frontiers
in Immunology

Check for updates

Multi-Omics Approaches in Immunological Research

Xiaojing Chu [1,2,3*†], Bowen Zhang [2,3†], Valerie A. C. M. Koeken [2,3,4], Manoj Kumar Gupta [2,3] and Yang Li [1,2,3,4*]

[1] Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, Netherlands,
[2] Department of Computational Biology for Individualised Medicine, Centre for Individualised Infection Medicine (CiiM), a joint venture between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany,
[3] TWINCORE, Centre for Experimental and Clinical Infection Research, a joint venture between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany, [4] Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, Netherlands

The immune system plays a vital role in health and disease, and is regulated through a complex interactive network of many different immune cells and mediators. To understand the complexity of the immune system, we propose to apply a multi-omics approach in immunological research. This review provides a complete overview of available methodological approaches for the different omics data layers relevant for immunological research, including genetics, epigenetics, transcriptomics, proteomics, metabolomics, and cellomics. Thereafter, we describe the various methods for data analysis as well as how to integrate different layers of omics data. Finally, we discuss the possible applications of multi-omics studies and opportunities they provide for understanding the complex regulatory networks as well as immune variation in various immune-related diseases.

Keywords: multi-omics, systems immunology, integrative analysis, immune-related diseases, immune variation

INTRODUCTION

Infections cause millions of deaths each year, and the current COVID-19 pandemic underlines the devastating effects of these communicable diseases. At the same time, the incidence of immune-related diseases such as atherosclerosis (1) and autoimmune diseases such as type 1 diabetes mellitus (2) have been increasing. All these diseases are related to or mediated by the immune system. Thus, the immune system plays a vital role in health and disease, and it is our defense mechanism against harmful substances, infectious diseases and cancer. Within a properly functioning immune system, immune responses should be kept in a certain range, as both hypo-activation and hyper-activation lead to disorders of the immune system. Understanding how the immune system works and what causes the immune system disorders may help us to efficiently fight against immune-related diseases.

However, getting a comprehensive understanding of the immune system is a challenging task. First of all, the immune response is mediated through a complex interactive network of many different immune cells and molecules, such as cytokines, immunoglobulins, and metabolites. At the same time, this network is highly variable depending on the exact threat of the wide variety of pathogens and other substances it's responding to. To make things even more complex, the immune

response to a certain stimulus or infection is highly variable between individuals, leading to population heterogeneity. This heterogeneity is exemplified by differences in severity of patients suffering from the same infectious disease (3), variability in vaccine efficacy (4), and variation in responses to the same medical treatment (5). Many factors contribute to the immune network and the inter-individual variation of immune responses, highlighting both the promise and the challenge of multi-omics studies.

Until now, omics data have been used in many immunological studies to identify the determinants of immune variation and molecular bases of the immune process in different population groups. Properly designed omics studies should make use of appropriate measurements as well as reasonable analytic approaches, which depend on their specific research question. Taking omics studies on COVID-19 as an example, a genome-wide association study revealed eight genetic regions to be associated with critical illness in COVID-19. By integrating both genome and transcriptome data, the authors prioritized one gene, *IFNAR2*, that might play a causal role in COVID-19 (6). Another study, focusing on transcriptome data of immune cells from the lung and blood, identified several pro-inflammatory immune pathways related to the pathogenesis of COVID-19 (7). A proteomics and metabolomics study investigated the changes in COVID-19 patient sera, and identified molecular changes implicating dysregulation in macrophage pathways, complement activation, and platelet degranulation, as well as suppression of metabolic pathways (8). A cellomics and single-cell transcriptome study also revealed dysregulation of the monocyte compartment as well as two neutrophils clusters specific to severe COVID-19 patients (9). Moreover, a study integrating single-cell transcriptome, cellomics, epigenome and proteome comprehensively characterized complex dynamic changes in immune cells. Their results disclose an elevation of IFN-activated megakaryocytes and erythroid cells, hypomethylations around immune signaling genes, and co-expression modules associated with clinical outcome (10). Additionally, a study on fecal fungal microbiota of COVID-19 patients showed enrichment of *Candia albicans* and a highly heterogeneous mycobiome configuration during hospitalization (11). From different angles, these studies make use of omics data to provide insights in the molecular pathology of COVID-19, which can eventually lead to improved therapeutic strategies.

In this review, we present an introduction to multi-omics studies to investigate immune function and variation. The review is split into three parts. In the first section, we describe in brief about the different layers of omics data relevant for immunological research, including genome, epigenome, transcriptome, proteome, metabolome, digestive system microbiota and cellomics (12) [also called cytomics (13)] (**Figure 1**), and the commonly used methodological approaches to measure these different types of omics data. We also discuss important considerations and recommendations for an appropriate study design. In the second section, we discuss how to analyze and integrate multiple omics platforms, including system genetic approaches to identify genetic factors, integration among multiple genetic profiles, as well as the integration and association with other omics data layers. We demonstrate how recent studies applied a multi-omics approach to the immune system researches, and we discuss the interpretation of results from different approaches and their importance in immunological studies. In the third section, we discuss the immunological subjects that need specific attention and may see progress in the next few years. As for detailed information on computational algorithms and models in multi-omics integration (14, 15), imputation on missing omics data (16), and strengths and limitations of system approaches in infectious diseases research (17), we refer readers to other recent reviews.
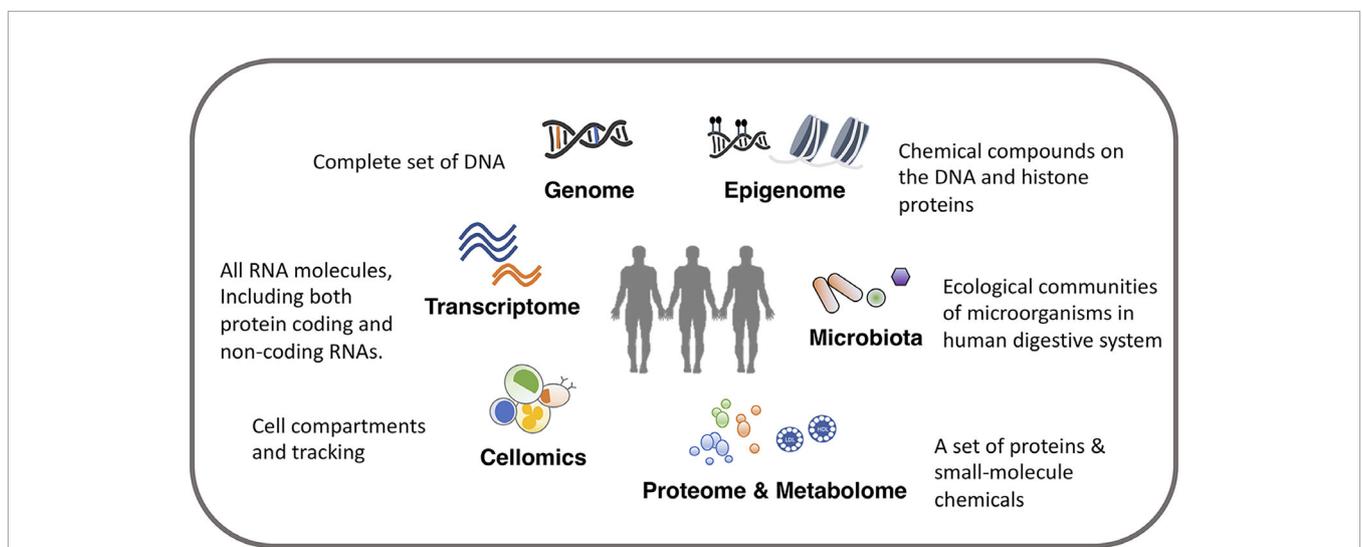

FIGURE 1 | Overview of omics data.

MEASUREMENTS OF OMICS DATA

We can identify potential immunological mediators and study immune phenotypes with a wide range of omics comprising of various molecular and cellular phenotypes including genome, epigenome, transcriptome, proteome, metabolome, digestive system microbiota and cellular phenotypes such as cell composition (**Table 1**). A single omics data layer characterizes a specific biological process from one aspect, for example, transcriptome, but this can only provide insights on genes at a transcriptional level. To achieve a holistic picture of the immune system, a systematic collection of multi-omics data is often required. The tissue (or source) to be measured is another important aspect to be considered. For example, the genome is usually regarded as a stable feature for each individual and collected from an easily accessible tissue, such as blood. Only in some specific contexts, somatic mutations acquired after birth have to be considered and measured in specific tissues (18). However, many other types of omics, such as transcriptome, proteome and metabolome, vary between cell types and tissues. Therefore, it is important to consider the tissue in your experimental design and aim to get as close to the relevant tissue as possible.

Given the complexity of the immune system, there is no golden standard for what to collect in multi-omics studies. The necessary data depends on the research question and subjects. Understanding the different layers of omics data is helpful for setting up an appropriate study design. Therefore, in this part, we introduce features and categories of different omics, and describe important considerations when generating these data.

GENOME VARIATION MEASUREMENT

Genotyping detects diversity in the genome. It describes small variations, such as single-nucleotide polymorphisms (SNPs), insertion/deletions (InDels) as well as large-scale mutations such as insertions, deletions and amplifications. Genetic diversity can lead to variation in individual immune function (19).

To date, many techniques can be used for detecting genotypes, including DNA sequencing, DNA microarrays (also known as genotyping chips) and PCR-based methods. These approaches can be categorized based on their measurement scales (high-throughput *vs.* low-throughput methods) or based on whether they include unknown variants (discovery *vs.* screening methods). Classical sequencing-based approaches detect genetic variants in a nearly unbiased manner on the genome (whole-genome sequencing) or within the exome regions (whole-exome sequencing), including known or novel SNPs as well as structural mutation such as short insertions, deletions, and copy number variations.

Considering the cost and effectiveness of genotyping scales and cohort sizes, most of the population-based association studies choose genotyping screening methods, such as DNA microarrays. These methods measure thousands to millions of known SNPs in well-studied organisms, such as humans and mice. The targeted polymorphisms depend on the chip designs. For example, Immunochip contains 196,524 polymorphisms (718 InDels and 195,806 SNPs) on most reported loci involved in autoimmune and inflammatory diseases (20), whereas other custom genotyping chips contain loci designed for specific research areas, such as Metabochip (21) or cardiovascular disease chip (22). The number of variants that can be detected using genotyping chips has increased over the years, but even the high-density 5 million SNPs chip (Illumina OMNI5) covers only a small fraction of the 3.3 billion bases in the human genome.

In order to improve the power in discovering genetic associations on the regions poorly covered by DNA microarrays, genotype imputation approaches are often used to expand the coverage. For example, a commonly used genetic imputation server (https://imputationserver.sph.umich.edu/index.html#)! takes the ~60,000 public available human haplotypes, covering ~40,000,000 SNPs, as a reference to impute millions of missing SNPs based on the measured genotypes and linkage disequilibrium (LD) structures (23).

Before association analysis, genotype data should pass a standard quality control (QC) at both individual level and SNP levels. Individuals with discordant sex information, outlying missing genotype or heterozygosity rate should be excluded

TABLE 1 | Typical approaches in omics measurements.

	sequencing-based	microarray-based	others
genetics	whole-genome-seq, whole-exome-seq	Illumina OMNI5, Immunochip etc.	–
epigenetics	ATAC-seq, whole-genome bisulfite-seq, RRBS-seq, DNase-seq, FAIRE-seq, ChIP-seq, etc.	MethylationEPIC BeadChip, ChIP-chip, etc.	–
3D chromosome	Hi-C, etc.	–	–
gene expression	RNA-seq, scRNA-seq, SLAM-seq	Affymetrix Genome U133 array, Illumina Whole-Genome Gene Expression BeadChips, etc.	–
protein level	–	–	Immunoassay, MS -based approaches
metabolites	–	–	NMR, MS-based approaches
microbiome	16s rRNA-seq, metagenomics, etc.	–	–
cellomics	single cell sequencing approaches	–	FCM, CyTOF

(24). Duplicates and relatives could be identified by calculating identity by descent (IBD), and a multi-dimensional scaling plot merging with reference data such as the 1000 Genomes project (25) could help with the identification of individuals of divergent ancestry. SNPs failed in genotyping and/or imputation and SNPs with low frequency and/or that deviate from the Hardy-Weinberg equilibrium are commonly removed before association analysis, especially in array-based studies, because those signals usually relate to bad genotyping quality. However, some SNPs with low frequency may also contribute to rare diseases or phenotypes. With the increase in genotyping quality, more and more recent studies focus on the function of rare alleles (minor allele frequency [MAF] < 0.01) (26–29).

