

## REVIEW OPEN



## Redefining intestinal immunity with single-cell transcriptomics

Kylie Renee James<sup>1,2</sup>, Rasa Elmentaite<sup>3</sup>, Sarah Amalia Teichmann<sup>3,4</sup> and Georgina Louise Hold<sup>5</sup>

© The Author(s) 2021, corrected publication 2022

The intestinal immune system represents the largest collection of immune cells in the body and is continually exposed to antigens from food and the microbiota. Here we discuss the contribution of single-cell transcriptomics in shaping our understanding of this complex system. We consider the impact on resolving early intestine development, engagement with the neighbouring microbiota, diversity of intestinal immune cells, compartmentalisation within the intestines and interactions with non-immune cells. Finally, we offer a perspective on open questions about gut immunity that evolving single-cell technologies are well placed to address.

*Mucosal Immunology* (2022) 15:531–541; <https://doi.org/10.1038/s41385-021-00470-y>

## INTRODUCTION

The intestinal tract contains a plethora of immune cells that are essential for normal physiology and defending the body against potential pathogens, but may also contribute to disease when their responses are exacerbated. Since the recognition of a localised intestinal immune system in 1919<sup>1</sup>, evolving technologies and experimental systems have helped refine our understanding of this complex cellular network.

The invention of single-cell RNA sequencing (scRNAseq) in 2009<sup>2</sup> has revolutionized the field of immunology, revealing an unappreciated complexity of immune cell subsets, identifying new cell types and states, redefining cellular ontogeny and enabling inference of cell fate trajectories and function<sup>3,4</sup>. ScRNAseq is able to piece together existing knowledge of cell markers, ontology and interactions into an integrative picture of the building blocks of human tissues. Applied to human mucosal immunity, scRNAseq is particularly powerful as it allows for systematic analysis of cells within these complex and highly-immunologically active tissues, thereby making the most of small and often difficult to obtain clinical samples. Although transcriptional expression is not a perfect readout of protein expression<sup>5</sup>, scRNAseq allows for the hypothesis-generating phase of research to begin with and be guided by tissue-specific clues. Targeted experiments in model systems can then be used to support findings and test biological mechanisms. In this way and spurred on by the conception of the Human Cell Atlas (HCA) initiative in 2016, scRNAseq has been applied with great effect to several human barrier tissues including skin<sup>6</sup>, reproductive organs<sup>7,8</sup> and mouth<sup>9</sup>, and recently in the context of SARS-CoV-2 infection<sup>10–14</sup>.

In this review, we focus on the immune system of the intestinal tract and specifically discuss how single-cell transcriptomics has advanced knowledge in this field. We provide an introduction to scRNAseq methods and analysis tools with particular use in this area and highlight studies that have shed light on the origins of intestinal immunity, cell diversity and plasticity, interactions with

non-immune cells and compartmentalisation within the tissue architecture.

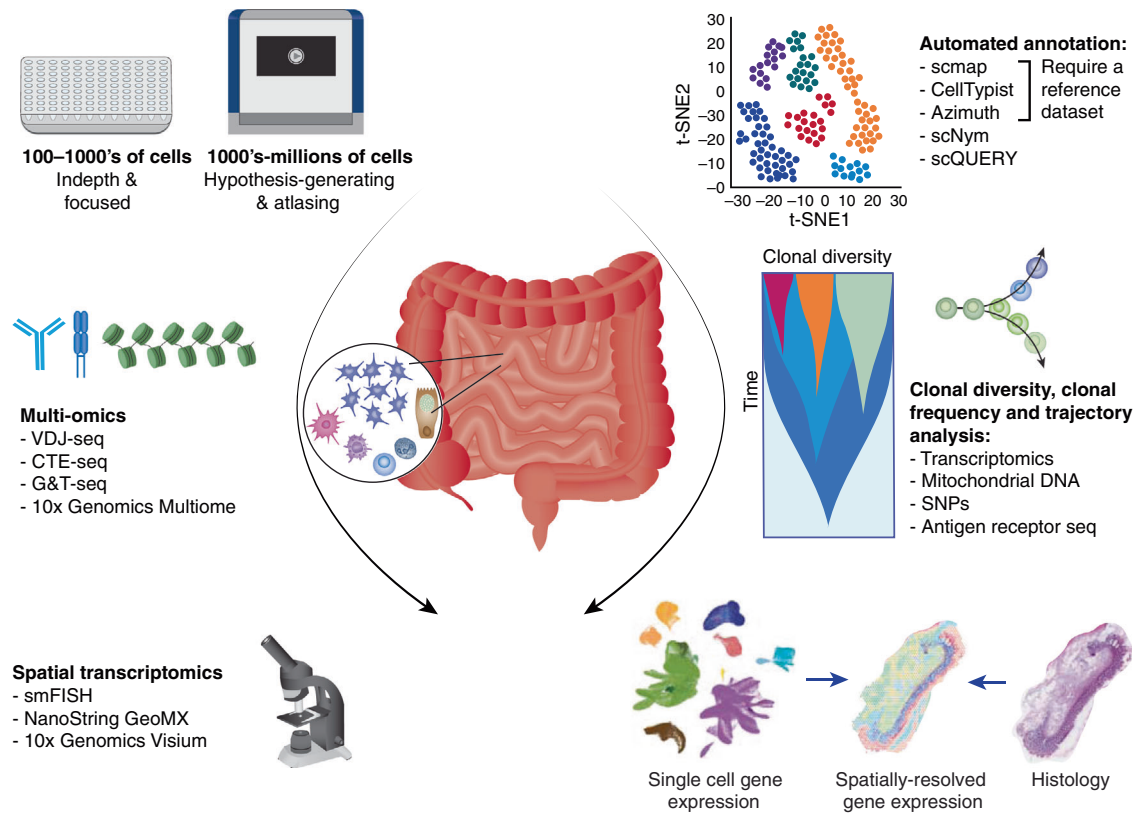
## SCRNASEQ APPROACHES TO STUDYING INTESTINAL IMMUNITY

The scRNAseq field is rapidly evolving, with the number of cells captured per experiment now in the millions. Approaches to single-cell profiling intestinal tissues vary between studies and depend on tissue availability and biological questions being asked. Current studies on intestinal immunity compare cells of healthy or IBD patients<sup>15–18</sup>, focus on regional differences<sup>19–21</sup> or investigate intestinal development<sup>22–24</sup> applying either in-depth or high-throughput methods, and increasingly combining other technologies such as V(D)J sequencing and spatial transcriptomics to better understand cell heterogeneity, lineage relationships and spatial locations in tissue<sup>20,21,24</sup>. Below we outline the current and emerging technologies and analysis tools for studying intestinal immunity.

## ScRNAseq platforms

There is a range of scRNAseq methods available with different benefits for studying mucosal immunology<sup>25–28</sup>. Platforms relying on the isolation of dissociated cells by fluorescence-activated cell sorting (FACS) or microfluidic devices include STRT-seq<sup>29</sup>, CEL2-seq<sup>30</sup>, MARS-seq<sup>31</sup> and SMART-seq2-3<sup>32–34</sup>. Use of FACS provides auxiliary information of proteins targeted by a select panel of fluorophore-tagged antibodies and can help unite transcriptional profiles to traditional cell type identities. These methods are lower in throughput due to limited capture sites, ranging from 100–1000's of cells per experiment (Fig. 1). A benefit of SMART-seq methods in particular is that they provide sequencing of full-length transcripts such that highly variable transcripts including B and T cell receptors (BCR and TCR respectively) are automatically included, and they generally have greater coverage of the

<sup>1</sup>Garvan-Weizmann Centre for Cellular Genomics, Garvan Institute of Medical Research, Sydney, NSW 2010, Australia. <sup>2</sup>School of Medical Sciences, University of New South Wales, Sydney, NSW 2006, Australia. <sup>3</sup>Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK. <sup>4</sup>Theory of Condensed Matter Group, Cavendish Laboratory/Department of Physics, University of Cambridge, Cambridge, NSW CB3 0HE, UK. <sup>5</sup>University of New South Wales Microbiome Research Centre, Sydney, NSW 2217, Australia. ✉email: k.james@garvan.org.au



**Fig. 1** ScRNAseq approaches and analysis tools for studying intestinal immunity. Left: ScRNAseq, multi-omics and spatial platforms; Right: mirrored analysis tools for analysing resulting data. sm-FISH: single molecule fluorescence in situ hybridization; SNPs: single nucleotide polymorphisms.

transcriptome compared to high-throughput approaches detailed below. Together, these approaches are especially useful for in-depth and targeted analysis of immune cell types.

High-throughput approaches rely on capture through microfluidic devices of single cells in water droplets in an oil phase (10x Genomics Chromium Gene Expression, Drop-Seq<sup>35</sup> and inDrops<sup>36</sup>) or in microwells (Seq-Well S<sup>37</sup> and STRT-seq-2i<sup>38</sup>). These methods tag either the 3' or 5' end of mRNA, incorporating a unique molecular identifier and applying a cell-specific barcode early after cell capture. Total mRNA can then be pooled for downstream library preparation allowing for processing of 1000's-millions of cells per experiment (Fig. 1). A major drawback of tagging either end of the mRNA is that highly variable transcripts such as splice variants and antigen receptors are not reliably captured. However, targeted amplification of TCRs and BCRs can be included as an additional step for the 10x Genomics Chromium 5' platform. The high-throughput, less targeted nature of these methods makes them ideal for tissue atlasing or hypothesis-generating experiments.

