This document highlights recent publications that demonstrate the use of Illumina technologies in immunology research. To learn more about the platforms and assays cited, visit www.illumina.com.
INTRODUCTION

Immunology is the field of study concerned with the recognition and disposal of foreign or “non-self” material that enters the body. This material is usually in the form of life-threatening infectious microorganisms or cancer but sometimes, unfortunately, in the shape of life-saving graft transplantation. The body can also be tricked into mobilizing the immune response against itself, to create autoimmune diseases. The NIH estimates that approximately 23.5 million Americans suffer from autoimmune disease and that the prevalence is rising. Recent progress in the treatment of autoimmune diseases amply illustrates the impact that advancements in immunology are having on human health and disease.

Next-generation sequencing technology is proving to be a powerful tool to map the vast repertoire of immune cells that are capable of recognizing the seemingly boundless array of targets. Repertoire sequencing has enabled researchers to identify unique receptor variants found in individuals with susceptibility to hematological malignancies, autoimmune diseases, and allergen response. This approach is rapidly gaining the attention of translational scientists who seek to improve patient care. Hematologists have led the repertoire sequencing effort and have demonstrated the reliability, cost-effectiveness, and medical value of repertoire sequencing in hematopoietic stem cell transplantation.

The major histocompatibility complex (MHC) is a locus that encodes a highly variable repertoire of cell surface proteins that present foreign antigens to T-cells. The encoded repertory of cell-surface molecules enables immune recognition and clearance of foreign agents. Genes within this locus are routinely assessed in matching patients and donors for solid organ transplantation and hematopoietic stem cell transplantation. By comparing variants of these genes between healthy and affected individuals, researchers are now able to elucidate the root causes of disease susceptibility (i.e. hematological, autoimmune, allergies, hypersensitivities, chronic inflammation, infectious diseases).

The development of Illumina’s next generation sequencing provides the quality, throughput and read lengths required by the research community to map the human immune response at high resolution. The emergence of new approaches such as phase-defined sequencing and single-cell sequencing can be expected to accelerate this knowledge base.

Next-generation sequencing has enabled researchers to make a significant impact in these applications (shown in blue). The list on the far right represents a subset of the human health and disease issues that can be addressed with these applications.

**Reviews**


Repertoire Sequencing of Lymphocyte Receptors

The B and T-cell lymphocytes constitute the adaptive branch of the immune system, which is capable of identifying a vast range of antigens. This diverse repertoire of recognition elements is created through unique arrangements of immunoglobulin molecules on B-cell and T-cell receptors. Successful recognition of antigens triggers both an effector immune response as well as a memory response. An effector response includes CD8+ T-cells that eliminate cells with foreign antigens and CD4+ T-cells that differentiate into several different kinds of effector cells, including those that can further activate macrophages, cytotoxic T-cells, and B cells.\textsuperscript{11,12} The B-cell effector response involves plasma cells that secrete antibodies capable of neutralizing or eliminating a foreign agent.\textsuperscript{13} The memory response occurs when B and T-cells are activated by exposure to a foreign antigen.\textsuperscript{14} Activation of these cells results in proliferation and preservation of the specific antigen receptor, such that secondary exposure to the foreign agent results in a robust immune response.\textsuperscript{15}

In comparison to somatic cells, B and T-cell lymphocytes are unique in that their development and maturation are determined by DNA sequences that are not encoded in the germline. Instead, during the maturation process, these cells undergo rearrangement of the variable (V), diversity (D) and joining (J) gene segments in order to create a unique sequence that can encode an exclusive receptor structure in the heavy immunoglobulin chain of B cells, the β chain of αβ T-cell receptors, and the δ chain of γδ T-cell receptors.

Simplified representation of TCR-β VDJ gene recombination resulting in TCR diversity. The TCR-β locus is located on chromosome 7 and is approximately 620 kb in length. Initially one of the two D regions is joined with one of 13 J regions (both randomly selected), followed by joining of the DJ region to one of more than 50 V regions (also randomly selected), yielding a final VDJ region that is approximately 500 bps in length. The mechanism by which gene segments are joined also introduces bp variability, which together with the combinatorial selection of these segments results in TCR diversity. A completely analogous process occurs for the TCR α chain, without the D gene segment included.

VDJ rearrangement in a B–cell generates the variable heavy chain of the immunoglobulin molecule. This immunoglobulin molecule is expressed on the surface of B cells and can also be freely secreted as an antibody.

The primary antibody heavy chain repertoire is created predominantly by the somatic recombination of variable (V), diversity (D) and joining (J) gene segments. Non-templated nucleotides (indicated in red) can also be added. The antigen-binding site of a heavy chain is formed by the juxtaposition of the hypervariable complementarity-determining regions (CDR-H1, H2 and H3) and the framework 3 region (FR3). After productive IgH rearrangement, recombination of the light chain (IgL) ensues, and the heterodimeric pairing of H and L chains forms the complete antibody of the IgM isotype that is expressed on the surface of a newly formed immature B cell.17
In addition to the aforementioned combinatorial diversity, made possible by the rearrangement of V, D, and J gene segments, splice variants contribute to this diversity via template-independent insertion and deletion of nucleotides at the V-D, D-J, and V-J splice junctions. The size of the repertoire is further increased in B-cell receptors by somatic hypermutations (SHM) of B-cell receptor genes during affinity maturation after initial antigen encounter.

This combinatorial mechanism has the potential to generate more than $10^{18}$ unique T-cell receptors in humans and a much more diverse B-cell repertoire. The entire human VDJ region is has been estimated to range from 300 to 400 nucleotides in length, which makes read length a critical parameter in high-throughput sequencing.

Complementarity determining regions (CDRs) are regions within antibodies or T cell receptors that complement an antigen’s shape. Of the three complementarity-determining regions, CDR3 is the most variable locus and the most critical determinant of antigenic specificity. The CDR3 region in the majority of rearranged functional TCR β and immunoglobulin heavy chains has a length ranging from 66 to 90 bps. Therefore, a sequencing depth of $1 \times 10^9$ successfully characterizes the entire B and T-cell repertoire.

Repertoire sequencing has applications in characterizing reconstitution of B and T-cell repertoires after hematopoietic stem cell transplantation, tracking lymphocytes in hematological malignancies, assessing vaccine efficacy, identifying lymphocyte repertoire variants associated with autoimmune diseases, and in identifying lymphocyte receptor variants in cancers such as colorectal cancer.

Reviews


References

The authors identify sequence features in the variable CDR3 region of the rearranged T cell receptor gene that distinguish CD4+ from CD8+ T cells. These features include variable gene usage and CDR3 region length. They estimate that as few as 1000 T cell receptor sequences are needed to accurately estimate the proportion of CD4+ and CD8+ T cells.

Illumina Technology: HiSeq 2000


The authors performed comparative analysis of these TCR repertoires of 3 mothers and 6 children. Thymic selection shapes the initial output of the TCR recombination machinery in both related and unrelated pairs, with minor effect from inherited differences. TCR profiling using characteristic TCR beta CDR3 variants as clonal identifiers also showed that mature T cells, transferred across the placenta during pregnancy, can expand and persist as functional microchimeric clones in their new host.

