Decoding changes in tumor-infiltrating leukocytes through dynamic experimental models and single-cell technologies
Lee, Colin YC; Clatworthy, Menna R; Withers, David R

DOI: 10.1111/imcb.12787
License: Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
Decoding changes in tumor-infiltrating leukocytes through dynamic experimental models and single-cell technologies

Colin YC Lee1, Menna R Clatworthy1 & David R Withers2

1 Cambridge Institute of Therapeutic Immunology and Infection Disease, University of Cambridge, Cambridge, UK
2 Institute of Immunology and Immunotherapy, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Keywords
cancer, immune response, in vivo models, single-cell technologies, temporal dynamics

Abstract
The ability to characterize immune cells and explore the molecular interactions that govern their functions has never been greater, fueled in recent years by the revolutionary advance of single-cell analysis platforms. However, precisely how immune cells respond to different stimuli and where differentiation processes and effector functions operate remain incompletely understood. Inferring cellular fate within single-cell transcriptomic analyses is now omnipresent, despite the assumptions typically required in such analyses. Recently developed experimental models support dynamic analyses of the immune response, providing insights into the temporal changes that occur within cells and the tissues in which such transitions occur. Here we will review these approaches and discuss how these can be combined with single-cell technologies to develop a deeper understanding of the immune responses that should support the development of better therapeutic options for patients.

INTRODUCTION
Determining how best to manipulate immune responses for the benefit of patients with cancer is fearsomely complex. Without invoking subsets or activation states, there is a wealth of immune cells to consider within the response. These cells utilize an array of effector mechanisms, many of which overlap among different cell types, creating challenges in defining specific cellular contributions to the response. Furthermore, while the response may be focused toward a single tumor site, its orchestration requires cells to traffic between, and interact within, multiple geographically distinct tissue environments. This is well illustrated by the reliance on draining lymphoid tissue to generate adaptive immune responses. Finally, cancers propagate a local environment that alters the functions of the incoming immune cells, blunting the immune response to facilitate tumor growth. Thus, not only do we need to characterize the nature of the responding immune cells in as much detail as possible, but also we need to understand where key interactions occur and how these impact immune cell fate and function.

Single-cell “omics” (transcriptomics, epigenomics and proteomics) have exponentially increased the number of parameters we can simultaneously measure compared with flow cytometry, and more recently, mass cytometry approaches that have been the workhorse of immunology for many years. Crucially, single-cell “omic” approaches are not biased by the need to preselect a limited panel of molecules for investigation. This has revolutionized our ability to describe immune cells, providing an unprecedented and transformative level of detail. Despite this, understanding the dynamics of the response, what happens over time to different cells and where this occurs remains a fundamental challenge. Here, we will briefly summarize current bioinformatics approaches designed to infer cellular differentiation and fate. We will then review experimental approaches that can be used to better track cell fate, with particular attention to in vivo models that provide exciting opportunities to assess immune cell dynamics when combined with single-cell technologies.
Finally, we will discuss potential experimental refinements that might be utilized in the future.

**In silico approaches to infer cell fate**

**Opportunities from single-cell data**

Before the introduction of single-cell assays, bulk microarrays and RNA sequencing (RNA-seq) offered a global view of the transcriptome to describe how cells and tissues changed over time or in response to a challenge. However, given the heterogeneous cell populations analyzed, accurate conclusions regarding individual cell fates were impossible. With the widespread application of single-cell RNA-seq (scRNA-seq) to tissue samples, not only are researchers able to examine constituent cell types in detail, but also distinguish similar cell states and decode their potential relationships. The application of scRNA-seq to tumor samples has enabled the profiling of tumor-infiltrating leukocytes at unprecedented resolution, driving forward our understanding of how immune states relate to clinical outcomes. For example, scRNA-seq analysis of tumor biopsies from patients with melanoma revealed that the presence of Tcf7-expressing CD8 T cells, now understood to contain a "stem-like" compartment, was predictive of clinical response to immune checkpoint therapy. The use of single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) identified cis-elements regulating the expression of PDCD1 and other inhibitory receptors in exhausted human CD8 T cells, building insights into epigenetic control of cellular differentiation.

The infiltration, differentiation and adaptation of leukocytes within tumors are nonstationary biological processes; therefore, sampled cells invariably traverse a continuum of cellular states. For example, cytotoxic CD8+ T cells progress through stages of activation followed by dysfunction in cancer, termed “exhaustion.” The high-dimensional nature of scRNA-seq (i.e. simultaneous measurement of thousands of transcripts) has been leveraged to order cells along a putative trajectory based on gradual differences in their transcriptomes. Representation of cells as a continuous manifold may help us understand the relationship between cell states and a temporal variable, and offers a compelling approach to reconstruct dynamic processes. These computational methods are referred to as trajectory inference or pseudotime analysis.

