Review Article

The Application of Single-Cell RNA Sequencing in Vaccinology

Andrés Noé [1], **Tamsin N. Cargill,**[2,3] **Carolyn M. Nielsen,**[1] **Andrew J. C. Russell,**[4] **and Eleanor Barnes**[2,3]

[1]*The Jenner Institute, University of Oxford, Old Road Campus Research Building, Oxford OX3 7DQ, UK*
[2]*Peter Medawar Building for Pathogen Research and Oxford NIHR Biomedical Research Centre, Nuffield Department of Medicine, University of Oxford, South Parks Road, Oxford OX1 3SY, UK*
[3]*Translational Gastroenterology Unit, John Radcliffe Hospital, Oxford OX3 9DU, UK*
[4]*Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK*

Correspondence should be addressed to Andrés Noé; andres.noe@gtc.ox.ac.uk

Single-cell RNA sequencing allows highly detailed profiling of cellular immune responses from limited-volume samples, advancing prospects of a new era of systems immunology. The power of single-cell RNA sequencing offers various opportunities to decipher the immune response to infectious diseases and vaccines. Here, we describe the potential uses of single-cell RNA sequencing methods in prophylactic vaccine development, concentrating on infectious diseases including COVID-19. Using examples from several diseases, we review how single-cell RNA sequencing has been used to evaluate the immunological response to different vaccine platforms and regimens. By highlighting published and unpublished single-cell RNA sequencing studies relevant to vaccinology, we discuss some general considerations how the field could be enriched with the widespread adoption of this technology.

1. Introduction

Vaccines are one of the most effective public health interventions in history and have been extremely successful in preventing illness and death from many infections. Much of this success can be attributed to the discovery of disease-causing agents and/or by the discovery of how to cultivate these pathogens to allow large-scale production of attenuated vaccines. While it is clear that effective vaccines induce protective immunological memory, the precise mechanisms by which this manifests are often poorly understood. Moreover, there are many diseases against which we have not developed successful vaccines, often a result of not fully understanding the "ideal" immune response and/or how to induce this with vaccination. Currently used techniques, such as ELISAs, ELISpots, flow cytometry, and growth inhibition assays, broadly measure responses in the T cell or humoral compartments after vaccination, but cannot agnostically measure differences in response between single immune cells [1–3]. Single-cell RNA sequencing (scRNA-seq) is a relatively novel tool which provides the advantage of understanding

responses to vaccination at the level of the individual cell in an unbiased manner.

RNA sequencing quantitatively profiles the cellular transcriptome. Polyadenylated messenger RNA (mRNA) molecules are often the target as the polyA tail is a convenient handle to selectively target the protein-coding mRNA (as opposed to other RNA types). In bulk RNA-seq studies, many thousand cells may be pooled together, obscuring heterogeneity. scRNA-seq (in contrast to bulk) allows the dissection of previously unappreciated levels of heterogeneity. This is a major motivation for embarking in scRNA-seq studies [4, 5]. Over 25 scRNA-seq techniques have been developed in just over a decade, all essentially following five steps: (1) single cell isolation, (2) cell lysis and RNA capture, (3) RNA reverse transcription to cDNA, (4) cDNA amplification, and (5) pooling and sequencing using library preparation, pooling, and next-generation sequencing techniques [5]. Some of the most used scRNA-seq techniques include Smart-seq2 [6], MARS-seq [7], 10x Genomics Chromium [8], inDrop [9], and Seq-Well [10]. The precise differences between these techniques have been discussed extensively by

Kolodziejczyk and colleagues [11], with the major differences relating to the resulting transcript data (including sensitivity, accuracy, and transcript portion profiled), throughput, single-cell isolation method, and sequencing platform.

The relative paucity of published reports of single-cell transcriptomic responses in the context of vaccination suggests that there remains much to be learned from scRNA-seq. As with all new techniques, there are difficulties in establishing robust, scalable, and cost-effective protocols for the generation and analysis of scRNA-seq data [12]. However, these obstacles are countered by the opportunity to elucidate complex networks of cell interactions and immune responses and the potential to identify novel or unanticipated response profiles, which have been beyond the scope of bulk RNA and other sequencing technologies. scRNA-seq can serve as the backbone for several other "omics" technologies, where the transcriptome can be profiled in the same cell as well as surface proteins (CITE-seq and REAP-Seq) [13, 14], chromatin accessibility (ATAC-seq) [15], and genomes (G&T-seq and DR-seq) [16–18]. The combination of these technologies allows new subpopulations to be revealed, which would not otherwise be possible by the use of each alone [19, 20], although in-depth discussion of these technologies is beyond the scope of this review.

This review considers the applications of scRNA-seq in prophylactic vaccine development, with a focus on infectious diseases. We use examples from several diseases to demonstrate the flexibility of the technology. We explore published and unpublished literature to highlight existing applications of this technology and provide recommendations and predictions as to how vaccinology could be enriched with its widespread adoption. To illustrate the adaptability of scRNA-seq, we present the case study of COVID-19 vaccine development and discuss the contribution unbiased transcriptional profiling could make.

2. Profiling Immune Responses to Infections

Our understanding of the mechanisms underlying immune responses in health and disease has important implications for vaccine design. Previously, targeted techniques have allowed us insights into specific parts of the immunological system during development, during infection, and after infection. scRNA-seq allows the immune system to be studied in an unbiased manner. Additionally, studying single cells allows quantitation of the heterogeneity in systems and to resolve time during dynamic processes. Studying the immune response to infection can provide a window to understanding the challenges that must be overcome by vaccination. This is particularly relevant in diseases such as influenza or malaria where natural infection does not engender complete protection [21, 22]. Here, we highlight uses of scRNA-seq for profiling different components of the immune response in the context of natural or artificial infection, as well as concurrent sequencing of pathogen and host, and responses in the specific context of controlled human infection studies.

2.1. Innate and Adaptive Responses to Pathogens. The particular innate cell types and pathways that trigger an effective adaptive immune response have been the focus of recent work by Blecher-Gonen et al. The authors used scRNA-seq to characterise the initial 48 hours of the cellular response to several fluorescently labelled inactivated pathogens [23]. As early as 24 hours after immunisation, relatively rare antigen-carrying cells showed pathogen-restricted programs of transcription. Fluorescent antigen-positive neutrophil and monocyte populations were found almost exclusively in inactivated Mycobacteria-immunised mice, whereas antigen-positive macrophages were mainly found in inactivated Candida-immunised mice. This analysis elucidates initial pathways after inactivated vaccine administration and shows that scRNA-seq can disentangle populations that vary widely in lineage, activation status, and antigen uptake. Further work is needed to identify whether these different innate pathways do in fact correlate with natural or vaccine-induced protection. If they did, regimens could include adjuvants and use specific platforms to trigger the type of innate activation that has been identified to be protective. Indeed, scRNA-seq would be a well-suited tool for this subsequent work.

Long-lived plasma cells are crucial to maintaining high levels of antibodies long after infection and vaccination [24]. Lam et al. prospectively sorted subsets of plasma cells formed in response to natural infections in mice and performed scRNA-seq to define plasma cell transcriptional heterogeneity [25]. In keeping with previous reports [26], approximately a third of plasma cell transcriptomes were made up of kappa light chain constant region transcripts. Despite there being significant metabolic differences in the subsets of plasma cells, metabolic state did not appear to correlate with transcriptional profile. Genes crucial for B cell longevity, including CD28, BLIMP-1, and B-cell maturation antigen, were not differentially expressed across plasma cells, regardless of longevity. There were no stable changes in transcription between long- and short-lived plasma cell subsets. Despite glucose uptake being shown to be important for plasma cell longevity and it being known that transcriptional changes are essential for the establishment of metabolic programs in plasmablasts [27], transcription does not seem to distinguish mature plasma cell subsets further. Indeed, a multiomic tool may assist in resolving these subsets—differences may be accounted for at the protein level by translational control. Analogous studies also need to be performed in humans to determine whether plasma cells generated in response to natural infection and vaccination behave similarly and whether further transcriptional investigation is necessary.

Peripherally circulating CD8+ T cells have been associated with immune control of HIV [28]. Elite controllers of HIV are people who, in the absence of antiretroviral therapy, suppress viral replication (in contrast to chronic progressors). Nguyen et al. compared the characteristics of HIV-tetramer-specific CD8+ T cells in the blood and lymph nodes of elite controllers and chronic progressors [29]. Weakly cytolytic CD8+ T cells preferentially homed to B cell follicles and vigorously suppressed replication of HIV in elite controller lymph nodes. These CD8+ T cells upregulated expression of numerous soluble factors and cytokines and downregulated inhibitory receptors. The authors argued that these features identified a CD8+ protective immune signature in

EC. These results provide guidance for vaccine design, towards the type of CD8+ T cell signature that may be necessary for effective vaccine-induced protection.

Memory CD4+ T cells are required for long-lived immunity and are induced by vaccination strategies, including against malaria and influenza [30, 31]. Additionally, influenza-specific CD4+ T cells correlate with protection against influenza challenge in humans [32]. To this end, Ciucci et al. profiled the transcriptional programs driving CD4+ T cell heterogeneity and memory T cell development [33]. The authors performed iterative scRNA-seq profiling of total T cells, CD4+ T cells, and virus-specific CD4+ T cells at seven days post lymphocytic choriomeningitis virus infection in mice. By comparing the gene expression differences in clusters, Ciucci et al. identified a specific T cell memory precursor signature. The establishment of the memory precursor signature and generation of a long-lived CD4+ response required the transcription factor Thpok. Knowledge of this signature and the mechanistic importance of Thpok in early CD4+ T cell memory could be used to predict longer term CD4+ responses induced by vaccination and, potentially, the generation or lack of vaccine-induced protection.

Rato and colleagues used scRNA-seq to investigate CD4+ T cell heterogeneity prior to HIV infection [34]. To explore the basis of CD4+ T cell permissiveness to HIV infection, they examined for the growth, and infection-permissiveness, of primary cells by a pseudotyped HIV-based vector. The main factor determining transcriptional heterogeneity was the degree of response to TCR stimulation, or cellular activation, which, in turn, manifested in varying degrees of HIV permissiveness. "HIV-permissive cells" identified prior to infection allowed discovery of a gene signature that divided populations into high- and low-permissive subsets. In a similar study, single-cell viral RNA quantitation was performed to demonstrate the correlation between cell gene expression and HIV latency [35]. Analogous immunoprofiling analyses could be performed in the context of vaccinated individuals to determine whether HIV-permissive or HIV-resistant cells are enriched and whether these populations correlate with vaccine-induced protection.

2.2. Dual scRNA-seq of Pathogen and Host.

scRNA-seq can be deployed to simultaneously interrogate both pathogen and host transcriptomes. Transcriptional profiling at high resolution has enabled an in-depth appreciation of the cellular diversity in biological organisms and the number of transcriptional states during infection. This can allow the interpretation of immune responses to intracellular pathogens at single-cell resolution, as bulk isolates are often heterogeneous [36]. Host-pathogen scRNA-seq has been comprehensively reviewed by Penaranda and Hung [37], but close consideration of how dual scRNA-seq, the simultaneous scRNA-seq analysis of a pathogen and its infected host, can be leveraged for specific vaccinology uses is warranted.

Several groups have recently used dual scRNA-seq to profile virally infected cells and draw insights from transcriptome information [38–42]. O'Neal et al. revealed the feasibility and value of West Nile Virus- (WNV-) inclusive scRNA-seq as a method for single-cell transcriptomics and WNV RNA detection [38]. There was extreme heterogeneity in viral quantity and antiviral gene expression among in vitro infected cells. The expression of IFN-stimulated genes in single cells negatively correlated with intracellular viral RNA abundance. Selecting vaccine adjuvants or platforms to generate IFN-stimulated gene products may therefore be a promising WNV vaccine approach. Similarly, Russell and colleagues profile host and influenza virus mRNAs from a variety of cell lines early after in vitro infection [39]. They demonstrate astonishing differences in the transcriptional load resulting from influenza infection between ostensibly identical cells in spite of a relatively pure influenza inoculum. These types of approaches could be implemented ex vivo in the case of human challenge-compatible pathogens such as influenza to provide insights into the cellular characteristics associated with vaccine-induced protective immunity.

