
Article:
Lönnberg, T., Svensson, V., James, K.R. et al. (20 more authors) (2017) Single-cell RNA-seq and computational analysis using temporal mixture modeling resolves TH1/TFH fate bifurcation in malaria. Science Immunology, 2 (9). eaal2192. ISSN 2470-9468

https://doi.org/10.1126/sciimmunol.aal2192

Published in final edited form as:
Sci Immunol. ; 2(9): . doi:10.1126/sciimmunol.aal2192.

Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria

Tapio Lönnberg[1,2,#], **Valentine Svensson**[1,#], **Kylie R. James**[3,#], **Daniel Fernandez-Ruiz**[4], **Ismail Sebina**[3], **Ruddy Montandon**[2], **Megan S. F. Soon**[3], **Lily G. Fogg**[3], **Arya Sheela Nair**[3], **Urijah Liligeto**[3], **Michael J. T. Stubbington**[1,2], **Lam-Ha Ly**[2], **Frederik Otzen Bagger**[1,2,5,6], **Max Zwiessele**[7], **Neil D. Lawrence**[7], **Fernando Souza-Fonseca-Guimaraes**[3], **Patrick T. Bunn**[3], **Christian R. Engwerda**[3], **William R. Heath**[4,8], **Oliver Billker**[2], **Oliver Stegle**[1,*,†], **Ashraful Haque**[3,*,†], and **Sarah A. Teichmann**[1,2,*,†]

[1]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, UK

[2]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK

[3]QIMR Berghofer Medical Research Institute, Herston, Brisbane, Queensland, Australia

[4]Department of Microbiology and Immunology, The Peter Doherty Institute, University of Melbourne, Parkville, Victoria, Australia

[5]Department of Haematology, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

[6]National Health Service (NHS) Blood and Transplant, Cambridge Biomedical Campus, Long Road, Cambridge, UK

[7]Department of Computer Science, University of Sheffield, Sheffield, UK

[8]The Australian Research Council Centre of Excellence in Advanced Molecular Imaging, The University of Melbourne, Parkville, Victoria, Australia

Abstract

Differentiation of naïve CD4$^+$ T cells into functionally distinct T helper subsets is crucial for the orchestration of immune responses. Due to extensive heterogeneity and multiple overlapping transcriptional programs in differentiating T cell populations, this process has remained a

*Correspondence to: st9@sanger.ac.uk, Ashraful.Haque@qimrberghofer.edu.au or stegle@ebi.ac.uk.
#denotes equal contribution
†denotes equal contribution

Author contributions

TL and KRJ performed the single-cell RNA-seq experiments. VS developed the GPfates model in collaboration with MZ, NDL, OS and SAT. DFR and WRH generated the PbTII mouse model. KRJ, RM, IS, MSFS, LGF, ASN, UL, FSFG, PTB and CRE performed the mouse experiments. TL, VS, KRJ, LHL and FOB analysed the data and interpreted the results
MJTS performed the TCR clonality analysis. TL, KRJ, RM, OB, AH and SAT designed the experiments. OS, AH and SAT co-supervised the study. TL, VS, KRJ, OS, AH and SAT wrote the manuscript. All authors have read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Data and materials availability

The data presented in this paper is publically available in the ArrayExpress database with accession number E-MTAB-4388.

Europe PMC Funders Author Manuscripts

challenge for systematic dissection *in vivo*. By using single-cell transcriptomics and computational analysis using a temporal mixtures of Gaussian processes model, termed GPfates, we reconstructed the developmental trajectories of Th1 and Tfh cells during blood-stage *Plasmodium* infection in mice. By tracking clonality using endogenous TCR sequences, we first demonstrated that Th1/Tfh bifurcation had occurred at both population and single-clone levels. Next, we identified genes whose expression was associated with Th1 or Tfh fates, and demonstrated a T-cell intrinsic role for Galectin-1 in supporting a Th1 differentiation. We also revealed the close molecular relationship between Th1 and IL-10-producing Tr1 cells in this infection. Th1 and Tfh fates emerged from a highly proliferative precursor that upregulated aerobic glycolysis and accelerated cell cycling as cytokine expression began. Dynamic gene expression of chemokine receptors around bifurcation predicted roles for cell-cell in driving Th1/Tfh fates. In particular, we found that precursor Th cells were coached towards a Th1 but not a Tfh fate by inflammatory monocytes. Thus, by integrating genomic and computational approaches, our study has provided two unique resources, a database www.PlasmoTH.org, which facilitates discovery of novel factors controlling Th1/Tfh fate commitment, and more generally, GPfates, a modelling framework for characterizing cell differentiation towards multiple fates.