**Trajectory inference and pseudotime analyses**

Many such methods exist, and each possesses unique characteristics. The specific comparison and benchmarking of different single-cell trajectory inference methods have been reviewed elsewhere by Saelens and colleagues. Notably, these methods vary in their accuracy, usability, scalability and may only be suitable for cells with specific population structures. We include a summary of several popular approaches used by the immunology community in Table 1. Attesting to their usefulness, these methods have been applied to tumor CD8 T cells and have recapitulated the progressive acquisition of an exhausted state. Moreover, in silico trajectory inference has helped uncover that antigen-specific “stem-like” PD-1+ CD8 T cells differentiate into functional effector cells followed by progression to a terminally exhausted state, indicating that PD-1+ expressing cells may be functional in human cancers. These studies highlight the usefulness of trajectory inference approaches in deriving biologically meaningful insight into tumor leukocyte dynamics.

However, pseudotime analyses are subject to limitations. Priors are often required as input variables (i.e. defining start and/or end points), which rely on assumptions from potentially incomplete domain knowledge. Complex lineage relationships are not always recapitulated because cell-state transitions may not strictly be “tree-like,” and many tools are limited to the inference of unidirectional linear or diverging topologies. For example, it is not possible to accurately represent the multidirectional relationship between naive, activated, stem-like memory and exhausted tumor CD8+ T-cell populations on a unidirectional manifold. Moreover, oscillating cell states, such as the cell cycle or germinal center B-cell dynamics, confound pseudotemporal ordering. Finally, feature selection for temporal ordering can affect output; insufficient or biased cell sampling may preclude representation of rare, transient cell states and result in incomplete trajectories; and more broadly, transcriptomic similarity may not imply a temporal or developmental relationship.