2.3. Human Challenge Studies.

Responses to infection can also be interrogated in a more regulated setting using controlled human infection models. These involve the direct inoculation of an infectious agent in order to evaluate the subsequent immune response and/or potential protective efficacy of interventions. Barton et al. have already discussed the use of transcriptomics in controlled human infection models [43]. Here, we discuss specific cases where scRNA-seq could be leveraged for particular benefit in vaccine development.

In the context of malaria, Tran et al. set out to profile the differences in the bulk blood transcriptome of challenge-protected and challenge-nonprotected volunteers during and after malaria immunisation [44]. The authors found robust transcriptomic changes, four weeks after immunisation, that were unique to protected volunteers. These changes included T cell, NK cell, protein synthesis, and mitochondrial signatures. The authors detected similar signatures three weeks after controlled human malaria infection (CHMI) and hypothesised that ongoing T cell memory response and clearance of antigen was driving this signal and mediating protection from malaria. scRNA-seq could extend these analyses by characterising immune subsets associated with protection and by identifying TCR clones, across time points, that share enrichment of hypothesised protective pathways to determine if clones are maintained and if there is a clonal dominance in protected individuals.

Mpina and colleagues assessed variations in NK, NKT, and MAIT cell populations using samples from a CHMI study of Tanzanian adults challenged with *P. falciparum* parasites [45]. CHMI decreased MAIT cell frequencies during blood-stage malaria and was followed by a "rebound" increase in circulating MAIT frequency for up to 168 days post CHMI. These cells showed distinct single-cell RNA expression profiles at each time point suggesting that MAIT cells respond to sporozoite challenge by day 9 and do not return to baseline transcriptional in the time period examined. Although MAIT cells express an invariant α chain, there are differences in the CDR3 junctions that allow tracking of cell clones. Interestingly, there were no changes in the relative distribution of MAIT cell clones by day 28 after CHMI. This paper demonstrates that the utility of scRNA-seq is not only restricted to interrogating the adaptive

immune response. Innate-like and innate cellular responses are likely critical, and underappreciated, to adjuvant and vaccine efficacy [46], as such scRNA-seq is a tool that can unlock insights from human challenge models.

It is clear that scRNA-seq could provide valuable insights for vaccine redesign and targeting in the context of controlled human infection models. Indeed, owing to the relatively common use of these models in malaria vaccine development, most of the published analyses relate to this pathogen. Investigation of other pathogens used in challenge models, such as influenza, *Salmonella typhi*, and BCG [47–49], will likely benefit from scRNA-seq adoption. There is likely to be a vast number of blood samples that have already been obtained in these types of studies, on which it is not possible to perform scRNA-seq, for example, blood sampled in PAX-gene Blood RNA Tubes or bulk RNA extracted from peripheral blood mononuclear cells (PBMC). CIBERSORTx [50], MuSIC [51], Bseq-SC [52], and deconvSeq [53] are tools that could be used to deconvolute bulk RNA-seq data derived from these samples. Use of CIBERSORT [54] suggests that *P. vivax* CHMI induces potent immunosuppression mediated by dendritic cells and that this immunosuppression conditions subsequent antimalarial immunity [55]. While these tools are incredibly useful in situations where bulk RNA has been isolated, they function by using known single-cell transcriptional signatures and do not serve as discovery tools for new pathways and cellular heterogeneity. Thus, there is importance of implementing and comprehension of scRNA-seq techniques by vaccinologists.

3. Vaccine Evaluation

The evaluation of the immune response to vaccination in both the preclinical and clinical phases is central to the prediction of success in disease protection. Antibody titres are correlates of protection for many, if not most, vaccines and vaccine candidates [62]. These can be measured with minimal sample preparation and equipment, and as a result, we rely heavily on titres as correlates of protection. Antibodies, however, are not always sufficient for protection. Cellular immunity can kill or suppress intracellular pathogens and may synergise with antibodies. T cell responses are indispensable and probably understudied as correlates of protection. Assessment of B and T cell clonality induced by vaccines, vaccine-induced cell phenotypes, and transcriptional signatures associated with protection are all important avenues of investigation that can be achieved through scRNA-seq. Here, we focus on examples of scRNA-seq used as a tool for comparison of vaccination routes and regimens, applications for BCR and TCR analyses, and also comment on cancer vaccines.

3.1. Comparing Vaccine Regimens and Responses. Immunogenicity of vaccines is modulated by a number of factors including vaccine antigen, vaccine platform, and adjuvant. scRNA-seq allows the impacts of these to be investigated with a high degree of specificity. scRNA-seq can also be used to characterise the heterogeneity in response to different vaccine regimens. Sheerin et al. performed comparative transcriptomics of the response to the capsular group B meningo-

coccal vaccine (4CMenB), administered concomitantly with other vaccines and on its own, and the response to its constituent antigens or several comparator antigens in mice ([63]; Table 1). The authors found neutrophil-specific genes were enriched at 24 hours following 4CMenB vaccination using bulk RNA-seq. To resolve whether concomitant immunisation resulted in differential transcriptional activity of neutrophils, scRNA-seq was performed on neutrophils isolated from 4CMenB + routine vaccines or routine vaccines-only immunised groups. Here, scRNA-seq has been used to characterise the specific innate immune receptor genes involved in different vaccine regimens providing mechanistic insight into differences in reactogenicity.

Vaccine human challenge studies (VHCS), a subset of controlled human infection models discussed above, involve the direct evaluation of vaccine efficacy by administration of an infectious agent to human volunteers after vaccination. Since the inception of VHCS such as those by Theodore Woodward in the 1940s [64], the means to evaluate vaccines have expanded past vaccine efficacy, clinical symptoms, and antibody titres. To date, however, there have been no published studies utilising scRNA-seq in the context of VHCS. This technique could be used to dissect the heterogeneity of cellular responses in participants with full, partial, or absent vaccine-induced protection after infectious challenge. Additionally, monitoring alterations in the single-cell gene expression profile of the challenge microbe itself could significantly improve understanding of host-pathogen interactions. Dunachie et al. performed whole transcriptome profiling of two groups receiving different malaria vaccine combinations [65]. The transcriptomes of stimulated PBMCs from three participants fully protected from malaria showed enrichment of modules associated with IFN induction and antigen presentation. This was conserved across vaccine regimens and suggests a common vaccine-induced protective pathway. scRNA-seq could extend these analyses by dissection of the particular cell types that contribute to these modules and their relative contributions.

In addition to effects modulated by vaccine regimen, platform, and immunisation route, there is ample evidence to suggest that gene expression following immunisation can be affected by adjuvant selection. Transcriptomic evaluation of nonhuman primates (NHP) and human responses to vaccine adjuvants, with and without vaccination, has largely been restricted to bulk and/or microarray analyses [66–69]. Using an NHP model to compare responses against the HIV Env antigen with eight different adjuvants, Francica et al. [70] demonstrated using microarrays with whole blood RNA that TLR4 or 7 agonists had differential effects on the upregulation of inflammatory and IFN genes. Furthermore, these gene signatures correlated across all adjuvant groups with antibody Fc functionality. Such comparative adjuvant studies in NHPs can support downselection of candidates for translation to humans and provide better targeting of pathogen-specific immune responses. There are thus opportunities for scRNA-seq approaches to further our understanding of adjuvant mode of action, which remains limited, and its association with immunogenicity and/or efficacy. Specifically, combining single-cell technologies with serial tissue sampling

Generating, curating, and characterising single-cell datasets of well-known pathogen isolates at various stages of infection will allow exploration of pathogen diversity and plasticity, which will ultimately aid in the identification of vaccine targets. The transcriptional variation of several malaria parasite species has been interrogated using scRNA-seq [56]. Transcriptomes derived from 10 life-cycle stages of a rodent malaria species, *P. berghei*, demonstrated 20 gene modules that encode proteins sharing related roles and that are cotranscribed in the parasite life cycle. This work allowed for the deconvolution of human disease-causing species of the malaria parasite isolated from clinical samples. The close grouping of functionally related genes and the downselection of proteins whose genes are not expressed in humans provide hope for better prioritisation of genes to develop vaccines. For example, targets for preerythrocytic malaria vaccines might be found in the gene expression data for the cluster that includes the well-studied circumsporozoite protein, the target of the RTS, S vaccine. Therefore, this is a clear example where systematic scRNA-seq can support applied research efforts in the identification of new targets for vaccine and drug development. This is especially the case in situations where widespread genetic knockdown/out screening is limited.

International collaborations including the Human Cell Atlas [57] and TCGA [58] are progressing, allowing comprehensive scRNA-seq profiling of a diverse range of cells. Szabo and colleagues used scRNA-seq of more than 50,000 T cells from multiple healthy tissues and peripheral blood to develop a reference map of T cell activation and homeostasis in various sites. The authors projected the single-cell expression profiles of tumour-associated T cells derived from several different human cancers onto the map of healthy T cell activation conditions. In doing this, the authors provide insights into the activation states of T cells in tumours, for example, tumour-associated T cells contain activated CD8+, but are devoid of activated CD4+, T cell states. This and other such reference maps could be used to measure and compare T cell activation in tissues and blood following vaccination.

Equally, new scRNA-seq reference maps following vaccination could provide vaccine-induced signatures that are known to correlate with long-lived protection from disease and/or infection. For example, one could envisage a viral vector response reference map, where the immunological responses to several virally vectored vaccines are profiled. More importantly, vaccination is almost exclusively performed in young children and early adolescence [59]. Adult, paediatric, and neonatal immune systems respond differently to infectious and inflammatory insults. It is therefore crucial to understand early life immune systems and develop appropriate transcriptomic reference atlases separately from that of adults. More broadly, there are scientific and ethical imperatives to be as globally inclusive as possible in the generation of these atlases, as failure to do so will only continue the trend of genetic studies involving and benefiting those of European descent [60, 61].

Box 1: scRNA-seq reference atlases.

Table 1: Summary of prophylactic vaccinology publications using scRNA-seq.

Reference	Cell type	Species	Vaccine pathogen	scRNA-seq method
Afik et al. 2017. *Nucleic Acids Res.* [78]	CD8+ T cells	Human	Yellow fever	Smart-seq
Upadhyay et al. 2018. *Genome Med.* [80]	Plasmablasts	Human; Rhesus	Influenza; SIV	Smart-seq
Neu et al. 2019. *J Clin Invest.* [92]	Plasmablasts	Human	Influenza	Smart-seq/Spec-Seq
Cirelli et al. 2019. *Cell.* [81]	B cells	Rhesus	HIV	Smart-seq
Waickman et al. 2019. *Nat Commun.* [87]	CD8+ T cells	Human	Dengue	10X
Sheerin et al. 2019. *Sci Rep.* [63]	Neutrophils	Mouse	Neisseria meningitidis serogroup B	10X
Darrah et al. 2020. *Nature.* [133]	T cells	Rhesus	Tuberculosis	Seq-Well

may allow superior insight into immune cell activation at the site of immunisation and in the draining lymph nodes, events which can be difficult to capture using other technologies.

3.2. BCR and TCR Analyses. For many pathogens, our understanding of protective epitopes is incomplete. In the study of B cells, peptide arrays and phage displays are methods that have been used for the discovery of linear epitopes [71, 72]. The search for T cell targets has frequently focused on HLA-A∗02-restricted epitopes, prevalent in individuals of Caucasian ancestry [73]. Epitope discovery relies on the identification of short peptide sequences within the longer peptide product of the pathogen that evoke an immune response. This can be determined experimentally using functional assays such as interferon gamma ELISpot to identify T cell responsiveness to different peptide sequences or computationally using peptide binding prediction algorithms (reviewed in [74, 75]).

scRNA-seq methods have the potential to improve antigen screening and selection, by providing a more accurate picture of the immune response generated by vaccines with different antigenic make-ups. The development of algorithms that reconstruct T cell receptor sequences from single-cell data allows parallel analysis of the T cell transcriptome and TCR clonotype in multimer sorted antigen specific cells [76, 77]. Redmond et al. describe single-cell TCRseq, a bioinformatic pipeline which identifies mRNA reads that align to genes coding for the variable (V) and constant (C) regions of the T cell receptor (TCR), to reconstruct paired alpha and beta TCRs. Eltahla et al. describe an alternative

computational method, VDJ puzzle, to reconstruct TCR sequences from scRNA-seq data. They show that this method can identify TCR clones that are not detected by single-cell PCR of the same sample and can link this analysis to the transcriptome of a cell. These algorithms can be paired with vaccine-reactive cell identification by fluorescent probes conjugated to vaccine antigens. For example, vaccine-reactive T cells were identified following a single dose of a yellow fever vaccine and were characterised by scRNA-seq with TCR sequence linkage [78].

