Received: 25 March 2022 | Revised: 2 June 2022 | Accepted: 24 June 2022

DOI: 10.1002/cyto.a.24668

ORIGINAL ARTICLE

MIFlowCyt | ISAC | CYTOMETRY Journal of Quantitative Cell Science PART A

TrackSOM: Mapping immune response dynamics through clustering of time-course cytometry data

Givanna H. Putri[1,2] | Jonathan Chung[3,4] | Davis N. Edwards[3,5] | Felix Marsh-Wakefield[3,5,6,7] | Irena Koprinska[3] | Suat Dervish[5] | Nicholas J. C. King[3,5,8,9] | Thomas M. Ashhurst[3,6,8,9] | Mark N. Read[2,3,4]

[1]School of Computer Science, The University of Sydney, Sydney, New South Wales, Australia

[2]Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia

[3]The Westmead Initiative, The University of Sydney, Sydney, New South Wales, Australia

[4]Viral Immunopathology Laboratory, Discipline of Pathology, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

[5]School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

[6]Vascular Immunology Unit, Department of Pathology, The University of Sydney, Sydney, New South Wales, Australia

[7]Sydney Cytometry Core Research Facility, The University of Sydney and Centenary Institute, Sydney, New South Wales, Australia

[8]Sydney Institute for Infectious Diseases, The University of Sydney, Sydney, New South Wales, Australia

[9]Sydney Nano, The University of Sydney, Sydney, New South Wales, Australia

Correspondence
Givanna H. Putri, School of Computer Science University of Sydney, Building J12/1 Cleveland St, Camperdown, NSW 2006, Australia.
Email: givanna.haryonoputri@sydney.edu.au

Mark N. Read, School of Computer Science University of Sydney, Building J12/1 Cleveland St, Camperdown, NSW 2006, Australia.
Email: mark.read@sydney.edu.au

Funding information
International Society for the Advancement of Cytometry (ISAC) Marylou Ingram Scholars Program; Merridew Foundation; University of Sydney's Westmead Initiative; Australian Government Research Training Program Scholarship

Abstract

Mapping the dynamics of immune cell populations over time or disease-course is key to understanding immunopathogenesis and devising putative interventions. We present TrackSOM, a novel method for delineating cellular populations and tracking their development over a time- or disease-course cytometry datasets. We demonstrate TrackSOM-enabled elucidation of the immune response to West Nile Virus infection in mice, uncovering heterogeneous subpopulations of immune cells and relating their functional evolution to disease severity. TrackSOM is easy to use, encompasses few parameters, is quick to execute, and enables an integrative and dynamic overview of the immune system kinetics that underlie disease progression and/or resolution.

KEYWORDS
bioinformatics, cytometry, immune response dynamics, immunokinetics, immunology, sequential clustering, single-cell analysis, West Nile virus

1 | INTRODUCTION

The immune response is a dynamic process in time, and inter-individual differences therein often underlie diverging clinical outcomes. Under challenge, the immune response deviates from homeostasis with the expansion of numerous cellular phenotypes that enact challenge-specific effector functions in a temporally-coordinated fashion. This is seen, for instance, with accelerated cellular differentiation in the bone marrow and the infiltration of immune cell populations into peripheral tissues. Yet, the immune response is fallible, as attested to by autoimmunity, chronic diseases, and lethal infections, all of which can be mediated by the immune response and may be ameliorated by immune intervention. Greater understanding of the temporal kinetics and interactions of immune cell populations would

Thomas M. Ashhurst and Mark N. Read are Co-senior authors.

present expanded opportunities to intervene for clinical benefit. Essential to delivering this vision is an explicit mapping of how immune system status varies with time ("*temporal mapping*") and how this in turn is associated with disease severity.

Cytometry is a popular and widely accessible technology for phenotypically characterizing and quantifying individual immune cells from a body fluid sample (e.g., blood). Taking samples progressively during disease development presents an opportunity for developing detailed mappings of immune response against disease-stage. Cytometry quantifies numerous genetically determined characteristics ("markers") on potentially millions of individual cells. Advancements in this technology have now enabled measurements of upwards of 45 markers per cell, and this will continue to grow [1–4]. A key analytical challenge in cytometry is aggregating ("gating") these many and heterogeneous cells into their distinct cellular phenotypes. Given the high-dimensional nature of such data, the task is challenging. It requires considerable domain expertise and is prone to subjectivity and poor reproducibility [5–7]. Many automated clustering algorithms have emerged to support gating [8–11]. Of these, FlowSOM has gained popularity, being simple to use, fast to execute, and highly congruent with manual gating efforts [12, 13].

Lacking, however, is algorithmic support for annotating the immune response dynamics with respect to time and disease-stage. This requires both (i) clustering to identify populations in each dataset in the time/disease-stage sequence, (ii) tracking how these populations evolve, and (iii) visualizing them. This can be complex as both the absolute number of cells for each phenotype, and their relative representation within the sample, are liable to change. So too are population marker expression levels, which change as populations undergo differentiation, maturation, and associated functional alteration. Furthermore, differentiation, such as from hematopoietic stem cells through common progenitors into distinct effector phenotypes, mean that phenotype trajectories through marker space can present as branching tree-like structures through the time/disease-stage sequence. Additionally, infiltration into and egress out of peripheral tissues can manifest as populations appearing and disappearing from the blood compartment. Identifying the existence and timing of phenotypic subpopulations that can impact on disease status is also highly relevant [14]. Required is a unified clustering-tracking-and-visualization algorithmic framework that can deliver these insights. Filling this knowledge gap, we previously developed ChronoClust [15], a clustering and tracking algorithm wherein the data themselves determine the number and shape (in marker space) of immune cell populations. While ChronoClust's clustering behavior is highly customizable, it is computationally cumbersome to execute and finding the appropriate parameters can prove challenging [12]. Critically, it lacks the crucial visualization techniques for interpreting the dynamics of such cell populations.

Here, we propose a new method, TrackSOM, for analyzing the immune response dynamics from time-course or disease-stage cytometry data. TrackSOM combines the high-quality clustering capacity and fast run time of FlowSOM with an improved cluster tracking capability adapted from ChronoClust to map immune cell population dynamics against time and/or disease severity status. Importantly, TrackSOM offers advanced visualization techniques which aid the visualization and interpretation of resulting immune cell populations' evolutions. We highlight these capabilities on a synthetic dataset and by recapitulating previous findings on a bone marrow dataset from West Nile Virus (WNV)-infected mice. We perform a parameter sensitivity analysis and demonstrate TrackSOM to have both an improved clustering performance and lower sensitivity to parameter value selections over ChronoClust. We offer usage and parameter selection advice to potential users with novel data. Lastly, we demonstrate usage on novel data by characterizing the evolving immune response in the brains of WNV-infected mice and relating this to disease severity.

2 | METHODS AND MATERIALS

2.1 | The TrackSOM algorithm

TrackSOM is a clustering-tracking-and-visualization algorithmic framework for clustering time/disease-course cytometry data into cellular populations, and for tracking and visualizing how these populations evolve over time/disease-stage. It takes as input a sequence of datasets (in either FCS or CSV file formats), where each dataset captures cytometric quantifications of heterogeneous cells, and the sequence represents time-points or disease-stages, Figure 1A. Critically, TrackSOM assumed each dataset has been appropriately preprocessed, that is, devoid of debris, dead cells, and doublets, and either transformed using logicle [16] or arc-sinh transformation, or linearly binned. TrackSOM determines clusters, representing phenotypes, within each dataset and then establishes links between them to reveal temporal evolutions between clusters across the sequence. Datasets from all time-points or disease-stages are first amalgamated and normalized using the *z*-score normalization. A Self-Organizing Map (SOM) is then built upon them, thus identifying the data's broad topological shape across the entire sequence, Figure 1B. Each data-point is uniquely associated with a single SOM node and the data that any given SOM node represents may be confined to either a single or several time-point(s) or disease-stage(s). TrackSOM's ability to track how clusters evolve is based on variation in which SOM nodes are supported by data-points as the dataset sequence progresses (Figure 1C) - some nodes will become more relevant, others less so.

Clusterings (phenotypes) of data-points at each time-point are captured through *meta-clusters*, which represent the collation of SOM nodes into discrete groupings (Figure 1C). Having built the SOM, TrackSOM then processes each dataset (time-point/disease-stage) independently, isolating SOM nodes containing data from that dataset only and forming meta-clusters accordingly. The centroids of these non-empty SOM nodes are re-calculated (temporarily disregarding data from other datasets) as per FlowSOM operation. Meta-clusters for the given dataset are derived through consensus as drawn from multiple efforts at sub-sampling and hierarchically clustering the SOM's nodes, again in accordance with FlowSOM. The number of

