Published in final edited form as:

Cytometry A. 2009 September ; 75(9): 734–742. doi:10.1002/cyto.a.20768.

Hierarchical Clustering of Monoclonal Antibody Reactivity Patterns in Nonhuman Species

Juan Pablo Pratt^{*}, Qing Zeng^{†,‡}, Dino Ravnic^{*}, Harold Huss^{*}, James Rawn[‡], and Steven J. Mentzer^{*}

^{*} Laboratory of Immunophysiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

[†] Decision Systems Group, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

[‡] Harvard-MIT Division of Human Sciences and Technology, Cambridge, MA

Abstract

Monoclonal antibodies are an important resource for defining molecular expression and probing molecular function. The characterization of monoclonal antibody reactivity patterns, however, can be costly and inefficient in nonhuman experimental systems. To develop a computational approach to the pattern analysis of monoclonal antibody reactivity, we analyzed a panel of 128 monoclonal antibodies recognizing sheep antigens. Quantitative single parameter flow cytometry histograms were obtained from five cell types isolated from normal animals. The resulting 640 histograms were smoothed using a Gaussian kernel over a range of bandwidths. Histogram features were selected by SiZer—an analytic tool that identifies statistically significant features. The extracted histogram features were compared and grouped using hierarchical clustering. The validity of the clustering was indicated by the accurate pairing of externally verified molecular reactivity. We conclude that our computational algorithm is a potentially useful tool for both monoclonal antibody classification and molecular taxonomy in nonhuman experimental systems.

Keywords

Other animals; Antibodies; Antigens/Epitopes

Introduction

The focus of the early Human Leukocyte Differentiation Antigens Workshops (HLDA) was the use of monoclonal antibodies (Mab) and flow cytometry to identify new molecules (1–7). The Workshops characterized the reactivity patterns of large panels of antibodies against a panel of cell populations. For practical purposes, most of the expression analyses were performed with widely available cell lines. Since multi-passaged cell lines generally represented uniform cell populations, the results of the analyses were typically binary; that is, the antibody either reacted with the cell line or it did not. Although the histograms derived from cell lines were typically amenable to parametric analyses, the antibody reactivity was effectively summarized as the percentage of cells with fluorescence intensity greater than the negative control. This “percent positive” data was subjected to hierarchical clustering to identify “clusters of differentiation” or CD specificities (1). In human biology, the ability to identify the gene corresponding to the target antigen has rendered this Workshop approach to

identifying new molecules obsolete. Mab, however, continue to be an important resource for defining molecular expression and probing molecular function.

Despite the utility of monoclonal antibodies in biologic investigations, attempts to develop antibody panels in other species has had limited success. A practical problem is that the HLDA approach requires widely available cell lines—a resource that is uncommon in most species. An alternative approach is to use isolated cell populations such as peripheral blood, thymocytes and splenocytes. These cell populations, however, are nonuniform. Flow cytometry analyses with most Mab produce histograms that cannot be effectively clustered using “percent positive” data. Furthermore, the visual inspection method of pairwise comparison becomes overwhelming with even a few dozen antibodies.

To compare reactivity patterns of 128 monoclonal antibodies (Mab) produced against sheep antigens, we developed an algorithm of heirarchical clustering based on nonparametric feature identification of quantitative single parameter flow cytometry histograms. After kernel smoothing, statistically significant features were compared by hierarchical clustering. The accurate pairing of known Mab suggest the utility of this approach in the characterization and comparison of Mab in nonhuman experimental systems.

Methods

Mice

Balb/c mice (Jackson Laboratory, Bar Harbor, Maine, USA), 25–33g, were used in all Mab production experiments. The care of the animals was consistent with guidelines of the American Association for Accreditation of Laboratory Animal Care (Bethesda, MD).

Sheep

Randomly bred female sheep (Parsons Farm, Springfield, Massachusetts, USA), 6–9 months old, ranging in weight from 35–50kg were used in these studies. They were given free access to food and water. Sheep were excluded from the study if there was any gross or microscopic evidence of infection. The care of the animals was consistent with the guidelines of the American Association for Accreditation of Laboratory Animal Care (AAALC) (Bethesda, Maryland, USA) and the Guide for the Care and Use of Laboratory Animals (Department of Health and Human Services, Publication No. 85-23).