EPIGENOME AND 3D CHROMOSOME MEASUREMENT

Epigenetics describes the study of chromatin traits (either in DNA or histones) that do not involve changes in the nucleotide sequence. Epigenetics measurements are mainly characterized by the changes in histone modification (methylation and acetylation), DNA methylation, chromatin modification, chromatin accessibility, and chromosome structure.

DNA methylation is the process of adding methyl groups to DNA molecules, almost exclusively in CpG dinucleotides with the cytosines on both strands being methylated. This process usually acts in promoter regions to repress gene transcription, and abnormal hypermethylation, which results in transcriptional silencing, is often associated with immune diseases or used as a biomarker (30). Genome-wide techniques, such as whole-genome bisulfite sequencing (WGBS) (31), reduce representation bisulfite sequencing (RRBS-seq) (32) and other non-targeted DNA methylation profiles, provide an opportunity to discover novel biomarkers. Other techniques, such as bisulfite-amplicon sequencing (BSAS) (33) and methylation arrays (34), detect the methylation status of CpG dinucleotides. Similar to genotyping arrays, the targeted regions from methylation arrays are based on the chip design. For the study of the human immune system, some well-established arrays can provide comprehensive coverage. For example, MethylationEPIC BeadChip covers over 850,000 methylation sites, making it ideal for an epigenome association study within big cohorts (35).

As the essential proteins that pack and order the DNA into structural units, histones play a role in gene regulation (36). Histone modification describes the post-translational modifications of histones, including methylation, acetylation and others. Histone methylation often occurs as arginine (R), lysine (K), or histidine (H) residues of histone H3 or H4 being monomethylated (me1), demethylated (me2), or trimethylated (me3). Array-based and sequencing-based approaches, such as ChIP-chip and ChIP-seq (37), are used to identify specific histone modifications that bind to DNA regions or domains.

Chromatin modifications and accessibility is another important aspect of epigenetic changes. One of the most widely-used techniques to capture chromatin accessibility is called Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq). A standard "bulk" ATAC-seq measurement detects genome-wide open chromatin within a pooled sample or tissue, while in order to capture cellular heterogeneity, single-cell ATAC-seq measures chromatin accessibility in thousands of individual cells, which can generate genome-wide profiles from 10k to 100k cells per experiment (38). Alternative techniques are also used to investigate chromatin phenomena, such as DNase-seq and FAIRE-seq, which measure open chromatin in regulatory regions, MNase, which identifies well-positioned nucleosomes, and ChIP-seq, which is used to detect binding sites of specific transcription factors (39).

Most epigenetic measurements also come with technical errors and biases. Biological replicates and technical replicates can help to characterize variability between samples and sequencing runs. Putting replicates of different conditions in the same batch is also important to avoid batch effects confounding treatment effects. Large projects, such as the Encyclopedia of DNA Elements (ENCODE), have provided standard pipelines for processing many types of epigenetic data, such as ChIP-seq and ATAC-seq. However, this is not applicable in all cases. Applying appropriate QC strategies and software that accounts for bias effects according to the experiment design is essential to obtain robust results. To increase the coverage of epigenetic measurements, several methods, such as ChromImpute (40), Melissa (41), Avocado (42), and SCALE (43), provide imputation approaches for different epigenetic markers. However, the existing imputation approaches have several limitations (16), and are not as widely applied as genotype imputation methods.

3D chromosome structure describes how chromosomes are folded, packaged, and organized into functional compartments, and how different compartments are interconnected. Orthogonal ligation-based approaches include DNA-FISH, which can help with nuclear architecture visualization, and chromosome conformation capture (3C) techniques. One of the 3C techniques, Hi-C, is the most widely used approach to detect interactions between different genome regions (in gigabase-scales) (39, 44). Single-cell adaptation of Hi-C methods are also used to investigate the interactions in individual cells (45).

Ligation-based approaches have the limitation of detecting DNA fragments connected with multiple genomic regions. To overcome this limitation, orthogonal ligation-free methods including genome architecture mapping (GAM) (46), split pool recognition of interactions by tag extension (SPRITE) (47) and chromatin-interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop) (48) were developed.

TRANSCRIPTOME MEASUREMENT

The transcriptome comprises all RNA molecules, both coding and non-coding transcripts, in a single or population of cells. Traditional qPCR techniques can only quantify a limited number of genes at the same time. The most commonly used high-

throughput techniques are RNA sequencing (RNA-seq) and microarray, and they can detect a large number of genes. Similar to genotyping methods, a sequencing-based approach (RNA-seq) can quantify the entire transcriptome, while microarray-based approaches (e.g., Affymetrix Genome U133 array and Illumina Whole-Genome Gene Expression BeadChips) are designed to target most known genes. In addition, a typical RNA-seq can detect alternative splicing and rare isoforms, which microarray-based techniques cannot.

Certain coverage is required for sequencing data, which depends on the aim of the study. For instance, a bulk RNA-seq study for human differential expression profiling requires 10-25 million reads per sample, while alternative splicing or allele-specific expression analysis need 50-100 million, and identifying novel transcripts requires >100 million reads per sample.

However, a "bulk" like measurement of transcriptome cannot deal with the cell heterogeneity and can be influenced by cell composition changes. Single-cell RNA sequencing (scRNA-seq) was designed to uncover the transcriptome diversity in heterogeneous samples, characterizing the transcriptome in cell resolution. There are several approaches of scRNA-seq, among them are plate-based (Smart-seq2) (49) and droplet-based (10x Genomics) the most commonly used ones. Usually, as few as 10,000 to 50,000 reads per cell are enough to detect cell types, and 500,000 reads can cover most of the genes (50).

In order to increase exonic coverage and accuracy of gene quantification, polyA selection library preparation is commonly applied in scRNA-seq approaches such as 10x scRNA-seq (51). This will, however, miss the important immune repertoire profiling, such as B-cell and T-cell receptors, which is mainly distinguished by their 5' mRNA sequences. Thus, sequencing facilities, such as 10x genomics, provide full length paired B-cell and T-cell repertoire sequencing, simultaneously, when examining cellular gene expression level. Combined with transcription measurement, this information can improve our understanding of clonal expansion and better characterize immune cell heterogeneity and functions (52).

SLAM-seq detects the newly synthesized RNAs using a metabolic RNA labeling approach. Compared to the other scRNA-seq techniques, this method can track the transcriptome dynamics (53). For example, scSLAM-seq was applied to characterize the onset of infection with lytic cytomegalovirus in single mouse fibroblasts (54).

The transcriptome reflects the dynamic changes in biological processes, which is much more unstable. So, an appropriate sampling strategy on transcriptome data is crucial. In addition to the quality control, normalization is usually performed within a sample and between samples. When considering comparison analysis, it is also necessary to have biological replicates and check for batch effects using clustering-based approaches. There are many computational tools handling batch effects. Of note, integration approaches (55), as included in Seurat (56) and Harmony (57) packages, are commonly used in scRNA-seq analysis which detect the consistent cell type signals from different batches or measurements. However, when the batch difference is confounded with other group information, it will be

tough to filter out the batch effects. In addition to batches, it is also important to consider other potential confounders in experiment design. For example, transcriptional differences were observed between males and females in COVID-19 patients (58), thus a gender-balanced design in a case-control study will lead to an unbiased conclusion for COVID-19. Moreover, when considering sampling tissues for immune responses, circulating leukocytes are often measured for systemic inflammatory responses, while inflamed tissues are measured for local inflammatory responses. In order to expand the capacity, deconvolution approaches have been applied to bulk RNA-seq data to characterize cell type compositions (59, 60), while expression recovery methods have been applied to single-cell RNA-seq data to reduce the dropout noise (61, 62). Like imputation approaches in genome and epigenome studies, one should be aware and careful with the potential false signals in these recovery or deconvolution approaches.

PROTEOME AND METABOLOME MEASUREMENT

Proteins are the major transcriptional products and functional units in the immune system. Immune molecules like immunoglobulins and cytokines are usually detected and/or quantified by immunoassays such as immunofluorescent staining, enzyme-linked immunosorbent assay (ELISA), enzyme multiplied immunoassay technique (EMIT), or mass spectrometry (MS)-based approaches.

In addition to independent measurement, proteins can be also measured together with RNA transcripts. CITE-seq provides an opportunity of identifying surface proteins along with RNA-seq. This approach is often used for cell labeling in scRNA-seq (63). Cells in different research groups (e.g., under different treatments, from different tissues) could be labeled with different antibodies as hashtags, then sequenced together as one pool. This process has two advantages: decreasing cost and excluding potential batch effects. In addition, as we also know that some immune cell types have specific cell markers, this approach can also be used to identify cell types. For example, the detection of CD3e, CD4 and CD8a proteins on the cell surface could help to distinguish CD4 T cells from CD8 T cells (64). Moreover, there is a new technique called INs-seq, which can measure intracellular protein activity along with scRNA-seq. This new technique shows a large potential of applications in immune-related studies (65).

The study of metabolic processes that regulate immune cell responses, which is referred to as immunometabolism, has become an exciting area in translational research, and is paving the way for novel therapies in immune-related diseases. The intermediate or end products of cellular metabolism are metabolites, which include, but are not limited to, lipids, fatty acids, amino acids, bile acids, and cholesterols. Considering the regulatory effects of metabolites on the immune response (12, 66, 67), the metabolome has become an important subject to study in immunological research.

Approaches to study the metabolome can be classified into targeted and non-targeted techniques. Nuclear magnetic resonance (NMR) spectroscopy is one of the most commonly used techniques, detecting specific nuclei in the target molecule (68). Compared to NMR, mass spectrometry (MS)-based approaches are more high-throughput and quantify metabolites in a non-targeted way, which detect mass-to-charge ratio (69). However, MS-based approaches have a limitation in annotating metabolites, which is the major drawback of this method in contrast to NMR. Metabolites data could be acquired from different sources of samples. Among them, circulating metabolites are the most commonly measured. There are also many studies about fecal and urine metabolites.

Similar to transcriptome analysis, a proper normalization (usually a log transformation) is required in both the proteome and metabolome data process. Secondly, biological replicates and batch effects have to be taken into consideration as well. In addition to linear regression, more advanced computational tools, such as ROIMCR (70), can also be used to reduce the batch effects and to identify metabolites that associate with immune responses. In terms of sampling tissues, in addition to blood cells and inflamed tissues, proteome and metabolome can also be measured in urine, which is thought to be a rich source but underestimated in recent studies (71–73). In addition, fecal metabolites are usually studied together with microbiota, which affects immune homeostasis and susceptibility of the host to immune-mediated diseases. Of note, there is a recent study reporting a reference map for serum metabolites (74), which can serve as a guide to control for irrelevant confounders in serum metabolite studies.

DIGESTIVE SYSTEM MICROBIOTA MEASUREMENT

Microbiota refer to all micro-organisms in a certain environment, for example the human digestive system. It has been reported to vary among individuals, to influence host immune functionality and to be involved in immune-mediated disease pathology (75–77). The commonly used approaches to study microbiota include 16s rRNA sequencing and metagenomics sequencing. After excluding host (human) reads, microbiota reads are aligned to the known microbiome genomes to identify the taxonomies and abundance. While there are also other omics approaches including metatranscriptomics, metaproteomics, and metabolomics, which target transcripts, proteins, or metabolites from microbiota (78).