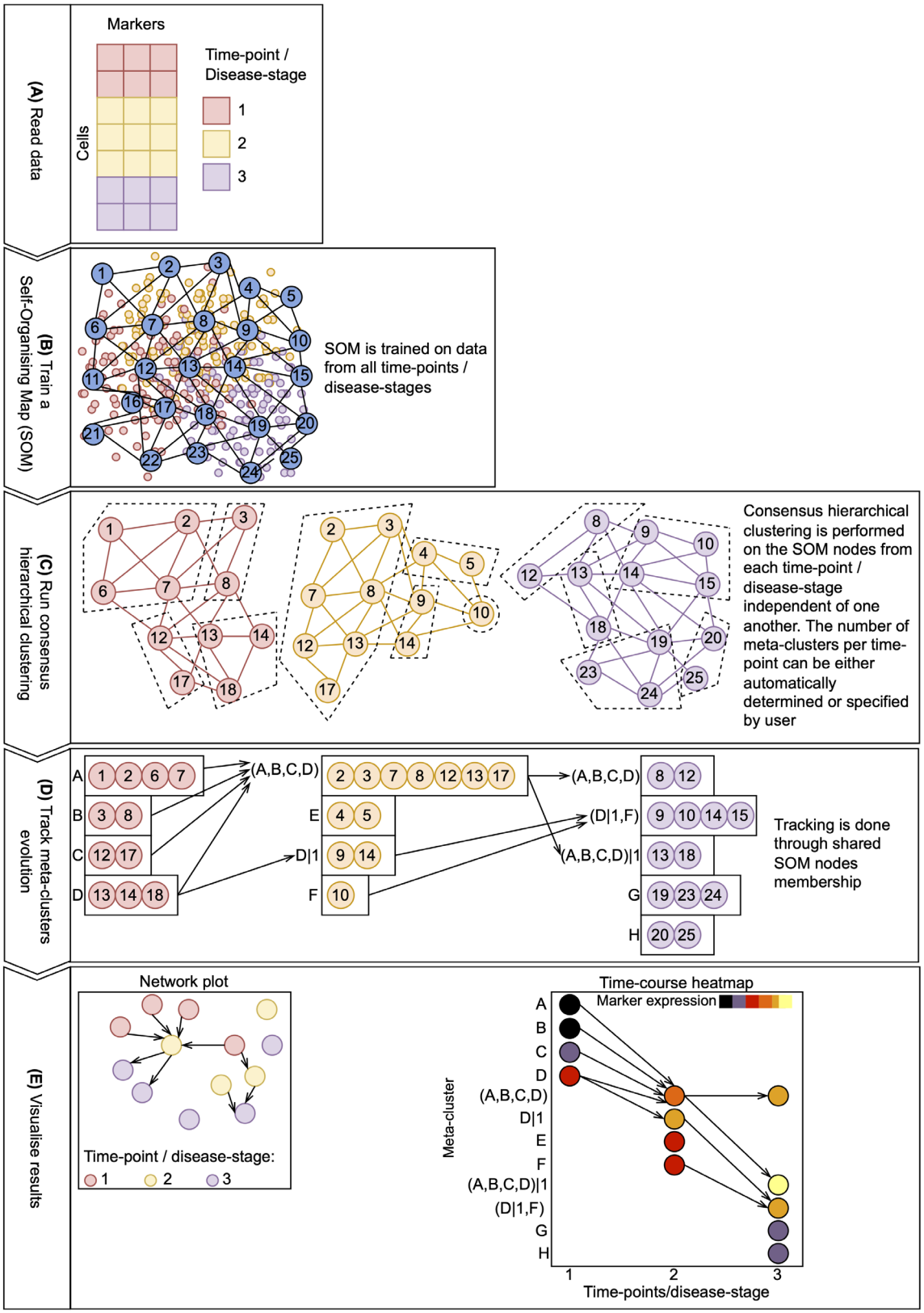

FIGURE 1 Schematic overview of TrackSOM's operation and usage. (A) Time-course or disease-course cytometry data is parsed by TrackSOM, resulting in a matrix of cells (rows) and marker values (columns) per time-point/disease-stage. (B) A Self-Organizing Map (SOM) is then trained on all the matrices. Numbered circles represent SOM nodes. (C) SOM nodes containing data from each given time-point/disease-stage are then isolated and "meta-clustered" (dotted polygons) using consensus hierarchical clustering. (D) The temporal evolutions of the resulting meta-clusters are tracked through their shared SOM node memberships. (E) Meta-cluster evolutions can be visualized as network plots (left) or time-course heatmaps (right). For clarity, the figure refers predominantly to "time-points", but these could instead be datasets in a series of escalating disease severities. We have shown only 3 time-points, but TrackSOM is not limited to this. [Color figure can be viewed at wileyonlinelibrary.com]

meta-clusters formed is parameter-defined. Under the Autonomous Adaptive option, the number of meta-clusters created at each time-point is automatically inferred using FlowSOM's elbow criterion which infers the optimal number of meta-clusters based on the point at which the rate of change in the variance captured in the meta-clusters suddenly decreases. Alternatively, users can specify the number of meta-clusters to be created as either a single number for all time-points/disease-stages ("*Prescribed Invariant*" operation) or a number for each time-point/disease-stage individually ("*Prescribed Variant*"). For these options, the specified number(s) of meta-clusters is passed into FlowSOM's consensus hierarchical clustering function for each time-point.

Next, meta-cluster linkages across adjacent time-points/disease-stages are determined to track meta-cluster evolutions. This is accomplished by identifying those meta-clusters in adjacent datasets that share common SOM nodes (Figure 1D). Briefly, meta-cluster evolutions are encoded into the meta-cluster IDs. New meta-clusters are defined as containing exclusively SOM nodes that did not belong to any meta-cluster(s) in the preceding datasets. This includes the first time-point/disease-stage. New meta-clusters are assigned unused alphabetic labels from an alphabetical sequence (A, B, …, Z, AA, BB, …). Meta-cluster evolutions are captured in how these labels are amended over time. A meta-cluster that contains only a subset of a preceding meta-cluster's SOM nodes is termed a split meta-cluster. A meta-cluster can simultaneously split multiple ways, and the child inheriting the majority of SOM nodes retains the parent's label. The remaining sibling meta-clusters are assigned A|1, A|2, and so forth. A meta-cluster which captured SOM nodes belonging to multiple preceding meta-clusters can be considered as either a merged meta-cluster and assigned a label that is the parenthesised concatenation of its parent IDs, for example, (A,B,C,D), or not. For the latter option, SOM nodes that belong to the same parent meta-cluster are assigned to a split meta-cluster (of the parent meta-cluster) while those which do not exist in previous time-points form new meta-clusters. Importantly, the ability to disable meta-cluster merging is unique to TrackSOM and better reflects the diverging nature of cellular differentiation. It can, however, inflate the number of meta-clusters above user-specified values.

TrackSOM provides several visualizations, Figure 1E. *Network plots* denote all meta-clusters at all time-points/disease-stages as bespoke nodes, linked via their evolutions. *Time-course heatmaps* capture time/disease-stage and meta-cluster IDs on separate axes. The visualizations are complementary: the network plot provides a fundamental global overview of the data while the time-course heatmap confers additional details for each meta-cluster for subsequent investigations. In both visualizations, node sizes indicate the proportion of a dataset's data-points that the corresponding meta-cluster captures. Meta-clusters can be colored by average marker expression level to depict movements through marker-space as the sequence progresses, by time-point/disease-stage, or originating meta-cluster for traceability of lineages.

For each cell, at every time-point, TrackSOM denotes which SOM node and meta-cluster it belongs to. This output coupled with TrackSOM's visualization techniques enable users to associate meta-clusters with cellular phenotypes and subsequently track how each meta-cluster evolved through the dataset sequence. Network plots are arranged through force-directed layouts. The coloring of meta-cluster nodes by marker expression level on plots represents raw values mapped to the color scales for each marker independently.

Notably, TrackSOM inherits FlowSOM's parameters and their respective default values, though these default values can be changed to suit the user's needs. A full list of adjustable parameters is available in Supplementary Tables S1 and S2.

2.2 | Datasets

We evaluated TrackSOM on a 3-dimensional (3D) synthetic dataset adapted from [15], and two flow cytometry datasets.

The synthetic dataset contains three distinct conglomerates (sprouting, splitting, and transient) of data points that evolve over five time-points (Supplementary Figure S1 and Supplementary Section S1). The dataset is manually-gated to provide "ground truth" labels of which data-points belong to the same population.

The first flow cytometry data (WNV bone marrow) is a previously published dataset [15], which quantifies leukocytes in the bone marrow of mice over 8 days of an infection time-course. Details on animal procedures are provided in Supplementary Section S2. The data comprise 14 markers: (1) FSC, (2) SSC, (3) Ly6C, (4) CD45, (5) CD48, (6) Ly6G, (7) CD117, (8) SCA-1, (9) CD11b, (10) CD11c, (11) B220, (12) CD115, (13) CD16/32, (14) CD3/CD19/NK1.1. Time-point 1 represents pre-infection (mock-infected), and subsequent time-points represent days post-infection. There are 190,000 data points (cells) per time-point. The dataset was manually-gated to assign each cell a label representing its cellular phenotype (Supplementary Figure S2). The following 16 cellular phenotypes were identified for each time-point, (1) Stem and progenitor cells, (2) B cells, (3) T and NK cells (indistinguishable given the markers in the flow panel), (4) Monoblasts, (5) Monocytes, (6) Eosinophils, (7) Neutrophils, (8) Plasmacytoid Dendritic cells (PDCs), with all 8 populations existing as "activated" or "un-activated" types based on the upregulation of SCA-1 expression [17–21].

The second flow cytometry data (WNV CNS) represents the immune cells in the brains of mock- and WNV-infected mice at day 7 post-infection. C57BL/6 female mice aged between 8 and 12 weeks were inoculated with PBS or a lethal dose of WNV (LD-100) and then euthanised for analysis 7 days later. Mice were stratified into 6 groups based on severity of symptoms at the time of sacrifice. As the WNV dose inoculated is lethal in 100% of the cohort, the differential disease severity observed between groups at day 7 represent mice progressing through disease at different stages. The corresponding cytometry datasets were organized into a sequence of increasing disease severities. Both the numbers of mice and the clinical scoring criteria are reported in Supplementary Table S3. Further details on animal procedures are provided in Supplementary Section S2. The following 18 markers were quantified: (1) FSC, (2) SSC, (3) Ly6C, (4) CD45, (5) CD62L, (6) CD4, (7) CD86, (8) CD11b, (9) B220, (10) Siglec-F, (11) I-A/I-E, (12) Ly6G, (13) CD8-alpha, (14) CD11c, (15) CD115, (16) CD80, (17) CD3e, (18) NK1.1.

2.3 | Algorithm parameter sweeps and evaluation

TrackSOM was evaluated on all three datasets and was compared against ChronoClust on the WNV bone marrow dataset.

For the datasets where ground truth labels were available (synthetic and WNV bone marrow datasets), the algorithms were executed over all available data-points in the dataset sequence, and evaluations were then performed on only those data-points (cells) for which ground-truth labels are available. For each time-point, each clustering solution was evaluated against the corresponding manually-gated ground-truth clustering using the Adjusted Rand Index (ARI) metric [22]. ARI values are obtained for each dataset in the sequence independently, and we then report the mean of these values. Details on the ARI metric are available in Supplementary Section S3.

To evaluate the quality of cluster evolution tracking, we adopted the tracking accuracy metric used in [15]. This metric computes the proportion of transitions which are valid (e.g., biologically plausible, such as un-activated B cells transitioning to activated B cells as characterized by upregulation of the marker SCA-1). Valid transitions for the synthetic and WNV bone marrow datasets are shown in Supplementary Figure S3. The metric requires each meta-cluster to be assigned a cell type (ground-truth) label. In our evaluation, each meta-cluster is assigned the majority label of the data-points it captured. In the event of a tie, alphabetical order of labels is used.

We broadly explore TrackSOM and ChronoClust parameter space when clustering the WNV bone marrow dataset, and TrackSOM for the synthetic dataset. For TrackSOM, we generated 100 sets of parameter values using Latin Hypercube sampling for each dataset [23, 23, 24]. For ChronoClust, we subsampled 100 results generated under Latin Hypercube from a prior publication using this dataset [25]. More details are available in Supplementary Section S5.

Latin Hypercube sampling generates parameter values as floating point numbers. However, because TrackSOM's SOM grid size and number of meta-clusters parameters are encoded as integers, we round the parameter samples to the nearest integer before passing them to TrackSOM. This rounding can resolve to identical points in parameter space, and we discard duplicates from the analysis. Nonsensical parameter samplings, which specify more meta-clusters than the number of non-empty nodes available for any given time-point, were also discarded.

The parameters and ranges of values explored are given in Supplementary Tables S2 and S4. The rationale for how parameter value ranges were chosen for each algorithm and dataset are provided in the Supplementary Sections S4 and S5.

