Vaccine 33 (2015) 5271–5281

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

New approaches to understanding the immune response to vaccination and infection

David Furman [a,b], Mark M. Davis [a,b,c,*]

[a] Institute for Immunity, Transplantation and Infection, School of Medicine, Stanford University, Stanford, CA, United States
[b] Department of Microbiology and Immunology, School of Medicine, Stanford University, Stanford, CA, United States
[c] Howard Hughes Medical Institute, School of Medicine, Stanford University, Stanford, CA, United States

ARTICLE INFO

Article history:
Available online 29 July 2015

Keywords:
Systems immunology
Immune profiling
High-throughput methods
Vaccinology
Computational immunology
Regularization
Feature selection
Elastic net
Human immunology

ABSTRACT

The immune system is a network of specialized cell types and tissues that communicates via cytokines and direct contact, to orchestrate specific types of defensive responses. Until recently, we could only study immune responses in a piecemeal, highly focused fashion, on major components like antibodies to the pathogen. But recent advances in technology and in our understanding of the many components of the system, innate and adaptive, have made possible a broader approach, where both the multiple responding cells and cytokines in the blood are measured. This systems immunology approach to a vaccine response or an infection gives us a more holistic picture of the different parts of the immune system that are mobilized and should allow us a much better understanding of the pathways and mechanisms of such responses, as well as to predict vaccine efficacy in different populations well in advance of efficacy studies. Here we summarize the different technologies and methods and discuss how they can inform us about the differences between diseases and vaccines, and how they can greatly accelerate vaccine development.

1. Introduction

The immune system is a complex adaptive system with emergent properties such as memory and self-regulation. Its complexity can be exemplified at the network level, lymphocyte receptor diversity, clonotype selection, cell migration, cell–cell interaction inside immunological tissues and long-distance communication via fluid dissemination throughout the body, homeostatic regulation and adaptation to changing environments. The net functionality of a healthy immune system likely depends upon the interaction between immune system components at the molecular level and considerable integration and regulation at the system level. Therefore, as in any complex system, it is likely that immune system function cannot be predicted from the behavior of any of its parts separately but rather, is context dependent and hard-wired in dynamic and functional networks involving hundreds of components. Nevertheless, our current knowledge of immunology has been built for many years using a relatively deterministic and reductionist approach, as was necessary given our relative ignorance of the many components that have only recently come to light. However, examples where the same cell or soluble cell product exert distinct and even opposing functions depending on the site of the immune reaction and the presence of other immune constituents are abundant. Thus, the robust functioning of the immune system likely relies on a highly complex multi-level interaction network, linking intracellular biochemical networks, intercellular communication networks and inter-organ cellular trafficking networks through space and time. In addition, for obvious ethical and practical reasons, we have utilized model antigens and largely relied on mouse models of health and disease, which while extremely useful for deciphering the cellular and molecular bases of many immune responses, are rarely predictive of human vaccine results. There are a number of possible explanations for this discrepancy, evolution being one, with an estimated 65 million years separating humans from mice [1] and second possibility being that mice are kept in a relatively sterile environment, whereas cage-free humans are exposed to a much broader range of pathogens over their (much longer) lifespan. A third factor is that the whole inbred mouse strains are homozygous at all their alleles whereas humans rarely if ever are.

Recent years have seen the emergence of 'systems biology' approaches that are now applicable to human studies [2,3]. Initially using gene expression analysis of white blood cells from

* Corresponding author at: 279 Campus Drive, B219 Beckman Center, Stanford, CA 94305, United States. Tel.: +1 650 725 4755; fax: +1 650 498 7771.
E-mail address: mmdavis@stanford.edu (M.M. Davis).

http://dx.doi.org/10.1016/j.vaccine.2015.06.117

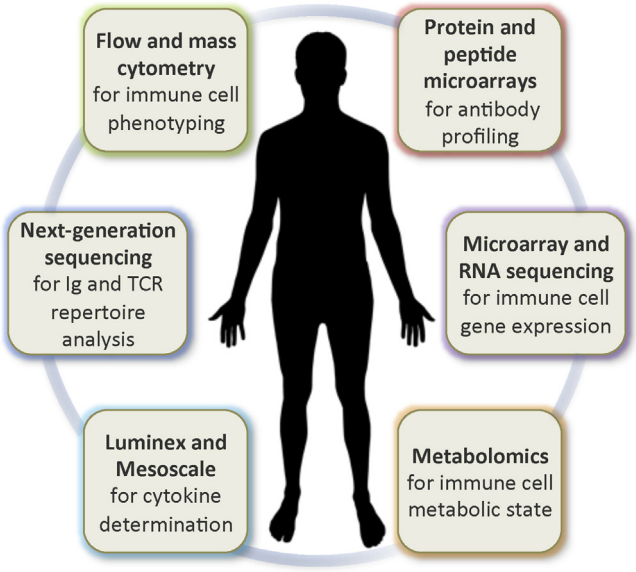

Fig. 1. Multi-level high-throughput analysis of the human immune system. Comprehensive immune profiling involves multiple technological platforms that allow us to capture and observe an important portion of the immune response. Peripheral blood is used to survey the perturbations in the immune system through a suite of available techniques including next-generation sequencing (NGS); gene and protein microarrays; multiparameter flow cytometry and mass cytometry (CyTOF); multiplex cytokine and chemokine analysis by Luminex and Mesoscale, and metabolomics that relies on major recent improvements in mass spectrometry now capable of resolving close to a thousand metabolites. This human-centered approach to immunology promise to improve our understanding of the immune response to vaccination and infection.

vaccinated subjects [4,5], this approach has now been extended to cover not only gene expression, but the basic components of the entire immune system, namely the hundreds of cell types and subsets, and many of the cytokines and chemokines that they communicate with [6] (Fig. 1). Here it should be noted that the "quanta" of the immune system are the many specialized cells that act relatively autonomously. T lymphocytes for example, can detect even a single molecule of a peptide antigen bound to an MHC molecule [7–9] and then act on that information by releasing cytokines in the case of $CD4^+$ cells [10]. Advances in DNA sequencing technology also make it possible to analyze the immunoglobulin (Ig) and T-cell receptor (TCR) repertoires responding to vaccines in great detail [11–13] and to do this on a single cell level as well [14,15]. It is also possible to obtain the exact DNA sequence of the whole genome in single individuals, the information about the genes that are expressed in a particular cell state, and the composition of hundreds of different metabolites from different tissues that provide extremely valuable information about a particular metabolic status in health and disease states (Fig. 1). This ability to generate high-throughput and high bandwidth data has co-evolved with technological advances in informatics to enable the generation of integrative models of the human immune response. This systems biology approach applied to infection and vaccination using people as a model is accelerating and will continue to accelerate our understanding of how the immune system works in humans, representing a necessary step to advance our basic understanding of immune system dysregulation and to enable future translational applications based on these basic findings. In this review paper, we discuss some of the most important technological advances to probe the immune system and the computational tools used to extract relevant biological information that can be used to gain mechanistic insights and/or to identify immune biomarkers.

2. Genes expressed in immune cells prior to (predictive of outcome) or in response to immune perturbation (microarray and RNA sequencing)

One of the most successful technologies originated from available genetic information generated by the Human Genome Project, is the gene expression microarray. Because this technology generates large amounts of expression data for a relatively low price, microarrays have gained extreme popularity during the last decade. The technology consists in hybridizing a nucleic acid sample (target) typically onto a glass surface containing microscopic spots with multiple identical strands of DNA that are printed by a robot, each of it representing one gene probe. Probe-target hybridization is typically quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target [16]. DNA microarrays allow for determination of genome-wide expression profiles and thus, are ideally suited for generating hypotheses to gene function that can help to identify appropriate targets for vaccine and therapeutic intervention. DNA microarrays have been used to systematically identify tumor antigens for tumor vaccine design [17]; to identify gene profiles or "signatures" in patients with bacterial pneumonia [18,19] and bacterial sepsis both in adults [20] and children [21]; in rhinovirus, respiratory syncytial virus, and influenza A infections [22]; malaria [23] and dengue virus infections [24]; in HIV patients [25,26]; as well as in different vaccination regimes such as influenza [27–30], yellow fever (YF) and meningococcus [31] vaccines. The use of this technique is accelerating our understanding of the bases of the host immune response to pathogenic insults and is also extending to the genetic characterization of genetically diverse infectious pathogens associated with a given disease.

3. Cells and biomarker proteins expressed in immune cells prior to (predictive of outcome) or in response to immune perturbation (flow cytometry, CyTOF)

3.1. Flow cytometry

Flow cytometry is probably the most commonly used technology in immunology research. It uses antibodies coupled with fluorophores to detect specific proteins expressed intracellularly and on the cell surface. It has been widely used for many decades to monitor immune responses to vaccination and infection in bulk cell populations as well as to track the phenotypic and functional characteristics of antigen-specific cells, but it has also been largely applied in routine clinical settings for the diagnosis, prognosis and monitoring of disease. For example, it helps in the diagnosis and staging of patients with a hematological diseases [32]; for the detection of minimal residual disease (disease beyond the limit of morphological detection using conventional microscopy) [33]; for stem cell enumeration during immunosuppressive therapies [34]; in solid organ transplantation to evaluate T cell cross-match [35]; to monitor changes in cell populations after cardiopulmonary bypass surgery for the prediction of infections in risk patients [36]; in HIV for the determination of $CD4^+$ T cell counts [37]; to predict hemolytic disease [38]; in primary [39] and secondary [40] immunodeficiencies; and largely used in blood transfusion [41].