Fusion of murine splenocytes

Of the 128 monoclonal antibodies (Mab), 112 were produced in our laboratory. Myeloma and spleen cells were infused at a ratio of five splenocytes per NS1 myeloma cell (8) with polyethylene glycol 4000 (Sigma; St. Louis, Missouri, USA) (50%, weight/volume) as a fusing agent (9). Mouse myeloma cells and immune splenocytes in serum-free medium were centrifuged at 300×g for 10 minutes. The cells were gently suspended in 1ml of fusogen for 30 seconds at 37°C and then slowly diluted with the 20ml of serum-free medium. The cells were again centrifuged for five minutes at 250×g and gently resuspended in culture medium containing hypoxanthine-aminopterin-thymidine (HAT; Gibco Laboratories, Grand Island, New York, USA) supplemented with fetal Serum (20%, v/v) and endothelial growth supplement (Sigma) (10). The cells were plated in 96-well flat-bottom microtiter culture plates (Costar; Corning, New York, USA) at approximately 5×10^4 cells per well. The cells were cultured at 37°C in a 5% CO₂ incubator and examined for hybridoma colonies at two-day intervals starting seven days after the fusion. The productive hybridomas, identified both by antibody concentration and antigen binding (11,12), were selected for further analysis.

Monoclonal antibodies

The Mab were cloned and subcloned by limiting dilution. The hybridoma colonies were monoclonal based on statistical analysis of both limiting dilution and multi-plexed bead isotyping (13). The anti-CD4 Mab 17D (14), the anti-LFA-1 Mab F10-150 (14), the beta1 integrin Mab ERD2/117 and FW4-101-1 (15) and the class I Mab T2/39 (16) have been previously described. Additionally, the anti-macrophage Mab M3/246 as well as the anti-lymphocyte Mab L2/34 and T4/268 were produced and characterized as previously described (12,17,18). In addition to flow cytometry, all antibodies were characterized by immunohistochemistry and immunoprecipitation. Affinity purified Mab or cultured supernatants used at five-fold saturating concentrations were used in all experiments.

Cell surface immunofluorescence

Cells were washed twice with phosphate-buffered saline containing 2.5% fetal calf serum and 0.02% sodium azide. Approximately 10^6 cells were incubated on ice for 30 minutes with an excess concentration of Mab. The cells were washed twice and stained with fluorescein-conjugated goat F(ab')² anti-mouse immunoglobulin (SouthernBiotech, Birmingham Alabama, USA) diluted 1:10. The cells were incubated on ice for another 30 minutes. After 3 washes, the cells were fixed in 1% para formaldehyde and analyzed on an Epics XL flow cytometer (Beckman Coulter, Miami, Florida, USA). The antibodies were tested against a panel of 5 sheep cell types: thymocytes, lymph node cells, alveolar macrophages, efferent lymphocytes and splenocytes.

Thymocytes

The thymus was harvested immediately after euthanasia through a median sternotomy and dissection of both lobes of the thymic gland. The thymus was dissected into 3mm cubes of tissue and placed in RPMI 1640, 10% fetal calf serum (FCS) and squeezed through a steel mesh. After mechanical dissociation, the thymocytes were washed twice in RPMI containing 2.5% FCS prior to immunostaining.

Efferent lymphocytes

Prescapular efferent lymph duct cannulations of normal 30–45 kg ewes were used as a source of lymphocytes (19). The lymphocyte-containing lymph plasma was collected in sterile polypropylene tubes (Falcon, Franklin Lakes, New Jersey, USA) or plastic bags (Baxter Healthcare, Deerfield, Illinois, USA) containing 300 units of heparin, 1000 units of penicillin and 1mg of streptomycin (Sigma-Aldrich, St. Louis, Missouri, USA). The cellular component was separated by velocity centrifugation and resuspended in cell media containing RPMI 1640 (Sigma-Aldrich) supplemented by 3% FCS (Sigma-Aldrich) and 1% deoxyribonuclease (Worthington Biochemical, Lakewood, New Jersey, USA) at a concentration of 5.0×10^6 cells/ml.

Lymph node cells

The sheep prescapular lymph node was harvested immediately after euthanasia and perinodal fat removed by sharp dissection. The lymph node was dissected into 3mm cubes of tissue and processed by mechanical shear using the glazed ends of two glass microscope slides. The cells were filtered and serially washed prior to preparation for immunostaining.

Alveolar macrophages

With airway protection using a 9Fr endotracheal tube (Bard, Covington, California, USA), an 18Fr Nasogastric tube (Bard) was passed into the mainstem bronchi. The tube was used to instill 1000cc of 27°C normal saline. A separate suction catheter was subsequently placed within

the endotracheal tube to remove the lavage fluid. The procedure was repeated for the contralateral lung. The samples were passed through a strainer and washed twice in culture media (RPMI 1640; Sigma) containing 10% FCS. The cells were counted and viability assessed prior to preparation for immunostaining.