Of note, studies on human microbiota usually have relatively low concordance compared to other omics data studies. A recent study has reported a number of host variables that could confound human gut microbiota researches. To be exact, body mass index (BMI), sex, age, geographical location, alcohol consumption, bowel movement quality (BMQ), and diet should be balanced in cases and controls when comparing gut microbiota compositions (79). In the context of sample collection, most of the microbiota samples are acquired from the stool, while urine and exhaled gas could be another important resource for microbiome detection (80, 81).

CELLOMICS MEASUREMENTS

Cellomics measurements often reveal the systemic responses at the level of cells and tissues, typically including cell composition, cellular localization and trafficking analyses. Cell composition is measured as cell type abundance or proportion, which is commonly quantified by flow and mass cytometry (82) (FCM and CyTOF) or single-cell sequencing techniques. With the help of cell surface markers or cellular-specific expression markers, both techniques can characterize hundreds of circulating cell subpopulations covering major immune cells involved in innate and adaptive immune responses (i.e., neutrophils, monocytes, lymphocytes, and their subtypes). Additionally, high-content screening (HCS) is commonly used to track cellular changes, including their localization, trafficking and morphologic phenotypes (83, 84).

SYSTEMS ANALYSIS ON OMICS DATA

After data collection and pre-processing with appropriate strategies, the next big challenge lies in linking different omics datasets and clinical phenotypes. For a certain trait or disease, a systems model can be built to specify the role and effect of different data layers. In this model, the qualitative or quantitative characteristics are linked by their relationships, which need to be estimated *via* comparison, association and other systems approaches. These links can simply be a correlation, or a regulatory or causal effect. In this section, we introduce general system approaches among different omics data and provide representative examples of how they can be applied in immunological studies.

GENOME-WIDE ASSOCIATION ANALYSIS AND QUANTITATIVE TRAIT LOCUS MAPPING

Genome-wide association studies (GWAS) aim to scan the whole genome to find genetic determinants of certain traits. When considering a binary trait (e.g., case-control), we compare allele frequency in two groups of individuals, for example one disease group and one healthy group. A chi-squared test is often applied to test for statistical significance. It is usually considered that there are ~1,000,000 independent loci in the human genome, so a p-value less than the Bonferroni corrected threshold of $0.05/1,000,000$ (5×10^{-8}) is regarded to be genome-wide significant (85).

To date, GWAS have identified ~5000 genetic risk loci of immune-related diseases in ~400 studies (86). Those findings improved our understanding of genetic factors influencing

immune-mediated diseases, further pointing to the genetic basis of pathology as well as treatment targets.

Generally, GWAS identify pathogenetic genetic factors contributing to phenotypes (diseases), though those variants will not cause disease directly but affect intermediate molecules. Quantitative trait locus (QTL) analysis is a statistical method to discover the genetic basis of the intermediated phenotypes, such as gene expression (eQTL) (87), splicing (sQTL) (88), metabolites (mQTL) (29), methylation (meQTL) (89, 90), and immune traits (91, 92).

After data normalization, a linear regression between each genetic variant and each quantitative trait is applied. Covariates are crucial aspects of the regression model of QTL analysis. Based on the type of omics, different covariates should be included in the model to correct the detected phenotypes. In general, basic host features such as age and sex are considered, and a population structure has to be additionally taken into account, especially in large cohorts with samples from admixed ancestry (93, 94).

eQTLs are the associations between SNPs and expression of genes, which provide insights of the function of genetic variants. eQTLs can explain 10% - 50% heritability of a phenotype/disease (95), which means that gene expression variation is one of the major consequences of genetic variants. It is very useful for prioritizing pathogenic genes when there is an association between a gene expression and a pathogenic genetic variant. Based on the position, eQTLs are classified into cis-eQTL (eQTL within 1Mb of the gene) and trans-eQTL (eQTL located outside 1Mb of the gene). Among them, trans-eQTLs are more tissue-specific than cis-eQTLs (88). Of note, tissue-specific eQTLs provide a way for prioritizing pathogenic tissues (96).

QTL analysis on epigenome identifies the associations between genetic variants and epigenetic modification. Most genome-wide significant disease-associated loci (~93%) are located in non-coding regions (97), particularly, regulatory elements identified by ENCODE (98) and Roadmap projects (99). These observations highlight the importance of epigenome in the genetic regulation of diseases and immune functionality. Similar to eQTL analysis, this analysis could help us find the potential epigenetic mechanism responsible for the association between genetic variants and immune traits/diseases. For example, a study investigated genetic variants that affect the activity of cis-regulatory domains (aCRD-QTLs) or correlation structure within cis-regulatory domains (sCRD-QTLs) in 317 lymphoblastoid and 78 fibroblast cell lines, and their consequence on gene expression (100). At the same time, genetic variants can also affect methylation (meQTL) by influencing the binding of DNA methyltransferase (DNA MTase). Large meQTL studies in blood samples showed significant enrichment in autoimmune diseases such as ulcerative colitis and Crohn's disease (101).

pQTL mirrors the associations between genetic variants and protein level. About 40% of cis-protein quantitative trait loci (pQTLs) are also eQTLs, as expected, indicating a sequential genetic regulation between gene expression level and protein levels. By applying pQTL analysis, we could identify the potential mechanism, at the protein expression level, behind the association from genetic variants to immune-related phenotypes. Same as with cis-eQTLs, cis-pQTLs are also located around transcription start sites (TSS). Notably, pQTL showed a significant enrichment on missense, 3UTR and splice region (102). pQTLs could also help with prioritizing causal proteins/genes of immune traits/diseases. For example, a pQTL of serum IL18R1 and IL1RL1 also associates with atopic dermatitis. This association between genetic locus and protein level indicates a possible involvement of IL18R1 and IL1RL1 in atopic dermatitis pathology (102).

Metabolites that mediate the association between genetic variants to immune functionality and immune diseases could be discovered in an mQTL analysis. More than 140 genomic loci are associated with circulating metabolite features explaining a median 6.9% heritability (103). Overlaps between mQTLs and immune traits QTLs suggest the role of metabolic processes in the genetic regulation of immune functionality. For instance, a mQTL study indicates that mQTL loci ARHGEF3 (rs1354034) and LRRC8A (rs13297295) also affect platelet function and neutrophil function, respectively (104).

Immune phenotypes such as circulating immune cell proportion and cytokine production capacity in response to stimulations are crucial parameters when characterizing immune activities. Understanding the genetic determinants of immune phenotypes can provide insights into immune function and immune-mediated diseases. A human functional genomics project has identified >20 genetic factors determining immune cell proportions and cytokine production upon stimulations, which provided a link between genetic control and inter-individual variation (92, 105).

INTEGRATION OF MULTIPLE GENETIC ASSOCIATION PROFILES

In the context of immunological research, multiple diseases, and molecular and cellular phenotypes can be regulated by the same genetic factors, indicating an internal association between them. Integration with multiple genetic profiles can provide insights and build connections between associated phenotypes. Ideally, such genetic profiles can be directly built from GWAS and QTL analysis of different layers from the same individuals. Otherwise, they can be also collected from different population-based cohorts. A number of computational approaches have been developed to discover the link. In particular, approaches like colocalization (106), genetic correlation (107) and Mendelian randomization (MR) (108) take genetic variants as the instrumental variables to infer the association or causality when multiple traits are associated with the same locus.

Colocalization analysis evaluates the association from each of the single locus, and it helps to identify the phenotypes that share the same genetic regulation. Examples of colocalization analysis include a study integrating genetics, epigenetics and transcription to identify colocalization of molecular traits from CD14+ monocytes, CD16+ neutrophils and naïve CD4+ T cells

(109). Results from this analysis illustrate molecular mechanism at autoimmune diseases-associated variants, including an alternative splicing signal around SP140 in T cells which might be involved in Crohn's disease pathology.

Genetic correlation considers the full summary statistics to describe to which extent the genetic background is shared between two phenotypes. An example from a LD regression-based genetic correlation approach showed a shared genetic basis of autoimmune diseases such as Celiac disease and type 1 diabetes (107). This indicates a similar pathological mechanism between these two diseases.

MR is a statistical method working on the step from association to causality. If one trait (exposure) is causal to another trait (outcome), then the genetic factors contributing to the exposure should also contribute to the outcome. This would be reflected in the correlation between effect sizes of the same genetic variant on exposure and outcome. There are many examples of immune-related studies that applied MR, which led to the identification of causal relationships between IL-6 signaling and rheumatoid arthritis (110), IL-18 and inflammatory bowel disease (111) and between eosinophilic indices and asthma (26).

COMPARISON AND ASSOCIATION OF EPIGENOME AND 3D CHROMOSOME STRUCTURES

Systems analysis of epigenetic changes can investigate their influence on and changes induced by immune functionality or variation as well as disease susceptibility and development (112, 113). As an example, the impact of cytokines was studied on the epigenome of insulin releasing cells (β cells) from type 1 diabetes pancreases. By measuring ATAC-seq, Chip-seq and RNA-seq, the authors identified proinflammatory cytokines induced neo/primed epigenetic events in human β cells (114). Moreover, in immune systems, the effects of epigenetic changes lead to long-term alterations in the metabolic and transcriptional pathways, and further induce immune memory (115) or immunological diseases (116). Thus, epigenomics is another vital area for better understanding of the personalized immune system.

While genetics is stable, the epigenome is subject to dynamic changes, which can be induced or affected by host and environmental factors, such as smoking, drug usage, diet, aging, inflammation, disease, and exposure to pets. Considering that epigenetic changes affect gene transcription levels, the epigenome is a pivotal part to study when trying to understand immunological networks.

In a case-control study, differential accessible regions (DARs) could be identified in an ATAC-seq data, as well as differential methylation positions/regions (DMP/DMRs) in bisulfite sequencing and methylation array. Instead of comparison analysis, association analysis is applied to continuous phenotypes to get associated regions. Upon the position of acquired regions, we could further map them to the

corresponding genes. More specifically, by checking which gene TSS regions are overlapped with the peaks/regions, the peaks/regions could be matched to genes, and then for pathway analysis to get more biological meanings. For example, in a multi-omics study on mixed-phenotype acute leukemia, researchers associated scATAC-seq with transcription responses from scRNA-seq and antibody captured from CITE-seq. Despite widespread epigenetic heterogeneity of chromatin accessibility within patients, they reported common malignant signatures across patients, and thus revealed both distinct and shared molecular mechanisms of mixed-phenotype acute leukemia (117).

Another application of epigenetic analysis is to annotate the function of the identified regions, based on the signals from epigenetic markers. A tool (118) used a multivariate hidden Markov model applied to annotate regulatory elements (e.g., Transcription starting sites, enhancers, promoters) with histone markers (e.g. H3K4me1, H3K4me3, H3K27me3, H3K9me3, H3K36me3) binds to the chromosomes. Applying this method, an example learnt the chromatin states in mice and humans, and reported the up-regulation of immune regulatory regions in Alzheimer's disease (119).

The analyses on 3D chromosomes are generally similar. In a case-control study with Hi-C data, we could get the compartment switches in a comparative analysis. We could further predict the interactions between those segments (120). Referring public epigenetic databases or genome annotations, we could check the overlap between switched compartments or interactions and known epigenetic markers or elements. Based on this information, we could again associate the changes with other immune profiles or annotate the involved regulatory elements. For example, in a study of lineage commitment of early T cells with Hi-C data, authors found wide compartment re-organizations across chromosomes from a transition between T cell double-negative-2 stage to double-negative-3 stage, and later double-negative-4 stage to double-positive stage. They annotated the changes with domain scores, and more interestingly, they found the changes in the domain scores between the two transitions are positively correlated, which suggests the re-organization at the former transition is actually reinforced at the later transition (121). Another example includes a study on activated T cells, that identified activation-sensitive interactions related to autoimmune diseases captured by Hi-C data (122).