The immunogenicity of vaccination and infection, and direct monitoring of the innate and adaptive immune responses can be measured by different methodologies. Flow cytometry can be used to analyze the contribution of innate immunity to vaccine efficacy and disease pathogenesis [42]. For T cells, intracellular cytokine staining (ICS) assays have proven to be useful to measure T-cell immunogenicity and there are numerous examples in the literature

[43–46]. Despite the fact that the T cell response to vaccination and infection has been widely studied, there is no robust T cell measure that correlates with vaccine-induced protection. For example, an HIV vaccine was shown to be highly immunogenic as measured by ICS, yet was not effective [47]. Additional measures for the T cell response include the assessment of cytotoxic potential, which can be achieved by measuring degranulation and granule contents in specific T cell subsets. Degranulation indicates cytotoxic potential and is typically measured by the expression of CD107a on the cell surface. This protein is normally expressed in internal granular membranes and is transiently expressed on the cell surface during degranulation [48]. Broader antibody panels such as those covering the analysis of multiple cytokines, cell surface markers, and other functional markers such as perforin, CD107a, and CD154 with up to 10-color resolution have also been presented [49]. For the $CD4^+$ T cell response, assessment of helper capacity by measurement of the expression of CD40 ligand is also common.

Several studies have analyzed the characteristics of antibody-producing B cells in the blood of vaccinated or infected individuals by flow cytometry. B cell differentiation into plasmablasts (PB) is most commonly monitored with the surface markers CD19, CD20, CD27, CD38, and CD138. The term "acute plasmablast" is often used to associate the $CD19^{low}$ $CD20^-$ $CD27^{high}$ $CD38^{high}$ $CD138^{+/-}$ cell population appearing after infection with the acute phase of the immune response and to differentiate them from steady-state plasmablasts [50]. The timing of acute PB appearance in the blood is consistent after immunization or infection with numerous studies showing that after vaccination with yellow fever vaccine (YF-17D) [4], inactivated influenza vaccine (TIV) [51–55], tetanus vaccine [56,57] and after infection with respiratory syncytial virus (RSV) [57] or dengue virus [58,59] PB numbers peak consistently at day 6 or 7 for recall responses and somewhat later ~day 10 for new responses [53,59]. Specific T cells are also circulating within that time frame, and this general principal also applies to at least some forms of autoimmunity, particularly celiac disease, where it has been shown that gluten specific $CD4^+$ T cells appear in the circulation six days post-gluten challenge [60] as well as other types of T cells [61]. This is a very fortunate circumstance for human work, since it means that a suitably timed blood sample can capture a population of lymphocytes highly enriched for cells specific for the antigenic stimulus (in the case of PB) [53] estimated that 50–80% were specific for the immunizing flu vaccine 6 days post-vaccination. This spike in specific PB in the blood after vaccination or gluten challenge, in the case of celiacs, is transient-down to baseline at 10+ days, whereas the duration of the response to infection depends on the persistence of the antigen. For example, after infection with influenza or dengue the PB numbers decreases to baseline level within 2–3 weeks after the onset of disease [58,59] whereas in RSV-infected patients, circulating PBs are produced as long as the virus is actively shed from infected cells [62].

3.2. Cytometry by time-of-flight (CyTOF)

The recently developed mass cytometer CyTOF instrument combines the advantages of single-cell analysis in flow cytometry with the ability to resolve approximately 45 metal probes with minimal signal overlap because the signals are detected as discrete line spectra in mass spectroscopy, with a clear separation between labels differing by only 1 Da in molecular weight. This enables an exponential leap in the amount of data collected and has revealed previously unknown complexities in lymphoid populations.

This technology has been utilized to track the signaling behavior of different cell subsets spanning defined hematopoietic hierarchies [63]; to characterize the phenotypical and functional status of T cells specific for CMV, EBV and influenza in combination with MHC-tetramer staining [64]; to monitor cellular states in response to multiple drugs [65]; to detect specific cell cycle phases of different immune cells [66]; to characterize the T cell response against over 100 epitopes [67]; to monitor antigen-specific $CD8^+$ αβ and γδ T cells in celiac disease [61]; to study the diversity of NK cells in humans [68,69]; to find the levels of antigen required to activate versus inhibit signaling cascades [70]; to characterize antiviral immunity to heterologous prime-boost vaccination strategies against hepatitis C virus [71]; to investigate varicella-zoster virus (VZV) infection of tonsil T cells [72]; to compare signaling pathway activation and cytokine production between the split influenza vaccine and a prototypical TLR response *ex vivo* [73]; to characterize signaling network relationships in $CD4^+$ T cells [74]; to identify phenotypic and functional immune responses to surgical trauma [75]; to better characterize the mucosal-associated invariant T (MAIT) cells [76]; to better understand the human B cell lymphopoiesis [77]; and more recently, in a study on twins, to decouple the effects of genetics versus environment in the composition of dozens of immune cells, among many other variables including the response to influenza vaccination [78]. Many of these data sets have been deposited in publicly available databases (for example, the Immunology Database and Analysis Portal, ImmPort).

The CyTOF instrument is a promising technology with huge advantages over common flow cytometry, especially for the identification of new cellular functions and cell markers important in the response to infection and vaccination. It is particularly suited to situations where the sample material is limited, such as pediatric samples, as only a few million PBMCs can yield a very comprehensive dataset [79] (Sigal et al., unpublished).

4. Immunoglobulin and T cell receptor repertoire analysis (Next Generation Sequencing)

Recent advances in nucleic acid sequencing have allowed the determination of the diversity and clonal expansion of responding Ig and TCR sequences in astonishing numbers and depth-with hundreds of thousands to millions of "reads" becoming common with the most advanced instruments. These "Next Generation Sequencing" (NGS) methodologies started with the breakthrough 454 instruments from Roche (introduced in the year 2004) but then has advanced to more high throughput instruments such as LifeTechnologies Ion Torrent and the Illumina MiSeq and HiSeq. These two technologies employ similar base methodology that includes template preparation, sequencing and imaging, and data analysis [80]. The process starts with the construction of a library of nucleic acids (DNA or complementary DNA (cDNA)) off of which new DNA fragments are synthesized. Then the sequencing occurs through a cycle of washing and flooding the fragments in a sequential order; as nucleotides incorporate into the growing DNA strand, they are digitally recorded as sequence. The PGM and the MiSeq each rely on a slightly different mechanism for detecting nucleotide sequence information. The PGM depends on the detection of pH changes (semiconductor sequencing) induced by the release of a hydrogen ion when the nucleotide is incorporated into a growing strand of DNA [81]. By contrast, the MiSeq relies on the detection of fluorescence generated by the incorporation of fluorescently labeled nucleotides into the growing strand of DNA. NGS performs massively parallel sequencing, during which millions or billions of DNA fragments from unique samples can be sequenced, minimizing the need for the fragment-cloning methods used in Sanger sequencing, thus facilitating high-throughput sequencing, which allows an entire genome to be sequenced in less than 1 day. NGS enables a very broad approach to Ig or TCR repertoire analysis. The applications of NGS are multiple but for the purpose of this review we highlight only those that pertain to immune variability, vaccinology and infection. NGS has been used to identify genetic variants

associated with immune cell phenotypes in healthy individuals and patients with autoimmune disease [82], as well as to study, in newborns, the variability in cytokine and chemokine expression, key soluble factors that regulate immune responsiveness [83]. It was also applied to characterize the diversity of human B cell or T cell repertoires in cases of HIV [84], influenza vaccination [11,12,85], T cell development [86], and in the context of common infections such as cytomegalovirus (CMV) and Epstein–Barr virus (EBV) [87]. Notable discoveries thus far are the observation of limited repertoires in the vaccine response repertoire of aging adults versus younger subjects [88], the presence of memory CD4(+) T cells specific to viral antigens in healthy adults who had never been exposed [154], and also the presence of clonal expansions unrelated to the vaccine responses, which correlate with a latent EBV infection [87]. A similar phenomenon has been seen in aging mice with respect to the TCR repertoire [89].

Also interesting and important is the observation of convergent antibody heavy chain sequences, especially in the CDR3 region, in at least some responses, such as Dengue infection and influenza vaccination [11,90]. These seem to occur in many although not all individuals, and suggest that there might be a useful uniformity in these responses detectable in bulk sequencing data alone.

A relatively new development in the use of NGS is its application on single T or plasmablast cells responding to a vaccine or infection [14,15]. Here, individual cells of the desired type are sorted into separate wells, lysed and the TCR or Ig chains amplified, with or without other genes of interest (such as cytokines or transcription factors in the case of T cells). After the amplification, all products from a given cell are ligated with a unique nucleic acid sequence "barcode" and then all the products are combined and sequenced *en masse*. The barcode enables all the products of each cell to be regrouped informatically, and thus a complete heavy and light chain form a plasmablast, or alpha and beta TCRs from a T cell can be reconstructed and used to determine the specificity and/or affinity of an antibody [91] or the specificity and phenotype of a T cell [15]. These methods allow one to go quickly from an immune response about which little or nothing is known.

A more recent and promising development of NGS is its application to epigenetic profiling. For example, a fast and highly sensitive method was developed based on direct *in vitro* transposition of sequencing adaptors into native chromatin by the bacterial enzyme transposase, which inserts only in regions of open chromatin, thereby generating sequencing-library fragments that can be PCR-amplified. Thus, amplifiable DNA fragments suitable for NGS are preferentially generated at locations of open chromatin allowing whole-genome profiling of active chromatin sites [92]. Although this technology has not yet been used in human immunology studies, it is important to note its great potential use in vaccination and infection settings.

Thus, NGS is rapidly becoming a core technology in vaccine analysis [93], which ultimate goal is to identify the pathways by which pathogens activate protective immune responses and potentially enhance those responses by vaccinating people based on identified genetic signatures predictive of both immunogenicity and safety.

5. Proteins produced prior to (predictive of outcome) or in response to immune perturbation

5.1. Antibody profiling (protein and peptide microarrays)

5.1.1. Protein microarrays

Molecular profiling at the protein level offers a rapid and high-throughput systems approach to scientific discovery compared to traditional single-protein or single-pathway studies. Emerging technologies providing global profiling represent exciting opportunities for the biomedical research community.