Methods using the transcriptome to identify vaccine-responsive T cells without knowing their epitope specificity have also been developed, which allows a broad understanding of the T cell repertoire in response to vaccination. Fuchs et al. used scRNA-seq to identify a genetic signature of virus-responsive cells by performing scRNA-seq on dye-labelled antigen-specific cells [79]. They identified key genes that are differentially expressed in virus-responsive cells compared to unresponsive cells and show the combined expression of TNFRSF9, XCL1, XCL2, and CRTAM can be used to identify antigen-specific T cells from an undifferentiated T cell population. This enables identification of antigen-responsive cells without the need to sort populations by fluorochrome-labelled multimer prior to sequencing. Further studies of vaccine responses in controlled trials could provide better evidence upon which to select the most immunogenic vaccine epitopes.

Upon antigen stimulation, B and T cells proliferate and undergo clonal expansion; the BCR or TCR sequences are effectively a "clonal barcode." This can provide information on antigen specificity and cell ancestry. A great strength of scRNA-seq is the ability to obtain unbiased transcriptome and V (D) J gene transcript usage information from the same cell. With the advent of new scRNA-seq workflows, a proliferation of bioinformatics tools to analyse these data has necessarily occurred. BALDR, an example of such a bioinformatic pipeline, is able to reconstruct the paired heavy and light chain immunoglobulin gene sequences from scRNA-seq data derived from Illumina short reads ([80]; Table 1). Upadhyay and colleagues applied BALDR to single plasmablasts, splenic germinal centre B cells, and memory B cells following vaccination of rhesus macaques with an HIV vaccine. The reconstruction accuracy was 100% for the plasmablasts and >80% for the other cell populations, demonstrating that even in rhesus macaques, a species with poor annotation of the immunoglobulin loci, BALDR can recreate paired antibody sequences. This tool has broad applications in vaccinology and has already been used in the preclinical assessment of HIV vaccines ([81]; Table 1).

In a similar fashion, TraCeR [82] and other tools [76–78] allow for the reconstruction of full-length, paired T cell receptor sequences from T cell scRNA-seq data. These methods, however, are limited by either their reliance on the plate-based Smart-seq2 method, Illumina short-read sequencing, the need for a large number of sequencing reads [83], or a failure to integrate switching diversity and alternative mRNA splicing involving the $3'$ end of immunoglobulin heavy chain mRNA [84]. Repertoire and Gene Expression by Sequencing (RAGE-Seq) provides a means to sequence full-length antigen receptor transcripts with Oxford Nanopore

sequencing and link this with short-read transcriptome profiling at single-cell level [85]. This can also be applied to droplet-based scRNA-Seq workflows. As the availability of commercial assays to analyse antigen receptors—such as 10x Genomics V (D) J library preparation kits—becomes more available, cost-effective, and user-friendly, antigen specificity and ancestry of T and B cells following will be more effectively understood.

In the context of dengue virus (DENV) vaccination, approaches centred solely on B cell-mediated protection have limitations [86]. CD4+ T cells are required to generate and maintain B cell responses and CD8+ T cells are critical for eliminating infected cells. Waickman et al. use scRNA-seq to assess the diversity and long-lived nature of CD8+ T cell responses to an experimental tetravalent DENV vaccine ([87]; Table 1). The investigators followed T cell responses from acute activation to memory time points and compared the overlap of TCR clonotypes. By looking at the occurrence of T cells with an expanded TCR clonotype and by enumerating the presence of these TCRs at late time points, the authors were able to identify a memory precursor subset of CD8+ T cells. These memory precursors had enrichment of cellular metabolism and proliferation gene pathways. Waickman and colleagues demonstrate the ability of scRNA-seq as a tool to accurately and longitudinally track vaccine-antigen-specific T cells across time to identify correlates of T cell-mediated immunity with single-cell resolution. Importantly, scRNA-seq is used in combination with flow cytometry as means for both validation of discovered markers and further hypothesis testing.

Using scRNA-seq data, it is possible to predict cell trajectories, that is, to computationally order cells along putative trajectories, by inferring how much progress an individual cell has made through a given process (such as cell differentiation). The above analyses by Waickman et al. could be extended by using pseudotime tools such as Monocle [88], Slingshot [89], or others (which are benchmarked by Saelens et al. [90]) to infer a trajectory of T cell differentiation. One could then infer as to whether further differentiation is correlated with shared or "public" TCR clones. RNA velocity [91], a high-dimensional vector predicting the future state of an individual cell at a scale of hours, could also be used to infer the directionality of these cell state progression trajectories.

Pairing TCR sequence and transcriptome information allows the discovery and exploration of new cell populations. Afik and colleagues performed scRNA-seq on Yellow Fever Virus (YFV) vaccine-reactive and other CD8+ T cells ([78]; Table 1). Unexpectedly, YFV-specific cells were found across two clusters: one containing effector memory CD8+ T cells and one containing naive CD8+ T cells, thus defining "naïve-like" and "effector memory-like" YFV-specific CD8+ T cells. Combined TCR-transcriptome analysis revealed that the CDR3 sequence was longer in naive-like compared to memory-like cells for both alpha and beta TCR chain. In addition, the authors compared scRNA-seq transcript abundance to protein level by reviewing index sort flow cytometry data, demonstrating the utility of coupling scRNA-seq with other omics technologies. This is an example of how the heterogeneity in the T cell compartment can be investigated

Recently, efforts to produce and characterise monoclonal antibodies (mAb) have made impressive progress. mAbs with broadly neutralising activity against specific antigens largely act through their Fab fragment specificities.

Our comprehension of the way antibody specificities interact with B cell function has remained limited due to the intricacies of polyclonal antibody responses. Neu et al. developed the Spec-seq protocol to tackle this challenge ([92]; Table 1). Spec-seq permits mAb production and transcriptional profiling from the same cell. This involves modifications to the Smart-seq2 scRNA-seq method [93] and the harmonisation of two mAb generation techniques [94, 95]. The authors apply Spec-seq to analyse plasmablasts following influenza vaccination, characterising differences in transcriptome according to BCR isotype and vaccine reactivity. IgA vaccine-reactive circulating plasmablasts were compared with IgG vaccine-reactive and IgA vaccine-nonreactive populations. The authors suggest that imprinting following initial B cell activation results in transcriptional similarities among clonally related plasmablasts. Total and IgA plasmablasts cluster in tSNE projections by their ability to bind the vaccine, suggesting transcriptional modules correlated with vaccine reactivity. These data highlight an unexpectedly high degree of transcriptional specialisation within plasmablasts, demonstrating that these cells are not terminally differentiated antibody-secreting cells. Spec-seq is a robust technique that allows the profiling of the transcriptome and mAbs of hundreds of cells.

With improvements in droplet-based scRNA-seq methods, the scale (thousands of single cells) and order of events (BCR sequence and transcriptome information first, mAb generation thereafter) could be changed. Rather than characterising every mAb "blind" (i.e., without any prior information on its cell of origin), mAbs could be selected on the basis of transcriptome information and clonal family position in a hope to only generate and characterise high affinity/avidity candidates. Combining this with fluorescence-assisted cell sorting (FACS) to isolate vaccine antigen-specific B cells could provide a powerful new workflow to produce monoclonal antibodies against specific pathogens. An example of this type of workflow is put forward by Goldstein and colleagues [96] and is already being put to use for mAb development in priority diseases (see Box 3) [97].

Box 2: scRNA-seq as a starting point for monoclonal antibody production.

following vaccination and accounted for by cell state and TCR properties.

3.3. Evaluating Vaccination in High-Risk Populations.
Immunocompromised individuals are at risk of higher acquisition and complication rates of many vaccine-preventable infectious diseases such as seasonal influenza, respiratory syncytial virus (RSV), and bacterial pneumonia. In parallel, immunological responses to vaccination are often less efficient compared to healthy adults [98–101].

Optimising vaccine immunogenicity in immunosuppressed populations is paramount, but heterogeneous underlying mechanisms of immunosuppression make this challenging. Causes of altered immune states range from pathological conditions (including primary immune deficiencies and/or acquisition of chronic viral infection with HIV or cytomegalovirus (CMV)) to physiological states (including neonates, pregnant, and older persons) and iatrogenic immunosuppression following organ transplant or treatment for autoimmune conditions. With respect to pathological conditions, coinfections may further complicate vaccination, for example, HIV with hepatitis B or hepatitis C virus.

Study of the immune response at the cellular level in conditions of immunosuppression has demonstrated the nuances of vaccination responses. For example, the contribution of humoral and cellular responses to both influenza and RSV vaccination is altered in older compared to younger adults [102–105]. Contrastingly, CMV infection is associated with attenuated vaccine responses across age groups [106–108]. scRNA-seq is ideally placed to discover how these observations fit together, by profiling heterogeneous cell populations in detail. Clusters of cellular functional networks can be identified, which may in turn uncover key transcriptional drivers that can be targeted to optimise vaccine responses.

scRNA-seq could be utilised in the study of alternative vaccine regimens in the context of high-risk populations, to evaluate changes in immunogenicity. For example, adjuvant and increased dose influenza vaccines have been trialled in multiple settings of immunosuppression including older persons and transplant recipients (whose immunosuppressive regimen can vary) [98]. Detailed immunological profiling has not been conducted comparing, head to head, the response to a given vaccine in these different groups. scRNA-seq would be especially useful at discovering common signatures of immunologically successful vaccination regimens, especially in less common settings such as organ transplantation.

3.4. Personalised Cancer Vaccines.
Cancer vaccines are different from those protecting against infectious diseases in many ways; most notably that they can be used in therapeutic and personalised capacities [109]. Samples from tumour biopsies or resections are composed of a variety of cell types including tumour, immune, and mesenchymal cells. This assortment of cell types can confound bulk RNA-seq data. Single-cell techniques can tackle this challenge.

Single-cell transcriptomic profiling can be incorporated into neoepitope selection and vaccine manufacture workflows. Petti and colleagues performed matched whole-genome sequencing and droplet-based scRNA-seq on samples from patients with acute myeloid leukaemia [110]. They were able to discriminate between tumour and wild-type cells, identify abnormally differentiated tumour cells, and discover mutation-associated transcriptional profiles. In doing so, the authors identified surface markers that could be used to purify and analyse subclones for downstream studies. The approach developed in their paper could theoretically be applied to any cancer type. In this case, scRNA-seq can narrow personalised cancer vaccine workflows to mutations

of expressed genes. This approach, albeit using bulk RNA-seq, has already gained traction in the field of melanoma vaccines [111, 112]. CITE-seq uses scRNA-seq in addition to oligonucleotide-labelled antibodies to allow simultaneous protein and transcriptome measurement at a single-cell level [13]. This technique may further enhance neoantigen discovery pipelines, by providing insight into protein expression given dynamic ranges of transcription [113].

Recognising that therapeutic responses are varied in pathologically identical tumours and even in genetically homogeneous cancer cells [114], scRNA-seq provides the ability to discern which cells respond favourably to vaccines. Single-cell sequencing has already been used in the optimisation of targeted drug treatment against metastatic renal cell carcinoma [115]. The authors' approach was able to identify the most effective drug combination for multiple subpopulations of tumour cells. Such approaches could potentially be used in patient xenograft models to determine the most effective and personalised vaccine combinations. Ott et al. analysed the transcriptome of neoantigen-reactive CD4+ single T cells using HLA class II tetramers before and after vaccination [112]. The main gene expression changes reflected transitions from naïve to effector and memory states. These changes included repression of genes promoting homeostatic survival of naïve T cells and upregulation of genes involved in Th1 fate polarisation. Therefore, implementation of scRNA-seq in evaluation of cancer vaccine responses is feasible and has already begun.

