Profiling B Cell Immune Responses by Microengraving

by

Eliseo Papa

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2008

Author .....
Department of Mechanical Engineering
May 9, 2008

Certified by .....
Hidde L. Ploegh
Professor, Whitehead/MIT Biology
Thesis Supervisor

Certified by .....
J. Christopher Love
Assistant Professor, MIT Chemical Engineering
Thesis Supervisor

Certified by .....
Robert Langer
Institute Professor, MIT Mechanical Engineering
Thesis Supervisor

Accepted by .....
Lallit Anand
Chairman, Department Committee on Graduate Students

Profiling B Cell Immune Responses by Microengraving

by

Eliseo Papa

Submitted to the Department of Mechanical Engineering
on May 9, 2008, in partial fulfillment of the
requirements for the degree of
Master in Mechanical Engineering

Abstract

The ability to monitor an immune response in the course of vaccination or disease progression is highly desirable. Currently, no technique is able to generate a comprehensive profile of the individual cells involved and the antibodies they produce at a particular point during the immune response.

The ability to obtain such detailed "snapshots" describing the immune response with a high level of resolution would have implications for diagnostics and biological discovery. Improvement in vaccination schemes, specific tailoring of anti-viral administrations, large-scale monitoring of complex latent infections in a population are all possibilities that would stem from a better understanding of the dynamics of immune responses.

Currently available methods for profiling of B cells that produce antigen-specific antibodies helped clarify humoral responses, but it remains a challenge to generate measurements capable of detailing the phenotypic changes and secretion patterns of individual lymphocytes. To address this need a soft lithographic approach termed microengraving (μEn) - previously used for the isolation and rapid selection of monoclonal antibodies[31] - was further developed and adapted to measure the affinity and isotype of secreted antibodies.

The objective of this thesis was to employ microengraving in conjunction with bioinformatics analysis to obtain routinely state-based comprehensive profiles detailing cellular and humoral immune responses to antigens to the level of clonal B cells.

Here I show how bioinformatics methods were employed to generate multidimensional datasets for large numbers of individual primary B cells (10^2 - 10^4). These data include three characteristics of the antibodies secreted by each cell: antigenic specificity, isotype, and affinity. These data are sufficient to classify individual cells into distinct groups of related cells using algorithms for data clustering. In a series of mice immunizations designed to mimic a multipart vaccination, I apply this method to profile the resulting B cell response with single cell resolution.

Thesis Supervisor: Hidde L. Ploegh

Title: Professor,Whitehead/MIT Biology

Thesis Supervisor: J. Christopher Love

Title: Assistant Professor, MIT Chemical Engineering

Thesis Supervisor: Robert Langer

Title: Institute Professor, MIT Mechanical Engineering

Acknowledgments

Chapters such as these, especially when improvised while nearing a deadline, have the tendency of being excessively circumstantial. Rarely I have been able to express my gratitude in words that were effective in relying what I sincerely felt. I apologize if I am not able to do any better in the words that follow.

I am indebted to many people who have made this work possible and, perhaps more importantly, to those who are responsible for the many opportunities I encountered along my path. It is my hope that one day I'll somehow be able to give back what has been an immense and perhaps unfair privilege.

I want to thank Hidde, one of my supervisors, for the unconditional and enthusiastic support he has given me and any path I am willing to take. The unconditional support, the time and enthusiasm I have had from Chris, my other supervisor, remains for me unbelievable. It has been my outmost privilege to work with him.

Many thanks also go to the members of the Ploegh lab, who have been incredibly helpful and patient in teaching me biology and immunology, while helping my thesis along. My gratitude in particular goes to my collaborators in this work – Craig Story, Chih-chi Andrew Hu and Jehnna Ronan – without whom this work wouldn't have happened.

On a more personal note, I am immensely thankful to all those people who have shaped and enriched my life. My brother, my great friends and most importantly Alessandro, who lies in between and above these two definitions.

Lastly, my immense love, my outmost gratitude and my undying admiration are for my parents. *Papá, c'é un'immensa parte di te in quel nome, Eliseo, che sono riuscito a mettere davanti al nome di una grande scuola e come sottotitolo di una vita felice, soltanto per merito tuo e della Mamma. Mamma... amore puro, bontá infinita... Io, cosi quanto Papá, non riesco a capacitarmi della fortuna che ho ricevuto. Il tuo sorriso carico d'amore é un tesoro che porteró sempre con me.*

Family, vi devo tutto. Spero di rendervi sempre piú orgogliosi e felici.

Contents

List of Figures

List of Tables

Chapter 1

Introduction: Vaccinations & Immunological Memory

1.1 Vaccination

Infectious diseases and vaccination Infectious diseases are the leading cause of death in humans worldwide[1]. Only the advent of vaccination and the improvement of sanitation measures has allowed us to combat effectively the consequences of infection. Since Jenner and Pasteur established these practices, vaccination has proven to be the single most effective public health measure to prevent mortality and morbidity from infections [52]. It is in fact those infectious diseases for which an effective vaccine is still lacking, such as malaria and HIV/AIDS, that represent the most severe global health issues and where much of the current research is focused. Contributing to the success of vaccination worldwide are the relatively low cost, the ease of distribution and administration and the fact that one administration suffices to give long-lasting effects, without requiring impractical and tedious maintenance regimens[52, 44, 49].

Basic principle Traditionally, vaccines rely on the exposure of the host to a weakened, less pathogenic form of the antigen. The consequent primary immune response of the host immune system to the antigen clears the infection and confers long lasting protection against susbsequent exposures to the same antigen. This is termed

immunological memory.

1.2 Immunological Memory

Immunological memory is the ability of the immune system to respond more rapidly and intensely to exposure to a previously encountered antigen. It reflects the existence of long-lived antigen-specific lymphocytes, which have been clonally expanded in the course of a primary response. This can be directly shown by adoptive transfer of memory cells into a naïve animal, which is sufficient to transfer immunological memory to the recipient[22].

Memory cells are generally regarded as a resting population, although it is possible to observe a small percentage of these cells undergoing division at any one time[36]. IL-7 and IL-15 are amongst the cytokines required for the maintenance of these populations [4, 53]. Fundamental to the generation of memory is the involvement of T cells in the primary immune response. Functional antigen specific B cells secreting antibodies persist only when accompanied by the primary T helper cell reaction to the antigen[22]. Once the primary phase of the immune response subsides, the frequency of B cells diminishes but remains higher than before immunization. These frequencies of memory B cells are then maintained for the life of an individual, independently of persisting immunizing agent [35] and are generally reflected in increased antibody titres against the immunogen[22].

Importantly, only those proteins that elicit the involvement of T cells are capable of generating significant immunological memory[37]. Memory responses are termed secondary, tertiary, and so on, depending on the number of repeated exposures to antigen. Repetitive challenge with an immunogen is called hyperimmunization. Subsequent responses can be distinguished from the primary response in terms of the frequency and type of antibodies produced, as well as differences in the cell surface markers of B and T cells involved (for more details see 4.2).

1.3 Following the course of an immunization

Monitoring the course of immunizations requires in principle the tracking of all participating mechanisms of immunity. However, humoral immunity and cellular immunity have been difficult to monitor simultaneously by means of a single technique. Independent techniques to measure antibodies titres, specificity, isotype and affinity are routinely used to monitor humoral immunity and are generally used to evaluate the level of protective memory induced by previous immunizations. Cellular immunity can instead be observed by enumerating lymphocytes, whether circulating or resident in specialized organs like the spleen or the lymph nodes.

1.3.1 Antibodies

For vaccines currently in use the titers of circulating antibodies are usually taken as a correlate of their level of protection. These titers may be assessed by their neutralizing capacity, microbicidal activity or simply antigen binding of serum antibodies.

Specific antibodies can be isolated from antiserum by affinity chromatography, while direct detection and quantification of antigen-specific antibodies are routinely performed by enzyme-linked immunosorbent assay (ELISA). Briefly, the antigen of interest is immobilized on a solid support: beads in affinity chromatography and the surface of adsorbing membranes for ELISA. While the antiserum is in contact with the solid support, the specific antibodies present in the antiserum bind to the immobilized antigen. In affinity chromatography, the unbound fraction is washed away, while the bound fraction is subsequently eluted from the column and collected. In ELISA, the unbound, non-specific antibodies are washed away in a similar fashion, but the amount of antibody captured is quantified by means of binding by a secondary antibody linked to an enzyme. The enzymatic reaction catalyzed by the secondary antibody provides a direct quantitative measure of the number of antibodies bound to antigen compared to control.

These technologies, however, are inherently limited in their throughput and the amount of information that they can provide. For instance, little is known about the

diversity of the antibodies elicited, except in those few cases where relatively large collections of monoclonal antibodies, produced by hybridomas generated from the immunized subjects, are available for analysis. The extent to which this approach biases the estimates of diversity of the B cell repertoire is not known; some B cells may be more readily immortalized as hybridomas than others. Neither does this approach reveal the emergence of antibody-based (humoral) immunity over time.

1.3.2 B Cell Frequency

Several methods exist to estimate the frequency of naïve and memory B cells that can bind specific haptens or antigens. These include the Jerne plaque assay[23], the splenic focus assay[20], a variant of the plaque assay that immobilizes colonies of cells on discs of filter paper [26], the enzyme-linked immunospot (ELISPOT) assay [13, 14], and fluorescence-activated cell sorting (FACS)[36].

The Jerne plaque assay or hemolytic plaque assay consists of an hemolytic reaction performed on living splenocytes immobilized in gel. It permits detection and enumeration of the B cell colonies that secrete antibodies. In the splenic focus assay, splenocytes are cultured in vitro and B cells secreting specific antibodies are stimulated to proliferate by addition of antigen. Although it is not possible to isolate monoclonal antibodies from single cells, populations of closely-related antibodies can be isolated and characterized. The filter paper technique improves upon this technique, allowing one to separate single lymphocyte precursors and stimulate their proliferation in order to obtain an array of monoclonal colonies ready to be characterized. Limiting dilution, ELISPOT, FACS all allow the enumeration and isolation of single antigen specific lymphocytes, which can then be further cultured and characterized.

However, none of these methods give information on the characteristics of the antibodies secreted by the lymphocytes. To date, no single method produces a comprehensive and correlated cellular profile that includes, among other traits, the frequency of antigen-specific B cells present, and multiple characteristics of the antibodies they produce (specificity, isotype, affinity) with single-cell resolution.

1.4 Cellular Immunity and Current Vaccines

Even though immunological memory involves both long-lasting cellular and humoral immunity, current vaccines are thought to be successful mostly due to their ability to elicit a long-lived humoral response [49].

Failure of classic vaccination However, classical vaccination strategies that rely on eliciting an humoral immune response have increasingly proven insufficient to deal with emerging infectious disease, where cellular immunity may play a greater role in host defense. HIV/AIDS represent a striking example of how classic approaches to vaccine discovery have repeatedly failed. Evolving viral evasion strategies, virus latency and the integration of viral genetic material into the host genome are some of the reasons why eliciting humoral immunity is not sufficient to clear or manage complex infections.

To address this issue, much of the attention in vaccine research is now focused on eliciting and maintaining specific cellular immunity as well as artificially stimulating native immunity. However, even though research has increasingly focused on manipulating such cellular immune responses, there remains a shortage of tools to monitor these responses in detail and at high resolution.

A single technique that could quantitatively describe both humoral and cellular components of the immune response at the single cell level would substantially help ongoing research efforts.

Chapter 2

High-throughput measurements of antibodies affinity

The process underlying the acquisition of immunity after vaccination is very complex and still remains to be understood in fine detail. The collection of cells of various types, specificities and activities that constitutes the immune system represents a fine and intricate network of interactions, regulations and balances, which remain difficult to explore using current methods. Arguably, the only way these interactions can be explored systematically is by large scale observations of the behaviour of the individual players of immunity. Our objective was to enable such observations in a quantitative fashion. In particular, we wanted to develop a method for the high-throughput measurement of the amount and quality of secreted antibodies from a population of lymphocytes. This in turn would allow to a) rapidly identify rare antibodies of particular specificity/affinity and b) obtain a comprehensive picture of an in vivo antibody repertoire. In order to address this need and enable the measurement of quantitative parameters in a large collection of individual immune cells, we decided to employ microfabrication techniques. The ability of manipulating single cells of the immune system afforded by the fabrication of biocompatible microchips, as well as the relative ease of use and scalability, has allowed high-throughput sampling of secreted products and surface markers with unprecedented resolution.

2.1 Soft Lithography

Microfabrication by soft lithography was selected for the purpose of this work. Soft lithography is a set of techniques for microfabrication based on the use of biocompatible elastomeric materials[55]. Soft lithography relies on printing and molding using elastomeric stamps with the patterns of interest in bas-relief. Soft lithography offers the ability to control the molecular structure of surfaces and to pattern the complex molecules relevant to biology, to fabricate channel structures appropriate for microfluidics, and to pattern and manipulate cells. For the relatively large feature sizes used in biology ($> 50\mu m$), production of prototype patterns and structures is convenient, inexpensive, and rapid.

Generally, microfluidic systems are fabricated with an elastomeric material, poly (dimethylsiloxane) or PDMS, starting from a micro-patterned silicon wafer employed as a mold. Briefly, features of the chip (eg. a network of microfluidic channels) are designed in a CAD program. This design is converted into a transparency by a high-resolution printer; this transparency is used as a mask in photolithography to create a master in positive relief photoresist. PDMS cast against the resulting mold yields a polymeric replica containing all the designed features in bas-relief. The chips are then oxidized in an oxygen plasma, in order to allow the surface of the chip to seal tightly and irreversibly when brought into conformal contact. For instance, oxidized PDMS seals irreversibly to other materials used in microfluidic systems, such as glass or silicon. Oxidation of the PDMS has the additional advantage that it yields channels whose walls are negatively charged when in contact with neutral and basic aqueous solutions, thereby rendering the chip hydrophilic and thus biocompatible.

2.2 Microengraving

Microengraving (μEn) is a soft lithographic technique previously developed in our laboratory [31] for the rapid selection of monoclonal antibodies from large hybrydoma populations. It is based on a microfabricated PDMS chip, consisting of an array of

approximately 100000 separated wells, each 50 $\mu m \times 50\mu m \times 50\mu m$ in size.