Illumina Technology: HiSeq 2000


Notes on Experimental Design

The primary challenges in CDR3 sequencing are the accumulation of PCR errors, sequencing errors, and ratio bias. These factors can result in the generation of false TCR diversity (artificial diversity) with the resultant inability to interpret sequence information accurately. For example, generating more sequencing reads may result in the expansion of erroneous sequence variants with one, two, or more mismatches. These amplified errors may be interpreted as evidence for sequence diversity.\textsuperscript{31,32}

To address artificial diversity, previous studies suggest withdrawal of the low-abundance CDR3 variants that differ from the high-abundance variants by a single nucleotide mismatch or blind elimination of low-abundance sequence variants that comprise a total of 4% of all sequencing reads. It has been shown that this approach can result in up to 50\% loss of sequencing reads and an even greater loss on non-Illumina platforms.\textsuperscript{33}

To eliminate PCR and sequencing errors the following recommendations have been made in the literature.\textsuperscript{34}

- From each sequencing read the CDR3 is extracted by aligning each sequence to the set of genomic VDJ segments from the IMGT/GENE-DB database. Low-quality nucleotides for VDJ segments are treated as allowable mismatches.
- Mapping low-quality reads. High-quality sequences at each nucleotide position within CDR3 form “core clonotypes.” These are merged with low-quality sequencing reads that have ≤3 low-quality nucleotides.
- Correcting PCR errors. Given that TCRs do not undergo somatic hypermutation, nucleotide mismatches with the VD, or J segments of CDR3 can only arise from PCR and sequencing errors. Low-abundant core clonotypes are merged with the more abundant (at least 5-fold more abundant) core clonotypes that differ by no more than 3 nucleotides.

mRNA is the preferable starting material for TCR profiling.\textsuperscript{35}

- T-cell contains multiple copies of RNA molecules that encode beta and alpha chains. These copies widen the bottleneck between the sampled T cells and the final TCR amplicon.
- Given that genomic DNA requires that the entire sample be amplified to compute TCR repertoire, this becomes technically challenging when studying sizable populations of T-cells, which would require unreasonably large aliquot volumes.

\textsuperscript{31} Warren R. L., Freeman J. D., Zeng T., Choe G., Munro S., et al. (2011) Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. Genome Res 21: 790-797

\textsuperscript{32} Nguyen P., Ma J., Pei D., Obert C., Cheng C., et al. (2011) Identification of errors introduced during high throughput sequencing of the T cell receptor repertoire. BMC Genomics 12: 106


Single-Cell Repertoire Sequencing

Analysis of the immunoglobulin variable region and T-cell receptor repertoires is of fundamental importance for our understanding of adaptive immunity in health and disease. However, the vast majority of repertoire studies yield data on only one of the two chains of immune receptors and thus cannot provide information about the identity of native receptor pairs encoded by single B or T-cell lymphocytes.37,38 Phage and yeast display technologies,39-41 although efficient for isolation of antigen-specific antibodies, rely on random pairing and do not provide information on the native pairs of chains. Methods which involve growing cultures of lymphocyte clones,42 or sorting of narrow antigen-specific populations of T cells43 or B cells44 are limited by the number of clones that can be identified, as well as by the complexity of biological samples.

New approaches to this problem take advantage of the sensitivity of next-generation sequencing to sequence single cells and identify multiple native TCR chain pairs in a single experiment.45

Previously VH:VL pairing in B-cell repertoire diversity was lost during lysis of B-cell populations. Here the authors employed a method of single-cell mRNA capture, reverse transcription and amplification by emulsion VH:VL linkage RT PCR of these pairings. The linked pairings were sequenced to identify unique antibody clonotypes in healthy peripheral blood IgG+ B-cells, peripheral antigen-specific plasmablasts isolated after tetanus toxoid immunization, and memory B-cell responses following influenza vaccination.


Celiac disease is an intestinal autoimmune disease caused by dietary gluten and gluten-specific CD4+ T-cell responses. Gluten exposure also induces the appearance of activated, gut-homing CD8+ αβ and γδ T cells in peripheral blood. Single-cell T-cell receptor sequence analysis indicates that both of these cell populations have highly focused T-cell receptor repertoires. Such a focused repertoire usually indicates that the induction is driven by an antigen.


The authors propose a single cell-based method to identify native pairs of alpha-beta T cell receptor (TCR) CDR3 chains within emulsion droplets by employing reverse-transcription of alpha and beta chain mRNA, PCR amplification, and subsequent fusion via overlap-extension. This PCR suppression technique resolves the issue of random overlap-extension of gene pairs that may create a high level of noise after the emulsion stage. The authors propose that this methodology can be applied to the identification of native pairs of variable heavy-light antibody chains.

Illumina Technology: MiSeq paired-end sequencing

LYMPHOCYTE DEVELOPMENT

T Cell Development

Multipotent or lymphoid-biased precursors enter the T cell developmental pathway in response to signals from the thymic microenvironment. Studies have shown that Notch, which has been classically associated with embryonic cell development, is an important trigger in T-cell lineage commitment. Notch signaling in the thymus causes hematopoietic precursors to commit to the T cell fate, mobilizes a T-cell gene expression program that prepares the cells for T-cell antigen receptor (TCR), TCR-based repertoire selection, and ultimately, prepares them for functional roles as immune effectors.

Many questions remain regarding the molecular mechanisms of this commitment. For example, precursors entering the thymus display regulatory genes that are either expressed or inducible, however upon commitment these genes are not only repressed but also irreversibly silenced.

Other questions relate to multiple regulatory requirements for successful deployment of the T-cell program. For example, there is a need to elucidate the functional role of additional transcription factors, including E2A, and HEB, TCF-1 and LEF-1, GATA-3, Myb, Runx1, Ikaros, and Gfi1.

Reviews

References


In this study the authors analyzed and compared the gene expression profiles and the genome-wide histone modification marks H3K4me3 (H3 lysine 4 trimethylation) and H3K27me3 (H3 lysine 27 trimethylation) in T-cells at various stages of development. They observed global changes of gene expression and the epigenetic profile for H3K4me3 and H3K27me3 at promoters.

Illumina Technology: Genome Analyzer for ChIP-Seq analysis


The authors mapped the activity of enhancer signatures using H3K4me1-high and p300-high regions in T helper 1 (Th1) and Th2 cells to interrogate active enhancer repertoires. They used RNA-Seq transcriptome profiling in T helper cells to identify the top 100 differentially expressed genes in each subset and did a comparative study between wild-type and STAT-deficient cells to assess STAT-dependency of positively regulated genes. They also used ChIP-Seq to demonstrate that STAT-deficient cells fail to fully recover the chromatin signature of STAT-dependent enhancers. Collectively, these findings suggest that STAT proteins play both a direct and indirect role in molding specialized enhancer architecture.

Illumina Technology: Genome Analyzer IIx for ChIP-Seq and HiSeq 2000 for RNA-Seq for 100 cycles (single read). RNA-Seq libraries were prepared using TruSeq sample prep kit


Notch pathway signaling prompts hematopoietic precursors to become committed to the T cell fate and prepares cells for TCR expression and TCR-based repertoire selection. The authors employed RNA-Seq and ChIP-Seq to identify the dynamic transformations in transcription and epigenetic marking that occur across the genome through five stages of T cell differentiation that span lineage commitment (FLDN1, FLDN2a, FLDN2b, ThyDN3, and ThyDP cells). They reported that the major genome-wide transcriptomic changes leading to T lineage identity occur during transition to the DN2b or DN3; thus, during commitment and β-selection stages. The authors also used ChIP-Seq to enrich DNA associated with three H3 modifications: H3K(9, 14)Ac, H3K4me2, and H3K27me3. They employed these histone-marking patterns on potential cis-regulatory elements to in vivo tracking data of GATA-3 and PU.1 (transcription factors with complementary roles in early T cell development) and demonstrated the functional relevance of these transcription factors at potential sites for different developmental cell subsets.