To capture the changes that occur in cellular activation and differentiation, time-series study is another hot topic in associating epigenome and 3D chromosome structures to immune responses. For example, a recent study elucidates the chromosome conformational changes in B lymphocytes as they differentiate and expand from a naive, quiescent state into antibody secreting plasma cells (123). The authors reveal that the changes to 3D chromatin structure occur in two discrete windows, associated with prolonged time in the G1 phase of the cell cycle. Their results also suggest chromosome reconfiguration is linked to a gene expression program that controls the differentiation process required for the generation of immunity.

COMPARISON AND ASSOCIATION OF TRANSCRIPTOME AND PROTEOME

As the downstream products of genetic and epigenetic regulation, transcriptome and proteome changes directly reflect the influence of genetic and epigenetic variants. Comparison and association studies of transcriptome and proteome have allowed researchers to estimate functional units and validate hypotheses in immune regulation.

As for a case-control study, the first and direct analysis is identifying differentially expressed genes/proteins (DEGs/DEPs), followed by pathway analysis. If the corresponding phenotypes are continuous, then associated genes/proteins will be identified before pathway analysis. Examples include many transcriptome/proteome studies upon the severe infectious disease COVID-19. Transcriptome measurement across samples from healthy, moderate patients and severe patients suggests an overall acute inflammatory response in COVID-19 patients, whereas transcriptional responses of high cytotoxic effector T cells are associated with moderate patients, and deranged interferon responses are associated with severe patients (124). Moreover, a urine proteome study identified 1986 urine proteins showing significant level changes in COVID-19 patients than in healthy controls (125).

Different from bulk RNA-seq, the adding information in scRNA-seq: cell composition, provides more analysis potentials. In a case-control study, in addition to DEGs and enriched pathways identification within each cell cluster/type, cell proportion could be compared between groups while novel cell subpopulation could also be identified in particular cases. For example, a scRNA-seq on two COVID-19 cohorts reported identical dysfunctional neutrophil clusters in severe patients' blood (9). When considering the TCR/BCR analysis, it would be interesting to explore the clonotype expansion and diversity under different conditions (126, 127), immune development stages (52), or antigen specificity (128). Usually, a clonal expansion means an adaptive immune response targeting certain stimulation, since a certain receptor is the mediator of specific antigen recognition.

Since transcriptome/proteome data is rapidly responding to environmental changes, with the transcriptome/proteome analysis in a time-series study, we could associate the dynamics with infection or stimulation to comprehensively understand the host immune responses. A nice example is demonstrated in a study of influenza vaccination efficiency, where authors measured the hemagglutination-inhibition (HAI) antibody titers and transcriptional responses at baseline and multiple time points post-vaccination. By comparing the profiles between day 28 and day 180, the authors describe individual categories as temporary and persistent responders and illustrate the underneath molecular mechanism (129). Many approaches have been developed for time-series studies, such as regression-based method like maSigPro (130) and a fusion method like O2-PLS (131). Of note, the dynamic study can also be achieved by applying a trajectory analysis such as pseudotime analysis (132, 133) and RNA velocity analysis (134) in scRNA-seq analyses. In

a recent study on COVID-19, researchers longitudinally measured samples at several time-point after symptoms, and applied pseudo-time trajectory inference on scRNA-seq data of epithelial cells from the upper respiratory tract. Based on the trajectory, they predicted a new, alternative differentiation pathway that is dependent on the interferon response and marked by interferon-stimulated genes, such as *ISG15*, *IFIT1*, and *CXCL10* (135).

Co-expression analysis among transcriptome or proteome provided information about gene co-regulation and interactions. These co-expression relationships are inferred by different association methods, such as a weighted gene co-expression network analysis (WGCNA) (136) applied on transcriptome to identify consistent expression patterns among genes. The identified associations among gene expression could be applied to predict gene co-regulatory networks, further to prioritize genes involved in the same pathways (137). At protein level, parts of these co-expression relationships could further be explained by protein-protein interactions, which are also collected by several protein-protein interaction databases, including the innateDB (138) who particularly focus on immune interactions. In application, similar to gene co-expression networks, protein-protein interaction relationships could help with functional/pathway enrichment analysis (139).

In the recent single-cell experiments, the co-expression relationships are further applied to predict the cell-cell interactions. By detecting the correlation between known ligand and receptor genes among different cell sub clusters, we could infer the potential communications between cell populations (140). This analysis fits well with immune network analysis. For example, by detecting ligand and receptor genes signals, a recent study identified cross-talks between CD8+ T cells and epithelial cells altered in the colon of ulcerative colitis patients compared to healthy controls (141). Additional methods, such as NicheNet (142), also take knowledge of gene regulatory networks or protein-protein interaction networks from public databases and literatures, then build a model to further predict the activated targets of the cell-cell interactions by correlating the ligands expression level with its potential downstream gene or protein level interactions. In an example study of cell-cell interaction underlying the tissue-specific imprinting of macrophages, the authors deciphered the interaction signals driving monocyte recruitment, engraftment, and acquisition of the Kupffer cells associated transcription factors, and they identified the contributions of different cells to Kupffer cell niche (143).

COMPARISON AND ASSOCIATION ON METABOLOME/MICROBIOME

Metabolome or microbiome are additional factors that reflect, or affect, a person's state of health (144, 145). Similar to transcriptome or proteome, comparison and association analysis could be applied on metabolome and microbiome data. However, metabolome can be hardly linked to genes,

which leads to different strategies of interpretation. Taking KEGG (146) and HMDB (147) as references, an online tool MetaboAnalyst performed metabolic pathway enrichment and network analysis on the identified metabolites (148). An example serum study on COVID-19 detected accumulation of 11 steroid hormones and suppression of amino acid metabolism in patients (8)

As for the gut microbiome, a diversity analysis could be applied to taxonomy data. There are different strategies available for functional profiling on the gut microbiome data. For example, HUMAnN takes metagenomic or metatranscriptomic sequencing data as input to identify gene families and abundances (149). Gene families could be further matched to broader functional categories, such as MetaCyc metabolic pathways and GO categories for functional interpretation. For example, a study associated gut microbiome features to cytokine production capacity, and found microbial metabolic pathways: palmitoleic acid metabolism and tryptophan degradation to tryptophol showed associations with TNFα and IFNγ production (150).

As in transcriptome and proteome analyses, time-series studies could provide valuable information in metabolome and microbiome data. For example, in a study of metabolic functions of gut microbes from patients with Inflammatory Bowel Diseases, fecal samples were collected at baseline and 2, 6, and 14/30 weeks after induction of therapy to collect metabolic and microbiota profiles. The observed association in dynamics of metabolites and diversity shifts of microbiota reveals the heterogeneity of the disease, and helps the authors to build a silico model that might be used to identify patients likely to achieve clinical remission from the therapy (151).

## INTEGRATION OF EPIGENOME, TRANSCRIPTOME, PROTEOME, METABOLOME, MICROBIOTA AND CELLOMICS

Besides associations between omics data and genetics, a simple association analysis between two different non-genetic omics

data could be applied to the data measured in the same cohort with a large sample size to find the co-regulations behind (**Table 2**). For instance, eQTMs (associations between methylation and gene expression) provide a resource to integrate methylation and gene expression. Highly methylation can block the binding of transcription factors on promoters and enhancers. In line with expectation, most eQTMs showed negative correlations between methylation and gene expression, and negatively correlated eQTMs are enriched in active TSS regions (152). For another example, a study carefully characterizes the changes in the gut microbiota of patients suffering inflammatory bowel diseases and the interplay between microbiome composition and gut metabolites (153).

In the situation of a more complex multi-omics integration, more advanced technique like building multivariable regression model could take features from different omics to evaluate the accumulative effects/prediction power on a certain phenotype. An example study integrates genomic, metagenomic, metabolomic, immune cell composition, hormone levels and platelet activation profiles with cytokine response profiles in a population-based cohort. Results from multivariable linear regression and machine learning approaches such as elastic net show the accumulative contribution and predict power of genetic and non-genetic factors on cytokine response (154).

On the other hand, if the sample size is not allowed for association analysis, it might be applicable to check the intersections between the findings from different omics. For example, we could easily compare the regions identified in ATAC-seq, methylation array and Hi-C data. In addition, by matching a DAR to genes, and intersecting with DEGs, we could further check whether an epigenetic change has the potential in regulating gene expression.

DISCUSSION AND PERSPECTIVES

In this review, we have discussed the multi-omics application for immunological studies, from measurements and analysis to comparison or association of several typical layers (**Figure 2**). For system studies – in particular newly discovered infectious diseases or rare diseases with fewer prior knowledge – the choice

TABLE 2 | System analysis between omics.

	binary traits	epigenetics	gene expresion	protein level	metabolites	microbiome	cellomics
genetics	GWAS	meQTL, CRDQTLs	eQTL, sQTL	pQTL	mQTL	mbQTL	cell proportion QTLs
epigenetics	DMRs/DARs/Compartment Switches/ Gained or lost Interactions	position-based overlap	gene-based overlap/ association	gene-based overlap/ association	association	association	association
gene expression	DEGs	–	co-expression	gene-based overlap/ association	association	association	association
protein level	DEPs	–	–	coexpression/ interaction	–	association	association
metabolites	different abundance	–	–	–	association	association	association
microbiome	different composition	–	–	–	–	association	–
cellomics	Different cell composition, etc.	–	–	–	–	–	association

FIGURE 2 | Central dogma and regulations of different omics layers.
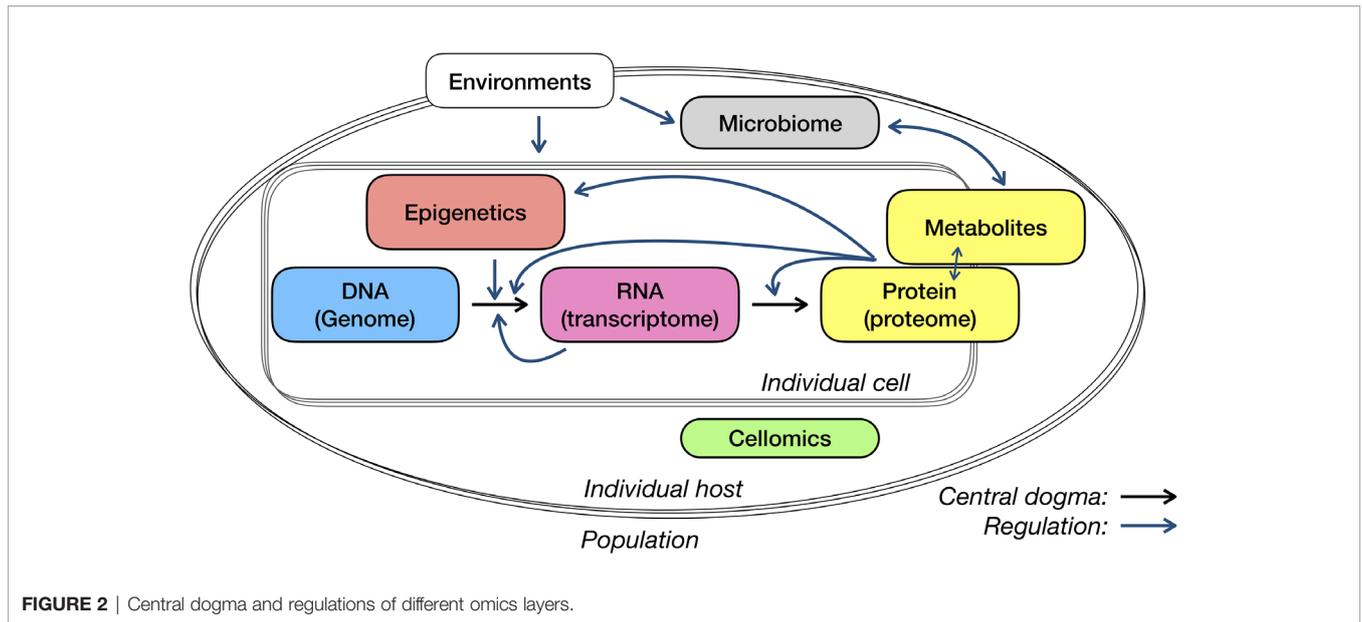
of data layers to collect and the selection of measuring approaches, target or non-target technique, bulk or single-cell level, can be as important as the analysis models and algorithms. Here, we discuss a few points that need specific attention in study design and interpretation, and subjects may see progress in the next few years.