**RNA velocity**

RNA velocity is another common approach to infer directionality in scRNA-seq data. Briefly, the ratio of unspliced (pre-messenger RNA) to spliced (mature messenger RNA) reads can be used to infer the underlying kinetics of gene expression, where gene activation is associated with increased abundance of unspliced transcripts and the converse in gene downregulation. The combined velocities over all features measured are used to estimate the future state of any individual cell. However, this approach also has its limitations. RNA velocity is susceptible to inaccuracies, such as those caused by cell populations exhibiting abrupt and discontinuous transcriptional activity (transcriptional bursts). Improbable trajectories have
<table>
<thead>
<tr>
<th>Method</th>
<th>Synopsis (i.e. what it does)</th>
<th>User notes (i.e. when to use)</th>
<th>Example applications to tumor immunity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monocle v1</td>
<td>• Constructs an MST on a reduced-dimension representation of scRNA-seq data&lt;br&gt;• Orders cells along the longest path through the MST</td>
<td>• Minimum priors: start cell or direction of trajectory, number of lineages&lt;br&gt;• Pioneering method for trajectory inference in scRNA-seq data&lt;br&gt;• Poor scalability for large data sets&lt;br&gt;• Requires many user-defined priors&lt;br&gt;• Deprecated—replaced by v3</td>
<td>CD8 T-cell fate in melanoma and checkpoint immunotherapy&lt;br&gt;CD8 T-cell lineage relationships in colorectal cancer</td>
</tr>
<tr>
<td>Monocle v2</td>
<td>• Dimensionality reduction using DDRTree, a reversed graph embedding algorithm, to improve learning of population topology&lt;br&gt;• DDRTree enables differentiation branching events to be detected in an unsupervised manner, instead of relying on user input (v1)</td>
<td>• Minimum priors: start cell&lt;br&gt;• Improved branch detection from v1 for cell populations with more complex topologies&lt;br&gt;• BEAM functionality for analysis of branch point-dependent genes&lt;br&gt;• Deprecated—replaced by v3, downstream analysis functionality transferred to v3</td>
<td>Neutrophil–monocyte lineage in multiple myeloma&lt;br&gt;Monocyte differentiation in hepatocellular carcinoma&lt;br&gt;HPV-specific PD-1–positive cells in head and neck cancer</td>
</tr>
<tr>
<td>Monocle v3</td>
<td>• Projects cells onto low-dimensional UMAP space&lt;br&gt;• Groups similar cells using Louvain community detection&lt;br&gt;• Learns principal graph in the UMAP space&lt;br&gt;• Resolves trajectories and pseudotime values using principal graph</td>
<td>• Minimum priors: start cell&lt;br&gt;• Improved scalability from v1 and v2, to enable analyses of millions of cells&lt;br&gt;• Supports trajectories with multiple roots to detect convergent trajectories&lt;br&gt;• Partitions cells (adopting approach from PAGA, see below) to learn disjointed or parallel trajectories</td>
<td>Neutrophil responses in immunotherapy&lt;br&gt;T-cell trajectories in metastatic ovarian cancer&lt;br&gt;T-cell dynamics in anti-PD-1–treated squamous cell carcinoma&lt;br&gt;CD8 T-cell differentiation in non-small-cell lung cancer</td>
</tr>
<tr>
<td>Slingshot</td>
<td>• Infers global lineage structure by applying an MST to cell clusters&lt;br&gt;• Estimates pseudotime variables for cells along each lineage using principal curves to smoothen lineages</td>
<td>• Minimum priors: start cell, end points optional&lt;br&gt;• Highest usability score in the benchmarking study (i.e. easiest to implement)&lt;br&gt;• Performs well on cell populations with simple topology&lt;br&gt;• Partially relies on or is affected by prior cell clustering</td>
<td>T-cell differentiation in anti-PD-1–treated breast tumors&lt;br&gt;Mutation-associated neoantigen-specific T-cell fates in lung cancer</td>
</tr>
<tr>
<td>Diffusion pseudotime</td>
<td>• Single-cell transcriptomes represented as a weighted kNN graph computed on a suitable low-dimensional representation (e.g. PCA)&lt;br&gt;• Determines the probability of cell-to-cell transitions using random walks of arbitrary length&lt;br&gt;• over the kNN graph</td>
<td>• Minimum priors: start cell&lt;br&gt;• Performs well on simple topologies, but not optimized to detect complex population structures&lt;br&gt;• Direct implementation in Scanpy to maximize usability&lt;br&gt;• Developers recommend a combination with PAGA (see below) for the detection of branching events</td>
<td>T-cell dynamics in anti-PD-1–treated squamous cell carcinoma&lt;br&gt;CD8 T-cell differentiation in non-small-cell lung cancer</td>
</tr>
<tr>
<td>PAGA</td>
<td>• kNN graph is partitioned at a desired resolution where partitions represent groups of connected cells; partitions computed using the Louvain community detection algorithm (default)</td>
<td>• Minimum priors: none, but does not infer ordered values or directionality&lt;br&gt;• Detects disconnected trajectories and resolves complex cell population topologies&lt;br&gt;• High scalability&lt;br&gt;• Direct implementation in Scanpy to maximize usability</td>
<td>T-cell trajectories in metastatic ovarian cancer&lt;br&gt;Neutrophil responses in immunotherapy</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Method</th>
<th>Synopsis (i.e. what it does)</th>
<th>User notes (i.e. when to use)</th>
<th>Example applications to tumor immunity</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEAM</td>
<td>Nodes (partitions) are connected by a weighted edge that represents a statistical measure of connectivity</td>
<td>Coarse-grained representation of trajectories—connectivity between partitions (clusters) not individual cells</td>
<td>T cells in metastatic ovarian cancer&lt;sup&gt;26&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Spurious edges with low weights are removed by thresholding to reveal the topology of data at a chosen resolution.</td>
<td>May be combined with distance measures or pseudotime analyses to infer pseudotime values</td>
<td>Dynamics of high-endothelial venules that provide the stem-like T-cell niche in tumors&lt;sup&gt;29&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Runs on any distance metric, including from imaging data&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Highly dependent on user-defined resolution input</td>
<td>T cells in anti–PD-1–treated breast tumors&lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>Palantir&lt;sup&gt;28&lt;/sup&gt;</td>
<td>Increases resolution and branch detection by using multiple diffusion components</td>
<td>Minimum priors: start cell, end points optional</td>
<td>Myeloid lineages in colorectal cancer&lt;sup&gt;35&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Iteratively refines the shortest path in diffusion space to compute pseudotime ordering</td>
<td>Prioritizes accuracy and resolution</td>
<td>Pan-cancer landscape and dynamics of T cells&lt;sup&gt;36&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Constructs a Markov chain using the kNN graph and pseudotime variables to model cell state transitions as a stochastic process</td>
<td>For each cell, computes differentiation potential and branch probability of differentiating into each terminal state in multifurcating topologies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Performs random walks over the Markov chain to infer terminal states</td>
<td>Able to identify points along a trajectory with drastic shifts in differentiation potential (i.e. marking key events in cell fate decision)</td>
<td></td>
</tr>
<tr>
<td>RNA velocity</td>
<td>Velocito detects the splicing state of each transcript from sequenced reads to infer the kinetics of gene expression</td>
<td>Minimum priors: none required</td>
<td></td>
</tr>
<tr>
<td>(velocito, scVelo)&lt;sup&gt;30,31&lt;/sup&gt;</td>
<td>scVelo applies a likelihood-based dynamical model to generalize RNA velocity estimation to transient systems or mixed cell populations with heterogeneous splicing kinetics</td>
<td>Uncovers inherent directionality in local cell neighborhoods from dynamic RNA transcription and processing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can be used to infer root and terminal cell states, to define inputs for trajectory analyses&lt;sup&gt;32,33&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Computationally intensive</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can be applied to single-cell epigenomes (scATAC-seq)&lt;sup&gt;34&lt;/sup&gt; to improve cell fate prediction and understand transcriptional regulation of fate decisions</td>
<td></td>
</tr>
</tbody>
</table>

BEAM, branch expression analysis modeling; HPV, human papillomavirus; kNN, k nearest neighbor; MST, minimum spanning tree; PAGA, partition-based graph abstraction; PCA, principal component analysis; UMAP, uniform manifold approximation and projection.
been inferred between independent, terminal cell populations, including unlikely directionality between blood leukocytes from distinct lineages, or trajectories in the opposite direction of known biology.\textsuperscript{42} Finally, velocity fields are only available in local cell neighborhoods, and do not provide global trends; however, a new approach, CellRank, aggregates local signals to compute longer-range trajectories.\textsuperscript{33} As with pseudotime, RNA velocity has also been successfully applied to capture state shifts in tumor-infiltrating immune cells, including CD8 T-cell exhaustion.\textsuperscript{36}