2.4 | Statistical methods

Latin Hypercube sampling and Partial Rank Correlation Coefficients (PRCC) were performed using the *Spartan* R package version 3.0.2 [26]. ARI implementation was provided by the *mclust* R package version 5.4.7 [27]. The *FIt-SNE* [28] plots and scatter plots for the WNV CNS data are provided through the *Spectre* R package version 0.4.0 [29].

2.5 | Availability of data and materials

TrackSOM is implemented in R, and is freely available to download under the GPL-3.0 open source license from https://github.com/ghar1821/TrackSOM. Code to reproduce all analyzes and figures in the manuscript are also available at the following GitHub repository: https://github.com/ghar1821/TrackSOM-evaluations. All three datasets used here are freely available to download under the GPL-3.0 open source license from the open science framework portal: https://osf.io/8dvzu/.

3 | RESULTS

3.1 | TrackSOM tracks splitting, merging and transient clusters

We verified TrackSOM's capacity to uncover cell population emergence, disappearance, splitting, and merging in time-series data by applying it to the synthetic dataset into which these phenomena had been engineered. To comprehensively assess TrackSOM's performance, we broadly explored each of TrackSOM's three clustering options by generating 100 parameter value sets (Supplementary Table S2) for each option using the Latin Hypercube sampling, varying the number of meta-clusters produced each day and the SOM grid size. Meta-cluster merging was permitted.

TrackSOM reproduced manually-gated populations and was robustly resistant to parametric perturbation. Solutions from Prescribed Variant operation consistently obtained mean ARI scores of at least 0.7, Figure 2A,B. TrackSOM's other operational modes, Autonomous Adaptive and Prescribed Invariant, fared less well. They scored ~0.55 and 0.35–0.65 for ARI, respectively. Autonomous Adaptive operation was strikingly invariant to parameter value variation. Closer manual inspection of the clustered data (not shown) revealed it to consistently produce 3–4 meta-clusters per day and, consequently, a sub-optimal capture of the sprouting conglomerate as a single meta-cluster. That Prescribed Variant operation obtained higher ARI scores than Prescribed Invariant likely reflects a better capture of the varying numbers of populations across time (Supplementary Figure S1), to which Prescribed Variant can be adapted. Thus, overall, Prescribed Variant operation more accurately reflected both manually-gated populations and their temporal evolutions.

TrackSOM excelled in correctly capturing temporal cluster dynamics, Figure 2A,C. For each parameter value set explored, 83%–100% of the cluster evolutions uncovered reflected valid population transitions. These were not simply self-referential transitions of the same populations across time-points; a quarter of the valid transitions

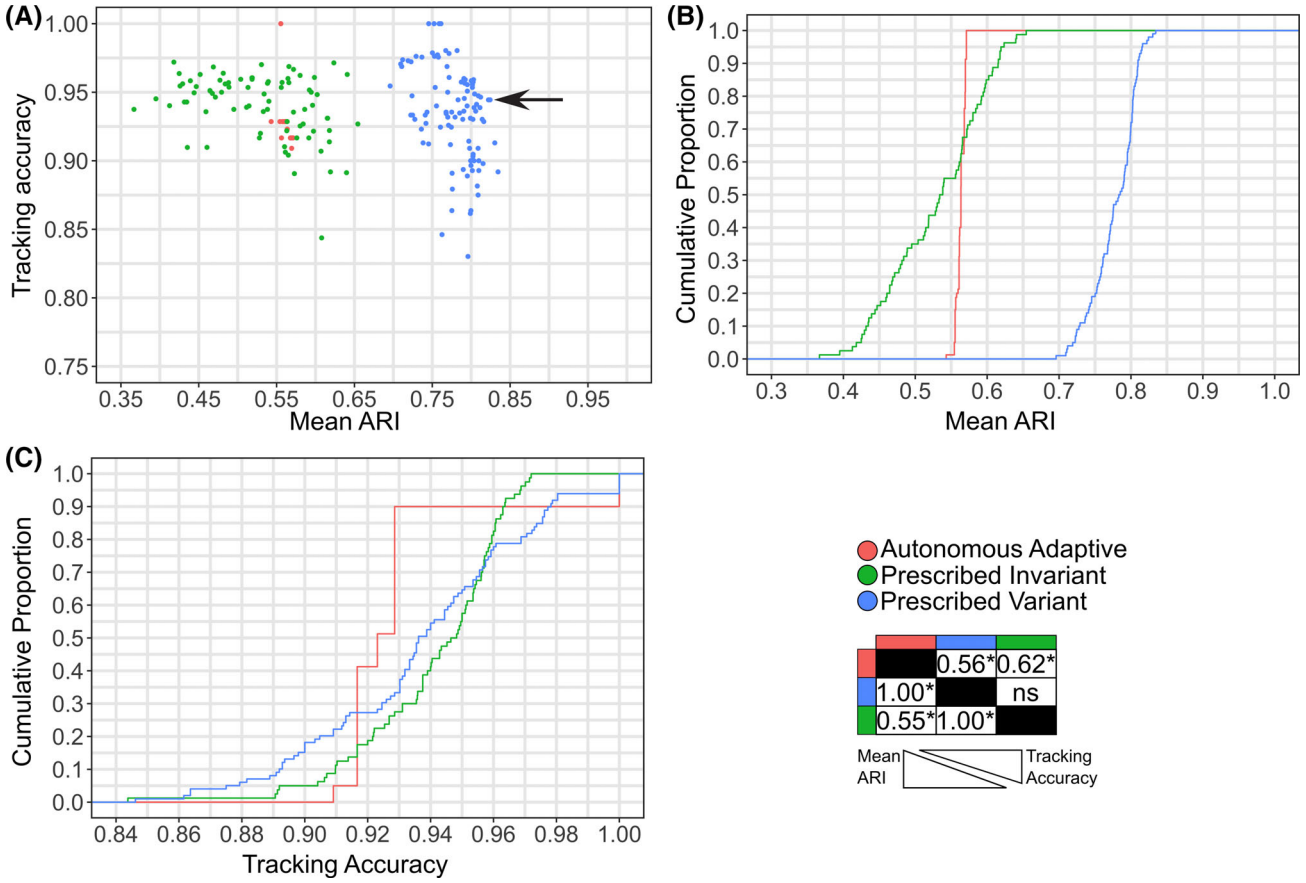

FIGURE 2 Evaluation of TrackSOM clustering solutions when applied to the synthetic dataset. We generated up to 100 unique parameter value combinations for each of TrackSOM's three modes of operation using Latin Hypercube sampling and then performed clustering on the synthetic dataset. Merging of meta-clusters was allowed. (A) A scatter plot of solutions in terms of mean clustering quality adjusted rand index (ARI) over the 5 time-points and the validity of the cluster temporal transitions generated (tracking accuracy). The solution explored further in the text is denoted by the arrow. (B-C) Cumulative distribution of TrackSOM's (B) mean ARI across time-points and (C) tracking accuracy. Statistical comparisons for (B-C) are done through the Kolmogorov-Smirnov test: *signifies *p*-value <0.005, ns denotes not significant. [Color figure can be viewed at wileyonlinelibrary.com]

were between distinct populations (e.g., meta-cluster (A,C), population II, splitting at day 2 to form meta-clusters representing populations III, IV, and V, Figure 3 and Supplementary Figures S1 and S3).

To demonstrate data exploration through TrackSOM visualizations, we focused on one solution that performed well in both metrics (Figure 2A, arrow). This solution was generated under Prescribed Variant, generating 3, 3, 11, 9, and 17 meta-clusters per day from an 8 × 8 SOM grid. Between days 2 and 3, population II "sprouts" into three different directions, each representing movements across X, Y or Z dimensions, Supplementary Figure S1. This divergence is clearly captured by TrackSOM and shown in Figure 3A and Supplementary Figure S4, where the branches stemming from meta-cluster (A,C) each move in only one dimension. Figure 3B,C allow meta-cluster lineages to be identified from their initial IDs, and can better visualize complex meta-cluster temporal transitions. For example, meta-cluster ((A,C)|4,(A,C)|3)|3 and ((A,C)|3|1,(A,C),(A,C|2|1),(A,C)|1)|2 in day 5 originates from meta-cluster A and C which has repeatedly merged and split across day 1–4. These complex temporal transitions most likely occurred due to over-clustering of the data and having permitted meta-cluster merging.

3.2 | Revealing the evolving immune response to West Nile Virus infection in mouse bone marrow

We next evaluated TrackSOM against the WNV bone marrow flow cytometry data. We executed each of TrackSOM's three modes of operation both with and without meta-cluster merging, resulting in six operational modes in total. As with the synthetic dataset, we generated up to 100 unique parameter value combinations for each operation.

All of TrackSOM's modes of operation were capable of generating high-quality clustering solutions. The majority of solutions scored over 0.7 for ARI, Figure 4A,B. Autonomous Adaptive operation was the least variable under parametric perturbation, but the ARI scores were suboptimal owing to its persistent under-clustering of the data: it typically generated eight meta-clusters despite manual gating defining 16 distinct cellular phenotypes (Supplementary Figure S2). Prescribed Variant and Prescribed Invariant operations together generated the best ARI scores but were more sensitive to parametric perturbation with meta-cluster merging disabled. This is because under no-merging operation the number of meta-clusters specified is a strict lower