Introduction

$CD4^+$ T cells are key instructors of the immune system. They can display extensive phenotypic and functional diversity, by differentiating into a range of T helper (Th) subsets including Th1, Th2, Th17, Tfh, Th22, Treg and Th9 cells, that are distinguished mainly by cytokine and transcription factor expression profiles. Since Th cells can control infections and drive immune-mediated diseases there remains tremendous interest in the molecular mechanisms that mediate their *in vivo* differentiation.

Malaria, caused by the protozoan parasite *Plasmodium*, afflicted 212 million humans in 2015 (1). Both Th1 responses (2) and Tfh-dependent antibody responses (3) can independently protect against malaria, and are elicited simultaneously in malaria-infected individuals (4), as well as in mice challenged with rodent-infective strains such as *P. chabaudi chabaudi* AS (*Pc*AS) (5). However, the molecular relationships between Th1 and Tfh cells remain unclear during *Plasmodium* infection, and more generally during any immune challenge. A recent study has demonstrated that the unique T cell receptor of a naïve $CD4^+$ T cell imparted a strong preference for either a Th1 or Tfh fate (6). Nevertheless, for many clones both fates could still emerge, implying that other mechanisms, such as internal stochasticity and cell-extrinsic factors also govern fate choices *in vivo*. Transcription factors including T-bet, Gata3, RORγT and Bcl6 have been reported to drive and stabilize Th fates, leading to their characterisation as “lineage-defining” molecules. This has tended to present Th differentiation as a choice between mutually exclusive linear pathways. However, transient co-expression of these transcription factors, e.g. of Bcl6/T-bet, and Foxp3/RORγT suggests that overlapping intermediate Th states also exist *in vivo*. Moreover, substantial heterogeneity occurs in the kinetics of $CD4^+$ T cell responses, resulting in a complex mixture of intermediate states during differentiation, which is not easily resolved via assessment of a small number of molecules.

Conventional dendritic cells (cDCs) are the dominant initial source of antigenic signalling to naïve $CD4^+$ T cells in secondary lymphoid tissues, for example in the spleens of *Plasmodium*-infected mice (7). In other models, it was shown that cDCs made long-lasting stable contacts with naïve $CD4^+$ T cells to initiate priming (8). Once activated, $CD4^+$ T cells regained motility, permitting further cellular interactions. Consistent with this observation, activated $CD4^+$ T cells required further antigenic stimulation to optimize clonal expansion and Th differentiation (9); cDC were considered the most likely candidates, (8, 10), with other cell types remaining less explored. Studies of mice with altered monocytic responses suggested roles for these cells in $CD4^+$ T cell priming, specifically in tissues with few cDCs (11). Other reports employed cDC-deficiency to illustrate that monocytes could activate naive $CD4^+$ T cells (12). However, few *in vivo* studies have explored roles for monocytes in Th differentiation where cDC responses remain intact.

Here, we used single-cell RNA sequencing (scRNA-seq) to study *Plasmodium*-specific TCR transgenic $CD4^+$ T (PbTII) cells during blood-stage *Pc*AS infection in mice. We then employed a computational modelling strategy to reconstruct the molecular trajectories of Th1 and Tfh cells. Finally, we investigated cell-cell interactions based on dynamic expression of chemokines and their receptors, and examined roles for inflammatory monocytes in supporting activated $CD4^+$ T cells towards a Th fate.

Results

scRNA-seq resolves Th1 and Tfh cell fates during *Plasmodium* infection in mice

We used scRNA-seq to elucidate the development and heterogeneity of Th1 and Tfh cells during *Pc*AS infection (Figure 1A, Figure S1). We transferred naïve, proliferative dye-labeled PbTII cells into congenic wild-type mice, and recovered them at days 2, 3, 4, and 7 post-infection (p.i.) by cell sorting those expressing the early activation marker, CD69, or displaying dilution of the proliferative dye (Figure S2). Flow cytometric measurements of the canonical Th1 markers, T-bet (coded by *Tbx21*) and Interferon-γ, and Tfh markers, CXCR5 and Bcl6, indicated that these subsets emerged in parallel by day 7 p.i. (13, 14) (Figure 1B-D and S3). Notably, markers of Th2, Th17 or Treg subsets were not upregulated by PbTII cells (Figure S4).