Recent advances in the realm of multi-omics technology, in which multiple cell features are simultaneously measured, are building on scRNAseq methods to also shed light on diversity in cell genotypes, transcriptional regulation and protein expression. These approaches include genotyping plus transcriptomics available as G&T-seq<sup>39</sup>, chromatin-accessibility with transcriptomics available via the 10x Genomics Multiome ATAC + Gene Expression platform and targeted protein quantification plus transcriptomics available as CITE-seq<sup>40</sup> or REAP-seq<sup>41</sup>. Spatial transcriptomics is another rapidly evolving area and has already provided spatial context for cell identities or cell-cell interactions identified from scRNAseq studies of the gut mucosa<sup>24,42</sup>. Available platforms include 10x Genomics Visium and nanoString GeoMx Digital Spatial Profiler, with the former currently offering whole

transcriptome capture of zones covering in the order of 10 cells, and the second providing simultaneous fluorescent imaging at single cell resolution and whole transcriptome profiling from tissue regions of interest.

#### Analysis of scRNAseq data

Pre-processing and analysis of scRNAseq data from low- and high-throughput platforms follows the same general workflow and is detailed in a number of review articles and online tutorials<sup>43,44</sup>. In short, raw sequencing data undergoes read quality control, assignment to cellular barcode, mapping to a reference genome and read quantification to obtain a cell by gene matrix, and can be done in pipelines such as Cell Ranger<sup>45</sup>, indrops<sup>46</sup>, SEQC<sup>47</sup>, or ZUMIs<sup>48</sup>. The data can then be handled with well documented computational packages as part of Seurat<sup>49</sup>, Scanpy<sup>50</sup> and OSCA<sup>51</sup> for quality control to remove empty droplets/poor quality cells, normalisation of the data and dimensionality reduction in preparation for visualisation. Downstream analysis of scRNAseq data typically involves cell clustering and cell type/state annotation, trajectory analysis<sup>52</sup> and ligand-receptor expression analysis<sup>53</sup>. Manual cell type annotation of clustered scRNAseq data is an iterative and laborious process. The most recent advances in the scRNAseq analysis include the development of automated tools for this step in the analysis pipeline. Amongst these methods, recently reviewed by ref. <sup>54</sup>, are correlation based methods that require a reference dataset (e.g., scmap, CellTypist<sup>55</sup> and Azimuth<sup>49</sup>) and neural network based algorithms without a prior reference (e.g., scNym<sup>56</sup> and scQUERY<sup>57</sup>) (Fig. 1). Reference-based methods are gaining popularity, but their output relies greatly on the relevance and quality of the cell type reference. For example, CellTypist<sup>55</sup> provides a collection of comprehensive and carefully curated immune cell profiles from multiple organs suited for the annotation of human tissue immune cells. Trajectory analysis for

tracing cell type development pathways or fate decisions is commonly carried out with packages including Monocle<sup>58</sup>, Wanderlust<sup>59</sup> and Slingshot<sup>60</sup>. Each of these methods order individual cells based on gene expression along a pseudotime trajectory describing a transient continuous biological process. Alternative or complementary approaches to lineage tracing implement paired antigen receptor sequences, single nucleotide polymorphisms<sup>61</sup> or variance in mitochondrial DNA<sup>62</sup> as natural barcodes to track differentiation or migration of clonally related cells. RNA velocity analyses such as scVelo<sup>63</sup> leverage splice variant information held within scRNAseq data to also map cellular response and developmental kinetics. Finally, recent tools to map cell signatures determined from scRNAseq data into spatially resolved transcriptomics data include cell2location (preprint available<sup>42</sup>) and within the Seurat framework.

## DEVELOPMENT OF INTESTINAL IMMUNITY

### Formation of gut-associated lymphoid structures

Intestinal immunity is established during development in utero. Prior to the single-cell genomics era, survey of immune populations during human development has been challenging, owing to tissue access, and there has been limited knowledge about the populations and markers expressed at this stage. Single-cell transcriptomics have been instrumental in understanding the diversity of immune and non-immune populations in these precious human intestinal developmental tissues<sup>22–24,64</sup>.

The intestinal immune system is supported and regulated by gut-associated lymphoid tissues (GALTs), including mesenteric lymph nodes, Peyer's Patches (PPs) and cryptopatches. Despite obvious differences in gut lymphoid tissue size and location between species<sup>65</sup>, our understanding of the development of these structures has previously relied on experiments in animals<sup>66</sup>. In mouse studies, early GALT formation has been described to involve interactions between mesenchymal lymphoid tissue organising (mLTo), endothelial LTo (eLTo) and lymphoid tissue inducer (LTI; related to innate lymphoid cells (ILCs)) cells<sup>66</sup>. Interactions between these cell types are critical for recruiting and retaining immune cells at the sites of developing lymphoid structures. Through scRNAseq of human fetal gut samples, central players in secondary lymphoid organ formation have been resolved in humans, and their communication programs in initiating PP formation are defined from as early as 12 weeks post-conception (Fig. 2)<sup>21,24</sup>. In addition, multiple subsets with LTI characteristics have been identified, proposing differences between human and mouse development<sup>21</sup>. Importantly, by comparing single-cell transcriptional profiles, equivalent stromal populations are predicted to be involved in the formation of ectopic lymphoid structures during inflammatory bowel disease (IBD), suggesting reactivation of developmental programs to support intestinal inflammation<sup>21,24</sup>. Fawkner-Corbett et al. described a population of mLTo-like stromal cells (i.e., with *CCL19*, *CCL21* and *CXCL13* expression) with similarities to a subtype of stromal cells expanded in ulcerative colitis (UC)<sup>17,24</sup>. Taking advantage of recent spatial transcriptomics technology, the authors showed localisation of these cells and confirmed the likelihood of relevant cell-cell interactions in lymphoid follicle formation *in situ*<sup>24</sup>. This shows that the formation of secondary organs is not restricted to development and is required for proper maturation and response of the immune system.

### First encounters with the microbiota

We live in an era where the relationship between our immune system and microbes has never received such unprecedented attention. Characterising human-associated microbiotas and their role in health and disease has become the holy grail of current medicine. It is well-established that the human-associated microbiota contains a wide and complex community of

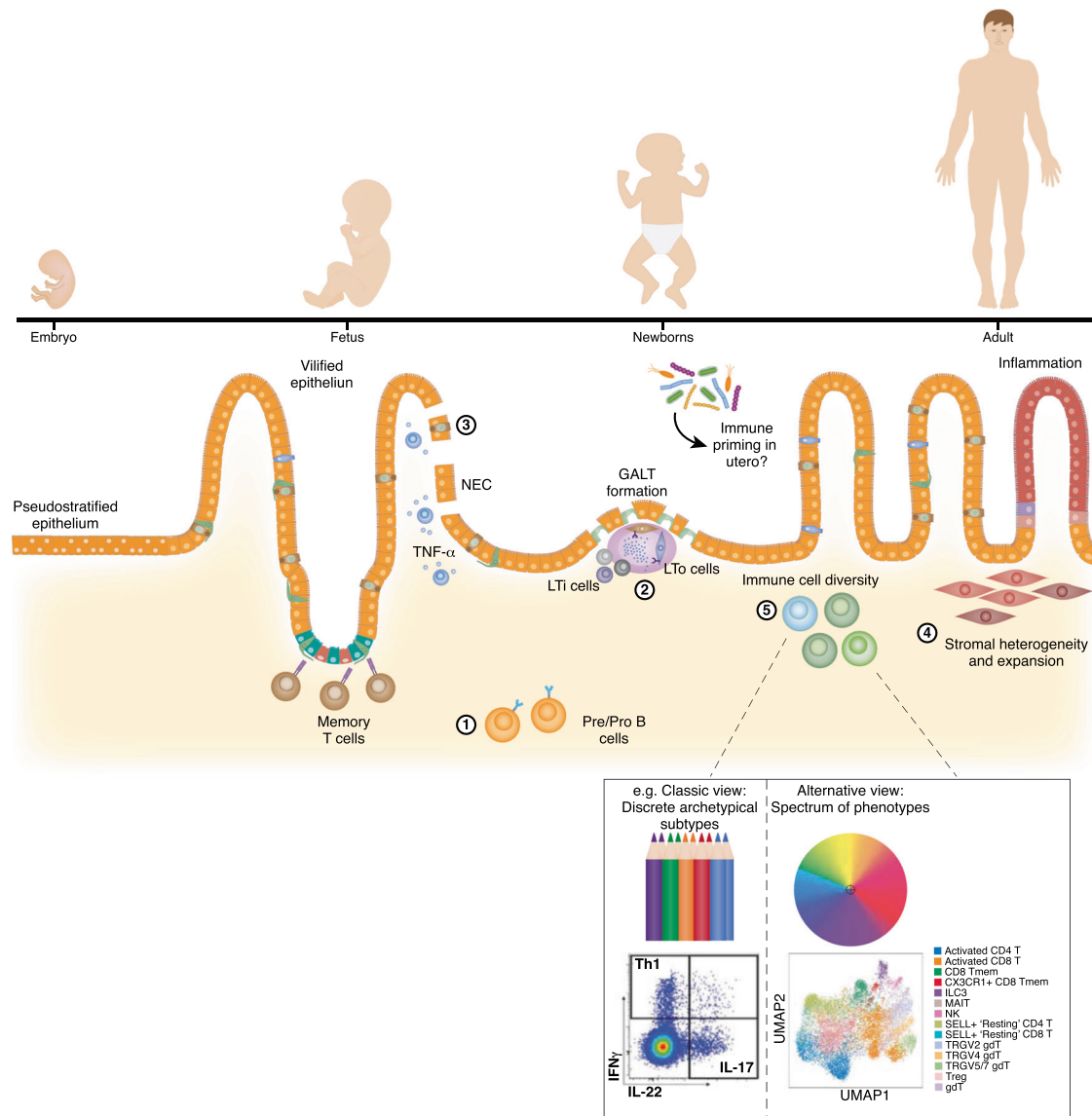
microorganisms that is unique to individuals and constantly evolves in response to its environment<sup>67,68</sup>. Microbial dysbiosis is well recognised in diseases such as IBD, colorectal cancer, metabolic disorders and in conditions including pregnancy although mechanistic insight into the host:microbial relationship remains in its infancy<sup>69</sup>. Whether the interaction is between the host and its resident microbiota or a direct response to a specific infectious entity, microbes communicate with the host through attachment to mucosal surfaces, binding specifically to host receptors, production of metabolites such as short chain fatty acids and bile acids or adapting their growth and metabolism based on changes we make to their local environment. In parallel host immune responses attempt to continuously decipher between microbial friend or foe.