Recently, protein microarrays comprised of cell lysates, antigen fractions, or highly purified proteins have been successfully used to profile serum antibody levels in various diseases [94–96]. The discovery of antibodies in different settings of vaccination and infection is one important aim in proteomic systems biology, and is a critical step for enabling personalized medicine.

Construction of protein arrays is similar to DNA microarrays, except that the immobilized species is a protein and the array often represents only partially the entire proteome. Several approaches for tracking immune responses have been undertaken using these technologies. For example, it has been used for the discovery of schistosomiasis vaccine antigens in chronically infected individuals [97]; to identify immunogenic *Plasmodium vivax* proteins in malaria patients [98]; in the identification of profiles that can distinguish *Helicobacter pylori*-seropositive and seronegative cancer patients [99]; to define immunodominant proteins within the *Salmonella typhi* membrane in Bangladeshis infected with acute typhoid [100]; to evaluate serum from autoimmune and immunodeficient patients for antibodies against cytokines, chemokines, and growth factors in systemic lupus erythematous (SLE) [101]; to identify conserved and polymorphic linear B-cell epitopes of *Trypanozoma cruzi* in Chagas disease [102]; and to characterize the antibody profiles elicited by smallpox vaccines [103] with follow-up studies involving the sera from more than 2000 smallpox-vaccinated humans [104].

Protein arrays have been also used to assess host antibody profiles in response to infection with *Brucella melitensis* [105], *Burkholderia pseudomallei* [106], Vaccinia virus [107] and *Coxiella burnetii* [108]. The use of protein microarrays has also aided the identification of different antibody isotype profiles for differentiating acute versus chronic Q-fever [109]. In addition to antibody profiling, protein microarrays can be used to identify antigens for diagnostic tools, and candidate antigens for vaccine development [109].

5.1.2. Peptide microarrays

A peptide array is a collection of peptides presented on a solid surface, usually made of glass or plastic. The assay principle is similar to an ELISA in that a biological sample is incubated and the presence of antibodies recognizing specific epitopes is detected. The peptides are typically linked to the surface of a chip, the shape of a microscope slide. This array is directly incubated with different biological samples like purified antibodies, patient or animal sera, cell lysates and cell supernatants, etc. After washing, a secondary antibody often tagged by a fluorescence label can then be detected by a fluorescence scanner. However, other detection methods such as chemiluminescence, colorimetric or autoradiography also exist.

Clinical applications of peptide arrays are biomarker discovery, profiling of serological responses of patients during infection or vaccination, monitoring clinical interventions, and development of diagnostic tools and vaccines. Peptide arrays were successfully used in the identification of specific epitopes on *Toxoplasma gondii* antigens, leading to an improvement in the serological diagnosis of toxoplasmosis [110]; to study the antibody diversity against linear HIV-1 sequences in HIV-1-infected humans and HIV-1-vaccinated humans [111]; to profile the pre-existing antibody repertoire to the seasonal influenza vaccine thus enabling to distinguish immune from pre-immune samples in young and older donors [6,112,113]; in Valley Fever patients [114]; to identify linear B epitopes of pertactin of *Bordetella pertussis* induced by immunization with whole and acellular vaccine in mice [115]; and to identify strain-specific B-cell epitopes in *T. cruzi* [102]. More recently photolithography

was used to synthesize arrays (Intel arrays) that contained every possible overlapping peptide within a linear protein sequence covering the N-terminal tail of human histone H2B individuals with SLE [116]. Such new methods provide powerful tools for rapid and accurate measurement of broad antibody-based immune responses that may be extremely useful in measuring response to vaccines and infectious agents.

5.2. Cytokines and chemokines (Luminex, MSD)

Cytokines and chemokines are soluble proteins secreted by immune cells that enable passage of information between immune cells and participate in cell activation, cell growth, migration, and differentiation. They are a large part of the grammar of the immune system, that is, the way it talks to itself and orchestrates an immune response or the resolution of a response. Cell–cell interaction is also import, but here as well, the message is often conveyed by particular cytokines.

Several technologies have been developed for multiplexing the analysis of cytokines and chemokines. Probably the most popular ones with the capability to measure several dozens of analytes are the xMap Luminex and the MesoScale Discovery (MSD) platforms. In the xMAP technology, microbeads are color-coded into 500 distinct sets each of which can be coated with a reagent specific for a particular assay. A light source then excites the internal dyes that identify each microsphere particle with many readings made on each bead set, which validate the results. This technology shares common components with general flow cytometry instruments such as lasers, fluidics, and optics. Using this process, the xMAP Technology allows multiplexing of up to 500 unique bioassays within a single sample. In contrast, the MesoScale Discovery (MSD) platform is similar to a multiplexed ELISA. Individual spots on a microtiter plate are coated with capture reagents for the analytes of interest, which are then detected with enzyme-linked detector antibodies. An electrochemiluminescent substrate is then used to create a light signal that is quantitated by the instrument. Advantages of this detection system include high sensitivity and a wide dynamic range, as well as minimal interference from matrix factors.

Studies addressing the changes in cytokine profiles upon immunization or infection are abundant. For example, these technologies have been used to determine the systemic cytokine pattern induced by vaccination with human papillomavirus (HPV) L1 virus-like particles (VLP) [117]; to evaluate the potential of cytokines present in plasma from patients with dengue in stratifying disease severity [118]; as potential biomarkers that differentiate healthy contacts from tuberculosis patients [119]; to investigate the responses to *Mycobacterium tuberculosis* purified protein derivative (PPD) in samples from BCG-vaccinated or -unvaccinated infants [120]; during influenza vaccination of patients undergoing tonsillectomy [121] or to compare cytokine production upon BCG vaccination in Malawian infants compared with UK infants [122]. In one study consisting of 141 healthy infants who had been immunized with hepatitis B vaccine (HBV), single-nucleotide polymorphisms in genes encoding for cytokines and cytokine receptors were investigated for their associations with variations in the immune response to vaccination [123]. Multiplex cytokine assays have been also used in the characterization of the $CD4^+$ T cell response to YF-17D vaccine [124]; to evaluate innate and adaptive immune responses to gut bacteria in HIV patients [125]; to study the effect of CMV in cytokine changes during cycling time trial in athletes [126]; to study the inflammatory response to TIV among pregnant women [127]; to find biomarkers of inflammation associated with mortality and hepatitis flares in persons coinfected with HIV and Hepatitis viruses [128]; and to determine the cytokine response to synthetic double-stranded RNA, as an agonist for toll-like receptor (TLR) 3 [129].

In summary, there are now excellent methods for assaying many cytokines at once in serum samples-up to 63 here in Stanford's Human Immune Monitoring Center, and probably many more in the future. This kind of data particularly combined with cell subset analysis and some of the other technologies described here, begin to capture a good portion of an individuals' immune system, and thus enable comprehensive immune monitoring.

6. Cellular metabolic states predictive of outcome and changes in cell metabolism in response to immune perturbation (metabolomics)

Metabolomics is the systematic measurement of small-molecule metabolite profile within a biological sample, such as urine, plasma or tissue. Surveying these small molecules allows for better understanding of an organism's phenotype and metabolism, and is a central hub for the influences of genes, microbiota and environmental influences.

There are two approaches in metabolomics: targeted metabolomics and untargeted metabolomics or metabolic profiling. In targeted metabolomics defined sets of structurally known and biochemically annotated metabolites are quantified, and is based on a previous understanding of biochemical pathways. One major advantage of targeted metabolomics is that it generally provides information about the molecular concentrations of metabolites involved in a pathway, which facilitates the immediate understanding of deviations from normal. Another advantage is that interpretation of targeted metabolomics data is straightforward. In contrast, untargeted metabolomics is the comprehensive analysis of all the measurable analytes in a sample including chemical unknowns. Many have recognized untargeted metabolomics as having unprecedented value to bringing clarity to complex "omics" data. The reason for this is that metabolites are a proxy to the phenotype and the metabolome is the very end product of the genetic setup of an organism, as well as the sum of all influences it is exposed to, such as nutrition, environmental factors, or treatment. Untargeted metabolomics includes analysis of large arrays of metabolites, thereby extracting biochemical information that reflects functional endpoints of biological events. In addition, metabolomics can serve as a diagnostic source of data for understanding the underpinnings of a particular phenotypic state whether that change is induced by a vaccine or an infectious agent.

In a metabolomics experiment, samples are subjected to methanol extraction and split into aliquots for analysis by ultrahigh performance liquid chromatography/mass spectrometry (UHPLC/MS) in the positive, negative or polar ion mode and by gas chromatography/mass spectrometry (GC/MS). Metabolites are identified by automated comparison of ion features to a reference library of chemical standards followed by visual inspection for quality control.

Metabolomics has been used as a tool for discovery of metabolic signature to distinguish sepsis from systemic inflammatory response syndrome in humans [130]; to profile cellular responses upon infection with Hepatitis C Virus [131]; to find biomarkers patients infected with tuberculosis (TB) [132] as well as to define metabolic networks in TB [133]. In addition, it has demonstrated potential for identification of biomarkers for disease subtype, stage and treatment response, as well as insight into the molecular basis of diseases. At the biochemical level, inflammatory mediators and biomarkers are represented by metabolites including lipids/fatty acids, amino acids, anti-oxidants and nucleotides

that report on key aspects of inflammation including cell trafficking, cell activation, oxidative stress and signaling.

Metabolomics signatures have been identified for prediction of death in septic patients [134]; and important metabolites were discovered in gum infections (periodontal disease) [135]. Metabolomics has been also used to distinguish patients with pneumonia from those suffering from sepsis [136]; and more recently, for the identification of important metabolites for Vaccinia virus replication [137].

The global analysis of soluble metabolites is starting to yield important insights with respect to the variability in the immune response in humans. In addition, these measurements enable the identification of biomarkers that can be used to distinguish patients infected with different pathogens and with various degrees of clinical presentations. Therefore, these technologies provide us with data that are potentially quite useful to understand immunity in humans.