There are some commonly used strategies of interpreting genetic associations. As the starting point of the central dogma of molecular biology (**Figure 2**), genetics has so far received a lot of attention and was associated with many types of data or phenotypes. In the interpretation of genome-wide associated loci, genes around them have also been regarded as the necessary and most essential compartments. The strategy to properly link loci with affected genes so far has been addressed on the position and associations between gene expression and genetic variants (i.e., eQTLs). In addition, functional annotation on identified loci, such as whether the variants are located on the regulatory elements or affected protein structure, may provide additional clues for loci interpretation in particular cases. Nevertheless, there are existing debates upon several aspects, for example, whether the host genome could influence the gut microbiome. It will never be nitpicking to be very careful with interpreting your microbiome QTLs.

Epigenetic could be used as a window to study environmental influence. In contrast to genetics, epigenetics often links the external factors to immune phenotypes. This is particularly true when considering the external effects as a risk to immune diseases, for example, smoking to asthma, because epigenetic modifications, such as methylation, are usually related to environmental exposures. Considering the various kinds of epigenetic changes, multiple types of epigenetic data are commonly used in one study and they often validate and complement each other. For example, an active TSS region could be identified by low methylation as well as high DNA accessibility (155), and the enhancer involved in a neo chromatin

interaction identified in Hi-C data could be characterized as a neo opening region in ATAC data (156). Considering the functional relationships, epigenetic data is commonly integrated with gene expression measurements. As the direct consequence of epigenetic modification, alteration in corresponding gene expression could be the best validation of the importance of your epigenetic studies.

scRNA-seq is usually applied together with Cellomics measurements. A cell composition discovered in scRNA-seq data could be validated with FCM-based approaches. FACS is also commonly used as a pre-filtering step to help with concentrating target cell types for scRNA-seq analysis. Especially, for the rare cell types (e.g., T regs in PBMCs), a pre-sorting process is necessary for concentrating on cells of interest.

Proteome, metabolome showing downstream immune functions require more attentions. As the downstream products of gene expression, protein or metabolites level measurements are not as popular as transcriptome measurements in current studies. This might because gene expression analysis takes advantage of the efficiency of next-generation sequencing and well-established microarray chips. Thus, there appears to be much room for further studies on proteome and metabolome in immune studies.

Proper measurement techniques and sampling tissues are crucial in an omics study. When considering the purpose of measurements, it is often appropriate to apply high-throughput and/or non-target approaches at the discovery stage, while single and/or target approaches are more commonly used for validation. Besides, except genome, all the other omics have tissue specificity. Data from the same tissue are more commonly associated. For example, associations between omics from blood samples could be easily interpreted, but it would be tricky and needs more biological basis to associate blood features with gut features.

A straight-forward joint visualization of multi-omics data is another challenge to better present and understand the

interconnections across molecular layers as well as to fully utilizing the increasingly available omics data. Integrated tools or platforms that combined a comprehensive analysis workflow and interactive visualizations were often more preferable to researchers. Some examples are: PaintOmics3 (157) and Metascape (158), which provide powerful online frameworks for the multi-omics pathway analysis and visualization; Seurat (56), which focuses on analysis and visualization of single-cell omics data and supports easy connections to other popular analysis tools; and Omics Playground (159), who provides a user-friendly and interactive self-service bioinformatics platform for analysis, visualization and interpretation of transcriptomics and proteomics data. Moreover, trials of combing data sharing and interactive visualization along with research publication have also been made to improve the data dissemination. For example, by accessing to Immgen (160), FastGenomics (161) or DeCovid (58), researchers can explore and visualize their interested immune signatures on the COVID-19 datasets, which significantly increases impact of the studies.

To fully elucidate the biological processes involved in the immune system, several aspects remain unknown in omics studies. Firstly, due to sample accessibility, fewer studies have been performed on tissues other than blood. Taking meQTLs as an example, several big studies have been carried out blood samples (101, 162, 163). However, there are very limited sample size and/or studies about meQTLs in other tissues (164). Secondly, considering the high dynamics, rapid response and spatial specificity of the immune system, temporal and spatial studies can provide more insights into the dynamic process and spatial heterogeneity in immune activities and/or immune-related diseases etiology. For example, the process that immune cells are activated by interacting physically and chemically with synapses is highly dynamic and depends on the spatial position of immune cells, neurons and glial cells. Despite its importance in immune functionality and immune-mediated diseases, our current knowledge is not sufficiently advanced, which calls for more comprehensive studies (165–167). Thirdly, as for population-based studies, there are much more of them in healthy individuals of European ancestry, while the studies in under-represented populations as well as in patients appeal for greater attention.

Considering the complexity of our immune system and patient heterogeneity, in terms of severity or treatment responses, for many immune-related diseases, the generation of personalized medicine is one of the most significant goals we can achieve through multi-omics studies (168). Personalized medicine stratifies a heterogeneous group of patients based on certain characteristics and provides treatment based on this stratification. In the case of infectious diseases, one of the personalized medicine trials is now being conducted for the treatment of sepsis using immunomodulatory interventions after stratification based on biomarkers identifying immunosuppression or hyper inflammation (169). In the field of tuberculosis, advances are being made too, as a clinical trial is now ongoing where tuberculous meningitis patients are being stratified based on genotype prior to treatment (170).

In conclusion, we systematically review measurements and analyses can be applied in immunological studies, which provide insights for personalized medicine. Through the development of high throughput techniques, e.g. single-cell RNA sequencing and mass cytometry, we now possess the tools to unravel the many complexities of the immune system in health and immune-related diseases, including infectious diseases, allergies and auto-immune diseases. With unbiased measurements and effective integration, multi-omics studies can help us understand the immune system and could lead to the development of personalized medicine.

AUTHOR CONTRIBUTIONS

XC made the conception and design of the review. XC and BZ drafted the manuscripts, supervised by YL. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

1. Topol EJ, Califf RM. *Textbook of cardiovascular medicine*. Philadelphia, USA: Lippincott Williams & Wilkins (2007).
2. Atkinson MA, Eisenbarth GS, Michels AW. Type 1 Diabetes. *Lancet* (2014) 383(9911):69–82. doi: 10.1016/S0140-6736(13)60591-7
3. Reyes M, Filbin MR, Bhattacharyya RP, Billman K, Eisenhaure T, Hung DT, et al. An Immune-Cell Signature of Bacterial Sepsis. *Nat Med* (2020) 26 (3):333–40. doi: 10.1038/s41591-020-0752-4
4. Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and Effectiveness of Influenza Vaccines: A Systematic Review and Meta-Analysis. *Lancet Infect Dis* (2012) 12(1):36–44. doi: 10.1016/S1473-3099 (11)70295-X
5. Warren RB, Smith CH, Yiu ZZ, Ashcroft DM, Barker JN, Burden AD, et al. Differential Drug Survival of Biologic Therapies for the Treatment of Psoriasis: A Prospective Observational Cohort Study From the British

Association of Dermatologists Biologic Interventions Register (Badbir). *J Invest Dermatol* (2015) 135(11):2632–40. doi: 10.1038/jid.2015.208
6. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic Mechanisms of Critical Illness in Covid-19. *Nature* (2020) 591.7848:92–98.. doi: 10.1101/2020.09.24.20200048
7. Xiong Y, Liu Y, Cao L, Wang D, Guo M, Jiang A, et al. Transcriptomic Characteristics of Bronchoalveolar Lavage Fluid and Peripheral Blood Mononuclear Cells in COVID-19 Patients. *Emerg Microbes infections* (2020) 9(1):761–70. doi: 10.1080/22221751.2020.1747363
8. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* (2020) 182(1):59–72.e15. doi: 10.1016/j.cell.2020.05.032
9. Schulte-Schrepping J, Reusch N, Paclik D, Baßler K, Schlickeiser S, Zhang B, et al. Severe COVID-19 is Marked by a Dysregulated Myeloid Cell Compartment. *Cell* (2020) 182(6):1419–40. doi: 10.1016/j.cell. 2020.08.001

10. Bernardes JP, Mishra N, Tran F, Bahmer T, Best L, Blase JI, et al. Longitudinal Multi-Omics Analyses Identify Responses of Megakaryocytes, Erythroid Cells, and Plasmablasts as Hallmarks of Severe Covid-19. *Immunity* (2020) 53(6):1296–314. doi: 10.1016/j.immuni.2020.11.017

11. Zuo T, Zhan H, Zhang F, Liu Q, Tso EY, Lui GC, et al. Alterations in Fecal Fungal Microbiome of Patients With COVID-19 During Time of Hospitalization Until Discharge. *Gastroenterology* (2020) 159(4):1302–10. doi: 10.1053/j.gastro.2020.06.048

12. Erkelens MN, Mebius RE. Retinoic Acid and Immune Homeostasis: A Balancing Act. *Trends Immunol* (2017) 38(3):168–80. doi: 10.1016/j.it.2016.12.006

13. Kriete A, Eils R. *Computational Systems Biology: From Molecular Mechanisms to Disease*. San Diego, USA: Academic Press (2013).

14. Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. *Genes* (2019) 10(2):87. doi: 10.3390/genes10020087

15. Jaumot J, Bedia C, Tauler R. *Data Analysis for Omic Sciences: Methods and Applications*. Amsterdam: Elsevier (2018).

16. Song M, Greenbaum J, Luttrell IVJ, Zhou W, Wu C, Shen H, et al. A Review of Integrative Imputation for Multi-Omics Datasets. *Front Genet* (2020) 11:570255. doi: 10.3389/fgene.2020.570255

17. Eckhardt M, Hultquist JF, Kaake RM, Hüttenhain R, Krogan NJ. A Systems Approach to Infectious Disease. *Nat Rev Genet* (2020) 21:339–54. doi: 10.1038/s41576-020-0212-5.

18. Savola P, Martelius T, Kankainen M, Koski Y, Eldfors S, Huuhtanen J, et al. Somatic Mutations in T Cells as Possible Regulators of Immunodeficiency. *Blood* (2018) 132(Supplement 1):515–. doi: 10.1182/blood-2018-99-110757

19. Netea MG, Joosten LA, Li Y, Kumar V, Oosting M, Smeekens S, et al. Understanding Human Immune Function Using the Resources From the Human Functional Genomics Project. *Nat Med* (2016) 22(8):831–3. doi: 10.1038/nm.4140

20. Cortes A, Brown MA. Promise and Pitfalls of the Immunochip. *Arthritis Res Ther* (2011) 13(1):101. doi: 10.1186/ar3204

21. Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, et al. The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, and Anthropometric Traits. *PloS Genet* (2012) 8(8): e1002793. doi: 10.1371/journal.pgen.1002793

22. Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, et al. Concept, Design and Implementation of a Cardiovascular Gene-Centric 50 K Snp Array for Large-Scale Genomic Association Studies. *PloS One* (2008) 3 (10):e3583. doi: 10.1371/journal.pone.0003583

23. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-Generation Genotype Imputation Service and Methods. *Nat Genet* (2016) 48 (10):1284–7. doi: 10.1038/ng.3656

24. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data Quality Control in Genetic Case-Control Association Studies. *Nat Protoc* (2010) 5(9):1564–73. doi: 10.1038/nprot.2010.116

25. Siva N. 1000 Genomes Project. *Nat Biotechnol* (2008) 26(3):256–. doi: 10.1038/nbt0308-256b

26. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* (2016) 167(5):1415–29.e19. doi: 10.1016/j.cell.2016.10.042

27. Emdin CA, Khera AV, Chaffin M, Klarin D, Natarajan P, Aragam K, et al. Analysis of Predicted Loss-of-Function Variants in UK Biobank Identifies Variants Protective for Disease. *Nat Commun* (2018) 9(1):1613. doi: 10.1038/s41467-018-03911-8.