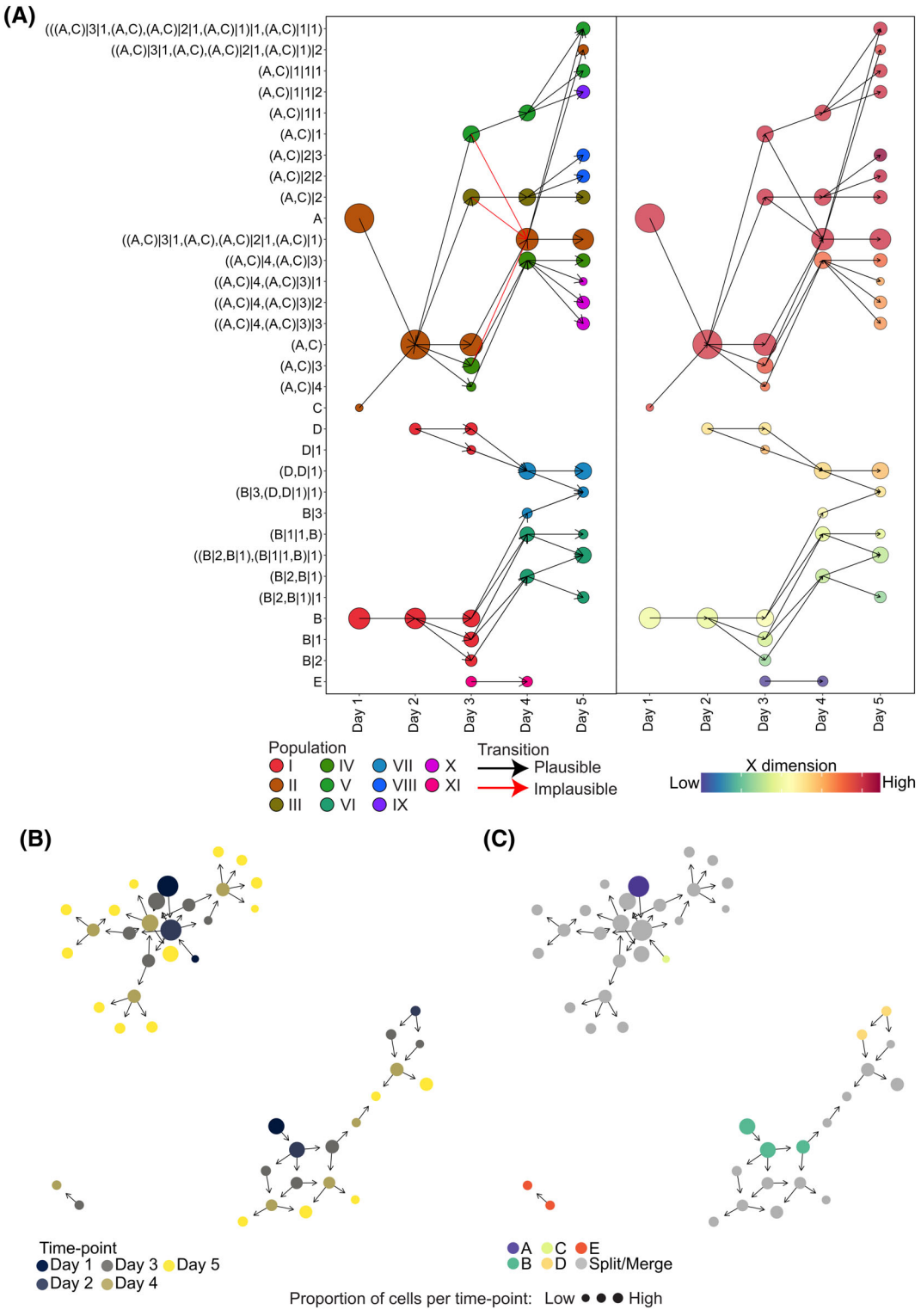

FIGURE 3 The TrackSOM's single high-performing solution for the synthetic dataset (denoted by the arrow in Figure 2) explored through several graphing formats. (A) Time-course heatmaps depict meta-clusters (dots), temporal evolutions (lines), relative meta-cluster sizes at each time-point (dot sizes), and either (left) the ground truth label it represents or (right) the meta-cluster locations and movements through the space (colors; X dimension shown. Y and Z dimension are available in Supplementary Figure S4). For A left, meta-clusters were assigned the ground truth label from which they captured the most data-points. Plausible and implausible transitions were determined based on the transition rules depicted in Supplementary Figure S3. (B-C) Network plots show meta-clusters colored by (B) time-point and (C) meta-cluster of origin (splits and merges are colored gray). Arrows indicate meta-cluster evolutions over consecutive time-points. [Color figure can be viewed at wileyonlinelibrary.com]

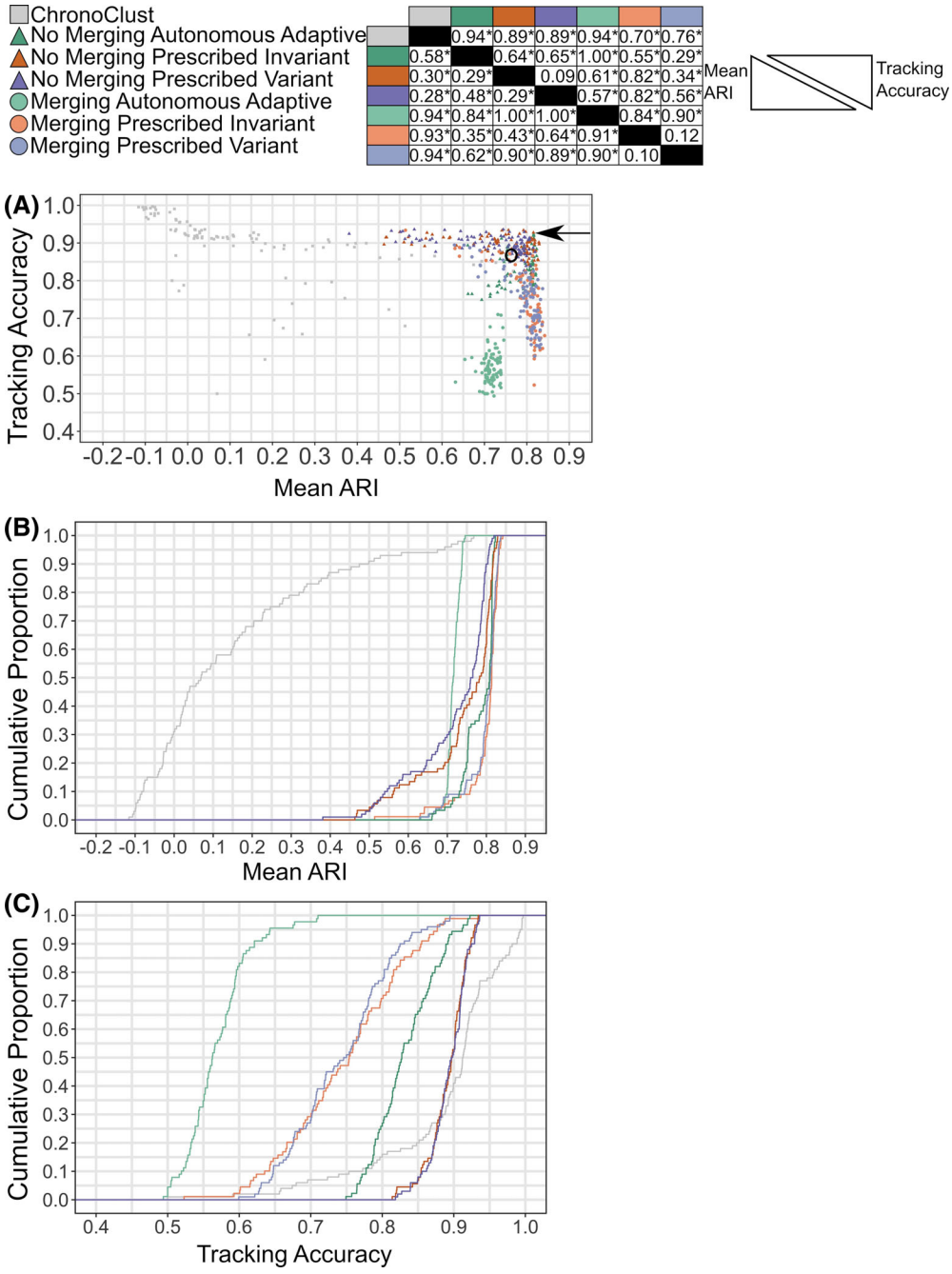

FIGURE 4 Evaluation of TrackSOM and ChronoClust clustering solutions when applied to the West Nile Virus bone marrow dataset. (A) Scatter plot of solutions in terms of mean clustering quality Adjusted Rand Index (ARI) over the 8 time-point dataset and the validity of the cluster temporal transitions generated (tracking accuracy). Solutions explored further in the text are denoted by the circle (ChronoClust) and the arrow (TrackSOM). (B-C) Cumulative distribution of TrackSOM's and ChronoClust's (B) Mean ARI and (C) Tracking accuracy. Statistical comparisons are done through the Kolmogorov-Smirnov test: * signifies *p*-value <0.005, ns denotes not significant. [Color figure can be viewed at wileyonlinelibrary.com]

bound and inappropriate parameter values can lead to an explosion in meta-cluster numbers, leading to reduced ARI scores. Conversely, with merging enabled, this parameter is not a strict lower bound and TrackSOM has greater capacity to correct for suboptimal parameter values.

TrackSOM performance in tracking evolutions of cell populations was more variable across its modes of operation, Figure 4A,C. Prescribed Variant and Prescribed Invariant were statistically indistinguishable and both outperformed Autonomous Adaptive, irrespective of whether merging was permitted. Interestingly, no-merging operation produced notably superior tracking accuracy scores than the same modes of operation when merging was permitted. This is because no-merging operation tended to generate more homogeneous meta-clusters (Supplementary Figure S5) and self-referential transitions are valid (Supplementary Figure S3). Conversely, merging operation has a greater capacity to amalgamate distinct cellular phenotypes and generate heterogeneous clusters, which can result in lower tracking accuracy scores.

Overall, the data revealed a possible trade-off between performance in clustering quality (mean ARI) and meta-cluster evolution tracking, as seen in the scatter plot of Figure 4A. The very best solutions by one metric are not necessarily optimal under the other. Notably, TrackSOM solutions at the leading edge of mean ARI scores (0.83) exhibited a wide range of tracking accuracy scores (0.6–0.9).

Conversely, solutions judged high quality by tracking accuracy (>0.9) varied widely in mean ARI (0.38–0.83). The latter can be explained by no-merging operation over-clustering the data and thus creating many self-referential transitions which are valid across time-points.