Splenocytes

The spleen was harvested immediately after euthanasia through a midline laparotomy incision. The splenic hilum was ligated and the spleen removed. The spleen was dissected into 3mm cubes of tissue and placed in RPMI 1640, 10% FCS and squeezed through a steel mesh. After mechanical dissociation, the cells were filtered and washed. The cells were centrifuged at 200g for 10 min and resuspended in an ammonium chloride lysis buffer for 10 minutes at 27°C. The splenocytes were washed twice in RPMI containing FCS with the final resuspension in preparation for immunostaining.

Quantitative flow cytometry

The immunostained cells were analyzed on an Epics XL (Beckman Coulter; Maimi, Florida, USA) equipped with a single laser with excitation wavelength at 488nm and three emission detectors. Gain settings were calibrated to 4 peak Rainbow calibration particles (Spherotech; Libertyville, Illinois, USA). During the experiments the laser power, photomultiplier tube voltage, light scatter and fluorescent gains were kept constant. A total of 10,000 events were acquired from each sample. The data was processed using WinList 5.0 (Verity; Topsham, Maine, USA) and exported to Microsoft Excel (Redmond, Washington, USA) for further analysis. The output is a summary graph, or histogram, of the frequency distribution of the data. In flow cytometry nomenclature, the data is grouped in “channels” which corresponds to “bins” in statistical nomenclature.

Expert panel

The correspondence between histogram features identified by the computational algorithm and visual inspection was assessed by panel of flow cytometry users judged to be “experts” by peer evaluation. The panel was independently shown the 640 histograms in random order and asked to identify the number of biologically meaningful peaks.

Kernel smoothing

Kernel smoothing is a general-purpose statistical technique for removing noise and highlighting structure in nonparametric datasets (20). The spatial resolution of kernel smoothing is determined by the bandwidth: a narrow bandwidth provides a highly resolved density estimation whereas a wide bandwidth produces a highly smoothed density estimation (21). In kernel smoothing, the estimation of the density at a particular location x and for bandwidth h is

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)$$

The kernel K in our application was a Gaussian kernel; one implication of which was that the number of zero crossings of the derivative is always a decreasing function of h (22).

SiZer: Significant Zero crossings

The statistical technique for defining statistically significant peaks has been described in detail (23). Briefly, the technique uses a family of curves

$$\{\widehat{f}_h(x): h \in [h_{\min}, h_{\max}]\}$$

to assess the significant features. The bandwidth range

$$[h_{\min}, h_{\max}]$$

is taken to be $h_{\min}=2(\text{bin width})$ and $h_{\max}=\text{the data range}$. The subsequent map was based on the confidence limits given by

$$\widehat{f}_h'(x) \pm qsd(\widehat{f}_h'(x))$$

The application of SiZer was possible using MatLab 7.0 (MathWorks, Natick, Massachusetts, USA) and software downloadable from Professor Marron's website http://www.stat.unc.edu/faculty/marron/marron_software.html.

Hierarchical clustering

First, the listmode standard FCS 2.0 files (24) were converted to a Matlab (ver. 7.0.4; MathWorks) array. In addition to the raw data, the FCS 2.0 structure specifies the instrument used for the measurement and the sample measured. The clustering process contains the following steps:

1. Feature extraction: on each organ, there is a histogram h corresponding to each antibody. The histogram is processed using SiZer, a kernel smoothing software. SiZer smoothes data at n even levels and generates a family n smoothed histograms h_p . We identify peak locations x_p ($h_p'(x)=0$ and $h_p(x)>h_p(x-1)$) in the smoothed histograms as well as peak height $h_p(x_p)$ as features. This condition applies to integer x values; a typical condition of flow cytometry datasets.
2. Feature representation: The features (height and location) of the first x peaks from each of the y smoothed histograms ($y<n$) are used for further clustering. When a histogram had fewer than the maximum number of peaks, the absent peak location and height values are set to zero. Thus, each antibody is represented by $2*x*y$ (or mathematically notated $2xy$) peak features. When clustering antibodies based on data from z organs, each antibody is represented by $2*x*y*z$ (or $2xyz$) features.
3. Clustering: the Euclidean distances between every pair of histograms are calculated based on the peak features; that is, the peak features of each individual histogram is used to represent differences between histograms. A hierarchical cluster tree is then created based on the unweighted average distance between clusters.