7. From data to function: making sense of the complexity

Generating large amounts of data encompassing multiple "layers" of the immune system is now straightforward and has been greatly facilitated by the creation at Stanford of a Human Immune Monitoring Core facility able to conduct deep immune profiling of human blood and other clinical samples [2]. This facility allows analysis protocols to be standardized and broadly applied across many studies, achieving economies of scale (much as the human genome project drove up sequencing throughput and drove down the cost). It also frees research personnel to pursue specific questions and analyses using the data provided by the facility. In particular, researchers are constantly looking for analytical tools that help them analyze and guide discovery of features that are either important mechanistically for a particular condition or outcome, and/or to identify biomarkers. The identification of predictors for a particular immunological and/or clinical outcome is the most common scenario in systems immunology. These types of associations fall into the category of supervised learning that contrasts with unsupervised learning, in which a suite of methodologies are available to identify hidden structure in data when the outcome is inexistent or unknown.

The most common types of associations are those between immune features (e.g. immune genes, cytokines, cell subsets, metabolites) and standard outcome variable of clinical importance (e.g. infected versus uninfected; responders versus non-responders to vaccination; treated versus untreated condition, etc.), and bioinformatics is key to this process with many hundreds of tools available. Important biological insights are emerging from these systems immunology approaches to vaccination and infection. For example, the role apoptosis [6], lipid metabolism and endocrine factors [29], the microbiome [138] and nutritional sensing pathways [139] on promoting vaccine-induced immunity have been reported.

The steps involved in this discovery adventure involve data processing and standardization, feature or variable selection, and functional annotation (interpretation of results). Below we mostly focus on feature selection since the downstream analyses and biological insights that can be drawn depend upon the variables that are discovered by a given computational model.

In feature selection, relevant parameters (subset of predictors) are discovered using appropriate statistical models. In order to be able to generalize an observation, that is, to find real associations between immune variables and a given outcome that are likely to be found in different data sets – encompassing diverse human populations – regularization methods are often used. In statistics and machine learning, regularization methods are used for model selection, in particular to prevent overfitting by penalizing models with extreme parameter values (see below).

Common questions in systems immunology involve continuous or binary outcomes, which can be solved using multiple linear regression, logistic regression, Cox partial likelihood, and others statistical procedures. To select an appropriate model that best explains the data using a relatively small number of predictors, commonly used statistical procedures involve model selection approaches based on Akaike Information Criteria, Bayesian Information Criteria, etc. Such model selection techniques are necessary since modern data sets require high-dimensional modeling often with many thousands of predictor variables such as in whole-genome gene expression microarrays (see above) and small number of observations and under these circumstances regression methods can overfit. In simple terms, overfitting refers to the propensity of models to capture noise in the data at the expense of a genuine relationship. In such cases, the set of identified predictors will likely fail to generalize to additional data sets. A regularization method imposes mathematical constraints on the regression coefficients in a regression analysis, encouraging "simpler" solutions – i.e. models with a restricted set of predictors. Regularization methods are used for model selection and thus, are often applied in biomarker discovery and in systems approaches to solve immunological questions. Specifically, regularization refers to tuning or selecting the preferred level of model complexity so the model is better at generalizing. To regularize, at least two considerations are required: (1) testing how good the models are at prediction, for example using cross-validation or a validation data set and (2) a tuning parameter which enables changing the complexity of the model. The complexity parameter is adjusted to find the value that gives the best model predictions.

The most common regularization variants are known as $l1$ and $l2$ regularization, which correspond to mathematical constraints imposed on the regression coefficients for each predictor variable such that only the most relevant associations are identified. As mentioned above, the purpose of this shrinkage is to prevent overfitting arising due to either collinearity of the covariates or high-dimensionality. Although both methods are shrinkage methods, the effects of $l1$ and $l2$ penalization are quite different in practice. Applying an $l2$ penalty tends to result in all small but non-zero regression coefficients, whereas applying an $l1$ penalty tends to result in many regression coefficients shrunk exactly to zero and a few other regression coefficients with comparatively little shrinkage. Combining $l1$ and $l2$ penalties tends to give a result in between, with fewer regression coefficients set to zero than in a pure $l1$ setting, and more shrinkage of the other coefficients (see the Elastic net below). The amount of shrinkage is determined by tuning parameters $\lambda 1$ and $\lambda 2$. A value of zero always means no shrinkage and a value of infinity means infinite shrinkage (= setting all regression coefficients to zero). $l1$ penalty can be also seen as setting a Laplacian prior on the regression coefficients – which results in a stringent penalty and sparse solutions. In contrast, $l2$ corresponds to Gaussian prior and results in a more permissive incorporation of features to the model since the regression coefficients are more spread across features (see below).

7.1. *Sparsity via* l1

As mentioned, $l1$ uses the Laplace distribution which is a continuous probability distribution that can be thought of as two exponential distributions spliced together back-to-back (see Fig. 2). A probability distribution links each outcome of a statistical experiment (in this case regression coefficients) with its probability of

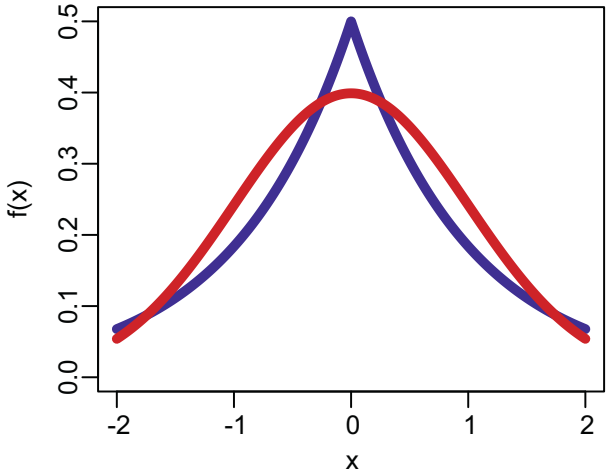

Fig. 2. Laplace versus Gaussian probability distribution. The figure shows the probability (y-axis) of occurrence of a regression coefficient (x-axis) with Laplace prior ($l1$) (blue line) or Gaussian prior ($l2$) (red line). As can be observed, with $l2$/Gaussian as one moves away from zero the probability for such regression coefficient becomes progressively smaller, whereas in $l1$/Laplace the probability of both very small and large coefficients is higher than with $l2$/Gaussian. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

occurrence. For simplicity, let's just consider the one-dimensional case where the $l1$ regularized loss function can be defined as

$$F(x) = f(x) + \lambda \|x\|_1$$

As it can be observed, $l1$ regularization combines a loss function $f(x)$ and a $l1$ penalty; it can be inferred that F would be 0 if λ is large enough, which produces a strong regularization effect.

Practically speaking, if x is a vector, use of $l1$ can lead to some component of optimal x being exactly zero while others may be relatively large (Fig. 2). The function is said to be non-smooth since the first derivative does not exist at $x = 0$. The implication of this is that the $l1$ regularization provides sparse estimates. Namely, in a high dimensional space, one obtains mostly zeros and a small number of non-zero coefficients. This is key since it incorporates variable selection to the modeling problem.

A number of methods have been proposed using $l1$ regularization. For example, the wavelet shrinkage and Basis pursuit [140,141], the Lasso [142], Least Angle Regression (LARS) [143], Component Selection and Smoothing (COSSO) in multivariate nonparametric regression [144], and others. Probably the most commonly used $l1$-based regularization method is the Lasso, which is a regularized version of least squares. The method of least squares is much applied in data fitting and seeks to find a solution that minimizes the sum of the squares of the errors made in the results of a given model. The best fit is the one that minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by a model.

As mentioned above, in Lasso a penalty parameter is applied such that most of the parameters' regression coefficients are driven to zero. This shrinkage procedure effectively culminates in the identification of a relatively small number of relevant predictors, which creates a model with high degree of sparsity. Despite this advantage over other regularization methods, the Lasso suffers from some disadvantages or limitations when applied to immunological questions. For example, (1) if the number of variables (or measurements) (p) is greater than the number of observations (samples) (n), which is typically the case in systems immunology problems, the Lasso selects at most n variables, this is, the number of selected immune features is bounded by the number of samples. This can be a restricting factor if important variables exceed the sample size, and (2) often there exist a degree of significant correlation between the measured variables (genes, cytokines, etc.) most frequently for features sharing the same biological pathway. These variables effectively form a group, and the Lasso fails to select for grouped variables, as it tends to pick one variable from a set of correlated ones and ignore the others. Thus, while being highly efficient and robust for feature selection, when applying the Lasso it is possible that (1) important features are ignored by the selected model and (2) features involving common pathways and functions are not necessarily selected in a joint fashion.

7.2. L2 regularization

$l2$-regularized loss function is defined as:

$$F(x) = f(x) + \lambda \|x\|_2^2$$

In this case, the function is smooth since it can have continuous derivatives (Fig. 2). Practically speaking, $l2$ regularization spreads error throughout the vector x, and jointly shrinks the corresponding coefficients minimally. Least-square penalized by $l2\hat{2}$ penalty is referred to as Ridge regression.

One of the prime differences between Lasso and Ridge regression is that in Ridge regression, as the penalty is increased, all parameters are reduced while still remaining non-zero, while in Lasso, increasing the penalty will cause an increasing number of the parameters to be driven to zero. Thus Lasso effectively deselects the features from the regression. Hence, it automatically selects more relevant features and discards the others, whereas Ridge regression never fully discards any features.

Additional feature selection techniques are developed based on the Lasso including Bolasso which, in addition, performs bootstrap on samples [145] and FeaLect [146] which generates a score based on a combination of bootstrapping procedures and the identification of the best relevance-ordering of the features for each sample. However, an ideal algorithm for efficient feature selection in systems immunology is one in which (1) variable selection is built into the procedure and (2) automatically includes whole groups into the model if one variable amongst them is selected.