The question of when host–microbe interactions become established has become a topic of intense investigation, with the presence of microbiota during in utero development still highly debated. A recent study of the meconium microbiota in human neonates (at term) before birth, controlling for process/delivery mode-induced contamination indicated that microbial colonization most likely occurs either during birth via maternal seeding or post-birth via environmental seeding<sup>70</sup>. Conversely, microscopic images of bacterial-like structures with mucin threads within the gut lumen during the second gestational trimester provide compelling evidence for in utero seeding as well as aligning with other studies that detail in utero antigenic priming of the fetal immune system<sup>71</sup>. However, these studies are caveated with the potential of contamination with environmental sources of microbes (reviewed in ref. <sup>72</sup>) making their physiological relevance questionable.

Nevertheless, priming of the immune system and unexpected activation of immune cells have been suggested to be linked to the early microbial colonisation in fetal organs, especially the gut<sup>73</sup>. In particular, multiple scRNAseq studies have shown that memory CD4+ and CD8+ T cells are present and clonally expanded in the intestines in the first and second trimester of development (Fig. 2)<sup>74–76</sup>. However, in a scRNA-seq study by Schreurs et al. fetal intestinal CD4+ T cells had a distinct gene expression profile from those in the post-natal intestine, and were characterized by high expression of genes regulating cell cycle, WNT signalling, and tissue development<sup>77</sup>. This supports the role of CD4+ T cells in fetal intestines promoting tissue development. A study using cytometry by time of flight (CyTOF) in combination with BCR sequencing showed that B cells are immature during second-trimester human development compared to those found in infants<sup>75</sup>. Our scRNAseq of human fetal intestines up to 17 weeks post conception also showed no evidence of B cell clonal expansion, class-switching or germinal centre formation<sup>21</sup>. Prenatal B cells may similarly be involved in development of lymphoid structures and have no need for class switching, while postnatal B cells undergo these events due to the presence of microbiome<sup>75</sup>. Through more precise analysis of cell phenotypes, these single cell studies promote the emerging concept that immune cell activation at least until second-trimester development is a product of their support of a highly controlled process of tissue generation rather than due to microbial seeding. Whether this is also the case in the third trimester of human development remains to be determined.

### Necrotising enterocolitis

It has been argued that the epithelial barrier in preterm infants is immature; unable to sustain the ensuing microbial colonisation due to epithelial leakiness. Necrotising enterocolitis (NEC) is a devastating intestinal disease that occurs primarily in premature infants, resulting in impairment of the epithelial barrier and in extreme cases causing intestinal perforation and tissue necrosis<sup>78</sup>. Studies have consistently highlighted differences in bacterial gut communities associated with NEC that result in an imbalance



**Fig. 2** **scRNAseq advances in understanding mucosal immunity throughout life and during disease.** (1) Building from observations of lymphocyte maturation in the prenatal intestines, single-cell studies have promoted the concept that T and B cell activation up to second-trimester development is in support of intestinal development rather than in response to microbial seeding<sup>21,24–76</sup>. (2) Key players in gut-associated lymphoid tissue (GALT) formation previously identified in developing mouse gut have been identified in scRNAseq studies of human fetal gut, with additional cell heterogeneity identified<sup>21,24</sup>. (3) TNF- $\alpha$ -producing CD4+ T cells resolved at single-cell level are enriched in the preterm intestines, likely contributing to epithelial damage observed in necrotising enterocolitis (NEC)<sup>77</sup>. (4) Intestinal stromal heterogeneity has been determined, and expanded subtypes in inflammatory bowel disease linked with pathology<sup>16–18</sup>. (5) Inlaid box depicts the classical idea of discrete immune cell subsets versus a spectrum of phenotypes determined through scRNAseq analyses<sup>88</sup>. LTI lymphoid tissue inducer, LTo lymphoid tissue organiser.

between pro- and anti-inflammatory gut immune mediators<sup>79,80</sup>. Work by Cho et al. used mouse models to highlight an imbalance within the adaptive immune system in the NEC intestinal environment typified by type 3/T helper (Th)17 polarization, with reduced Th1, Th2, and Treg responses<sup>81</sup>. These findings were further supported by scRNAseq studies showing preferential presence of TNF- $\alpha$ -producing CD4+ T cells in early intestinal development and an enrichment for these cells in the intestines of preterm infants with NEC (Fig. 2)<sup>77</sup>. The TNF- $\alpha$  overloaded microenvironment likely contributes to NEC-associated epithelial damage<sup>77</sup>. In addition to the effects of IL-10 in promoting self-renewal of stem cells, the potential of T cell cytokines IFN $\gamma$ , IL-17A and IL-13 in promoting differentiation of epithelial cells towards mature cell types has been shown in adult mice<sup>82</sup>. The capacity of T cells to inform epithelial cell differentiation and maturation

provides the opportunity to harness this interaction in clinical practice for treatment of NEC and other gut disorders.

Bacterial dysbalance in the premature intestines is considered one of the key factors contributing to NEC. No single microbe has been identified as the *mea culpa* for NEC although increased abundance of Proteobacteria are frequently reported in NEC infants<sup>83,84</sup>. A recent study analysed microbial features predictive of NEC and identified Enterobacteriaceae overgrowth; including specifically *Klebsiella*—known to possess secondary metabolite gene clusters related to quorum sensing and bacteriocin production to be replicating more rapidly in the days prior to NEC diagnosis<sup>85</sup>. The transcriptional and proportional cell changes are likely reflected in these preterm infants and future single-cell studies will be instrumental in defining these changes.

## HETEROGENEITY AND PLASTICITY OF INTESTINAL IMMUNE CELLS

The immune system must exhibit diversity and plasticity to respond to the countless challenges incurred throughout life. The conventional approach to studying diversity in immune cells has been top down—focussing on a cell type and iteratively subdividing it into more distinct subsets based on marker gene expression. This approach has been essential in understanding the intestinal immune system, but relies on pre-selection of markers and is limited in resolving heterogeneity within distinct cell groups. A strength of scRNAseq is its ability to explore heterogeneity from the bottom up—dividing cells into distinct groups and then defining molecular profiles that best describe each population<sup>4</sup>. In this way scRNAseq has refined classical immune cell type labels, defined new populations and predicted the role of cell types and states in the intestinal immune system of both mice<sup>86</sup> and humans<sup>16,87–89</sup>.

ILCs are innate immune cells that defend against both intra- and extracellular infections and are particularly abundant in mucosal tissues<sup>90</sup>. While they do not possess a functional TCR, they draw parallels in function and subtype classification with Th cells. ILCs are typically divided into 5 types- ILC1, ILC2, ILC3, natural killer and LT $\alpha$  cells<sup>91</sup>. A study by Muzzurana et al. compared sorted CD127+ ILCs from human blood, lung, colon and a past tonsil dataset<sup>92</sup> using the Smartseq2 platform<sup>87</sup>. Adopting a bottom-up approach to classifying ILC subtypes, they performed unbiased clustering followed by differential expression analysis and correlation analysis on the pooled data. ILCs subdivided into 20 subsets, clustering largely by tissue origin and FACS phenotype. Highest ILC3-associated gene expression was detected in the colon as expected<sup>90</sup>, but also the highest degree of diversity covering a spectrum of signatures ranging from migratory (expressing *SELL*, *S1PR1*, *ITGAX* and *GPR183*) to activation and tissue residence (expression of *IL22*, *NCR2*, *GRM7* and *LTA4H*)<sup>87</sup>. ILC heterogeneity has similarly been shown with scRNAseq of mouse intestines<sup>86</sup>. To show the influence of the neighbouring microbiota on ILC signatures, the authors of this study treated mice with antibiotics prior to scRNAseq analysis. In antibody-treated mice, profiles of ILC1 and ILC2 more closely resembled ILC3 cells (with increased *Atf5*, *Cxcl9* and *Gpx1*) compared to mice with an intact microbiota<sup>86</sup>. This points towards ILC3 representing the “default” phenotype with environmental factors driving diversification.