**Shortcomings of in silico approaches**

The application of pseudotime or RNA velocity analyses to single-cell data has helped infer the temporal dynamics of tumor-infiltrating leukocytes. However, when applied in isolation, there are major shortcomings. First, and most importantly, in silico methods lack an empirically measured ground truth, including where trajectories begin and end. This is particularly relevant where user-defined parameters are required, and therefore, the resulting output is completely dependent on prior assumptions. Next, scRNA-seq measurements are typically derived from a static “snapshot” description of tumor tissues, so inferred temporal dynamics do not reflect the universal time and lack inherent directionality. In other words, these methods do not provide information on the real-time rate of observed state transitions, and pseudotime coordinates can only be interpreted as a measure of the cell’s differentiation or maturation state relative to its counterparts. Combined with traversing multiple tissue compartments, these factors point to the importance of combining computational analyses with different experimental tools and in vivo models to refine investigations of tumor-infiltrating leukocytes.

**Time-series sampling**

Time-series experiments (i.e. sampling of a biological system at various time points such as after disease onset or intervention) provide additional temporal coordinates that help contextualize inferred trajectories in real-time, which may complement in silico analyses.\textsuperscript{43} Furthermore, repeated sampling can extend the absolute timescale captured by pseudotime or velocity-based applications compared with “snapshot” experiments. Indeed, this approach has been applied to investigate temporal dynamics of tumor-infiltrating leukocytes.

Chronological sampling of T cells has enabled the precise identification of the time point at which T-cell transcriptomic responses diverged in acute versus chronic viral infection and the onset of exhaustion.\textsuperscript{44} Understanding these temporal dynamics led to the identification of TOX as a key transcription factor involved in the exhaustion program. ScRNA-seq of a murine model of melanoma over successive time points revealed a continuum of T-cell states from lymph nodes (LNs) to tumors with progressive upregulation of exhaustion markers, even in early tumorigenesis, restricted to the tumor but not LN T cells.\textsuperscript{45} Independently, using mass cytometry to profile tumor macrophages in sequential days following immune checkpoint blockade, a shift in the differentiation trajectory of monocytes toward inflammatory iNOS\textsuperscript{+} macrophages was described.\textsuperscript{46} Finally, repeat biopsies and scRNA-seq of human tumors have provided important insights into the evolving tumor immune landscape and how therapeutic interventions influence leukocyte dynamics beyond cancer models in the laboratory.\textsuperscript{20,47}

A challenge in time-series data is that scRNA-seq experiments are cell-destructive and only represent a fraction of the profiled tissues. Hence, constructing trajectories requires the alignment of biologically distinct cells from different experimental samples with overlapping cell states. Computational strategies to address these added complexities have recently been described.\textsuperscript{48,49} However, sampling at sufficient density to guarantee overlap between captured cell states at sequential time points is crucial. Notably, time-series sampling provides opportunities to resolve the biological processes that underlie shifts in cell-state frequencies over time, for example, changes to proliferation rate, differentiation dynamics or death rates, particularly in response to environmental cues.\textsuperscript{40,50}

Tumor leukocytes are inadvertently at different points of their immune activation response, and indeed, it is this asynchrony that is leveraged by computational methods to predict trajectories through heterogeneous cell populations. Interestingly, a recent study attempted to synchronize the “starting point” of tumor-responding T cells to investigate their fate in precise chronology.\textsuperscript{51} This temporal calibration was achieved by infusing tumor antigen–specific in vitro–activated CD8 T cells into tumor-bearing mice. Mice were killed at various dense time points following T-cell injection, the infused cells (expressing a fluorescent reporter protein) reisolated and their transcriptomes profiled by scRNA-seq. This enabled the authors to identify a “cytotoxic window” of T-cell activity and demonstrate that immune checkpoint therapy directly modulates the real-time differentiation rate of antigen-specific T cells.

**TCR/BCR sequences provide insights into the dynamics of clonal T- or B-cell responses**

Several single-cell transcriptomic platforms enable adaptive immune receptor repertoires to be profiled
simultaneously. Specifically, B cell receptor (BCR) or T cell receptor (TCR) sequences of individual cells, paired with their transcriptome, can help delineate relationships between lymphocytes in cancer tissues. The development of B or T cells involves extensive somatic rearrangement events that generate an immense repertoire of unique BCRs or TCRs. Therefore, shared V(D)J sequences between BCRs or TCRs captured from distinct lymphocytes reflect a direct clonal relationship that may span cell states or tissue sites.52

In B cells, somatic hypermutation and affinity maturation of BCRs from germline sequences, or class-switch recombination from the default heavy-chain isotype classes, are biological processes with inherent directionality. Thus, additionally capturing these changes significantly refines analyses of lymphocyte fate and enables cross-tissue comparisons and tracking of lymphocyte dissemination. Somatic sequence alterations have been leveraged to construct phylogenetic trees of malignant clone evolution in B-cell–derived cancers,53 but importantly, they can also be used to understand B- or T-cell dynamics in epithelial tumors, particularly when performed with scRNA-seq.54–59 Several examples are outlined.

