REVIEW

Single-cell transcriptomics to explore the immune system in health and disease

Michael J. T. Stubbington,[1]* Orit Rozenblatt-Rosen,[2]* Aviv Regev,[2,3]†‡ Sarah A. Teichmann[1,4]†‡

The immune system varies in cell types, states, and locations. The complex networks, interactions, and responses of immune cells produce diverse cellular ecosystems composed of multiple cell types, accompanied by genetic diversity in antigen receptors. Within this ecosystem, innate and adaptive immune cells maintain and protect tissue function, integrity, and homeostasis upon changes in functional demands and diverse insults. Characterizing this inherent complexity requires studies at single-cell resolution. Recent advances such as massively parallel single-cell RNA sequencing and sophisticated computational methods are catalyzing a revolution in our understanding of immunology. Here we provide an overview of the state of single-cell genomics methods and an outlook on the use of single-cell techniques to decipher the adaptive and innate components of immunity.

The immune system is composed of different cell lineages that reside in primary and secondary lymphoid organs and tissues throughout the body and that transit through the peripheral blood and lymphatic systems. The cells in these lineages are primary responders to changes in the environment, eliciting a complex network of intracellular circuits and intercellular interactions that result in transient responses within and between cells and cell states, more permanent differentiation choices, and flexible adaptation to their tissue of residence. Thus, the cells of each lymphoid and nonlymphoid tissue are key members of diverse cellular ecosystems composed of multiple immune and nonimmune cell types, which together maintain and protect tissue function, integrity, and homeostasis upon changes in functional demands, including insults and injuries. Hence, immunity involves innate and adaptive immune cells interacting with additional cells to form dynamic cellular communities in tissues.

Profiling the immune system

In seminal studies, immunologists have developed an extensive taxonomy of the cells of the immune system, integrating and unifying their functional characteristics, cell fate, and lineage relations with molecular markers. This effort was enabled by tools ranging from microscopy and flow cytometry to functional assays, animal models, and, most recently, genomics.

[1]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK. [2]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [3]Howard Hughes Medical Institute, Koch Institute for Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Theory of Condensed Matter, Cavendish Laboratory, 19 JJ Thomson Ave, Cambridge CB3 0HE, UK. *These authors contributed equally to this work. †These authors contributed equally to this work. ‡Corresponding author. Email: aregev@broadinstitute.org (A.R.); st9@sanger.ac.uk (S.A.T.)

However, the immune cell census remains incomplete. The immune system harbors a breadth of cell types and states, each of which can be at different stages of differentiation or response to environmental cues such as pathogens. In addition, because of the immune system's distributed nature, the "same" cell types and states are present in locations throughout the body but are modified by adaptations that reflect the distinctive niche and functional demand of their tissue of residence [reviewed in (*1*)]. Immune cells pose a further challenge: Lymphocytes with particular antigen receptor sequences [such as classical T and B cells but also invariant natural killer T (iNKT) cells, γδ T cells, and other populations] are clonal in nature, which introduces subtle yet important genetic diversity into these cell populations. Recent advances in single-cell genomics technologies are beginning to allow us to fill in these gaps by inspecting the immune system one cell at a time.

Technologies for characterization of the cells of the immune system

Over the years, three major techniques have established themselves for the categorization of immune cells. The most prominent is immunophenotyping through flow cytometry, which can identify cells of the immune system (while in suspension) by the single-cell expression of both cell-surface and intracellular proteins, including cytokines, and their posttranslational modifications [reviewed in (*2*)]. In addition, staining, sorting, and enrichment or depletion of specific viable cell subsets, including rare cell types, can then be used for downstream experiments. Advances in instrumentation, expansion of the number of parameters measured, and standardization of assays have increased the power, resolution, and impact of flow cytometry.

These assays of immune cell suspensions have been complemented by histological assays in situ,

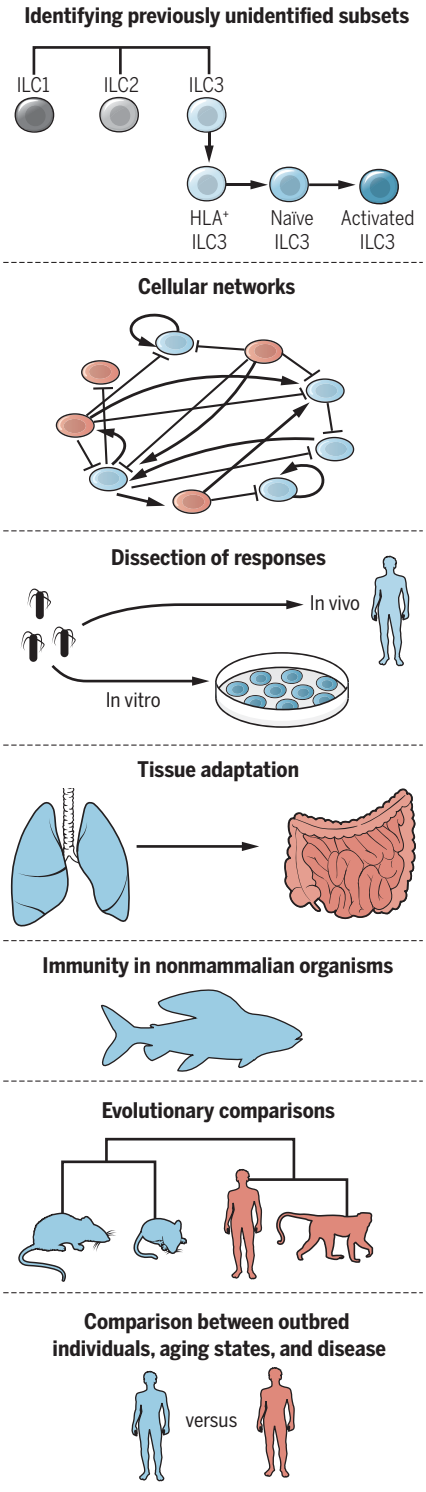

Fig. 1. Single-cell genomics in immunology. The immune system is well suited to studies at single-cell resolution. Immune responses involve a wide variety of cell types that can be further subdivided into more fine-grained subtypes and distinct cell states that can be followed throughout the duration of a response. Furthermore, single-cell analyses provide insights into intercellular networks and allow us to compare immune responses across individuals and species. HLA, human leukocyte antigen.