Amongst the most diverse immune cells are CD4+ T cells, which have classically been partitioned into discrete subsets according to their expression of key transcription factors and cytokines (e.g., Th1 and Th2 cells expressing IFN $\gamma$ /TBET and IL4/IL5/IL13/GATA3, respectively). However, plasticity or merging between these subsets has been a frequent observation in mice and humans<sup>93</sup>. A scRNAseq study by Kiner et al. observed extensive heterogeneity and blended signatures of colonic T cells in specific pathogen-free mice<sup>88</sup>. In an attempt to drive Th differentiation they infected mice with *Citrobacter rodentium*, an inducer of Th17 cells (determined by IL-17 expression), *Heligmosomoides polygyrus* and *Nippostrongylus brasiliensis*, both inducers of Th2 (IL-5 and IL-13) responses or *Salmonella enterica*, a bacterial infection inducing Th1 (IFN $\gamma$ ) responses. While FACS of T cells from infected mice confirmed the expected skewing of Th differentiation, scRNAseq analysis and unbiased clustering separated cells by infection system rather than characteristic Th genes. Expression of canonical Th cytokines dominated opposing sides of the same clusters in their data, arguing against discrete subsets and in favour of a polarised continuum of Th phenotypes driven by the infection setting<sup>88</sup> (Fig. 2, inlay). Other scRNAseq studies of acute immune responses in mice have reported skewed Th signatures including Th1 and Tfh in peripheral blood of mice infected with *Plasmodium*<sup>94</sup>, Th2 in lungs of mice exposed to dust mites<sup>95</sup> and Th2 in spleen and lymph nodes of mice infected with *Nippostrongylus brasiliensis*<sup>96</sup>, but have also showed heterogeneity

and blending of canonical marker genes between clusters. Spectrums of Th phenotypes have also been resolved at single-cell level within the human breast cancer tumour microenvironment<sup>47</sup>, asthmatic lung<sup>97</sup> and in blood of SARS-CoV-2 infected individuals<sup>98</sup>.

ScRNAseq studies of human intestinal disease have similarly added to our extensive understanding of the diversity in T cell phenotypes and highlighted specific enriched populations of likely significance to pathology. One such study observed a Th17-like population of CD4+ CD8+ cells expanded in UC<sup>16</sup>. Given the known association between Th17 cells and IBD<sup>99</sup>, the authors hypothesised the role of the Th17-like cells in driving inflammation, although this remains to be confirmed<sup>16</sup>. In contrast, an independent UC study showed expansion of an IL26-expressing subset of Th17-like CD8+ T cells with an immunoregulatory signature<sup>89</sup>. Trajectory and TCR sequencing analysis of the single cell profiles further characterised this population as clonally expanded and arising from tissue-resident T cells or representing a post-effector state. To understand the significance of IL26 expression by these cells during inflammation the authors compared pathology of dextran sodium sulfate (DSS)-induced acute colitis in wild-type mice to humanised *IL26*-expressing transgenic mice - the *IL26* gene does not naturally exist in rodents<sup>89</sup>. The *IL26* transgenic mice experienced less severe disease, a phenotype that could be reversed with the administration of an IL26-antibody<sup>89</sup>. This suggests a possible role for IL26 in protecting against inflammation. In the context of colorectal carcinoma, paired TCR and transcriptome sequencing identified 8 distinct populations of CD8+ T cells, with signatures ranging from naive, central and effector memory cells, recently activated effector memory/effector cells (TEMRA; with *PRF1*, *GZMB* and *GZMH* expression) to dysfunctional exhausted cells (expressing *PDCD1* and *HAVCR2*)<sup>100</sup>. Distinct TCR clonal populations and trajectory mapping supported two possible differentiation paths for T effector memory cells- either towards TEMRA or exhausted T cell states. The authors suggest that skewing differentiation towards beneficial TEMRA and away from a state of exhaustion could represent a possible avenue for therapeutic intervention<sup>100</sup>. Furthermore, this study showed tumour-specific T cell responses, with enrichment and clonal expansion of pro-inflammatory *CXCL13*+ *BHLHE40*+ TH1-like cells in tumours with microsatellite instability, but moderate enrichment for Th17 cells in those with microsatellite stability<sup>100</sup>. The enrichment of *CXCL13*+ T cells in tumours with high mutational burden was supported by a second scRNAseq study and offers a possible explanation for why this patient cohort responds better to checkpoint blockade therapy<sup>100,101</sup>. Together these studies highlight how scRNAseq can assist in understanding the complexity of intestinal diseases and the nuanced involvement of cell types and states in disease progression and control.

## ZONATION OF INTESTINAL IMMUNITY

Immune cells do not act in isolation, rather their phenotype and response is shaped by their local environment. The intestinal tract in particular comprises unique microenvironments at the macro-anatomical level in terms of distinct tissue regions and at the microanatomical level within the cross-sectional layers of the intestinal wall. Single-cell transcriptomics studies have built upon knowledge of zonation of cells within the human and mouse intestinal tract through providing the full breadth of molecular profiles of cell states, suggesting distinct roles for cells between different zones and how these contribute to the physiological functions of the intestines.

### Zonation between anatomical gut regions

With roles in segregation of luminal contents and gut tissue and mediating absorption of nutrients and transfer of signals<sup>102</sup>, the

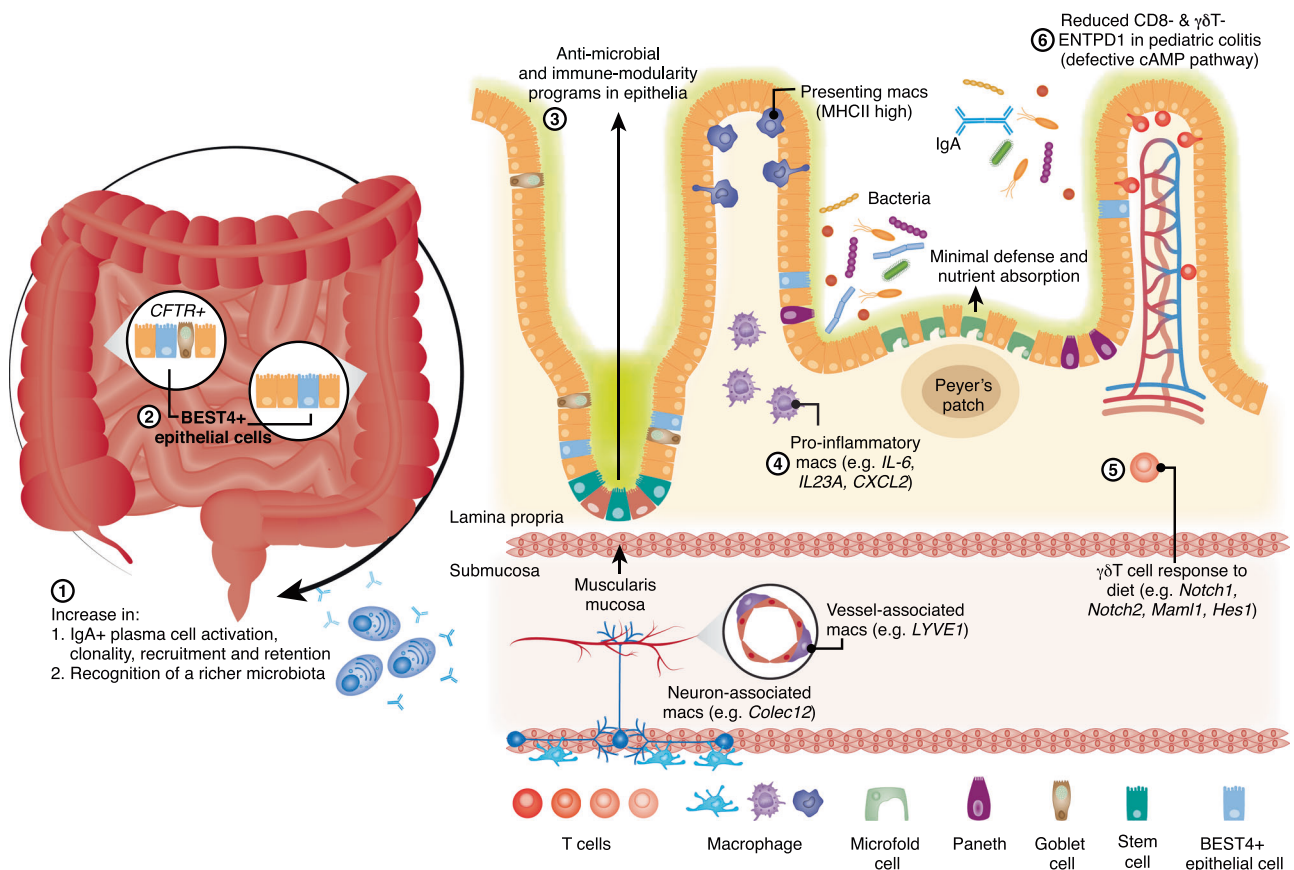
epithelial barrier cells have notable variability between small and large intestines. For example, small intestinal epithelium forms villi and crypts while large intestines form only crypts, and Paneth cells that secrete antimicrobial peptides are only present in the small intestine<sup>103</sup>, while mucus secreting goblet cells are more abundant in the large intestine where they maintain a thicker mucus layer<sup>104</sup>. Through unbiased analysis of gene expression, scRNAseq has shown variability in the expression of nutrient absorption and antimicrobial defence genes by the epithelia between small and large intestines, leading to identification of a Paneth-like cell in the latter<sup>19</sup>. ScRNAseq has also resolved further rare subtypes based on distinct gene expression and shown that these change by gut regions. BEST2+ goblet cells that are restricted to the colon<sup>105,106</sup>, have been deeply profiled at single cell level in humans. This analysis revealed their specific expression of Kallikreins *KLK15* and *KLK3*, and protease inhibitors *WFDC2* and *WFDC3* compared to other colonocytes<sup>21</sup>. Similarly BEST4+ epithelial cells, first identified in the intestinal tract by Ito et al.<sup>105</sup>, have been shown to be transcriptionally distinct from other epithelial cells in the human intestines<sup>15,16</sup>. Building on previous work in which a rare subset of small intestinal epithelial cells was reported to highly express *CFTR*, encoding a key channel mutated in cystic fibrosis<sup>107,108</sup>, further scRNAseq studies showed that BEST4+ epithelial cells of the human small, but not large, intestine co-expressed *CFTR* (Fig. 3)<sup>21,109</sup>. Based on their transcriptional profile and