A study by Ari Hakimi and colleagues54 showed that immune checkpoint blockade altered the fate of tissue-resident T cells in renal cell carcinoma, by performing Palantir pseudotime analyses on only tissue-resident T cells expressing identical TCR sequences. This approach lends confidence to the cellular relationships inferred by pseudotime computation because all cells are clonally related, but they depend on the capture of sufficient TCRs. Similarly, RNA velocity analysis applied to scRNA/TCR-seq data identified a single hyperexpanded clone of tumor T cells that spanned the temporal spectrum of activation states.47 Further exemplifying the power of multimodal single-cell analyses, Satpathy and colleagues60 combined RNA velocity analysis of scRNA-seq, scATAC-seq and TCR-seq to map exhausted T cells in chronic viral infection, concluding that TCR signaling avidity influenced T-cell differentiation into terminally exhausted versus more functional subsets. Hence, these approaches helped inform why T cells adopt different fates. Elsewhere, Hao and colleagues65 combined Monocle pseudotime, CytoTRACE,61 a method that uses the number of detectably expressed genes as a raw signal of differentiation status, and BCR analyses to investigate the dynamics of B and plasma cells in lung adenocarcinoma. Considering the order of the genomic coordinates (5'-to-3') of immunoglobulin subclass loci, the authors showed a preferential development of immunoglobulin A plasma cell clones in lung adenocarcinoma tumors; and an analysis of somatic hypermutation frequency showed increased affinity maturation in lung B cells with increased proximity to the tumor, but this was impaired in epidermal growth factor receptor (EGFR)–mutant lung adenocarcinoma. Multiregion sampling of scTCR/BCR sequences has also been used to infer dynamic relationships between B- or T-cell clones spanning tissue compartments, including the primary tumor, metastases, LNs and circulation.15,57

Recent efforts to develop computational frameworks that integrate transcriptomics with BCR or TCR sequence–encoded data will further support efforts to decode lymphocyte dynamics in tumors. These include tools such as sciCSR, which improves the accuracy of B-cell trajectory inference using class-switch recombination dynamics,62 and Benisse, which uses B-cell gene expression to inform BCR-seq analyses, thereby improving the functional interpretation of BCR repertoires.63 Moreover, spatial adaptive immune receptor repertoire profiling has offered further locational context to the temporal dynamics of intratumoral lymphocytes.64,65

**Time stamping a cell’s immune response through activity-induced reporter systems**

How immune cells respond in the tumor microenvironment is highly dependent on the ligation of surface receptors and the accumulated signals this conveys. For T cells, it is evident that both the TCR and the array of costimulatory and coinhibitory receptors expressed at the cell surface determine cell fate and function. Of particular relevance to the optimization of immunotherapy combinations is determining where these signals are received, and their impact on responding T cells. However, dissecting this in vivo is challenging. Responding T cells will not be perfectly synchronized, as can be more easily modeled in vitro, impeding efforts to map out how interventions really alter T-cell fate and function. To this end, changes in gene expression can be visualized using transcription-activated fluorescent reporter mice, including for genes expressed specifically following TCR stimulation.66 However, because the half-life of most fluorescent proteins is several days, the expression does not directly demonstrate current or even very recent signaling events.

To more accurately capture responding T cells, the Ono laboratory pioneered the development of transgenic mice exploiting a fluorescent “Timer” protein.67 This initially forms a short-lived blue fluorescence–emitting version, before maturing into a red form that then slowly breaks down. Nr4a3-Tocky transgenic mice contain the Timer protein inserted into the Nr4a3 gene, which is immediately induced by TCR signaling.68 By comparing
levels of the blue and red forms of the Timer protein, it is possible to visualize the spectrum of responding T cells, from those newly signaling through their TCR to those in which transcription driven by TCR signaling has ceased. This supports dynamic analyses of individual responding T cells. Furthermore, the nature of Timer protein expression can further inform the strength or duration of the signal. Recently, the Ono laboratory\textsuperscript{69} has used scRNA-seq analysis of B16 tumors implanted into Nr4a3-Tocky mice to map the dynamic changes in the transcriptome associated with TCR–antigen recognition in different intratumoral T-cell populations. These initial experiments highlight the possibilities for using scRNA-seq to interrogate transcriptomic changes in specific T-cell populations at defined times after TCR signaling. Finally, the authors showed that combined blockade of both cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed cell death protein 1 (PD-1) altered the balance of actively responding T cells within the tumor and observed a loss of the highly activated effector regulatory T cells compartment. Nr4a3-Tocky mice have also been used to investigate how anti–PD-1 treatment alters gene expression in responding CD4 T cells, an approach that identified a TCR response signature able to better stratify outcomes of patients with melanoma to anti–PD-1 therapy.\textsuperscript{70} Using Nr4a3-Tocky mice to identify responding T cells and to distinguish the nature of the response, such as new \textit{versus} persistent signaling, has significant potential for studying T-cell responses \textit{in vivo}, especially when combined with single-cell technologies. New transgenic models exploiting the Timer protein approach\textsuperscript{71} will enable further pathways downstream of immune receptor activation to be considered, and extend investigations into the control of leukocyte responses and the immediate impacts of different immunotherapies.