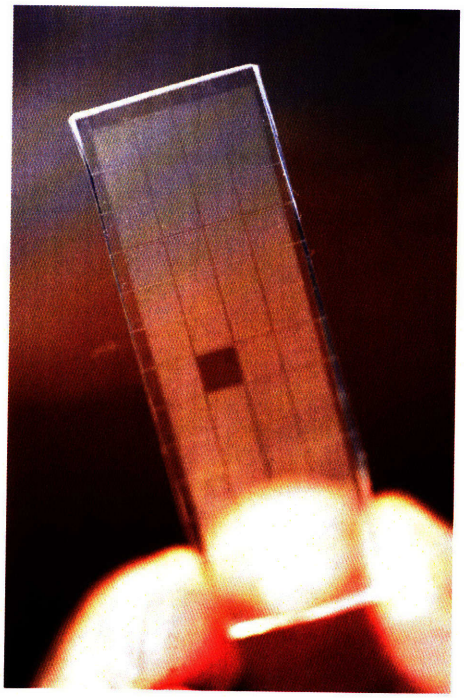

Figure 2-1: Microwell chip employed for microengraving. Routinely fabricated in PDMS by soft lithography, it consists of an array of 50 μm wells

The size of the wells is chosen such that when cells are seeded in the device and settle in each well by means of gravity, each of the wells contains on average[1] a single cell. Non-specific interactions maintain the cells on the bottom of the wells and ease the manipulation of the chip loaded with cells. The PDMS chip is then inverted and placed in firm contact on a glass slide, whose surface is in turn functionalized with a capture antibody. Each individual cell's secreted products are then captured on the glass slide by the deposited antibody (or other appropriate moiety), forming a spatially defined spot. The resulting array of spots, each containing the secreted products of an individual cell, effectively makes up a protein microarray. Once the chip is separated by the produced microarray, the process can be repeated multiple

[1] "Average" here is a misnomer. The process of cells filling the well can be described in terms of an occupancy problem, a subject of interest in random combinatorics. Depending on the starting assumptions (eg. uniform cell distribution, low reynolds number, etc.), the distribution can take many different forms. If the number of cells reaching the chip can be described by a poisson distribution, then the well occupancies become independent Bernoulli random variables and the overall probability of all wells being occupied by n cells at time t is a binomial random variable with parameters n and $\exp -t$.

times (the current protocol allows up to 6-7 replicates without significant loss of cells and engraving quality). Each microarray can then be stained with fluorescent moieties of interest and the resulting fluorescent microarray can then be analyzed by means of the same instrumentation routinely employed in the analysis of DNA and protein microarrays. Cells contained in the chip can be easily stained in parallel for relevant surface markers and the resulting data can be rapidly collected by high-throughput microscopy. Moreover, cells in the chip remain viable and can be retrieved by custom automated instruments, allowing further expansion and molecular analysis.

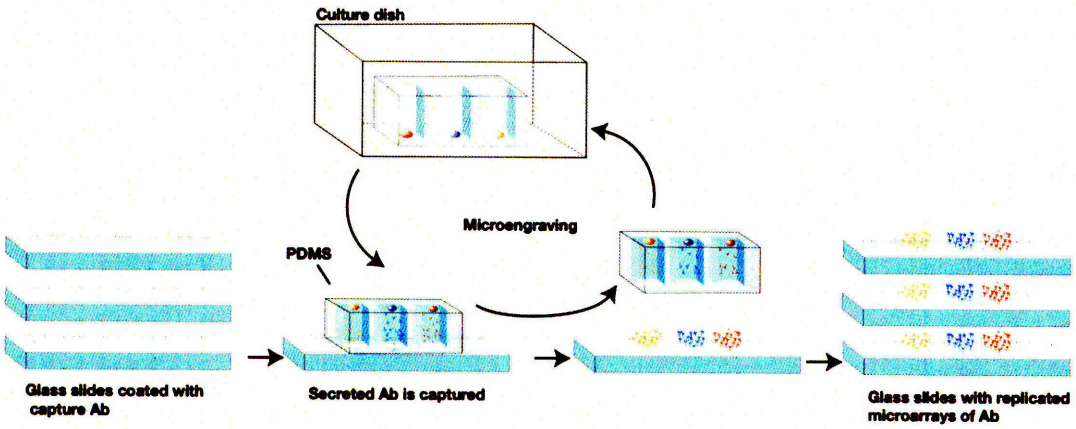

Figure 2-2: Scheme for generating multiple replicates of engraved microarrays of antibodies from a population of cells. After producing microarrays, microwells containing cells are transferred to media for subsequent enumeration.

2.3 Affinity Measurements

To test the feasibility of the microengraving process for the high-throughput measurements of antibodies affinity, we employed a monoclonal hybridoma line specific for a known and available antigen. A key feature of μEn is its ability to produce replicate microarrays of antibodies from a given set of cells with excellent reproducibility. We measured the affinity of the secreted monoclonal hybridoma by producing microarray replicates from the secreted antibodies and staining the microarrays with increasing concentration of the antigen. In principle, by increasing the concentrations of antigen

applied to these replicates, one can define a binding curve (Ag/Ab fractional occupancy vs. concentration) for individual elements in the array, and, in turn, derive the apparent dissociation constants K_d for the antibodies produced by individual cells in the array[47]. The precision and consistency of the affinity measurements is then determined simply by comparing a large number of individual measurements from the same experiment.

2.3.1 Methods

The microengraving system was employed as described above. Super-Epoxy 2 (ArrayIt SME2) functionalized glass slides were pre-coated with a 1:1 mix of anti-mouse IgG capture antibody (Southern Biotech 1030-01) and anti IgG (H+L) (Zymed 1010-01), and then blocked with BSA. These capture antibodies were deposited using coating buffer (50 mM Borate pH 9, 80 mM Trehalose, 50 mM NaCl).

Approximately 100,000 hybridoma cells specific for ovalbumin (OVA) were then seeded into one chip. The hybridoma cell line 099-01, secreting anti-OVA was obtained from Statens Serum Institut and maintained in DMEM (Gibco) supplemented with 10% FBS, 50 units penicillin/50 g streptomycin, 20 mM HEPES, 50 M 2-mercaptoethanol, 1 mM sodium pyruvate and 0.1 mM nonessential amino acids (Gibco). Cells were maintained at 5% CO2, at 37 C, and split every 2-3 days.

The device was then inverted onto the functionalized glass slide and secreted antibodies were captured onto the slide during a 30-minute incubation. The process was repeated 5 times, to obtain 5 replica prints of the same microarray.

Each resulting microarray was exposed to a different concentration of fluorescent OVA. Because several factors —including variation in the rates of secretion, and the efficiency of capture and retention of the antibodies for each replicate print— could affect the quantity of antigen bound at a given site, we also measured the amount of antibody present at a given element in the microarray using a fixed concentration of anti-mouse Ig secondary antibody. As shown in Figure 2-3, non-specific spots (where the secreted antibodies do not bind OVA) are labelled exclusively by the antibody,

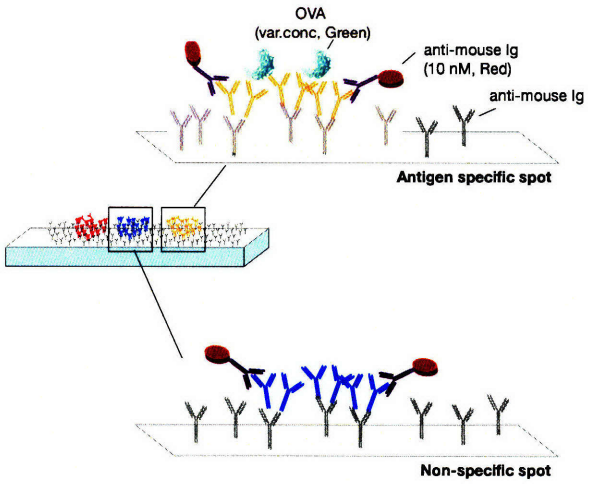

Figure 2-3: Labeling strategy for assessing fractional occupancy of a captured monoclonal antibody by antigen at a given element in the microarray. Replicate prints are exposed to antigen (green) over a range of concentrations (e.g., 10 pM to 100 nM).

while antigen-specific spot show both fluorescent dyes colocalized.

The slides, once probed with fluorescent dye reagents, were scanned by Genepix 4000B and 4200AL microarray scanners, employing the accompanying software Genepix Pro (Molecular Devices, Sunnyvale, CA) in order to extract features and fluorescent data. Raw numerical data was then exported to MATLAB (Mathworks, Natick,MA) for filtering and subsequent statistical analysis.

2.4 Results

2.4.1 Microengraving can be employed to measure fractional occupancy

Assessing the affinity of an antibody immobilized on a microarray requires measuring the fractional occupancy of the antibody at several different antigen concentrations (as in Figure 2-4). In this experiment we employed concentrations ranging from 10 pM to 0.1 μM.

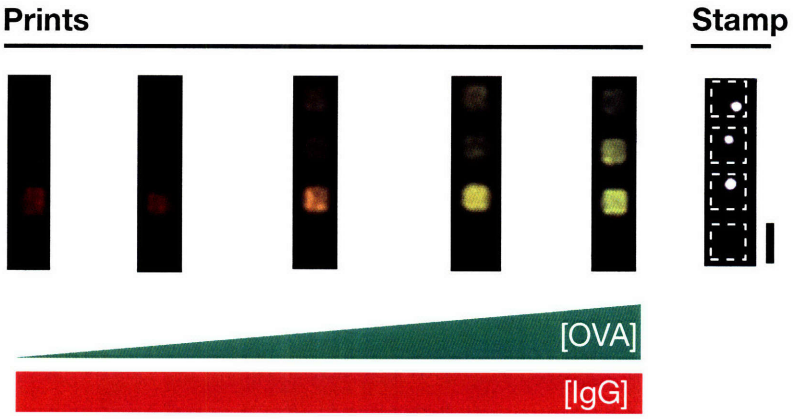

Figure 2-4: Representative composite fluorescent images from five antigen-labeled microarrays produced by microengraving, and the anti-ovalbumin-specific hybridomas (stained with CFSE) corresponding to each element in the arrays. Scale bar is 50 m.

The ratio of the median fluorescent intensities of antigen to antibody gave a measure of the fractional occupancy of the element in the microarray at the applied concentration of ovalbumin.

2.4.2 Occupancy measurements defined affinity curves with high accuracy

When the concentration of antibody is kept fixed, while the concentration of ovalbumin is increased across a number of replicate microarrays, an affinity curve can be constructed for each element of the microarray. The resulting data is presented in Figure 2-5, showing a random sample of approximately 50 microarray spots, each corresponding to an individual hybridoma clone.

The measured fractional occupancy curves are a direct estimate of the binding curve of the immobilized antibody. A random sample of 50 clones shows how the experimental measurements are centered around a median curve, with relatively lit-

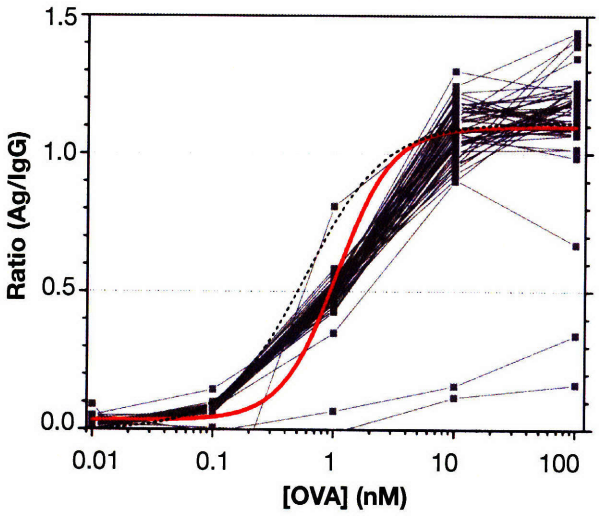

Figure 2-5: Plot of the fractional occupancies measured for 50 cells (grey lines, selected randomly from the complete dataset of 4,426 single cells) as a function of the concentration of ligand applied to five replicate microarrays. The average binding curve fit to these clones (red) and the curve measured by ELISA for the purified antibody (dashed black line) are shown.

tle variation, demonstrating the consistency of the affinity measurements across the array. Importantly, no normalization was necessary here to compare these individual independent elements. It is important to notice how the estimate of affinity here is not biased by effect of bivalent binding as it could be with surface immobilized ELISA, since the antibody fraction here is bound, while the antigen is in solution. This leads to more precise estimate of binding pocket affinity, rather than effective avidity of the antibody.

2.4.3 Conclusions

Microengraving, with the ability of analyzing secreted products of individual cells in a highly parallel fashion, allows the high-throughput quantitation of secretion levels and characterization of secreted products. By exploiting the ability to replicate the protein microarrays with high fidelity, we have shown how precise measurements of antibodies affinity can be routinely performed over a large number of monoclonal cells. This can in turn allows the screening of desirable high-binding clones in a population of antibody-producing cells or, as further explored in this thesis, the compilation of affinity "maps" detailing the repertoire responsible for the humoral immune responses.

Chapter 3

Resolving clonal B cell populations by measuring affinity of secreted antibodies

The consistency of the binding curves produced by each OVA-specific hybridoma in the experiment described above, suggested the potential of identifying individual clones within a mixed population of cells by their relative affinities.

To test the feasibility of this approach we attempted to distinguish a limited and known number of hybridoma lines secreting antibodies for a common antigen. By measuring affinity curves in a high-throughput fashion, we aimed to determine whether the measurements were sensitive enough to allow discrimination between subtly different antibody producing cells.

3.1 Experimental methods

We repeated the experiment described in 2.3 with a mixed population of three different hybridomas. We choose three hybridoma cell lines specific for the same antigen (H2Kb), albeit with varying affinity.

The Y3 cell line, secreting anti-H2Kb antibody was a kind gift of Dr. P Cresswell. Anti-Kb hybridoma lines (JCL127 and JCL136) were described recently. All hybrido-

mas were maintained in DMEM (Gibco) supplemented with 10% FBS, 50 units penicillin/50 μg streptomycin, 20 mM HEPES, 50 μM 2-mercaptoethanol, 1 mM sodium pyruvate and 0.1 mM nonessential amino acids (Gibco). Cells were maintained at 5% CO2, at 37 C, and split every 2-3 days.

The three populations were differentially labelled with cytosolic dyes (c136 with CellTracker Blue CMAC, c127 with CellTracker Red CMTPX, Y3 with CFSE), so that information on their original identity was retained, and then deposited into the microwell chip (as above, 100000 wells each 50 x 50 x 50 μm^3).

Seven replicate microarrays were generated by μEn, and probed with anti-mouse IgG (red) and tetrameric Kb-streptavidin (green). Each hybridoma was labeled with a unique cytosolic dye prior to their deposition in microwells, and after the final print, all of the cells were stained in situ for DNA.

The intracellular stains did not interfere with antibody secretion or cellular viability, and allowed us to correlate the identity of each cell with their measured binding curves.

The experiment yielded a model set of data comprising some 3700 cells. A representative random sampling of these curves is shown in Figure 3-1.

Even when looking at a single block in the microarray, it is already possible to observe how three very distinguishable curves are observable. This visual classification is rapidly verified by coloring the curves based on the information extracted from the cytosolic dyes.

The quantification of the antigen/antibody ratio was here done manually for every clone, an approach that quickly limited the throughput of our analysis. To increase the yield of each experiment and to allow routine quantitative observations, we expanded and automated the data mining process.