Single-cell sequencing must overcome a number of challenges prior to its wholesale adoption in the field of cancer vaccines. These hurdles are present in many scRNA-seq experiments but have specific consequences in cancer vaccine discovery. A major difficulty is "drop-out." This happens when a transcript or an allele in a heterozygous mutation is not captured or amplified and can occur at 10–50% of mutation sites [5, 116]. While there are computational models that can correct for this drop-out [117], this is of critical importance in cancer vaccinology as it may be impossible to determine whether a cell is truly wild type for a given mutation. Partial transcript coverage is confined to end-biased platforms such as the 10x Genomics Chromium platform. Coverage decreases nonlinearly across the length of a transcript, so some variants are much more easily detectable than others. The utility of scRNA-seq to detect expressed mutations is therefore dependent on the specific mutational composition of the tumour and sample in question. As a result, scRNA-seq will likely perform better in mutation detection for cancers with high mutation burdens [110]. The increase in throughput of techniques such as G&T-seq, allowing simultaneous transcriptomic and genomic profiling from single cells, could potentially circumvent some of these hurdles to scRNA-seq adoption in cancer vaccine workflows.

4. Challenges Associated with the Uptake of scRNA-seq

The wholesale adoption of scRNA-seq in vaccinology is limited by hurdles relating to analysis, technical issues, and the experimental questions that can be asked with the technology (see Appendix A for further considerations). For further discussion of the challenges in driving scRNA-seq and other single-cell techniques forward, we direct you to the review by Lähnemann and colleagues [128].

While there is increasing development of computational tools and reference databases [129, 130], it will become progressively more difficult to compare studies based on primary analyses of raw data. It is impractical to suggest that a single analysis pipeline could satisfy the needs of all primary analyses; however, there is a huge potential for studies performing secondary analyses on multiple published raw datasets. This is of particular relevance when considering comparisons of vaccine regimens, screening for correlates of vaccine-induced protection, and evaluating adjuvants. This will necessitate open data access and thorough annotation of experimental methods and parameters.

There are several challenges related to the scRNA-seq technology itself. Every scRNA-seq protocol begins with the preparation of a single-cell suspension of the tissue of interest. When making inferences from scRNA-seq experiments, there are two inherent assumptions related to this step. These are that (i) the cellular composition of the suspension is a faithful representation of the original tissue and (ii) sample preparation results in insignificant (or no) transcriptional changes. Enzymatic, as well as mechanical, dissociation can result in biases of cellular representation as certain cells may be more sensitive to enzymes or dissociation [131]. The latter assumption has also been challenged [132]. Several sources have suggested that isolation of nuclei and their use for single-nucleotide RNA-seq could help to mitigate both of these challenges [131, 133, 134] (see Appendix A).

By beginning reverse transcription using a poly (T)-oligonucleotide, the majority of scRNA-seq technologies use the mRNA polyA tail to synthesise the first strand of cDNA. Challenges with this approach include an inability to capture nonpolyadenylated microRNAs and regulatory RNAs [135], although techniques have been developed to overcome these hurdles [136]. This is of particular relevance in vaccinology as miRNAs are important in B cell antibody affinity maturation, may modulate infection susceptibility and vaccine responses, and can be used as immune modulators and adjuvants [137]. Additionally, scRNA-seq is estimated to detect 30% of mRNA molecules present in a cell, while other approaches such as seq-FISH detect >80% of targeted transcripts in situ, with recent iterations sure to increase this coverage [20, 138]. However, seq-FISH can be difficult to adopt in nonspecialist laboratories due to the microfluidic and microscopic expertise needed. In time, improvements in the sensitivity of conventional scRNA-seq methods and/or increased usability of in-situ hybridisation based scRNA-seq methods will undoubtedly lead to greater insights.

Ultimately, most cell-cell and extracellular cell-pathogen interaction is protein-mediated. For scRNA-seq, inferences about cell-cell interactions occurring between receptor-ligand pairs can be made using repositories of ligands, receptors, and their interactions, such as CellPhoneDB v2.0 [139]. This, and other methods [140], can be used to characterise communication among cells in homeostatic as well as pathological conditions. In many scRNA-seq studies, the

In December 2019, a cluster of patients with pneumonia of unknown cause was linked to a seafood wholesale market in Wuhan, China, later confirmed to be associated with infection with a new beta-coronavirus, now known as SARS-CoV-2 [118, 119]. As of 9 June 2020, over 7,131,261 cases of and 406,807 deaths due to the COVID-19 pandemic have occurred, with over 188 countries affected. Public health measures required to limit transmission have been drastic and are unlikely to be completely lifted without a vaccine (or highly efficacious prophylaxes and therapeutics). scRNA-seq can contribute to COVID-19 vaccine design, development, and evaluation.

Using Existing Reference Atlases
Understanding how the virus interacts with the host has been aided by the COVID-19 Cell Atlas (http://www.covid19cellatlas.org). This collaborative has investigated the virus's tropism by plotting the expression of genes encoding viral entry proteins in previous scRNA-seq datasets [120, 121]. SARS-CoV-2 entry genes have the highest expression in nasal goblet and airway ciliated cells [120]. Using these data to infer cell susceptibility to infection could aid in vaccine design: it would be reasonable to consider intranasal/mucosal and/or inhaled administration routes for COVID-19 vaccines. This would be in the hope of inducing local, site-specific immunity stopping the virus from entering the vasculature and organs. Indeed, Wen et al. used scRNA-seq to show that IgA, a secretory immunoglobulin, was overrepresented in early COVID-19 responses, compared to healthy controls, and may have been associated with a reduced recovery time [122].

Profiling Cellular Immunity to Tailor Vaccine Responses
A logical next step would be to generate a high-resolution multiomic cell atlas of the host immune response to COVID-19 infection in the periphery and in lung tissue [123]. Liao et al. profiled bronchoalveolar lavage immune cells in COVID-19 patients [124]. Moderate cases were characterised by the presence of highly clonally expanded CD8+ T cells with tissue-resident features. These data suggest that SARS-CoV-2-specific tissue-resident memory T cells (TRM) enable control of the virus, protecting against disease progression. Further investigation is necessary to see whether TRM-targeted vaccination is a safe and efficacious strategy against SARS-CoV-2.

BCR Analyses to Support mAb Development
Using scRNA-seq to understand humoral immunity and guide vaccine-mediated antibodies against SARS-CoV-2 has shown particular promise. Wen and colleagues profiled PBMC of convalescing COVID-19 patients [122]. The BCR repertoire contained a number of highly expanded clones, notably biased usage of IGV genes and high pairing frequencies of IGHV3-23-IGHJ4 genes. Several groups have already reported the rapid identification of SARS-CoV-2-neutralising antibodies through scRNA-seq [97, 125]. Investigation of these antibodies is worthwhile as it can assist in identifying sites of vulnerability on SARS-CoV-2, antigen/epitope prioritisation and, eventually, rational design and redesign of a vaccine.

Research on COVID-19 is evolving quickly; many studies have not yet been peer reviewed and there have been concerns about the robustness of the peer review process of studies that have [126, 127]. Even if scRNA-seq has significant potential to aid in getting a vaccine to licensure, time will tell whether these and other studies provide proof of concept for using scRNA-seq in outbreak pathogen vaccine development.

Box 3: COVID-19 vaccine development using scRNA-seq.

transcriptome is used as an index of protein expression; however, several discrepancies limit comparisons that can be made (see Figure 2 in [141]). For example, the differences in half-lives of mRNA and the proteins they encode, intracellular trafficking, and posttranslational modifications all have the potential to alter the function of cells that ostensibly have comparable mRNA levels [141, 142]. Technologies such as CITE-seq, REAP-seq, and Ab-seq complement restrictions of solely RNA-based analyses. Combination techniques also present us with an opportunity to understand the dynamics of these systems as we can take advantage of the very different half-lives of the measured molecules; RNA levels can change at a timescale of seconds, while proteins can persist for days. Furthermore, this highlights scRNA-seq as a high-yield adjunct to be used in conjunction with more conventional vaccinology techniques, rather than a panacea.

The spatial position of cells in tissues strongly influences function, yet there remains no truly single-cell, unbiased, spatial transcriptomics approach. Several approaches, however, are reaching cellular resolution (e.g., Slide-seq with 10μm voxels) and/or are extremely highly multiplexed (e.g., SeqFISH+ which can measure up to 10,000 genes in single cells) [143–146]. The coming single-cell spatial transcriptomic approaches may be of particular use in the evaluation of vaccines that are mediated by site-specific immunity, for example, through lung tissue-resident memory T cells following tuberculosis vaccination [147].

5. Conclusion and Recommendations

Profiling the immune response to both natural and artificial pathogen exposure by scRNA-seq has advanced our ability to identify favourable immunological profiles. The capability of scRNA-seq to concurrently examine the global gene expression, antigen-specificity, clonality, and individual copy number variants (CNVs) and infer the developmental trajectory of immune cells offers a powerful toolbox to appraise host responses to vaccine candidates. Certain areas have not yet been tackled by scRNA-seq, including critical confounders of immunogenicity such as coinfections and age, vaccine platforms, and adjuvants. In addition, we were unable to find any studies primarily using scRNA-seq to

Hypothesis generation: areas in vaccinology that lend themselves to investigation by scRNA-seq
    (i) Cellular responses to various adjuvants and vaccines
    (ii) Transcriptional and antigen-specific responses to adjuvants and vaccines
    (iii) Site-specific immunity induced by vaccinations (for example, at mucosal surfaces following HIV vaccination, in the liver following liver-stage malaria vaccination, or in the lungs following tuberculosis vaccination [147])
        (a) Spatial transcriptomics, soon to provide resolution at the single-cell level, could also be used
    (iv) Single-cell transcriptomic signatures associated with neutralising antibody responses
    (v) The accordance between protective transcriptional signatures in vaccine human challenge studies in nonendemic and endemic countries

Experimental design: exploit the technology and make the most of precious samples
    (i) Assess the need for scRNA-seq as the primary experimental technique and contemplate whether the question can be answered by established techniques (e.g., ELISAs, ELISpots, and flow cytometry)
    (ii) Consider using bulk RNA-seq as an adjunct to scRNA-seq (e.g., bulk RNA-seq on all samples, with scRNA-seq on a subset to allow computational deconvolution)
    (iii) Tailor the particular type of scRNA-seq to the experimental question (e.g., will alternative splicing be of interest? Use full-length transcript profiling if so)
    (iv) Longitudinal gene expression and TCR/BCR profiling to track antigen-specific clones
    (v) Define heterogeneity of cellular response in protected individuals
    (vi) Generate monoclonal antibodies from TCR/BCR sequences for rational vaccine redesign
    (vii) Use scRNA-seq as "backbone" to other technologies (e.g., G&T seq/CITE-seq)
    (viii) Combine scRNA-seq with flow cytometry/FACS and/or magnetically assisted cell separation to isolate rare or specific cell types [148]
    (ix) Multiplex samples according to genotype to reduce sample preparation time, reagent and sequencing costs, and batch effects [149, 150]
    (x) Preserve leftover cells either by sorting into a plate or by preserving in fixative for later use (for example, in the reanalysis of an interesting sample)

Analytical considerations: bioinformatic requirements/suggestions
    (i) Plan with, budget, and include bioinformaticians who are capable of working with scRNA-seq data from study conception onwards
    (ii) Use freely available packages that are regularly maintained. See Table 1 in the review by Zeng and Dai [129] for a list of computational tools for scRNA-seq analyses
    (iii) Upload data files to Gene Expression Omnibus [151] with as much metadata as possible (and not contravening human subject guidelines, if using human samples) as both processed and raw data
    (iv) Consider submission of data to other relevant databases such as Human Cell Atlas and/or TCGA
    (v) Consider analysing a similar published dataset with the experimental dataset to increase statistical power and/or to ensure a novel pipeline reproduces results in a previously published study
    (vi) Make freely available the code which relates to bespoke analyses

Box 4: Recommendations for the implementation of scRNA-seq in vaccine research

better understand adverse events following vaccination, or any other vaccine safety metrics, with the exception of one study discussed above [63].

scRNA-seq is now a widespread research tool; the number of diseases and areas of research in which it is being applied is growing. Improvements in the scale of adoption, robustness and ease of use of reagents, instrumentation, and computational tools mean that scRNA-seq will continue to be used more. The utility of this tool in vaccine design and development is contingent on the particular questions which are being asked (see Appendix A).