In this experiment, we chose 11 smoothing levels and extracted features from level 2 to 9, which led to $y=8$. The number of peaks per histogram used for clustering is set to 4, i.e. $x=4$.

Results

Mab panel and quantitative flow cytometry

Meaningful clustering requires that the histogram features are both varied and reproducible. To provide this variability and avoid duplication, we selected a panel of 128 unique monoclonal

antibodies. When tested against naturally occurring sheep cells, the single parameter flow cytometry analysis of Mab staining demonstrated a range of histogram features (Figure 1). More than one Mab recognizing LFA-1 (N=2), CD29 (N=2), class II (N=2), and class I (N=3) were included in the panel because these antibodies recognized unique epitopes as demonstrated by crossblocking experiments (not shown). These antibodies provided an externally verified test of clustering accuracy.

To simplify machine and reagent calibration, the analysis was performed with single parameter flow cytometry. The use of a single fluorochrome 1) eliminated the need to normalize the integral scale of different fluorochromes (e.g. fluorescein and phycoerythrin), and 2) facilitated quantitative flow cytometry data collection. Flow cytometry histograms obtained over a 6 year period were highly reproducible (Figure 2A). Biologic variability, as reflected in the population dynamics of tissues such as the efferent lymph, lymph node, alveolar macrophage, spleen and thymus (Figure 2B–F) was minimized by selecting a modal histogram representative of the results obtained from multiple animals.

Kernel smoothing and feature identification

To identify the statistically significant underlying structure in flow cytometry histograms, the histogram was smoothed using a Gaussian kernel (22) and a confidence interval was determined for the derivative of the smoothed curve at each location along the histogram. This technique, developed by Chaudhuri and Marron (23), identifies a statistically significant peak when the confidence interval is completely above zero; hence, “SiZer” defines the Significant Zero crossing. Of note, only 152 of the 640 histograms (23.8%) were judged to be unimodal by an expert panel.

Although bin width is fixed by the cytometer calibration, data analysis must still choose the level of smoothing or bandwidth (h). Since there is no broadly applicable “data-driven” method for bandwidth selection (20), we included a range of smoothing levels in the data analysis by constructing a confidence interval for the derivative of the smoothed curve at each location along the histogram and at a range of bandwidths (Figure 3A,B). Since the location of significant zero crossings varies with location on the intensity (x) axis, the SiZer output can be plotted graphically. The location of significant histogram features can be identified by color transition (Figure 3C,D).

Hierarchical clustering of anti-sheep Mab

To test the utility of SiZer-derived feature identification in data clustering, the reactivity of a panel of 128 Mab was tested against five sheep cell types: efferent lymphocytes, lymph node cells, alveolar macrophages, splenocytes and thymocytes. These cell types were chosen because of their availability and our interest in leukocyte biology. The dataset from hierarchical clustering was comprised of the modal histogram of three or more replicates. By convention, the resulting dendrogram is presented with the horizontal position reflecting the combined metric information (Figure 4). Given the cell types analyzed in our panel, antibodies recognizing target molecules with high cell surface density, such as major histocompatibility complex class I or CD45 molecules, are displayed rightward. In contrast, antibodies with little or no reactivity with these cell types would be plotted on the left side of the dendrogram.

Dissimilarity index

The utility of hierarchical clustering depends upon the accurate estimation of similarity or dissimilarity between two objects. In the clustering of flow cytometry histograms, the similarity/dissimilarity was calculated as the Euclidean distance between Mab histograms. This distance measure can be expressed as a “dissimilarity index” with lower distance measures reflecting greater similarity and higher distance measures reflecting greater dissimilarity.

(Figure 4B). To validate the clustering result, antibodies recognizing the same target molecule—independently confirmed by biochemical and genetic studies (15,25)—were examined. Mab recognizing class I, LFA-1 and $\beta 1$ integrin demonstrated high similarity indices (Figure 4B: yellow, green and red). Visual inspection of the corresponding flow cytometry results demonstrated correspondence between the histograms and the dissimilarity indices (Figure 5: yellow, green and red). For comparison, the pair with the highest dissimilarity index (Figure 4B: blue) demonstrated markedly disparate flow cytometry histograms (Figure 5: blue).