3.2 Data Mining

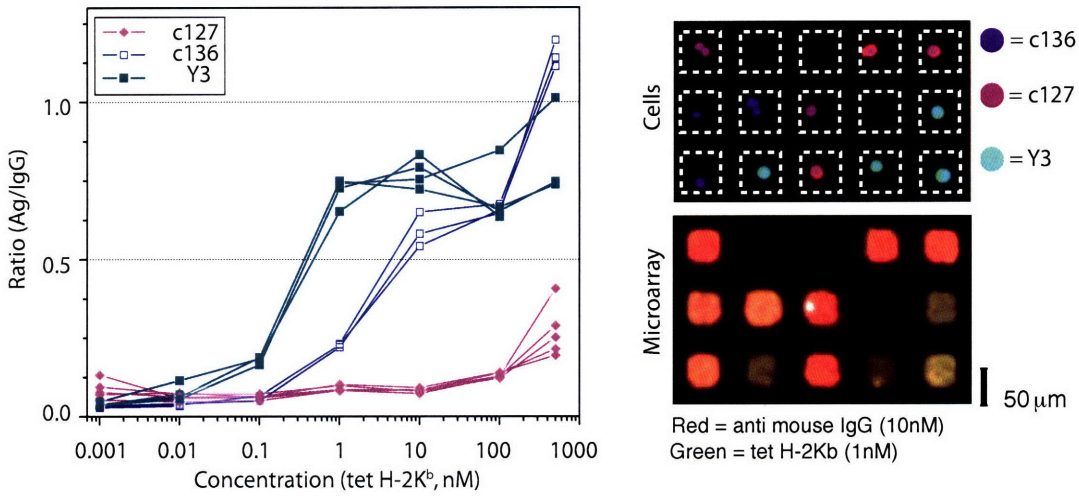

Figure 3-1: A plot showing the fractional occupancy, as a function of the concentration of tetrameric H-2Kb, measured for a set of cells. The colors of the traces correspond with the verified identity of the clone (Y3, aqua; c127, magenta; c136, purple). Representative images of one microarray and the corresponding composite fluorescent image of the cells are also shown.

Imaging Arrays of microwells were placed face down with a thin layer of PBS onto the surface of a glass slide, and imaged on an inverted epifluorescence microscope (Nikon Eclipse TE2000-E) equipped with a Hamamatsu Orca AG camera and an automated x-y translation stage (Prior Instruments). Overlapping images were acquired for the entire array using the ScanSlide module in the Metamorph software package (v7.1, MDS). Images of complete blocks of microwells (40 x 40 wells x 52 blocks per device) were assembled using the Scan Slide module in Metamorph for each data channel acquired, as exemplified by 3-2.

Grid alignment & data extraction Fluorescence intensities, signal-to-noise ratio, and other various parameters were extracted from the microarray scans by means of the commercial software Genepix. Due to the slight deformability of PDMS and the poor tolerance built into this software, grids had to be aligned manually. We expect this process to be exponentially reduced by the development of custom software (see below)

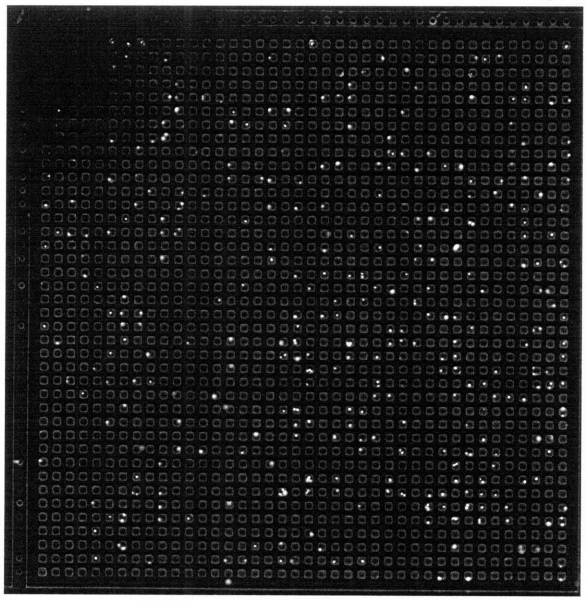

Figure 3-2: Representative fluorescent image of a 40 × 40 microwell block

In order to extract the required data from each of the acquired blocks and correlate it with data from the microarray scans, it was necessary to preserve the given spatial information. Grid alignment was performed semi-automatically in metamorph to produce working images such as the one depicted in Figure 3-3

Once the grids were aligned, semi-automated tabulation of the number of cells found in each well, and their expressed surface markers, was also generated using custom macros written in Metamorph (available upon request).

As an alternative to this tedious process, automated algorithms are under active development. In collaboration with Bjorn Nilson (Broad Institute), we are developing custom code able to perform completely automated alignment. The sophisticated algorithm can be described briefly:

1. binarize the image
2. projection of the information-containing pixels of the transmission image onto the cartesian axis of the image.
3. rotation of the image consistent with the distribution of the projected intensities
4. elastical stretching of a pre-defined grid to match the distribution of the pixel axis

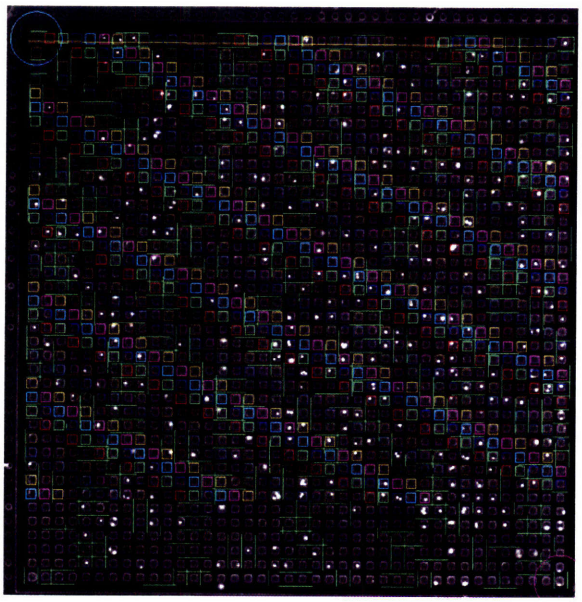

Figure 3-3: Grid alignment in Metamorph software

5. transfer of the fitted grid to all previously segmented fluorescent images, associated with the same element of the microarray
6. quantitation of parameters

Data Correlation Filtered information from a) affinity slides and b) cellular stains (imaged by epifluorescence microscopy) was correlated on the basis of relative spatial coordinates. Raw data collected by a) Metamorph analysis of the PDMS chip images and b) microarray scans was imported into MATLAB (Mathworks, NA) for further processing. The data is represented as closed `struct` data structures in matlab, with shared array coordinate systems. Initial selection of the features and successive aggregation of the dataset in a more manageable form (ie. multidimensional array) was done by means of routine data manipulation scripts. Selected parameters for further analysis were a) 568/635 fluorescent ratios for antibody microarray data and b) boolean variable classifying each cell in the chip as positive or negative for any given marker.

Filtering, however, had to be performed to restrict the size of the dataset prior to these processing steps.

Filtering

Wells that contained more or fewer than 1 cell - as determined by counting the number of nuclei in the DAPI fluorescent channel - were excluded from subsequent analysis. Similarly, cells that failed to show any staining in any of the three channels corresponding to cytosolic dyes were also excluded.

Scanned spots with poor signal quality (high spot covariance, low signal to noise ratio and/or high saturation levels) were filtered from the dataset. Good quality spots were then ranked as Ig+, by setting a threshold on the background corrected median intensity, total intensity and signal to noise ratio in the corresponding channel. A detailed MATLAB script, outlining the algorithm for flagging good quality spots is contained in the Appendix (see A).

3.3 Data Clustering

The filtered dataset which is mined from a single microengraving experiment is extensive and multidimensional in nature. Interestingly, it resembles in its characteristic the data routinely collected by oligonucleotide microarrays: gene expression signatures are replaced here by affinity curves corresponding of the secreted antibody. Likewise, approaches employed in gene expression arrays to identify relevant patterns in the dataset are also similar. In particular, this thesis is concerned with the use of unsupervised clustering methods to the detection and monitoring of biologically significant population in an unknown sample.

3.3.1 Clustering Methods

Current clustering algorithms can be subdivided into two major categories: a) supervised methods, which rely on some *a priori* knowledge of the expected pattern and b) unsupervised methods, which attempt to identify significant clusters without prior knowledge regarding the expected data [7].

Although unsupervised methods such as nearest-neighbour analysis, neural net-

works or support vector machines could undoubtedly be applied to the data collected by microengraving, they require a larger number of experiments to be performed before any substantial knowledge regarding the characteristic of the dataset is obtained. The proof-of-principle analysis presented in this thesis is therefore based on unsupervised methods, such as nearest-neighbor clustering, self-organizing maps, hierarchical clustering and k-means clustering. These clustering methods build on an array of dissimilarity measures to create groups of features with similar patterns[7].

Principle Each pattern, be it a gene expression pattern or an affinity pattern, is described by a point in a multidimensional space, where each dimension corresponds to an experiment or, in this case, to a different antigen concentration. The coordinates of the point are determined by the relative measurements of the same well at the various antigen concentrations. The similarity between patterns is judged quantitatively by a *distance* measurement over this multidimensional space. Various alternative distance metrics exist and is good norm to use multiple ones while clustering the same dataset, to pick out potential biases due to the metric employed[40].

Hierarchical Clustering Closely related patterns are joined together to form a new combined leaf by a *linkage* method. Many choices exist here as well and should be routinely explored for any given experiment.

In hierarchical clustering, the linking process is repeated iteratively until all nodes are connected, resulting in a hierarchical relationship structure. Dendrograms are generally used to represent the resultant hierarchical tree, showing branches between interrelated patterns and thus visually highlighting dominant clusters. The quality of the description is usually judged by calculating the *cophenetic distances* and the corresponding cophenetic correlation coefficient. The cophenetic correlation coefficient is a measure of how faithfully a dendrogram preserves the pairwise distances between the original unmodeled data points. Higher cophenetic coefficients correspond to better quality descriptions and are used to select the best *distance/linkage* combination for the given dataset.

K-means clustering K-means clustering falls into the category of unsupervised *partitioning* methods, where the number of cluster is determined a priori. The dataset is split consequently, in order to maximize the distinction between the given number of clusters. It has the disadvantage of requiring some knowledge of the dataset, although in general K-means clustering is repeated varying the number of desired clusters and comparing the clustering results, in order to identify the cluster number that has produced the highest quality separation (and thus more likely resembles the number of clusters effectively present in the given dataset)

Principal Component Analysys The quality of the linkage method and the feature of the dataset can be visually explored by means of principal component analysis (PCA). This mathematical technique allows one to find the linear combination of the multidimensional dimensions that would explain most of the variation in the dataset. Effectively, this allows a remapping of the multidimensional space in which every affinity pattern lies in a more treatable two- or three-dimensional space. By making possible the actual visualization of the dataset, PCA allows to identify relevant features in the data and, potentially, detect significant clusters from the measurements taken.

3.4 Results

Microengraving analysis of the antibodies secreted by a population of three mixed hybridoma populations lead to a total of 5,000 individual cells with a complete dataset associated with it. Attention was restricted to the best quality dataset, consisting of approximately 3,600 single cells.

3.4.1 Clustering of the *mixed hybridoma* dataset

Unsupervised clustering, either by hierarchical or partitioning methods, was performed in MATLAB to resolve distinct biologically meaningful classes of clones. Cophenet distances and coefficients were computed to assess the quality of differ-

ent hierarchical methods. Only hierarchical clustering methods yielding an average cophenetic correlation coefficient > 90% were employed. Silhouette and gap statistics were similarly employed to identify the adequate number of clusters in partitioning methods (ie. k-means). Principal component analysis (PCA) was routinely used to monitor the distribution of the selected dataset and to examine the quality of the clustering results.

3.4.2 Heatmap visualization of affinity measurements

To evaluate the results of clustering algorithms and allow rapid visual inspection of the dataset, it was necessary to develop a novel representation system. The ensemble of affinity measurements can be represented as a series of independent curves, but it rapidly becomes confusing to distinguish individual curves when their number is greater than 50. An alternate way of displaying these curves is to center each curve such that the median value is set at zero, and map the data as a density plot where points above and below the median are shown in contrasting colors, as shown in Figure 3-4.

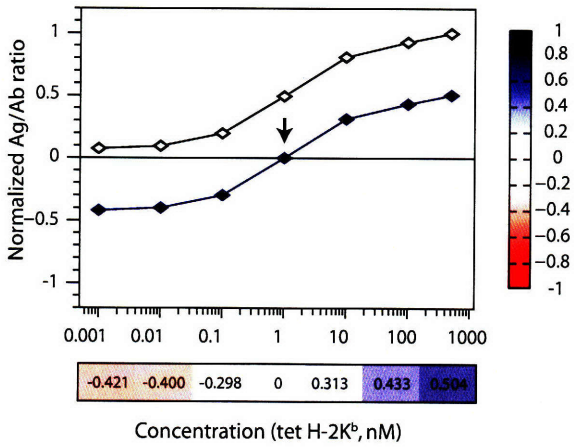

Figure 3-4: Effect of the median centering of each affinity curve on the color mapping scheme employed. High affinity binders show brighter red and blue bands. Non saturated antibodies ($K_D > 100$nM) show only blue coloring. Low affinity antibodies are deprived of any coloring.

Depicted in this manner, visual inspection permits an estimate of two characteristics of the antibodies: the white region indicates the median value, and thus it approximates Kd within an order of magnitude, while the intensities and progression of the colored regions indicates the strength of the binding interaction and suggests whether the antibody is fully saturated or not in the range of concentrations analyzed.

3.4.3 Affinity curves were sufficient to distinguish clonality of antibodies producing cells

Unsupervised clustering of the individual binding curves allowed a clear distinction of the different clones within the dataset, as shown here in Figure 3-5. Both unsupervised, hierarchical and partitional (k-means) clustering algorithms showed a high degree of accuracy when sorting cells into groups of identical clones (see Table 3.1).

We used the identities of the cells determined by immunofluorescence as a reference to assess the accuracy of the assignments made by clustering. Principal component analysis (PCA) of the dataset also yielded three dominant clusters, evident by vi-

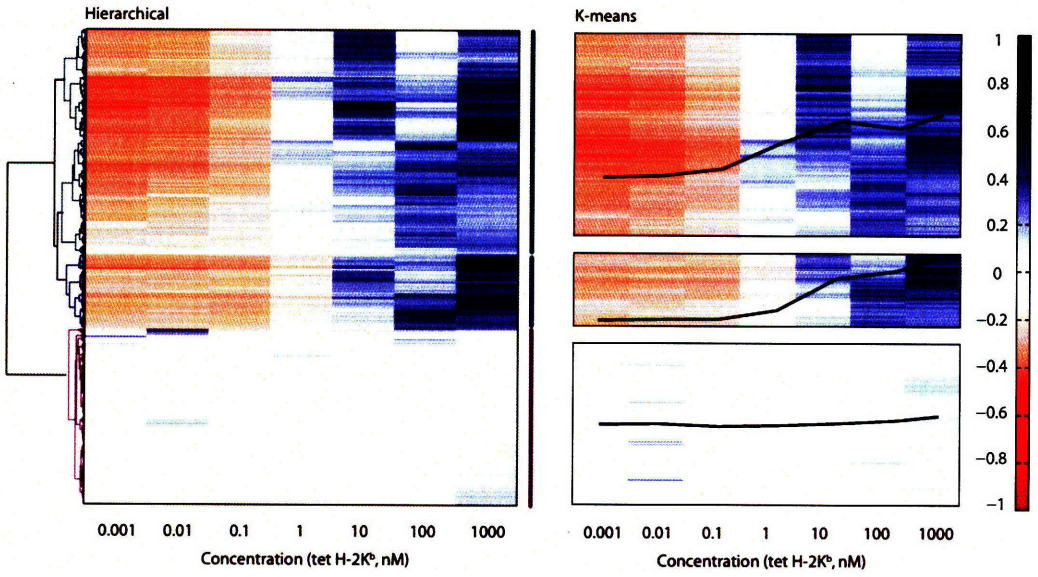

Figure 3-5: Hierarchical (city/ward) and k-means (n=3) cluster analysis of the binding profiles measured for the three clones (3,711 cells). Median profiles for each group in the k-means clusters are superimposed.

sual inspection of Figure 3-6. All three analyses resolved the difference in measured avidities between two clones, Y3 (0.60 nM) and c136 (3.77 nM).