28. Ferraro NM, Strober BJ, Einson J, Abell NS, Aguet F, Barbeira AN, et al. Transcriptomic Signatures Across Human Tissues Identify Functional Rare Genetic Variation. *Science* (2020) 369(6509):eaaz5900. doi: 10.1126/science.aaz5900

29. Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, Menni C, et al. Whole-Genome Sequencing Identifies Common-to-Rare Variants Associated With Human Blood Metabolites. *Nat Genet* (2017) 49(4):568–78. doi: 10.1038/ng.3809

30. Lund G, Andersson L, Lauria M, Lindholm M, Fraga MF, Villar-Garea A, et al. Dna Methylation Polymorphisms Precede Any Histological Sign of Atherosclerosis in Mice Lacking Apolipoprotein E. *J Biol Chem* (2004) 279 (28):29147–54. doi: 10.1074/jbc.M403618200

31. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences. *nature* (2009) 462(7271):315–22. doi: 10.1038/nature08514

32. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced Representation Bisulfite Sequencing for Comparative High-Resolution DNA Methylation Analysis. *Nucleic Acids Res* (2005) 33 (18):5868–77. doi: 10.1093/nar/gki901

33. Masser DR, Berg AS, Freeman WM. Focused, High Accuracy 5-Methylcytosine Quantitation With Base Resolution by Benchtop Next-Generation Sequencing. *Epigenet chromatin* (2013) 6(1):1–12. doi: 10.1186/1756-8935-6-33

34. Mallik S, Odom GJ, Gao Z, Gomez L, Chen X, Wang L. An Evaluation of Supervised Methods for Identifying Differentially Methylated Regions in Illumina Methylation Arrays. *Briefings Bioinf* (2019) 20(6):2224–35. doi: 10.1093/bib/bby085

35. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical Evaluation of the Illumina Methylationepic Beadchip Microarray for Whole-Genome Dna Methylation Profiling. *Genome Biol* (2016) 17(1):1–17. doi: 10.1186/s13059-016-1066-1

36. Greer EL, Shi Y. Histone Methylation: A Dynamic Mark in Health, Disease and Inheritance. *Nat Rev Genet* (2012) 13(5):343–57. doi: 10.1038/nrg3173

37. Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH. An Integrated Software System for Analyzing Chip-Chip and Chip-Seq Data. *Nat Biotechnol* (2008) 26(11):1293–300. doi: 10.1038/nbt.1505

38. Chen H, Lareau C, Andreani T, Vinyard ME, Garcia SP, Clement K, et al. Assessment of Computational Methods for the Analysis of Single-Cell Atac-Seq Data. *Genome Biol* (2019) 20(1):1–25. doi: 10.1186/s13059-019-1854-5

39. Meyer CA, Liu XS. Identifying and Mitigating Bias in Next-Generation Sequencing Methods for Chromatin Biology. *Nat Rev Genet* (2014) 15 (11):709–21. doi: 10.1038/nrg3788

40. Ernst J, Kellis M. Large-Scale Imputation of Epigenomic Datasets for Systematic Annotation of Diverse Human Tissues. *Nat Biotechnol* (2015) 33(4):364–76. doi: 10.1038/nbt.3157

41. Kapourani C-A, Sanguinetti G. Melissa: Bayesian Clustering and Imputation of Single-Cell Methylomes. *Genome Biol* (2019) 20(1):1–15. doi: 10.1186/s13059-019-1665-8

42. Schreiber J, Durham T, Bilmes J, Noble WS. Avocado: A Multi-Scale Deep Tensor Factorization Method Learns a Latent Representation of the Human Epigenome. *Genome Biol* (2020) 21(1):1–18. doi: 10.1186/s13059-020-01977-6

43. Xiong L, Xu K, Tian K, Shao Y, Tang L, Gao G, et al. Scale Method for Single-Cell Atac-Seq Analysis Via Latent Feature Extraction. *Nat Commun* (2019) 10(1):1–10. doi: 10.1038/s41467-019-12630-7

44. Kempfer R, Pombo A. Methods for Mapping 3d Chromosome Architecture. *Nat Rev Genet* (2020) 21(4):207–26. doi: 10.1038/s41576-019-0195-2

45. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-Cell Hi-C Reveals Cell-to-Cell Variability in Chromosome Structure. *Nature* (2013) 502(7469):59–64. doi: 10.1038/nature12593

46. Beagrie RA, Scialdone A, Schueler M, Kraemer DC, Chotalia M, Xie SQ, et al. Complex Multi-Enhancer Contacts Captured by Genome Architecture Mapping. *Nature* (2017) 543(7646):519–24. doi: 10.1038/nature21411

47. Vangala P, Murphy R, Quinodoz SA, Gellatly K, McDonel P, Guttman M, et al. High-Resolution Mapping of Multiway Enhancer-Promoter Interactions Regulating Pathogen Detection. *Mol Cell* (2020) 80(2):359–73. doi: 10.1016/j.molcel.2020.09.005

48. Koch L. Getting the Drop on Chromatin Interaction. *Nat Rev Genet* (2019) 20(4):192–3. doi: 10.1038/s41576-019-0103-9

49. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-Seq2 for Sensitive Full-Length Transcriptome Profiling in Single Cells. *Nat Methods* (2013) 10(11):1096–8. doi: 10.1038/nmeth.2639

50. Haque A, Engel J, Teichmann SA, Lönnberg T. A Practical Guide to Single-Cell Rna-Sequencing for Biomedical Research and Clinical Applications. *Genome Med* (2017) 9(1):1–12. doi: 10.1186/s13073-017-0467-4

51. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D. Evaluation of Two Main RNA-Seq Approaches for Gene Quantification in Clinical Rna Sequencing: Polya+ Selection Versus Rrna Depletion. *Sci Rep* (2018) 8 (1):1–12. doi: 10.1038/s41598-018-23226-4

52. Park J-E, Botting RA, Conde CD, Popescu D-M, Lavaert M, Kunz DJ, et al. A Cell Atlas of Human Thymic Development Defines T Cell Repertoire Formation. *Science* (2020) 367(6480). doi: 10.1101/2020.01.28.911115

53. Herzog VA, Reichholf B, Neumann T, Rescheneder P, Bhat P, Burkard TR, et al. Thiol-Linked Alkylation of RNA to Assess Expression Dynamics. *Nat Methods* (2017) 14(12):1198–204. doi: 10.1038/nmeth.4435

54. Erhard F, Baptista MA, Krammer T, Hennig T, Lange M, Arampatzi P, et al. Scslam-Seq Reveals Core Features of Transcription Dynamics in Single Cells. *Nature* (2019) 571(7765):419–23. doi: 10.1038/s41586-019-1369-y

55. Chen W, Zhao Y, Chen X, Yang Z, Xu X, Bi Y, et al. A Multicenter Study Benchmarking Single-Cell RNA Sequencing Technologies Using Reference Samples. *Nat Biotechnol* (2020) 1–12. doi: 10.1038/s41587-020-00748-9

56. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WMIII, et al. Comprehensive Integration of Single-Cell Data. *Cell* (2019) 177(7):1888–902.e21. doi: 10.1016/j.cell.2019.05.031

57. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, Sensitive and Accurate Integration of Single-Cell Data With Harmony. *Nat Methods* (2019) 16:1289–96. doi: 10.1038/s41592-019-0619-0

58. Liu T, Balzano-Nogueira L, Lleo A, Conesa A. Transcriptional Differences for COVID-19 Disease Map Genes Between Males and Females Indicate a Different Basal Immunophenotype Relevant to the Disease. *Genes* (2020) 11 (12):1447. doi: 10.3390/genes11121447

59. Wang X, Park J, Susztak K, Zhang NR, Li M. Bulk Tissue Cell Type Deconvolution With Multi-Subject Single-Cell Expression Reference. *Nat Commun* (2019) 10(1):1–9. doi: 10.1038/s41467-018-08023-x

60. Aguirre-Gamboa R, de Klein N, di Tommaso J, Claringbould A, van der Wijst MG, de Vries D, et al. Deconvolution of Bulk Blood Eqtl Effects Into Immune Cell Subpopulations. *BMC Bioinf* (2020) 21(1):1–23. doi: 10.1186/s12859-020-03576-5

61. Eraslan G, Simon LM, Mircea M, Mueller NS, Theis FJ. Single-Cell RNA-Seq Denoising Using a Deep Count Autoencoder. *Nat Commun* (2019) 10(1):1–14. doi: 10.1038/s41467-018-07931-2

62. Van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, et al. Recovering Gene Interactions From Single-Cell Data Using Data Diffusion. *Cell* (2018) 174(3):716–29. doi: 10.1016/j.cell.2018.05.061

63. Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM, et al. Cell Hashing With Barcoded Antibodies Enables Multiplexing and Doublet Detection for Single Cell Genomics. *Genome Biol* (2018) 19(1):1–12. doi: 10.1186/s13059-018-1603-1

64. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Large-Scale Simultaneous Measurement of Epitopes and Transcriptomes in Single Cells. *Nat Methods* (2017) 14(9):865. doi: 10.1038/nmeth.4380

65. Katzenelenbogen Y, Sheban F, Yalin A, Yofe I, Svetlichnyy D, Jaitin DA, et al. Coupled Scrna-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer. *Cell* (2020) 182(4):872–85.e19. doi: 10.1016/j.cell.2020.06.032

66. Kumar NG, Contaifer D, Madurantakam P, Carbone S, Price ET, Van Tassell B, et al. Dietary Bioactive Fatty Acids as Modulators of Immune Function: Implications on Human Health. *Nutrients* (2019) 11(12):2974. doi: 10.3390/nu11122974

67. Loftus RM, Finlay DK. Immunometabolism: Cellular Metabolism Turns Immune Regulator. *J Biol Chem* (2016) 291(1):1–10. doi: 10.1074/jbc.R115.693903

68. Barba I, Fernandez-Montesinos R, Garcia-Dorado D, Pozo D. Alzheimer's Disease Beyond the Genomic Era: Nuclear Magnetic Resonance (Nmr) Spectroscopy-Based Metabolomics. *J Cell Mol Med* (2008) 12(5a):1477–85. doi: 10.1111/j.1582-4934.2008.00385.x

69. Dettmer K, Aronov PA, Hammock BD. Mass Spectrometry-Based Metabolomics. *Mass spectrometry Rev* (2007) 26(1):51–78. doi: 10.1002/mas.20108

70. Gorrochategui E, Jaumot J, Tauler R. Roimcr: A Powerful Analysis Strategy for LC-MS Metabolomic Datasets. *BMC Bioinf* (2019) 20(1):1–17. doi: 10.1186/s12859-019-2848-8

71. Fernández-Ochoa Á, Brunius C, Borrás-Linares I, Quirantes-Piné R, Cádiz-Gurrea M, Consortium PC, et al. Metabolic Disturbances in Urinary and Plasma Samples From Seven Different Systemic Autoimmune Diseases Detected by HPLC-ESI-QTOF-MS. *J Proteome Res* (2020) 19(8):3220–9. doi: 10.1021/acs.jproteome.0c00179

72. Kolmert J, Gómez C, Balgoma D, Sjödin M, Bood J, Konradsen JR, et al. Urinary Leukotriene E4 and Prostaglandin D2 Metabolites Increase in Adult and Childhood Severe Asthma Characterized by Type 2 Inflammation. A Clinical Observational Study. *Am J Respir Crit Care Med* (2021) 203(1):37–53. doi: 10.1164/rccm.202101-0208LE

73. Souter I, Bellavia A, Williams PL, Korevaar T, Meeker JD, Braun JM, et al. Urinary Concentrations of Phthalate Metabolite Mixtures in Relation to Serum Biomarkers of Thyroid Function and Autoimmunity Among Women From a Fertility Center. *Environ Health Perspect* (2020) 128(6):067007. doi: 10.1289/EHP6740

74. Bar N, Korem T, Weissbrod O, Zeevi D, Rothschild D, Leviatan S, et al. A Reference Map of Potential Determinants for the Human Serum Metabolome. *Nature* (2020) 588(7836):135–40. doi: 10.1038/s41586-020-2896-2