co-localisation with goblet cells, it has been proposed that these cells specifically in the upper intestinal tract support mucus secretion<sup>21,109</sup> and could be implicated in intestinal symptoms experienced by many cystic fibrosis patients<sup>110</sup>.

Zonation in plasma B cells is similarly described between small and large intestines in humans, with previous studies describing a dominance of IgA1 isotype in small intestine versus IgA2 in the large intestine and an overall trend toward greater abundance of dimeric IgA plasma cells in the latter<sup>111</sup>. A recent single-cell study looked more closely at how these cells changed within the healthy human colon<sup>20</sup>. ScRNAseq of multiple colonic regions from the same individuals not only showed the increasing abundance of IgA+ plasma cells from proximal to distal colon, but transcriptional signatures suggesting this was at least in part due to increased retention/recruitment (Fig. 3). BCR repertoire analysis of the same cells indicated that distal colonic plasma cells were also more clonally expanded and somatically mutated, demonstrating the wealth of information that can be simultaneously obtained through scRNAseq approaches. Paired analysis of the neighbouring microbiota linked the increasing gradient of plasma cell response to recognition of a richer microbiota<sup>20</sup>.

### Zonation at the microanatomical level

The intestinal mucosa can be divided into three compartments-epithelium, lamina propria and muscularis mucosae<sup>112</sup>. These



**Fig. 3 Key scRNAseq finding of zonation of intestinal immunity.** (1) ScRNAseq of distinct anatomical regions of the human intestines has revealed increasing activation of plasma cells linked to recognition of a richer microbiota from proximal to distal colon<sup>20</sup> and (2) small intestine-specific expression of *CFTR* by BEST4+ epithelial cells<sup>21,109</sup>. (3) At the microanatomical level, scRNAseq has shown distinctions in gene expression of the epithelium, with cells at the villus tip expressing immune-modulatory genes<sup>114</sup> and follicle-associated cells prioritising efficient sampling of microbial antigens<sup>132</sup>. (4) Unbiased clustering of single intestinal macrophages has revealed their unique signatures within the layers of the intestinal wall indicative of specific functions and the influence of interactions with tissue-resident cells<sup>113</sup>. (5) ScRNAseq of physically separated intraepithelial and lamina propria  $\gamma\delta$  T has implicated the latter in supporting intestinal epithelium remodelling in mice due to change in diet<sup>115</sup>. (6) In the context of colitis, scRNAseq has revealed decreased abundance of CD8- and  $\gamma\delta$ -ENTPD1 cells in the intraepithelial space and implicated defective cAMP signalling in disease pathogenesis<sup>116</sup>.

layers are colonised by distinct communities of cells, with substantial interaction and movement between them. The majority of intestinal CD8<sup>+</sup> and  $\gamma\delta$  T cells exist within the intraepithelial layer, while CD4<sup>+</sup> T cells typically reside in the lamina propria. Separating these two compartments prior to scRNAseq processing has revealed further surprising details of the zonation of these cell types<sup>113,114</sup> and their adaptations<sup>115</sup> and contributions to disease<sup>116,117</sup>.

A study by Sullivan et al. of the mouse small intestinal epithelium showed up-regulation of enteric and pancreatic genes involved in digestion and absorption in response to a high-carbohydrate diet<sup>115</sup>. This gene program was defective in mice depleted of  $\gamma\delta$  T cells. Following this observation, the authors performed scRNAseq on sorted intraepithelial and lamina propria  $\gamma\delta$  T cells and identified four transcriptionally distinct populations across both compartments. Surprisingly, while cells of the intraepithelial space would have better access to the epithelium and luminal content, it was  $\gamma\delta$  T of the lamina propria with the necessary transcriptional profile (i.e., *Notch1*, *Notch2*, *Maml1* and *Hes1*) to permit communication with the epithelium and support its remodelling in response to diet (Fig. 3)<sup>115</sup>. In human coeliac disease, scRNAseq has shown reorganisation of the lamina propria lymphocytes, with natural killer cells of this compartment during health, completely absent during disease<sup>117</sup>. While results of both studies required further validation, they point to finer grain variability of immune cells between intestinal compartments.

ScRNAseq has similarly been applied to better resolve the compartmentalisation of T cells and expression of known risk factors<sup>118</sup> during Crohn's disease (CD). Th17 cells and their cytokines are known to be key mediators of the pathogenesis of CD<sup>119</sup>. ScRNAseq has further shown that Th17 cells accumulate within the intraepithelial space at the expense of CD8<sup>+</sup> T,  $\gamma\delta$  T, Tfh and T regulatory cells during active CD compared to controls<sup>120</sup>. Studies of pediatric colitis reported a decreased abundance of CD8-*ENTPD1* (expressing the gene encoding CD39) and  $\gamma\delta$ -*ENTPD1* cells in the intraepithelial compartment<sup>116</sup>. The transcriptional profile of these specific cell subsets led the authors to hypothesise that a defective cAMP pathway was at play and contributing to disease pathogenesis. To test this theory, the phosphodiesterase inhibitor and anti-platelet drug, dipyridamole, was used to drive the cAMP pathway in a mouse model of colitis and in patients, resulting in a dose-dependent increase in T cell CD39 expression and improved epithelial integrity and decreased colitis severity<sup>116</sup>.

Separate populations of macrophages exist within the lamina propria, submucosa and muscularis propria. A wealth of earlier research has described diverse roles for these populations appropriate to their microenvironment- lamina propria macrophages phagocytose bacterial antigens and produce mediators that drive epithelial cell renewal and muscularis macrophages interact closely with the enteric nervous system<sup>121</sup>. Bulk RNA sequencing (RNAseq) analysis of macrophages from these physically separated compartments showed separate expression profiles<sup>122</sup>. However, while fluorescence microscopy of mouse intestinal tissue highlighted at least two morphologically distinct populations of muscularis macrophages<sup>122</sup>, the nature and origins of further subsets of macrophages within each compartment remained a mystery. An unpublished study by Domanska et al. implemented scRNAseq of adjacent normal colorectal cancer tissues to address these questions<sup>113</sup>. They showed that lamina propria macrophages comprise 13 transcriptionally distinct subsets with a spectrum of proinflammatory signature (*IL-1B*, *IL-1A*, *IL-6*, *IL23A*, *CXCL2*, *CXCL3* and *CXCL8* or *CXCL9*, *CXCL10*, *CXCL11*, *IDO1*, *GBP1*, *GBP2*, *GBP4* and *GBP5*) or high antigen presenting and phagocytic capacity (high levels of HLA class II genes and gene ontology pathways enrichment for endocytosis). Trajectory analysis predicted the majority of these subtypes arise from bone

marrow-derived monocytes<sup>113</sup>. In the submucosal space, the majority of macrophages expressed *LYVE1* (associated with vasculature) and *COLECT12* (associated with neurons) and had low antigen presenting capacity, but high chemotactic and tissue-protective properties (Fig. 3). Twelve transcriptionally distinct populations of macrophages were present in the muscularis propria with proinflammatory properties (e.g., expression of *IL1A*, *IL1B*, *CXCL2*, *CXCL3*, *CXCL8*, *CCL3* and *CCL4*) and homeostatic properties (e.g., expression of *LILRB5*, *MARCO*, *LYVE1*, *FOLR2* and *COLECT12*). Homeostatic muscularis macrophages were also positive for *PMP22* and *EMP1*, genes expressed by Schwann cells, suggesting these macrophages phagocytose Schwann cells and are in close contact with neurons<sup>113</sup>. Macrophages in both compartments showed ligand/receptor expression enabling them to interact extensively with tissue resident cells indicating that their expression profile is heavily influenced by their local microenvironment<sup>113</sup>.