3.4.4 Clustering based on affinity-resolved populations with high sensitivity and specificity

We estimated the sensitivity of the affinity clustering method by comparing it with the data obtained by immunofluorescence of the same cells (and thus the known cell identity). For instance, true positives were those cells that were classified by unsupervised method as being of the cell type previously determined by cytosolic stain. The average sensitivity (ratio of true positives to combined true positives and false) was greater than 87% for both methods, and the average specificity (ratio of true negatives to combined true negative and false positive) was greater than 94% as detailed in Table 3.1, together with estimates of the positive and negative predictive value of the test.

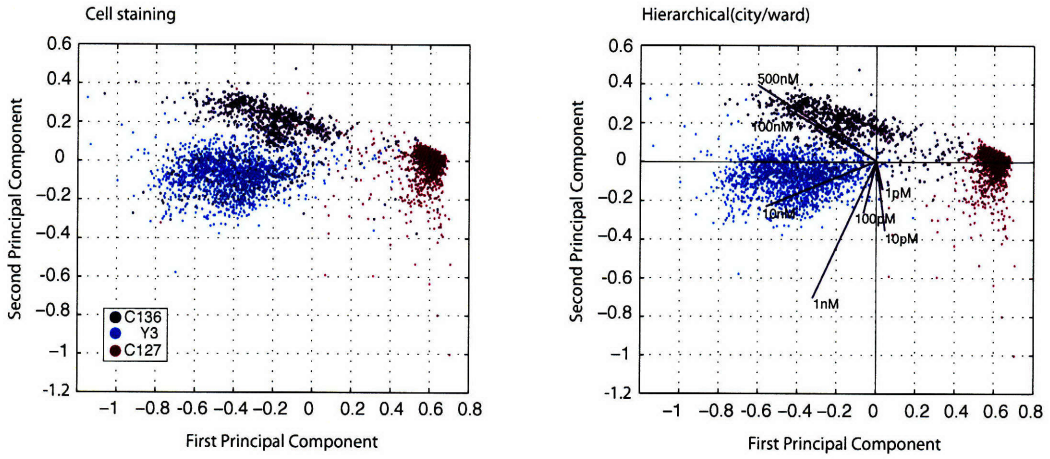

Figure 3-6: Principal-components analysis of binding profiles. Single cells are plotted as a function of the first two principal components. Colors indicate the identity of the cells when determined (A) by intracellular staining and (B) by hierarchical clustering. The vectors show the two-dimensional projection of the original multidimensional basis of the principal-component space.

Table 3.1: Sensitivity and specificity of clustering algorithms to hybridoma data

Classifier	**Y3**	**c136**	**c127**	**Avg. Sensitivity**	**Avg. Specificity**	**Avg. PPV**	**Avg. NPP**
Cellmarkers	1527	856	1328	—	—	—	—
K-means	1482	831	1398	87.0%	94.3%	87.0%	94.4%
Hierarchical	1617	744	1350	88.3%	95.2%	89.8%	95.6%

3.4.5 Conclusions

This work demonstrates the ability of distinguishing antibody-secreting cells on the basis of the affinity curve of their secreted antibody. Perhaps surprisingly, the correlation between affinity of the secreted antibody and identity of the secreting cell is a robust one. Therefore, antibodies affinity curves could potentially be employed as "signatures" of the antibodies-producing cells, in a manner analogous to the use of gene expression signatures as univocal identifiers of genes in oligonucleotides microarrays.

Along the same lines, analysis of microengraving data may lead to phenotypic studies of cellular populations (eg. the immune system) resembling in character the molecular studies that microarray technology has enabled thus far.

In particular, the ability to *profile* the cellular and humoral immune system at any given state may open up many opportunities for biological discovery and diagnostic applications.

Chapter 4

Multivariate Profiling of Immune Responses during the Course of a Vaccination

Having shown the ability of distinguishing immunoglobulin produced by antibody-secreting cells of similar specificities by their antigen binding curves, it should be possible to analyze the diversity among a more heterogeneous population of primary B cells. The ability to map in vivo B cell responses is of particular interest in the field of vaccinology and, thus we decided to test the feasibility of the microengraving approach in describing and monitoring a classic immunization.

The objective of the experiments was to test our ability to determine frequencies of antigen-specific cells, while at the same time measuring the affinity and isotype of secreted antibodies. Manipulation and correlation of these data by bioinformatics methods would in turn demonstrate the ability to compile profiles or "snapshots" of the in vivo humoral immune response to a protein antigen. Moreover, the ability to gather an highly multidimensional dataset might permit the detection of biologically significant groups of interrelated immune cells.

4.1 Immunizations

Immunization schemes vary considerably depending on i) the nature of the antigen, ii) its immunogenic potential (measured as the capacity to elicit protective level of circulating antibodies following exposure), iii) the dose, iv) the adjuvant employed and v) the delivery route.

The immunogen employed can either be an a) inactivated pathogen, b) an attenuated organism with reduced pathogenicity (eg. by passing the virus in cell cultures), c) a conserved protein component found in the pathogen, d) pathogens which have been genetically engineered to reduce mutations and potential for infection or e) peptides. Antigens can also be used in conjugated form (eg. to carrier proteins) in order to elicit more than one mechanism of immunity. Adjuvants, such as inorganic salts or emulsions containing killed bacteria, are routinely used to enhance the immunogenicity of antigens.

The immunogenic potential of an antigen can depend on a variety of factors. Larger, complex aggregated proteins are generally more immunogenic than soluble small molecules. Similarly, antigens that largely differ from self proteins and are well processed in the MHC pathway tend to have increased immunogenicity.

Doses of antigen employed in immunization should not be too high or too low. If the dose of immunogen administered in a primary immunization is too low, little or no immune response can be detected. If the dose is too high, on the other hand, the immune system can develop tolerance to the antigen, failing to develop immunological memory. Doses employed in subsequent immunization are usually progressively lower, since the presence of memory is responsible for immune responses at lower antigen doses.

The use of adjuvant is generally required to elicit sufficient immune responses and to favour the involvement of multiple mechanisms of immunity. Adjuvants enhance the immunogenicity of an antigen by a) converting soluble proteins into particulate material and b) by means of their microbial components, which are capable of stimulating the production of cytokines, induce inflammatory responses and activate TLR

receptors pathways.

Routes of immunogen administration change the rapidity by which the antigen is released and in turn is recognized by the cells of the immune system. In general, slower release favors immunogenicity, with the subcutaneous route being the most immunogenic, followed by intraperitoneal and intravenous.

Lastly, immunization schemes employed in human vaccinations are also affected by collateral considerations such as age, the potential infectivity and the risk of exposure for the particular individual.

4.2 Adaptive Immune Responses following Immunizations

Primary exposure to antigen and complete adjuvant triggers a *primary immune response*, which is characterized by the the production of specific antibodies and specific effector T cells. The intensity of the response, as measured by the production of antibodies, peaks around 6-7 days following injection and then slowly decreases but does not return to base level. Detectable levels of specific antibodies remain in the circulation as one of the signs of acquired immunological memory.

Subsequent responses can be distinguished from the primary response in terms of the frequency and type of antibodies produced, as well as differences in the cell surface markers of B and T cells involved.

The natural course of a B cell response to an antigen involves an initial burst of IgM production. Upon secondary stimulation, or if antigen persists and a suitable source of T cell help is available, antigen-specific B cells undergo antigen-driven proliferation (frequency of antigen-specific B cells in immunized animals can be up to 100 fold higher than in immunized animals), somatic hypermutation (with consequent affinity maturation of the secreted antibodies) and isotype switching (resulting in increasing levels of IgG antibodies being produced versus less mature IgM antibodies).

4.3 Collecting *snapshots* of the humoral immune responses

To test the feasibility of affinity-based profiling in monitoring an in-vivo immune response, we analyzed the antibody secretion of stimulated splenocytes harvested by mice undergoing an immunization series. The objective of this work was to determine whether microengraving would allow the description of known processes occurring during immunizations, such as antigen-specific B cell proliferation, affinity maturation and isotype switching. Furthermore, this work was directed to the application of clustering techniques for the detection of relevant B cell population previously unobserved.

Enhancements in the microengraving process allow to profile primary cells
Further refinements of the microengraving technique allowed us to analyze the secretory pattern of primary B cells, which are generally smaller (and thus more difficult to image), require stricter culture conditions and secrete less antibodies than hybrydoma lines. Specifically,

- Improvements in the signal to noise ratio:
  - Buffer optimization for the functionalization of glass slides with capture antibody
  - Optimization of the capture antibody mix used for slides functionalization
  - Optimization of engraving time vs. number of microarray replicates
- Faster and more accurate image analysis and data mining, by means of optimized scripts
- Choice of minimal cell separation methods prior to cell seeding in the PDMS device.
- Non-specific stimulation by lypopolysaccharide (LPS) to boost antibody secretion levels

Hyperimmunization Hybrid strain mice (between 129 and Balb/c) were immunized with a mixture of OVA (50 μg) and HEL (50 μg) in a CFA emulsion, and boosted at day 14, and day 27 with IFA antigen emulsion. Splenocytes (1×10^6 in RPMI 1640) were harvested from unimmunized mice, as well as immunized mice 18 days after the initial immunization, and stimulated with Lipopolysaccharide (LPS; 20 g/ml) for three days, and seeded into grids for microengraving analysis on day 21. On day 31, splenocytes from a second mouse, subjected to 1st and 2nd booster immunizations on day 14 and 27 respectively, were similarly collected and analyzed on day 34 after the primary immunization (see Figure 4-3).

To classify individual B cells, the three parameters we considered were specificity, affinity, and isotype. Specificity and isotype can be readily scored by application of a single reagent: a fluorescently labeled antigen (specificity) or an appropriate secondary antibody (isotype) to microarrays of captured antibodies produced by μEn[31]. Six replicate microarrays were generated by μEn. One replicate was stained with anti IgM (Alexa 594), anti IgG1 (Alexa 647), anti IgG2a (Alexa 532) and anti IgG2b (Alexa 488) antibodies to distinguish the isotype of the secreted antibodies. Five replicates were probed with anti-mouse Ig(H+L) (Alexa 647, red) and increasing concentrations of ovalbumin-Alexa 555 (green). After engraving, the cells were fixed with 4% paraformaldehyde, labeled with a nuclear stain (Hoescht 33342, 1 g/ml) and two primary antibodies (anti-IgM-Alexa 647 and anti-B220-rhodamine), and imaged for enumeration.

Multiparametric profiling of individual cells:surface markers and secreted antibodies Information obtained from data mining included 1) phenotypic surface-expressed markers for identifying individual B cells (IgM and B220), 2) the isotype of the antibody produced by a cell (IgM, IgG1, IgG2a, IgG2b), and 3) the binding curve of that antibody for ovalbumin.

The resulting multidimensional dataset was correlated on the basis of relative spatial coordinates. Once the full multidimensional dataset was assembled, a subset of wells was selected for further analysis.

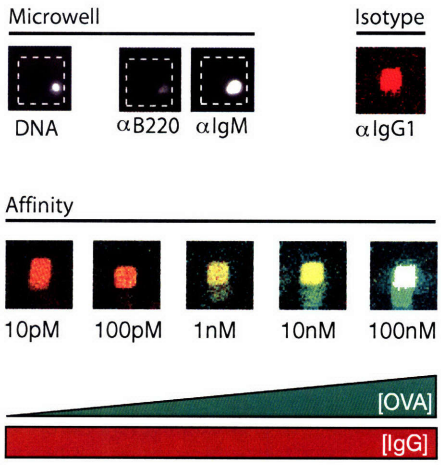

Figure 4-1: A representative set of raw data for a single splenocyte.

Filtering of the multidimensional dataset While some of the non-secreting B cells may be specific for ovalbumin, we had to limit our analysis to those cells induced to secrete antibodies. We further limited our dataset to those cells for which a complete binding curve was also available. All such datasets included five distinct and high quality replicates of the antibody taken from all microarrays interrogated with ovalbumin. The uniformity and signal-to-noise ratios for a given element on one or more replicates were factors that excluded some cells; both technical and biological variability contributed to the exclusions. While these stringent quality conditions further reduced the datasets, there remained at least 645 individual antibody-secreting B cells to analyze for each mouse.

4.4 Results

4.4.1 Antibodies with detectable affinities revealed antigen-specific B cells

Antigen specific clones in the mouse dataset were identified fitting a langmuir binding isotherm by non linear least squares (95% confidence interval) to every affinity curve

obtained and sorting the dataset in ascending order, based on the fitted dissociation constant K_d.

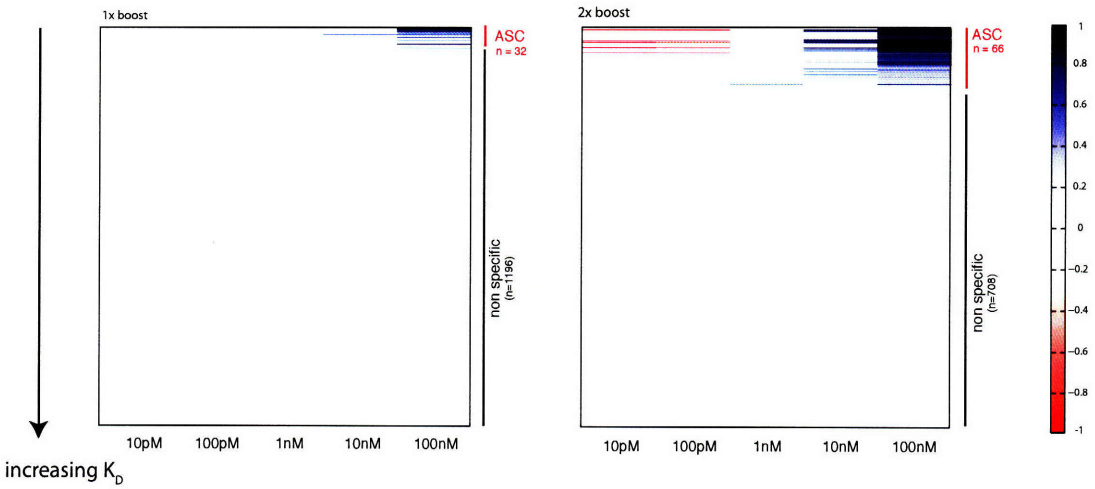

Figure 4-2: Heatmaps for binding curves measured for Ig+ cells with good quality data from immunized mice. The cells are sorted by dissociation constant. A threshold is set to distinguish a subset of "antigen-specific" clones.