Illumina Technology: Genome Analyzer for RNA-Seq and ChIP-Seq


B Cell Development

When mature B cells encounter an antigen they undergo a programmed DNA recombination event known as class switch recombination (CSR), which alters the effector function of the antibody molecule. During class switching, one constant region gene (typically Cμ) is replaced with another (either Cy3, Cy1, Cy2b, Cy2a, Cε, or Cα) via the introduction of double strand breaks (DSBs) and subsequent deletion of intervening sequences.51-53

In B lymphocytes, V(D)J recombination, class switch recombination (CSR) and somatic hypermutation (SHM) produce obligate single and double-strand DNA break intermediates that can become substrates for translocations.54,55 These rearrangements could trigger cancer development.56 This is supported by the observation that genetic ablation of the enzymes that create DNA lesions during V(D)J recombination (RAGs) or CSR and SHM (AID) has a significant protective effect on B-cell transformation.50,57

Nuclear architecture is another potential contributor to the incidence of chromosomal translocations.59 Spatial organization of the genome is compartmentalized into chromosome territories as well as transcriptionally active and silent sub nuclear environments.52,62 These compartments are believed to impact the frequency with which genes from different chromosomes interact and recombine.64


An Overview of Publications Featuring Illumina® Technology
Chromatin conformation is one of many mechanisms for regulating gene expression. In developing B cells, the immunoglobulin heavy chain (IgH) locus undergoes a scheduled genomic rearrangement of the V, D, and J gene segments. In this study, an allele-specific chromosome conformation capture sequencing technique (4C-Seq) was applied to unambiguously follow the individual IgH alleles in mature B lymphocytes. The authors found that IgH adopts a lymphoid-specific nuclear location, and in mature B cells the distal VH regions of both IgH alleles position themselves away from active chromatin.

Illumina Technology: Genome Analyzer IIx
for paired-end 4C-seq and ChIP-Seq


Approximately 95% of lymphomas are of B cell origin and many of these are attributed to aberrant rearrangements generated by activation-induced cytidine deaminase (AID) mediated breaks outside of the IgH locus. The authors used chromosome conformation capture-on-chip with massively parallel sequencing (4C-seq) to show that loci that have a significant interaction with IgH are specifically enriched for RNA Pol II, Spt5, H3K4me3, and AID. They employed a domain-centric approach for analyzing the 4C-seq data and found that chromosomal regions, which contact IgH at significant frequency, contain the vast majority (90%) of known sites identified as hotspot AID target genes. This study provides insight into the role of nuclear organization in AID targeting and the maintenance of genomic stability.

Illumina Technology: Genome Analyzer IIx, for single-read 72-cycle run for 4C-seq and Illumina HiSeq 2000 to resequence a Cγ1 library


Reviews

References

Chromatin conformation is one of many mechanisms for regulating gene expression. In developing B cells, the immunoglobulin heavy chain (IgH) locus undergoes a scheduled genomic rearrangement of the V, D, and J gene segments. In this study, an allele-specific chromosome conformation capture sequencing technique (4C-Seq) was applied to unambiguously follow the individual IgH alleles in mature B lymphocytes. The authors found that IgH adopts a lymphoid-specific nuclear location, and in mature B cells the distal VH regions of both IgH alleles position themselves away from active chromatin.

Illumina Technology: Genome Analyzer IIx
for single-read 72-cycle run for 4C-seq and Illumina HiSeq 2000 to resequence a Cγ1 library


Notes on experimental design

4C-seq is the preferred chromosomal conformation capture technique when analyzing the DNA contact profile of individual genomic sites. This assay has been particularly useful for investigating the associations of specific genes with long-range regulatory elements.65,66 (See Chromatin Structure and Rearrangement for more details).

4C is currently limited to the assessment of long-range contacts with larger regions elsewhere on the chromosome (in cis) or on other chromosomes (in trans). For example, local interactions (<50 kb distance) between a gene and its enhancer are not readily detected. Most 4C strategies use restriction enzymes with a 6-nucleotide recognition sequence, which cut once every few kilobases. This creates fragments that are much larger than the average regulatory sequences, which are no larger than several hundred bps. Increased resolution may depend on the use of more selective restriction enzymes that can generate shorter fragments, which can enable detection of de novo local regulatory interactions.

Circular chromatin conformation capture (4-C)67, allows the unbiased detection of all genomic regions that interact with a particular region of interest.68 In this method, DNA-protein complexes are crosslinked using formaldehyde. The sample is fragmented, and the DNA is ligated and digested. The resulting DNA fragments self-circularize, followed by reverse PCR and sequencing. Deep sequencing provides base-pair resolution of ligated fragments.
INNATE IMMUNITY

Innate immunity is the frontline of host defense. It prompts the rapid and local response against pathogens and is also important for the symbiotic partnership between the host and its microbiota. The innate immune system and adaptive immunity comprise the binary-classification of the immune response. Historically the distinction between these two branches of immunity relied on the consensus that innate immunity is nonspecific and lacks memory whereas the adaptive immunity is characterized by specific antigen recognition and subsequent memory response. However, new evidence that demonstrates innate immune features of B cells and T cells and adaptive immune properties of natural killer (NK) cells are now blurring this conventional binary-classification.69

Innate immunity involves the coordinated action of families of receptors, known as pattern-recognition receptors (PRRs) or microbial sensors that respond to a wide range of microorganisms through the detection of specific conserved microbial patterns or molecules.70-73 This innate immune response is activated by specialized sets of receptors found on macrophages, mast cells, dendritic cells, natural killer cells, and polymorphonuclear leukocytes. These receptors include the membrane-bound Toll-like receptors (TLRs) and C-type lectin receptors (CLRs), and the cytosolic RIG-I-like receptors (RLRs), NOD-like receptors (NLRs) and other DNA sensors.74-78 Furthermore, complement recognition molecules are circulating proteins that constitute the humoral (free in serum and body fluids) arm of innate immunity.79 Following ligand binding, receptors induce the activation of distinct signaling pathways that involve effector molecules, such as interferons (IFNs) or antimicrobial peptides (AMPs), which are required for the eradication of pathogens or danger signals.

Reviews

74. Kawai T. and Akira S. (2011) Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 34: 637-650
References
The authors employed single-cell RNA sequencing to investigate heterogeneity in the response of mouse bone-marrow-derived dendritic cells (BMDCs) to lipopolysaccharides. Maturation of BMDCs can occur as a response to pathogen-derived ligands, such as lipopolysaccharides, which results in co-expression of defense cytokines. Here the authors identified a cluster of 137 highly variable, yet co-regulated, antiviral response genes. 100 out of the clusters’ genes were bimodally expressed across single cells despite being strongly expressed at the population level. The authors stimulated and profiled BMDCs from interferon receptor knockout (IfnR^-/-) mice and discovered markedly reduced expression in these cells for both Stat2 and Irf7, as well other measured genes in the cluster. They concluded that interferon signaling is required for the induction of Stat2 and Irf7, which subsequently act to induce variable antiviral cluster genes.
Illumina Technology: HiSeq 2000 for single-cell RNA-Seq with average depth of 27 million read pairs


CANCER AND THE IMMUNE RESPONSE

The development from a normal hematopoietic cell to a cancerous cell involves a multistep process of clonal evolution driven by a series of somatic mutations. These mutations progressively transform the cell from normal growth to a precancerous state and finally a cancerous state, where all checkpoints designed to regulate cell growth have been surmounted.