GRAPHIC: ADAPTED BY K. SUTLIFF/*SCIENCE*

for both RNA and proteins, including in situ hybridization and single-molecule RNA–fluorescence in situ hybridization [reviewed in (*3*)] for RNA and immunohistochemistry for proteins. Microscopy methods provide high-definition spatial representation of cell types, cell boundaries, neighbors or interacting cells, niches, and tissue contexts and have been used to characterize immune cells [reviewed in (*4*)]. More recently, comprehensive profiling of selected bulk populations of large numbers of cells, including of entire transcriptomes and proteomes, helped discover additional markers (*5*).

Although these approaches provided invaluable insights, they suffer from complementary limitations. Single-cell approaches, such as flow cytometry and fluorescence-activated cell sorting, or immunofluorescence and in situ hybridization, have been limited to probing a few selected RNAs or proteins, hindering our ability to study comprehensive profiles and to uncover previously unrecognized factors due to a bias toward precharacterized genes. Conversely, genomic analyses have relied on either profiling heterogeneous mixtures, whose ensemble average obscures the diversity of cells in the sample, or first sorting subpopulations and then profiling them. The latter sorting strategy is limited to known subpopulations and sorting panels and can be difficult to implement for small samples while still masking variation within the subpopulation.

Recent advances in single-cell genomics and spatial profiling methods

Over the past few years, the revolution in single-cell genomics has enabled an unbiased genome-wide quantification of molecules in thousands of individual cells, as well as multiplex spatial analysis of proteins and RNA in situ.

Among the single-cell profiling approaches, the most mature and widely disseminated method is single-cell RNA sequencing (scRNA-seq), which aims to measure the expression levels of genes in cells in a comprehensive way. scRNA-seq can be both sensitive and accurate (*6*, *7*), despite the minute amounts of starting material in an individual cell. scRNA-seq has grown over time in both accuracy and scale, from a handful of individual cells to hundreds of thousands in a single experiment [reviewed in (*8*)]. Despite the relative sparsity of the measured profile from any individual cell, many single-cell profiles can be analyzed with statistical methods that can quantify even subtle changes in gene expression between individual cells (cell-to-cell variation) and also any dynamic expression changes during a response. This has been particularly important in the analysis of profiles of the genomes [reviewed in (*9*)] and epigenomes as either single or simultaneous "multi-omics" modalities [reviewed in (*10*)] of single cells.

Though extremely useful, the measurement of RNA expression does not provide information about protein abundance or posttranslational modifications; in a key parallel advance, mass cytometry by time of flight (CyTOF) (*11*) allows the detection of dozens of proteins at once in each of millions of individual cells, by mass spectrometric detection of metal-labeled antibodies. These can also be applied by staining in situ, for spatial measurements [reviewed in (*12*)]. Protein and RNA measurements can also be combined, either through index sorting followed by scRNA-seq (*13*) or through mass cytometry [e.g., PLAYR (*14*)] or by first staining with DNA-barcoded antibodies and then performing scRNA-seq, to both detect the DNA barcode and profile the transcriptome [CITE-seq (*15*)].

We expect spatially resolved methods [reviewed in (*3*)] to considerably enhance our understanding of the composition and function of the immune system in future studies, but we do not discuss them further here. Instead, we focus on the insights provided by scRNA-seq–based approaches.

The scale, specific characteristics, and throughput of single-cell genomics and spatial techniques are rapidly leading to the accumulation of complex data, which presents analytical challenges and exciting opportunities. These are active and exciting areas of research, reviewed elsewhere (*16*, *17*).

Insights from single-cell analysis of the immune system

Single–cell genomics studies can inform on many key aspects of the immune system (Fig. 1). As the cost of performing single-cell experiments has