A threshold was determined following visual inspection of the data, allowing to separate clones with measurable K_d (<100 μM) from low binders and non-specific clones.

4.4.2 Isotype information defined the antibody secreting population

We took the isotype slide as the reference measure of the Ig secreting cells in a given grid. Scans of isotype slides were consistently of better quality (ie. higher signal to noise ratio, higher number of positive spots) then the scans obtained for the affinity slides. This can be accounted for by the fact that isotype staining was performed on the first print obtained by the grid (with the fewest number of dislodged cells) and four separate anti Ig antibody were employed.

4.4.3 Microengraving detected an increase in the fraction of antigen-specific B cells

B cells (B220+ and/or IgM+) comprised the major fraction of single cells deposited in microwells after stimulation with LPS. Some 20-30% of those cells, however, actively secreted antibodies that corresponded to one of the four isotypes scored (IgM, IgG1, IgG2a, or IgG2b). For splenocytes from immunized mice that received no stimulation ex vivo, the percentage of immunoglobulin secreting cells was 0.1 1%. Immunization followed by a booster injection yielded an increase in the percentage of antigen-specific B cells.

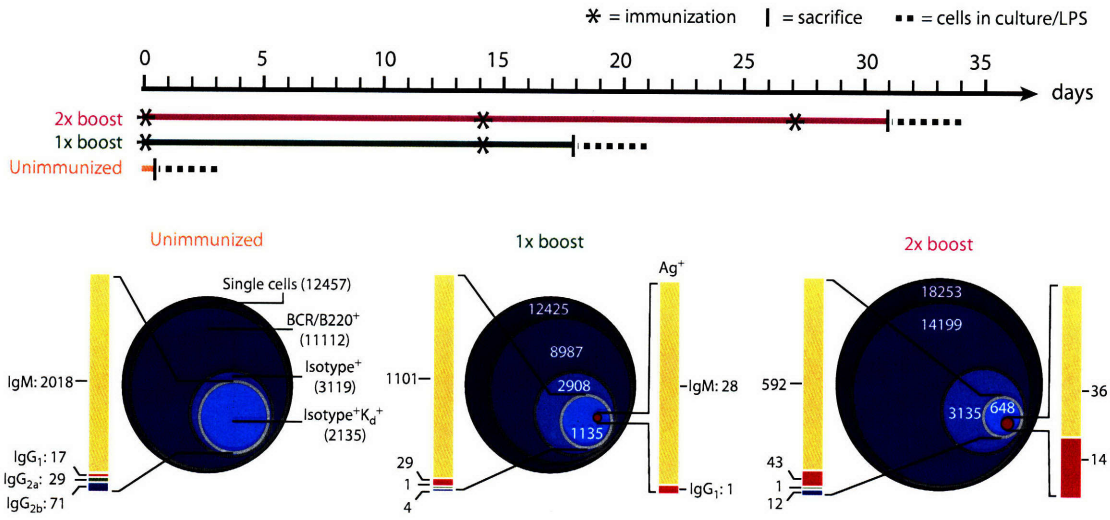

Figure 4-3: (top) Graphical representation of the immunization schedule used for the three mice profiled. (bottom) Populations of B cells scored for immunized mice. The total area of each circle is proportional to the number of cells enumerated with the phenotypes indicated. Single cells expressing either B220 or IgM on their surfaces were classified as B cells. The innermost green circle represents the subset of single secreting B cells for which a complete set of affinity data was obtained. The isotype distribution for this subset is indicated in the left column. Red circles represent the subset of single secreting B cells classified as antigen specific (with detectable dissociation constant, ie. Kd ¡ 100 nM). The isotype distribution in each of these subsets is shown in the right column. The quality of the microengraved microarrays can be estimated by comparing the number of single secreting B cells which yielded a complete set of affinity data to the total number of single secreting B cells detected (ie. relative size of the green circle with respect to the enclosing isotype+ circle). Here, from left to right, 68.5%, 39% and 20.7%.

Following a single immunization, the percentage of antigen-specific B cells was 2.5% of the Ig-secreting cells captured in our analysis. This percentage increased to 7.7% for a mouse receiving two immunizations (Figure 4-3). While the frequencies for antigen-specific cells reported in literature vary with schedules of immunizations, antigen, and genotype of the animal used, our values are similar to those previously determined by flow cytometry for actively responding B cells and memory B cells surface-stained with antigen.

4.4.4 Microengraving detected an increase in class-switched antibody producing B cells

As expected, the frequency of class-switched antibodies (IgG1) within the collection of antigen-specific antibodies was higher in the mouse receiving three immunizations (28%) than the mouse with two (3%). Although isotype switching should not occur in non specific clones, we have consistently measured a small proportion of IgG1 secreters in the non-antigen-specific population. These baseline IgG1 levels can be attributable to antigen specific clones, which show low affinity for antigen and therefore remain undetectable with our approach. IgG1 clones in the non specific population can also be due to exposure to other parasitic antigen in the individual mouse considered or impurities in ova immunization.

4.4.5 Microengraving detected affinity maturation

The median dissociation constant for the antigen-specific antibodies decreased with the number of immunizations. The majority of the measured affinities, however, were centered primarily around a relatively narrow range of values (20-30 nM). Only a small number of cells were producing antibodies with affinities 10 nM (Fig. S2).

This result raises the possibility that B cells producing antibodies with high affinities for an antigen may be present and active before titers of their antibodies are sufficiently high to detect in sera.

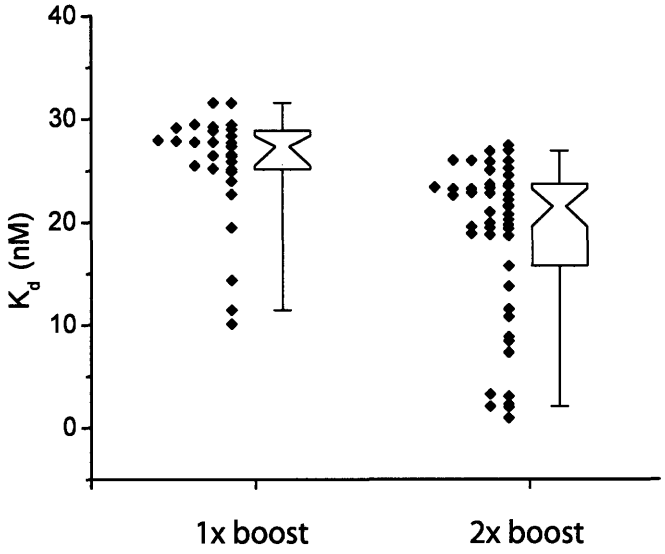

Figure 4-4: Plot of dissociation constants calculated from individual binding curves for immunized mice.

4.4.6 Closely related subgroups can be individuated amongst splenic B cells

To examine further the similarities among the antigen-specific B cells identified in immunized mice, we applied unsupervised hierarchical clustering (euclidean/average) to classify the cells on the basis of both the isotype and binding curves for their antibodies (Fig. 3C). For both immunized mice, dendrograms from the cluster analysis, and the calculated values for Kd, indicated the antigen-specific IgMs were largely indistinguishable (over the range of concentrations tested) by their binding curves. One noteworthy exception was observed in the population of cells from the mouse that received a primary immunization and one secondary challenge. The IgM with the lowest calculated value of Kd was close to that for the lone antigen-specific IgG1 observed, and both exhibited similar binding curves. The calculated Kd values for the IgG1 antibodies identified in the mouse immunized three times ranged from 1 nM to 25 nM, and analysis by clustering classified the cells into five populations based on

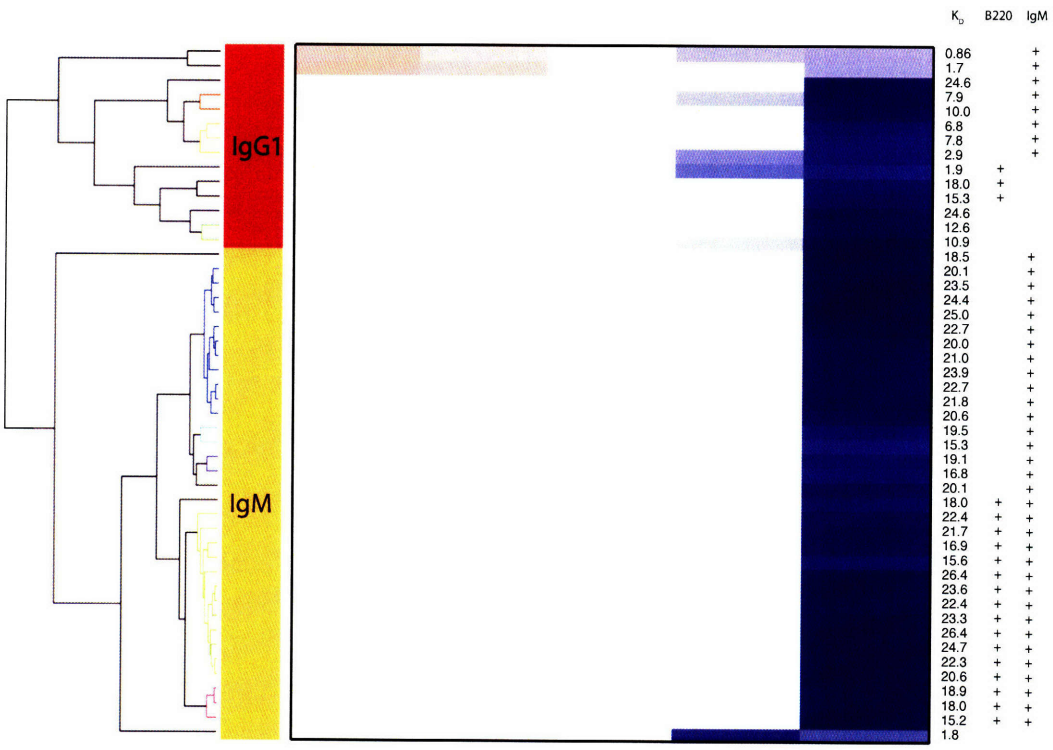

Figure 4-5: Unsupervised hierarchical clustering of antigen-specific cells identified after one or two boosters. Isotype data was included in the clustering, but set apart graphically for clarity. The colored branches in the dendrograms indicate cells with closely related binding curves. The dissociation constants (nM) calculated from a langmuir curve fit for each clone are listed to the right of the data for each cell.

shape and intensities of their binding curves.

4.4.7 Conclusions

These results show the appearance and expansion of antigen-specific cells, affinity maturation, and isotype switching in immunocompetent mice sampled at different times before and during immunization with chicken ovalbumin, and effectively give a "snapshot" of the quantities and quality of individual antibody-producing B cells at a given point in time. It is important to emphasize that not only microengraving-based profiling of an immune responses is successful in showing the expected pattern of a general immune response, but also that these *observations are done for the first time at the level of single antibody-secreting splenocytes*, a compartment of the humoral immunity previously inaccessible.

Chapter 5

Discussion

5.1 Conclusions

This work describes the development and application of the microengraving approach to the high-throughput affinity analysis of antibody repertoires and the multivariate profiling of in vivo immune responses. The microengraving approach yields quantitative data on the frequencies, specificities, affinities, and isotypes of the antibodies secreted by primary B cells from immunized mice and also enables direct correlation of that data to the phenotypes of the cells secreting them.

5.1.1 Microengraving allows routine state based profiling of immune responses

The use of bioinformatics techniques allows harnessing the multivariate datasets generated by microengraving for the generation of true "snapshots" of cellular immunity, capable of describing an humoral immune response by detailing the characteristics and effects of each immune cell involved. The profiling of a *large* number of *single* cells by microengraving permits us to describe the immune response to an antigen both at the level of the individual players of immunity and at the level of the overall response.

Table 5.1: Differential expression of surface biomarkers allows to distinct different B cell phenotypes and stages in the development of memory B cells. *Adapted from [36]*

Surface Marker	Naive B cells	Memory B cell	Plasma Cells
Antigen	-	++	++
B220	++	++	+
IgD	+++	+	-
CD138	-	-	++++

5.1.2 Clusters of B cells can be routinely individuated

B lymphocytes are generally classified in distinct subgroups (eg. germinal center cells, memory cells, etc.) by means of cell sorting techniques. The expression of determined surface markers has been associated with distinct in vivo phenotypes and is employed to map the development stages of memory B cells under a host of different biological conditions and pathogens exposures.

This work demonstrates how microengraving can be used in analogous manner to FACS, but with the added advantage of providing data on the antibodies secreted by each B cell. The ability to collect in a single experiment correlated data on the surface phenotype and the secreted products of individual lymphocytes allows us to highlight much finer differences amongst cells with similar expression patterns. Clusters of interrelated (or interdependent) immune cells can be identified by unsupervised methods, in a manner analogous to the classification of genes into families on the basis of their microarray expression profiles.

Moreover, differently from FACS analysis or ELISPOT analysis, data generated by microengraving is multivariate in nature and immediately correlated for each individual cell under analysis. Normally, the percentage of cells being part of a group is determined by means of intersecting boolean categorizations (eg. B220+ *and* IgD+, instead of B220+ *and* IgD-). Microengraving allows us to perform the same categorization on the basis of surface markers, integrating it with multidimensional information about the secreted antibodies (or cytokines). Each single cell can thus be described by a complex multidimensional *signature* that unequivocally characterizes

it. The ability to obtain such detailed signatures may lead to the development of finer and more precise biomarkers of B cell development.

5.1.3 Microengraving can be routinely employed in immunological studies

The simplicity of the microengraving process and the relative inexpensive materials employed (PDMS chips and fluorescent antibodies) make it possible to perform these measurement routinely in a laboratory setting, adding to the tools routinely employed by immunologists in studying immune responses. A single microengraving experiment requires a very limited sample to be performed, thus allowing parallel analysis of the samples by means of other techniques.

5.1.4 Advantage over similar methods

The ability to map antibodies to single cells in spatially-defined arrays of microwells offers an advantage over plaque-based assays that require dilutions of cells over relatively large areas (100 cm^2) to ensure spatial isolation. In particular, μEn does not require assuming clonality by limiting dilution, and does not require analysis of multicellular colonies. In principle, cells of interest can be retrieved by micromanipulation for genetic analysis[31].