Induction of malignant transformations appears to involve at least two distinct phases: initiation and promotion. Initiation involves changes in the genome but does not, in itself, lead to malignant transformation. Malignant transformation requires a secondary step, termed promotion. Promotion can occur during the aggressive cell division that follows the initiation phase, and results from the accumulation of new DNA alterations, typically affecting proto-oncogenes, tumor-suppressor genes or apoptotic genes, that result in unregulated cellular growth.

The ability of next-generation sequencing to detect mutations in rare clonal types, or cells, through deep sequencing makes it possible to study the role of immune effector functions in the pathogenesis of hematological malignancies. A notable example has been the influx of reports that implicate autoreactive T-cell clones in the pathogenesis of clonal stem cell disorders such as myelodysplastic syndromes (MDS) and aplastic anemia (AA). These studies have been supported by the widely consolidated understanding that impairment of anti-tumor immunity, which is physiologically mediated by T-cells, can predispose the development of hematological malignancies. Collectively these T-cell repertoire studies and new reports that implicate immunoglobulin heavy chain rearrangements in clonal evolution of acute lymphoblastic leukemia have quickly become one of the most exciting research areas in hematology.

Reviews
Gottgens B. (2012) Genome-scale technology driven advances to research into normal and malignant haematopoiesis. Scientifica (Cairo) 2012: 437966

For more information, see: Cancer and the Immune System.
MICROBIOTA AND THE IMMUNE SYSTEM

Microbiota refers to the extraordinarily large and diverse reservoir of microorganisms that has a co-evolved relationship with the mammalian immune system. These complex microbial communities inhabit the body surfaces of virtually all vertebrates. The immune system plays an essential role in maintaining homeostasis with resident microbial communities to ensure that the mutualistic nature of the host-microbial relationships is sustained. The co-evolution of the vertebrate immune system has therefore been driven by the need to protect the host from pathogens and to foster complex microbial communities for their protective and metabolic benefits. Given that alterations in host-microbiota homeostasis have been implicated in viral infections,\textsuperscript{84,85} autoimmune diseases,\textsuperscript{86} cancer, metabolic diseases, and cardiovascular diseases, this is an exciting opportunity for researchers to examine the interactions between the microbiota and the host-immune response.

Next generation sequencing technologies have enabled researchers to define the construction of these microbiota by operationally defining polymorphisms of bacterial genes; especially those encoding the 16S ribosomal RNA sequences. Sequencing the human microbiome is now enabling researchers to examine the interactions between microbial communities and host immunity. This has illuminated the significant role that the immune system plays in mediating this homeostatic relationship.

The human intestine harbors over 100 trillion microbes, which represent approximately 500 different species of bacteria.

---

The host and microbiota have co-evolved mutually beneficial outcomes and the immune system plays a critical role in preserving homeostasis. The host provides a nutrient-rich environment and protected niches for the microbiota. The microbiota provides the host with vitamins and nutrients as by-products of microbial digestion and protects the host from pathogens. The microbiota enhances the innate and adaptive immune response. Conversely, there is a need for the host to promote a tolerant immune response that enables the microbiota to inhabit the niches of the gut.

This network illustrates the various roles of the gut microbiota in extra-intestinal autoimmune diseases.87

Reviews

An Overview of Publications Featuring Illumina® Technology


References


Changes to bacterial taxa may be associated with disease pathogenesis in inflammatory bowel disease. The authors performed deep sequencing analysis on the variable region 4 (V4) of bacterial 16S rRNA to investigate the mucosal microbiota communities in pinch biopsies from a cohort of ulcerative colitis (UC) patients. This deep sequencing study when combined with data generated by flow cytometry demonstrates that the depletion of Th22 cells, a subset of CD4+ helper T cells that produce IL-22, during active inflammation in UC patients is associated with reduced populations of Clostridiales and increased population of Proteobacteria among other specific alteration of the mucosal microbiota

Illumina Technology: MiSeq of 16S rRNA


Microbial factors, in particular the gut microbiota, are thought to influence susceptibility to type 1 diabetes. In the non-obese diabetic (NOD) mouse model of type 1 diabetes, female mice are significantly more susceptible to disease than males. This difference is not apparent under germ-free conditions, which indicates a role for microbes in type 1 diabetes susceptibility. The authors sequenced bacterial 16S rRNA libraries prepared from cecal contents to identify microbiome composition differences between sexes, through maturation and post-cecal transplantation. Transfer of cecal contents from male NOD mice to female NOD mice prior to disease onset protected against pancreatic islet inflammation, reduced autoantibody production, prevented the onset of diabetes, and was associated with increased testosterone in female mice. The protection conferred by M→F microbiome transfer was attenuated when androgen receptors were blocked. This study suggests that the microbiota may have a regulatory role on sex hormones and may influence autoimmune disease fate in individuals with high genetic risk.

Illumina Technology: MiSeq paired-end sequencing 16S rRNA


Fusarium species can cause papulopustular lesions with abscesses, ulcerations, and even necrosis in immunocompromised and immunocompetent patients. The authors report a case study of a 7 year old girl with treatment-resistant cutaneous fusariosis. They employed exome sequencing and identified a single heterozygous missense mutation in the signal transducer and activator of transcription 1 (STAT1) gene, which is the most likely genetic defect underlying the fusariosis in this patient.

Illumina Technology: HiSeq 2000 exome sequencing

Schwab C., Berry D., Rauch I., Rennisch I., Ramesmayer J., et al. (2014) Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. ISME J


Maslanki T., Tannura K., Mahaffey L., Loughridge A. B., Benninson L., et al. (2012) Commensal bacteria and MAMPs are necessary for stress-induced increases in IL-1beta and IL-18 but not IL-6, IL-10 or MCP-1. PLoS ONE 7: e50636

An Overview of Publications Featuring Illumina® Technology 23
Although both T and B cells use surface receptors to recognize antigens, they accomplish this in two different ways. In contrast to antibodies or B-cell receptors, which can directly recognize antigens, T-cell receptors only recognize antigens that are presented on the surface of antigen presenting cells, such as dendritic cells and macrophages. These antigen peptides reside within the groove of a cell surface protein called the major histocompatibility complex (MHC) molecule.

In humans, the MHC locus is referred to as the Human Leukocyte Antigen (HLA) and encodes a collection of genes that span a contiguous 4 Mb region on the short arm of chromosome 6. Moreover, the extended MHC, termed (xMHC), spans an even larger 7.6 Mb region comprising more than 400 annotated genes and pseudogenes.

This is a map of the human MHC loci. The MHC class I genes are colored red, MHC class II genes are colored blue, and genes in MHC III are colored green. The 6 loci that are outlined encode the peptide binding sites of class I and II MHC molecules. These loci are routinely assessed in matching donors and recipients in hematopoietic cell and solid organ transplantation.
Six of the HLA genes (HLA-A, -B, -C, -DQA1, -DQB1 and –DRB1) are extremely polymorphic and constitute a set of important markers that are routinely employed in matching patients and donors for solid organ transplantation\(^{90}\) and hematopoietic stem cell transplantation.\(^{91,92}\) HLA genes also play an important role in infectious diseases (HIV, Hep C and CMV), autoimmune diseases (diabetes, rheumatoid arthritis, and celiac disease), and drug hypersensitivity.\(^{93}\)

With conventional technologies, only the most polymorphic regions of HLA class I (exons 2 and 3) and II (exon 2), which encode the peptide binding sites, are assessed in clinical settings. Next-generation sequencing provides clinical researchers with the capability to sequence the entire gene, resulting in phase-resolved, unambiguous HLA typing. Clinical studies have shown that matching these regions of the 6 major HLA loci provides the best clinical outcomes with decreased incidence of rejection and graft versus host disease (GVHD) in solid organ and hematopoietic stem cell transplantation. However, even when these regions are matched, approximately 30% of recipients experience adverse events within 5 years.\(^{94}\)

The source of these imperfect matches is unknown, but there are several possibilities. Adverse events may reflect mismatches in regions that lie outside of the regions that are currently analyzed. Given the high degree of polymorphisms, ambiguous combination of alleles may arise during HLA typing. These may result from cis/trans ambiguities or due to a particular allele combination being identical over the regions commonly analyzed. For example, in conventional sequencing both heterozygous alleles are coamplified and sequenced. Combination ambiguity can occur when two or more alleles share identical sequences in the targeted exons but exhibit differences in non-sequenced exons.\(^{95}\) Next-generation sequencing provides phase information, which may significantly improve the analysis of the HLA cohort (See Phase-Defined HLA Sequencing for more details).