Initially, we used principal component analysis (PCA) to explore the overall transcriptomic landscape of the PbTII cells (Figure S5A). The top principal components were strongly associated with the number of detected unique transcripts (reflective of mRNA content and proliferative status (Figure S5B)), and differentiation (Figures S5C and S6-7, Table S1). As expected, the variability related to previously established Th1 and Tfh gene expression signatures became more prominent with time, separating two subpopulations at day 7 (15) (Figure 1E-1F). Taken together, these results suggested a progressive commitment to Th1 and Tfh fates, and indicated that single-cell transcriptomes could be used for estimating both proliferative states and degrees of differentiation of individual cells.

Delineation of Th1 and Tfh trajectories using a Mixture of Gaussian Processes model

The results from the PCA suggested that variation in PbTII transcriptomes could be used to reconstruct the transcriptional programs that is underlying the Th1 and Tfh differentiation. To more explicitly model the temporal dynamics of this differentiation process, we developed GPfates, a temporal mixture model that builds on the Gaussian Process Latent Variable Model (GPLVM) (16) and Overlapping Mixtures of Gaussian Processes (OMGP) (17). Briefly, this approach is based on first reconstructing the differentiation trajectory from the observed data ("pseudotime", Figure 2A-B), thereby establishing an order for the cells. While our model uses the sample time as prior information, the inferred orderings did not strictly adhere to the experimental time points (Figure S8). For example, cells from day 4 p.i. were mixed with some of the cells from day 3 and day 7 at either end of the day 4 pseudotime distribution. This result is consistent with the idea that bulk assessments of cells at specific time points fail to account for the heterogeneity and differential kinetics of responses made by single cells. To assess the robustness of the established ordering, we repeated this analysis without supplying the experimental sampling times to the model, finding overall consistent results (Comp. Supp. Figure 8).

In a second step, GPfates uses the inferred temporal orders as input for a non-parametric time series mixture model (OMGP, (17)). This approach revealed two simultaneous trends emerging during pseudotime (Figure 2C-D), which separated from each other, indicating that a developmental bifurcation occurred.

In a third step, GPfates employs a change point model (Section 4.2 of the Computational Supplement), thereby facilitating to annotate pseudotime *after* bifurcation. The cell fate split appeared to initiate amongst early day 4 p.i. cells (in pseudo time, Figure 2C-D), an inference that was robust when using bootstrapped subsets of cells (Section 6.2 of Computational Supplement).

We found that differentially expressed genes between the identified trajectories were in agreement with known Th1/Tfh signature genes (15) (Figure 3A-B, and S9), strongly suggesting that the fitted mixture components corresponded to cells with Th1 and Tfh phenotypes. Notably, these bifurcation trends could not be identified by other published methods for reconstructing bifurcating single-cell trajectories (18–22) (Comp. Supp. Figure 14). We also successfully applied GPfates to resolve bifurcation events in other published datasets, (Comp. Supp. Figure 11) (23, 24) (Comp. Supp. Figure 12), suggesting that our approach is more generally applicable for studying cellular differentiation using scRNA-seq data.

Lineage barcoding using endogenous TCR sequences reveals Th1/Tfh bifurcation from single CD4$^+$ T cells

While the TCR transgenic approach employed in this study minimized the influence of TCR sequence variability on cell fate determination (6), the strain was *Rag*-sufficient, thus retaining potential for expression of diverse endogenous TCR chains in addition to the transgenic TCR. Sequence analysis of TCR transcripts in single PbTII cells confirmed universal expression of the PbTII Vα2 and Vβ12 chains (Supplementary Tables 2 & 3), as

well as highly diverse, though lower levels of expression of endogenous TCRα chains in many cells (Figure S10). Importantly, *Rag*-sufficient PbTII cells differentiated as effectively as *Rag1*$^{-/-}$ PbTII cells into both Th1 and Tfh cells (Figure S11), indicating that endogenous TCR sequences had not influenced Th fate bifurcation.

