometry.

It is important to note that, although this technology has advantages over flow cytometry, it measures a small number of cDNAs at a time (96 per plate) – i.e., 96 cells per plate. The power of this technology comes from the use of flow cytometry sorters depositing single desired cells in the assay wells. Additionally, by coupling sorting technology with indexed sorting (the exact identification of each cell deposited into a well), the immunofluorescence profile of every single sorted cell can be included in the correlations of mRNA expression levels. Such studies provide an unprecedented look at the gene and protein level expression of molecules associated with an immune response.

An example of data obtained through a coordinated flow cytometry and Fluidigm experiment is provided in Figure 5. Here, CD8⁺ T-cells specific for the NLV epitope of CMV were identified, phenotyped, and sorted by flow cytometry; the sorted cells were then analyzed by Fluidigm. Remarkably, incredible diversity was observed within the antigen-specific CD8⁺ T-cell population, which included cells making most transcripts, and cells making very few transcripts. This diversity underlies the relatively homogeneity of this population in terms of differentiation (most cells were effector memory T-cells by flow cytometry). Thus, the analysis of additional parameters reveals that even cells specific for a single antigenic peptide can be remarkably varied.

CyTOF is a second novel technology for cytometric analysis⁴⁰. In a manner analogous to flow cytometry, CyTOF experiments begin with the conjugation of antibodies to polymers carrying isotopes of a family of rare earth elements (lanthanides). Thus, “metal” isotopes replace the fluorescent dyes used in flow cytometry, and time-of-flight mass spectrometry replaces laser-excited fluorescence emission for quantitation of cell-bound antibody.

Antibody conjugates are used to stain cells, using the same procedures as flow cytometric cell staining.

There are three potential advantages to this technology. First, by virtue of the range of unique atomic mass elements, it allows measurement of 33 (and potentially up to 100) markers simultaneously. Second, it has the potential for high sensitivity, because the metal probes are not present naturally in our cells and therefore have no background signal. This is in stark contrast to flow cytometry, where auto fluorescence of cells in some regions of the spectrum significantly reduces sensitivity (i.e., the ability to detect dim staining). Third, the signals from metal tags does not overlap, leading to simple “panel” design (the steps in the fluorescence panel design process are largely irrelevant). In this way particularly, CyTOF can be used as a simple screening tool when multiple markers are identified as potentially interesting in Fluidigm experiments. As a new technology, the staining patterns obtained with this technology must be validated by flow cytometry. Moreover, given the unusual background characteristics (nominally zero for most measurements), typical data visualization tools such as the biexponential transform are inadequate. Future development of these, and other early instrumentation and reagent issues, is likely to be highly informed by the early years of flow cytometric analysis.

Conclusion

With the evolution of flow cytometry technology, we have developed a remarkable understanding of the complex, and varied, facets of human immunity. Early achievements were driven by the introduction of new hardware, while the ability to multiplex measurements was supported later on by the introduction of new fluorochromes. To a large extent, we now possess the tools to perform high complexity, polychromatic experiments routinely and this gives us the potential to examine human immunity and disease with incredible resolution. However, associated methods (e.g., for QC, for data analysis) require more significant development, so that they can be efficient and accessible to general users. As new technologies emerge, like CyTOF and Fluidigm, the experience gained from developing flow cytometry reagents, QC methods, and data analysis tools will be invaluable. Most importantly, however, the ability to validate results and acquire complementary data easily using flow cytometry will be absolutely critical. Thus, despite the emergence of new cytometry technologies, the future of polychromatic flow cytometry development and experimentation is likely to remain quite vibrant.

References

1. Hulett HR, Bonner WA, Barrett J, Herzenberg LA. Cell sorting: automated separation of mammalian cells as a function of intracellular fluorescence. *Science*. 1969; 166:747–749. [PubMed: 4898615]
2. Wider ME, Cram LS. Differential fluorochromasia of human lymphocytes as measured by flow cytometry. *J Histochem Cytochem*. 1977; 25:888–891. [PubMed: 70458]
3. Chattopadhyay PK, Roederer M. Good cell, bad cell: flow cytometry reveals T-cell subsets important in HIV disease. *Cytometry A*. 2010; 77:614–622. [PubMed: 20583275]
4. De Rosa SC, Herzenberg LA, Roederer M. 11-color, 13-parameter flow cytometry: identification of human naive T cells by phenotype, function, and T-cell receptor diversity. *Nat Med*. 2001; 7:245–248. [PubMed: 11175858]
5. De Rosa SC, et al. Vaccination in humans generates broad T cell cytokine responses. *J Immunol*. 2004; 173:5372–5380. [PubMed: 15494483]
6. Chattopadhyay PK, et al. Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. *Nat Med*. 2006; 12:972–977. [PubMed: 16862156]