5.1.5 Antibodies secretion at the level of the spleen can be directly measured

The ability of evaluating B cell secretion allows better evaluation of the interplay between the cellular immune response to antigens (here limited to primary B cells, but in principle expandable to every immune cell) and the accompanying humoral immune response (as determined by serum antibody levels). Our data, for instance, appears to suggest important differences between the distribution of isotypes in the antibodies secreted at the level of the spleen and the circulating antibodies following

immunization. Secretion of IgM antibody is abundant even after hyperimmunization, when the response is expected to shift towards the IgG1 isotype. Considerations regarding the half life of the circulating antibodies and their trafficking in and out of the spleen may play an important role and should be part of future work.

5.1.6 Summary

In summary profiling of immune responses by microengraving makes it is possible to compare the quality and diversity of a humoral immune response with a level of resolution not possible with other single assays. It not only provides data equivalent to both cytometry (phenotype of cells) and immunosorbant assays (isotype, specificity, and frequency) with one-to-one correlations, but also characteristics not easily attained by other existing methods with single-cell resolution (ie. affinity).

Routinely available profiles describing the state and evolution of a cellular immune response should enhance the quality of predictive diagnosis and understanding in the pathogenesis of diseases, and help determine the efficacy of vaccines.

5.2 Future Directions

This thesis describes the experimental and bioinformatic methods involved in the development of the microengraving process. Future work will span multiple directions, ranging from technological development of the process to the profiling of immune responses to viral pathogens. The following sections outline some suggestions for further development in some of these areas.

5.2.1 Technological developments

Even though microengraving is a relatively simple experimental procedure, it still requires some amount of training and suffers from poor automation. Substantial

technological development should be directed toward:

1. the standardization of the microengraving protocol
2. the creation of specific devices and engineering solutions that simplify the current process
3. the streamlining of data mining instrumentation and algorithms.

Once more standardized protocols are established and a set of appropriate tools is developed, the quality and repeatability of the process should improve drastically. In particular, a) staining of glass slides could be done by means of commercially available automation, b) chip fabrication could be standardized by the fabrication of an injection mold and c) the placement of the chip in contact with the glass slide could be performed by means of an appropriate mechanical device instead of relying on manual placement. In order to streamline the process of data mining it is necessary to:

1. improve and customize the chip design
2. develop dedicated instrumentation for imaging and acquisition
3. compile custom software for deployment on computational clusters

New chips should be designed according to the imaging equipment which will be employed in the analysis: i) Repeating clusters of wells should fit in a single field of view of the microscope employed, avoiding the necessity of stitching images via software; ii) More regular spacing between these clusters would aid automated image acquisition, automated grid alignment by software and database storage; iii) An internal coding system that contains fiduciary markers and allows the coordinate of each cluster to be determined without looking at any elements outside the microscope view will also facilitate downstream software image analysis.

Dedicated instrumentation would reduce the required setup time and the actual acquisition time currently required to image microwell chips and produced protein microarrays. Instruments based on spectral imaging are currently able to resolve up to 11 distinct fluorescent signals and would allow a substantial increase in the information density of each experiment.

Grid alignment, image analysis, correlation of the multidimensional dataset and subsequent analysis are currently performed by relatively slow proprietary software (metamorph and genepix) and custom MATLAB routines (see appendix A). Substantial software development is possible and appears relatively straightforward at this stage. Hinging on the ongoing development of Cell Profiler (Harvard/MIT Broad Institute, Anne Carpenter), it should be possible to implement a system capable of aligning grids, analyze image and create a database of the acquired images in one single program, capable of running on dedicated computational clusters. These improvements should speed up the data mining exponentially, rendering it a matter of few minutes instead of days. The resulting database would also allow the filtering, exploration and analysis of the complex multidimensional dataset by the end user.

These efforts should enable the routine collection of data by microengraving, adding the ability to select relevant antibody producing cells and profile the status of humoral immune responses to the panel of techniques at the disposal of the immunologist.

5.2.2 Cytokine Profiling

The work presented in this thesis is exclusively concerned with antibodies when considering materials secreted by the cell under analysis. Although antibodies are of paramount significance in the study of primary B lymphocytes, other secreted products are of interest when profiling the dynamic of immune responses. Cytokines in particular are responsible for lymphocytes activity and homeostasis, as well as the regulation of dendritic cells and antigen presenting cells in general.

Cytokine secretion measurements would allow to detect and profile other fundamental players of the immune response (such as T lymphocytes and dendritic cells), while looking at antibodies secretion allows to fully characterize only secreting B lymphocytes. Moreover, by measuring secretion of a panel of cytokines and integrating the data with antibody secretion and surface expression of relevant markers, the dimensionality - and thus the sensitivity - of the data collected by microengraving can increase substantially.

Cytokine secretion measurements will require a) substantial improvements in the signal-to-noise ratio for the printed protein microarrays currently obtainable by the process (which should stem in part by a more standardized and streamlined process) and b) a cheaper source of cytokine-specific capture antibodies.

5.2.3 Immune responses to pathogens

This thesis presents a proof-of-principle study outlining the potential to profile in vivo immune responses by means of detailed quantitative multidimensional data. However, this study is limited to the observation of immunity developed following exposure to a protein antigen. Actual viral or bacterial infections are known to generate more complex and extensive immune responses in the host. The application of microengraving to the monitoring and profiling of such immune responses will be paramount for further validation of the technology.

Direct comparison with results obtained by FACS and ELISPOT techniques, as well with observations reported in literature, would allow to further assess the precision of the technology and may uncover important biological observations previously unaccessible by these methods. Cytokine/antibodies profiling of responses to different classes of pathogens (bacterial or viral, acute or chronic, latent microorganisms, etc.) will allow to assess the importance of rare populations of immune cells to the progression and maintenance of infections.

5.2.4 Meta-analysis & Predictive Modeling

Profiles of immune responses based on microengraving data are extremely rich in information. These profiles, which describe a given immune system in a particular state with some degree of completeness, can be correlated with other informations. In principle, in an animal infected by a pathogen, it is possible to link such profiles at a given time during infection with information relating to the causative pathogen, the duration, the ultimate course of the infection,etc.

The progressive accumulation of profiles compiled based on microengraving data

would then allow the development of predictive algorithms, capable of determining the probable outcomes of immune responses (and corresponding infections) on the basis of a single microengraving measurement. This is analogous to the demonstrated use of gene microarray in the prognosis of various form of cancer. Obviously, such use will depend on further validation of the technique, as well as on the accumulation of a large number of standardized dataset (which in turn would depend on the succesful creation of a standard format and a public database system).

5.2.5 Antibody repertoires

The work described in this thesis limits its attention to the antibodies directed against a single antigen. Work directed at mapping in vivo antibody repertoires, by employing multiple antigens, will lead to profiles of even higher dimensionality. Such information-dense descriptions of ongoing immune responses will permit more precise evaluations and differentiation of immune statuses and may grant microengraving profiling higher diagnostic and predictive power.

Epitope specificity of the secreted antibodies is also not explored in the work presented here. Microengraving could be employed to map in a high-throughput fashion the contribution of different epitopes to the immunogenicity of particular antigens. Similarly, cross-reactivities of produced and circulating antibodies to different strains and mutation of the same pathogen could be carefully mapped. Data generated in this fashion may allow to individuate conserved epitopes or factors required for the propagation and maintenance of infections.

5.2.6 Memory B cells

Lastly, the ability to obtain data on antibody secretion of rare primary B cells, previously unattainable by other methods, makes microengraving an ideal tool in the study of memory B cell development. Different stages of memory B cell generation in the spleen could be mapped with greater precision and correlated to humoral immunity. The secretory dynamic of memory B cells before and after antigen stimulation

could be directly observed, allowing to evaluate the magnitude of recruitment and new memory generation. It also becomes possible to individuate and enumerate rare antigen-primed memory B cells without the necessity of antigen stimulation. These studies have the potential to shed light on important immunological questions, such as the mechanism of tolerance, the persistence of latent infections and the manifestation of autoimmune disease.

Appendix A

MATLAB code

A.1 Data acquisition and manipulation

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


%% extract stats from AFF data
suffix = 'agabratio';
%!remember to choose whether we use igpos only or doublepos ova/ig


exps = {'D0_W','D0_M','D0_X',
            'D15_W','D15_M','D15_X',
                'D33_W','D33_M','D33_X'};
stats = [repmat({'Ratio of Medians (488/532)'},1,3),
            repmat({'Ratio of Medians (532/635)'},1,6)];
logs = {'WTd0','MD4d0','XBPd0',
            'WTd15','MD4d15','XBPd15',
                'WTd33','MD4d33','XBPd33'};


for e =1:numel(exps)

    %extract data from 5 slides with increasing concentration
    datoarray =[];
    for conc = 2:6
        dato = [];
        flagdata = [];
        flags = [];
        eval(char(strcat('tempslide = ',exps(e),int2str(conc),';')))
        dato = tempslide.Data(:,find(strcmp(tempslide.ColumnNames,stats(e))));
        flagdata = ...
            tempslide.Data(:,find(strcmp(tempslide.ColumnNames,'Flags')));
        if isempty(dato)
            warning('did not find the column')
        end
```

```
        dato(¬(flagdata == 0),1) = NaN;
        datoarray = [datoarray,dato];
    end


    %save array into properly named variable after initializing it
    eval(char(strcat(exps(e),suffix,' = [];')))
    eval(char(strcat(exps(e),suffix,' = datoarray;')))

end
```

A.2 Filtering of data points

A.2.1 Mouse data: Affinity

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst



function flaggedaff = flag_aff_d0_m(mastruct)
% colors
%d0
% 488 ¬ ova
% 532 ¬ igG & IgM

    TI532 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 Total Intensity')));
    TI488 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 Total Intensity')));

    dia = mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'Dia.')));

    SNR532 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 532')));
    SNR488 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 488')));

    CV532 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 CV')));
    CV488 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 CV')));

    sat532 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 % Sat.')));
    sat488 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 % Sat.')));

    pxabbg532 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B532+2SD')));
    pxabbg488 = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B488+2SD')));
```

```
    Smed = ...
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'Sum of Medians (488/532)')));

%% rulesets
% here we do care about saturation.... b/c it skews the ratio towards lower
% values
    satthres = 2;

%because of the nature of these signals, both the SNR ratio
%AND the px above bkgrnd
%in these datapts is NOT usable!


%md relaxed
  decent = (dia ≥ 50);
  TIthres532 = 500000;
  covthres = 100;
  Smedthres = 1750;
  TIthres488 = 120000;

  snrthres532 = 5;
  pxthres532 = 35;
  snrthres488 = 1.5;
  pxthres488 = 25;

%md tight
%   decent = (dia > 55);
%   TIthres532 = 250000;
%   covthres = 80;
%   Smedthres = 2200;
%   TIthres488 = 180000;
% pxthres532 = 55;
% snrthres532 = 5;
% snrthres488 = 2;
% pxthres488 = 30;

igpos = decent & ...
    ((TI532 > TIthres532) | (SNR532 > snrthres532)) & ...
    (CV532 < covthres) & (pxabbg532 > pxthres532) & (sat532 < satthres);

%use the sum of the medians before we consider the pt
% we couldn't really put a threshold on the TI or we'll skew the ratios
% towards high values. Sum of the medians would discriminate less.
% However, given the nature of the signal and the crosstalk between stains
% sum of the medians would just pick pts that are high in Ig stain, paradoxically
% skewing the ratio down! Hence...

ovapos = decent & ...
    ((TI488 > TIthres488) & (SNR488 > snrthres488)) &...
    (CV488 < covthres)  & (pxabbg488 > pxthres488) & (sat488 < satthres);
%ovapos is not anymore a subset of igpos but a completely separate set,
%which pleases a little mroe my mathematically inclined mind
```

```
%the "decent" spots are given by the union of these two sets

%the spots "considered" are the intersection of these two.
%NOTE!! this gives a baseline to the ratios... ie. undetectable OVA is not
%counted.

flaggedaff = mastruct;
flaggedaff.igpos = igpos;
flaggedaff.ovapos = ovapos;

end
```

A.2.2 Mouse data: Isotype

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


%% ISO ruleset

%%colors are
% 594 ¬ IgM
% 488 ¬ IgG2b
% 532 ¬ IgG2a
% 635 ¬ IgG1

%order of input is
%[IgM, IgG1, IgG2a, IgG2b]
%[594   635   532    488 ]

%for 594 SNR is better than TI in discriminating.
%for other colors is equally valid.
%in terms of CALLS:
% %%%%%%%%%%xbp%%%%%%%%%%%%%
% relaxed:
% D0_XISO =
% flag_iso_tighter(D0_XISO,[150000 150000 200000 220000],[3 1.5 1.5 2],25)
% D15_XISO =
%  flag_iso_tighter(D15_XISO, [200000 300000 275000 300000],[2.5 2 3 2.5],25)
% D33_XISO =
%  flag_iso_tighter(D33_XISO,[150000 200000 200000 220000],[1.5 1 5 1.5],25)
%
%
% D0_MISO =
% flag_iso_tighter(D0_MISO,[250000 400000 650000 450000],[1.7 1.5 3 4], 25)
% D15_MISO =
%  flag_iso_tighter(D15_MISO,[200000 520000 250000 350000],[2.5 1.5 2 4],25)
% D33_MISO =
%  flag_iso_tighter(D33_MISO,[400000 300000 250000 300000],[1.8 1.5 3 5],25)
%
%
```

```
% D0_WISO =
% flag_iso_tighter(D0_WISO,[270000 450000 600000 375000],[2 2 1.75 1.2],25)
% D15_WISO =
%  flag_iso_tighter(D15_WISO, [450000 400000 220000 300000],[2 2 3 4.5],25)
% D33_WISO =
%  flag_iso_tighter(D33_WISO,[250000 300000 275000 275000],[2 1.7 9 3],25)

%


%% Eli Papa
% filtering out Ig positive spots from 4-color ISO data
% order of input is [IgM, IgG1, IgG2a, IgG2b]
function flaggediso = flag_iso_tighter(mastruct,TIthres,snrthres,pxthres)

    TI532 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 Total Intensity')));
    TI635 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F635 Total Intensity')));
    TI594 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F594 Total Intensity')));
    TI488 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 Total Intensity')));


    dia =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'Dia.')));


    SNR532 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 532')));
    SNR635 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 635')));
    SNR594 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 594')));
    SNR488 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'SNR 488')));


    CV532 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 CV')));
    CV635 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F635 CV')));
    CV594 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F594 CV')));
    CV488 =
mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 CV')));


    sat532 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F532 % Sat.')));
    sat635 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F635 % Sat.')));
    sat594 =
```

```
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F594 % Sat.')));
    sat488 =
 mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'F488 % Sat.')));


    pxabbg532 =
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B532+2SD')));
    pxabbg635 =
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B635+2SD')));
    pxabbg594 =
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B594+2SD')));
    pxabbg488 =
    mastruct.Data(:,find(strcmp(mastruct.ColumnNames,'% > B488+2SD')));