Intestinal epithelial cells arise from a common stem cell at the crypt base and transdifferentiate as they move towards the villus tip. Although the positions of enterocytes along the villus axis correlate with their age<sup>123</sup>, exposure to morphogen gradients<sup>124</sup>, and hypoxia<sup>125</sup>, low-resolution approaches were unable to determine the positional effects on enterocyte function in mice<sup>126,127</sup> or humans<sup>128</sup>. A study by Moor et al. applied laser capture microdissection of mouse enterocytes followed by scRNAseq to elegantly resolve a continuous gradient of transdifferentiation along the villus axis<sup>114</sup>. The villus tip enterocytes expressed an immune-modulatory program with the capacity to modulate immune reaction to the microbiota in the gut lumen (Fig. 3)<sup>114</sup>. Follicle-associated epithelium covering the lymphoid structures (i.e., PPs) possess characteristics distinct from villus epithelium<sup>129-131</sup>. Microdissection and RNAseq of mouse intestinal epithelium followed by single-cell validation of gene expression with single molecule fluorescence in situ hybridization showed that follicle-associated epithelium expresses lower levels of antimicrobial and nutrient absorption genes<sup>132</sup>. This suggests that epithelium at these sites is tuned for the optimal and efficient sampling of bacterial antigens by M cells and immune cells, rather than nutrient absorption and antimicrobial activity (Fig. 3).

## INTESTINAL IMMUNITY SHAPED BY NON-IMMUNE INTERACTIONS

Key components of tissue microenvironments are the resident non-hematopoietic cells that have multiple established roles in immune responses and inflammation in mucosal surfaces. While previously this involvement was thought to be passive, with research focusing on fibrosis, tumour progression and wound healing, scRNAseq studies are highlighting the extent of active engagement of non-immune cells in shaping mucosal immunity with implication for health and disease progression<sup>133,134</sup>.

Mesenchymal or stromal cells of the intestine reside in the subepithelial layers and contribute largely to structural integrity. Three recent studies comparing healthy and IBD intestinal tissue have applied scRNAseq to not only map the diversity of intestinal stromal subtypes, but also pinpoint which cell subtypes and interactions are at play during inflammation<sup>16-18</sup>. Kinchen et al. defined four distinct stromal populations with unique transcriptional signatures<sup>17</sup>. One of these stromal types termed stromal 4 cells, marked by expression of genes involved in cytokine signalling, T cell activation and cell adhesion, was scarce in healthy controls and enriched in UC. Crucially, *IL-6* and *TNFSF14* were additionally expressed by stromal 4 cells during disease and shown to prevent epithelial regeneration in follow up intestinal organoids experiments. Martin et al. similarly observed stromal cells contributing to the cellular response of CD<sup>18</sup>. However, here

they defined a collective cell module (termed GIMATS) consisting of IgG plasma cells, inflammatory mononuclear phagocytes, activated T cells, and stromal cells that corresponded with failure to achieve durable corticosteroid-free remission upon anti-TNF therapy in a fraction of patients. A real strength of scRNAseq here was the capacity to compare ligand receptor pair expression between equivalent cells of patient groups. In this way, the authors showed that enriched cellular interactions between myeloid cells, activated endothelial cells and activated CCL2+ CCL7+ stromal cells were generating a positive inflammatory feedback loop in the GIMATS samples<sup>18</sup>. Last, Smillie et al. identified a population of inflammation associated fibroblasts (IAFs) that were expanded 189-fold in biopsies from UC patients versus controls<sup>16</sup>. The profile of IAFs was comparable to cancer-associated fibroblasts, a key player in creating an immune tolerant tumour environment. IAFs also highly express *OSMR*, a predictor of resistance to anti-TNF therapy in UC patients, and ranked high for a resistance gene signature determined from bulk RNAseq data. The gene encoding the ligand for OSMR, *OSM* was most strongly expressed by inflammatory monocytes and cDC2s in the scRNAseq data, implicating interactions between these cells in resistance to treatment<sup>16</sup>.

Tuft cells are chemosensors of the gut epithelium, transmitting messages in the form of a spectrum of biological effector molecules to immune and neuronal cells<sup>135</sup>. Previous bulk RNAseq had identified neuronal and inflammation gene signatures from these cells<sup>136</sup>, but was unable to resolve whether these programs were from one population of cells or distinct subtypes. Using scRNAseq Haber et al. carried out unbiased clustering of Tuft cells from the small intestines of mice and identified two distinct subsets contributing these profiles<sup>137</sup>. Tuft-2 cells were enriched for immune-related genes particularly those supporting Th2 responses (*Il4ra* and *Il13ra1* and *Il17rb*). Incredibly, this population also expressed *Ptprc* (encoding the pan-immune marker CD45), the first recording of this in non-hematopoietic cells and blurring the lines of the traditional definition of immune cells<sup>137</sup>. While equivalent findings have not been made from single cell analysis of human intestinal tuft cells, a fraction of human and mouse Tuft cells were shown to express immune signalling machinery, specifically activating and inhibitory Fc gamma receptors and downstream mediators for IgG signalling<sup>21</sup>. This could facilitate direct activation of Tuft cells in response to signals from plasma cells. These findings formed the basis of experiments in mouse models of intestinal colitis in which Tuft cells upregulated the inhibitory receptor suggested their potential as a rheostat of intestinal inflammation<sup>21</sup>.

Antigen presentation is a critical step in the transmission of immune activation to the adaptive immune system with primary antigen presenters regarded as conventional DCs, macrophages and naive B cells. A body of prior work has extended this role to various epithelial cell types via MHC-II expression, with particular roles for microfold cells localised to PPs<sup>138</sup>. Recent scRNAseq experiments have taken this further to pinpoint exact sub-populations. Work from the Xavier and Regev laboratories combined scRNAseq, flow cytometry and immunofluorescence assays to define three novel subtypes of Lgr5+ intestinal stem cells in the mouse small intestine<sup>82</sup>. Although not as high as DCs, two of these populations expressed MHC-II at significant levels and were capable of presenting antigen to antigen-specific T cells in co-culture experiments. While the exact role of antigen presentation by these cells is unknown, the authors speculated that it could be a non-essential means for the epithelial layer to respond to infection or be a means by which T helper cells can interact with ISCs and shape their appropriate differentiation into mature epithelial cell types. The latter explanation is particularly interesting in light of further results showing Tregs cells promote ISC renewal while Th1 and Th17 cells promote differentiation<sup>82</sup>.

### Box 1. Areas of open investigation

- Resolving the positioning of intestinal cells using in situ transcriptomics.
- Integration of scRNAseq with other modalities e.g., metatranscriptomics, metabolomics, proteomics.
- Systematic analysis of therapeutic mechanisms in in vitro and models systems.
- Cross tissue analysis i.e., common cell types and movement of cells between organ systems.

### CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Single-cell studies have provided a wealth of knowledge about the complex cellular landscape of the intestinal immune system. In their short history, they have detailed the spectrum of cell phenotypes, provided resolution of zonation of immune cells and shown the impact of their engagement with the neighbouring microbiota in health and disease. As these methods continue to evolve, there is little doubt that they will continue to provide insights into the field of intestinal immunity (Box 1).

Spatial transcriptomics, while not yet at the single-cell with whole transcriptome level, has already placed gut immune signatures in their tissue context<sup>21,24</sup> and will be a key feature of future studies. The application of scRNAseq to in vitro systems and experimental models will offer the ability to look in detail at the mechanism of therapeutic and biological agents (e.g., faecal microbiota) on intestinal immune cells. ScRNAseq has already been adapted for capture of bacterial RNA in pioneering studies<sup>139</sup>, opening the possibility to study the function of specific bacteria. Integration of modalities for example combining scRNAseq of host cells with single-cell metatranscriptomics and patient genotype data will also provide the opportunity to study the interaction between these factors in shaping intestinal immune environments. Last, as the first chapter of the HCA approaches completion<sup>140</sup>, studies of individual organ systems will be combined to provide a global picture of human biology. We anticipate that this will bring with it studies of the contribution of gut immune cells to human biology and disease at a systems-wide level.

### REFERENCES

1. Besredka, A. De la vaccination contra les ctsats typhoid par la voic buccale. *Ann. Inst. Pasteur* **33**, 882 (1919).
2. Tang, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
3. Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A. & Teichmann, S. A. Single-cell transcriptomics to explore the immune system in health and disease. *Science* **358**, 58–63 (2017).
4. Papalexi, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* **18**, 35–45 (2018).
5. Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550 (2016).
6. Reynolds, G. et al. Developmental cell programs are co-opted in inflammatory skin disease. *Science* **371**, eaba6500 (2021).
7. Li, Y. et al. Single-cell transcriptome profiling of the vaginal wall in women with severe anterior vaginal prolapse. *Nat. Commun.* **12**, 87 (2021).
8. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* **563**, 347–353 (2018).
9. Williams, D. W. et al. Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. *Cell* **184**, 4090–4104.e15 (2021).
10. Qi, F., Qian, S., Zhang, S. & Zhang, Z. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochem. Biophys. Res. Commun.* **526**, 135–140 (2020).
11. Ziegler, C. G. K. et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **181**, 1016–1035.e19 (2020).
12. Sungnak, W. et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020).
13. Lukassen, S. et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J.* **39**, e105114 (2020).