Beyond the anticipated improvements in system efficiencies and the increased availability of reagents, what is the best way that scRNA-seq can be applied systematically in vaccine design, development, and evaluation? This will depend on specific hypothesis-driven, experimental, and analytical considerations. In Box 3, we present the specific case study of COVID-19 vaccine development to demonstrate the potential of this technology, and in Box 4, we illustrate more general applications of scRNA-seq in vaccine development and evaluation. Ultimately, scRNA-seq, and its integration with other single-cell systems, will elucidate further information that can be used to drive favourable responses following therapeutic and/or prophylactic vaccination. The context provided by information from classic measurements, such as growth inhibition assays, antibody levels, and cytokine secretion, and controlled human infection studies can further enhance the value of the observations that are made. The generation of integrated datasets is computationally intensive but can provide comprehensive characterisations of vaccine responses by increasing the power of statistical calculations or capturing a greater amount of heterogeneity in systems. These insights can be used to identify novel populations, previously unappreciated correlates, and biomarkers of favourable and protective immune responses for use in the systematic and streamlined assessment of vaccine candidates.

Appendix

A. Experimental Considerations for scRNA-seq in Vaccinology

(1) Research question considerations

   (a) Ample consideration of the particular research question to be answered is required for any experimental design, but in particular scRNA-seq experiments

   (b) Potential questions that can be asked in scRNA-seq experiments

   (i) Does vaccination result in previously uncharacterised single-cell states?

   (ii) What is the temporal sequence of cellular processes taking place after vaccination?

   (iii) How does T cell and/or B cell clonal diversity change in response to vaccination?

   (iv) What are the transcriptional differences among vaccine reactive/antigen-specific cells?

   (v) To what extent are vaccine-induced transcriptional changes reflected at the cellular (bulk vs. single-cell sequencing) and protein level (transcriptomics vs. proteomics)?

   (vi) What are the single-cell transcriptional differences in vaccine response between vaccines A and B?

   (vii) What is the single-cell transcriptional profile in peripheral lymphocytes (or organ) given a particular vaccine platform (e.g., virus-like particles), regardless of the antigen that is delivered?

   (c) Other specific experimental considerations will influence experimental design

   (i) Breadth vs. depth

   (1) Are lowly expressed genes of particular interest? For transcripts that are lowly expressed (e.g., transcription factors), full length sequencing approaches may be better than $3'$ methods [152]. In this case, the amount of sequencing should also be considered such that "saturation" is reached; that is, further sequencing does not lead to the discovery of more unique transcripts

   (2) Are rare cell types to be profiled? Increasing cell number and maintaining read depth relatively low allows more power to detect rare cell populations (that may exist at <1% in frequency) [5]. It has been suggested that to estimate several important gene characteristics, the most favourable sequencing depth is around one read per cell per gene [153]. Alternatively, enrich the cell type of interest by using FACS, followed by scRNA-seq and use methods such as GateID to predict "nonintuitive" gating strategies based on scRNA-seq data [154]

   (ii) Will comparisons be made between different conditions (e.g., prime alone vs. prime-boost)?

   (iii) What are the qualities and expression levels of the marker genes of cell types of interest? Transcriptional bursting can result in substantially different transcript quantities and apparent gene expression levels [155]

   (iv) What is the overall budget of the project? Cost/cell profiled is an often-used metric for budgeting scRNA-seq experiments

   (v) What facilities and expertise are available?

   (vi) Has technical and experimental advice been sought by nonconventional means (consider the active scRNA-seq community on Twitter (particularly #scRNAseq and #scQA), ResearchGate, medRxiv, bioRxiv, EMBI-EBI training (https://www.ebi.ac.uk/training), and the Galaxy platform (https://usegalaxy.org/))

(2) Sample preparation and processing considerations

   (a) Processing samples up to the point where scRNA-seq can be performed is a process that can greatly affect the outcome of the experiment

   (i) Are there unchangeable constraints on sample collection/processing times? What effects, if any, will these have on the results?

   (ii) Do samples have to be processed immediately or is there a window where gene expression will not be affected, without specific preservation?

   (iii) Will the samples be preserved (e.g., flash frozen, fresh frozen, or formalin-fixation and paraffin-embedding) [156]?

   (b) scRNA-seq usually requires mechanical or enzymatic dissociation of samples to produce a single-cell suspension. Certain factors will affect this process

   (i) How fragile/robust are the samples?

   (ii) How well is the tissue dissociated?

   (iii) Which single-cell isolation procedure will be used (e.g., microdissection, reverse emulsion

droplets, FACS into plates, and/or nanowell isolation)?

(iv) Are there preexisting protocols in the published literature that describe isolation techniques from the sample of interest (consider published, preprint, and commercial (e.g., 10x Genomics technical notes) literature)?

(v) Will the cells be preserved (e.g., cryopreservation using DMSO, methanol fixation, or storage in commercially available formulations to preserve cells and their RNA) [157]?

(vi) Can a method be used that does not dissociate tissue (e.g., Slide-seq) as a baseline to check for dissociation effects? This will likely add extra cost and require additional technical expertise

(c) Once cells (or nuclei) are in a suspension, the starting point for RNA capture is achieved

(i) Which is more apt and feasible to answer your specific research question, nuclei or cells isolated from the sample [134]? Nuclei-derived transcripts have a higher intronic/exonic read ratio as they contain proportionally more pre-mRNAs and there is potential that a narrower time period is profiled as mRNAs in the cytoplasm may have existed for longer

(ii) What are the characteristics of the cells of interest (e.g., size and adherability)? Volume can vary widely cell-to-cell, this affects the absolute number of transcripts and can be reflected in the detected number of genes per cell [158]. Adherability of a cell may affect processes before sequencing, such as FACS

(iii) What quantity of input material (cells or nuclei) is there?

(d) Assessing RNA quality. Before embarking on costly library preparation for every sample, it is important to ensure that the RNA has not significantly degraded:

(i) RNA quality is typically measured using the RNA Integrity Number (RIN) algorithm [159]. This test is performed by isolating RNA from a sample of interest (typically bulk or 10s-100s of cells) and performing RNA microcapillary electrophoresis. The algorithm uses multiple features from the resultant electropherogram trace to score quality of the RNA from 1 (most degraded) to 10 (least degraded-highest integrity)

(ii) Is there enough of the sample to produce a RIN score test on each sample? If the site, conditions, or timing of sample collections is variable, it may lead to differences in RNA degradation between samples (and batch effects, discussed below). If a RIN score step can be built into each experiment before embarking on library preparation, then it can prevent spending money on low-quality samples

(3) Replicates, scale, sequencing, and batch considerations

(a) Batch effects

(i) Batch effects are random technical artefacts which occur during handling/processing. If batches correspond to different biological conditions, then it is largely impossible to determine what differences are biological vs. artefacts

(ii) Avoid batch effects by

(1) sorting cells from different biological conditions into different wells of the same plate

(2) using genetic variants to post hoc assign sequenced cells back to their genetically unique donor

(3) using the expression of an inserted genetic construct (not recommended)

(4) using barcoded antibodies (Cell Hashing) to label samples after dissociation, but before cell-capture step, to multiplex samples [13, 160]

(iii) Batch effects may be corrected by a number of bioinformatics tools and/or packages [161, 162]

(b) Experiment scale

(i) How many cells will be tested largely depends on the level of heterogeneity of the sample and on the number of available cells

(ii) Plate-sorted single cells are limited in the amount that can be handled compared to microfluidic platforms, which enable studies with several thousands of cells (Figure 1)

(c) Method of amplification

(i) Either exponential PCR-based amplification or linear in vitro transcription (IVT) amplification is usually used. IVT incorporates less PCR bias and erroneous bias as it is based on an unamplified RNA template [164]
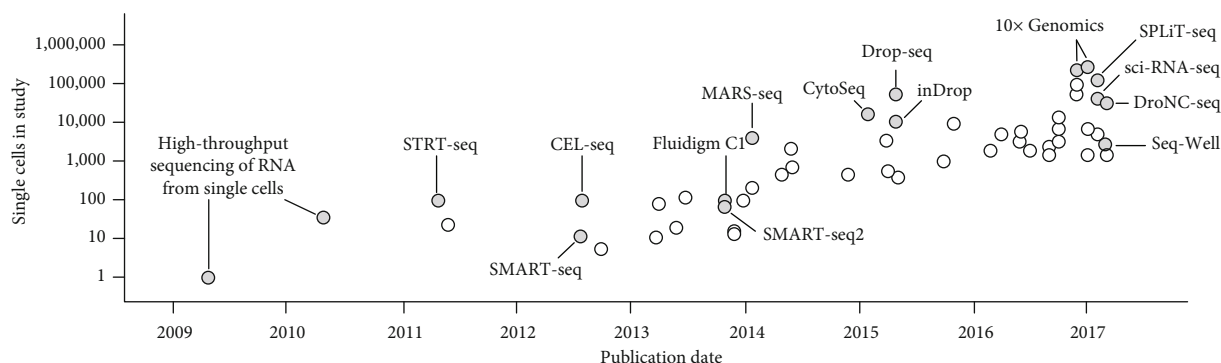
(d) Transcript position

FIGURE 1: scRNA-seq technologies that have been critical to allowing increments in experiment scale. Achievements over the past three years have more or less continued this pace; for example, combinatorial fluidic preindexing has increased the throughput of droplet-based single-cell RNA sequencing up to 15-fold. Figure adapted from references [133, 163].

(i) Some protocols provide full-length transcript data, whereas others amplify only the $3'$ or $5'$ ends of the transcripts

(ii) Non-full-length transcript methods allow increased throughput of cells, while full-length transcripts are advantageous if splice variants are important, looking to detect genetic variants or when studying species that have poorly annotated genomes

(e) Sensitivity

(i) This is the ability of an assay to capture an mRNA molecule from a single cell within the final library. Low sensitivity protocols have a disproportionate effect on weakly expressed genes (e.g., genes encoding cytokines)

(ii) If weakly expressed genes are to be evaluated, consider higher sensitivity methods or consider "clean-up" procedures such as rRNA removal [165]

(f) Ultimately, the particular protocol must be decided on an individual experiment basis. It is also important to note that while scRNA-seq methods have greater sensitivity than bulk RNA-seq methods, bulk methods have higher accuracy [166]

B. Review Search Strategy

We performed an initial scoping review of the literature using MEDLINE/PubMed to identify major themes present in the vaccinology literature. Thereafter, we performed predefined searches in the context of the vaccinology themes we identified: controlled human infection studies, correlates of protection, profiling the immune response to infection, cancer vaccines, understanding host-pathogen interactions, comparing vaccine regimens and responses, adjuvants, animals as models for human diseases and natural infections of livestock, antigen screening/selection, and effects of coinfec-

tion on vaccination. Predefined searches included using MeSH terms: "Sequence Analysis, RNA", "Vaccinology", "Vaccines." In January and February 2020, we searched for published literature in MEDLINE/PubMed and Embase and for preprint/not yet peer-reviewed literature in bioRxiv, medRxiv, Wellcome Open Research and arXiv. Further searches were conducted using free text to expand our capture of scRNA-seq (e.g., ("scRNA-seq" OR "single cell sequencing" OR "single cell RNA sequencing")), vaccine (e.g., (vaccin∗ OR immunis∗ OR immuniz∗)) and theme-specific studies (e.g., ("vaccine correlate∗" OR "correlate of protection" OR "correlates" OR "correlate of vaccine protection")).

After performing the searches, we refined the scope of our review to include only scRNA-seq studies—as opposed to those exclusively considering bulk RNA-seq—performed in humans or animal models of human disease published after 2015, unless they were considered sufficiently relevant to the narrative of our review.

Disclosure

Conflicts of Interest

Acknowledgments

References

[1] B. Kaijser, "Pneumococcal infections and the possible need for a vaccine," *Scandinavian Journal of Infectious Diseases*, vol. 11, no. 3, pp. 179–184, 1979.