Reviews

de Bakker P. I. and Raychaudhuri S. (2012) Interrogating the major histocompatibility complex with high-throughput genomics. Hum Mol


---
References


The authors propose a high-throughput HLA genotyping method that employs the use of a single long range PCR to amplify the genomic DNA spanning the majority of the coding regions of four polymorphic HLA genes (HLA-A, -B, -C, and -DRB1). This extensive coverage method enhances allelic resolution, which reduces combination ambiguity that results from off-phase heterozygous sequences. Their coverage of non-polymorphic regions increases the chance of identifying previously undescribed alleles with mismatches, insertions, and deletions.

Illumina Technology: MiSeq and HiSeq 2000 for 150 and 100 bp paired-end sequencing. Used Genome Analyzer IIx to sequence 150 bases from both ends


Phase-Defined HLA Sequencing

Phase-defined sequencing indicates which of the two parental chromosomes a particular allele is derived from. Paired-end sequencing inherently provides information for phasing. When a read encompasses two or more heterozygous genotypes of an individual, the phase of the heterozygous genotypes is determined since each fragment from which a read or pair of reads is obtained in a single allele. Therefore, if read lengths have sufficiently high coverage a substantial amount of phase information can be obtained. This has vast implications in understanding the interplay of genetic variation and disease, imputing untyped genetic variation, calling genotypes in sequence data, detecting genotype error, inferring human demographic history, inferring points of recombination, detecting recurrent mutation, signatures of selection, and modeling cis-regulation of gene expression.

Phase-Defined HLA Sequencing

Combination ambiguity occurs when a non-phased consensus sequence is generated. When a single consensus sequence is generated this conceals which of the parental chromosome variants are derived from. Phasing analysis enables researchers to generate two identifiable sequences that correspond to both parental chromosomes. This resolves combination ambiguity and enables researchers to identify which of the two parental chromosomes variants are derived from.
Although there has been a strong emergence of next generation sequencing efforts in HLA-genotyping, these comprehensive analyses omit non-coding HLA regions and mRNA-spliced data,\(^{109-111}\) which may have an impact on gene regulation.\(^ {112,113}\) Moreover, allele determination is conventionally based on sequence alignment to the reference library of HLA sequences in the IMGT/HLA database,\(^ {114}\) which prevents the identification of novel phase-defined HLA gene haplotypes.

**Reviews**


**References**


This is the first study to report a complete sequence of the HLA region. Here the authors were able to determine the phase-defined entire HLA gene sequences, regardless of whether the alleles were rare or novel. They sequenced long-range PCR products of HLA genes spanning from promoter to 3’-UTRs and employed a gene-tagging method to generate two HLA gene haplotype sequences based on phase-defined SNVs. Paired end reads of 2 x 250 bps allowed them to demonstrate phase-defined allele determination for 33 HLA homozygous samples, 11 HLA heterozygous samples, and 3 parent-child families.

Illumina Technology: MiSeq 2 x 250 bp and Nextera DNA Sample Prep Kit for library construction


\(^{113}\) Thomas R., Apps R., Qi Y., Cao X., Male V., et al. (2009) HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet 41: 1290-1294

\(^{114}\) Hosomichi K., Jinam T. A., Mitsunaga S., Nakaoka H. and Inoue I. (2013) Phase-defined complete sequencing of the HLA genes by next-generation sequencing. BMC Genomics 14: 355
**SELF VS NON-SELF AND ANTIGEN DISCRIMINATION**

**Tolerance**

Tolerance refers to the many layers of protection imposed by the immune system to prevent the reaction of its cells and antibodies against host components. An important form of tolerance is self-tolerance, which refers to the lack of response of the immune system to self-antigens.

Recent studies report a more active role of immune cells in the selective inhibition of responses to self-antigens. For example, the study of regulatory T cells (TREG), which in fact recognize self-proteins, have revolutionized the field of tolerance and autoimmunity, not to mention transplantation.

Linked suppression represents a way in which regulatory T cells (TREG) support local self-tolerance. TREG cells inhibit antigen-presenting cells (APCs) presenting their cognate antigen. They can also inhibit bystander T cells, of the same and different antigen specificity, through soluble inhibitory factors.
Central tolerance occurs in the primary lymphoid organs: the bone marrow for B cells and the thymus for T cells. In the first step of this process, T or B-cell clones that recognize self-antigens with high affinity are not allowed to mature. Peripheral tolerance is a secondary precaution in the event that some self-reactive lymphocytes do find their way into the periphery and secondary lymphoid tissues. The peripheral tolerance will render some self-reactive lymphocytes in secondary lymphoid tissues inactive and generates others that actively inhibit immune responses against self. Furthermore, induced cell death, or apoptosis adds a further protective measure by limiting the lifespan of activated lymphocytes.

T cell anergy has been characterized as a hyporesponsive state, or unresponsiveness to an antigenic stimulus, induced by TCR engagement in the absence of costimulation.\textsuperscript{115,116} Conversely, when the same antigen is presented with appropriate costimulatory molecules it can become a potent immunogen and mount an immunologic response. Indirect evidence suggests that T cell dysfunction in the tumor microenvironment and establishment of transplant tolerance is partially attributed to T cell anergy.\textsuperscript{117} Despite the advances in the characterization of T cell anergy, there are gaps in our knowledge base of the anergic phenotype. This is due to the lack of surface markers that might be useful in identifying anergic T cells. Furthermore, it is unclear teleologically why T cells that are subjected to anergy-inducing conditions are not deleted from the repertoire, in order to eliminate T cells of undesired specificities.

Reviews


References


T-cell anergy contributes to peripheral tolerance and plays a role in tumor growth and in promoting transplant allograft acceptance. Egr2 is a critical transcriptional regulator of T-cell anergy. To identify the direct transcriptional targets of Egr2, the authors performed a ChIP-Seq analysis of anti-Egr2 Ab immunoprecipitated nuclear extracts derived from untreated and anergized Th1 T cells. By merging this data with gene expression profiling analyses, the authors revealed 49 targets that are directly regulated by Egr2. They unexpectedly identified cell surface molecules and secreted factors, including lymphocyte-activation gene 3 (Lag3), Class-I-MHC-restricted T-cell associated molecule (Crtam), Semaphorin 7A (Sema7A), and chemokine CCL1. These results indicate that the anergic state may have a functional role through interactions with other immune cells.

Illumina Technology: Genome Analyzer\textsuperscript{116} for ChIP-Seq of DNA fragments ranging from 200 and 400 bp


Genetic polymorphisms within a single locus encoding the transcription factor BACH2 are associated with numerous autoimmune and allergic diseases. The authors use mice that had the BACH2 gene disrupted. They find that BACH2 is a regulator of immune activation.