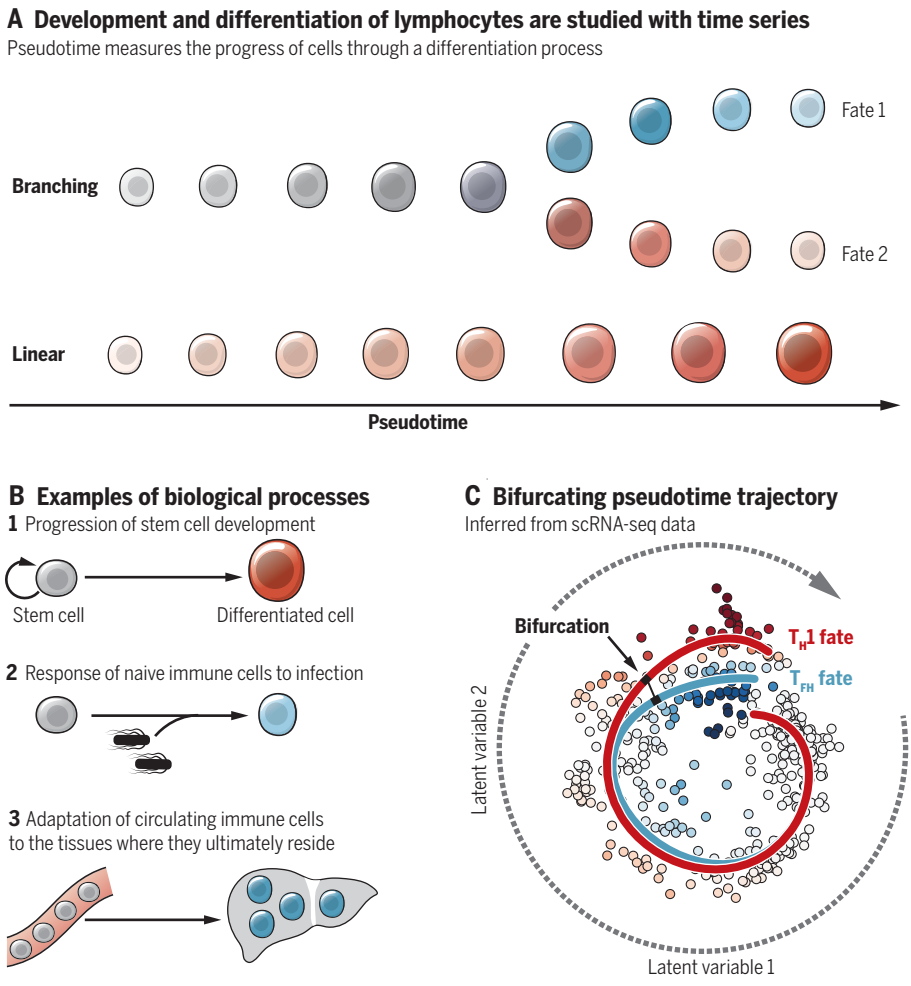

Fig. 2. Inferring cellular trajectories from single-cell data. (**A**) During a differentiation process, individual cells can be aligned along "pseudotime," which represents their progression within the differentiation pathway. The processes described in this way can be linear or can involve branches to multiple eventual fates. (**B**) Examples of biological processes analyzed in terms of cellular trajectories include the progression of stem cells to terminally differentiated fates, the response of naïve immune cells to infection, and the adaptation of circulating immune cells to the tissues where they ultimately reside. (**C**) A bifurcating pseudotime trajectory inferred from scRNA-seq data generated from a mouse malaria infection model [adapted from (*53*), reprinted with permission from AAAS, and modified with permission from the authors]. Each point represents an individual cell following dimensionality reduction using a Bayesian Gaussian process latent variable model. These cells are then ordered in pseudotime, and two simultaneous and bifurcating developmental trajectories (red and blue lines) are inferred using overlapping mixtures of Gaussian processes. The color of each point indicates the probability that a cell belongs to either the red or the blue trend.

decreased, it has allowed profiling of a larger number of immune cells across different conditions. Studies have explored both the innate and adaptive branches of the vertebrate immune system with single-cell genomics techniques, providing important insights. These will guide biomedical research and inspire the next generation of therapies directed toward inflammation, autoimmunity, and cancer.

Innate immunity

Innate immune responses represent the first barrier of fast-acting defense against pathogens. These intrinsic, cell-autonomous responses are mounted upon stimulation with various pathogen-associated molecules [e.g., viral elements such as double-stranded RNA (dsRNA) or bacterial components such as lipopolysaccharide (LPS)]. Single-cell approaches have been used to study these responses' stochasticity—for example, in dendritic cells (DCs) (*18*, *19*)—and to determine their evolutionary architecture in fibroblasts (*20*). Further studies of innate immune cells that mediate a rapid but nonadaptive immune response have illustrated how single-cell genomics can identify previously unidentified subpopulations and increase our insight into immune mechanisms.

Dendritic cells, first responders in the immune system

DCs are sentinels of the immune system and are present in the blood, lymphoid organs, and across many tissues. These cells engulf antigens, then migrate to lymph nodes to present the antigens to T cells. Early scRNA-seq experiments in immune cells were performed on just 18 murine bone marrow–derived dendritic cells (BMDCs) after 4 hours of stimulation with LPS, a component of Gram-negative bacteria (*18*). This revealed cellular heterogeneity in cell subsets, coexpressed gene programs, and splicing variants. In particular, hundreds of key immune genes were bimodally expressed across cells 4 hours after stimulation: Although some cells expressed them at very high levels, there was little, if any, expression in other cells. A following study profiled 1700 mouse BMDCs by scRNA-seq along a time course of stimulation with each of three pathogen-associated molecular patterns (*19*). This charted how heterogeneity across cells represents temporal changes (asynchronicity) in the response and that cells are more or less synchronous in their activation and shut-off of different gene modules on different time scales. In particular, a "core" module of antiviral genes is expressed very early by a few "precocious" cells in response to uniform LPS stimulation but is later activated in all cells. By stimulating and then profiling individual cells in complete isolation, this response was shown to be coordinated by interferon-mediated paracrine signaling from the precocious cells to other immune cells via cell-cell interactions.

Perturb-seq experiments combine large-scale CRISPR-based perturbations in cells with single-cell RNA-seq and have been used to study this response in the context of dozens of genetic perturbations, especially in transcription factors (TFs) (*21*, *22*). This clarified the genetic programs controlling two different subsets of cells (*21*), controlled by different TFs. These experiments also identified new functions for regulators of differentiation, the antiviral response, and mitochondrial function during immune activation (*21*) and detected gene interactions between TFs. Another study (*22*) further identified interactions and redundancies between developmental and signaling-dependent factors in other mononuclear phagocyte populations.

Single-cell genomics can also help to dissect the subsets of human innate immune cells. scRNA-seq studies using human DCs have revealed previously unreported DC subsets and shed light on the complexity of the lineage of these cells (*23*–*25*). For example, analysis of human tonsils and synovial fluid from rheumatoid arthritis patients showed that cells from one previously unrecognized DC subset are in direct proximity to T cells, supporting their role in mounting an immune response and in autoimmunity (*24*). Such initial profiling studies should lead to immunophenotyping of newly discovered markers and subsequent functional assays.

> ***"The high-resolution mapping of cell states by scRNA-seq can infer developmental processes in detail."***

Innate lymphoid cells and their development

Innate lymphoid cells (ILCs) reside at mucosal barriers and are important for host defense and tissue homeostasis. ILCs are the innate immune system's functional equivalent of T cell subsets of the adaptive immune system, composed of cytotoxic and noncytotoxic cells and resembling T lymphocytes in cytotoxicity and cytokine production but lacking antigen-specific receptors (*26*). The complete spectrum of immune tasks performed by ILCs is still unknown. Although there is evidence that ILCs are important in the processes of inflammation and tissue homeostasis [reviewed in (*27*)], they may be at least partly dispensable in the context of functional T and B cells (*28*).