[2] C. Lue, A. Tarkowski, and J. Mestecky, "Systemic immunization with pneumococcal polysaccharide vaccine induces a predominant IgA2 response of peripheral blood lymphocytes and increases of both serum and secretory antipneumococcal antibodies," *Journal of Immunology*, vol. 140, no. 11, pp. 3793–3800, 1988.

[3] E. De Gregorio and R. Rappuoli, "From empiricism to rational design: a personal perspective of the evolution of vaccine development," *Nature Reviews Immunology*, vol. 14, no. 7, pp. 505–514, 2014.

[4] T. Gomes, S. A. Teichmann, and C. Talavera-López, "Immunology Driven by Large-Scale Single-Cell Sequencing," *Trends in Immunology*, vol. 40, no. 11, pp. 1011–1021, 2019.

[5] A. Haque, J. Engel, S. A. Teichmann, and T. Lönnberg, "A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications," *Genome Medicine*, vol. 9, no. 1, p. 75, 2017.

[6] S. Picelli, Å. K. Björklund, O. R. Faridani, S. Sagasser, G. Winberg, and R. Sandberg, "Smart-seq2 for sensitive full-length transcriptome profiling in single cells," *Nature Methods*, vol. 10, no. 11, pp. 1096–1098, 2013.

[7] D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul et al., "Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types," *Science*, vol. 343, no. 6172, pp. 776–779, 2014.

[8] G. X. Y. Zheng, J. M. Terry, P. Belgrader et al., "Massively parallel digital transcriptional profiling of single cells," *Nature Communications*, vol. 8, no. 1, article 14049, 2017.

[9] A. M. Klein, L. Mazutis, I. Akartuna et al., "Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells," *Cell*, vol. 161, no. 5, pp. 1187–1201, 2015.

[10] T. M. Gierahn, M. H. Wadsworth, T. K. Hughes et al., "Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput," *Nature Methods*, vol. 14, no. 4, pp. 395–398, 2017.

[11] A. A. Kolodziejczyk, J. K. Kim, V. Svensson, J. C. Marioni, and S. A. Teichmann, "The Technology and Biology of Single-Cell RNA Sequencing," *Molecular Cell*, vol. 58, no. 4, pp. 610–620, 2015.

[12] G. Chen, B. Ning, and T. Shi, "Single-Cell RNA-Seq Technologies and Related Computational Data Analysis," *Frontiers in Genetics*, vol. 10, p. 317, 2019.

[13] M. Stoeckius, C. Hafemeister, W. Stephenson et al., "Simultaneous epitope and transcriptome measurement in single cells," *Nature Methods*, vol. 14, no. 9, pp. 865–868, 2017.

[14] V. M. Peterson, K. X. Zhang, N. Kumar et al., "Multiplexed quantification of proteins and transcripts in single cells," *Nature Biotechnology*, vol. 35, no. 10, pp. 936–939, 2017.

[15] M. Reyes, K. Billman, N. Hacohen, and P. C. Blainey, "Simultaneous Profiling of Gene Expression and Chromatin Accessibility in Single Cells," *Advanced Biosystems*, vol. 3, no. 11, p. 1900065, 2019.

[16] I. C. Macaulay, W. Haerty, P. Kumar et al., "G&T-seq: parallel sequencing of single-cell genomes and transcriptomes," *Nature Methods*, vol. 12, no. 6, pp. 519–522, 2015.

[17] S. S. Dey, L. Kester, B. Spanjaard, M. Bienko, and A. van Oudenaarden, "Integrated genome and transcriptome sequencing of the same cell," *Nature Biotechnology*, vol. 33, no. 3, pp. 285–289, 2015.

[18] I. C. Macaulay, M. J. Teng, W. Haerty, P. Kumar, C. P. Ponting, and T. Voet, "Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq," *Nature Protocols*, vol. 11, no. 11, pp. 2081–2103, 2016.

[19] M. Baron and I. Yanai, "New skin for the old RNA-Seq ceremony: the age of single-cell multi-omics," *Genome Biology*, vol. 18, no. 1, 2017.

[20] L. Chappell, A. J. C. Russell, and T. Voet, "Single-Cell (Multi)omics Technologies," *Annual Review of Genomics and Human Genetics*, vol. 19, no. 1, pp. 15–41, 2018.

[21] F. Krammer, "The human antibody response to influenza A virus infection and vaccination," *Nature Reviews Immunology*, vol. 19, no. 6, pp. 383–397, 2019.

[22] A. F. Cowman, J. Healer, D. Marapana, and K. Marsh, "Malaria: Biology and Disease," *Cell*, vol. 167, no. 3, pp. 610–624, 2016.

[23] R. Blecher-Gonen, P. Bost, K. L. Hilligan et al., "Single-Cell Analysis of Diverse Pathogen Responses Defines a Molecular Roadmap for Generating Antigen-Specific Immunity," *Cell Systems*, vol. 8, no. 2, pp. 109–21.e6, 2019.

[24] A.-K. E. Palm and C. Henry, "Remembrance of Things Past: Long-Term B Cell Memory After Infection and Vaccination," *Frontiers in Immunology*, vol. 10, 2019.

[25] W. Y. Lam, A. Jash, C.-H. Yao et al., "Metabolic and Transcriptional Modules Independently Diversify Plasma Cell Lifespan and Function," *Cell Reports*, vol. 24, no. 9, pp. 2479–2492.e6, 2018.

[26] W. Shi, Y. Liao, S. N. Willis et al., "Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells," *Nature Immunology*, vol. 16, no. 6, pp. 663–673, 2015.

[27] M. J. Price, D. G. Patterson, C. D. Scharer, and J. M. Boss, "Progressive Upregulation of Oxidative Metabolism Facilitates Plasmablast Differentiation to a T-Independent Antigen," *Cell Reports*, vol. 23, no. 11, pp. 3152–3159, 2018.

[28] Y. Fukazawa, R. Lum, A. A. Okoye et al., "B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers," *Nature Medicine*, vol. 21, no. 2, pp. 132–139, 2015.

[29] S. Nguyen, C. Deleage, S. Darko et al., "Elite control of HIV is associated with distinct functional and transcriptional signatures in lymphoid tissue CD8+ T cells," *Science Translational Medicine*, vol. 11, no. 523, article eaax4077, 2019.

[30] S. A. Valkenburg, N. H. L. Leung, M. B. Bull et al., "The Hurdles From Bench to Bedside in the Realization and Implementation of a Universal Influenza Vaccine," *Frontiers in Immunology*, vol. 9, p. 1479, 2018.

[31] B. Mordmüller, G. Surat, H. Lagler et al., "Sterile protection against human malaria by chemoattenuated PfSPZ vaccine," *Nature*, vol. 542, no. 7642, pp. 445–449, 2017.

[32] T. M. Wilkinson, C. K. F. Li, C. S. C. Chui et al., "Preexisting influenza-specific CD4$^+$ T cells correlate with disease

protection against influenza challenge in humans," *Nature Medicine*, vol. 18, no. 2, pp. 274–280, 2012.

[33] T. Ciucci, M. S. Vacchio, Y. Gao et al., "The Emergence and Functional Fitness of Memory CD4+ T Cells Require the Transcription Factor Thpok," *Immunity*, vol. 50, no. 1, pp. 91–105.e4, 2019.

[34] S. Rato, A. Rausell, M. Muñoz, A. Telenti, and A. Ciuffi, "Single-cell analysis identifies cellular markers of the HIV permissive cell," *PLoS Pathogens*, vol. 13, no. 10, article e1006678, 2017.

[35] T. Bradley, G. Ferrari, B. F. Haynes, D. M. Margolis, and E. P. Browne, "Single-Cell Analysis of Quiescent HIV Infection Reveals Host Transcriptional Profiles that Regulate Proviral Latency," *Cell Reports*, vol. 25, no. 1, pp. 107–17.e3, 2018.

[36] M. Ackermann, "A functional perspective on phenotypic heterogeneity in microorganisms," *Nature Reviews Microbiology*, vol. 13, no. 8, pp. 497–508, 2015.

[37] C. Penaranda and D. T. Hung, "Single-Cell RNA Sequencing to Understand Host–Pathogen Interactions," *ACS Infectious Diseases*, vol. 5, no. 3, pp. 336–344, 2019.

[38] J. T. O'Neal, A. A. Upadhyay, A. Wolabaugh, N. B. Patel, S. E. Bosinger, and M. S. Suthar, "West Nile Virus-Inclusive Single-Cell RNA Sequencing Reveals Heterogeneity in the Type I Interferon Response within Single Cells," *Journal of Virology*, vol. 93, no. 6, 2019.

[39] A. B. Russell, C. Trapnell, and J. D. Bloom, "Extreme heterogeneity of influenza virus infection in single cells," *eLife*, vol. 7, 2018.

[40] F. Zanini, M. L. Robinson, D. Croote et al., "Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 52, pp. E12363–E12369, 2018.

[41] Y. Steuerman, M. Cohen, N. Peshes-Yaloz et al., "Dissection of Influenza Infection In Vivo by Single-Cell RNA Sequencing," *Cell Systems*, vol. 6, no. 6, pp. 679–91.e4, 2018.

[42] S. Y. Kupke, L.-H. Ly, S. T. Börno et al., "Single-Cell Analysis Uncovers a Vast Diversity in Intracellular Viral Defective Interfering RNA Content Affecting the Large Cell-to-Cell Heterogeneity in Influenza A Virus Replication," *Viruses*, vol. 12, no. 1, p. 71, 2020.

[43] A. J. Barton, J. Hill, A. J. Pollard, and C. J. Blohmke, "Transcriptomics in Human Challenge Models," *Frontiers in Immunology*, vol. 8, 2017.

[44] T. M. Tran, E. M. Bijker, M. C. Haks et al., "Whole-blood transcriptomic signatures induced during immunization by chloroquine prophylaxis and Plasmodium falciparum sporozoites," *Scientific Reports*, vol. 9, no. 1, 2019.

[45] M. Mpina, N. J. Maurice, M. Yajima et al., "Controlled Human Malaria Infection Leads to Long-Lasting Changes in Innate and Innate-like Lymphocyte Populations," *Journal of Immunology*, vol. 199, no. 1, pp. 107–118, 2017.

[46] B. Pulendran, "Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology," *Nature Reviews Immunology*, vol. 9, no. 10, pp. 741–747, 2009.

[47] S. A. Harris, J. Meyer, I. Satti et al., "Evaluation of a human BCG challenge model to assess antimycobacterial immunity induced by BCG and a candidate tuberculosis vaccine,

MVA85A, alone and in combination," *The Journal of Infectious Diseases*, vol. 209, no. 8, pp. 1259–1268, 2014.

[48] T. C. Darton, C. Jones, C. J. Blohmke et al., "Using a Human Challenge Model of Infection to Measure Vaccine Efficacy: A Randomised, Controlled Trial Comparing the Typhoid Vaccines M01ZH09 with Placebo and Ty21a," *PLOS Neglected Tropical Diseases*, vol. 10, no. 8, article e0004926, 2016.

[49] A. C. Sherman, A. Mehta, N. W. Dickert, E. J. Anderson, and N. Rouphael, "The Future of Flu: A Review of the Human Challenge Model and Systems Biology for Advancement of Influenza Vaccinology," *Frontiers in Cellular and Infection Microbiology*, vol. 9, 2019.

[50] A. M. Newman, C. B. Steen, C. L. Liu et al., "Determining cell type abundance and expression from bulk tissues with digital cytometry," *Nature Biotechnology*, vol. 37, no. 7, pp. 773–782, 2019.

[51] X. Wang, J. Park, K. Susztak, N. R. Zhang, and M. Li, "Bulk tissue cell type deconvolution with multi-subject single-cell expression reference," *Nature Communications*, vol. 10, no. 1, 2019.

[52] M. Baron, A. Veres, S. L. Wolock et al., "A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure," *Cell Systems*, vol. 3, no. 4, pp. 346–360.e4, 2016.

[53] R. Du, V. Carey, and S. T. Weiss, "deconvSeq: deconvolution of cell mixture distribution in sequencing data," *Bioinformatics*, vol. 35, no. 24, pp. 5095–5102, 2019.

[54] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.