Given the vast combinatorial diversity of the endogenous TCR sequences, we employed these as unique molecular barcodes to identify ancestrally-related PbTII clones. We identified six clones comprising multiple sibling cells. Of these, two consisted of sibling cells that mapped close to the bifurcation point. For the remaining four clones, siblings exhibited highly diverging patterns of differentiation, with three sibling pairs falling at the extremities of the Th1-Tfh phenotype spectrum (Figure 3C). These results demonstrated that Th1/Tfh bifurcation had occurred at both population and at single clone levels in our system, with the progeny of a single cell populating both Th1 and Tfh compartments.

Transcriptional signatures associated with bifurcation of Th1 and Tfh fates

Next, we sought to identify genes whose expression differed between the Th1 and Tfh branches. We derived a *bifurcation statistic* to estimate the concordance with bifurcation for individual genes (see section 4.2 of the Computational Supplement text for details, Figure 3D). Among the highest-ranking genes, the most common pattern was up-regulation along the Th1 branch (Figure 3D). This suggested that Tfh cells were in fact developmentally closer to the shared progenitor state than Th1 cells, since the Th1 fate involved up-regulation of numerous genes not expressed in either the progenitor or Tfh states.

To validate the robustness of these gene signatures and the timing of the bifurcation, we repeated the infection, and at days 0, 4 and 7 sequenced additional PbTII cells using the Smart-seq2 protocol (Figure 1A & S12A). A non-linear dimensionality reduction indicated that the single cells from both experiments populated similar transcriptional landscapes (Figure S12B) and that the subset-characteristic co-expression patterns of the bifurcating genes identified by GPfates emerged by day 7 (Figure S12C). Notably, the day 7 cells from each mouse could be separated into distinct Th1 and Tfh subpopulations using the top bifurcating genes (Figure S12D). These results indicated that the bifurcation-associated gene expression patterns were reproducible across experiments and sequencing platforms.

The highest-ranking transcription factors for the bifurcation included *Tcf7* for the Tfh fate, and *Id2* for the Th1 fate (Figure 3D & 3E). *Tcf7* is required for T cell development, and has been recently shown to be instrumental for Tfh differentiation (25, 26). It also represented one of the rare genes defined by a decrease in expression when moving towards the Th1 fate. *Id2* is an antagonist of *Tcf7* and was recently identified as a key driver of Th1 responses (27). As expected, the hallmark Tfh transcription factor *Bcl6* was also strongly associated with the Tfh fate. In Th1 cells, many bifurcating genes encoded immune-related receptors (Figure 3D-E), such as *Cxcr6* (Figure S13A and S13B), *Ifngr1* and *S1pr1*, which mediate egress from secondary lymphoid organs. This was consistent with the notion that Th1 cells can migrate to peripheral tissues and remain receptive to external signals. In contrast, the only bifurcating chemokine receptor associated with a Tfh fate was *Cxcr5*, which is important for trafficking into B cell follicles (28).

Many of the bifurcating genes had no known role in Th differentiation. For example, *Igals1* (encoding Galectin-1), a molecule generally implicated in cDC (29) and Treg function (30), was unexpectedly upregulated in PbTII cells around bifurcation, and maintained along the Th1, but not the Tfh trajectory (Figure S14A). This observation was confirmed at protein level (Figure S14B). Next, comparison of Th1/Tfh fates in co-transferred WT and *Igals1*$^{-/-}$ PbTII cells during *Pc*AS infection (Figure S14C) revealed a specific role for Galectin-1 in supporting Th1 but not Tfh fate (Figure S14D). Together, these data illustrate the potential for the GPfates model to enable identification of factors controlling Th1 and Tfh fates. Further examination of bifurcating genes is facilitated by the online database, www.PlasmoTH.org, accompanying this paper (Figure 3F).

