Molecular Immunology 103 (2018) 191–199

Contents lists available at ScienceDirect

Molecular Immunology

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

Using single cell analysis for translational studies in immune mediated diseases: Opportunities and challenges

Siddhartha Sharma[a], Louis Gioia[b], Brian Abe[a], Marie Holt[a], Anne Costanzo[a], Lisa Kain[a], Andrew Su[b], Luc Teyton[a,*]

[a] Department of Immunology and Microbiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA, 92037, USA
[b] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA

ARTICLE INFO

Keywords:
Single cell analysis
Translational studies
Immune mediated diseases
Ulcerative colitis
MAIT cells

ABSTRACT

The difficulty of studying small tissue samples and rare cell populations have been some of the main limitations in performing efficient translational studies of immune mediated diseases. Many of these conditions are grouped under the name of a single disease whilst there are strong suggestions that disease heterogeneity leads to variable disease progression as well as therapeutic responses. The recent development of single cell techniques, such as single cell RNA sequencing, gene expression profiling, or multiparametric cytometry, is likely to be a turning point. Single cell approaches provide researchers the opportunity to finally dissect disease pathology at a level that will allow mechanistic classifications and precision therapeutic strategies. In this review, we will give an overview of the current and developing repertoire of single cell techniques, the benefits and limitations of each, and provide an example of how single cell techniques can be utilized to understand complex immune mediated diseases and their translation from mouse to human.

1. Introduction

One of the major hurdles in studying the immune status of human diseases is the access to informative samples. Only two routes are available, biopsies for solid organs, and/or blood draw, now also called "liquid biopsy". However, both modes of sampling have inherent limitations: is the biopsy from an affected area? Is the biopsy representative of the entire organ? Will there be affected and unaffected tissue in the same sample? What control should be used? How many circulating immune cells are coming from the diseased organ? How often can the tissue and/or blood be sampled without affecting the patient?

Additionally, a consistent challenge is the low number of immune cells recovered from each sample. Up until now, most of the available and established techniques in immunology relied on bulk, population analysis that required a large number of cells defined by a limited set of markers. In very practical terms, biopsies are usually examined by immunohistochemistry, whereas peripheral blood mononuclear cells (PBMCs) are enumerated and phenotyped by flow cytometry. While immunohistochemistry investigates anatomical features, its resolution is low. Flow cytometry provides single cell resolution but is limited by the small set of phenotypic markers that can be used; this approach hinders the analysis of low frequency populations, and is ultimately only as good as the quality of the reagents used for staining (Chattopadhyay et al., 2014). In addition, these "bulk techniques" average out the signal over multiple cells, potentially obscuring rare disease-specific cells (Chattopadhyay et al., 2014). While bulk genomic techniques face the same issues, they are additionally limited in their interrogation of lymphocyte specificity as defined by T cell and B cell receptors, both of which rely on the co-expression of two chains, heavy and light for B cells, α and β for T cells, as it loses the information that pairing provides.

Antigen specificity of T and B cells is one of the most informative aspects of studying the immune system in cancer and autoimmunity as it directly links a cell to its function. Most, if not all, functionally informative gene expression observed in activated lymphocytes will be downstream of idiotypic receptor engagement. To add further complexity, heterogeneity has been observed in the gene and protein expression of cells within these populations. For resting cells, the steady state analysis demonstrates variability in single cell RNA expression that reflects stochastic gene expression, or "allele intrinsic" variability, as well as "allele-extrinsic" variability (Raj et al., 2006; Wagner et al., 2016). This variability is often significant because beyond differentiating two cells of the same type and same specificity within the

* Corresponding author.
E-mail address: lteyton@scripps.edu (L. Teyton).

https://doi.org/10.1016/j.molimm.2018.09.020
Received 29 May 2018; Received in revised form 11 September 2018; Accepted 27 September 2018

same tissue, it may influence their functions in response to a pathogen (Haque et al., 2017). Finally, it has been shown that in humans, each patient with an autoimmune disease can exhibit progression of disease and clinical features that are unique to that individual (Coppieters et al., 2012; Roep et al., 2012; van der Helm-van Mil et al., 2005).

In this context, single cell analysis permits the interrogation of samples of small size (biopsies) and the dissection of complex mixtures of cells found in blood and tissues. The first high throughput single cell technique to be developed was flow cytometry and while it provides single cell resolution, it is limited by the small number of parameters that can be simultaneously measured. The development of flexible and cheap microfluidic systems a decade ago was a breakthrough for the single cell field. Microfluidics provided access to a single cell's transcriptome in a high throughput format and allowed the field to expand on the pioneering work of *Eberwine et al.* in 1992. In that particular study, the authors demonstrated that morphologically similar cells have distinct patterns of gene expression and that some cells had expression of several mRNAs that were not found at the population level (Eberwine et al., 1992; Grun and van Oudenaarden, 2015; Svensson et al., 2018).

Nearly two decades later, and via intermediate steps such as single cell qPCR, the first single cell RNA sequencing paper and protocol were published by Tang et al. in (2009,2010) (Tang et al., 2010, 2009). Within three years, Nature Methods declared single cell RNA sequencing (scRNAseq) the method of the year (Editorial, 2014). As proof of how far single cell RNA sequencing has come, in 2017, single cell RNA sequencing is being discussed as a tool to bridge personalized medicine with cancer diagnostics (Shalek and Benson, 2017).