[55] A. F. Vallejo, R. C. Read, M. Arevalo-Herrera, S. Herrera, T. Elliott, and M. E. Polak, "Malaria systems immunology: Plasmodium vivax induces tolerance during primary infection through dysregulation of neutrophils and dendritic cells," *The Journal of Infection*, vol. 77, no. 5, pp. 440–447, 2018.

[56] V. M. Howick, A. J. C. Russell, T. Andrews et al., "The Malaria Cell Atlas: Single parasite transcriptomes across the completePlasmodiumlife cycle," *Science*, vol. 365, no. 6455, article eaaw2619, 2019.

[57] A. Regev, S. A. Teichmann, E. S. Lander et al., "The Human Cell Atlas," *eLife*, vol. 6, 2017.

[58] G. F. Gao, J. S. Parker, S. M. Reynolds et al., "Before and After: Comparison of Legacy and Harmonized TCGA Genomic Data Commons' Data," *Cell Systems*, vol. 9, no. 1, pp. 24–34.e10, 2019.

[59] O. Levy, S. Goriely, and T. R. Kollmann, "Immune response to vaccine adjuvants during the first year of life," *Vaccine*, vol. 31, no. 21, pp. 2500–2505, 2013.

[60] A. R. Martin, M. Kanai, Y. Kamatani, Y. Okada, B. M. Neale, and M. J. Daly, "Clinical use of current polygenic risk scores may exacerbate health disparities," *Nature Genetics*, vol. 51, no. 4, pp. 584–591, 2019.

[61] "Genetics for all," *Nature Genetics*, vol. 51, no. 4, 2019.

[62] S. A. Plotkin, "Correlates of protection induced by vaccination," *Clinical and Vaccine Immunology*, vol. 17, no. 7, pp. 1055–1065, 2010.

[63] D. Sheerin, D. O'Connor, C. Dold et al., "Comparative transcriptomics between species attributes reactogenicity pathways induced by the capsular group B meningococcal vaccine, 4CMenB, to the membrane-bound endotoxin of its

outer membrane vesicle component," *Scientific Reports*, vol. 9, no. 1, 2019.

[64] R. B. Hornick, W. E. Woodward, and S. E. Greisman, "Doctor T. E. Woodward's Legacy: From Typhus to Typhoid Fever," *Clinical Infectious Diseases*, vol. 45, Supplement 1, pp. S6–S8, 2007.

[65] S. Dunachie, T. Berthoud, A. V. S. Hill, and H. A. Fletcher, "Transcriptional changes induced by candidate malaria vaccines and correlation with protection against malaria in a human challenge model," *Vaccine*, vol. 33, no. 40, pp. 5321–5331, 2015.

[66] R. A. van den Berg, L. De Mot, G. Leroux-Roels et al., "Adjuvant-Associated Peripheral Blood mRNA Profiles and Kinetics Induced by the Adjuvanted Recombinant Protein Candidate Tuberculosis Vaccine M72/AS01 in Bacillus Calmette–Guérin-Vaccinated Adults," *Frontiers in Immunology*, vol. 9, 2018.

[67] T. S. M. Mathan, J. Textor, A. E. Sköld et al., "Harnessing RNA sequencing for global, unbiased evaluation of two new adjuvants for dendritic-cell immunotherapy," *Oncotarget*, vol. 8, no. 12, pp. 19879–19893, 2017.

[68] T. Olafsdottir, M. Lindqvist, and A. M. Harandi, "Molecular signatures of vaccine adjuvants," *Vaccine*, vol. 33, no. 40, pp. 5302–5307, 2015.

[69] A. C. D. Salyer and S. A. David, "Transcriptomal signatures of vaccine adjuvants and accessory immunostimulation of sentinel cells by toll-like receptor 2/6 agonists," *Human Vaccines & Immunotherapeutics*, vol. 14, no. 7, pp. 1686–1696, 2018.

[70] J. R. Francica, D. E. Zak, C. Linde et al., "Innate transcriptional effects by adjuvants on the magnitude, quality, and durability of HIV envelope responses in NHPs," *Blood Advances*, vol. 1, no. 25, pp. 2329–2342, 2017.

[71] K. L. Hess and C. M. Jewell, "Phage display as a tool for vaccine and immunotherapy development," *Bioengineering & Translational Medicine*, vol. 5, no. 1, 2020.

[72] T. Jaenisch, K. Heiss, N. Fischer et al., "High-density Peptide Arrays Help to Identify Linear Immunogenic B-cell Epitopes in Individuals Naturally Exposed to Malaria Infection," *Molecular & Cellular Proteomics*, vol. 18, no. 4, pp. 642–656, 2019.

[73] D. Middleton, L. Menchaca, H. Rood, and R. Komerofsky, "New allele frequency database: http://www.allelefrequencies.net," *Tissue Antigens*, vol. 61, no. 5, pp. 403–407, 2003.

[74] W. Fleri, S. Paul, S. K. Dhanda et al., "The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design," *Frontiers in Immunology*, vol. 8, 2017.

[75] C. B. Palatnik-de-Sousa, I. d. S. Soares, and D. S. Rosa, "Editorial: Epitope Discovery and Synthetic Vaccine Design," *Frontiers in Immunology*, vol. 9, 2018.

[76] D. Redmond, A. Poran, and O. Elemento, "Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq," *Genome Medicine*, vol. 8, no. 1, p. 80, 2016.

[77] A. A. Eltahla, S. Rizzetto, M. R. Pirozyan et al., "Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells," *Immunology and Cell Biology*, vol. 94, no. 6, pp. 604–611, 2016.

[78] S. Afik, K. B. Yates, K. Bi et al., "Targeted reconstruction of T cell receptor sequence from single cell RNA-seq links CDR3 length to T cell differentiation state," *Nucleic Acids Research*, vol. 45, no. 16, article e148, 2017.

[79] Y. F. Fuchs, V. Sharma, A. Eugster et al., "Gene Expression-Based Identification of Antigen-Responsive CD8+ T Cells on a Single-Cell Level," *Frontiers in Immunology*, vol. 10, 2019.

[80] A. A. Upadhyay, R. C. Kauffman, A. N. Wolabaugh et al., "BALDR: a computational pipeline for paired heavy and light chain immunoglobulin reconstruction in single-cell RNA-seq data," *Genome Medicine*, vol. 10, no. 1, p. 20, 2018.

[81] K. M. Cirelli, D. G. Carnathan, B. Nogal et al., "Slow Delivery Immunization Enhances HIV Neutralizing Antibody and Germinal Center Responses via Modulation of Immunodominance," *Cell*, vol. 177, no. 5, pp. 1153–1171.e28, 2019.

[82] M. J. T. Stubbington, T. Lönnberg, V. Proserpio et al., "T cell fate and clonality inference from single-cell transcriptomes," *Nature Methods*, vol. 13, no. 4, pp. 329–332, 2016.

[83] S. Rizzetto, A. A. Eltahla, P. Lin et al., "Impact of sequencing depth and read length on single cell RNA sequencing data of T cells," *Scientific Reports*, vol. 7, no. 1, p. 12781, 2017.

[84] E. Azizi, A. J. Carr, G. Plitas et al., "Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment," *Cell*, vol. 174, no. 5, pp. 1293–1308.e36, 2018.

[85] M. Singh, G. al-Eryani, S. Carswell et al., "High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes," *Nature Communications*, vol. 10, no. 1, p. 3120, 2019.

[86] Y. Tian, A. Grifoni, A. Sette, and D. Weiskopf, "Human T Cell Response to Dengue Virus Infection," *Frontiers in Immunology*, vol. 10, 2019.

[87] A. T. Waickman, K. Victor, T. Li et al., "Dissecting the heterogeneity of DENV vaccine-elicited cellular immunity using single-cell RNA sequencing and metabolic profiling," *Nature Communications*, vol. 10, no. 1, 2019.

[88] C. Trapnell, D. Cacchiarelli, J. Grimsby et al., "The dynamics and regulators of cell fate decisions are revealed by pseudo-temporal ordering of single cells," *Nature Biotechnology*, vol. 32, no. 4, pp. 381–386, 2014.

[89] K. Street, D. Risso, R. B. Fletcher et al., "Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics," *BMC Genomics*, vol. 19, no. 1, 2018.

[90] W. Saelens, R. Cannoodt, H. Todorov, and Y. Saeys, "A comparison of single-cell trajectory inference methods," *Nature Biotechnology*, vol. 37, no. 5, pp. 547–554, 2019.

[91] G. L. Manno, R. Soldatov, A. Zeisel et al., "RNA velocity of single cells," *Nature*, vol. 560, no. 7719, pp. 494–498, 2018.

[92] K. E. Neu, J. J. Guthmiller, M. Huang et al., "Spec-seq unveils transcriptional subpopulations of antibody-secreting cells following influenza vaccination," *The Journal of Clinical Investigation*, vol. 129, no. 1, pp. 93–105, 2019.

[93] S. Picelli, O. R. Faridani, A. K. Björklund, G. Winberg, S. Sagasser, and R. Sandberg, "Full-length RNA-seq from single cells using Smart-seq2," *Nature Protocols*, vol. 9, no. 1, pp. 171–181, 2014.

[94] K. Smith, L. Garman, J. Wrammert et al., "Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen," *Nature Protocols*, vol. 4, no. 3, pp. 372–384, 2009.

[95] I. Y. Ho, J. J. Bunker, S. A. Erickson et al., "Refined protocol for generating monoclonal antibodies from single human

and murine B cells," *Journal of Immunological Methods*, vol. 438, pp. 67–70, 2016.

[96] L. D. Goldstein, Y.-J. J. Chen, J. Wu et al., "Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies," *Communications Biology*, vol. 2, no. 1, 2019.

[97] Y. Cao, B. Su, X. Guo et al., "Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High- Throughput Single-Cell Sequencing of Convalescent Patients' B Cells," *Cell*, vol. 182, no. 1, pp. 73–84.e16, 2020.

[98] M. Bosaeed and D. Kumar, "Seasonal influenza vaccine in immunocompromised persons," *Human Vaccines & Immunotherapeutics*, vol. 14, no. 6, pp. 1311–1322, 2018.

[99] M. K. Andrew, S. K. Bowles, G. Pawelec et al., "Influenza Vaccination in Older Adults: Recent Innovations and Practical Applications," *Drugs & Aging*, vol. 36, no. 1, pp. 29–37, 2019.

[100] S. B. Drysdale, R. S. Barr, C. S. Rollier, C. A. Green, A. J. Pollard, and C. J. Sande, "Priorities for developing respiratory syncytial virus vaccines in different target populations," *Science Translational Medicine*, vol. 12, no. 535, article eaax2466, 2020.

[101] S. N. Crooke, I. G. Ovsyannikova, G. A. Poland, and R. B. Kennedy, "Immunosenescence and human vaccine immune responses," *Immunity & Ageing*, vol. 16, no. 1, 2019.

[102] J. E. McElhaney, C. Ewen, X. Zhou et al., "Granzyme B: Correlates with protection and enhanced CTL response to influenza vaccination in older adults," *Vaccine*, vol. 27, no. 18, pp. 2418–2425, 2009.

[103] Z. Shahid, A. Kleppinger, B. Gentleman, A. R. Falsey, and J. E. McElhaney, "Clinical and immunologic predictors of influenza illness among vaccinated older adults," *Vaccine*, vol. 28, no. 38, pp. 6145–6151, 2010.

[104] E. E. Walsh and A. R. Falsey, "Age related differences in humoral immune response to respiratory syncytial virus infection in adults," *Journal of Medical Virology*, vol. 73, no. 2, pp. 295–299, 2004.

[105] A. Cherukuri, K. Patton, R. A. Gasser Jr. et al., "Adults 65 years old and older have reduced numbers of functional memory T cells to respiratory syncytial virus fusion protein," *Clinical and Vaccine Immunology*, vol. 20, no. 2, pp. 239–247, 2013.

[106] M. Weltevrede, R. Eilers, H. E. de Melker, and D. van Baarle, "Cytomegalovirus persistence and T-cell immunosenescence in people aged fifty and older: A systematic review," *Experimental Gerontology*, vol. 77, pp. 87–95, 2016.

[107] G. Bowyer, H. Sharpe, N. Venkatraman et al., "Reduced Ebola vaccine responses in CMV+ young adults is associated with expansion of CD57+KLRG1+ T cells," *Journal of Experimental Medicine*, vol. 217, no. 7, 2020.