**Precise temporal labeling of cells in specific tissue compartments**

Models utilizing Timer protein expression may support studies of leukocyte response dynamics, but are currently limited to the investigation of T cells within days following TCR activation. Tracking TCR or BCR sequences can establish the dissemination of lymphocyte clones, but such analyses often lack directionality or evidence for the site of cellular changes. To overcome these limitations, tissue- or location-specific labeling of cells \textit{in vivo} provides an exciting approach to help resolve the source or fate of immune populations. Photoactivatable fluorescent proteins that undergo pronounced spectral changes (from green to red fluorescence) upon exposure to ultraviolet–violet light have been described for over 20 years\textsuperscript{72,73} (Figure 1a). While often used for developmental studies, the creation of transgenic mice expressing these proteins, such as the Kaede and KikGR models, enabled site-specific labeling of immune cells at tissues including LNs, skin and colon and tracking of their subsequent migration to other locations.\textsuperscript{74–81}

Beyond tracking the dissemination of cells from the site of labeling, a more refined use of photo-labeling is to time stamp the cellular compartment of tissues.\textsuperscript{77} This requires that essentially all of the cells within the target tissue (e.g. tumors) are photo-labeled at a given time through careful optimization of the photoconversion process. Although challenging for some tissues, if achieved, it enables cells entering after photoconversion to be distinguished from those cells present at the time of labeling, visualized by the absence or presence of the photoconverted (red) label, respectively (Figure 1b). This means that the nature of cells that have only recently entered the tissue can be captured and compared with longer-term residents. Furthermore, longitudinal studies assessing the labeled tissue at different times after photoconversion can resolve how infiltrating cells change over time (Figure 1c).

As previously discussed, the major limitation of \textit{in silico} tools to computationally infer cellular trajectories in scRNA-seq data is the absence of empirical measurements that represent “real-time.” Hence, the application of tissue-labeling models combined with scRNA-seq experiments is a powerful method to directly address this and leverage the throughput of single-cell “omics” technologies to study highly dynamic systems, such as the cancer immune response. For computational approaches where priors are required, for example, choosing the root or end points for pseudotime analyses, cells with the greatest fluorescent signals relating to recent infiltration (i.e. Kaede-green) or signals relating to the longest duration of tumor residence (i.e. Kaede-red) can be used to define these inputs.\textsuperscript{82} In the absence of these empirical measurements, the definition of such priors is biased by user assumptions. The presence of dysfunctional immune cells within tumors is well established. Accurate tracking of cell fate within cancer models has the potential to pinpoint how these populations emerge, which in turn can support the design of interventions to prevent or overcome such changes. Our initial studies of photo-labeling tumors utilized these approaches to clarify how rapidly T cells established an exhausted program after entry into the tumor.\textsuperscript{82} For approaches that do not rely on user-defined priors, consistency between orthogonal computational outputs and empirically observed experimental data strongly increases confidence in the resulting conclusions.
Figure 1. Temporal labeling of specific tissue compartments. (a) Illustration of the shift in the emission spectrum of Kaede protein upon violet light exposure. (b) Cartoon illustrating the dynamic labeling of tumors to track how immune cells change over time, alongside egress to peripheral tissues. (c) Illustration of how the proportion of Kaede-green and Kaede-red cells changes over time after labeling and how these data can be interpreted. (d) Cartoon summarizing the temporal labeling of cells within the circulation pioneered in the recent “zman-seq” methodology. Images were generated using BioRender. APC, allophycocyanin; FITC, fluorescein isothiocyanate; IV, intravenous.
More recently, we have used time stamping of the tumor immune compartment to consider how other immune populations respond to entering the tumor microenvironment. For example, using the relative Kaede-green/red signals in scRNA-seq of tumor NK cells, we found that CD11b and CD49a expression defined how long these cells had been present in the tumor. Circulating NK cells transcriptionally resembled CD11b+ NK cells, which were predominantly Kaede-green, pointing to the source of continual NK cell influx. Particularly striking within these data was the speed at which NK cells become dysfunctional within the tumor, a process that our analyses indicate can be achieved in only 24 h. Hence, site-specific labeling of cells enables the identification of time-associated changes in tumor immune cells with greater sensitivity and accuracy.

Notably, the rapid loss of NK function identified within our photo-labeling analyses was similarly identified in a recent study that combined labeling of circulatory cells with scRNA-seq, the so-called zman-seq, in a murine model of glioblastoma. Here, intravenous injection of anti-CD45 antibodies was used to specifically label the circulating immune cells, with an injection of different fluorescently tagged antibodies at regular time intervals (e.g. every 12 h) used to label cohorts of circulating cells (Figure 1d). The injection of intravenous antibodies minutes before culling mice for tissue isolation has been a key approach to distinguishing cells in the vasculature and assessing tissue residency for many years. The “zman-seq” approach builds upon this methodology to effectively time stamp cells within the circulation, supporting the subsequent tracking of what happens to cells after entry into tissues and how quickly cells adapt to the tumor environment. The authors analyzed the ratio of fluorescent labels injected at different time points within each cellular neighborhood to assign a statistic relating to the duration of “tumor exposure,” which was then used to define a trajectory connecting cell states.