To determine the relative advantage of examining multiple cell types, we compared the dissimilarity Indices of the anti-LFA-1, anti-CD29, class II and anti-Class I Mab. For controls, we chose two Mab pairs with comparable surface expression to the anti-LFA-1 Mab on sheep splenocytes. As expected, the hierarchical clustering of individual cell types was less discriminating than the combination of all five cell types: the combined analysis demonstrated high dissimilarity in the control antibody comparisons and low dissimilarity indices in antibodies with shared specificities (Figure 6). Of interest, Mab recognizing related molecular subunits (e.g. CD29 and CD49a) showed marked similarity in some tissues and dissimilarity on others (Figure 6). Mab recognizing polymorphic epitopes on class II showed comparable variability (Figure 6). The intermediate results with antibodies recognizing members of a molecular family—the heterodimeric subunits of the CD29 family and polymorphic epitopes of class II molecules—suggest that the dissimilarity index, when compared across cell types, may also provide a nuanced measure of molecular relatedness.

Discussion

Characterizing and comparing Mab reactivity patterns continues to be important for both basic and applied biologic research. In this report, we used a anti-sheep Mab panel to illustrate the application of statistical and informatics tools in Mab pattern analysis. Our approach has similarities to traditional workshops; that is, Mab were used as specific probes of target molecules, flow cytometry was used as a quantitative measure of molecular expression, and hierarchical clustering was performed with the cytometry results. Because of the absence of cell lines in sheep, however, Mab reactivity was tested against naturally occurring cell types. A disadvantage of using isolated cell populations is that the resulting histograms are frequently multimodal and difficult to model using parametric families—a data structure incompatible with a “percent positive” analysis. Identifying the underlying structure in these histograms required the application of kernel smoothing and an advanced statistical tool called SiZer. Using SiZer to identify significant features of the histograms from 5 different cell types, we performed hierarchical clustering of 128 Mab. The validity of the clustering algorithm was inferred by the reliable identification of externally verified molecular reactivity.

Our analysis of single parameter histograms is different from, but complementary to, current trends in flow cytometry (26). Recent technologic advances have focused on increasing the number of dimensions that can be simultaneously measured by flow cytometers. This multi-dimensional approach takes advantage of the unique ability of flow cytometry to analyze cell populations on a “per cell” basis. Implicit in this approach are the assumptions that 1) the analysis will yield a binary result if sufficient subdivisions are analyzed, and 2) the result will reflect static cell populations with a stable phenotype. An advantage of this approach is that it reduces a multi-dimensional dataset to a string of binary descriptors (e.g. CD45RO⁺CD44⁺L-selectin⁺) (27). A disadvantage is that the analysis is relatively inflexible to independently regulated molecular expression and evolving phenotypes. An additional limitation of this approach is that the discriminating dimensions (parameters) must be “built into” the analysis. As a result, conceptual distinctions may be driven more by methodologic decisions than biologic patterns.

In contrast, our data analysis is based on simple single parameter histograms. Implicit in our approach are the assumptions that 1) combinatorial complexity—multiple cell types—can be used to effectively discriminate molecular expression patterns, and 2) stochastic properties of cell populations may reflect insights that are nonidentical to the detailed phenotypic description of individual cells (28–30). An advantage of our approach is that single parameter analyses are currently more technically reproducible than multi-parameter analyses (31). Similarly, the use of a single fluorochrome preserves the integral scale of the distance measures used for clustering. Third, our approach limits the *a priori* judgements required for the analysis—perhaps enhancing the knowledge discovery process. Finally, it is important to underscore that patterns observed in our analysis of single parameter histograms can and should be explored using contemporary multi-dimensional (multi-parameter) flow cytometry techniques.

A major accomplishment of our work is the development of a computational approach to the reliable identification of “real” peaks. The underlying structure of nonparametric flow cytometry histograms can be masked by both biologic and technical variability. A statistical technique designed to preserve the underlying histogram structure, while limiting channel-to-channel variability, is kernel smoothing (20). In studies of smoothing techniques based on a Gaussian kernel, we found that clustering results varied substantially depending upon the bandwidth or degree of smoothing. Narrow bandwidths risked undersmoothing and the inclusion of unmeaningful peaks in the clustering algorithm. Alternatively, wide bandwidths risked oversmoothing and the exclusion of small, but meaningful, peaks. To address this problem we used the statistical methodology referred to as SiZer (Significant Zero crossings) (23).

SiZer provided three major advantages. First, SiZer produced an assessment of the significance of a histogram feature. A peak was present when there was a zero crossing of the derivative of the smoothed curve (or “density estimate”). The peak was statistically significant when the derivative of the estimate was significantly positive to the left and significantly negative to the right. Second, SiZer enabled data analysis and clustering without the arbitrary selection of a bandwidth parameter. Although an appropriate bandwidth could be reasonably selected by visual inspection when analyzing one histogram, this approach became impractical when clustering dozens of histograms. Using a family of bandwidths avoided the necessity of arbitrary bandwidth selection. Third, SiZer provided an objective tool for differentially weighting prominent peaks. Since the clustering algorithm used data from the entire family of SiZer curves, prominent peaks that were present in all bandwidths were more influential in the hierarchical clustering analysis than peaks that were present in only 1 or 2 bandwidths.