14. Zhang, H. et al. Specific ACE2 expression in small intestinal enterocytes may cause gastrointestinal symptoms and injury after 2019-nCoV infection. *Int. J. Infect. Dis.* **96**, 19–24 (2020).
15. Parikh, K. et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49–55 (2019).
16. Smillie, C. S. et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* **178**, 714–730.e22 (2019).
17. Kinchen, J. et al. Structural remodeling of the human colonic mesenchyme in inflammatory bowel disease. *Cell* **175**, 372–386.e17 (2018).
18. Martin, J. C. et al. Single-cell analysis of Crohn's Disease Lesions identifies a pathogenic cellular module associated with resistance to Anti-TNF Therapy. *Cell* **178**, 1493–1508.e20 (2019).
19. Wang, Y. et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. *J. Exp. Med.* **217**, e20191130 (2020).
20. James, K. R. et al. Distinct microbial and immune niches of the human colon. *Nat. Immunol.* **21**, 343–353 (2020).
21. Elmentaite, R. et al. Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250–255 (2021).
22. Gao, S. et al. Publisher Correction: Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. *Nat. Cell Biol.* **20**, 1227 (2018).
23. Elmentaite, R. et al. Single-cell sequencing of developing human gut reveals transcriptional links to childhood Crohn's Disease. *Dev. Cell* **55**, 771–783.e5 (2020).
24. Fawcner-Corbett, D. et al. Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* **184**, 810–826.e23 (2021).
25. Gomes, T., Teichmann, S. A. & Talavera-López, C. Immunology driven by large-scale single-cell sequencing. *Trends Immunol.* **40**, 1011–1021 (2019).
26. Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
27. Haque, A., Engel, J., Teichmann, S. A. & Lönnberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* **9**, 75 (2017).
28. Efrimova, M., Vento-Tormo, R., Park, J.-E., Teichmann, S. A. & James, K. R. Immunology in the Era of single-cell technologies. *Annu. Rev. Immunol.* **38**, 727–757 (2020).
29. Islam, S. et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11**, 163–166 (2013).
30. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* **2**, 666–673 (2012).
31. Jaitin, D. A. et al. Massively parallel single-cell RNA-seq for marker-free deconvolution of tissues into cell types. *Science* **343**, 776–779 (2014).
32. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
33. Hagemann-Jensen, M. et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* **38**, 708–714 (2020).
34. Mamanova, L. et al. High-throughput full-length single-cell RNA-seq automation. *Nat. Protoc.* **16**, 2886–2915 (2021).
35. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
36. Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).
37. Hughes, T. K. et al. Second-strand synthesis-based massively parallel scRNA-seq reveals cellular states and molecular features of human inflammatory skin pathologies. *Immunity* **53**, 878–894.e7 (2020).
38. Hochgerner, H. et al. STRT-seq-2i: dual-index 5' single cell and nucleus RNA-seq on an addressable microwell array. *Sci. Rep.* **7**, 8 (2017).
39. Macaulay, I. C. et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **12**, 519–522 (2015).
40. Stoekius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
41. Peterson, V. M. et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).
42. Kleshchevnikov, V. et al. Comprehensive mapping of tissue cell architecture via integrated single cell and spatial transcriptomics. *bioRxiv* <https://doi.org/10.1101/2020.11.15.378125>.
43. Andrews, T. S., Kiselev, V. Y., McCarthy, D. & Hemberg, M. Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. *Nat. Protoc.* **16**, 1–9 (2020).
44. Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15**, e8746 (2019).
45. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
46. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
47. Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308.e36 (2018).
48. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* **7**, giy059 (2018).
49. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
50. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
51. Amezquita, R. A. et al. Orchestrating single-cell analysis with Bioconductor. *Nat. Methods* **17**, 137–145 (2020).
52. Herring, C. A., Chen, B., McKinley, E. T. & Lau, K. S. Single-cell computational strategies for lineage reconstruction in tissue systems. *Cell Mol. Gastroenterol. Hepatol.* **5**, 539–548 (2018).
53. Armingol, E., Officier, A., Harismendy, O. & Lewis, N. E. Deciphering cell-cell interactions and communication from gene expression. *Nat. Rev. Genet.* **22**, 71–88 (2021).
54. Pasquini, G., Rojo Arias, J. E., Schäfer, P. & Busskamp, V. Automated methods for cell type annotation on scRNA-seq data. *Comput. Struct. Biotechnol. J.* **19**, 961–969 (2021).
55. Conde, C. D. et al. Cross-tissue immune cell analysis reveals tissue-specific adaptations and clonal architecture in humans. *bioRxiv* <https://doi.org/10.1101/2021.04.28.441762>.
56. Kimmel, J. C. & Kelley, D. R. Semisupervised adversarial neural networks for single-cell classification. *Genome Res.* <https://doi.org/10.1101/gr.268581.120> (2021).
57. Alavi, A., Ruffalo, M., Parvanga, A., Huang, Z. & Bar-Joseph, Z. A web server for comparative analysis of single-cell RNA-seq data. *Nat. Commun.* **9**, 4768 (2018).
58. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
59. Bendall, S. C. et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell* **157**, 714–725 (2014).
60. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
61. Ding, J., Lin, C. & Bar-Joseph, Z. Cell lineage inference from SNP and scRNA-Seq data. *Nucleic Acids Res.* **47**, e56 (2019).
62. Xu, J. et al. Single-cell lineage tracing by endogenous mutations enriched in transposase accessible mitochondrial DNA. *Elife* **8**, e45105 (2019).
63. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
64. Haniffa, M. et al. A roadmap for the human developmental cell atlas. *Nature* **597**, 196–205 (2021).
65. Haley, P. J. The lymphoid system: a review of species differences. *J. Toxicol. Pathol.* **30**, 111–123 (2017).
66. Krishnamurthy, A. T. & Turley, S. J. Lymph node stromal cells: cartographers of the immune system. *Nat. Immunol.* **21**, 369–380 (2020).
67. Rothschild, D. et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).
68. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
69. Marchesi, J. R. et al. The gut microbiota and host health: a new clinical frontier. *Gut* **65**, 330–339 (2016).
70. Kennedy, K. M. et al. Fetal meconium does not have a detectable microbiota before birth. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-021-00904-0> (2021).
71. Rackaityte, E. et al. Viable bacterial colonization is highly limited in the human intestine in utero. *Nat. Med.* **26**, 599–607 (2020).
72. Silverstein, R. B. & Mysorekar, I. U. Group therapy on in utero colonization: seeking common truths and a way forward. *Microbiome* **9**, 7 (2021).
73. Mishra, A. et al. Microbial exposure during early human development primes fetal immune cells. *Cell* <https://doi.org/10.1016/j.cell.2021.04.039> (2021).
74. Schreurs, R. R. C. E. et al. Intestinal CD8 T cell responses are abundantly induced early in human development but show impaired cytotoxic effector capacities. *Mucosal Immunol.* **14**, 605–614 (2021).
75. Stras, S. F. et al. Maturation of the human intestinal immune system occurs early in fetal development. *Dev. Cell* **51**, 357–373.e5 (2019).
76. Li, N. et al. Memory CD4 T cells are generated in the human fetal intestine. *Nat. Immunol.* **20**, 301–312 (2019).
77. Schreurs, R. R. C. E. et al. Human fetal TNF- $\alpha$ -Cytokine-Producing CD4 effector memory T cells promote intestinal development and mediate inflammation early in life. *Immunity* **50**, 462–476.e8 (2019).
78. Mihi, B. et al. Interleukin-22 signaling attenuates necrotizing enterocolitis by promoting epithelial cell regeneration. *Cell Rep. Med.* **2**, 100320 (2021).
79. Tanner, S. M. et al. Pathogenesis of necrotizing enterocolitis: modeling the innate immune response. *Am. J. Pathol.* **185**, 4–16 (2015).