The initial discovery of noncytotoxic ILCs coincided with the era of single-cell RNA-seq, and these methods have helped reveal the heterogeneity, developmental pathways, regulatory circuits, and responses to microbiota of ILCs. An analysis of 648 ILCs from human tonsils used gene expression profiles to separate the cells into the expected four distinct populations [ILC1, ILC2, ILC3, and natural killer (NK) cells] with unbiased clustering, without relying on any previous knowledge (*29*). The transcriptomes of the cells in the four known types helped to identify previously unknown marker genes. Additionally, on the basis of their cytokine responses and cytokine secretion profiles, they revealed three previously unreported subsets of ILC3 cells that are predicted to have distinct functions. Another study analyzed scRNA-seq profiles of ILC subsets in the colons of mice under different microbiotic conditions (specific pathogen–free, antibiotic-treated, and germ-free facilities) (*30*). This revealed subsets within the ILC1, -2, and -3 populations, as well as a hybrid set of cell states that may represent plasticity between the ILC1 and ILC3 states.

Single-cell RNA-seq has also been instrumental in resolving details of the trajectory and regulation of the developmental pathway of ILCs (*31*) (Fig. 2, A and B). Analysis of mouse bone marrow progenitors revealed precursor subsets, and delineated distinct ILC development stages and pathways. This study highlighted markers—such as the PD1 receptor on both progenitor cells and cells stimulated after influenza infection and lung inflammation—and regulation by the TF Bcl11b. As the PD1 receptor's ligand is targeted in tumor immunotherapy, these findings may have implications for disease and therapy.

The monocyte-macrophage lineage and its dynamics in health and disease

Macrophages, a broad family of phagocytic cells, not only play a role in encounters with pathogens but also have diverse roles in tissue homeostasis (e.g., microglia in the brain, osteoclasts in the bone) (*32*–*34*). Single-cell profiling has shed light on both types of roles, as well as on tissue adaptation.

Profiling mouse macrophages infected with fluorescently labeled *Salmonella enterica* serovar Typhimurium revealed diverse macrophage states correlating with bacterial survival and proliferation and linking pathogen state to host signaling pathways (*35*). A similar approach, focused on pathogen proliferation and macrophage response, identified an M2-like state in macrophages infected with actively proliferating bacteria (*36*). Developing scRNA-seq methods that measure the transcriptional changes in both the immune cell and the pathogen will help inform on the function and mechanisms employed by these cells in defense.

Microglia are a particular lineage of cells within the macrophage family—the tissue-resident macrophages in the brain. Surveys with scRNA-seq of the developmental progression of microglia identified their activation in a mouse model of Alzheimer's disease (*22*, *37*). A previously unrecognized microglia subset was associated with neurodegenerative diseases and linked with relevant markers, spatial localization, and the activity of specific cellular pathways. From the specific sequence of disease progression in animal models, this distinct microglia type showed the potential to restrict neurodegeneration.

Comparative studies of innate immune cells and cell-intrinsic responses

One of the challenges in immunology has been to understand how the immune system has evolved

across different animal models relative to humans. This was exacerbated by the facts that both compositional and cell-intrinsic changes affect such comparisons based on bulk profiles, that subsets defined by cell surface markers may be challenging to match, and that antibodies and markers may be lacking in nonmodel organisms.

scRNA-seq helps to circumvent these challenges. For example, T cells, NK cells, and myeloid-like cell transcriptomes were characterized in zebrafish, a vertebrate model organism (*38*). In another example, scRNA-seq profiled the cell-intrinsic innate immune response of fibroblasts across four mammalian species stimulated with dsRNA (*20*). This highlights the relationships among divergent gene expression across species, individual cells, and protein sequence divergence.

Adaptive immunity

Previously unrecognized lymphocyte subsets

Immunologists have gained an understanding of the diversity of lymphocyte subpopulations and their roles in immunity, characterizing populations of lymphocytes such as B cells, $CD4^+$ and $CD8^+$ T cells, NK cells, and iNKT cells. Leveraging these advances, scRNA-seq is aiding in the discovery of previously unreported lymphocyte subpopulations and states, their molecular underpinnings, and their relationship to physiology and disease.

For example, the plasticity and complexity of the $CD4^+$ T cell compartment makes it difficult to understand with bulk techniques but particularly amenable to investigation by scRNA-seq methods. scRNA-seq analysis of a mouse $CD4^+$ T cell population enabled correlation of a steroidogenic enzyme with a cell surface marker that was used to enrich this subpopulation for further functional characterization (*39*).

Other studies have focused on the characterization of variation within an immune subset, which can be continuous rather than involving further discrete subsets. For example, scRNA-seq helped to characterize the variation within T helper 17 (T_H17) populations, which can span a continuous spectrum correlated with distinct levels of "pathogenicity," or the ability to induce an autoimmune disease in an animal model upon adoptive transfer (*40*, *41*). scRNA-seq of T_H17 cells from in vivo autoimmune models and in vitro polarization experiments helped to characterize the genetic programs that underlie this diversity and to shed light on the mechanisms that control them.

These strategies are proving beneficial in the context of immunooncology, in both mouse models and human tumors. Studies of human melanoma (*42*) and hepatocellular carcinoma (*43*), and a B16 melanoma mouse model (*44*), characterized diverse T cell states in the tumor ecosystem (see also below). These studies explored the relationship between cytotoxicity and cell exhaustion, as well as the roles of particular proteins in suppressive T cell states.