We selected one no-merging Prescribed Invariant solution that performed particularly well in both metrics for closer examination (arrow in Figure 4A). Figures 5A,B illustrate the solution as time-series heatmaps and network plots respectively. TrackSOM was able to track the progressive upregulation of SCA-1 expression by several meta-clusters, and these mapped onto manually-gated cell populations transiting from un-activated to activated phenotypes. For instance, the transition from un-activated to activated monocytes can be observed in meta-clusters E, E|1, and E|3 at day 5, 3, and 5 respectively. The same can be observed for meta-cluster K (stem and

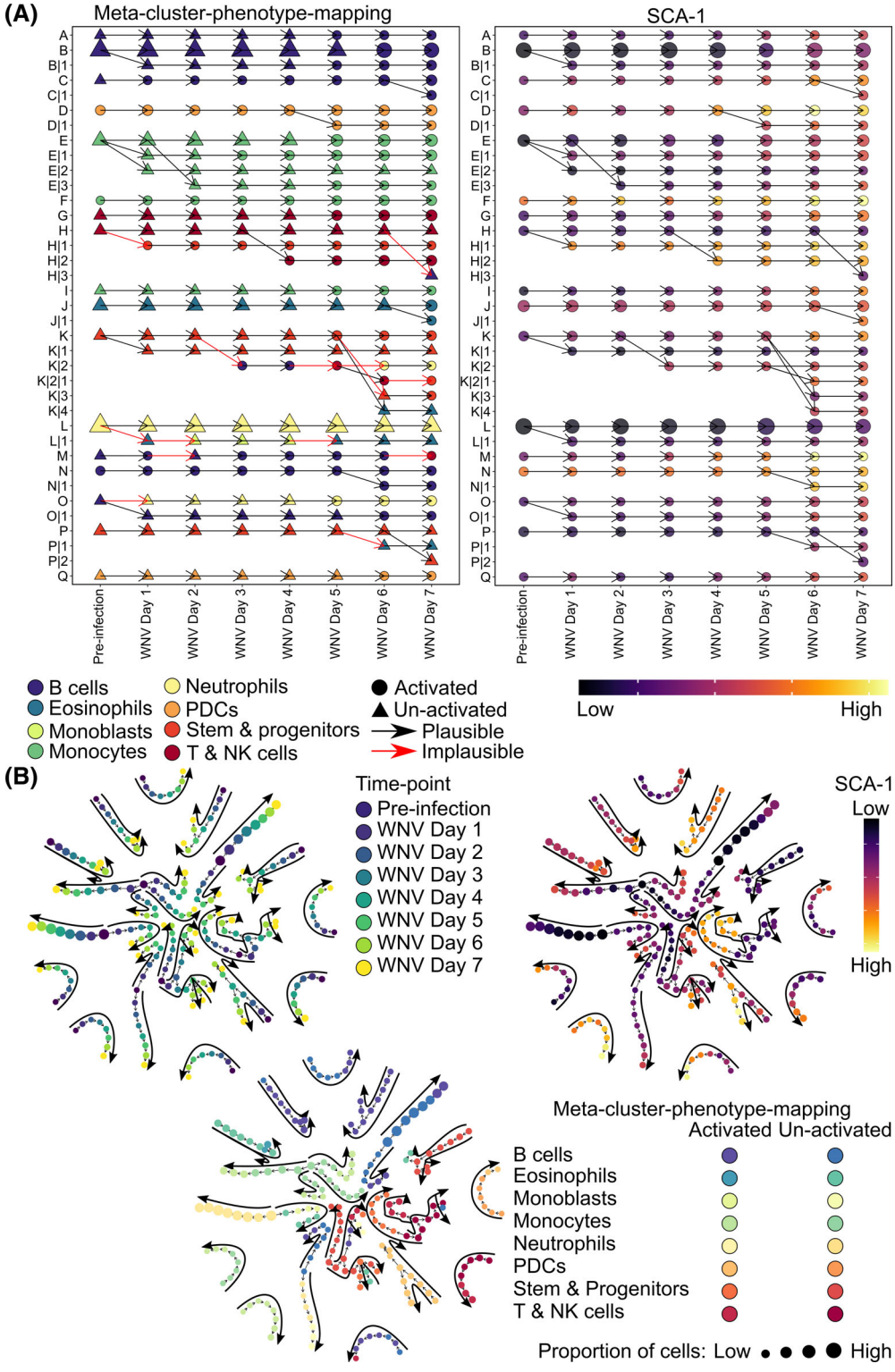

FIGURE 5 TrackSOM clustering of West Nile Virus (WNV) bone marrow dataset using no-merging Prescribed Invariant operation. (A) Time-series heatmaps illustrating meta-clusters' (symbols) temporal changes (lines). Meta-clusters on the left plot are colored by the majority cell population they each represent. Biological plausibility of transitions are indicated based on known cellular differentiation pathways (see Supplementary Figure S3). The right plot captures the average expression of SCA-1 marker for each meta-cluster. (B) Network plots with meta-clusters colored by time-point, cell population represented, and average SCA-1 marker expression. [Color figure can be viewed at wileyonlinelibrary.com]

progenitor cells) and meta-cluster G (T and NK cell populations) at day 5. Not all the meta-clusters capturing a given population show upregulation of SCA-1 marker: clusters P and K|1 also capture stem and progenitor cells, but show no such transition, demonstrating a capacity to uncover heterogeneity within phenotypes in response to infection. Further, being the bone marrow, activated forms of all the cell populations exist in small numbers even at pre-infection, and this is reflected in the cluster population assignments, showing the sensitivity of TrackSOM to uncovering minor phenotypes in the data. For instance, meta-cluster N captures pre-infection activated B cells, and meta-cluster D does likewise for PDCs. The majority of cluster evolutions are biologically plausible (see Supplementary Figure S3), although there are some exceptions, for example, in meta-cluster M, B cells transit into T and NK cells. Such implausible transitions occur only in meta-clusters that capture a minority of data points (indicated by small symbols) and these likely reside on the periphery between several high-density regions of cell populations in marker-space.

Together, these data demonstrate that TrackSOM can accurately identify, without external guidance, both majority and minority population phenotypes and track their evolution through time in terms of changing marker expression levels. Further, it offers excellent performance in doing so without being overly sensitive to exact parameter value choices.

3.2.1 | Comparison with ChronoClust

We contrasted TrackSOM's performance with that of ChronoClust on the WNV bone marrow dataset. For a fair comparison with TrackSOM, we collected 100 unique parameter value combinations for ChonoClust (see Methods and Materials) and obtained their mean ARI and tracking accuracy scores.

All TrackSOM modes of operation vastly out-performed ChronoClust, based on mean ARI score, Figure 4B. All TrackSOM solutions held ARI scores >0.38, whereas 30% of ChronoClust solutions scored <0, indicative of worse than random clustering. Curiously, these very poor ARI scoring ChronoClust solutions also exhibited very high tracking accuracy scores (Figure 4A). Indeed, ChronoClust tended to generate higher tracking accuracy scores than TrackSOM, Figure 4C. Manual inspection of representative ChronoClust solutions revealed a pattern of severely over-clustering the data (Supplementary Figure S6 and Table S5) into highly homogeneous clusters (poor ARI) that exhibited self-referential temporal transitions which are technically valid and thus promote high tracking accuracy scores. Examining Figure 4A more closely, the best ChronoClust solution (black circle on figure) held an ARI score of 0.76 and a tracking accuracy of 0.87. These scores are impressive, but still suboptimal compared to what TrackSOM was readily able to produce in many solutions under parameteric perturbation. Indeed, the very wide distribution of ChronoClust solution scores suggests the algorithm to be quite sensitive to its parameter values, mirroring prior findings elsewhere [12]. Obtaining high-quality performance would require either considerable exploration of ChronoClust parameter space or prior knowledge of good parameter value choices. Overall, we consider TrackSOM the vastly superior algorithm.

3.3 | Advice on using TrackSOM

Key to the successful use of clustering algorithms on novel data is the selection of appropriate parameter values. We performed a global sensitivity analysis to relate TrackSOM's two most influential parameters, SOM grid size and the number of meta-clusters generated, to its clustering and tracking performance. Accordingly, we calculated PRCC of these parameters against TrackSOM performance on the WNV bone marrow dataset. PRCC ascertains the effect of one parameter on TrackSOM performance while controlling for the effect of other parameters whose values will simultaneously be varying given the Latin Hypercube design [30]; this is a form of global sensitivity analysis [31].

Examining the SOM grid size parameter first, Table 1, we find frequent positive correlations of grid size with mean ARI, yet negative correlations with tracking accuracy. For the Prescribed modes of operation, we find a non-linear diminishing increase in mean ARI as the SOM grid size increases (Supplementary Figures S7A,S8A,S9A,S10A). With SOM nodes being the fundamental unit of clustering, excessively small SOM grid sizes will capture distinct cellular phenotypes into the same nodes, ultimately reducing ARI scores. This effect was worse for no-merging operation, where small SOM grid sizes generate more variable and lower ARI scores (Supplementary Figures S7A vs. S8A, and S9A vs. S10A), suggesting that merging operation has some capacity to correct for poor choices of SOM grid size. Merging Autonomous Adaptive operation exhibited an opposite pattern, with increasing SOM grid size reducing ARI scores, Supplementary Figure S11A. We suspect this relates to an interaction between SOM grid size and how FlowSOM chooses the number of meta-clusters to generate, see Supplementary Section S6.

There is a consistent pattern of increasing SOM grid sizes reducing the tracking accuracy. Tracking accuracy can diminish when the number of meta-clusters drops, because numerous meta-clusters tend toward self-referential transitions which are always valid. However, the presently observed pattern is independent of the number of meta-clusters created: PRCC controls for this parameter, and we find no discernible differences between merging and no-merging operations (the latter tends to raise meta-cluster numbers). This suggests that the pattern is driven by incorrect meta-cluster label re-assignments across time-points, which result in invalid population transitions. SOM nodes do not aim to capture equal numbers of data-points, instead they emphasize a comprehensive coverage of space, even between high-density regions. As such, increasing the SOM grid size will result in more nodes supported by fewer data points. Such nodes will be more sensitive to how the distribution of the data changes between time-points, and for minority phenotypes (which are represented by relatively few nodes), these dynamics under large grid sizes could result in strong fluctuations of nodes entering and leaving the meta-cluster. We believe this could result in assignments of

incorrect labels to meta-clusters and thus invalid transitions that reduce tracking accuracy scores.

As expected, increasing meta-cluster numbers tended to increase tracking accuracy for Prescribed modes of operation, though we found no effect for Autonomous Adaptive operation (Table 1, graphs in Supplementary Figures S11 and S13). This likely reflects increasingly homogenous meta-clusters, many of which will generate self-referential transitions as explored in the datasets above. For Prescribed Invariant operational modes, increasing the number of meta-clusters reduced the ARI scores, likely due to the resultant over-clustering relative to manually-gated populations. Curiously, Autonomous Adaptive operation again exhibited a converse pattern, wherein the maximum meta-cluster number explored increased ARI scores; see Supplementary Section S6.

TABLE 1 Partial Rank Correlation Coefficients (PRCC) of TrackSOM parameters against clustering (mean ARI) and tracking accuracy performance metrics for the indicated modes of operation. Data represents samplings of TrackSOM parameter space and corresponding clustering of the WNV bone marrow dataset. Only parameters with PRCC *p*-values <0.005 are indicated (*p*-values not shown); ns, not significant

Operational mode	PRCC: Mean ARI	PRCC: Tracking
Parameter: SOM grid size		
Autonomous Adaptive, merging	−0.72	−0.50
Autonomous Adaptive, no-merging	ns	ns
Prescribed Invariant, merging	0.62	−0.93
Prescribed Invariant, no-merging	0.71	−0.71
Prescribed Variant, merging	0.78	−0.95
Prescribed Variant, no-merging	0.82	−0.66
Parameter: Maximum number of meta-clusters considered		
Autonomous Adaptive, merging	0.68	ns
Autonomous Adaptive, no-merging	0.38	ns
Parameter: Number of meta-clusters		
Prescribed Invariant, merging	−0.55	0.88
Prescribed Invariant, no-merging	−0.83	ns
Parameter: Number of meta-clusters per time-point		
Prescribed Variant (Pre-infection), merging	−0.29	ns
Prescribed Variant (day 1), merging	−0.31	ns
Prescribed Variant (day 2), merging	−0.35	ns
Prescribed Variant (day 3), merging	−0.39	0.52
Prescribed Variant (day 4), merging	ns	0.50
Prescribed Variant (day 6), merging	ns	0.32
Prescribed Variant (day 1), no-merging	−0.37	ns
Prescribed Variant (Pre-infection), no-merging	−0.52	ns

Abbreviations: ARI, Adjusted Rand Index; PRCC, Partial Rank Correlation Coefficients; SOM, Self-Organizing Map.