%we don't care about saturation so much
    satthres = 10;
    covthres = 100;

    decent = (dia ≥ 50);

igG1pos = decent & ...
    (SNR635 > SNR594) & (SNR635 > SNR532) & (SNR635 > SNR488) & ...
    ((TI635 > TIthres(2)) & (SNR635 > snrthres(2))) & ...
    (CV635 < covthres) & (pxabbg635 > pxthres) & ...
    (sat594 < satthres) & (sat532 < satthres) & (sat488 < satthres);
% want to exclude saturation in other channels, b/c usually when there is
% saturation is more than one, dust particles are responsible. on the other
% hand, if just one channel saturates, we'd like to keep the data and
% consider that channel the "winner" (ie. the stronger signal which in turn
% determines the isotype of the spot)
igMpos = decent & ...
    (SNR594 > SNR635) & (SNR594 > SNR532) & (SNR594 > SNR488) & ...
    ((TI594 > TIthres(1)) & (SNR594 > snrthres(1))) & ...
    (CV594 < covthres)  & (pxabbg594 > pxthres) & ...
    (sat635 < satthres) & (sat532 < satthres) & (sat488 < satthres);

igG2apos = decent & ...
    (SNR532 > SNR635) & (SNR532 > SNR594) & (SNR532 > SNR488) & ...
    ((TI532 > TIthres(3)) & (SNR532 > snrthres(3))) &...
    (CV532 < covthres)  & (pxabbg532 > pxthres) & ...
    (sat594 < satthres) & (sat635 < satthres) & (sat488 < satthres);

igG2bpos = decent & ...
    (SNR488 > SNR635) & (SNR488 > SNR594) & (SNR488 > SNR532) & ...
    ((TI488 > TIthres(4)) & (SNR488 > snrthres(4))) &...
    (CV488 < covthres)  & (pxabbg488 > pxthres) & ...
    (sat594 < satthres) & (sat532 < satthres) & (sat635 < satthres);



flaggediso = mastruct;
flaggediso.decent = decent;
```

```
flaggediso.igG1pos = igG1pos;
flaggediso.igMpos = igMpos;
flaggediso.igG2apos = igG2apos;
flaggediso.igG2bpos = igG2bpos;

end
```

A.2.3 Hybridoma data

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst

% flagging features in a two color hybridoma expt

function flaggedmastruct = flag_goodspots_3xKb(mastruct,thres)


    colSnrG = find(strcmp(mastruct.ColumnNames,'SNR 532'));
    colSnrR = find(strcmp(mastruct.ColumnNames,'SNR 635'));
    colCVG = find(strcmp(mastruct.ColumnNames,'F532 CV'));
    colCVR = find(strcmp(mastruct.ColumnNames,'F635 CV'));
    colGsat = find(strcmp(mastruct.ColumnNames,'F532 % Sat.'));
    colRsat = find(strcmp(mastruct.ColumnNames,'F635 % Sat.'));
    colSmed = find(strcmp(mastruct.ColumnNames,'Sum of Medians (532/635)'));
    colpxabbgG = find(strcmp(mastruct.ColumnNames,'% > B532+2SD'));
    colpxabbgR = find(strcmp(mastruct.ColumnNames,'% > B635+2SD'));

SnrG = mastruct.Data(:,colSnrG);
SnrR = mastruct.Data(:,colSnrR);
CVG = mastruct.Data(:,colCVG);
CVR = mastruct.Data(:,colCVR);
Gsat = mastruct.Data(:,colGsat);
Rsat = mastruct.Data(:,colRsat);
Smed = mastruct.Data(:,colSmed);
pxabbgG = mastruct.Data(:,colpxabbgG);
pxabbgR = mastruct.Data(:,colpxabbgR);

decentspots = (CVG < 100) & (CVR < 100) & (Gsat < 2) & (Rsat < 2);

maxSmed = max(Smed(decentspots)); %trying to exclude saturated pixel
%%%%%%%%%%%%%%%
%autothres = 0.018*maxSmed;  %get output, just to check
%thres ~ 650 for 1pM; 850 for 500nM

%just set it in advance for everyone:

goodspots = decentspots & (Smed > thres);


Ig_pos = (pxabbgR > 50) | ((pxabbgR > 20) & (SnrR > 1.5));

%px above bg R:  in kb1pm 50 seems a good value,
```

```
%                           20 is just for extreme cases,
%                           but does not do much;
%                 use it just bc we have cell data

%Snr R: could be more restrictive, but you get some good spots for 1.5

%NB - spots that are green but not red may be due to some strange occurence
%in the expt, but they should be ruled out by the constraints on the Ig

%NB - we decided to keep spots that seems to have no Ag signal. This may
%increase the noise, but makes sure we are detecting very low conc of Ag.
%If the cell continues to not bind any antigen, it will be excluded anyway
%when we align all affinities (or it be clustered together)


%save what's done
flaggedmastruct = mastruct;
flaggedmastruct.goodspots = goodspots;
flaggedmastruct.Ig_pos = Ig_pos;

% and bring everything together
flaggedmastruct.goodratios = goodspots & Ig_pos;
%this still needs to pass through cell counts masks

end
```

A.2.4 Cells

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst

%% from cell logs, extract secreters, single cells, etc.

function logstruct = classifywells(logarray)
    logstruct.source = logarray;
    logstruct.populatedwell = (logarray(:,4) ≥ 1);
    logstruct.singleb220 = ...
        (logarray(:,5) == 1) & (logarray(:,6) == 0) & (logarray(:,7) == 0);
    logstruct.singleantibcr = ...
        (logarray(:,5) == 0) & (logarray(:,6) == 1) & (logarray(:,7) == 0);
    logstruct.singledoublest = ...
        (logarray(:,5) == 0) & (logarray(:,6) == 0) & (logarray(:,7) == 1);
    logstruct.singlebcell = ...
        logstruct.singleb220 |
            logstruct.singleantibcr |
                logstruct.singledoublest;


    logstruct.singlenonsecreter = (logarray(:,4) == 1) & ...
        (logarray(:,5) == 0) & (logarray(:,6) == 0) & (logarray(:,7) == 0);

    logstruct.singlecell =  ...
```

```
        logstruct.singlenonsecreter | logstruct.singlebcell;
end


% example call
%%%%%%%%%%%%%%%%%
% logs = whos('*log');
%
% for k = 1:numel(logs)
% eval([logs(k).name, ' = classifywells(',logs(k).name ,');'])
% end

%call from structs
% logs = whos('*log');
%
% for k = 1:numel(logs)
% eval([logs(k).name, ' = classifywells(',logs(k).name ,'.source);'])
% end
```

A.3 Visualization

A.3.1 Create *affinity maps* visualizations

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


function affmap(data,varargin)
figure

pcolor(padarray(data,[1 1],0,'post'))
shading flat
colormap(french(256,2))
set(gca,'Clim',[-1 1], 'XTick',[1.5:1:5.5],'XTickLabel', ...
            {'10pM','100pM','1nM','10nM','100nM'})

if ~isempty(varargin) && varargin{1} ==1
    set(gca,'YTick',[1.5:1:size(data,1)+1],'YTickLabel',...
        [1:1:size(data,1)])
    if ~isempty(varargin{2})
        set(gca,'YTickLabel',varargin{2})
    end
else
    axis off
end

end
```

A.3.2 Associate information with *affinity maps* visualizations

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst

% isolate subset and calculate things needed to make a pretty heatmap

%eg call
% d3310nmap =
% makelabeledsubsetmaps(D33_Wagabratio,(D33W_idh_10nmbndr & ¬isnan(D33_WISOs)));

function mapstruct = ...
    makelabeledsubsetmaps(concs,affdata,isodata,isonumdata,iddata,mask)

%save basis of data
mapstruct.mask = mask;
mapstruct.aff = affdata(mask,:);

%row normalize! and center too..
for r=1:size(mapstruct.aff,1)
        mapstruct.normaff(r,:) = ...
            mapstruct.aff(r,:) + min(mapstruct.aff(r,:));
        mapstruct.normaff(r,:) = ...
            mapstruct.normaff(r,:) / max(mapstruct.normaff(r,:));
        mapstruct.normcenteraff(r,:) = ...
            mapstruct.normaff(r,:) - median(mapstruct.normaff(r,:));
    end

%calculate and save Kd's
mapstruct.Kd = fitBindK(concs,mapstruct.normaff);
%save other stuff
mapstruct.iso = isodata(mapstruct.mask);
mapstruct.isonum = isonumdata(mapstruct.mask);
mapstruct.ids = iddata(mapstruct.mask);



%sort based on isotype and save reordering
[mapstruct.sorted mapstruct.sortindex] = ...
    sortrows([mapstruct.isonum,mapstruct.aff]);
mapstruct.affsorted = mapstruct.sorted(:,2:end);
mapstruct.isonumsorted = mapstruct.sorted(:,1);
clear mapstruct.sorted

    %sorter the other stuff
    mapstruct.isosorted = mapstruct.iso(mapstruct.sortindex);
    mapstruct.Kdsorted = mapstruct.Kd(mapstruct.sortindex);
    mapstruct.normaffsorted = mapstruct.normaff(mapstruct.sortindex,:);
    mapstruct.normcenteraffsorted = ...
        mapstruct.normcenteraff(mapstruct.sortindex,:);
    mapstruct.idsorted = mapstruct.ids(mapstruct.sortindex);

%adjust isotype "colors"
mapstruct.isonumsorted(mapstruct.isonumsorted == 1) = -0.5;
mapstruct.isonumsorted(mapstruct.isonumsorted == 2) = 0.5;
```

```
%plot reference maps
affmap(mapstruct.normcenteraffsorted,1,mapstruct.isosorted)
affmap(mapstruct.normcenteraffsorted,1,mapstruct.Kdsorted)
title(['median Kd (norm): ', num2str(median(mapstruct.Kdsorted))])
affmap(mapstruct.normcenteraffsorted,1,mapstruct.idsorted)

end


function Kest = fitBindK(concs, ratiodata)

langmuir = @(k,xdata) (xdata)./(k + xdata);
Kest = [];

for row = 1:size(ratiodata,1)

    ydata = ratiodata(row,:);

    if any(isnan(ydata))
        warning('missing data pt')
    end

    k = lsqcurvefit(langmuir,rand(1),concs,ydata);
    Kest = [Kest; k];

end
end
```

A.3.3 Color Mapping

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


function h = french(m,flag)
%FRENCH    French's flag color map.


if nargin < 1
    m = size(get(gcf,'colormap'),1);
else
    if isempty(m)
        m = size(get(gcf,'colormap'),1);
    end
end
if nargin < 2, flag = 1; end


n1 = fix(3*m/8);
n2 = fix(m/4);
n3 = fix(m/2);
```

```
switch flag
    case 1
        r = [ones(n3,1); (sqrt(1-((1:n3)/n3).^2))'];
        g = [flipud( (sqrt(1-((1:n3)/n3).^2))'); (sqrt(1-((1:n3)/n3).^2))'];
        b = [flipud((sqrt(1-((1:n3)/n3).^2))'); ones(n3,1)];
    case 2
        r = [ones(n1+n2,1); (n1-1:-1:0)'/n1;];
        g = [(0:n1-1)'/n1; ones(n2,1); (n1-1:-1:0)'/n1;];
        b = [(0:n1-1)'/n1; ones(n1+n2,1);];

end


h = [r g b];

if size(h,1) < m
    h(ceil(m/2)+1:m,:) = h(ceil(m/2):end,:);
    h(ceil(m/2),:) = 1;
end
```

A.4 Compare clustering methods to assess quality of distance/linkage combinations

A.4.1 Hierarchical

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


function [coeffcell,trees, dists] = compareclustermethod(data)

coeffcell = {};
distancemethod = {
    'euclidean';
    'seuclidean';
 %   'mahalanobis';
    'cityblock';
 %   'minkowski';
  %   'cosine';
   %  'correlation';
     'spearman';
%     'hamming';
%     'jaccard';
    'chebychev';
    };
linkmethod = {
 %   'single';
    'complete';
    'average';
    'weighted';
    'median';
```

```
%   'centroid';
    'ward';
    };

trees = cell(numel(distancemethod),numel(linkmethod));
dists = cell(numel(distancemethod),2);
coeffcell = cell(numel(distancemethod),numel(linkmethod));



%find distances and store them in a table with a descriptor
for j = 1:numel(distancemethod)
    dists{j,2} = pdist(data,distancemethod{j});
    dists(j,1) = {distancemethod{j}};
end


%perform linkage to form trees and evaluate the cophenet coeff as
%determinant of quality
for l = 1:numel(linkmethod)
    for d = 1:numel(distancemethod)
%         if (strcmp(linkmethod{l},'ward') |
%           strcmp(linkmethod{l},'median')) &
%               (¬strcmp(distancemethod{d},'euclidean'))
%             continue %these linkages fcns need euclidean distance input
%         end

        trees{d,l} = linkage(dists{d,2},linkmethod{l});
        coeffcell{d,l} = cophenet(trees{d,l},dists{d,2});

    end
end


%fill in the rows and columns of the coeff and trees matrix
trees = [distancemethod, trees];
coeffcell = [distancemethod, coeffcell];

trees = [{'DIST\LINK'},linkmethod'; trees];
coeffcell = [{'DIST\LINK'},linkmethod';coeffcell];

% %example call
% exps = {'D0_W','D0_M','D0_X','D15_W','D15_M','D15_X','D33_W','D33_M','D33_X'};
% for e = 1:numel(exps)
%     eval(char(strcat('[',exps(e),'coph , ',exps(e),'trees , ',exps(e),'dist]
%       %= compareclustermethods(',exps(e),'agabratio);')))
% end


%
%
% set(0,'RecursionLimit',3000)
% figure
% plt = 1;
```

```
% for i = 1:7
%     for j = 1:5
%     subplot(7,5,plt);
%     plt = plt+1;
%     if isempty(trees{i,j})
%         continue
%     end
%     dendrogram(trees{i,j},'colorthreshold','default');
%     set(gca,'TickDir','out','TickLength',[.002 0],'XTickLabel',[]);
%     end
% end
```

Dendrograms

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


set(0,'RecursionLimit',3000)
figure
plt = 1;
for i = 1:7
    for j = 1:5
    subplot(7,5,plt);
    plt = plt+1;
    if isempty(trees{i,j})
        continue
    end
    dendrogram(trees{i,j},'colorthreshold','default');
    set(gca,'TickDir','out','TickLength',[.002 0],'XTickLabel',[]);
    end
end
```

A.4.2 K-means

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


distmeth = {'cityblock','sqEuclidean','correlation'};

for meth = 1:numel(distmeth)

    for numclust = 2:6
    [temp,tempmeans] = kmeans(allratiodata_single_good,numclust,'dist',...
            distmeth{meth},'display','final','replicates',5);
        eval([ 'idx',int2str(numclust),'_',distmeth{meth},' = temp;']);
        eval([ 'idx',int2str(numclust),'mean_',distmeth{meth},' = temp;']);
%     figure
%     [tempsilh,h] = silhouette(allratiodata_single_good,temp,distmeth{meth});
```

```
%     title( [distmeth{meth},' ',int2str(numclust)]);

    eval(['silh',int2str(numclust),'_',distmeth{meth},' = tempsilh;']);
    end
end
```

A.5 Extract frequency statistics