Coinciding with Th1/Tfh differentiation, we also noted up-regulation of *Il10* particularly in the Th1 branch (Figure S15A). Most of the *Il10*-expressing cells also expressed *Ifng* at equal or higher levels as those expressing *Ifng* alone (Figure S15B-C). These data revealed the development of Tr1 cells, defined as IL-10/IFNγ co-expressing CD4^{+} T cells. Given that *Il10* expression was associated with the Th1 branch, this suggested that Tr1 cells were developmentally related to Th1 cells. Surprisingly, we found that aside from *Il10*, only two genes, *Trib2* and *BC017643*, were differentially expressed between *Il10/Ifng* co-expressing Tr1 cells and *Ifng*-expressing Th1 cells (Figure S15D). Furthermore, a comparison of gene expression frequencies between Tr1 and Th1 cells, revealed a substantial degree of similarity across the genome (Figure S15E). Together these data strongly suggest that Tr1 cells derive directly from Th1 cells during blood-stage *Plasmodium* infection.

Pseudotemporal relationships between cell cycling, aerobic glycolysis and cytokine expression

Clonal expansion, increased aerobic glycolysis, and cytokine expression, are hallmarks of Th cell development whose temporal relationships with each other remain to be fully resolved *in vivo*. We noted that PbTII cells became highly proliferative around bifurcation, as shown by the upregulation of *Mki67* (Figure 4A-B and S16A) and other known proliferation marker genes (31) (Figure S16B). This correlated with cell cycle activity, as inferred from the scRNA-seq data using the Cyclone tool, and confirmed by flow cytometric measurements of DNA content and cell size (Figure 4C-D, Figure S16C). On Day 4 p.i., the cells also increased expression of genes associated with aerobic glycolysis but not oxidative phosphorylation (Figure 4F), an indication of increased metabolic requirements being met by glucose metabolism and increased mTORC1 activity. Consistent with this was the observed elevated levels of ribosomal protein S6 phosphorylation on day 4 p.i. (Figure 4E).

By day 4 p.i., PbTII cells had gone through several rounds of cell division with differing kinetics, and with some cells expressing IFNγ. By comparing *Ifng*-expressing and non-expressing cells on day 4 p.i., we noted that early *Ifng*-expressing cells cycled faster and expressed aerobic glycolysis genes more highly than non-cytokine expressing counterparts (Figure 4G). Taken together, our data suggest that around bifurcation, PbTII cells exhibited a highly proliferative and metabolically active state, with those cells cycling fastest and exhibiting most glycolytic activity being the first to acquire the capacity to secrete IFNγ.

Gene dynamics identifies potential decision-making molecules

To elucidate how PbTII cells transitioned from the proliferative precursor state to Th1 and Tfh fates we sought to resolve the hierarchy of gene expression before and during cell fate bifurcation. In addition to genes directly following the bifurcation trend, we reasoned that expression of genes encoding key decision-making molecules might also be likely to be dynamic and peak prior to the bifurcation. To identify these, we selected those genes displaying interesting non-linear trends in their expression patterns over pseudotime by Gaussian Process regression. This was achieved firstly via a *D-statistic* (model likelihood ratio) where each gene' s expression pattern over pseudotime was tested for variation unexplained by random noise (32). Based on the D-statistic (> 50.0, figure 5C) we identified 2061 dynamic genes (Figure 5A). Secondly, we ordered these genes according to their peak expression time to provide a temporal overview (Figure 5A), and noted that a substantial fraction of them peaked around bifurcation. These included the Th1-driving genes *Tbx21, Il2ra,* and *Il2rb*, supporting our initial hypothesis. Moreover, cells around bifurcation also transcribed the highest number of genes compared to those at all other points in pseudotime (Figure 5B).

This model also infers the *lengthscale* of the dynamic model, namely the degree of fast-acting behaviour over pseudotime (Figure 5C). Using this additional feature, we noted roughly equivalent dynamics for *Tbx21, Il2ra* and *Il2rb*. Furthermore, we noted similar dynamics, though with slightly later peak times, for the chemokine receptors, *Cxcr5* and *Cxcr3*. Closer examination of all chemokine receptor genes also revealed peak expression around bifurcation for *Ccr4*, but not others (Figure S17). Given that *Cxcr5* and *Cxcr3* have been associated with Tfh and Th1 cells, respectively (10, 33, 34), and since they exhibited similar dynamics, we hypothesized these were competing receptors that directly influenced Th1/Tfh fate (Figure 5D). Indeed, assessment of *Cxcr3*/*Cxcr5* co-expression around bifurcation revealed a substantial portion of cells expressing both receptors (Figure S18). Thus, our examination of gene expression dynamics revealed large numbers of genes being expressed and peaking around bifurcation, including not only those associated with clonal expansion, but also numerous sequentially-expressed transcription factors and receptors with potential to influence Th fate.