7.3. The Elastic net regularization

The Elastic net is a regularized regression method that linearly combines the $l1$ and $l2$ penalties of the Lasso and Ridge methods. As mentioned previously, the Ridge penalty shrinks the coefficients of correlated predictors toward each other while the Lasso tends to pick one of them and discard the others.

The Elastic net regularized function can be defined as:

$$\hat{\beta} = \underset{\beta}{arg\,min}(\left\|y - X\beta\right\|^2 + \lambda_2 \left\|\beta\right\|^2 + \lambda_1 \left\|\beta\right\|_1)$$

where the $l1$ part of the penalty generates a sparse model, and the quadratic part of the penalty: (1) removes the limitation on the number of selected variables, (2) encourages "grouping effect", or the joint selection of correlated predictors and (3) stabilizes the $l1$ regularization path. As observed form the function, the Elastic net penalty mixes both $l1$ and $l2$ penalties.

The glmnet package (R Bioconductor) is an extremely efficient procedure for fitting the entire Lasso or Elastic net regularization path for linear regression, logistic and multinomial regression models, Poisson regression and the Cox model. With a penalty defined as:

$$(1 - \alpha)/2 \left\|\beta\right\|^2 + \alpha \left\|\beta\right\|_1$$

the Elastic net mixing parameter (alpha, α) can be set to 1, which is equal to the Lasso penalty, or set to 0, which is equal to the Ridge penalty.

Here, if predictors are correlated in groups, an $\alpha = 0.5$ will tend to select the groups in or out together (see below). This is a higher-level parameter, and users might pick a value upfront, or experiment with a few different values.

Probably one of the most interesting characteristics of the glmnet package is the ability to separate penalty factors that are applied to each coefficient (or data-type). The penalty factor allows differential shrinkage; it can be set to 0 for a defined set of variables, which implies no shrinkage, or 1 for a different set of variables, which will be penalized assuming $l1$ penalty.

This is extremely helpful in studies where the size for each data type is substantially different. For example, broad and deep immune profiling [6,29,147] often involves measuring genes expressed in blood cells in addition to circulating cytokines, metabolites, etc. (Fig. 1). These data sets greatly differ in size, with large-size data types such as whole-genome gene expression microarrays, which contain some 20,000 features; intermediate-size data types, such as broad-coverage (untargeted) metabolomics consisting in less than 1000 features; and small-size data types such as multiplexed cytokine assays typically consisting in dozens of features. These disparities can affect the feature type that is selected with genes having the highest likelihood and cytokines the least. To account for this and encourage that features from different data types are selected more parsimoniously, glmnet can set differential penalty factors for each data type. For example, an alpha = 0.2 applied for cytokines will promote a Ridge penalty; alpha = 0.5 applied for metabolites will weight Lasso and Ridge penalties to a similar extent, and alpha = 1 can be set for genes, which will apply full Lasso penalty for genes only.

The Elastic net has been used by our group and others in a number of studies in immunology in general, and in vaccination and infection in particular. For example, in the analyses of gene–gene interaction in diabetic patients [148]; for deconvolving transcriptomic data in the context of acute kidney allograft rejection [149]; to identify correlates of protection to anthrax vaccination in macaques [150]; to identify features able to classify microbial communities associated with clinical outcomes in bacterial vaginosis [151,152]; for the identification of factors associated with influenza infections [153] and those predicting influenza vaccine responsiveness [6,113].

8. Concluding remarks

These are exciting times in human immunology and vaccine work where we are able to produce data of unprecedented scale and depth, spanning much of the immune system. This systems immunology approach is starting to provide us with a much more global perspective on the complex molecular and cellular events that occur during an immune response. The technological breakthroughs that have transformed genomics, epigenetics, proteomics and multiparameter flow cytometry have also provided immunologists with new tools with which to examine immunological responses. With the many recent vaccine failures or near failures (HIV, Dengue, TB), there is also the realization that we have unfinished business in the development of vaccines to the most difficult diseases, where long established methods are not working; and so we need a better understanding of vaccines that work, to start with. In this review, in addition to these new technologies, we have also focused on data analysis; specifically regularization and feature selection, which enable researchers to better understand biological data collected from systems approaches. Indeed, mining and interpreting high-throughput data sets remains the major challenge to understanding immunity, especially now that technologies have become robust and fairly inexpensive. Importantly, there exist disproportionate rates with comparison of data generation with data analysis, interpretation and knowledge extraction. This generates a bottleneck since data is produced in a much faster way than can be analyzed. One of the reasons for this is that data is generated almost in a completely automated fashion with little intervention of a rather technical human work; and in contrast, data analysis, interpretation and knowledge extraction is a much slower process since requires thought intervention. In addition, data analysis and interpretation is typically achieved by a team of qualified personnel, this is, individuals with an understanding of immunology, physiology and disease as well as those trained in data mining and bioinformatics. To exemplify this unbalance in the rate of data generation to knowledge extraction, in our experience, multidimensional data that can be collected in a few months usually takes over a year to become coherent information and be ultimately interpreted to generate knowledge. For instance, a research paper reporting results from a system biology study typically filters the many thousands of hypotheses that can be drawn from multidimensional data. Hundreds of solutions arise that are then hierarchized by a set of parameters predefined on an individuals' own set of values and prior knowledge. However, the published report consists of one or two main messages that may or may not be based on prior knowledge. Therefore, enormous quantities of under-analyzed data have accumulated over the last few years, many of which are available in public repositories, such as gene expression omnibus (GEO), ImmPort, ArrayExpress, etc. and thus provide fruitful substrates for subsequent and meta analyses.

More often than not, for any of our complex traits, the phenotype that we observe does not come from a single factor acting alone, but rather from a system of features that are connected to each other and it is the output of the system that produces the phenotype, or trait that we are interested in. Correlating the huge amounts of genotypic and phenotypic data generated by high-throughput methods can create the knowledge needed to speed development of harder and more productive immune systems; how the immune system evolves, how it diversifies, how it interacts with other immune systems and with other entities, and how it responds to environmental changes, so we can learn from better adapted systems as well.

It is a fact that one can test enough different correlations and fluke results will drown out the real discoveries. Propitiously, there are various ways to deal with this problem, which is more serious in large data sets, because there are vastly more possible comparisons than there are data points to compare. Without careful analysis and the application of regularization methods such as the $l1$, $l2$ or the Elastic net, the ratio of genuine patterns to spurious patterns – of signal to noise – quickly tends to zero.

In summary, systems immunology represents a new approach to the analysis of infections and vaccinations that examines how associations between different immune components give rise to the collective behavior of the immune system and how it interacts with its environment.

Conflict of interest statement

All authors declare that no conflicts of interest.

Acknowledgements

We would like to thanks Drs. Rob Tibshirani and Vladimir Jojic for helpful discussions concerning this manuscript. This work was supported by grants from the Ellison Medical Foundation

(AG-SS-1788), the National Institutes of Health (U19s AI057240 and AI090019), and the Howard Hughes Medical Institute.