Lymphocyte differentiation

The high-resolution mapping of cell states by scRNA-seq can infer developmental processes in detail. In bulk transcriptomics methods, development and differentiation are studied with time series. In contrast, the single-cell resolution of scRNA-seq allows the inference of time series from static snapshots and reveals the continuum of cell states across time points (Fig. 2, A and B). Known as "pseudotime," this unobserved dimension measures the progress of cells through a transition. For example, such an approach was used to resolve developmental progressions from single-cell quantitative polymerase chain reaction data (*45*, *46*) and was later applied to scRNA-seq data in zebrafish hematopoiesis, revealing a continuous spectrum of differentiation in hematopoietic cells (*47*). An early example of pseudotime inference

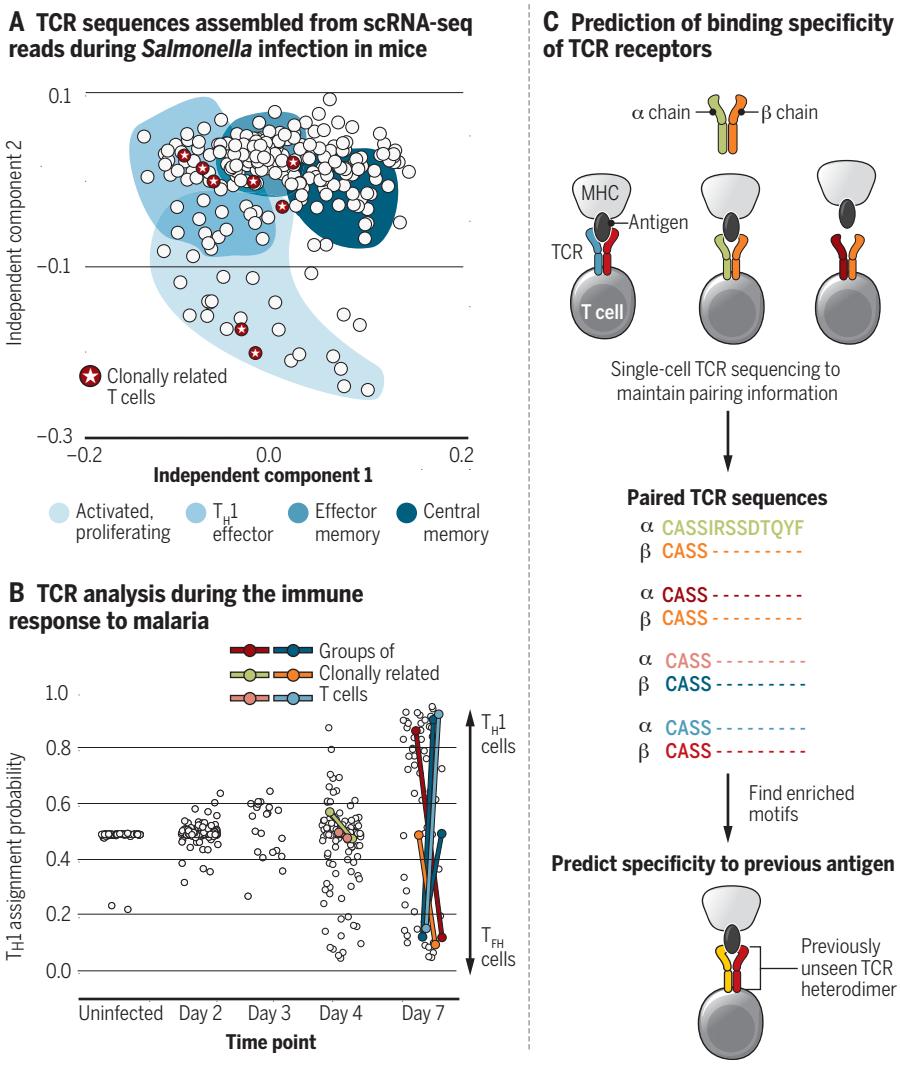

Fig. 3. Single-cell analysis of antigen receptor sequences reveals clone distributions between transcriptional states. (**A**) Independent component analysis of scRNA-seq data from mouse splenic $CD4^+$ T cells during *Salmonella* infection. Each point represents an individual cell. Shaded areas indicate likely functional identities associated with each region of reduced dimensionality space. Starred points indicate cells that are clonally related and share one particular set of TCR sequences. Clonally related cells are distributed throughout the gene expression space [adapted from (*56*)]. (**B**) Cell type assignment of cells with scRNA-seq data from mouse splenic $CD4^+$ T cells at varying time points during malaria infection. Each point represents an individual cell, with *y*-axis position indicating the likelihood that it is a T_H1 cell rather than a T_{FH} cell (high values imply a T_H1 identity, low values a T_{FH} identity). Colored points indicate pairs of cells inferred to be clonally related due to shared TCR sequences. Sibling cells can be found such that one is a T_H1 cell, whereas the other is a T_{FH} cell [adapted from (*53*), reprinted with permission from AAAS, and modified with permission from the authors]. (**C**) Data sets (*59*, *60*) of linked α and β TCR chains provide enough power to allow machine learning inference of common sequence motifs from diverse T cells recognizing the same antigen. This opens the door for the future possibility to more systematically associate TCR sequences with their cognate peptide–MHC sequences. [Figure adapted by the authors from (*66*)]

GRAPHIC: ADAPTED BY K. SUTLIFF/*SCIENCE*

in the immune system focused on B cells undergoing V(D)J recombination. In this work, the Wanderlust method inferred trajectories on the basis of similarity of cells according to CyTOF profiles (*48*).

Notably, when dynamic processes occur concurrently with cell activation or differentiation—for example, the induction of cell proliferation in many immune processes—this specific contribution can be either identified and studied (*49*, *50*) or isolated to highlight other phenomena (*45*, *51*). This latter approach may be challenging when cell proliferation is coupled to other immune responses (*17*).

Layered on top of the progression of differentiation are cell fate decisions. These can be single bifurcations, such as $CD4^+$ T cell fate choices, or complex hierarchies and other relations, as in the whole of hematopoiesis. With appropriate computational analyses, these cell fate decisions can be explored from both snapshot (when differentiation is asynchronous and ongoing) and time-course data. Such approaches were applied to study both T cell development (*52*) and T helper cell differentiation into T_H1 and T follicular helper (T_{FH}) cell fates (*53*) (Fig. 2C).