For users with unanalyzed data we offer the following usage advice (Figure 6). If the number of cellular phenotypes to be discovered is completely unknown, Autonomous Adaptive operation can give a solid performance. Prescribed Invariant operation can be more effective, but is dependent on parameter values that such users are unlikely to have a basis to choose well; this is seen for both synthetic (Figure 2) and bone marrow datasets (Figure 4). Autonomous Adaptive operation can accommodate a large number of meta-clusters to explore (we attempted up to 40 with no deterioration in performance, Supplementary Figures S11B and S13A), and we recommend users setting this to a large number as TrackSOM will adjust as needed. If the number of populations sought is known, then the Prescribed operations are more appropriate. Prescribed Variant is more appropriate where the number of cell populations to be discovered varies over time, for instance via tissue infiltrates. If the number of populations is instead constant over time, then Prescribed Invariant is more appropriate and has fewer parameters to set.

Preventing meta-cluster merging arguably better reflects the biology of cellular differentiation, and it conferred better performance in all operational modes. However, this came at the possible expense of less meaningful clustering through an explosion of meta-cluster numbers if parameters are poorly chosen. The alternative, enabling merging, restricts the number of meta-clusters created, but can instead generate convoluted meta-cluster transitions. We suggest users try both modes and select that which better suits their circumstance. For selecting a SOM grid size, values between 10 and 15 consistently delivered a good balance between ARI and tracking accuracy scores (Supplementary Figures S7–S11).

3.4 | Mapping the CNS immune response onto WNV clinical disease severity

TrackSOM is not constrained to analyzing time-course data, it can also relate immune system status to the spectrum of clinical disease severities that present at a single time-point. We demonstrated this in a novel, ungated dataset representing the immune cells in the brains of mock- and WNV-infected mice at day 7 post-infection. We ran TrackSOM over the sequence of cytometry datasets indicative of increasing disease severities to relate changes in cellular phenotypes to disease-stage at this single time-point. In accordance with the above TrackSOM usage advice, we employed Prescribed Variant operation with meta-cluster merging disabled (see methods and materials) and conducted some manual exploration of meta-cluster and grid size numbers before manually selecting a solution that we judged to best separate the cell phenotypes. We anticipated that the number of cellular populations would increase with disease severity. This was supported by preliminary unguided analysis through FIt-SNE, where increased severity was associated with a larger number of distinct clusters of cells (Supplementary Figure S14).

TrackSOM successfully distinguished numerous meta-clusters of cells, that we could resolve to immune cell phenotypes, and related

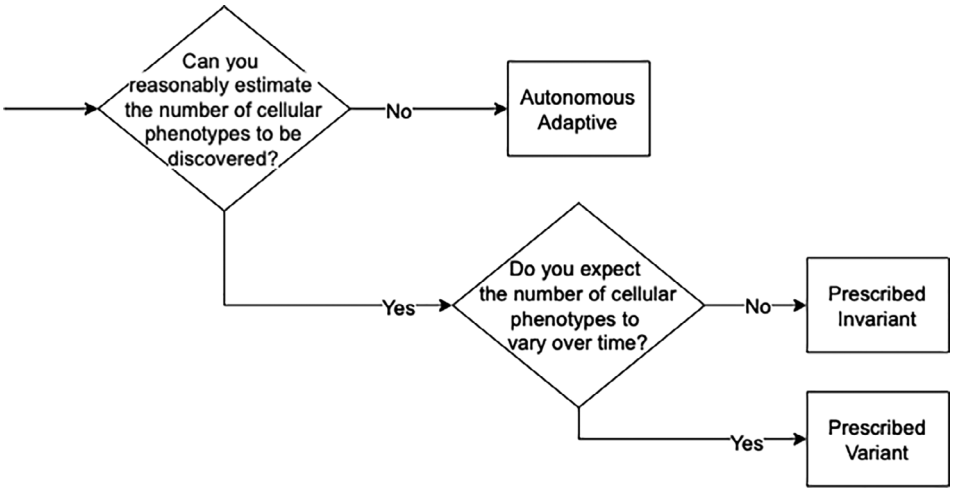

FIGURE 6 Practical guidelines for choosing TrackSOM's mode of operation. As the separability and tracking of cellular phenotypes mostly depends on the number of meta-clusters for each time-point, the choice of operation mode is best guided by the user's prior knowledge of the expected number of phenotypes in the data.

their phenotypic evolution to increasing disease severity, Figure 7 and Supplementary Figure S15. TrackSOM tracked how the relative (Figure 8) and absolute counts (Supplementary Figure S16) of cellular populations, and their marker expression levels (Figure 7B and Supplementary Figure S15) shifted with disease severity. As disease severity rose from mock-infection to the 2nd most severe stage WNV-04, we observed substantial rises in infiltrating macrophage, CD4 and CD8 T cell, and NK cell populations. Further, we tracked a drop in the proportion of microglia from 82% to 10% of total cells, consistent with the concomitant substantial increases over this time in infiltrating macrophages. The proportion of infiltrating macrophages rose sharply from 9% to 65% of total cells by WNV-04. Several of these cell populations are not ordinarily found in the brain but were identified by TrackSOM in the mock-infection dataset, likely representing either the capture of post-flushing remnant intravascular cells in the samples and/or the recognition of central nervous system-associated macrophage populations in the homeostatic brain [32].

Notably, TrackSOM identified the infiltration of phenotypic subpopulations with worsening disease severity, indicated by the branching meta-clusters in Figure 7. Macrophages were represented by 11 meta-clusters in mock-infection, curiously high given their low relative abundance (Figure 8), and 4 of these underwent extensive branching into distinct subpopulations as disease progressed. Most of these macrophage subphenotypes upregulated the CD80 activation marker with increasing disease severity, indicative of response to inflammation. NK cells were captured by one meta-cluster which expressed the NK1.1 marker expression. Both CD4 and CD8 T cell populations were captured by meta-clusters that bifurcated at the point of mild (WNV-01) symptoms in accordance with CD4 and CD8 marker expression levels respectively, presumably capturing differentiating functions under immune challenge [33]. Microglia were captured by a single meta-cluster at mock infection that branched into several sublineages with heightening symptoms, varying in CD11b, Ly6C, and CD45 expression.

Overall, TrackSOM was able to elucidate and map the changing profile of CNS immune cell populations onto a progression of clinical symptom severities emanating from WNV infection.

4 | DISCUSSION

Immune responses are fallible. Identical inoculation doses can cause disease that resolves in some individuals but are lethal in others [20], and autoimmunity perhaps represents quintessential immune system error [34]. In addition, many non-communicable diseases, of which prevalence rates are rising globally, have an immune system component [35, 36]. Many such diseases are characterized by periods of remission and relapse, often with localized exacerbation of inflammation, such as the skip legions that characterize Crohn's disease [37] or the localized lesions in multiple sclerosis. The tone of the immune response, and its influence on health status, is a complex consequence of broad interactions between numerous dynamic cell populations [34, 36].

Thus, the identification and characterization of novel cell subpopulations has potential therapeutic benefit [14]. Mapping the immune response, at the level of cell population dynamics and interactions, over time and in relevant organs of the body, and relating these maps to disease outcomes, is essential for discovering possible intervention targets.

We developed TrackSOM with a view to enable mapping of the immune response against time and/or disease severity. TrackSOM excelled in unveiling the temporal evolution of clusters engineered into a synthetic dataset. When applied to pregated, time-series bone marrow cytometry data from WNV-infected mice, TrackSOM correctly tracked the activation of several immune cell populations and the upregulation of the SCA-1 marker. On ungated data, TrackSOM identified and tracked the infiltration and functional evolution of distinct subpopulations of key immune cell phenotypes and mapped these against clinical disease-stage. Further, TrackSOM readily revealed how the phenotypic composition, in terms of both absolute cell counts and relative abundance, varied with disease severity. TrackSOM is a powerful tool for studying coordinated immune population dynamics and how these pertain to disease status; it has already been used to identify cell populations and their shifting functional status over time in COVID-19 infection in humans [38].