References

[1] Mestas J, Hughes CC. Of mice and not men: differences between mouse and human immunology. J Immunol 2004;172:2731–8.
[2] Davis MM. A prescription for human immunology. Immunity 2008;29:835–8.
[3] Shen-Orr SS, Furman D. Variability in the immune system: of vaccine responses and immune states. Curr Opin Immunol 2013;25:542–7.
[4] Querec TD, Akondy RS, Lee EK, Cao W, Nakaya HI, Teuwen D, et al. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat Immunol 2009;10:116–25.
[5] Gaucher D, Therrien R, Kettaf N, Angermann BR, Boucher G, Filali-Mouhim A, et al. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. J Exp Med 2008;205:3119–31.
[6] Furman D, Jojic V, Kidd B, Shen-Orr S, Price J, Jarrell J, et al. Apoptosis and other immune biomarkers predict influenza vaccine responsiveness. Mol Syst Biol 2013;9:659.
[7] Davis MM, Krogsgaard M, Huse M, Huppa J, Lillemeier BF, Li QJ. T cells as a self-referential, sensory organ. Annu Rev Immunol 2007;25:681–95.
[8] Irvine DJ, Purbhoo MA, Krogsgaard M, Davis MM. Direct observation of ligand recognition by T cells. Nature 2002;419:845–9.
[9] Purbhoo MA, Irvine DJ, Huppa JB, Davis MM. T cell killing does not require the formation of a stable mature immunological synapse. Nat Immunol 2004;5:524–30.
[10] Huang J, Brameshuber M, Zeng X, Xie J, Li QJ, Chien YH, et al. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. Immunity 2013;39:846–57.
[11] Jackson KJ, Liu Y, Roskin KM, Glanville J, Hoh RA, Seo K, et al. Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. Cell Host Microbe 2014;16:105–14.
[12] Jiang N, He J, Weinstein JA, Penland L, Sasaki S, He XS, et al. Lineage structure of the human antibody repertoire in response to influenza vaccination. Sci Transl Med 2013;5:171ra19.
[13] Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR. Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc Natl Acad Sci U S A 2013;110:13463–8.
[14] Robinson WH. Sequencing the functional antibody repertoire-diagnostic and therapeutic discovery. Nat Rev Rheumatol 2015;11(3):171–82, http://dx.doi.org/10.1038/nrrheum.2014.220.
[15] Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat Biotechnol 2014;32:684–92.
[16] Heller MJ. DNA microarray technology: devices, systems, and applications. Annu Rev Biomed Eng 2002;4:129–53.
[17] Bedognetti D, Wang E, Sertoli MR, Marincola FM. Gene-expression profiling in vaccine therapy and immunotherapy for cancer. Expert Rev Vaccines 2010;9:555–65.
[18] Parnell G, McLean A, Booth D, Huang S, Nalos M, Tang B. Aberrant cell cycle and apoptotic changes characterise severe influenza A infection – a meta-analysis of genomic signatures in circulating leukocytes. PLoS ONE 2011;6:e17186.
[19] Davenport EE, Antrobus RD, Lillie PJ, Gilbert S, Knight JC. Transcriptomic profiling facilitates classification of response to influenza challenge. J Mol Med 2015;93:105–14.
[20] Parnell GP, Tang BM, Nalos M, Armstrong NJ, Huang SJ, Booth DR, et al. Identifying key regulatory genes in the whole blood of septic patients to monitor underlying immune dysfunctions. Shock 2013;40:166–74.
[21] Wynn JL, Cvijanovich NZ, Allen GL, Thomas NJ, Freishtat RJ, Anas N, et al. The influence of developmental age on the early transcriptomic response of children with septic shock. Mol Med 2011;17:1146–56.
[22] Zaas AK, Chen M, Varkey J, Veldman T, Hero 3rd AO, Lucas J, et al. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. Cell Host Microbe 2009;6:207–17.
[23] Ockenhouse CF, Hu WC, Kester KE, Cummings JF, Stewart A, Heppner DG, et al. Common and divergent immune response signaling pathways discovered in peripheral blood mononuclear cell gene expression patterns in presymptomatic and clinically apparent malaria. Infect Immun 2006;74:5561–73.
[24] Kwissa M, Nakaya HI, Onlamoon N, Wrammert J, Villinger F, Perng GC, et al. Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. Cell Host Microbe 2014;16:115–27.
[25] Geiss GK, Bumgarner RE, An MC, Agy MB, van 't Wout AB, Hammersmark E, et al. Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. Virology 2000;266:8–16.
[26] Ockenhouse CF, Bernstein WB, Wang Z, Vahey MT. Functional genomic relationships in HIV-1 disease revealed by gene-expression profiling of primary human peripheral blood mononuclear cells. J Infect Dis 2005;191:2064–74.
[27] Bucasas KL, Franco LM, Shaw CA, Bray MS, Wells JM, Nino D, et al. Early patterns of gene expression correlate with the humoral immune response to influenza vaccination in humans. J Infect Dis 2011;203:921–9.
[28] Henn AD, Wu S, Qiu X, Ruda M, Stover M, Yang H, et al. High-resolution temporal response patterns to influenza vaccine reveal a distinct human plasma cell gene signature. Sci Rep 2013;3:2327.
[29] Furman D, Hejblum BP, Simon N, Jojic V, Dekker CL, Thiebaut R, et al. Systems analysis of sex differences reveals an immunosuppressive role for testosterone in the response to influenza vaccination. Proc Natl Acad Sci U S A 2014;111:869–74.
[30] Nakaya HI, Wrammert J, Lee EK, Racioppi L, Marie-Kunze S, Haining WN, et al. Systems biology of vaccination for seasonal influenza in humans. Nat Immunol 2011;12:786–95.
[31] Li S, Rouphael N, Duraisingham S, Romero-Steiner S, Presnell S, Davis C, et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat Immunol 2014;15:195–204.
[32] Craig FE, Foon KA. Flow cytometric immunophenotyping for hematologic neoplasms. Blood 2008;111:3941–67.
[33] Jaso JM, Wang SA, Jorgensen JL, Lin P. Multi-color flow cytometric immunophenotyping for detection of minimal residual disease in AML: past, present and future. Bone Marrow Transpl 2014;49:1129–38.
[34] Sutherland DR, Anderson L, Keeney M, Nayar R, Chin-Yee I. The ISHAGE guidelines for CD34+ cell determination by flow cytometry. International Society of Hematotherapy and Graft Engineering. J Hematotherapy 1996;5:213–26.
[35] Rebibou JM, Carvalho Bittencourt M, Saint-Hillier Y, Chabod J, Dupont I, Chalopin JM, et al. T-cell flow-cytometry cross-match: influence in renal transplantation. Transpl Proc 1998;30:2800–1.
[36] Strohmeyer JC, Blume C, Meisel C, Doecke WD, Hummel M, Hoeflich C, et al. Standardized immune monitoring for the prediction of infections after cardiopulmonary bypass surgery in risk patients. Cytom B: Clin Cytom 2003;53:54–62.
[37] Barnett D, Walker B, Landay A, Denny TN. CD4 immunophenotyping in HIV infection. Nat Rev Microbiol 2008;6:S7–15.
[38] Martini C, Patroncini S, Mighetto L, Audino A, Girardi P, Giachino O, et al. Cytometry and rhesus: not only for haemolytic disease of the newborn. Blood Transfus 2011;9:104–6.
[39] O'Gorman MR, Zollett J, Bensen N. Flow cytometry assays in primary immunodeficiency diseases. Methods Mol Biol 2011;699:317–35.
[40] Nicholson JK. Use of flow cytometry in the evaluation and diagnosis of primary and secondary immunodeficiency diseases. Arch Pathol Lab Med 1989;113:598–605.
[41] Garratty G, Arndt P. Applications of flow cytofluorometry to transfusion science. Transfusion 1995;35:157–78.
[42] Bolton DL, Roederer M. Flow cytometry and the future of vaccine development. Expert Rev Vaccines 2009;8:779–89.
[43] Spearman P, Lally MA, Elizaga M, Montefiori D, Tomaras GD, McElrath MJ, et al. A trimeric, V2-deleted HIV-1 envelope glycoprotein vaccine elicits potent neutralizing antibodies but limited breadth of neutralization in human volunteers. J Infect Dis 2011;203:1165–73.
[44] De Rosa SC, Thomas EP, Bui J, Huang Y, deCamp A, Morgan C, et al. HIV-DNA priming alters T cell responses to HIV-adenovirus vaccine even when responses to DNA are undetectable. J Immunol 2011;187:3391–401.
[45] Keefer MC, Frey SE, Elizaga M, Metch B, De Rosa SC, Barroso PF, et al. A phase I trial of preventive HIV vaccination with heterologous poxviral-vectors containing matching HIV-1 inserts in healthy HIV-uninfected subjects. Vaccine 2011;29:1948–58.
[46] Davis ID, Chen W, Jackson H, Parente P, Shackleton M, Hopkins W, et al. Recombinant NY-ESO-1 protein with ISCOMATRIX adjuvant induces broad integrated antibody and CD4(+) and CD8(+) T cell responses in humans. Proc Natl Acad Sci U S A 2004;101:10697–702.
[47] Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D, et al. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the step study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet 2008;372:1881–93.
[48] Betts MR, Brenchley JM, Price DA, De Rosa SC, Douek DC, Roederer M, et al. Sensitive and viable identification of antigen-specific CD8+ T cells by a flow cytometric assay for degranulation. J Immunol Methods 2003;281:65–78.
[49] Maecker HT. Multiparameter flow cytometry monitoring of T cell responses. Methods Mol Biol 2009;485:375–91.
[50] Mei HE, Yoshida T, Sime W, Hiepe F, Thiele K, Manz RA, et al. Blood-borne human plasma cells in steady state are derived from mucosal immune responses. Blood 2009;113:2461–9.
[51] Cox RJ, Brokstad KA, Zuckerman MA, Wood JM, Haaheim LR, Oxford JS. An early humoral immune response in peripheral blood following parenteral inactivated influenza vaccination. Vaccine 1994;12:993–9.
[52] Moldoveanu Z, Clements ML, Prince SJ, Murphy BR, Mestecky J. Human immune responses to influenza virus vaccines administered by systemic or mucosal routes. Vaccine 1995;13:1006–12.
[53] Wrammert J, Smith K, Miller J, Langley WA, Kokko K, Larsen C, et al. Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature 2008;453:667–71.
[54] He XS, Sasaki S, Narvaez CF, Zhang C, Liu H, Woo JC, et al. Plasmablast-derived polyclonal antibody response after influenza vaccination. J Immunol Methods 2011;365:67–75.
[55] Halliley JL, Kyu S, Kobie JJ, Walsh EE, Falsey AR, Randall TD, et al. Peak frequencies of circulating human influenza-specific antibody secreting cells correlate with serum antibody response after immunization. Vaccine 2010;28:3582–7.

[56] Odendahl M, Mei H, Hoyer BF, Jacobi AM, Hansen A, Muehlinghaus G, et al. Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. Blood 2005;105:1614–21.

[57] Lee FE, Halliley JL, Walsh EE, Moscatiello AP, Kmush BL, Falsey AR, et al. Circulating human antibody-secreting cells during vaccinations and respiratory viral infections are characterized by high specificity and lack of bystander effect. J Immunol 2011;186:5514–21.

[58] Balakrishnan T, Bela-Ong DB, Toh YX, Flamand M, Devi S, Koh MB, et al. Dengue virus activates polyreactive, natural IgG B cells after primary and secondary infection. PLoS ONE 2011;6:e29430.

[59] Wrammert J, Onlamoon N, Akondy RS, Perng GC, Polsrila K, Chandele A, et al. Rapid and massive virus-specific plasmablast responses during acute dengue virus infection in humans. J Virol 2012;86:2911–8.

[60] Qiao SW, Raki M, Gunnarsen KS, Loset GA, Lundin KE, Sandlie I, et al. Posttranslational modification of gluten shapes TCR usage in celiac disease. J Immunol 2011;187:3064–71.

[61] Han A, Newell EW, Glanville J, Fernandez-Becker N, Khosla C, Chien YH, et al. Dietary gluten triggers concomitant activation of CD4+ and CD8+ alphabeta T cells and gammadelta T cells in celiac disease. Proc Natl Acad Sci U S A 2013;110:13073–8.

[62] Lee FE, Falsey AR, Halliley JL, Sanz I, Walsh EE. Circulating antibody-secreting cells during acute respiratory syncytial virus infection in adults. J Infect Dis 2010;202:1659–66.

[63] Bendall SC, Simonds EF, Qiu P, Amir el AD, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 2011;332:687–96.