```
% Eliseo Papa
% Harvard/MIT Health Science & Tech. Inst


logs = {'WTd0','MD4d0','XBPd0',
            'WTd15','MD4d15','XBPd15',
                'WTd33','MD4d33','XBPd33'};
exps = {'D0_W','D0_M','D0_X',
            'D15_W','D15_M','D15_X',
                'D33_W','D33_M','D33_X'};

for grid = 1:numel(exps)

    jointigpos = [];
    jointovapos = [];
    jointbothpos = [];
    for concs = 2:6
        eval(char(strcat('tempslide = ',exps(grid),int2str(concs),';')))
        jointigpos = [jointigpos, tempslide.igpos];
        jointovapos = [jointovapos, tempslide.ovapos];
        jointbothpos = [jointbothpos, (tempslide.ovapos & tempslide.igpos)];
    end


    eval(char(strcat('grid_celllog = ',logs(grid),'_celllog;')))


    %compare ig+ trough slides with cellcounts
    igpositives_foraff = (sum(jointbothpos,2) ≥ 3);
    truthtable = [ sum(grid_celllog.singlecell & igpositives_foraff),
                    sum(¬grid_celllog.singlecell & igpositives_foraff);
                   sum(grid_celllog.singlecell & ¬igpositives_foraff),
                    sum(¬grid_celllog.singlecell & ¬igpositives_foraff)];

    b220truthtable = [ sum(grid_celllog.singleb220 & igpositives_foraff),
                        sum(¬grid_celllog.singleb220 & igpositives_foraff);
                   sum(grid_celllog.singleb220 & ¬igpositives_foraff),
                    sum(¬grid_celllog.singleb220 & ¬igpositives_foraff)];

    antiBCRtruthtable = [ sum(grid_celllog.singleantibcr & igpositives_foraff),
                            sum(¬grid_celllog.singleantibcr & igpositives_foraff);
                   sum(grid_celllog.singleantibcr & ¬igpositives_foraff),
                    sum(¬grid_celllog.singleantibcr & ¬igpositives_foraff)];
```

```
    eval(char(strcat(exps(grid),'_ttable = truthtable;')))
    eval(char(strcat(exps(grid),'_ttable = b220truthtable;')))
    eval(char(strcat(exps(grid),'_antiBCRttable = antiBCRtruthtable;')))


end
```

A.6 Binding curve fit

```
function estimates = fitbindingcurve(xdata,ydata)
% Call fminsearch with a random starting point.
start_point = rand(1, 3);
model = @bindingcurve;
options.MaxFunEvals = 1e7;
estimates = fminsearch(model, start_point, options);

    function sse = bindingcurve(params)
        Kd = params(1);
        max = params(2);
        min = params(3);
        FittedCurve = min + max*(xdata ./ (Kd + xdata));
        ErrorVector = FittedCurve - ydata;
        sse = sum(ErrorVector .^ 2);
    end
end
```

Bibliography

[1] *Global atlas of infectious diseases: an interactive information and mapping system.* World Health Organization.

[2] Rafi Ahmed, Michael B A Oldstone, and Peter Palese. Protective immunity and susceptibility to infectious diseases: lessons from the 1918 influenza pandemic. *Nat Immunol*, 8(11):1188–93, Nov 2007.

[3] D Barouch, N Letvin, and R Seder. The role of cytokine dnas as vaccine adjuvants for optimizing cellular immune responses. *Immunological Reviews*, Jan 2004.

[4] Nadia L Bernasconi, Elisabetta Traggiai, and Antonio Lanzavecchia. Maintenance of serological memory by polyclonal activation of human memory b cells. 298(5601):2199–202, Dec 2002.

[5] Timothy J Bradford, Xiaoju Wang, and Arul M Chinnaiyan. Cancer immunomics: using autoantibody signatures in the early detection of prostate cancer. *Urol Oncol*, 24(3):237–42, Jan 2006.

[6] Ulisses M Braga-Neto and Ernesto T A Marques. From functional genomics to functional immunomics: new challenges, old problems, big rewards. *PLoS Comput Biol*, 2(7):e81, Jul 2006.

[7] Atul Butte. The use and analysis of microarray data. *Nature reviews Drug discovery*, 1(12):951–60, Dec 2002.

[8] Atul J. Butte, Pablo Tamayo, Donna Slonim, Todd R. Golub, and Isaac S. Kohane. Discovering functional relationships between rna expression and chemotherapeutic susceptibility using relevance networks. *Proc Natl Acad Sci USA*, 97(22):12182, Oct 2000.

[9] Michel Caron, Geneviève Choquet-Kastylevsky, and Raymonde Joubert-Caron. Cancer immunomics using autoantibody signatures for biomarker discovery. *Mol Cell Proteomics*, 6(7):1115–22, Jul 2007.

[10] A Casadevall, M Feldmesser, and L Pirofski. Induced humoral immunity and vaccination against major human fungal pathogens. *Current Opinion in Microbiology*, Jan 2002.

[11] Rodrigo F Chuaqui, Robert F Bonner, Carolyn J M Best, John W Gillespie, Michael J Flaig, Stephen M Hewitt, John L Phillips, David B Krizman, Michael A Tangrea, Mamoun Ahram, W Marston Linehan, Vladimir Knezevic, and Michael R Emmert-Buck. Post-analysis follow-up and validation of microarray experiments. *Nat Genet*, 32 Suppl:509–14, Dec 2002.

[12] Irun R Cohen. Real and artificial immune systems: computing the state of the body. *Nat Rev Immunol*, 7(7):569–74, Jul 2007.

[13] C C Czerkinsky, L A Nilsson, H Nygren, O Ouchterlony, and A Tarkowski. A solid-phase enzyme-linked immunospot (elispot) assay for enumeration of specific antibody-secreting cells. *J Immunol Methods*, 65(1-2):109–21, Dec 1983.

[14] C C Czerkinsky, A Tarkowski, L A Nilsson, O Ouchterlony, H Nygren, and C Gretzer. Reverse enzyme-linked immunospot assay (relispot) for the detection of cells secreting immunoreactive substances. *J Immunol Methods*, 72(2):489–96, Sep 1984.

[15] Susmita Datta and Somnath Datta. Comparisons and validation of statistical clustering techniques for microarray gene expression data. *Bioinformatics*, 19(4):459–66, Mar 2003.

[16] Thomas Dörner and Andreas Radbruch. Antibodies and b cell memory in viral immunity. *Immunity*, 27(3):384–92, Sep 2007.

[17] J Foote and H N Eisen. Kinetic and affinity limits on antibodies produced during immune responses. *Proc Natl Acad Sci USA*, 92(5):1254–6, Feb 1995.

[18] L Hangartner, R Zinkernagel, and H Hengartner.... Antiviral antibody responses: the two extremes of a wide spectrum. *Nat. Rev. Immunol*, Jan 2006.

[19] Julie Hardouin, Jean-Paul Lasserre, Loïk Sylvius, Raymonde Joubert-Caron, and Michel Caron. Cancer immunomics: from serological proteome analysis to multiple affinity protein profiling. *Ann N Y Acad Sci*, 1107:223–30, Jun 2007.

[20] K Hayakawa, R Ishii, K Yamasaki, T Kishimoto, and R R Hardy. Isolation of high-affinity memory b cells: phycoerythrin as a probe for antigen-binding cells. *Proc Natl Acad Sci USA*, 84(5):1379–83, Mar 1987.

[21] P Hegde, R Qi, K Abernathy, C Gay, S Dharap, R Gaspard, J E Hughes, E Snesrud, N Lee, and J Quackenbush. A concise guide to cdna microarray analysis. *BioTechniques*, 29(3):548–50, 552–4, 556 passim, Sep 2000.

[22] Charles Janeway, Kenneth P Murphy, Paul Travers, and Mark Walport. *Janeway's immunobiology*. Garland Science, New York, NY, 7th ed. edition, 2008.

[23] N Jerne and A Nordin. Plaque formation in agar by single antibody-producing cells. *Science*, 140(3565):405, Apr 1963.

[24] S M Kaech and R Ahmed. Memory cd8+ t cell differentiation: initial antigen encounter triggers a developmental program in naïve cells. *Nat Immunol*, 2(5):415–22, May 2001.

[25] G Kelsoe. Cloning of mitogen- and antigen-reactive b lymphocytes on filter paper discs. i. a description of the technique and of methods for the analysis of colonies. *J Immunol Methods*, 76(2):345–63, Feb 1985.

[26] G Kelsoe. Cloning of mitogen- and antigen-reactive b lymphocytes on filter paper disks: phenotypic and genotypic analysis of b cell colonies. *Meth Enzymol*, 150:287–304, Jan 1987.

[27] G Kelsoe and J T Stout. Cloning of mitogen- and antigen-reactive b lymphocytes on filter paper discs. ii. paratope frequencies within the mitogen-selected repertoire. *Cell Immunol*, 98(2):506–16, Apr 1986.

[28] N R Klinman and G Aschinazi. The stimulation of splenic foci in vitro. *J Immunol*, 106(5):1338–44, May 1971.

[29] Antonio Lanzavecchia, Nadia Bernasconi, Elisabetta Traggiai, Claudia R Ruprecht, Davide Corti, and Federica Sallusto. Understanding and making use of human memory b cells. *Immunol Rev*, 211:303–9, Jun 2006.

[30] Yuk Fai Leung and Duccio Cavalieri. Fundamentals of cdna microarray data analysis. *Trends Genet*, 19(11):649–59, Nov 2003.

[31] J Christopher Love, Jehnna L Ronan, Gijsbert M Grotenbreg, Annemarthe G van der Veen, and Hidde L Ploegh. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nat Biotechnol*, 24(6):703–7, Jun 2006.

[32] B Maecker, von Bergwelt-Baildon, K S Anderson, R H Vonderheide, and J L Schultze. Linking genomics to immunotherapy by reverse immunology–'immunomics' in the new millennium. *Curr Mol Med*, 1(5):609–19, Nov 2001.

[33] Vinay S Mahajan, Ilya B Leskov, and Jian Zhu Chen. Homeostasis of t cell diversity. *Cell Mol Immunol*, 2(1):1–10, Feb 2005.

[34] M Juanita Martinez, Anthony D Aragon, Angelina L Rodriguez, Jose M Weber, Jerilyn A Timlin, Michael B Sinclair, David M Haaland, and Margaret Werner-Washburne. Identification and removal of contaminating fluorescence from commercial and in-house printed dna microarrays. *Nucleic Acids Research*, 31(4):e18, Feb 2003.

[35] M Maruyama, K P Lam, and K Rajewsky. Memory b-cell persistence is independent of persisting immunizing antigen. *Nature*, 407(6804):636–42, Oct 2000.

[36] Louise J McHeyzer-Williams and Michael G McHeyzer-Williams. Antigen-specific memory b cell development. *Annu Rev Immunol*, 23:487–513, Jan 2005.

[37] M G McHeyzer-Williams, M J McLean, G J Nossal, and P A Lalor. The dynamics of t cell-dependent b cell responses in vivo. *Immunol Cell Biol*, 70 (Pt 2):119–27, Apr 1992.

[38] Amy S McKee, Michael W Munks, and Philippa Marrack. How do adjuvants work? important considerations for new generation adjuvants. *Immunity*, 27(5):687–90, Nov 2007.

[39] R Nayak, S Mitra-Kaushik, and M S Shaila. Perpetuation of immunological memory: a relay hypothesis. *Immunology*, 102(4):387–95, Apr 2001.

[40] J Quackenbush. Computational analysis of microarray data. *Nat Rev Genet*, 2(6):418–27, Jun 2001.

[41] John Quackenbush. Microarray data normalization and transformation. *Nat Genet*, 32 Suppl:496–501, Dec 2002.

[42] Francisco J Quintana, Gad Getz, Guy Hed, Eytan Domany, and Irun R Cohen. Cluster analysis of human autoantibody reactivities in health and in type 1 diabetes mellitus: a bio-informatic approach to immune complexity. *J Autoimmun*, 21(1):65–75, Aug 2003.

[43] Francisco J Quintana, Peter H Hagedorn, Gad Elizur, Yifat Merbl, Eytan Domany, and Irun R Cohen. Functional immunomics: microarray analysis of igg autoantibody repertoires predicts the future response of mice to induced diabetes. *Proc Natl Acad Sci USA*, 101 Suppl 2:14615–21, Oct 2004.

[44] Rino Rappuoli. Bridging the knowledge gaps in vaccine design. *Nat Biotechnol*, 25(12):1361–6, Dec 2007.

[45] M Regner and P Lambert. Autoimmunity through infection or immunization? *Nature Immunology*, Jan 2001.

[46] J W Rohrer, K Vasa, and R G Lynch. Myeloma cell immunoglobulin expression during in vivo growth in diffusion chambers: evidence for repetitive cycles of differentiation. *J Immunol*, 119(3):861–6, Sep 1977.

[47] Victor C Rucker, Karen L Havenstrite, and Amy E Herr. Antibody microarrays for native toxin detection. *Anal Biochem*, 339(2):262–70, Apr 2005.

[48] R A Schultz, T Nielsen, J R Zavaleta, R Ruch, R Wyatt, and H R Garner. Hyperspectral imaging: a novel approach for microscopic analysis. *Cytometry*, 43(4):239–47, Apr 2001.

[49] R Seder and A Hill. Vaccines against intracellular infections requiring cellular immunity. *Nature*, Jan 2000.

[50] Michael B Sinclair, Jerilyn A Timlin, David M Haaland, and Margaret Werner-Washburne. Design, construction, characterization, and application of a hyperspectral microarray scanner. *Applied optics*, 43(10):2079–88, Apr 2004.

[51] Seng-Lai Tan, Gopinath Ganji, Bryan Paeper, Sean Proll, and Michael G Katze. Systems biology and the host response to viral infection. *Nat Biotechnol*, 25(12):1383–9, Dec 2007.

[52] Andreas Wack and Rino Rappuoli. Vaccinology at the beginning of the 21st century. *Curr Opin Immunol*, 17(4):411–8, Aug 2005.

[53] Nan-Ping Weng, Kebin Liu, Marta Catalfamo, Yu Li, and Pierre A Henkart. Il-15 is a growth factor and an activator of cd8 memory t cells. *Ann N Y Acad Sci*, 975:46–56, Dec 2002.

[54] E John Wherry, Daniel L Barber, Susan M Kaech, Joseph N Blattman, and Rafi Ahmed. Antigen-independent memory cd8 t cells do not develop during chronic viral infection. *Proc Natl Acad Sci USA*, 101(45):16004–9, Nov 2004.

[55] G M Whitesides, E Ostuni, S Takayama, X Jiang, and D E Ingber. Soft lithography in biology and biochemistry. *Annual review of biomedical engineering*, 3:335–73, Jan 2001.

[56] D Wraith, M Goldman, and P Lambert. Vaccination and autoimmune disease: what is the evidence? *The Lancet*, Jan 2003.

[57] R M Zinkernagel, M F Bachmann, T M Kündig, S Oehen, H Pirchet, and H Hengartner. On immunological memory. *Annual Review of Immunology*, 14:333–67, Jan 1996.