Immune repertoire analyses

The adaptive immune response requires interactions between nonself antigens and antigen-specific receptor molecules expressed by T and B lymphocytes. T cell receptors (TCRs) detect antigenic peptides presented as complexes with major histocompatibility complex (MHC) proteins. B cell receptors bind to proteins or peptides without presentation and are also secreted from B cells as antibodies. The immune system can specifically recognize the necessarily large range of antigenic molecules, thanks to the enormous potential diversity of antigen receptor (AgR) amino acid sequences. This diversity is generated by DNA rearrangement in each developing lymphocyte that combines randomly chosen gene segments and introduces additional variability at the junctions between segments. The vast diversity of AgR molecules allows us to assume that cells with identical receptor DNA sequences arose from the same original developing lymphocyte.

The distinctiveness and diversity of AgRs mean that single-cell techniques are ideal to study antigen receptor repertoires, as well as their relation to the cells' subtype and state. Furthermore, AgRs are heterodimeric proteins that comprise two independently encoded protein chains that both typically determine each receptor's antigen specificity. Single-cell measurements identify the paired receptor chains in each cell.

One line of work focuses on the repertoire of TCR sequences at the single-cell level and provides insights into clonal fates and dynamics. Targeted sequencing of TCR sequences in individual cells, combined with immunophenotyping of selected marker genes, related fate and clonality in tumor-infiltrating lymphocytes in colorectal cancer (*54*). Cells from two populations of interleukin 17–expressing cells ($FOXP3^-RORC^+$ and $FOXP3^+RORC^+$), often shared TCR sequences, implying a common ancestry. Furthermore, the use of droplet-based encapsulation and targeted AgR amplification detected paired sequences from hundreds of thousands of lymphocytes (*55*), along with two phenotyping markers.

More recently, AgR sequences have been studied in the context of scRNA-seq, which can simultaneously report on the receptor of a cell and its expression profile (*42*, *53*, *56*–*58*). TCR sequences assembled from scRNA-seq reads (*56*) (Fig. 3A) were used to demonstrate the breadth of transcriptional states within a single expanded T cell clonotype during *Salmonella* infection in mice. In another example, TCR analysis was combined with pseudotime and branching inference during the immune response to malaria to show that clonally related sibling cells can be found in both the T_H1 and T_{FH} fates (*53*) (Fig. 3B). In human melanoma samples (*42*), similar analysis allowed the relation of the exhaustion state of the cell to the extent of a clone's expansion.

The ideal next step would be to incorporate AgR detection into massively parallel droplet-based whole-transcriptome scRNA-seq. This would allow clonality to be linked to phenotypic state in thousands of cells, at least at a coarse-grained level, and would hence illuminate the development of lineages of lymphocyte cell states at unprecedented resolution.

Finally, the systematic and large-scale single-cell analysis of AgRs has recently provided evidence that it may be possible to predict a receptor's cognate antigen using only the receptor sequence, potentially leading to the development of a clinical diagnostic tool. Paired single-cell TCR sequences in a variety of antigen-specific T cell populations from multiple donors were determined, along with enriched sequence motifs associated with binding to each particular antigen (*59*, *60*). The presence of these enriched motifs in previously unseen TCRs enabled prediction of binding specificity of these receptors (Fig. 3C). Although these studies still require large training sets of TCR sequences with preexisting knowledge of specificity, they are likely to be an early step on the road to generalizable prediction of TCR specificity from sequence.

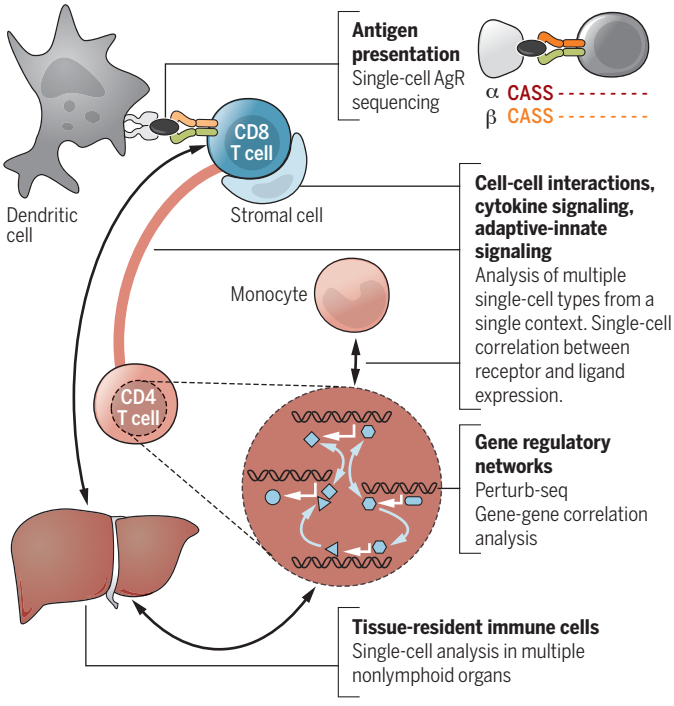

Fig. 4. Network interactions within the immune system. Immune responses involve networks at multiple scales, ranging from intracellular gene regulatory networks to long-distance intercellular communication mediated by cytokines or chemokines. A systems approach to understanding these networks will be crucial if we are to fully understand immune biology and will be accelerated by the application of multiple, different single-cell analysis methods.

Cellular ecosystems: An integrative view of cell-cell interactions in immunity

The immune system requires a complex network of interactions to perform its functions (Fig. 4). An integrated view of individual cells, cell types, and their spatial organization and functional interaction is needed to achieve a systems-level view of immune function. Single-cell approaches are already accelerating our integrated understanding of immunity, and we anticipate much future work in this area.