Tracking photo-converted cells beyond the photo-labeled tissue further enables transcriptomic interrogation of leukocyte dynamics across organs. For example, following photo-labeling of the tumor, scRNA-seq of Kaede-red tumor dendritic cell (DC) emigrants in the draining LNs allowed us to directly compare migrated LN DCs with “unsuccessful” DC emigrants in the tumor. These data indicated that some activated DCs become trapped within the tumor microenvironment, subverted by further signals that potentially impede the antitumor T-cell response. Expanding these analyses to studying immune cell migration across other cancer compartments, including metastases, is an exciting future area of research.

There are limitations to temporal labeling approaches. Integral to the success of “zman-seq” is the labeling of cells only within the blood, with no leaching of the antibody into the surrounding tissue. However, analysis of other tissues remains important to confirm that no labeling of cells beyond the intended tissue occurs to support the wider use of this methodology. Labeling studies, whether with light or injected antibodies, typically only allow reliable marking of cells for short windows of time, impeding assessment of genuine tissue residency. While modification of protein turnover can prolong the half-life of the fluorescent protein, for example, through histone-tagging, proliferation-induced dilution of the fluorescent label can further reduce the time over which all cells can be tracked. An obvious further point is that photo-labeling of some tissues may be technically challenging. There are clear constraints regarding the depth of light penetration of tissue and the efficiency of photoconversion. Thus, tumors beyond a certain size, in our hands approximately 8 mm in diameter, cannot be fully labeled. Tissues with a high red blood cell compartment (e.g. spleen) or pigmentation (e.g. B16 melanoma models) are resistant to labeling. Finally, care must be taken not to damage the tissue (e.g. skin) covering the site of labeling. In our experience, optimizing photo-labeling of a given tissue or site requires a significant investment of time and careful assessment.

**Future opportunities for dynamic analyses of immune responses**

The combination of new experimental models to support dynamic studies, advances in single-cell technologies and the refinement of in silico analyses makes this a hugely exciting time to study antitumor immunity. We conclude this review by offering some opinions on how such investigations might evolve in the future.

**Refine existing models**

Bringing together the experimental approaches outlined above can generate new insights into the antitumor response. For example, crossing photoconvertible and Tocky mouse strains would help visualize T cells responding to a specific immunotherapy alongside information on how long such cells had been within the tumor. The temporal labeling of cells in the circulation, through rounds of anti-CD45 antibody injection, could be performed in conjunction with tumor photo-labeling to pinpoint which cells recruited into the tumor were present in the blood versus present in peripheral lymphoid tissues (and shielded from antibody labeling).
However, to really improve on the current photo-labeling methodology, time stamping of cells needs to extend beyond a few days and should support long-term tracking of cell fate for weeks, enabling studies using slower-growing or more mature tumors that better model human cancers. Genetic-tagging, the so-called fate-mapping, which exploits inducible cre recombinase expression to perpetually mark cells and their progeny, is usually achieved through systemic tamoxifen administration to enable cre recombinase activity. Site-specific fate mapping of cells would solve the transient labeling afforded by photoconvertible mice, although progeny cells would also be marked that may not have been in the original location. Local administration of tamoxifen can support site-specific labeling, although nagging concerns with dissemination of the tamoxifen will remain. Combining photo-labeling with cre recombinase activity is an exciting potential approach to site-specifically label cells, and photoactivatable cre recombinase models have been described. More efficient in terms of applicability is the potential use of photoactivatable tamoxifen-like molecules where ultraviolet–violet light “de-cages” the molecules to then drive gene expression.

Bring new approaches to cancer immunology

The importance of resolving temporal dynamics within single-cell studies is pushing the development of new experimental approaches that can help better order cellular states and intercellular heterogeneity. To our knowledge, the following approaches have not yet been applied to study anticancer responses in vivo, but offer clear potential.

Metabolic RNA labeling utilizes the exposure of cells to artificial nucleosides, which are taken up as cell-permeable metabolic precursors, to label nascent RNA. Thus, newly synthesized RNA containing the tagged nucleotides can be distinguished from “older” RNA generated before exposure to the artificial nucleosides. Building on bulk RNA studies, approaches combining metabolic labeling with single-cell analyses have been developed to support a more robust ordering of cellular states. For example, Sci-fate used incubation with 4-thiouridine (a thymidine analog) to assess the dynamics of gene expression in thousands of single cultured cells using scRNA-seq. Using scEU-seq, cultured cells were exposed to 5-ethynyl-uridine to estimate RNA transcription and degradation rates in single cells. Most analyses so far have focused on short-term labeling of in vitro cultured cells, including organoids, with toxicity of the modified nucleosides a key concern. Of note, SLAM-ITseq, which labels RNA with 4-thiouracil, was performed using transgenic mice. While the analysis performed was bulk, rather than scRNA-seq, it seems likely that single-cell analyses are feasible in the future.

Genetic barcoding enables detailed lineage tracing in a cell-ambiguous manner. By utilizing Cas9 to introduce insertions and deletions of varying lengths that “scar” the genome, unique and heritable cellular barcodes can be introduced to support lineage tracing in vivo in single cells. Pioneered in zebrafish, LINNAEUS, ScarTrace and scGESTALT are all approaches that enable the profiling of lineages at the individual cell level using scRNA-seq. Bowling et al. moved this approach beyond zebrafish, developing the CARLIN (CRISPR array repair lineage tracing) mouse. Here, 10 single guide RNAs cut target sites in a 276-bp array, when Cas9 is expressed. Repair of these breaks in the target array results in a range of altered DNA sequences that are stably inherited. Through doxycycline-dependent Cas9 expression, this approach can be used to introduce tractable genetic scars in any tissue at any stage of development, which can then be identified from scRNA-seq to track lineage trajectories and interrogate gene expression. Applying such approaches to study the tumor microenvironment would help unlock cellular fate across the immune cell landscape.