Illumina Technology: HiSeq for MiRNA sequencing and ChIP-Seq


Autoimmunity

Autoimmunity is caused from the failure of tolerance to protect the host from autoreactive T or B-cell clones. The pathogenesis of these diseases manifests itself in the destruction of proteins, cells, and organs by self-reactive lymphocytes. The onset and pathogenesis of autoimmunity not only depends on intrinsic factors of T and B-cells, such as germline or somatic mutations\textsuperscript{118-120} but also on environmental factors such as microbiota or infections,\textsuperscript{121} the cytokine milieu, and the presence of other immune cells in the microenvironment.\textsuperscript{122}

For most chronic autoimmune and inflammatory diseases, patient populations are heterogeneous and do not uniformly respond to a given therapy. As a result the therapeutic decisions for most autoimmune and inflammatory diseases are based mainly on trial-and-error observations. The development of "actionable biomarkers" may potentially improve the design of clinical trials and inform treatment decisions.\textsuperscript{123}

For example high-throughput DNA sequencing facilitates the tacking of disease-associated clones of T and B-cells in autoimmune diseases. Furthermore, changes in these cell populations can be correlated with a patient's response to therapies.\textsuperscript{124}

The greatest hope in treating these diseases lie in a greater understanding of the functional roles of genetic and epigenetic variants in autoimmune pathogenesis.\textsuperscript{125}

Next-generation sequencing of whole-exomes and whole-genomes has become an essential tool in identifying rare genetic variants in large cohorts of autoimmune disease patients. In addition, recent advances in the sequencing of epigenetic markers are adding more information to elucidate the subtle interplay among epigenetic modifications, genetic factors, and environmental signals that predispose individuals to autoimmune risk.

Reviews


References


Scleroderma is an autoimmune connective tissue disease in which patients make antibodies to a limited group of autoantigens. Patients with scleroderma and antibodies against RPC1 are at increased risk for cancer. The authors sequenced the tumor and normal coding sequences of the POLR3A, TOP1, and CENPB genes in 16 patients. The results suggest that POLR3A mutations triggered cellular immunity and cross-reactive humoral immune responses.

Illumina Technology: Genome Analyzer IIx


The authors utilize exome-sequencing analysis to identify rare and novel variants in known inflammatory bowel disease (IBD) susceptibility genes. Of a panel of 169 known IBD susceptibility genes, approximately 300 non-synonymous, truncating and frameshift mutations were identified from eight pediatric IBD patients. After excluding HLA variants, they uncovered 58 variants across 39 genes, of which 17 were not previously reported. Of the cohort, both patients with severe ulcerative colitis (UC) displayed a distinct profile; both carried potentially deleterious unique variation in the B-cell regulatory gene BACH2 and IL10 genes, which was not seen in the other IBD patients. Variation in BACH2 has not been reported in GWAS of UC.

Illumina Technology: HiSeq 2000 for exome sequencing


DNA methylation changes have been implicated in T cell differentiation in systemic lupus erythematosus patients. The authors performed a genome-wide DNA methylation study and gene expression profiling in naïve CD4+ T cells from lupus patients and controls. Among 86 CG sites that are differentially methylated in naïve CD4+ T cells from lupus patients, they revealed that 21 out of 35 hypo methylated genes are regulated by type-1 interferon - including IFIT1, IFIT3, MX1, STAT1, IFI44L, USP18, TRIM22 and BST2. These results indicate that abnormal DNA methylation exists in lupus T cells prior to activation and differentiation and provide an epigenetic explanation for hyper-responsiveness to type-1 interferon in lupus T cells.

Illumina Technology: Infinium Human-Methylation450 BeadChip array for DNA methylation studies and HumanHT-12 v4 Expression BeadChip array for gene expression studies. DNA methylation analysis was performed with GenomeStudio methylation analysis package


Celiac disease is an intestinal autoimmune disease caused by dietary gluten and gluten-specific CD4+ T cell responses. Gluten exposure also induces the appearance of activated, gut-homing CD8+ αβ and γδ T cells in peripheral blood. Single-cell T cell receptor sequence analysis indicates that both of these cell populations have highly focused T cell receptor repertoires. Such a focused repertoire usually indicates that the induction is driven by an antigen.

Illumina Technology: MiSeq paired-end sequencing


BACH2 is expressed in B cells where it acts as a transcriptional repressor of Blimp-1 and other class switch recombination genes. Polymorphisms within a single locus encoding the transcription factor BACH2 are associated with numerous autoimmune and allergic diseases. By studying mice in which the BACH2 gene was disrupted, the authors found that BACH2 is a key regulator of CD4+ T-cell differentiation that prevents inflammatory disease by controlling the balance between tolerance and immunity. They stimulated BACH2 knockout murine naïve CD4+ T cells, and subsequently performed massively parallel RNA sequencing to show that the majority of differentially expressed genes were unregulated in BACH2-deficient cells. The authors measured genome-wide BACH2 binding in iTreg cells by chromatin immunoprecipitation with massively parallel sequencing. They determined that BACH2 bound 43.6% of all derepressed genes, including 408 derepressed effector lineage-associated genes.

Illumina Technology: HiSeq 2000 and used TruSeq Sample Prep Kit to prepare RNA-Seq libraries

Sjögren’s syndrome is a common autoimmune disease that typically presents as inflammation of the cornea (keratoconjunctivitis sicca) and dry mouth syndrome (xerostomia). In this publication the authors performed a genome-wide association study using Illumina Omni1 Quad array, Illumina ImmunoChip and performed gene expression profiling on the Illumina Human WG-6 v.3.0 BeadChip. Using bioinformatics tools, the combination of genome-wide significance thresholding and suggestively associated variants provided evidence of direct and indirect protein-protein interaction and enrichment of genes involved in immune signaling processes including: TNFAIP3, PTG1, PRDM1, DGKQ, FCGR2A, IRAK1BP1, ITS1N2 and PHIP, among others.

Illumina Technology: HumanOmni1-Quad, Human ImmunoChip


In this meta-analysis of two genome-wide association studies (GWAS) the authors searched for common genetic susceptibility loci for systemic sclerosis (SSc) and lupus erythematosus (SLE). Both diseases are archetypical autoimmune diseases and previous studies have shown that several autoimmune diseases have a common genetic basis. In this study of a total of ~21,000 samples, one new associated locus was identified and two previously described SLE loci were found to be shared with SSc.

Illumina Technology: HumanCNV370-Duo, HumanHap550, Human610-Quad, Human Gene Expression - BeadArray


Solid Organ Transplantation

Graft rejection in solid organ transplantation is attributed to histoincompatible tissues. One type of transplanted tissue is an allograft, which is transferred between genetically different members of the same species. Because an allograft is genetically dissimilar to the host and therefore expresses unique antigens, these are often not recognized as self-antigens by the immune system and result in graft rejection.

Tissues that share sufficient antigenic similarity, allowing transfer without immunologic rejection, are said to be histoincompatible, as is the case when the transfer occurs between identical twins. Most transplants are conducted between individuals with a matching ABO blood group and HLA matching. However, even when MHC antigens are identical, the transplanted tissue can be rejected because of differences at various other loci, including the minor histocompatibility locus.

Currently, for heart transplant recipients, the endomyocardial biopsy (EMB) has been employed as the ‘gold standard’ for rejection surveillance. However, the endomyocardial biopsy is an expensive and invasive procedure that is limited by sampling error, interobserver variability in grading, late detection of rejection, and risk of morbidity. Therefore, there has been a considerable effort to develop noninvasive techniques that might replace or reduce the need for EMB, with much focus placed on monitoring the recipient’s immune response to detect the onset of rejection.