Existing single-cell studies have begun to demonstrate the power of addressing multiple parts of the immune response at the same time. An analysis of 4645 individual cells from melanoma patient samples revealed the complexity of the cellular ecosystem present in the tumors (*42*) and the extent to which it can be compared across individual patients. Whereas malignant cell types exhibit substantial diversity between patients, nonmalignant cells [including T cells, B cells, macrophages, endothelial cells, cancer-associated fibroblasts (CAFs), and NK cells] were each grouped by type rather than tumor origin. Similar phenomena were observed in other tumor types, such as glioma (*61*). Moreover, when the single-cell profiles are combined with bulk profiles from hundreds of patients, collected by The

Cancer Genome Atlas (*62*) (http://cancergenome.nih.gov/), they reveal dependencies between different cell types and the molecules that may mediate them: for example, the correlation between high levels of $CD8^+$ T cells in a tumor and the expression of complement proteins by its CAFs (*42*).

Another study analyzed splenic $CD4^+$ T cells, monocytes, and DCs in the mouse response to malaria (*53*), comparing uninfected mice to mice at day 3 postinfection. Integrated analyses of the changes in expression of chemokine ligands and receptors across these populations predicted that monocytes would support the differentiation of activated $CD4^+$ T cells toward a T_H1 fate rather than the alternative T_{FH} fate. This was confirmed in an experiment in which monocytes were depleted during T cell activation and before fate choice.

Both studies inferred processes of intercellular communication from the expression of cognate receptors, co-receptors, and/or ligands in scRNA-seq data. Systematic and generalizable approaches for such connections, and relating them to covariation in cell proportions and states, will be of great value. Furthermore, it is clear that methods that analyze single-cell gene expression in a spatially resolved context will be very important to understanding the interactions between cells of the immune system. Indeed, imaging mass cytometry has identified the presence of immune cells within the spatial context of breast cancer tissues (*63*).

Future applications in immunity

The immune system is composed of numerous cell types that work in concert to sense and appropriately respond to foreign challenges and physiological changes in order to monitor and maintain health. If the carefully orchestrated functioning of the immune system is perturbed, diseases such as infectious disease, autoimmune disease, and cancer can arise.

The rich taxonomies and cell fate maps generated in immunology over the past several decades relate cells by cellular function, differentiation potential, and expression of marker proteins. However, to date there is still no complete reference map of immune cells. The additional comprehensive profiles produced by single-cell genomic approaches provide an important new tool in this endeavor by helping to address some of the fundamental questions in immunology—from the taxonomy of cells, histological structure in tissues, recruitment to tissues, developmental biology, and cell fate and lineage to physiology and homeostasis and their underlying molecular mechanisms.

Moreover, to deeply understand the full scope and function of immune cells, it is most informative to study them in a challenged state—that is, as manifested during disease, infection, development, aging, and environmental changes. Studying immune cells in humans often requires handling minute samples, and single-cell genomic approaches are highly compatible with such limitations in input material.

Finally, genetic perturbations (natural or engineered) can also elicit changes in gene expression that may differ in response to environmental cues or changes such as aging (*18*, *19*, *64*). As the cost of single-cell experiments goes down, profiling more immune cells across conditions will be possible via pooled genetic screens and by economically multiplexing a large number of individuals, using their sequence variations as a natural genetic barcode (*65*).

Combining single-cell genomics, emerging spatial approaches, immune repertoire analysis, and multiplex immunophenotyping, together with established approaches for functional analysis, could also affect the next generation of diagnostics and therapies. For diagnostics, the white blood cell count might be reimagined from a tally of major cell populations to an assay that identifies cell signatures (defined by single-cell genomics) of cell types and states and their proportions. For therapy, comparing the role and mechanisms of immune cells in cellular ecosystems between healthy and diseased tissues can help to identify new therapeutic targets, as well as to better assess the effect of current therapies in the context of clinical trials.

To help usher in this future, an effort to generate an Immune Cell Atlas (ICA) is emerging as part of the international Human Cell Atlas initiative (www.humancellatlas.org). The ICA will assess the immune system at different stages of differentiation, across different tissues, and in the context of a wide range of diseases. To properly survey the spectrum of immune cells, even the initial pilot effort will include samples from small numbers of patients with a diversity of diseases. Such a systematic characterization of immunity requires an international collaboration among clinicians, immunologists, genomics experts, and computational biologists.

Overall, these types of approaches and projects stand to radically transform our knowledge of immune function and dysfunction in infection, autoimmunity, allergy, inflammatory disorders, and cancer, as well as to affect therapeutic developments.