7. Perfetto SP, Chattopadhyay PK, Roederer M. Seventeen-colour flow cytometry: unravelling the immune system. *Nat Rev Immunol.* 2004; 4:648–655. [PubMed: 15286731]
8. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4+ CD25+ regulatory T cells. *Nat Immunol.* 2003; 4:330–336. [PubMed: 12612578]
9. Crotty S. Follicular helper CD4 T cells (TFH). *Annu Rev Immunol.* 2011; 29:621–663. [PubMed: 21314428]
10. McAleer JP, Kolls JK. Mechanisms controlling Th17 cytokine expression and host defense. *J Leukoc Biol.* 2011; 90:263–270. [PubMed: 21486905]
11. Mahnke YD, Roederer M. OMIP-001: Quality and phenotype of Ag-responsive human T-cells. *Cytometry A.* 2010; 77:819–820. [PubMed: 20722005]
12. Chattopadhyay PK, Hogerkorp CM, Roederer M. A chromatic explosion: the development and future of multiparameter flow cytometry. *Immunology.* 2008; 125:441–449. [PubMed: 19137647]
13. Chattopadhyay PK, et al. Techniques to improve the direct ex vivo detection of low frequency antigen-specific CD8+ T cells with peptide-major histocompatibility complex class I tetramers. *Cytometry A.* 2008; 73:1001–1009. [PubMed: 18836993]
14. Zhou T, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science.* 2010; 329:811–817. [PubMed: 20616231]
15. Itoh M, et al. Thymus and autoimmunity: production of CD25+ CD4+ naturally anergic and suppressive T cells as a key function of the thymus in maintaining immunologic self-tolerance. *J Immunol.* 1999; 162:5317–5326. [PubMed: 10228007]
16. Thornton AM, Shevach EM. CD4+ CD25+ immunoregulatory T cells suppress polyclonal T cell activation in vitro by inhibiting interleukin 2 production. *J Exp Med.* 1998; 188:287–296. [PubMed: 9670041]
17. Suri-Payer E, Amar AZ, Thornton AM, Shevach EM. CD4+CD25+ T cells inhibit both the induction and effector function of autoreactive T cells and represent a unique lineage of immunoregulatory cells. *J Immunol.* 1998; 160:1212–1218. [PubMed: 9570536]
18. Fontenot JD, et al. Regulatory T cell lineage specification by the forkhead transcription factor foxp3. *Immunity.* 2005; 22:329–341. [PubMed: 15780990]
19. Kochetkova I, Thornburg T, Callis G, Pascual DW. Segregated Regulatory CD39+CD4+ T Cell Function: TGF- β -Producing Foxp3 $^{-}$ and IL-10-Producing Foxp3 $^{+}$ Cells Are Interdependent for Protection against Collagen-Induced Arthritis. *J Immunol.* 2011; 187:4654–4666. [PubMed: 21967895]
20. Liu W, et al. CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J Exp Med.* 2006; 203:1701–1711. [PubMed: 16818678]
21. Chattopadhyay PK, Perfetto SP, Roederer M. The colorful future of cell analysis by flow cytometry. *Discov Med.* 2004; 4:255–262. [PubMed: 20704956]
22. Darrah PA, et al. Multifunctional TH1 cells define a correlate of vaccine-mediated protection against *Leishmania major*. *Nat Med.* 2007; 13:843–850. [PubMed: 17558415]
23. Betts MR, et al. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood.* 2006; 107:4781–4789. [PubMed: 16467198]
24. Duvall MG, et al. Polyfunctional T cell responses are a hallmark of HIV-2 infection. *Eur J Immunol.* 2008; 38:350–363. [PubMed: 18200635]
25. Precopio ML, et al. Immunization with vaccinia virus induces polyfunctional and phenotypically distinctive CD8(+) T cell responses. *J Exp Med.* 2007; 204:1405–1416. [PubMed: 17535971]
26. Gaucher D, et al. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *J Exp Med.* 2008; 205:3119–3131. [PubMed: 19047440]
27. Sun Y, et al. Recombinant vector-induced HIV/SIV-specific CD4+ T lymphocyte responses in rhesus monkeys. *Virology.* 2010; 406:48–55. [PubMed: 20667574]
28. Roederer M. Spectral compensation for flow cytometry: visualization artifacts, limitations, and caveats. *Cytometry.* 2001; 45:194–205. [PubMed: 11746088]
29. Perfetto SP, Ambrozak D, Nguyen R, Chattopadhyay P, Roederer M. Quality assurance for polychromatic flow cytometry. *Nat Protoc.* 2006; 1:1522–1530. [PubMed: 17406444]