Reviews


References

The authors employed transcriptome sequencing of peripheral blood mononuclear cells (PBMCs) derived from 6 quiescent and 6 severe rejection heart transplant recipients. Through digital gene expression (DGE) profiling, a measurement of expression based on the number of reads of the same or similar sequences, they identified a 10-gene PBMC signature capable of distinguishing patients with acute cardiac allograft rejection. Based on a protein-protein interaction network analysis, the authors indicate that CXCR4 and HLA-A are the most informative genes based on a higher degree of control over information flowing to the other 10 genes in the cooperative network.

Illumina Technology: Genome Analyzer IIx for 85 bp reads RNA-Seq


The human leukocyte antigen (HLA) region, the 3.8-Mb segment of the human genome at 6p21, has been associated with more than 100 different diseases, mostly autoimmune diseases. Due to the complex nature of HLA genes, there are difficulties in elucidating complete HLA gene sequences especially HLA gene haplotype structures by the conventional sequencing method. This study presents a new method for... read more cost-effective phase-defined complete sequencing of HLA genes using indexed multiplexed samples on Illumina MiSeq. The method was demonstrated on 53 samples showing high resolution for HLA typing.

Illumina Technology: MiSeq, Nextera DNA Sample Prep


Due to increased cell death in the organ during graft rejection, increased donor molecules are expected to be present in the blood at these times. Here the authors genotyped the donor and recipient to establish a unique donor “genetic fingerprint,” which was subsequently detected by high-throughput sequencing of the cell-free DNA in peripheral blood of heart transplant recipients. Reads with donor and recipient SNP calls were identified to determine a % Donor DNA. This study establishes a mean value below 1% as indicative of a healthy normal level of donor-derived cell-free DNA. In contrast, during organ rejection the level of donor DNA signal rises to a mean value ranging from 3-4% of the total cell-free DNA.

Illumina Technology: Genome Analyzer IIx and Omni1-Quad Beadchip
INFECTIONOUS DISEASES AND VACCINES

New technological advances in T cell isolation and T receptor sequencing have enabled greater understanding of the basic structure of immune T cell repertoires, the diversity of responses within and between individuals, and temporal changes in repertoires and in response to infectious conditions.

Reviews

Viral Infections

Viral infections such as HIV, a retrovirus, are able to perturb and alter gene expression through several mechanisms. Studies have profiled the expression of cellular miRNA and some snRNA post HIV infection using next generation sequencing. Emerging studies have focused on the novel mechanisms of gene expression regulation, central to recently discovered players between HIV and the immune system. For example, the human leukocyte antigen (HLA) family of proteins plays a key role in retroviral progression because it is a crucial modulator of the immune response.

Ultimately, understanding how immune cells, such as naive virus-specific CD8 T cells, influence the type of immune response generated after virus infections is critical to the development of enhanced therapeutic and vaccination strategies to exploit CD8+ T cell-mediated immunity.

Reviews

References


In this paper the authors investigated the impact of HIV-1 infection on the intestinal microbiome and its association with mucosal T-cell and dendritic cell (DC) frequency and activation, as well as with levels of systemic T-cell activation, inflammation, and microbial translocation. They found that HIV-1-related change in the microbiome that was associated with increased mucosal cellular immune activation, microbial translocation, and blood T-cell activation.

Illumina Technology: MiSeq with 250 bp paired-end kit


The IFNL3 genotype predicts the clearance of hepatitis C virus (HCV), spontaneously and with interferon (IFN)-based therapy. The authors identified an association between a cluster of ISGs, the metallothioneins (MTs) and IFNL3 genotype. They found that metallothioneins (MTs) were significantly upregulated (in contrast to most other ISGs) in HCV-infected liver biopsies of IFNL3 genotype rs8099917 responders.

Illumina Technology: HiSeq 2000 TruSeq RNA sample preparation and Human HT-12 V3


The authors studied miRNA expression in 158 cervical specimens, including 38 normal, 52 cervical intraepithelial neoplasia (CIN), and 68 cervical cancer (CC) tissues. They found an increase of miR-25, miR-92a, and miR-378 expression with lesion progression but no obvious change of miR-22, miR-29a, and miR-100 among the HPV-infected tissues. An expression ratio ≥1.5 of miR-25/92a group over miR-22/29a group could serve as a cutoff value to distinguish normal cervix from CIN and from CIN to CC.

Illumina Technology: HiSeq 2000


The authors investigated the effects of HIV infection on small RNA expression in CD4-expressing T lymphoblastoid cells at 5, 12, and 24 h post infection (hpi). The authors focused on the host response at the level of a single infected cell type and profiled this system over time to detect a phased pattern of microRNA expression. Small RNA-Seq identified 14 differentially expressed microRNA at 5 and 12 hpi; many of which displayed initial suppressed expression followed by rebound later by 24 hpi. They also identified a novel microRNA, an 18-mer encoded in the first intron of the EPB41L2 gene, which was highly expressed in uninfected cell and down regulated by 90% at 24 hpi.

Illumina Technology: Genome Analyzer IIx for RNA-Seq of 54 bp reads. Small RNA libraries were prepared with a small RNA version 1.5 sample preparation kit


The question of how HIV-1 interfaces with cellular microRNA (miRNA) biogenesis and effector mechanisms has been highly controversial. In this paper, the authors used Illumina HiSeq 2000 for deep sequencing of small RNAs in two different infected cell lines and two types of primary human cells to unequivocally demonstrate that HIV-1 does not encode any viral miRNAs.

Illumina Technology: HiSeq 2000 with TruSeq RNA kit deep sequencing of small RNAs and PAR-CLIP to find miRNA binding sites in the HIV-1 genome


Vaccine Development

The new generation of sequencing technology holds tremendous promise in the areas of systems biology and vaccinomics for developing a deeper understanding of the host response to both vaccines and viral infections.

Studies, which report differential gene expression patterns between high and low responders to vaccines, provide insight into the divergent immunoregulatory processes between high and low responders. Further investigation of these loci may lead to important findings regarding the genetic control of immune responses, which can inform the engineering of new vaccine candidates.

Human monoclonal antibodies have a high potential to serve as potential therapeutic tools. Until recently, single antibodies capable of neutralizing a broad array of evolving viruses, such as influenza or HIV, were considered extremely rare and nearly impossible to isolate. By employing high-throughput technologies, careful screening processes and clever selection of infected donors, researchers are now able to isolate and characterize these broadly neutralizing antibodies. There is now a strong effort to preferentially target the epitopes of these antibodies.

Reviews


References


There is a known association between pre-existing antibodies and poor vaccination response, which has been attributed to pre-existing flu-specific memory CD4+ T cells that inhibit antigen-presentation by dendritic cells, and subsequently suppress B-cell response. The authors used whole-genome DNA microarrays to assess the baseline levels of immune parameters that correlate with the hemagglutinin inhibition titer response in influenza vaccination. They reported 109 gene modules, sets of coexpressed genes to which the same set of transcription factors binds. Nine variables could predict the antibody response with 84% accuracy. This is the first study to report an association between apoptosis of reactive memory cells and robust antibody response to a vaccine.

Illumina Technology: HumanHT-12v3 Expression BeadChip and GenomeStudio software


Vaccinia virus (VACV) is an immunologically cross-protective virus that used in the smallpox vaccine. The authors used mRNA-Seq transcriptome profiling to identify host and viral gene expression patterns in peripheral blood mononuclear cells (PBMCs) from smallpox vaccine recipients after VACV stimulation. Of the over 1200 genes that exhibited differential gene expression, they identified a number of chemokine, cytokines, interferon and macrophage-associated genes with significant down regulation upon vaccinia infection. Conversely, they identified upregulation in genes encoding histone, IFNβ, IFNγ and heat-shock proteins.