80. Neu, J. & Allan Walker, W. Necrotizing Enterocolitis. *N. Engl. J. Med.* **364**, 255–264 (2011).
81. Cho, S. X. et al. Characterization of the pathoimmunology of necrotizing enterocolitis reveals novel therapeutic opportunities. *Nat. Commun.* **11**, 5794 (2020).
82. Biton, M. et al. T helper cell cytokines modulate intestinal stem cell renewal and differentiation. *Cell* **175**, 1307–1320.e22 (2018).
83. Morrow, A. L. et al. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. *Microbiome* **1**, 13 (2013).
84. Pammi, M. et al. Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome* **5**, 31 (2017).
85. Olm, M. R. et al. Necrotizing enterocolitis is preceded by increased gut bacterial replication, Klebsiella, and fimbriae-encoding bacteria. *Sci. Adv.* **5**, eaax5727 (2019).
86. Gury-BenAri, M. et al. The spectrum and regulatory landscape of intestinal innate lymphoid cells are shaped by the microbiome. *Cell* **166**, 1231–1246.e13 (2016).
87. Mazzurana, L. et al. Tissue-specific transcriptional imprinting and heterogeneity in human innate lymphoid cells revealed by full-length single-cell RNA-sequencing. *Cell Res.* **31**, 554–568 (2021).
88. Kiner, E. et al. Gut CD4 T cell phenotypes are a continuum molded by microbes, not by TH archetypes. *Nat. Immunol.* **22**, 216–228 (2021).
89. Corridoni, D. et al. Single-cell atlas of colonic CD8 T cells in ulcerative colitis. *Nat. Med.* **26**, 1480–1490 (2020).
90. Bernink, J. H. et al. Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. *Nat. Immunol.* **14**, 221–229 (2013).
91. Dieffenbach, A., Colonna, M. & Koyasu, S. Development, differentiation, and diversity of innate lymphoid cells. *Immunity* **41**, 354–365 (2014).
92. Björklund, Å. K. et al. The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing. *Nat. Immunol.* **17**, 451–460 (2016).
93. Brucklacher-Waldert, V., Carr, E. J., Linterman, M. A. & Veldhoen, M. Cellular plasticity of CD4+ T cells in the intestine. *Front. Immunol.* **5**, 12 (2014).
94. Lönnberg, T. et al. Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria. *Sci Immunol* **2**, 9 (2017).
95. Tibbitt, C. A. et al. Single-Cell RNA sequencing of the T helper cell response to house dust mites defines a distinct gene expression signature in airway Th2 cells. *Immunity* **51**, 169–184.e5 (2019).
96. Proserpio, V. et al. Single-cell analysis of CD4+ T-cell differentiation reveals three major cell states and progressive acceleration of proliferation. *Genome Biol.* **17**, 103 (2016).
97. Vieira Braga, F. A. et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* **25**, 1153–1163 (2019).
98. Stephenson, E. et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.* **27**, 904–916 (2021).
99. Gálvez, J. Role of Th17 cells in the pathogenesis of human IBD. *ISRN Inflamm.* **2014**, 1–14 (2014).
100. Zhang, L. et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268–272 (2018).
101. Pelka, K. et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* **184**, 4734–4752.e20 (2021).
102. Okumura, R. & Takeda, K. Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Exp. Mol. Med.* **49**, e338–e338 (2017).
103. Cunliffe, R. N. & Mahida, Y. R. Expression and regulation of antimicrobial peptides in the gastrointestinal tract. *J. Leukoc. Biol.* **75**, 49–58 (2004).
104. Bowcutt, R. et al. Heterogeneity across the murine small and large intestine. *World J. Gastroenterol.* **20**, 15216–15232 (2014).
105. Ito, G. et al. Lineage-specific expression of bestrophin-2 and bestrophin-4 in human intestinal epithelial cells. *PLoS ONE* **8**, e79693 (2013).
106. Yu, K., Lujan, R., Marmorstein, A., Gabriel, S. & Criss Hartzell, H. Bestrophin-2 mediates bicarbonate transport by goblet cells in mouse colon. *J. Clin. Invest.* **120**, 1722–1735 (2010).
107. Jakab, R. L., Collaco, A. M. & Ameen, N. A. Characterization of CFTR High Expresser cells in the intestine. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G453–G465 (2013).
108. Ameen, N. A., Ardito, T., Kashgarian, M. & Marino, C. R. A unique subset of rat and human intestinal villus cells express the cystic fibrosis transmembrane conductance regulator. *Gastroenterology* **108**, 1016–1023 (1995).
109. Busslinger, G. A. et al. Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution. *Cell Rep.* **34**, 108819 (2021).
110. Henen, S., Denton, C., Teckman, J., Borowitz, D. & Patel, D. Review of gastrointestinal motility in cystic fibrosis. *J. Cyst. Fibros.* <https://doi.org/10.1016/j.jcf.2021.05.016> (2021).
111. Brandtzaeg, P. & Johansen, F.-E. Mucosal B cells: phenotypic characteristics, transcriptional regulation, and homing properties. *Immunol. Rev.* **206**, 32–63 (2005).
112. Rao, J. N. & Wang, J.-Y. Intestinal architecture and development. In *Regulation of Gastrointestinal Mucosal Growth* (Morgan & Claypool Life Sciences, 2010).
113. Domanska, D. et al. Spatiotemporal single cell transcriptomic analysis of human gut macrophages reveals multiple functional and niche-specific subsets. *bioRxiv* <https://doi.org/10.1101/2021.05.11.443586>.
114. Moor, A. E. et al. Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *Cell* **175**, 1156–1167.e15 (2018).
115. Sullivan, Z. A. et al.  $\gamma\delta$  T cells regulate the intestinal response to nutrient sensing. *Science* **371**, eaba8310 (2021).
116. Huang, B. et al. Mucosal profiling of pediatric-onset colitis and IBD reveals common pathogenics and therapeutic pathways. *Cell* **179**, 1160–1176.e24 (2019).
117. Atlasy, N. et al. Single cell transcriptome atlas of immune cells in human small intestine and in celiac disease. *bioRxiv* <https://doi.org/10.1101/721258>.
118. Uniken Venema, W. T. et al. Single-Cell RNA sequencing of blood and ileal T cells from patients with Crohn's disease reveals tissue-specific characteristics and drug targets. *Gastroenterology* **156**, 812–815.e22 (2019).
119. Gálvez, J. et al. Role of Th17 cells in the pathogenesis of human IBD. *ISRN Inflamm.* **2014**, 928461 (2014).
120. Jaeger, N. et al. Single-cell analyses of Crohn's disease tissues reveal intestinal intraepithelial T cells heterogeneity and altered subset distributions. *Nat. Commun.* **12**, 1921 (2021).
121. De Schepper, S., Stakenborg, N., Matteoli, G., Verheijden, S. & Boeckstaens, G. E. Muscularis macrophages: key players in intestinal homeostasis and disease. *Cell. Immunol.* **330**, 142–150 (2018).
122. Gabanyi, I. et al. Neuro-immune Interactions drive tissue programming in intestinal macrophages. *Cell* **164**, 378–391 (2016).
123. Moor, A. E. & Itzkovitz, S. Spatial transcriptomics: paving the way for tissue-level systems biology. *Curr. Opin. Biotechnol.* **46**, 126–133 (2017).
124. Crosnier, C., Stamatakis, D. & Lewis, J. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat. Rev. Genet.* **7**, 349–359 (2006).
125. Zheng, L., Kelly, C. J. & Colgan, S. P. Physiologic hypoxia and oxygen homeostasis in the healthy intestine. A review in the theme: cellular responses to Hypoxia. *Am. J. Physiol. Cell Physiol.* **309**, C350–C360 (2015).
126. Mariadason, J. M. et al. Gene expression profiling of intestinal epithelial cell maturation along the crypt-villus axis. *Gastroenterology* **128**, 1081–1088 (2005).
127. Stegmann, A. et al. Metabolome, transcriptome, and bioinformatic cis-element analyses point to HNF-4 as a central regulator of gene expression during enterocyte differentiation. *Physiol. Genomics* **27**, 141–155 (2006).
128. George, M. D. et al. In vivo gene expression profiling of human intestinal epithelial cells: analysis by laser microdissection of formalin fixed tissues. *BMC Genomics* **9**, 209 (2008).
129. Neutra, M. R., Pringault, E. & Kraehenbuhl, J. P. Antigen sampling across epithelial barriers and induction of mucosal immune responses. *Annu. Rev. Immunol.* **14**, 275–300 (1996).
130. Owen, R. L. Uptake and transport of intestinal macromolecules and microorganisms by M cells in Peyer's patches—a personal and historical perspective. *Semin. Immunol.* **11**, 157–163 (1999).
131. Kato, T. & Owen, R. L. Structure and function of intestinal mucosal epithelium. *Mucosal Immunology* **131–151** <https://doi.org/10.1016/b978-012491543-5/50012-7> (2005).
132. Cohen, N. et al. Spatial gene expression maps of the intestinal lymphoid follicle and associated epithelium identify zoned expression programs. *PLoS Biol.* **19**, e3001214 (2021).
133. Elmentaite, R., Teichmann, S. A. & Madisson, E. Studying immune to non-immune cell cross-talk using single-cell technologies. *Curr. Opin. Syst. Biol.* **18**, 87–94 (2019).
134. Owens, B. M. J. & Simmons, A. Intestinal stromal cells in mucosal immunity and homeostasis. *Mucosal Immunol.* **6**, 224–234 (2013).
135. Schneider, C., O'Leary, C. E. & Locksley, R. M. Regulation of immune responses by tuft cells. *Nat. Rev. Immunol.* **19**, 584–593 (2019).
136. Bezençon, C. et al. Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. *J. Comp. Neurol.* **509**, 514–525 (2008).
137. Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
138. Wosen, J. E., Mukhopadhyay, D., Macaubas, C. & Mellins, E. D. Epithelial MHC Class II expression and its role in antigen presentation in the gastrointestinal and respiratory tracts. *Front. Immunol.* **9**, 2144 (2018).
139. Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol.* **5**, 1202–1206 (2020).

140. Lindeboom, R. G. H., Regev, A. & Teichmann, S. A. Towards a human cell atlas: taking notes from the past. *Trends Genet.* **37**, 625–630 (2021).

### AUTHOR CONTRIBUTIONS

All authors contributed to the conception and refinement of this review and approve the final version. The review was written and revised by K.R.J., R.E. and G.L.H.

### COMPETING INTERESTS

In the past three years, S.A.T has consulted or been a member of Scientific Advisory Boards at Roche, Qiagen, Genentech, Biogen, GlaxoSmithKline and ForeSite Labs. G.L.H. has consulted or been a member of Scientific Advisory Boards at Ferring, Procter and Gamble and Enterobiotix. The remaining authors declare no conflicts of interest.

### ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Kylie Renee James.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2022