[108] C. M. Nielsen, M. J. White, C. Bottomley et al., "Impaired NK Cell Responses to Pertussis and H1N1 Influenza Vaccine Antigens in Human Cytomegalovirus-Infected Individuals," *Journal of Immunology*, vol. 194, no. 10, pp. 4657–4667, 2015.

[109] Z. Hu, P. A. Ott, and C. J. Wu, "Towards personalized, tumour-specific, therapeutic vaccines for cancer," *Nature Reviews Immunology*, vol. 18, no. 3, pp. 168–182, 2018.

[110] A. A. Petti, S. R. Williams, C. A. Miller et al., "A general approach for detecting expressed mutations in AML cells using single cell RNA-sequencing," *Nature Communications*, vol. 10, no. 1, 2019.

[111] U. Sahin, E. Derhovanessian, M. Miller et al., "Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer," *Nature*, vol. 547, no. 7662, pp. 222–226, 2017.

[112] P. A. Ott, Z. Hu, D. B. Keskin et al., "An immunogenic personal neoantigen vaccine for patients with melanoma," *Nature*, vol. 547, no. 7662, pp. 217–221, 2017.

[113] E. M. Lancaster, D. Jablons, and J. R. Kratz, "Applications of Next-Generation Sequencing in Neoantigen Prediction and Cancer Vaccine Development," *Genetic Testing and Molecular Biomarkers*, vol. 24, no. 2, pp. 59–66, 2020.

[114] A. A. Cohen, N. Geva-Zatorsky, E. Eden et al., "Dynamic proteomics of individual cancer cells in response to a drug," *Science*, vol. 322, no. 5907, pp. 1511–1516, 2008.

[115] K.-T. Kim, H. W. Lee, H.-O. Lee et al., "Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma," *Genome Biology*, vol. 17, no. 1, 2016.

[116] A. R. Wu, N. F. Neff, T. Kalisky et al., "Quantitative assessment of single-cell RNA-sequencing methods," *Nature Methods*, vol. 11, no. 1, pp. 41–46, 2014.

[117] J. M. Alves and D. Posada, "Sensitivity to sequencing depth in single-cell cancer genomics," *Genome Medicine*, vol. 10, no. 1, 2018.

[118] R. Lu, X. Zhao, J. Li et al., "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding," *Lancet*, vol. 395, no. 10224, pp. 565–574, 2020.

[119] N. Zhu, D. Zhang, W. Wang et al., "A Novel Coronavirus from Patients with Pneumonia in China, 2019," *The New England Journal of Medicine*, vol. 382, no. 8, pp. 727–733, 2020.

[120] W. Sungnak, H. C. A. L. B. Network, N. Huang et al., "SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes," *Nature Medicine*, vol. 26, no. 5, pp. 681–687, 2020.

[121] S. Lukassen, R. L. Chua, T. Trefzer et al., "SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells," *The EMBO Journal*, vol. 39, no. 10, article e105114, 2020.

[122] W. Wen, W. Su, H. Tang et al., "Immune cell profiling of COVID-19 patients in the recovery stage by single- cell sequencing," *Cell Discovery*, vol. 6, no. 1, 2020.

[123] A. J. Wilk, A. Rustagi, N. Q. Zhao et al., "A single-cell atlas of the peripheral immune response to severe COVID-19," *medRxiv*, 2020.

[124] M. Liao, Y. Liu, J. Yuan et al., "Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19," *Nature Medicine*, vol. 26, no. 6, pp. 842–844, 2020.

[125] S. J. Zost, P. Gilchuk, R. E. Chen et al., "Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein," *bioRxiv*, 2020.

[126] M. R. Mehra, F. Ruschitzka, and A. N. Patel, "Retraction—Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis," *Lancet*, vol. 395, no. 10240, p. 1820, 2020.

[127] M. R. Mehra, S. S. Desai, S. Kuy, T. D. Henry, and A. N. Patel, "Retraction: Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. N Engl J Med. DOI:

10.1056/NEJMoa2007621," *New England Journal of Medicine*, vol. 382, no. 26, p. 2582, 2020.

[128] D. Lähnemann, J. Köster, E. Szczurek et al., "Eleven grand challenges in single-cell data science," *Genome Biology*, vol. 21, no. 1, p. 31, 2020.

[129] T. Zeng and H. Dai, "Single-Cell RNA Sequencing-Based Computational Analysis to Describe Disease Heterogeneity," *Frontiers in Genetics*, vol. 10, p. 629, 2019.

[130] P. A. Szabo, H. M. Levitin, M. Miron et al., "Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease," *Nature Communications*, vol. 10, no. 1, 2019.

[131] E. Hedlund and Q. Deng, "Single-cell RNA sequencing: Technical advancements and biological applications," *Molecular Aspects of Medicine*, vol. 59, pp. 36–46, 2018.

[132] S. C. van den Brink, F. Sage, Á. Vértesy et al., "Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations," *Nature Methods*, vol. 14, no. 10, pp. 935-936, 2017.

[133] V. Svensson, R. Vento-Tormo, and S. A. Teichmann, "Exponential scaling of single-cell RNA-seq in the past decade," *Nature Protocols*, vol. 13, no. 4, pp. 599–604, 2018.

[134] B. B. Lake, S. Codeluppi, Y. C. Yung et al., "A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA," *Scientific Reports*, vol. 7, no. 1, 2017.

[135] A. M. Ranzoni, P. M. Strzelecka, and A. Cvejic, "Application of single-cell RNA sequencing methodologies in understanding haematopoiesis and immunology," *Essays in Biochemistry*, vol. 63, no. 2, pp. 217–225, 2019.

[136] Z. Xiao, G. Cheng, Y. Jiao et al., "Holo-Seq: single-cell sequencing of holo-transcriptome," *Genome Biology*, vol. 19, no. 1, p. 163, 2018.

[137] R. E. Drury, D. O'Connor, and A. J. Pollard, "The Clinical Application of MicroRNAs in Infectious Disease," *Frontiers in Immunology*, vol. 8, 2017.

[138] C.-H. L. Eng, M. Lawson, Q. Zhu et al., "Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+," *Nature*, vol. 568, no. 7751, pp. 235–239, 2019.

[139] M. Efremova, M. Vento-Tormo, S. A. Teichmann, and R. Vento-Tormo, "CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes," *Bio Rxiv*, vol. 15, no. 4, pp. 1484–1506, 2020.

[140] D. A. Skelly, G. T. Squiers, M. A. McLellan et al., "Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart," *Cell Reports*, vol. 22, no. 3, pp. 600–610, 2018.

[141] C. Vogel and E. M. Marcotte, "Insights into the regulation of protein abundance from proteomic and transcriptomic analyses," *Nature Reviews Genetics*, vol. 13, no. 4, pp. 227–232, 2012.

[142] J. W. B. Hershey, N. Sonenberg, and M. B. Mathews, "Principles of Translational Control: An Overview," *Cold Spring Harbor Perspectives in Biology*, vol. 4, no. 12, 2012.

[143] E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, and L. Cai, "Single-cell in situ RNA profiling by sequential hybridization," *Nature Methods*, vol. 11, no. 4, pp. 360-361, 2014.

[144] P. L. Ståhl, F. Salmén, S. Vickovic et al., "Visualization and analysis of gene expression in tissue sections by spatial transcriptomics," *Science*, vol. 353, no. 6294, pp. 78–82, 2016.

[145] S. G. Rodriques, R. R. Stickels, A. Goeva et al., "Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution," *Science*, vol. 363, no. 6434, pp. 1463–1467, 2019.

[146] V. Svensson, "A method for transcriptome-wide gene expression quantification in intact tissues," *Immunology and Cell Biology*, vol. 97, no. 5, pp. 439–441, 2019.

[147] P. A. Darrah, J. J. Zeppa, P. Maiello et al., "Prevention of tuberculosis in macaques after intravenous BCG immunization," *Nature*, vol. 577, no. 7788, pp. 95–102, 2020.

[148] R. Hanamsagar, T. Reizis, M. Chamberlain et al., "An optimized workflow for single-cell transcriptomics and repertoire profiling of purified lymphocytes from clinical samples," *Scientific Reports*, vol. 10, no. 1, 2020.

[149] A. Serin Harmanci, A. O. Harmanci, and X. Zhou, "CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data," *Nature Communications*, vol. 11, no. 1, 2020.

[150] H. M. Kang, M. Subramaniam, S. Targ et al., "Multiplexed droplet single-cell RNA-sequencing using natural genetic variation," *Nature Biotechnology*, vol. 36, no. 1, pp. 89–94, 2018.

[151] T. Barrett, S. E. Wilhite, P. Ledoux et al., "NCBI GEO: archive for functional genomics data sets–update," *Nucleic Acids Research*, vol. 41, no. Database issue, pp. D991–D995, 2013.

[152] C. Ziegenhain, B. Vieth, S. Parekh et al., "Comparative Analysis of Single-Cell RNA Sequencing Methods," *Molecular Cell*, vol. 65, no. 4, pp. 631–643.e4, 2017.

[153] M. J. Zhang, V. Ntranos, and D. Tse, "Determining sequencing depth in a single-cell RNA-seq experiment," *Nature Communications*, vol. 11, no. 1, p. 774, 2020.

[154] C. S. Baron, A. Barve, M. J. Muraro et al., "Cell type purification by single-cell transcriptome-trained sorting," *Cell*, vol. 179, no. 2, pp. 527–42.e19, 2019.

[155] A. J. M. Larsson, P. Johnsson, M. Hagemann-Jensen et al., "Genomic encoding of transcriptional burst kinetics," *Nature*, vol. 565, no. 7738, pp. 251–254, 2019.

[156] J. M. Vahrenkamp, K. Szczotka, M. K. Dodson, E. A. Jarboe, A. P. Soisson, and J. Gertz, "FFPEcap-seq: a method for sequencing capped RNAs in formalin-fixed paraffin-embedded samples," *Genome Research*, vol. 29, no. 11, pp. 1826–1835, 2019.

[157] C. T. Wohnhaas, G. G. Leparc, F. Fernandez-Albert et al., "DMSO cryopreservation is the method of choice to preserve cells for droplet- based single-cell RNA sequencing," *Scientific Reports*, vol. 9, no. 1, p. 10699, 2019.

[158] C. A. Vargas-Garcia, K. R. Ghusinga, and A. Singh, "Cell size control and gene expression homeostasis in single-cells," *Current Opinion in Systems Biology*, vol. 8, pp. 109–116, 2018.

[159] A. Schroeder, O. Mueller, S. Stocker et al., "The RIN: an RNA integrity number for assigning integrity values to RNA measurements," *BMC Molecular Biology*, vol. 7, no. 1, p. 3, 2006.

[160] M. Stoeckius, et al.S. Zheng, B. Houck-Loomis et al., "Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics," *Genome Biology*, vol. 19, no. 1, p. 224, 2018.

[161] H. T. N. Tran, K. S. Ang, M. Chevrier et al., "A benchmark of batch-effect correction methods for single-cell RNA sequencing data," *Genome Biology*, vol. 21, no. 1, p. 12, 2020.

[162] L. Haghverdi, A. T. L. Lun, M. D. Morgan, and J. C. Marioni, "Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors," *Nature Biotechnology*, vol. 36, no. 5, pp. 421–427, 2018.

[163] P. Datlinger, A. F. Rendeiro, T. Boenke, T. Krausgruber, D. Barreca, and C. Bock, "Ultra-high throughput single-cell RNA sequencing by combinatorial fluidic indexing. Genomics," *bio Rxiv*, p. 90, 2019.

[164] S. Hen-Avivi and R. Avraham, "Immune cell type 'fingerprints' at the basis of outcome diversity of human infection," *Current Opinion in Microbiology*, vol. 42, pp. 31–39, 2018.

[165] J. Ding, X. Adiconis, S. K. Simmons et al., "Systematic comparative analysis of single cell RNA-sequencing methods. Genomics," *bio Rxiv*, p. 439, 2019.

[166] V. Svensson, K. N. Natarajan, L.-H. Ly et al., "Power analysis of single-cell RNA-sequencing experiments," *Nature Methods*, vol. 14, no. 4, pp. 381–387, 2017.