Gene set analysis revealed that genes with lowest expression values were expressed ‘late’ in the viral life cycle, whereas genes classified as ‘early’ were expressed at significantly higher levels. The patient cohort of high and low vaccinia-specific responders enabled the identification of differential gene regulation patterns between robust humoral immunity and weaker humoral immune responses.

Illumina Technology: Genome Analyzer IIx and Single Read Cluster Generation Kit (v2) and 50 cycle Sequencing Kit (v3). cDNA libraries were created using mRNA-Seq 8 sample prep kit


miRNA and noncoding RNAs

Only a small fraction of the transcriptome is translated, leaving most of the transcriptional output as ncRNAs, which is classified into two broad categories: small and long RNAs. MicroRNA (miRNAs) is a sub-class of the small noncoding RNA (ncRNA) family, which are small endogenously expressed molecules that regulate the expression of proteins encoded by their mRNA targets. miRNAs have been associated with central roles in growth, development, and immune response \textit{in vivo}.\textsuperscript{139-141} They primarily target gene expression at the post-transcriptional level by adjoining to the RNA-induced silencing complex (RISC), which targets the 3’-untranslated region (3’-UTR) of complementary mRNAs and results in the transcript’s repression or degradation.\textsuperscript{142,143}

Recent studies have shown that miRNAs have unique expression profiles in cells of the innate and adaptive immune systems, CNS, and cancers.\textsuperscript{146-149} Furthermore, new evidence implicates a central role of miRNAs in altering mRNA expression in HIV-target cells in response to viral replication.\textsuperscript{150} Improvements in high-throughput sequencing technologies, with respect to depth and sensitivity, are enabling researchers to profile known and novel miRNAs, and identify their exact sequence and length, which provides insights on RNA editing processes and mutational events.\textsuperscript{151} This allows researchers to decode the networks of non-coding RNA control in the development of the adaptive and innate immune systems and their functional response.

\textit{“Wherever the requirement for miRNAs has been tested in the immune system, essential roles have been found”}
Ansel et al. 2013

\textsuperscript{145} Lewis B. P., Burge C. B. and Bartel D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120: 15-201
Chromatin isolation by RNA purification (ChIRP-Seq) is a protocol to detect the locations on the genome where non-coding RNAs (ncRNAs), such as long non-coding RNAs (lncRNAs), and their proteins are bound. In this method, samples are first crosslinked and sonicated. Biotinylated tiling oligos are hybridized to the RNAs of interest, and the complexes are captured with streptavidin magnetic beads. After treatment with RNase H the DNA is extracted and sequenced. With deep sequencing the lncRNA/protein interaction site can be determined at single-base resolution.

Reviews

References
The non-protein–coding parts of the mammalian genome encode thousands of large intergenic non-coding RNAs (lncRNAs). To identify lncRNAs associated with activation of the innate immune response, this study applied custom microarrays and Illumina RNA sequencing for THP1 macrophages. A panel of 159 lincRNAs was found to be differentially expressed following innate activation. Further analysis of the RNA-Seq data revealed that linc1992 was required for expression of many immune-response genes, including cytokines and regulators of TNF-alpha expression.

Illumina Technology: HiSeq 2000
The authors used next generation sequencing to construct a comprehensive miRNA atlas of T cell development, which reveals the dynamic nature of miRNA gene transcription and processing throughout this developmental pathway starting from murine hematopoietic stem cells to mature CD4 and CD8 thymocytes. They found that many of the miRNAs that were highly expressed in bone marrow progenitor populations were down regulated by 2-3 orders of magnitude in thymocyte populations. They also used ChIP-seq and polyA RNA enrichment (RNA-Seq) to map the structures of miRNA genes expressed in CD4 T cells.

Illumina Technology: Genome Analyzer system, RNA-Seq and ChIP-Seq
Authors used next-generation sequencing to perform smRNA expression profiling of human CD56+CD3- Natural Killer (NK) cells during the process of cytokine activation. Of the >200 novel miRNAs identified, they report two abundant miRNA, miRNA-378 and miRNA-30e, are down-regulated in IFN-α activated NK cells. They also found that these two miRNA directly target granzyme B and perforin, respectively, which demonstrates that miR-378 and miR-30e suppress human NK cell cytotoxicity during NK cell activation.

Illumina Technology: smRNA-Seq between 18 and 30 bp
An Overview of Publications Featuring Illumina® Technology


ChIP-Seq

Many transcription factors and chromatin modifiers linked to innate and adaptive immunity.\textsuperscript{154-157} The identification and characterization of the genome-wide locations of transcription factors and chromatin-modifying enzymes and the modification status of histones has been accelerated by the application of chromatin immunoprecipitation techniques in next-generation sequencing analysis (ChIP-Seq). This method employs antibodies directed against a target protein to isolate a DNA-protein complex. Purified DNA is obtained from the immunoprecipitated DNA-protein complexes and is subsequently ligated with sequencing adaptors, amplified by PCR and sequenced on a next-generation sequencing platform.\textsuperscript{158} Ultimately, the need to comprehend global transcriptional regulation of the immune system positions ChIP-Seq as a powerful application, which informs our understanding of the dynamic processes of stem cell differentiation, formation of immunological memory, disease progression, and response to environmental stimuli.\textsuperscript{158,159}


\textsuperscript{156.} Northrup D. L. and Zhao K. (2011) Application of ChIP-Seq and related techniques to the study of immune function. Immunity 34: 830-842


\textsuperscript{159.} Vahedi G., Takashashi H., Nakayama S., Sun H. W., Sartorelli V., et al. (2012) STATs Shape the Active Enhancer Landscape of T Cell Populations. Cell 151: 981-993

\textbf{The ChIP-Seq workflow.} Chromatin immunoprecipitation sequencing (ChIP-Seq) is a well-established method to map specific protein-binding sites. In this method, DNA-protein complexes are crosslinked in vivo. Samples are then fragmented and treated with an exonuclease to trim unbound oligonucleotides. Protein-specific antibodies are used to immunoprecipitate the DNA-protein complex. The DNA is extracted and sequenced, giving high-resolution sequences of the protein-binding sites.
Reviews


References


BIBLIOGRAPHY


Browning B. L. and Browning S. B. (2013) Next-generation sequencing of small RNAs from HIV-infected cells identifies phased microRNA expression patterns and candidate microRNA differently expressed upon infection. MBio 4: e00549-00512


A new HLA-B44 allele and isolation of rare memory B cells with antigen tetramers. Blood 118: 348-357


Gottgens B. (2012) Genome-scale technology driven advances to research into normal and malignant haematopoiesis. Scientifica (Cairo) 2012: 437-556


Kawai T. and Akira S. (2011) Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity 34: 637-650


Le S. Q. and Durbin R. (2011) SNP detection and genotyping from low-coverage sequencing data on multiple diploid samples. Genome Res 21: 952-960


Maslanik T., Tannura K., Mahaffey L., Loughridge A. B., Benninger L. L., et al. (2012) Commensal bacteria and MAMPs are necessary for stress-induced increases in IL-1beta and IL-18 but not IL-6, IL-10 or MCP-1. PLoS ONE 7: e50636


Schwab C., Berry D., Rauch I., Rennisch I., Ramesmayer J., et al. (2014) Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. ISME J


Walsh D., McCarthy J., O’Driscoll C. and Melgar Walsh D., McCarthy J., O’Driscoll C. and Melgar