75. Al Bander Z, Nitert MD, Mousa A, Naderpoor N. The Gut Microbiota and Inflammation: An Overview. *Int J Environ Res Public Health* (2020) 17 (20):7618. doi: 10.3390/ijerph17207618

76. Fitzgibbon G, Mills KH. The Microbiota and Immune-Mediated Diseases: Opportunities for Therapeutic Intervention. *Eur J Immunol* (2020) 50 (3):326–37. doi: 10.1002/eji.201948322

77. Jiao Y, Wu L, Huntington ND, Zhang X. Crosstalk Between Gut Microbiota and Innate Immunity and its Implication in Autoimmune Diseases. *Front Immunol* (2020) 11:282. doi: 10.3389/fimmu.2020.00282

78. Zhang X, Li L, Butcher J, Stintzi A, Figeys D. Advancing Functional and Translational Microbiome Research Using Meta-Omics Approaches. *Microbiome* (2019) 7(1):1–12. doi: 10.1186/s40168-019-0767-6

79. Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host Variables Confound Gut Microbiota Studies of Human Disease. *Nature* (2020) 587(7834):448–54. doi: 10.1038/s41586-020-2881-9

80. Frølund M, Wikström A, Lidbrink P, Abu Al-Soud W, Larsen N, Harder CB, et al. The Bacterial Microbiota in First-Void Urine From Men With and Without Idiopathic Urethritis. *PloS One* (2018) 13(7):e0201380. doi: 10.1371/journal.pone.0201380

81. Winters BR, Pleil JD, Angrish MM, Stiegel MA, Risby TH, Madden MC. Standardization of the Collection of Exhaled Breath Condensate and Exhaled Breath Aerosol Using a Feedback Regulated Sampling Device. *J breath Res* (2017) 11(4):047107. doi: 10.1088/1752-7163/aa8bbc

82. Cheung RK, Utz PJ. Cytof—the Next Generation of Cell Detection. *Nat Rev Rheumatol* (2011) 7(9):502–3. doi: 10.1038/nrrheum.2011.110

83. Pal A, Glaß H, Naumann M, Kreiter N, Japtok J, Sczech R, et al. High Content Organelle Trafficking Enables Disease State Profiling as Powerful Tool for Disease Modelling. *Sci Data* (2018) 5(1):1–15. doi: 10.1038/sdata.2018.241

84. Kiyoi T, Liu S, Sahid MNA, Shudou M, Maeyama K, Mogi M. High-Throughput Screening System for Dynamic Monitoring of Exocytotic Vesicle Trafficking in Mast Cells. *PloS One* (2018) 13(6):e0198785. doi: 10.1371/journal.pone.0198785

85. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the Multiple Testing Burden for Genomewide Association Studies of Nearly All Common Variants. *Genet Epidemiol* (2008) 32(4):381–5. doi: 10.1002/gepi.20303

86. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI Gwas Catalog of Published Genome-Wide Association Studies, Targeted Arrays and Summary Statistics 2019. *Nucleic Acids Res* (2019) 47(D1):D1005–12. doi: 10.1093/nar/gky1120

87. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic Identification of Trans Eqtls as Putative Drivers of Known Disease Associations. *Nat Genet* (2013) 45(10):1238–43. doi: 10.1038/ng.2756

88. Consortium G. Genetic Effects on Gene Expression Across Human Tissues. *Nature* (2017) 550(7675):204–13. doi: 10.1038/nature24277

89. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and Phenomic Insights From an Atlas of Genetic Effects on DNA Methylation. *medRxiv* (2020) 1:1–30. doi: 10.1101/2020.09.01.20180406

90. Xu C-J, Bonder MJ, Söderhäll C, Bustamante M, Baïz N, Gehring U, et al. The Emerging Landscape of Dynamic Dna Methylation in Early Childhood. *BMC Genomics* (2017) 18(1):1–11. doi: 10.1186/s12864-016-3452-1

91. Li Y, Oosting M, Deelen P, Ricaño-Ponce I, Smeekens S, Jaeger M, et al. Inter-Individual Variability and Genetic Influences on Cytokine Responses to Bacteria and Fungi. *Nat Med* (2016) 22(8):952–60. doi: 10.1038/nm.4139

92. Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KT, et al. A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell* (2016) 167(4):1099–110.e14. doi: 10.1016/j.cell.2016.10.017

93. Hellwege JN, Keaton JM, Giri A, Gao X, Velez Edwards DR, Edwards TL. Population Stratification in Genetic Association Studies. *Curr Protoc Hum Genet* (2017) 95(1):1. doi: 10.1002/cphg.48

94. Martin ER, Tunc I, Liu Z, Slifer SH, Beecham AH, Beecham GW. Properties of Global-and Local-Ancestry Adjustments in Genetic Association Tests in Admixed Populations. *Genet Epidemiol* (2018) 42(2):214–29. doi: 10.1002/gepi.22103

95. Gamazon ER, Segrè AV, van de Bunt M, Wen X, Xi HS, Hormozdiari F, et al. Using an Atlas of Gene Regulation Across 44 Human Tissues to Inform

Complex Disease-and Trait-Associated Variation. *Nat Genet* (2018) 50 (7):956–67. doi: 10.1038/s41588-018-0154-4

96. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A Gene-Based Association Method for Mapping Traits Using Reference Transcriptome Data. *Nat Genet* (2015) 47(9):1091. doi: 10.1038/ng.3367

97. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic Localization of Common Disease-Associated Variation in Regulatory Dna. *Science* (2012) 337(6099):1190–5. doi: 10.1126/science.1222794

98. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA Elements (Encode): Data Portal Update. *Nucleic Acids Res* (2018) 46(D1):D794–801. doi: 10.1093/nar/gkx1081

99. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative Analysis of 111 Reference Human Epigenomes. *Nature* (2015) 518(7539):317–30 doi: 10.1038/nature14248

100. Delaneau O, Zazhytska M, Borel C, Giannuzzi G, Rey G, Howald C, et al. Chromatin Three-Dimensional Interactions Mediate Genetic Effects on Gene Expression. *Science* (2019) 364(6439):eaat8266. doi: 10.1126/science.aat8266

101. McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE, et al. Identification of 55,000 Replicated Dna Methylation Qtl. *Sci Rep* (2018) 8 (1):1–9. doi: 10.1038/s41598-018-35871-w

102. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic Atlas of the Human Plasma Proteome. *Nature* (2018) 558 (7708):73–9. doi: 10.1038/s41586-018-0175-2

103. Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J, Santos R, Huang J, et al. An Atlas of Genetic Influences on Human Blood Metabolites. *Nat Genet* (2014) 46(6):543–50. doi: 10.1038/ng.2982

104. Nath AP, Ritchie SC, Byars SG, Fearnley LG, Havulinna AS, Joensuu A, et al. An Interaction Map of Circulating Metabolites, Immune Gene Networks, and Their Genetic Regulation. *Genome Biol* (2017) 18(1):1–15. doi: 10.1186/s13059-017-1279-y

105. Aguirre-Gamboa R, Joosten I, Urbano PC, van der Molen RG, van Rijssen E, van Cranenbroek B, et al. Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits. *Cell Rep* (2016) 17(9):2474–87. doi: 10.1016/j.celrep.2016.10.053

106. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian Test for Colocalisation Between Pairs of Genetic Association Studies Using Summary Statistics. *PloS Genet* (2014) 10(5): e1004383. doi: 10.1371/journal.pgen.1004383

107. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An Atlas of Genetic Correlations Across Human Diseases and Traits. *Nat Genet* (2015) 47(11):1236. doi: 10.1038/ng.3406

108. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base Platform Supports Systematic Causal Inference Across the Human Phenome. *Elife* (2018) 7:e34408. doi: 10.7554/eLife.34408

109. Chen L, Ge B, Casale FP, Vasquez L, Kwan T, Garrido-Martín D, et al. Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. *Cell* (2016) 167(5):1398–414. doi: 10.1016/j.cell.2016.10.026

110. Rosa M, Chignon A, Li Z, Boulanger M-C, Arsenault BJ, Bossé Y, et al. A Mendelian Randomization Study of IL6 Signaling in Cardiovascular Diseases, Immune-Related Disorders and Longevity. *NPJ genomic Med* (2019) 4(1):1–10. doi: 10.1038/s41525-019-0097-4

111. McGowan LM, Davey Smith G, Gaunt TR, Richardson TG. Integrating Mendelian Randomization and Multiple-Trait Colocalization to Uncover Cell-Specific Inflammatory Drivers of Autoimmune and Atopic Disease. *Hum Mol Genet* (2019) 28(19):3293–300. doi: 10.1093/hmg/ddz155

112. Baccarelli A, Bollati V. Epigenetics and Environmental Chemicals. *Curr Opin Pediatr* (2009) 21(2):243. doi: 10.1097/MOP.0b013e32832925cc

113. Martin DI, Cropley JE, Suter CM. Epigenetics in Disease: Leader or Follower? *Epigenetics* (2011) 6(7):843–8. doi: 10.4161/epi.6.7.16498

114. Ramos-Rodríguez M, Raurell-Vila H, Colli ML, Alvelos MI, Subirana-Granés M, Juan-Mateu J, et al. The Impact of Proinflammatory Cytokines on the β-Cell Regulatory Landscape Provides Insights Into the Genetics of Type 1 Diabetes. *Nat Genet* (2019) 51(11):1588–95. doi: 10.1101/560193

115. Netea MG, Joosten LA, Latz E, Mills KH, Natoli G, Stunnenberg HG, et al. Trained Immunity: A Program of Innate Immune Memory in Health and Disease. *Science* (2016) 352(6284):aaf1098. doi: 10.1126/science.aaf1098

116. Mazzone R, Zwergel C, Artico M, Taurone S, Ralli M, Greco A, et al. The Emerging Role of Epigenetics in Human Autoimmune Disorders. *Clin Epigenet* (2019) 11(1):1–15. doi: 10.1186/s13148-019-0632-2

117. Granja JM, Klemm S, McGinnis LM, Kathiria AS, Mezger A, Corces MR, et al. Single-Cell Multiomic Analysis Identifies Regulatory Programs in Mixed-Phenotype Acute Leukemia. *Nat Biotechnol* (2019) 37(12):1458–65. doi: 10.1038/s41587-019-0332-7

118. Ernst J, Kellis M. ChromHMM: Automating Chromatin-State Discovery and Characterization. *Nat Methods* (2012) 9(3):215–6. doi: 10.1038/nmeth.1906

119. Gjoneska E, Pfenning AR, Mathys H, Quon G, Kundaje A, Tsai L-H, et al. Conserved Epigenomic Signals in Mice and Humans Reveal Immune Basis of Alzheimer's Disease. *Nature* (2015) 518(7539):365–9. doi: 10.1038/nature14252

120. Cairns J, Freire-Pritchett P, Wingett SW, Várnai C, Dimond A, Plagnol V, et al. Chicago: Robust Detection of DNA Looping Interactions in Capture Hi-C Data. *Genome Biol* (2016) 17(1):1–17. doi: 10.1186/s13059-016-0992-2

121. Hu G, Cui K, Fang D, Hirose S, Wang X, Wangsa D, et al. Transformation of Accessible Chromatin and 3D Nucleome Underlies Lineage Commitment of Early T Cells. *Immunity* (2018) 48(2):227–42. doi: 10.1016/j.immuni.2018.01.013

122. Burren OS, García AR, Javierre B-M, Rainbow DB, Cairns J, Cooper NJ, et al. Chromosome Contacts in Activated T Cells Identify Autoimmune Disease Candidate Genes. *Genome Biol* (2017) 18(1):1–19. doi: 10.1186/s13059-017-1285-0

123. Chan WF, Coughlan HD, Zhou JH, Keenan CR, Bediaga NG, Hodgkin PD, et al. Pre-Mitotic Genome Re-Organisation Bookends the B Cell Differentiation Process. *Nat Commun* (2021) 12(1):1–13. doi: 10.1038/s41467-021-21536-2