[64] Newell EW, Sigal N, Bendall SC, Nolan GP, Davis MM. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. Immunity 2012;36:142–52.

[65] Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat Biotechnol 2012;30:858–67.

[66] Behbehani GK, Bendall SC, Clutter MR, Fantl WJ, Nolan GP. Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry A 2012;81:552–66.

[67] Newell EW, Sigal N, Nair N, Kidd BA, Greenberg HB, Davis MM. Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. Nat Biotechnol 2013;31:623–9.

[68] Horowitz A, Strauss-Albee DM, Leipold M, Kubo J, Nemat-Gorgani N, Dogan OC, et al. Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry. Sci Transl Med 2013;5:208ra145.

[69] Strauss-Albee DM, Horowitz A, Parham P, Blish CA. Coordinated regulation of NK receptor expression in the maturing human immune system. J Immunol 2014;193:4871–9.

[70] Wolchinsky R, Hod-Marco M, Oved K, Shen-Orr SS, Bendall SC, Nolan GP, et al. Antigen-dependent integration of opposing proximal TCR-signaling cascades determines the functional fate of T lymphocytes. J Immunol 2014;192:2109–19.

[71] Swadling L, Capone S, Antrobus RD, Brown A, Richardson R, Newell EW, et al. A human vaccine strategy based on chimpanzee adenoviral and MVA vectors that primes, boosts, and sustains functional HCV-specific T cell memory. Sci Transl Med 2014;6:261ra153.

[72] Sen N, Mukherjee G, Sen A, Bendall SC, Sung P, Nolan GP, et al. Single-cell mass cytometry analysis of human tonsil T cell remodeling by varicella zoster virus. Cell Rep 2014;8:633–45.

[73] O'Gorman WE, Huang H, Wei YL, Davis KL, Leipold MD, Bendall SC, et al. The split virus influenza vaccine rapidly activates immune cells through Fcgamma receptors. Vaccine 2014;32:5989–97.

[74] Krishnaswamy S, Spitzer MH, Mingueneau M, Bendall SC, Litvin O, Stone E, et al. Systems biology. Conditional density-based analysis of T cell signaling in single-cell data. Science 2014;346:1250689.

[75] Gaudilliere B, Fragiadakis GK, Bruggner RV, Nicolau M, Finck R, Tingle M, et al. Clinical recovery from surgery correlates with single-cell immune signatures. Sci Transl Med 2014;6:255ra131.

[76] Fergusson JR, Smith KE, Fleming VM, Rajoriya N, Newell EW, Simmons R, et al. CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. Cell Rep 2014;9:1075–88.

[77] Bendall SC, Davis KL, Amir el AD, Tadmor MD, Simonds EF, Chen TJ, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell 2014;157:714–25.

[78] Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJ, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell 2015;160:37–47.

[79] Yao Y, Liu R, Shin MS, Trentalange M, Allore H, Nassar A, et al. CyTOF supports efficient detection of immune cell subsets from small samples. J Immunol Methods 2014;415:1–5.

[80] Metzker ML. Sequencing technologies – the next generation. Nat Rev Genet 2010;11:31–46.

[81] Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 2012;13:341.

[82] Orru V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell 2013;155:242–56.

[83] Cho P, Gelinas L, Corbett NP, Tebbutt SJ, Turvey SE, Fortuno 3rd ES, et al. Association of common single-nucleotide polymorphisms in innate immune genes with differences in TLR-induced cytokine production in neonates. Genes Immun 2013;14:199–211.

[84] Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science 2011;333:1593–602.

[85] Laserson U, Vigneault F, Gadala-Maria D, Yaari G, Uduman M, Vander Heiden JA, et al. High-resolution antibody dynamics of vaccine-induced immune responses. Proc Natl Acad Sci U S A 2014;111:4928–33.

[86] Sherwood AM, Desmarais C, Livingston RJ, Andriesen J, Haussler M, Carlson CS, et al. Deep sequencing of the human TCRgamma and TCRbeta repertoires suggests that TCRbeta rearranges after alphabeta and gammadelta T cell commitment. Sci Transl Med 2011;3:90ra61.

[87] Wang C, Liu Y, Xu LT, Jackson KJ, Roskin KM, Pham TD, et al. Effects of aging, cytomegalovirus infection, and EBV infection on human B cell repertoires. J Immunol 2014;192:603–11.

[88] Wu YC, Kipling D, Dunn-Walters DK. Age-related changes in human peripheral blood IGH repertoire following vaccination. Front Immunol 2012;3:193.

[89] Shifrut E, Baruch K, Gal H, Ndifon W, Deczkowska A, Schwartz M, et al. CD4(+) T cell-receptor repertoire diversity is compromised in the spleen but not in the bone marrow of aged mice due to private and sporadic clonal expansions. Front Immunol 2013;4:379.

[90] Parameswaran P, Liu Y, Roskin KM, Jackson KK, Dixit VP, Lee JY, et al. Convergent antibody signatures in human dengue. Cell Host Microbe 2013;13:691–700.

[91] Tan YC, Kongpachith S, Blum LK, Ju CH, Lahey LJ, Lu DR, et al. Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. Arthritis Rheumatol 2014;66:2706–15.

[92] Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 2013;10:1213–8.

[93] Poland GA, Kennedy RB, Ovsyannikova IG. Vaccinomics and personalized vaccinology: is science leading us toward a new path of directed vaccine development and discovery? PLoS Pathog 2011;7:e1002344.

[94] Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, Gillespie JW, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene 2001;20:1981–9.

[95] Nam MJ, Madoz-Gurpide J, Wang H, Lescure P, Schmalbach CE, Zhao R, et al. Molecular profiling of the immune response in colon cancer using protein microarrays: occurrence of autoantibodies to ubiquitin C-terminal hydrolase L3. Proteomics 2003;3:2108–15.

[96] Robinson WH, Fontoura P, Lee BJ, de Vegvar HE, Tom J, Pedotti R, et al. Protein microarrays guide tolerizing DNA vaccine treatment of autoimmune encephalomyelitis. Nat Biotechnol 2003;21:1033–9.

[97] Gaze S, Driguez P, Pearson MS, Mendes T, Doolan DL, Trieu A, et al. An immunomics approach to schistosome antigen discovery: antibody signatures of naturally resistant and chronically infected individuals from endemic areas. PLoS Pathog 2014;10:e1004033.

[98] Lu F, Li J, Wang B, Cheng Y, Kong DH, Cui L, et al. Profiling the humoral immune responses to *Plasmodium vivax* infection and identification of candidate immunogenic rhoptry-associated membrane antigen (RAMA). J Proteomics 2014;102:66–82.

[99] Lian G, Wei C, Wang D, Cui M, Wang Z, Liu X, et al. Protein profiling of *Helicobacter pylori*-associated gastric cancer. Am J Pathol 2014;184:1343–54.

[100] Charles RC, Liang L, Khanam F, Sayeed MA, Hung C, Leung DT, et al. Immunoproteomic analysis of antibody in lymphocyte supernatant in patients with typhoid fever in Bangladesh. Clin Vaccine Immunol 2014;21:280–5.

[101] Price JV, Haddon DJ, Kemmer D, Delepine G, Mandelbaum G, Jarrell JA, et al. Protein microarray analysis reveals BAFF-binding autoantibodies in systemic lupus erythematosus. J Clin Investig 2013;123:5135–45.

[102] Mendes TA, Reis Cunha JL, de Almeida Lourdes R, Rodrigues Luiz GF, Lemos LD, dos Santos AR, et al. Identification of strain-specific B-cell epitopes in *Trypanosoma cruzi* using genome-scale epitope prediction and high-throughput immunoscreening with peptide arrays. PLoS Negl Trop Dis 2013;7:e2524.

[103] Davies DH, McCausland MM, Valdez C, Huynh D, Hernandez JE, Mu Y, et al. Vaccinia virus H3L envelope protein is a major target of neutralizing antibodies in humans and elicits protection against lethal challenge in mice. J Virol 2005;79:11724–33.

[104] Hermanson G, Chun S, Felgner J, Tan X, Pablo J, Nakajima-Sasaki R, et al. Measurement of antibody responses to Modified Vaccinia virus Ankara (MVA) and Dryvax((R)) using proteome microarrays and development of recombinant protein ELISAs. Vaccine 2012;30:614–25.

[105] Cannella AP, Lin JC, Liang L, Atluri V, Gotuzzo E, Felgner PL, et al. Serial kinetics of the antibody response against the complete *Brucella melitensis* ORFeome in focal vertebral brucellosis. J Clin Microbiol 2012;50:922–6.

[106] Suwannasaen D, Mahawantung J, Chaowagul W, Limmathurotsakul D, Felgner PL, Davies H, et al. Human immune responses to *Burkholderia pseudomallei* characterized by protein microarray analysis. J Infect Dis 2011;203:1002–11.

[107] Davies DH, Molina DM, Wrammert J, Miller J, Hirst S, Mu Y, et al. Proteome-wide analysis of the serological response to vaccinia and smallpox. Proteomics 2007;7:1678–86.