Another important advantage of our approach is that it permits the development of a longitudinal database independent of pairwise comparisons. Because features are identified relative to an absolute calibration standard, pairwise comparisons used in many approaches (32)(33) are unnecessary. As a result, the database is cumulative; that is, new antibody data can be added to the database without the necessity of testing the entire antibody “panel.” A cumulative database is a particular advantage in nonhuman species because international antibody workshops have become increasingly impractical. A challenge to our technique, shared by almost all approaches to Mab pattern recognition, is the requirement for “real” peaks on which to base the clustering analysis. In circumstances in which the histogram is only slightly shifted from the “negative” peak, the reliable identification of a significant peak depends on minimizing technical and biologic variability. Similarly, when target molecules are infrequently expressed in normal tissues, the analysis must systematically include cell populations that express the appropriate target antigens.

A theoretical advantage of clustering data derived from naturally occurring cell populations, rather than immortalized cells lines, is biological relevance. We speculate that similarities and

dissimilarities of cell surface molecule expression identified by hierarchical clustering of naturally occurring cell populations will provide insights into a variety of biological processes. Parallel molecular expression levels may reflect receptor-ligand relationships, common regulation, or membrane linkage. Similarly, a parallel increase in cell surface molecule expression after cytokine exposure may imply a related function or the common participation in a biologic process. Of practical significance, at least for nonhuman experimental systems, will be the comparison of hierarchical clustering results between species. Cross-species comparisons may facilitate Mab characterization. Further, these clustering comparisons may enhance the connection between animal models and human disease.

Acknowledgments

Supported in part by NIH Grant HL47078 and HL75426

The authors would like to thank Drs. Howard Shapiro and Edgar Milford for their helpful comments and insights, and Drs. Matt Wand and J. Stephen Marron for their introduction to SiZer.

Abbreviations

CD	clusters of differentiation
FCS	fetal calf serum
HLDA	Human Leukocyte Differentiation Antigens Workshops
Mab	monoclonal antibodies
SiZer	Significant Zero crossing

References

1. Bernard A, Boumsell L. The clusters of differentiation (CD) defined by the First International Workshop on Human Leucocyte Differentiation Antigens. *Hum Immunol* 1984;11(1):1–10. [PubMed: 6434495]
2. Bernard, A.; Boumsell, L.; Dausset, J.; Milstein, C.; Schlossman, SF. *Leukocyte Typing*. Berlin: Springer-Verlag; 1984.
3. Reinherz, EL.; Haynes, BF.; Nadler, LM.; Bernstein, ID. *Leucocyte Typing II*. Berlin: Springer-Verlag; 1986.
4. McMichael, AJ.; PCLB; Cobbold, S.; Crumpton, MJ.; Gilks, W.; Gotch, FM.; Hogg, N.; Milstein, C.; Spiegelhalter, DJ.; Waldmann, H., editors. *Leucocyte Typing III: White Cell Differentiation Antigens*. Oxford: Oxford University Press; 1987.
5. Knapp, W.; Dorken, B.; Gilks, W.; Rieber, EP.; Schmidt, RE.; Stein, H.; von dem Borne, AEGK., editors. *Leucocyte Typing IV*. Oxford: Oxford University Press; 1989.
6. Schlossman, SF.; Boumsell, L.; Gilks, W.; Harlan, JM.; Kishimoto, T.; Morimoto, C.; Ritz, J.; Shaw, S.; Silverstein, RL.; Springer, T., et al. *Leucocyte Typing V*. Oxford: Oxford University Press; 1995.
7. Kishimoto, T.; Kikutani, H.; von dem Borne, AEGK.; Goyert, SM.; Mason, DY.; Miyasaka, M.; Moretta, L.; Okumura, K.; Shaw, S.; Springer, TA., et al., editors. *Leucocyte Typing VI: White Cell Differentiation Antigens*. New York: Garland Publishing, Inc; 1997.
8. Kohler G, Milstein C. Derivation of specific antibody-producing tissue culture and tumor lines by cell fusion. *Eur J Immunol* 1976;6:511–519. [PubMed: 825377]
9. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 1975;256:495–497. [PubMed: 1172191]
10. Astaldi GC. Use of human endothelial culture supernatant (HECS) as a growth factor for hybridomas. *Methods Enzymol* 1983;92:39–47. [PubMed: 6855620]
11. Li X, Abdi K, Mentzer SJ. o-phthaldehyde fluorescence microassay for the determination of antibody concentration. *J Immunol Methods* 1994;172(2):141–145. [PubMed: 8034969]