However, single cell techniques still have important technical issues that require resolving including the validation of gene expression analysis at the protein level and the incorporation of spatial distribution of heterogeneous cell populations in tissues and lesions in single cell analysis.

In this review, we will briefly discuss in a non-exhaustive way the main available techniques in the single cell field, highlight the strength and weaknesses of some techniques, and discuss an approach we have taken that combines various single cell techniques to examine very small populations of cells in human biopsies and peripheral blood.

2. Important single cell technologies

While most single cell experiments isolate individual cells via flow cytometry, a single cell proteomic technique, the term "single cell analysis" most often refers to quantification of RNA and sequencing of DNA. To utilize the limited quantities of material extractable from individual cells, most single cell transcriptomic and genomic techniques rely on the ability of the polymerase chain reaction (PCR) to amplify a single, or a few molecules of DNA. Therefore, all approaches will be necessarily limited by the quality of primer pairs, the variable efficiencies of some primer pairs in multiplex reactions, the necessity to perform a reverse transcription (RT) step to examine RNA expression, and the fidelity of the RT and DNA polymerase enzymes. The addition of linkers for sequencing, bar codes for identification, and molecular identifiers for normalization can compound these technical limitations.

Currently, single cell genomics techniques start by the isolation of single cells in reaction chambers, or reaction droplet, using microfluidics instruments, or in open wells in multi-well plates using fluorescence-activated cell sorting (FACS). While the latter is reasonably efficient (> 95%), the former approach is still limited in its ability to isolate single cells and not doublets; success rates vary from 60 to 90%, depending on the instrument (Holt et al., 2018).

2.1. Single-cell analysis beyond single cell RNA sequencing

The most common single cell genomics technique is single cell RNA sequencing (scRNAseq) that examines and quantifies the transcriptional landscape of a single cell (Linnarsson and Teichmann, 2016; Tang et al., 2009). Other single cell technologies examine the genome, the epigenome, the transcriptome or precisely quantify the expression of a particular set of genes. We will briefly discuss some of these techniques and direct the reader to reviews that cover each topic in more detail.

Single cell DNA sequencing studies genomic DNA and holds the potential of tracking somatic mutations, substitutions, insertions/deletions, copy number variants, and structural rearrangements (Grun and van Oudenaarden, 2015). One particularly promising use for single cell DNA genomics is to be able to sequence the entire genome of individual cancer cells to ascertain the copy number variants and/or the single nucleotide variants, gene translocation, and the rate of mutation in driver and non-driver genes within each tumor cell (Gawad et al., 2016). The determination of this tumor landscape and the intratumoral diversity, has the potential to tailor treatments that best target each group of cancerous cells (Gawad et al., 2016; Shalek and Benson, 2017). Additionally, this same approach can be used to diagnose cancer via the detection of circulating tumor cells (Gawad et al., 2016). However, for "liquid biopsy" diagnostic technique to be fully successful, the issue of whole genome amplification fidelity must be overcome. Currently, challenges of whole gene amplification include the loss of genomic coverage that can restrict de novo assembly of a cell genome, allelic dropout or imbalance, and errors during genomic amplification (Gawad et al., 2016).

Single cell epigenomics which examines epigenetic changes, has the theoretical capacity to examine DNA methylation, histone modifications, as well as changes in conformation and compaction of chromatin. As many of these techniques rely on antibodies, their sensitivity is highly linked to the quality of the antibodies used (Cheung et al., 2018). As a consequence, the number of single cell epigenetic studies are still limited but increasing. The first single cell epigenomic technique published examined DNA methylation in single mouse embryonic stem cells (mESCs), mouse sperm and mouse oocytes via reduced representation bisulfite sequencing (Guo et al., 2013). This technique queried DNA methylation at different sites, such as CpG islands, without averaging and compared the results to pooled mESCs of different cell numbers (Guo et al., 2013). A recently described technique, called epigenetic landscape profiling using cytometry by time of flight (EpiTOF), measures 8 classes of histone marks and 4 histone variants in immune cell subsets (Cheung et al., 2018) to distinguish the main cell types and lineages by assigning patterns of histone marks (Cheung et al., 2018). As a result of the technical limitations mentioned above, the major challenge of this technology is to improve the quality of the antibodies needed and to increase the low number of reads obtained by these methods that are otherwise robust techniques (Hyun et al., 2015).

Another recently published study paired assay for transposase-accessible chromatin with sequencing (ATAC-seq) and TCR sequencing at the single cell level. This study identified epigenomic signatures that were unique to clonal cancerous T cells (Satpathy et al., 2018). Applied in conjunction with single cell RNA analysis, single cell epigenomics has the potential to directly correlate epigenetic modifications to changes in gene expression (Wagner et al., 2016). Within the past few years, protocols have been developed to do exactly this including a protocol called single cell genome-wide methylome and transcription sequencing (scM&T-seq) which interrogates DNA methylation and transcriptome analysis (Angermueller et al., 2016). Additionally, another protocol has been developed by the same group that integrates chromatin accessibility with DNA methylation and transcriptome analysis in a protocol called single cell nucleosome, methylation and transcription sequencing (scNMT-seq) (Clark et al., 2018). However, it should also be noted that single cell epigenomic techniques are still primarily performed by specific labs and are very challenging. As a testimony to this assertion, no manufacturer has yet released a kit and/or protocols to perform any of these experiments.

From a technical standpoint, ***single cell transcriptomics*** are by far the most developed single cell techniques to qualitatively and quantitatively measure gene expression, and they can be carried out by most