REFERENCES AND NOTES

1. *Nat. Immunol.* **14**, 977 (2013).
2. S. P. Perfetto, P. K. Chattopadhyay, M. Roederer, *Nat. Rev. Immunol.* **4**, 648–655 (2004).
3. E. Lein, L. E. Borm, S. Linnarsson, *Science* **358**, 64–69 (2017).
4. M. Bajénoff, R. N. Germain, *Eur. J. Immunol.* **37** (suppl. 1), S18–S33 (2007).
5. T. Shay, J. Kang, *Trends Immunol.* **34**, 602–609 (2013).
6. V. Svensson *et al.*, *Nat. Methods* **14**, 381–387 (2017).
7. A. A. Kolodziejczyk, J. K. Kim, V. Svensson, J. C. Marioni, S. A. Teichmann, *Mol. Cell* **58**, 610–620 (2015).
8. V. Svensson, R. Vento-Tormo, S. A. Teichmann, Exponential scaling of single-cell RNA-seq in the last decade. arXiv:1704.01379 [q-bio.GN] (5 April 2017).
9. C. Gawad, W. Koh, S. R. Quake, *Nat. Rev. Genet.* **17**, 175–188 (2016).
10. G. Kelsey, O. Stegle, W. Reik, *Science* **358**, 69–75 (2017).
11. D. R. Bandura *et al.*, *Anal. Chem.* **81**, 6813–6822 (2009).
12. B. Bodenmiller, *Cell Syst.* **2**, 225–238 (2016).
13. N. K. Wilson *et al.*, *Cell Stem Cell* **16**, 712–724 (2015).
14. A. P. Frei *et al.*, *Nat. Methods* **13**, 269–275 (2016).
15. M. Stoeckius *et al.*, *Nat. Methods* **14**, 865–868 (2017).
16. R. Rostom, V. Svensson, S. A. Teichmann, G. Kar, *FEBS Lett.* **591**, 2213–2225 (2017).
17. A. Wagner, A. Regev, N. Yosef, *Nat. Biotechnol.* **34**, 1145–1160 (2016).
18. A. K. Shalek *et al.*, *Nature* **498**, 236–240 (2013).
19. A. K. Shalek *et al.*, *Nature* **510**, 363–369 (2014).
20. T. Hagai *et al.*, bioRxiv 137992 [Preprint]. 15 May 2017.
21. A. Dixit *et al.*, *Cell* **167**, 1853–1866.e17 (2016).
22. D. A. Jaitin *et al.*, *Cell* **167**, 1883–1896.e15 (2016).
23. G. Breton *et al.*, *J. Exp. Med.* **212**, 401–413 (2015).
24. A.-C. Villani *et al.*, *Science* **356**, eaah4573 (2017).
25. P. See *et al.*, *Science* **356**, eaag3009 (2017).
26. G. Eberl, M. Colonna, J. P. Di Santo, A. N. J. McKenzie, *Science* **348**, aaa6566 (2015).
27. D. Artis, H. Spits, *Nature* **517**, 293–301 (2015).
28. F. Vély *et al.*, *Nat. Immunol.* **17**, 1291–1299 (2016).
29. Å. K. Björklund *et al.*, *Nat. Immunol.* **17**, 451–460 (2016).
30. M. Gury-BenAri *et al.*, *Cell* **166**, 1231–1246.e13 (2016).
31. Y. Yu *et al.*, *Nature* **539**, 102–106 (2016).
32. Y. Okabe, R. Medzhitov, *Nat. Immunol.* **17**, 9–17 (2016).
33. Y. Lavin, A. Mortha, A. Rahman, M. Merad, *Nat. Rev. Immunol.* **15**, 731–744 (2015).
34. I. Amit, D. R. Winter, S. Jung, *Nat. Immunol.* **17**, 18–25 (2016).
35. R. Avraham *et al.*, *Cell* **162**, 1309–1321 (2015).
36. A.-E. Saliba *et al.*, *Nat. Microbiol.* **2**, 16206 (2016).
37. H. Keren-Shaul *et al.*, *Cell* **169**, 1276–1290.e17 (2017).
38. S. J. Carmona *et al.*, *Genome Res.* **27**, 451–461 (2017).
39. B. Mahata *et al.*, *Cell Reports* **7**, 1130–1142 (2014).
40. J. T. Gaublomme *et al.*, *Cell* **163**, 1400–1412 (2015).
41. C. Wang *et al.*, *Cell* **163**, 1413–1427 (2015).
42. I. Tirosh *et al.*, *Science* **352**, 189–196 (2016).
43. C. Zheng *et al.*, *Cell* **169**, 1342–1356.e16 (2017).
44. M. Singer *et al.*, *Cell* **166**, 1500–1511.e9 (2016).
45. J. C. H. Tsang *et al.*, *Genome Biol.* **16**, 178 (2015).
46. F. Buettner, F. J. Theis, *Bioinformatics* **28**, i626–i632 (2012).
47. I. C. Macaulay *et al.*, *Cell Reports* **14**, 966–977 (2016).
48. S. C. Bendall *et al.*, *Cell* **157**, 714–725 (2014).
49. A. Scialdone *et al.*, *Methods* **85**, 54–61 (2015).
50. N. Leng *et al.*, *Nat. Methods* **12**, 947–950 (2015).
51. F. Buettner *et al.*, *Nat. Biotechnol.* **33**, 155–160 (2015).
52. M. Setty *et al.*, *Nat. Biotechnol.* **34**, 637–645 (2016).
53. T. Lönnberg *et al.*, *Sci. Immunol.* **2**, eaal2192 (2017).
54. A. Han, J. Glanville, L. Hansmann, M. M. Davis, *Nat. Biotechnol.* **32**, 684–692 (2014).
55. A. W. Briggs *et al.*, bioRxiv 134841 [Preprint]. 5 May 2017.
56. M. J. T. Stubbington *et al.*, *Nat. Methods* **13**, 329–332 (2016).
57. S. Canzar, K. E. Neu, Q. Tang, P. C. Wilson, A. A. Khan, *Bioinformatics* **33**, 425–427 (2017).
58. A. A. Eltahla *et al.*, *Immunol. Cell Biol.* **94**, 604–611 (2016).
59. P. Dash *et al.*, *Nature* **547**, 89–93 (2017).
60. J. Glanville *et al.*, *Nature* **547**, 94–98 (2017).
61. A. S. Venteicher *et al.*, *Science* **355**, eaai8478 (2017).
62. K. Tomczak, P. Czerwińska, M. Wiznerowicz, *Contemp. Oncol.* **19**, A68–A77 (2015).
63. C. Giesen *et al.*, *Nat. Methods* **11**, 417–422 (2014).
64. C. P. Martinez-Jimenez *et al.*, *Science* **355**, 1433–1436 (2017).
65. H. M. Kang *et al.*, bioRxiv 118778 [Preprint]. 15 May 2017.
66. S. T. Reddy, *Nature* **547**, 36–38 (2017).

ACKNOWLEDGMENTS

We thank J. E. Rood, R. Majovski, V. Svensson, L. Gaffney, and K. Meyer for help with preparation of this manuscript. A.R. was supported by funds from the Howard Hughes Medicine Institute, National Institute of Allergy and Infectious Diseases grants U24AI118672 and R24AI072073, the Manton Foundation, and the Klarman Cell Observatory. M.J.T.S. and S.A.T. were supported by the Wellcome Trust grant 206194. A.R. is a scientific advisory board member for ThermoFisher Scientific and Syros Pharmaceuticals and a consultant for Driver Group; the other authors have no conflicts of interest.

10.1126/science.aan6828

Single-cell transcriptomics to explore the immune system in health and disease

Michael J. T. Stubbington, Orit Rozenblatt-Rosen, Aviv Regev and Sarah A. Teichmann

Science **358** (6359), 58-63.
DOI: 10.1126/science.aan6828