12. Li X, Abdi K, Mentzer SJ. Hybridoma screening using an amplified fluorescence microassay to quantify immunoglobulin concentration. *Hybridoma* 1995;14(1):75–78. [PubMed: 7768535]
13. Jiang X, Pratt JP, Huss HT, Ravnic DJ, Mentzer SJ. Structural mapping of immunoglobulin subclass using multiplexed bead flow cytometry. *Hybridoma*. 2006 In press.
14. Mackay CR, Marston WL, Dudler L, Spertini O, Tedder TF, Hein WR. Tissue-specific migration pathways by phenotypically distinct subpopulations of memory T cells. *Eur J Immunol* 1992;22:887–895. [PubMed: 1372559]
15. Su M, He C, West CA, Mentzer SJ. Generation of sheep x (sheep x mouse) heterohybridoma cell line expressing the beta 1 integrin membrane molecule. *Hybridoma* 2000;19:81–88. [PubMed: 10768844]
16. Li X, Abdi K, Rawn J, Mackay CR, Mentzer SJ. LFA-1 and L-selectin regulation of recirculating lymphocyte tethering and rolling on lung microvascular endothelium. *Am J Respir Cell Mol Biol* 1996;14(4):398–406. [PubMed: 8600945]
17. Li X, Abdi K, Mentzer SJ. Cloning hybridomas in a reversible three-dimensional alginate matrix. *Hybridoma* 1992;11:645–652. [PubMed: 1459587]
18. Li X, Abdi K, Herren T, Faller DV, Mentzer SJ. Affinity membrane identification of immunoglobulin subclass in hybridoma screening. *Hybridoma* 1994;13(5):431–435. [PubMed: 7860099]
19. He C, Young AJ, West CA, Su M, Konerding MA, Mentzer SJ. Stimulation of regional lymphatic and blood flow by epicutaneous oxazolone. *J Appl Physiol* 2002;93(3):966–973. [PubMed: 12183492]
20. Wand, MP.; Jones, MC. Kernel Smoothing. CRC Press; 1995.
21. Hardle W, Marron JS, Wand MP. Bandwidth choice for density derivatives. *Journal of the Royal Statistical Society* 1990;52:223–232.
22. Aldershof B, Marron JS, Wand MP. Facts about the gaussian probability density function. *Applicable Analysis* 1996;59:289–306.
23. Chaudhuri P, Marron JS. SiZer for exploration of structures in curves. *J Am Stat Assoc* 1999;94:807–882.
24. Dean PN, Bagwell CB, Lindmo T, Murphy RF, Salzman GC. Introduction to flow cytometry data file standard. *Cytometry* 1990;11(3):321–2. [PubMed: 2340768]
25. Zhao T, He C, Su M, West CA, Swanson SJ, Young AJ, Mentzer SJ. Cell adhesion molecule expression in the sheep thymus. *Dev Comp Immunol* 2001;25:519–530. [PubMed: 11356231]
26. Shapiro, HM. Practical Flow Cytometry. New York: John Wiley & Sons, Inc; 2003. p. 736
27. Sprent J. T and B memory cells. *Cell* 1994;76:315–322. [PubMed: 8293465]
28. Zeng Q, Young AJ, Boxwala A, Rawn J, Long W, Wand M, Salganik M, Milford EL, Mentzer SJ, Greenes RA. Molecular identification using flow cytometry histograms and information theory. *Proc AMIA Symp* 2001:776–80. [PubMed: 11825291]
29. Zeng Q, Wand M, Young AJ, Rawn J, Milford EL, Mentzer SJ, Greenes RA. Matching of flow-cytometry histograms using information theory in feature space. *Proc AMIA Symp* 2002:929–33. [PubMed: 12463961]
30. Kim EY, Zeng Q, Rawn J, Wand M, Young AJ, Milford E, Mentzer SJ, Greenes RA. Using a neural network with flow cytometry histograms to recognize cell surface protein binding patterns. *Proc AMIA Symp* 2002:380–4. [PubMed: 12463851]
31. Zeng QT, Pratt JP, Pak J, Ravnic DJ, Huss H, Mentzer SJ. Single parameter feature-guided clustering for population identification in flow cytometry data. *J Biomed Informatics* 2007;40:325–331.
32. Salganik MP, Milford EL, Hardie DL, Shaw S, Wand MP. Classifying antibodies using flow cytometry data: Class prediction and class discovery. *Biometrical Journal* 2005;47(5):740–754. [PubMed: 16385913]
33. Salganik, MP.; Hardie, DL.; Swart, B.; Dandie, GW.; Zola, H.; Shaw, S.; Shapiro, H.; Tinkam, K.; Milford, EL.; Wand, MP. Elsevier Science Bv; 2005. Detecting antibodies with similar reactivity patterns in the HLDA8 blind panel of flow cytometry data; p. 67-74.