30. Chattopadhyay PK. Quantum dot technology in flow cytometry. *Methods Cell Biol.* 2011; 102:463–477. [PubMed: 21704850]
31. Zarkowsky D, et al. Heavy metal contaminants can eliminate quantum dot fluorescence. *Cytometry A.* 2010
32. Mahnke YD, Roederer M. Optimizing a multicolor immunophenotyping assay. *Clin Lab Med.* 2007; 27:469–485. v. [PubMed: 17658403]
33. Roederer M, Moody MA. Polychromatic plots: graphical display of multidimensional data. *Cytometry A.* 2008; 73:868–874. [PubMed: 18629844]
34. Roederer M, Nozzi JL, Nason MC. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytometry A.* 2011; 79:167–174. [PubMed: 21265010]
35. Aghaeepour N, Nikolic R, Hoos HH, Brinkman RR. Rapid cell population identification in flow cytometry data. *Cytometry A.* 2011; 79:6–13. [PubMed: 21182178]
36. Roederer M, Hardy RR. Frequency difference gating: a multivariate method for identifying subsets that differ between samples. *Cytometry.* 2001; 45:56–64. [PubMed: 11598947]
37. Roederer M, Moore W, Treister A, Hardy RR, Herzenberg LA. Probability binning comparison: a metric for quantitating multivariate distribution differences. *Cytometry.* 2001; 45:47–55. [PubMed: 11598946]
38. Qiu P, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat Biotechnol.* 2011; 29:886–891. [PubMed: 21964415]
39. Bagwell CB. Breaking the dimensionality barrier. *Methods Mol Biol.* 2011; 699:31–51. [PubMed: 21116977]
40. Bendall SC, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science.* 2011; 332:687–696. [PubMed: 21551058]

Figure 1.
A timeline illustrating coordinates advances in flow cytometry technology and understanding of the complexity of the T-cell compartment.

Figure 2. Bulk measurements introduce noise into assays, thereby reducing the power to detect populations with disease or vaccine relevance

For example, if the population that correlates strongly with a disease or vaccine outcome consists of IFN γ + IL2- TNF+ CD107a+ cells, the measurement of these four markers will reveal the correlation. However, if only three markers are measured, an irrelevant cell type is included (TNF- cells), introducing noise into the assays, and reducing the correlations. When additional markers are omitted, even more noise is introduced, further reducing the power to detect correlations.

Figure 3. Common fluorochromes used in flow cytometry, and the variety of reagents commercially available for each

Note the dearth of commercially available reagents that can be excited by the violet laser. Abbreviations for commercial fluorochromes: QD, quantum dot; A, Alexa (e.g., Alexa-488); Cy, Cyanine; PerCP, peridinin chlorophyll; PE, phycoerythrin; ECD, energy coupling dye (also known as PE-Texas Red); D, DyLight; HL, HiLyte.

Figure 4. Properties and applications of BV421

A) Schematic illustrating polymer structure of BV421, as conjugated to antibody. Energy from a violet laser ($h\nu_a$) excites electrons in the polymer and blue light ($h\nu_e$) is emitted. Excitation (purple) and emission spectra (blue) for BV421 are also shown. B) Staining patterns for BV421 (left), Pacific Blue (read on the same channel as BV421, middle), and PE (right) conjugates of anti-human CD8. C) Identification of CMV-NLV specific CD8⁺ T-cells using pMHCImultimers made with streptavidin-BV421 (left), -APC (middle), or -QD605 (right).

Figure 5. Fluidigm analysis of CD8+ T-cells specific for HLA-A2 NLV epitope of CMV PBMC were stimulated with NLV peptide for three hours (without monensin or brefeldin A). Results are presented as a heat map, where each row represents the gene expression profile of a single cell (e.g., cell A and cell B, pink box), and each column represents the expression level of one of the assayed genes (e.g., transcript I and transcript T, pink arrows). Levels of each transcript are potted on a gray scale, from low expression (white) to high expression (black). To simplify presentation, individual gene names are not shown; instead, genes are grouped into modules, as indicated at the bottom of the figure. The final four columns in the map depict composite expression profiles, which show the total level of expression for all the measured transcripts (1), and for those transcripts associated with cellular activation (2), cytolysis (3), and cytokines (4), within each cell.