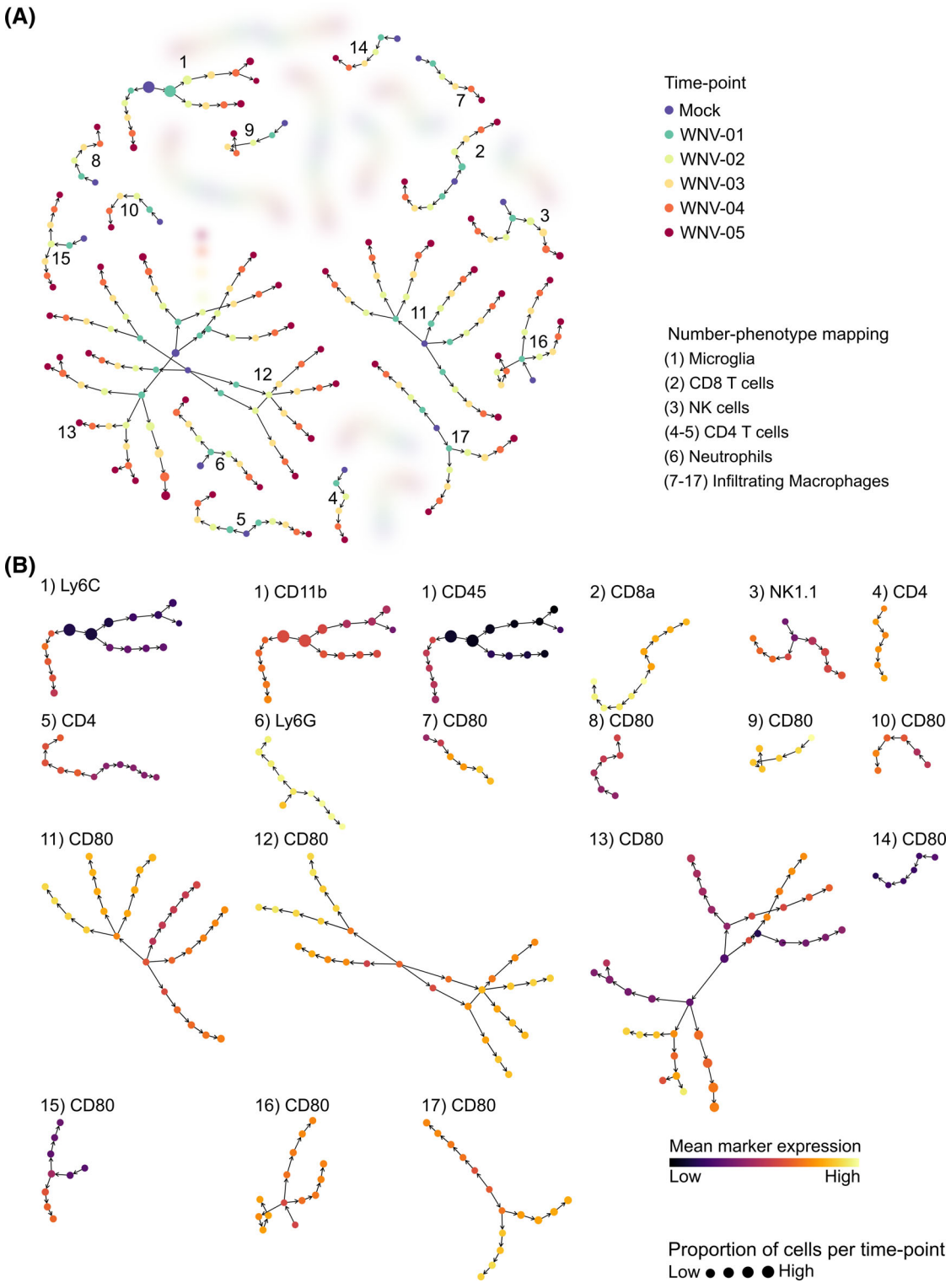

FIGURE 7 TrackSOM-enabled exploration of the immune response in the brains of West Nile Virus (WNV)-infected mice over 6 graded clinical disease severities. All data represent day 7 post-inoculation. Shown is a network plot of TrackSOM meta-clusters and their evolution with worsening clinical disease severity at this time-point. Analysis is focused on meta-clusters given high-confidence manual phenotype labels (not blurred). Meta-clusters in (a) are colored by disease severity and arrows indicate the sequence of progression. Meta-clusters in (B) are colored by the indicated mean marker expression level. The numbering in (B) refers to the number-phenotype mapping in (a). For (B), color bars capture the full range of values for each marker independently. Node sizes relate to the proportion of cells captured in the meta-cluster at given disease severity. [Color figure can be viewed at wileyonlinelibrary.com]

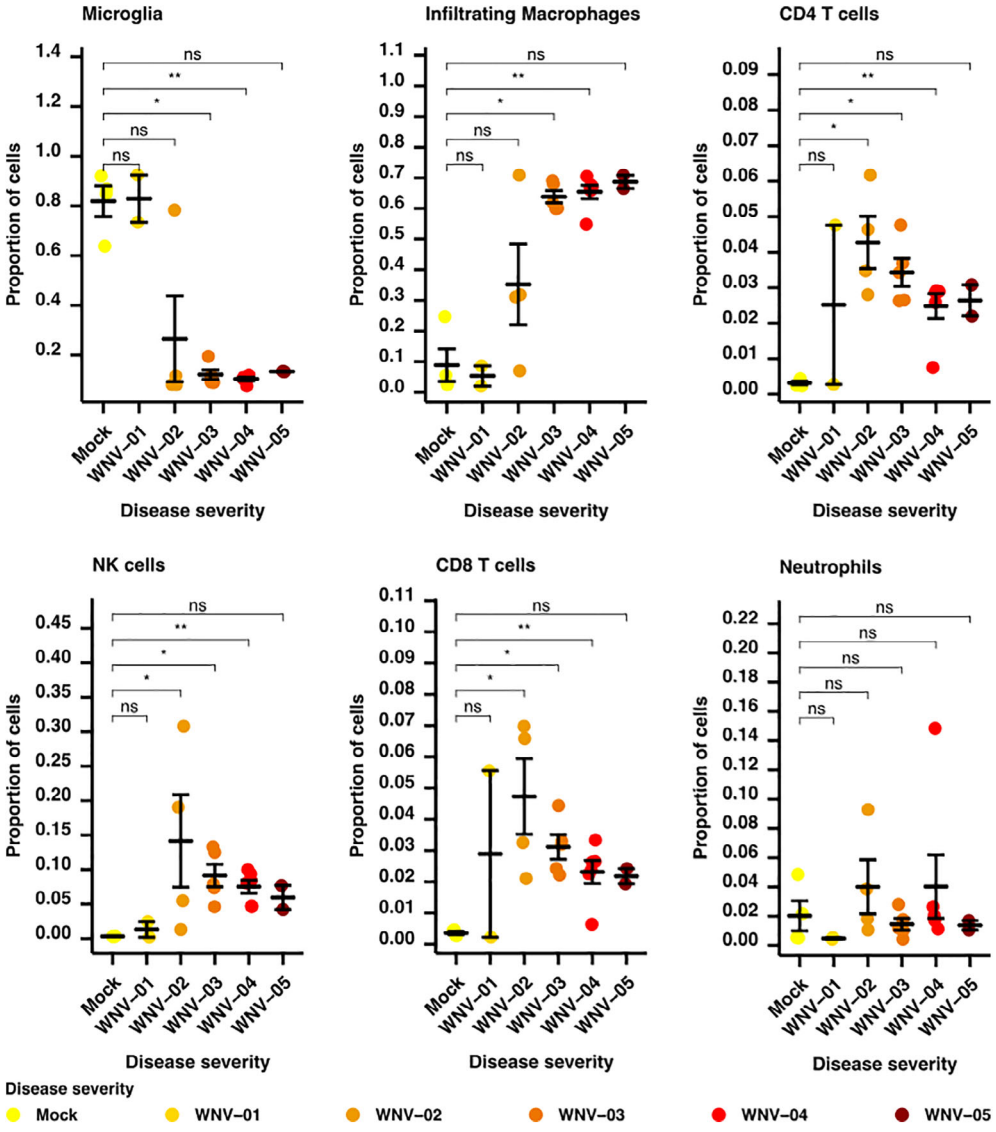

FIGURE 8 The profile of immune cell phenotypes varies with increasing disease severity in the brains of West Nile Virus (WNV)-infected mice. Mapped are the proportion of the given cell phenotypes per mouse. Only cells (and TrackSOM meta-clusters) that we could label with high confidence are included in the calculation of proportions. Data represent between $n = 2$ and $n = 6$ mice per group (Supplementary Table S3). Note that ranges for y-axes vary across panels. Pairwise statistical comparisons between mock and other disease severities are through the Wilcoxon test; * signifies p-value <=0.05, ** signifies p-value <=0.01, ns signifies not significant. [Color figure can be viewed at wileyonlinelibrary.com]

When evaluating clustering solutions, we observed a degree of trade-off across our two algorithm performance metrics. Solutions judged optimal for cluster quality tended to be poorer for cluster evolution tracking, and vice versa, particularly for Chronoclust. This is seemingly contradictory; how could poor quality clusters be well-tracked? The tracking accuracy metric is somewhat ignorant of cluster quality, and judges specifically only if cluster transitions are valid. Over-clustering, wherein single phenotypes are split across numerous meta-clusters, will tend to produce valid, self-referential transitions, thus scoring low on ARI but high on tracking accuracy. Under-clustering, wherein single meta-clusters amalgamate several phenotypes, will adopt and retain the majority population label, again valid for tracking accuracy but poor for ARI. Conversely, high quality clusters could also be poorly tracked. This may be due to misclustered data-points (leading to non-homogeneous meta-clusters) suppressing the ARI score to the same extent, irrespective of which meta-clusters they occur in, but not tracking accuracy where all meta-clusters contribute equally to the score irrespective of their size in terms of data-points. Misclusterings that are concentrated in a single, small meta-cluster at a given time-point could switch the assigned label of the meta-cluster and lead to invalid transitions, thus impacting tracking accuracy. Importantly, if misclusterings are instead confined to large meta-clusters or broadly distributed across many meta-clusters, tracking accuracy is unlikely to be affected, though ARI would be. Hence, the location of misclustered data-points could impact tracking accuracy while minimally affecting ARI.

There are advantages to multi-metric evaluation, as single metrics (and thus values) struggle to comprehensively describe all aspects of a complex phenomenon. Complementary perspectives through multiple metrics can offer a robust and comprehensive evaluation framework [12], and is the approach we took here. We identified and focused on solutions that exhibited minimal trade-offs between the two metrics, those solutions that generated an optimal balance of the conflicting criteria. With only two metrics, scatter plots such as Figures 2A and 4A were sufficient to identify such solutions. Where more metrics are employed, and the potential trade-offs more severe, a formalized framework through Pareto fronts can be employed [12].

TrackSOM's Autonomous Adaptive operation carries considerable appeal, being largely insensitive to exact parameter value choices and thus simplifying the task of choosing appropriate values. However, we consistently found it to under-cluster the data, failing to fully distinguish all the subtly different cell populations discovered through expert manual gating. Its performance lagged behind the best that the Prescribed Invariant operations could produce when given optimal parameter values. Autonomous Adaptive operation draws upon FlowSOM's procedure for self-determining an appropriate number of meta-clusters. The selection is made by identifying the point at which the reducing variance captured per meta-cluster sharply diminishes as the number of meta-clusters generated increases [11] (a standard approach in clustering called the "elbow method"). FlowSOM's method is simple and quick to execute, however it evidently does not converge on a clustering that an expert would manually produce. Identifying an alternative method that more closely aligns with expert opinion in cytometry specifically would be transformative for the discipline, enabling a heuristic search for appropriate parameter values that algorithm users must currently best-guess at and laboriously evaluate. This is a worthy avenue of future work, though prior efforts indicate that it is non-trivial [39].

A key investigative theme in immunology is to relate the kinetics of immune cell subpopulations with disease outcomes, and to identify the key events at which outcomes bifurcate; these point to potential therapeutic targets. TrackSOM's primary contribution is as an algorithm that automatically identifies and tracks the dynamics of cellular subpopulations, mapping these over time and/or progressing disease status. A salient line of future investigation is the automated extraction of features in the rich data TrackSOM generates that constitute key distinguishers of clinical outcomes. For instance, TrackSOM could be run twice, over datasets of patients who recover versus those who do not. Within a supervised machine learning context, classifiers could be built over TrackSOM-originating features describing subpopulation kinetics to highlight those of particular relevance. Further, the classifiers would be of natural clinical benefit, potentially informing clinical decisions for novel patients by predicting their future clinical outcomes. A contemporary possibility could include for example, predicting in advance which COVID-sufferers would require ventilators or admission into intensive care units. We are currently pursuing such technologies.