Figure 1.

Segmented density profile of the panel of anti-sheep Mab staining of naturally occurring sheep cells. The flow cytometry histograms were arbitrarily segmented into 3 levels for presentation purposes (A). The panel of 128 anti-sheep Mab were analyzed on 5 cell types: B) thymocytes, C) alveolar macrophages, D) lymph node cells, E) splenocytes and F) efferent lymph cells.

Figure 2.

Quantitative flow cytometry histograms of five cell types in sheep. Particle calibration (A) histograms selected from a similar date in six consecutive years is shown. To illustrate biologic variability, monoclonal antibody staining of different tissues was similarly selected from A) efferent lymph (17D), B) lymph node (ERD2/117), C) alveolar macrophages (M3/246), D) splenocytes (F10–150), and F) thymocytes (T2/39). Each color reflects the same year in the different tissues.

Figure 3.

The single parameter flow cytometry histogram of sheep efferent lymph stained with the monoclonal antibody L2/34. The quantitative histogram (A) was smoothed with a range of bandwidths (B). The green “jitter plot” shows relative data density along the X-axis (arbitrary Y-axis). By including the bandwidth (h) dimension, the smoothed histograms can be plotted in 3-dimensions (C). To facilitate the presentation, the significant zero crossing is color coded (C and D): blue represents a slope significantly upward (derivative CI above 0), red represents a slope significantly downwards (derivative CI below 0) and the purple represents insignificant slope (derivative CI contains 0). Gray denotes regions with insufficient data. The white dotted line (D) shows the effective window widths for each bandwidth.

Figure 4.

The anti-sheep Mab dendrogram. A) By convention, the ordering of the Mab is based on Euclidean distance; a reflection of total fluorescence staining. This is illustrated by anti-class I Mab (T2/39, yellow), anti-CD45 Mab (T2/51, green), anti-CD4 Mab (17D, red) and anti-TCR-gamma/delta Mab (86D, blue). B) The dissimilarity index is illustrated by similar anti-class I (yellow), anti-LFA-1 (green) and anti-Beta1 integrin Mab staining. The extreme of this index is illustrated by two dissimilar Mab (blue). The corresponding color-coded histograms are shown in Figure 5.

Figure 5.

Flow cytometry histograms of similar and dissimilar Mab. The color-coded Mab identified in Figure 3B are shown. The similar anti-class I Mab (T2/39, B1/41 and B1/366; yellow column), anti-LFA-1 Mab (I6, F10-150; green column), and anti-Beta-1 integrin Mab (ERD2/117, FW4-101-1; red column) are shown. The dissimilar Mab (T4/268 and E2/85) are also shown. The rows represent different cell types: A) efferent lymph cells, B) lymph node cells, C) alveolar macrophages, D) splenocytes, and E) thymocytes.

Figure 6.

Effect of combining multiple cell types in the computational algorithm. The dissimilarity index of pairs of Mab pairs were compared when the computational algorithm was applied to single cell types—thymocytes (thymus), splenocytes (spleen), lymph node cells (lymph node), alveolar macrophages (alv macs) or the combined cell types (gray rectangle). For each antibody pair, the dissimilarity index was calculated. The Mab recognizing LFA-1 (I6 and F10-150), CD29 (ERD2/117 and FW4-101-1) and class I (T2/39 and B/41) were specific for the same molecule. The Mab recognizing the alpha (CD49a) and beta (CD29) chains of the beta-1 integrin family (CD49/CD29) showed a higher dissimilarity index reflecting the differential expression of alternative alpha subunits. Two Mab shown by biochemical analysis to recognize distinct polymorphic class II epitopes (M2/36 and M4/68) demonstrated a relatively high dissimilarity index. As a control comparison, two pairs of Mab (Control #1 and Control #2) were chosen because of comparable expression levels on splenocytes to the anti-LFA-1 Mab, but demonstrated a high dissimilarity index when all 5 cell types were combined.