[108] Xiong X, Wang X, Wen B, Graves S, Stenos J. Potential serodiagnostic markers for Q fever identified in *Coxiella burnetii* by immunoproteomic and protein microarray approaches. BMC Microbiol 2012;12:35.
[109] Vigil A, Chen C, Jain A, Nakajima-Sasaki R, Jasinskas A, Pablo J, et al. Profiling the humoral immune response of acute and chronic Q fever by protein microarray. Mol Cell Proteomics 2011;10. M110 006304.
[110] Maksimov P, Zerweck J, Maksimov A, Hotop A, Gross U, Pleyer U, et al. Peptide microarray analysis of in silico-predicted epitopes for serological diagnosis of *Toxoplasma gondii* infection in humans. Clin Vaccine Immunol 2012;19:865–74.
[111] Stephenson KE, Neubauer GH, Reimer U, Pawlowski N, Knaute T, Zerweck J, et al. Quantification of the epitope diversity of HIV-1-specific binding antibodies by peptide microarrays for global HIV-1 vaccine development. J Immunol Methods 2015;416:105–23, http://dx.doi.org/10.1016/j.jim.2014.11.006.
[112] Legutki JB, Magee DM, Stafford P, Johnston SA. A general method for characterization of humoral immunity induced by a vaccine or infection. Vaccine 2010;28:4529–37.
[113] Price JV, Jarrell JA, Furman D, Kattah NH, Newell E, Dekker CL, et al. Characterization of influenza vaccine immunogenicity using influenza antigen microarrays. PLoS ONE 2013;8:e64555.
[114] Navalkar KA, Johnston SA, Woodbury N, Galgiani JN, Magee DM, Chicacz Z, et al. Application of immunosignatures for diagnosis of valley fever. Clin Vaccine Immunol 2014;21:1169–77.
[115] da Silva FR, Napoleao-Pego P, De-Simone SG. Identification of linear B epitopes of pertactin of *Bordetella pertussis* induced by immunization with whole and acellular vaccine. Vaccine 2014;32:6251–8.
[116] Price JV, Tangsombatvisit S, Xu G, Yu J, Levy D, Baechler EC, et al. On silico peptide microarrays for high-resolution mapping of antibody epitopes and diverse protein–protein interactions. Nat Med 2012;18:1434–40.
[117] Garcia-Pineres A, Hildesheim A, Dodd L, Kemp TJ, Williams M, Harro C, et al. Cytokine and chemokine profiles following vaccination with human papillomavirus type 16 L1 virus-like particles. Clin Vaccine Immunol 2007;14:984–9.
[118] Bozza FA, Cruz OG, Zagne SM, Azeredo EL, Nogueira RM, Assis EF, et al. Multiplex cytokine profile from dengue patients: MIP-1beta and IFN-gamma as predictive factors for severity. BMC Infect Dis 2008;8:86.
[119] Anbarasu D, Raja CP, Raja A. Multiplex analysis of cytokines/chemokines as biomarkers that differentiate healthy contacts from tuberculosis patients in high endemic settings. Cytokine 2013;61:747–54.
[120] Smith SG, Lecher S, Blitz R, Locht C, Dockrell HM. Broad heparin-binding haemagglutinin-specific cytokine and chemokine response in infants following *Mycobacterium bovis* BCG vaccination. Eur J Immunol 2012;42:2511–22.
[121] Eriksson JC, Cox RJ, Szyszko E, Davidsson A, Brokstad KA. Local and systemic cytokine and chemokine responses after parenteral influenza vaccination. Influenza Other Respir Viruses 2007;1:139–46.
[122] Lalor MK, Floyd S, Gorak-Stolinska P, Ben-Smith A, Weir RE, Smith SG, et al. BCG vaccination induces different cytokine profiles following infant BCG vaccination in the UK and Malawi. J Infect Dis 2011;204:1075–85.
[123] Yucesoy B, Johnson VJ, Fluharty K, Kashon ML, Slaven JE, Wilson NW, et al. Influence of cytokine gene variations on immunization to childhood vaccines. Vaccine 2009;27:6991–7.
[124] James EA, LaFond RE, Gates TJ, Mai DT, Malhotra U, Kwok WW. Yellow fever vaccination elicits broad functional CD4+ T cell responses that recognize structural and nonstructural proteins. J Virol 2013;87:12794–804.
[125] Lozupone CA, Li M, Campbell TB, Flores SC, Linderman D, Gebert MJ, et al. Alterations in the gut microbiota associated with HIV-1 infection. Cell Host Microbe 2013;14:329–39.
[126] LaVoy EC, Nieman DC, Henson DA, Shanely RA, Knab AM, Cialdella-Kam L, et al. Latent cytomegalovirus infection and innate immune function following a 75 km cycling time trial. Eur J Appl Physiol 2013;113:2629–35.
[127] Christian LM, Porter K, Karlsson E, Schultz-Cherry S, Iams JD. Serum proinflammatory cytokine responses to influenza virus vaccine among women during pregnancy versus non-pregnancy. Am J Reprod Immunol 2013;70:45–53.
[128] Andrade BB, Hullsiek KH, Boulware DR, Rupert A, French MA, Ruxrungtham K, et al. Biomarkers of inflammation and coagulation are associated with mortality and hepatitis flares in persons coinfected with HIV and hepatitis viruses. J Infect Dis 2013;207:1379–88.
[129] Caskey M, Lefebvre F, Filali-Mouhim A, Cameron MJ, Goulet JP, Haddad EK, et al. Synthetic double-stranded RNA induces innate immune responses similar to a live viral vaccine in humans. J Exp Med 2011;208:2357–66.
[130] Langley RJ, Tipper JL, Bruse S, Baron RM, Tsalik EL, Huntley J, et al. Integrative "omic" analysis of experimental bacteremia identifies a metabolic signature that distinguishes human sepsis from systemic inflammatory response syndromes. Am J Respir Crit Care Med 2014;190:445–55.
[131] Roe B, Kensicki E, Mohney R, Hall WW. Metabolomic profile of hepatitis C virus-infected hepatocytes. PLoS ONE 2011;6:e23641.
[132] Weiner 3rd J, Parida SK, Maertzdorf J, Black GF, Repsilber D, Telaar A, et al. Biomarkers of inflammation, immunosuppression and stress with active disease are revealed by metabolomic profiling of tuberculosis patients. PLoS ONE 2012;7:e40221.
[133] Galagan JE, Minch K, Peterson M, Lyubetskaya A, Azizi E, Sweet L, et al. The *Mycobacterium tuberculosis* regulatory network and hypoxia. Nature 2013;499:178–83.
[134] Langley RJ, Tsalik EL, van Velkinburgh JC, Glickman SW, Rice BJ, Wang C, et al. An integrated clinico-metabolomic model improves prediction of death in sepsis. Sci Transl Med 2013;5:195ra95.
[135] Barnes VM, Kennedy AD, Panagakos F, Devizio W, Trivedi HM, Jonsson T, et al. Global metabolomic analysis of human saliva and plasma from healthy and diabetic subjects, with and without periodontal disease. PLOS ONE 2014;9:e105181.
[136] Seymour CW, Yende S, Scott MJ, Pribis J, Mohney RP, Bell LN, et al. Metabolomics in pneumonia and sepsis: an analysis of the GenIMS cohort study. Intensive Care Med 2013;39:1423–34.
[137] Fontaine KA, Camarda R, Lagunoff M. Vaccinia virus requires glutamine but not glucose for efficient replication. J Virol 2014;88:4366–74.
[138] Oh JZ, Ravindran R, Chassaing B, Carvalho FA, Maddur MS, Bower M, et al. TLR5-mediated sensing of gut microbiota is necessary for antibody responses to seasonal influenza vaccination. Immunity 2014;41:478–92.
[139] Ravindran R, Khan N, Nakaya HI, Li S, Loebbermann J, Maddur MS, et al. Vaccine activation of the nutrient sensor GCN2 in dendritic cells enhances antigen presentation. Science 2014;343:313–7.
[140] Donoho DL, Johnstone IM, Kerkiacharian G, Picard D. Wavelet shrinkage: asymptopia? J R Stat Soc 1995;57:301–69.
[141] Chen SS, Donoho DL, Saunders MA. Atomic decomposition by Basis Pursuit. SIAM Rev 2001;43:129–59.
[142] Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc 1996;58:267–88.
[143] Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. Ann Stat 2004;32:407–99.
[144] Lin Y, Zhang HH. Component selection and smoothing in multivariate nonparametric regression. Ann Stat 2006;34:2272–97.
[145] Bach FR. Bolasso model consistent Lasso estimation through the bootstrap. In: Proceedings of the 25th International Conference on Machine Learning. 2008.
[146] Zare H, Haffari G, Gupta A, Brinkman RR. Scoring relevancy of features based on combinatorial analysis of Lasso with application to lymphoma diagnosis. BMC Genomics 2013;14(Suppl. 1):S14.
[147] Furman D. Sexual dimorphism in immunity: improving our understanding of vaccine immune responses in men. Expert Rev Vaccines 2015;14:461–71.
[148] Han D, Cai X, Wen J, Kenyon NS, Chen Z. From biomarkers to a clue of biology: a computation-aided perspective of immune gene expression profiles in human type 1 diabetes. Front Immunol 2012;3:320.
[149] Shannon CP, Balshaw R, Ng RT, Wilson-McManus JE, Keown P, McMaster R, et al. Two-stage, in silico deconvolution of the lymphocyte compartment of the peripheral whole blood transcriptome in the context of acute kidney allograft rejection. PLOS ONE 2014;9:e95224.
[150] Chen L, Schiffer JM, Dalton S, Sabourin CL, Niemuth NA, Plikaytis BD, et al. Comprehensive analysis and selection of anthrax vaccine adsorbed immune correlates of protection in rhesus macaques. Clin Vaccine Immunol 2014;21:1512–20.
[151] Srinivasan S, Hoffman NG, Morgan MT, Matsen FA, Fiedler TL, Hall RW, et al. Bacterial communities in women with bacterial vaginosis: high resolution phylogenetic analyses reveal relationships of microbiota to clinical criteria. PLoS ONE 2012;7:e37818.
[152] Beck D, Foster JA. Machine learning techniques accurately classify microbial communities by bacterial vaginosis characteristics. PLOS ONE 2014;9:e87830.
[153] Mansiaux Y, Carrat F. Detection of independent associations in a large epidemiologic dataset: a comparison of random forests, boosted regression trees, conventional and penalized logistic regression for identifying independent factors associated with H1N1pdm influenza infections. BMC Med Res Methodol 2014;14:99.
[154] Su LF, Kidd BA, Han A, Kotzin JJ, Davis MM. Virus-specific CD4(+) memory-phenotype T cells are abundant in unexposed adults. Immunity 2013;38:373–83.