A key aim of tissue labeling approaches is to study the environments within which cells may encounter each other. While cellular associations can be identified using immunofluorescence or spatial transcriptomic approaches, there are limits to how far these data can be interpreted. Actually demonstrating physical cellular interactions and what this means for the cells involved remain hugely challenging. To interrogate leukocyte interactions, Victoria and colleagues recently developed “uLIPSTIC” (universal labeling immune partnerships by SorTagging intercellular contacts), an exciting model system where a peptide substrate tag is directly transferred between closely interacting cells. Here, new mouse models enabled ubiquitous expression of the labeling machinery initially used to study T cell–DC interactions, and therefore, support the study of cellular interactions in vivo, regardless of the molecules driving this. scRNA-seq then provided detailed insights into the nature of the cells that have recently interacted. Combining the uLIPSTIC methodology with tissue-labeling approaches and dynamic trajectory inference will provide a powerful opportunity to resolve how cellular interactions alter the differentiation and fate of tumor-infiltrating leukocytes.

Informing studies of human cancer

The various animal model systems described provide detailed insights into tumor-infiltrating leukocyte
dynamics, and enable the precise identification of cell-state transitions. However, the relevance of findings from animal models to human disease has been a major obstacle in the history of immunology research. The identification of analogous cell populations across species or cancer types has typically relied on the expression of a handful of surface proteins that do not reflect the complex cancer immune landscape. Detailed and high-dimensional scRNA-seq data now provide a critical framework on which to map tumor-infiltrating leukocytes. Large-scale efforts to unify immune cell states between murine cancer models,\textsuperscript{103} between human samples across cancer types\textsuperscript{36,104,105} and importantly, between murine and human tumors\textsuperscript{106,107} using scRNA-seq data have already begun, leveraging thousands of measurements of transcriptomic features per cell. Crucially, these studies will identify parallel differentiation processes, and their regulation, in tumors across species and contexts, thereby accelerating the translation of discoveries from preclinical models to the clinic.

Advances in computational frameworks

New algorithms for \textit{in silico} trajectory inference are being continually developed, often leveraging computational methodologies from other academic disciplines. While each approach undoubtedly possesses merits, there remain challenges in selecting the “best” approach, or more importantly, evaluating the biological accuracy of the resulting output. For the former, benchmarking studies\textsuperscript{13} are helpful, but biological processes are diverse and tools that work well in a particular system may perform poorly elsewhere. To accurately decode the dynamics of tumor-infiltrating leukocytes, we believe that computational tools to infer dynamical processes are best complemented by experimental models that provide empirical data relating to the chronology of the biological process or system being investigated. Therefore, computational approaches capable of incorporating multiple experimental priors, in various modalities, will provide a powerful framework to study temporal dynamics at maximum resolution. The updated \textit{CellRank} framework\textsuperscript{35,108} facilitates the combination of complementary approaches to study biological processes, including pseudotime, RNA velocity, experimental time points, metabolic labeling and more, to inform a unified temporal trajectory. Bayesian Gaussian process latent variable model approaches also enable pseudotime estimates to be informed by experimental time data.\textsuperscript{109,110} However, cross-tissue non-linear cellular dynamics introduce additional complexity that cannot be easily incorporated into present models. Finally, single-cell technologies have the potential to uncover convergent gene programs across disease/tumor types and treatment regimens, but the scalability of softwares to increasingly large volumes of data is key to this.

With all the experimental models discussed above, there are many exciting opportunities to generate single-cell data to unravel tumor leukocyte dynamics. Our ability to maximize the utility of these high-dimensional data will rely on collections of stable, robust and user-friendly informatics tools accessible to the community, for example, \textit{scverse}, which aims to unify a computational ecosystem for the analysis of “omics” data.\textsuperscript{111}

ACKNOWLEDGMENTS

DRW is supported by a Wellcome Discovery Award (227890/Z/23/Z). MRC is supported by a Wellcome Trust Investigator Award (220268/Z/20/Z), Wellcome Discovery Award (227890/Z/23/Z) and by the NIHR Cambridge Biomedical Research Centre. CYCL is supported by the Gates Cambridge scholarship.

AUTHOR CONTRIBUTIONS

\textbf{David R Withers:} Conceptualization; writing – original draft; writing – review and editing. \textbf{Menna R Clatworthy:} Conceptualization; writing – original draft; writing – review and editing. \textbf{Colin YC Lee:} Conceptualization; writing – original draft; writing – review and editing.

CONFLICT OF INTEREST

The authors declare no competing interests.

REFERENCES

Decoding immune cell dynamics

CYC Lee et al.


38. Giles JR, Globig AM, Kaech SM, Wherry EJ. CD8+ T cell replacement of tumor-specific T cells following PD-1 replacement of tumor-specific T cells following PD-1.