We focused here on cytometry because it is relatively cheap to run, accessible and widely adopted. Single-cell RNA sequencing (scRNAseq) is a related technology with falling costs and growing popularity. There is interest in identifying phenotype differentiation pathways through scRNAseq using trajectory inference algorithms [40–45]. This resembles TrackSOM's usage context, and while trajectory inference algorithms have been applied over amalgamated time-series datasets [46, 47], the algorithms themselves do not account for time, and thus do not relate changes to specific time-points. Pseudodynamics [48] added an explicit accounting for time to trajectory inference methods. ScRNAseq differs from cytometry in that the data generated typically cover many thousands of features (genes) and hundreds to thousands of cells, whereas cytometry captures under 100 features and up to millions of cells. How well pseudodynamics can accommodate cytometry data, and how well TrackSOM performs over scRNAseq data is worthy of future investigation.

AUTHOR CONTRIBUTIONS

Givanna H. Putri, Suat Dervish, Irena Koprinska, Nicholas J. C. King, Thomas M. Ashhurst, Mark N. Read conceived the project. Givanna H. Putri, Jonathan Chung, Davis N. Edwards, Felix Marsh-Wakefield implemented the TrackSOM algorithm. Thomas M. Ashhurst performed the experimental work for the WNV datasets. Givanna H. Putri, Thomas M. Ashhurst, Mark N. Read interpreted results. Givanna H. Putri, Jonathan Chung, Mark N. Read wrote the manuscript. Suat Dervish, Irena Koprinska, Nicholas J. C. King, Thomas M. Ashhurst, Mark N. Read supervised the project. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank the Sydney Informatics Hub (University of Sydney) for providing access to their High-Performance Computing facility. We thank the Sydney Cytometry Core Research Facility for assisting in data generation, and providing access to computational analysis resources.

FUNDING INFORMATION

Givanna H. Putri was supported by the Australian Government Research Training Program Scholarship. Davis N. Edwards was supported by a summer scholarship awarded by the University of Sydney's Westmead Initiative. Jonathan Chung is supported by the Australian Government Research Training Program Scholarship. Nicholas J. C. King was funded by a grant from the Merridew Foundation. Thomas M. Ashhurst and Felix Marsh-Wakefield are supported by the International Society for the Advancement of Cytometry (ISAC) Marylou Ingram Scholars Program. Mark N. Read received support from the University of Sydney Westmead Initiative.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/cyto.a.24668.

ORCID

Givanna H. Putri https://orcid.org/0000-0002-7399-8014
Felix Marsh-Wakefield https://orcid.org/0000-0002-6839-7628
Irena Koprinska https://orcid.org/0000-0001-9479-4187
Nicholas J. C. King https://orcid.org/0000-0002-3877-9772
Thomas M. Ashhurst https://orcid.org/0000-0001-7269-7773
Mark N. Read https://orcid.org/0000-0002-1481-4780

REFERENCES

1. Bendall SC, Davis KL, Amir ED, Tadmor MD, Simonds EF, Chen TJ, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell. 2014;157:714–25.

2. Park LM, Lannigan J, Jaimes MC. OMIP-069: forty-color full spectrum flow cytometry panel for deep immunophenotyping of major cell subsets in human peripheral blood. Cytom A. 2020;97:1044–51.
3. Mair F, Prlic M. OMIP-044: 28-color immunophenotyping of the human dendritic cell compartment. Cytom A. 2018;93:402–5.
4. Nettey L, Giles AJ, Chattopadhyay PK. OMIP-050: a 28-color/30-parameter fluorescence flow cytometry panel to enumerate and characterize cells expressing a wide array of immune checkpoint molecules. Cytom A. 2018;93:1094–6.
5. Mair F, Hartmann FJ, Mrdjen D, Tosevski V, Krieg C, Becher B. The end of gating? An introduction to automated analysis of high dimensional cytometry data. Eur J Immunol. 2016;46:34–43.
6. Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the human immunology project. Nat Rev Immunol. 2012;12:191–200.
7. Saeys Y, Van Gassen S, Lambrecht BN. Computational flow cytometry: helping to make sense of high-dimensional immunology data. Nat Rev Immunol. 2016;16:449–62.
8. Liu P, Liu S, Fang Y, Xue X, Zou J, Tseng G, et al. Recent advances in computer-assisted algorithms for cell subtype identification of cytometry data. Front cell. Dev Biol. 2020;8:234.
9. Qiu P, Simonds EF, Bendall SC, Gibbs KD Jr, Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol. 2011;29:886–91.
10. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir EAD, Tadmor MD, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell. 2015;162:184–97.
11. van Gassen S, Callebaut B, van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data: flowSOM. Cytom A. 2015;87:636–45.
12. Putri GH, Koprinska I, Ashhurst TM, King NJC, Read MN. Using single-cell cytometry to illustrate integrated multi-perspective evaluation of clustering algorithms using Pareto fronts. Bioinformatics. 2021;37(14):1972–81.
13. Weber LM, Robinson MD. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data: comparison of high-dim. cytometry clustering methods. Cytom A. 2016;89:1084–96.
14. Krzywinska E, Cornillon A, Allende-Vega N, Vo DN, Rene C, Lu ZY, et al. CD45 isoform profile identifies natural killer (NK) subsets with differential activity. PLoS ONE. 2016;11:1–18.
15. Putri GH, Read MN, Koprinska I, Singh D, Röhm U, Ashhurst TM, et al. ChronoClust: density-based clustering and cluster tracking in high-dimensional time-series data. Knowl-Based Syst. 2019;174:9–26.
16. Parks DR, Roederer M, Moore WA. A new "Logicle" display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. Cytom Part J Int Soc Anal Cytol. 2006;69:541–51.
17. Essers MAG, Offner S, Blanco-Bose WE, Waibler Z, Kalinke U, Duchosal MA, et al. IFNalpha activates dormant haematopoietic stem cells in vivo. Nature. 2009;458:904–8.
18. Pietras EM, Lakshminarasimhan R, Techner J-M, Fong S, Flach J, Binnewies M, et al. Re-entry into quiescence protects hematopoietic stem cells from the killing effect of chronic exposure to type I interferons. J Exp Med. 2014;211:245–62.
19. de Bruin AM, Voermans C, Nolte MA. Impact of interferon-γ on hematopoiesis. Blood. 2014;124:2479–86.
20. King NJC, Getts DR, Getts MT, Rana S, Shrestha B, Kesson AM. Immunopathology of flavivirus infections. Immunol Cell Biol. 2007;85:33–42.
21. Morcos MNF, Schoedel KB, Hoppe A, Behrendt R, Basak O, Clevers HC, et al. SCA-1 expression level identifies quiescent hematopoietic stem and progenitor cells. Stem Cell Rep. 2017;8:1472–8.
22. Hubert L, Arabie P. Comparing partitions. J Classif. 1985;2:193–218.
23. McKay MD, Beckman RJ, Canover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a a computer code. Dent Tech. 1979;21:239–45.
24. Read MN, Alden K, Timmis J, Andrews PS. Strategies for calibrating models of biology. Brief Bioinform. 2020;21:24–35.
25. Putri GH, Read MN, Koprinska I, Ashhurst TM, King NJC. Dimensionality reduction for clustering and cluster tracking of cytometry data. Int Conf Artif Neural Netw Springer. 2019;624–40.
26. Alden K, Read M, Timmis J, Andrews P.S., Veiga-Fernandes H., Coles M. Spartan: a comprehensive tool for understanding uncertainty in simulations of biological systems. Loewe L, editor. PLoS Comput Biol 2013;9:e1002916.
27. Scrucca L, Fop M, Murphy TB, Raftery AE. Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J. 2016;8:289–317.
28. Linderman GC, Rachh M, Hoskins JG, Steinerberger S, Kluger Y. Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data. Nat Methods. 2019;16:243–5.
29. Ashhurst TM, Marsh-Wakefield F, Putri GH, Spiteri AG, Shinko D, Read MN, et al. Integration, exploration, and analysis of high-dimensional single-cell cytometry data using spectre. Cytom A. 2022;101:237–53.
30. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254:178–96.
31. Read M, Andrews PS, Timmis J, Kumar V. Techniques for grounding agent-based simulations in the real domain: a case study in experimental autoimmune encephalomyelitis. Math Comput Model Dyn Syst. 2012;18:67–86.
32. Ivan DC, Walthert S, Berve K, Steudler J, Locatelli G. Dwellers and trespassers: mononuclear phagocytes at the borders of the central nervous system. Front Immunol. 2021;11:3892.
33. Xiao Z, Mescher MF, Jameson SC. Detuning CD8 T cells: Down-regulation of CD8 expression, tetramer binding, and response during CTL activation. J Exp Med. 2007;204:2667–77.
34. Li R, Patterson KR, Bar-Or A. Reassessing B cell contributions in multiple sclerosis. Nat Immunol. 2018;19:696–707.
35. De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. Nat Rev Cancer. 2006;6:24–37.
36. Zeng B, Shi S, Ashworth G, Dong C, Liu J, Xing F. ILC3 function as a double-edged sword in inflammatory bowel diseases. Cell Death Dis. 2019;10:315. https://doi.org/10.1038/s41419-019-1540-2
37. Roda G, Chien Ng S, Kotze PG, Argollo M, Panaccione R, Spinelli A, et al. Crohn's disease. Nat Rev Dis Primer. 2020;6:1–19. https://doi.org/10.1038/s41572-020-0156-2
38. Koutsakos M, Rowntree LC, Hensen L, Chua BY, van de Sandt CE, Habel JR, et al. Integrated immune dynamics define correlates of COVID-19 severity and antibody responses. Cell Rep Med. 2021;2:100208.
39. Wiwie C, Baumbach J, Röttger R. Comparing the performance of biomedical clustering methods. Nat Methods. 2015;12:1033–8.
40. Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat Biotechnol. 2016;34:637–45.
41. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics. 2018;19:477.
42. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32:381–6.
43. Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. Nat Biotechnol. 2019;37:547–54.
44. Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 2019;20:1–9.

45. van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeys Y, et al. Trajectory-based differential expression analysis for single-cell sequencing data. Nat Commun. 2020;11:1–13.
46. Byrnes LE, Wong DM, Subramaniam M, Meyer NP, Gilchrist CL, Knox SM, et al. Lineage dynamics of murine pancreatic development at single-cell resolution. Nat Commun. 2018;9:1–17.
47. Scavuzzo MA, Hill MC, Chmielowiec J, Yang D, Teaw J, Sheng K, et al. Endocrine lineage biases arise in temporally distinct endocrine progenitors during pancreatic morphogenesis. Nat Commun. 2018;9:1–21.
48. Fischer DS, Fiedler AK, Kernfeld EM, Genga RMJ, Bastidas-Ponce A, Bakhti M, et al. Inferring population dynamics from single-cell RNA-sequencing time series data. Nat Biotechnol. 2019;37:461–8.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Putri GH, Chung J, Edwards DN, Marsh-Wakefield F, Koprinska I, Dervish S, et al. TrackSOM: Mapping immune response dynamics through clustering of time-course cytometry data. Cytometry. 2023;103(1):54–70. https://doi.org/10.1002/cyto.a.24668