124. Zhang J-Y, Wang X-M, Xing X, Xu Z, Zhang C, Song J-W, et al. Single-Cell Landscape of Immunological Responses in Patients With Covid-19. *Nat Immunol* (2020) 21(9):1107–18. doi: 10.1038/s41590-020-0762-x

125. Tian W, Zhang N, Jin R, Feng Y, Wang S, Gao S, et al. Immune Suppression in the Early Stage of COVID-19 Disease. *Nat Commun* (2020) 11(1):1–8. doi: 10.1038/s41467-020-19706-9

126. Wu TD, Madireddi S, de Almeida PE, Banchereau R, Chen Y-JJ, Chitre AS, et al. Peripheral T Cell Expansion Predicts Tumour Infiltration and Clinical Response. *Nature* (2020) 579(7798):274–8. doi: 10.1038/s41586-020-2056-8

127. Miller BC, Sen DR, Al Abosy R, Bi K, Virkud YV, LaFleur MW, et al. Subsets of Exhausted Cd8+ T Cells Differentially Mediate Tumor Control and Respond to Checkpoint Blockade. *Nat Immunol* (2019) 20(3):326–36. doi: 10.1038/s41590-019-0312-6

128. Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* (2019) 179(7):1636–46.e15. doi: 10.1016/j.cell.2019.11.003

129. Nakaya HI, Hagan T, Duraisingham SS, Lee EK, Kwissa M, Rouphael N, et al. Systems Analysis of Immunity to Influenza Vaccination Across Multiple Years and in Diverse Populations Reveals Shared Molecular Signatures. *Immunity* (2015) 43(6):1186–98. doi: 10.1016/j.immuni.2015.11.012

130. Conesa A, Nueda MJ, Ferrer A, Talón M. Masigpro: A Method to Identify Significantly Differential Expression Profiles in Time-Course Microarray Experiments. *Bioinformatics* (2006) 22(9):1096–102. doi: 10.1093/bioinformatics/btl056

131. Bouhaddani S, Houwing-Duistermaat J, Salo P, Perola M, Jongbloed G, Uh HW. Evaluation of O2PLS in Omics Data Integration. *BMC Bioinf* (2016) 17: S11. doi: 10.1186/s12859-015-0854-z.

132. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The Dynamics and Regulators of Cell Fate Decisions are Revealed by Pseudotemporal Ordering of Single Cells. *Nat Biotechnol* (2014) 32(4):381. doi: 10.1038/nbt.2859

133. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed Graph Embedding Resolves Complex Single-Cell Trajectories. *Nat Methods* (2017) 14(10):979. doi: 10.1038/nmeth.4402

134. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al. Rna Velocity of Single Cells. *Nature* (2018) 560(7719):494–8. doi: 10.1038/s41586-018-0414-6.

135. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. Covid-19 Severity Correlates With Airway Epithelium–Immune Cell Interactions Identified by Single-Cell Analysis. *Nat Biotechnol* (2020) 38(8):970–9. doi: 10.1038/s41587-020-0602-4

136. Langfelder P, Horvath S. Wgcna: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9(1):559. doi: 10.1186/1471-2105-9-559

137. Deelen P, van Dam S, Herkert JC, Karjalainen JM, Brugge H, Abbott KM, et al. Improving the Diagnostic Yield of Exome-Sequencing by Predicting Gene–Phenotype Associations Using Large-Scale Gene Expression Analysis. *Nat Commun* (2019) 10(1):2837. doi: 10.1038/s41467-019-10649-4.

138. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, et al. Innatedb: Systems Biology of Innate Immunity and Beyond—Recent Updates and Continuing Curation. *Nucleic Acids Res* (2013) 41(D1): D1228–D33. doi: 10.1093/nar/gks1147

139. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING Database in 2017: Quality-Controlled Protein–Protein Association Networks, Made Broadly Accessible. *Nucleic Acids Res* (2017) 45(D1):D362–8. doi: 10.1093/nar/gkw937

140. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. Cellphonedb: Inferring Cell–Cell Communication From Combined Expression of Multi-Subunit Ligand–Receptor Complexes. *Nat Protoc* (2020) 15(4):1484–506. doi: 10.1038/s41596-020-0292-x

141. Corridoni D, Antanaviciute A, Gupta T, Fawkner-Corbett D, Aulicino A, Jagielowicz M, et al. Single-Cell Atlas of Colonic Cd8+ T Cells in Ulcerative Colitis. *Nat Med* (2020) 26(9):1480–90. doi: 10.1038/s41591-020-1003-4

142. Browaeys R, Saelens W, Saeys Y. Nichenet: Modeling Intercellular Communication by Linking Ligands to Target Genes. *Nat Methods* (2020) 17(2):159–62. doi: 10.1038/s41592-019-0667-5

143. Bonnardel J, T'Jonck W, Gaublomme D, Browaeys R, Scott CL, Martens L, et al. Stellate Cells, Hepatocytes, and Endothelial Cells Imprint the Kupffer Cell Identity on Monocytes Colonizing the Liver Macrophage Niche. *Immunity* (2019) 51(4):638–54. doi: 10.1016/j.immuni.2019.08.017

144. Cullen CM, Aneja KK, Beyhan S, Cho CE, Woloszynek S, Convertino M, et al. Emerging Priorities for Microbiome Research. *Front Microbiol* (2020) 11:136. doi: 10.3389/fmicb.2020.00136

145. Dorrestein PC, Mazmanian SK, Knight R. From Microbiomess to Metabolomes to Function During Host-Microbial Interactions. *Immunity* (2014) 40(6):824. doi: 10.1016/j.immuni.2014.05.015

146. Hattori M, Okuno Y, Goto S, Kanehisa M. Development of a Chemical Structure Comparison Method for Integrated Analysis of Chemical and Genomic Information in the Metabolic Pathways. *J Am Chem Soc* (2003) 125 (39):11853–65. doi: 10.1021/ja036030u

147. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. Hmdb 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Res* (2018) 46(D1):D608–17. doi: 10.1093/nar/gkx1089

148. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. Metaboanalyst 4.0: Towards More Transparent and Integrative Metabolomics Analysis. *Nucleic Acids Res* (2018) 46(W1):W486–94. doi: 10.1093/nar/gky310

149. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-Level Functional Profiling of Metagenomes and Metatranscriptomes. *Nat Methods* (2018) 15(11):962–8. doi: 10.1038/s41592-018-0176-y

150. Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, et al. Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* (2016) 167(4):1125–36.e8. doi: 10.1016/j.cell.2016.10.020

151. Aden K, Rehman A, Waschina S, Pan W-H, Walker A, Lucio M, et al. Metabolic Functions of Gut Microbes Associate With Efficacy of Tumor Necrosis Factor Antagonists in Patients With Inflammatory Bowel Diseases. *Gastroenterology* (2019) 157(5):1279–92. doi: 10.1053/j.gastro.2019.07.025

152. Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, et al. Disease Variants Alter Transcription Factor Levels and Methylation of Their Binding Sites. *Nat Genet* (2017) 49(1):131–8. doi: 10.1038/ng.3721

153. Ananthakrishnan AN, Khalili H, Pan A, Higuchi LM, de Silva P, Richter JM, et al. Association Between Depressive Symptoms and Incidence of Crohn's Disease and Ulcerative Colitis: Results From the Nurses' Health Study. *Clin Gastroenterol Hepatol* (2013) 11(1):57–62. doi: 10.1016/j.cgh.2012.08.032

154. Bakker OB, Aguirre-Gamboa R, Sanna S, Oosting M, Smeekens SP, Jaeger M, et al. Integration of Multi-Omics Data and Deep Phenotyping Enables Prediction of Cytokine Responses. *Nat Immunol* (2018) 19(7):776–86. doi: 10.1038/s41590-018-0121-3

155. Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* (2007) 129(4):823–37. doi: 10.1016/j.cell.2007.05.009

156. Marco A, Meharena HS, Dileep V, Raju RM, Davila-Velderrain J, Zhang AL, et al. Mapping the Epigenomic and Transcriptomic Interplay During Memory Formation and Recall in the Hippocampal Engram Ensemble. *Nat Neurosci* (2020) 23:1606–17. doi: 10.1038/s41593-020-00717-0

157. Hernández-de-Diego R, Tarazona S, Martínez-Mira C, Balzano-Nogueira L, Furió-Tarí P, Pappas GJJr., et al. Paintomics 3: A Web Resource for the Pathway Analysis and Visualization of Multi-Omics Data. *Nucleic Acids Res* (2018) 46(W1):W503–9. doi: 10.1093/nar/gky466

158. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1–10. doi: 10.1038/s41467-019-09234-6

159. Akhmedov M, Martinelli A, Geiger R, Kwee I. Omics Playground: A Comprehensive Self-Service Platform for Visualization, Analytics and Exploration of Big Omics Data. *NAR Genomics Bioinf* (2020) 2(1):lqz019. doi: 10.1093/nargab/lqz019

160. Aguilar SV, Aguilar O, Allan R, Amir EAD, Angeli V, Artyomov MN, et al. Immgen at 15. *Nat Immunol* (2020) 21(7):700–3. doi: 10.1038/s41590-020-0687-4

161. Scholz CJ, Biernat P, Becker M, Baßler K, Günther P, Balfer J, et al. Fastgenomics: An Analytical Ecosystem for Single-Cell RNA Sequencing Data. *bioRxiv* (2018) 1:272476. doi: 10.1101/272476

162. Szymczak S, Dose J, Torres GG, Heinsen F-A, Venkatesh G, Datlinger P, et al. Dna Methylation Qtl Analysis Identifies New Regulators of Human Longevity. *Hum Mol Genet* (2020) 29(7):1154–67. doi: 10.1093/hmg/ddaa033

163. Huan T, Joehanes R, Song C, Peng F, Guo Y, Mendelson M, et al. Genome-Wide Identification of DNA Methylation Qtls in Whole Blood Highlights Pathways for Cardiovascular Disease. *Nat Commun* (2019) 10(1):1–14. doi: 10.1038/s41467-019-12228-z

164. Morrow J, Qiu W, Make B, Regan E, Han M, Hersh C, et al. DNA Methylation Is Predictive of Mortality in Current and Former Smokers. *Am J Respir Crit Care Med* (2020) 201(9):1099–109. doi: 10.1164/rccm.201902-0439OC

165. Carrier M, Robert M-È, González Ibáñez F, Desjardins M, Tremblay M-È. Imaging the Neuroimmune Dynamics Across Space and Time. *Front Neurosci* (2020) 14:903. doi: 10.3389/fnins.2020.00903

166. Chu C, Artis D, Chiu IM. Neuro-Immune Interactions in the Tissues. *Immunity* (2020) 52(3):464–74. doi: 10.1016/j.immuni.2020.02.017

167. Stakenborg N, Viola MF, Boeckxstaens GE. Intestinal Neuro-Immune Interactions: Focus on Macrophages, Mast Cells and Innate Lymphoid Cells. *Curr Opin Neurobiol* (2020) 62:68–75. doi: 10.1016/j.conb.2019.11.020

168. Delhalle S, Bode SF, Balling R, Ollert M, He FQ. A Roadmap Towards Personalized Immunology. *NPJ Syst Biol Appl* (2018) 4(1):1–14. doi: 10.1038/s41540-017-0045-9

169. Karakike E, Giamarellos-Bourboulis EJ. Macrophage Activation-Like Syndrome: A Distinct Entity Leading to Early Death in Sepsis. *Front Immunol* (2019) 10:55. doi: 10.3389/fimmu.2019.00055

170. Donovan J, Phu NH, Le Thi Phuong Thao NH, Lan NTHM, Trang NTM, Hiep NTT, et al. Adjunctive Dexamethasone for the Treatment of HIV-Uninfected Adults With Tuberculous Meningitis Stratified by Leukotriene A4 Hydrolase Genotype (Last ACT): Study Protocol for a Randomised Double Blind Placebo Controlled non-Inferiority Trial. *Wellcome Open Res* (2018) 3:32. doi: 10.12688/wellcomeopenres.14007.1