Monocytes support activated PbTII cells towards a Th1 but not a Tfh fate

Given similar dynamics and peak expression times for *Cxcr3* and *Cxcr5*, and peak expression around bifurcation for *Ccr4* (Figure S17), we reasoned that cell-cell interactions via these receptors controlled Th1/Tfh fate. Hence we considered cell-types that could control Th fate, specifically around bifurcation. Since B cells supported a Tfh fate (Figure S19), we hypothesized that co-ordinated action by myeloid cells provided competing signals to support a Th1 fate.

To study this, we examined splenic cDCs and inflammatory monocytes prior to PbTII bifurcation. We sorted $CD8\alpha^+$ and $CD11b^+$ cDCs and $Ly6C^{hi}$ monocytes from naïve and infected mice (Figure S20) and performed scRNA-seq. PCA of cDCs distinguished the two naïve cell types along PC2 (Figure 6A & S21) with an efficiency consistent with recent data (35), and further highlighted a number of expected and previously unknown cDC subset-

specific genes (Figure S22). We next compared naïve cDCs with those from infection (Figure 6A & S21), and separated these along PC6 (Figure 6A). Analysis of differential gene expression in cDCs due to infection identified 30 genes, 29 of which were upregulated (Figure 6B & S23), including transcription factors *Stat1* and *Irf1*, and CXCR3-attractant chemokines *Cxcl9* and *Cxcl10*. Notably, gene expression patterns amongst individual cDCs varied according to the gene. For example, *Stat1* and *Irf1* were expressed by several naïve cDCs, and further upregulated during infection (Figure 6C). This was similar for *Cxcl9,* which was expressed by CD8α$^+$ cDCs in naive mice, while *Cxcl10* was induced only upon infection (Figure 6C). These data suggested interactions between cDCs and uncommitted CXCR3$^+$ PbTII cells, consistent with a recent study (10). Next, PCA of Ly6C^{hi} monocytes from naïve and infected mice distinguished them from each other along PC2 (Figure 6D & S24). Differential gene expression analysis between naïve and infected groups uncovered ~100 genes, both up- and down-regulated during infection (Figure 6E & S25). A high proportion (~40%) of genes upregulated in cDCs were also induced in Ly6C^{hi} monocytes, including *Stat1, Irf1*, and *Cxcl10* (Figure 6E & F), suggesting possible overlapping functionality. In addition, monocytes expressed other chemokines, including *Cxcl2, Ccl2* and *Ccl3* (Figure 6E & F). Furthermore, specific examination of all immune cellular interaction genes (Figure S26) revealed emerging variable expression of *Tnf, Cd40, Pdl1*, *Ccl4, Ccl5, Cxcl16, Cxcl9* and *Cxcl11* in monocytes, thus suggesting complex cell-cell interactions for Ly6C^{hi} monocytes during infection.

Since *Cxcl9-11, Ccl2, Ccl3* and *Ccl5* signal through *Cxcr3* or *Ccr4*, which were expressed by activated PbTII cells, we next hypothesized that Ly6C^{hi} monocytes, in addition to cDCs (10), interacted with PbTII cells and influenced Th1/Tfh fate. To test this, we first assessed chemokine expression at protein level by Ly6C^{hi} monocytes (Figure 6G). Kinetics of CXCL9 production was similar in cDCs and Ly6C^{hi} monocytes. Next, we employed *LysM^{Cre}* x *iDTR* mice, in which Ly6C^{hi} monocytes were depleted after PbTII cell activation, but before bifurcation (Figure 6H, Figure S27). We noted a modest reduction in CD68$^+$ splenic macrophages using this approach (Figure S27B). Using this approach, we found that monocytes/macrophages supported a Th1 but not a Tfh fate (Figure 6H). Together, our data support a model in which activated, PbTII cells are supported towards either a Tfh fate by B cells (Figure S19), or a Th1 fate by chemokine-expressing myeloid cells, including Ly6C^{hi} inflammatory monocytes.

Discussion

By capturing single CD4$^+$ T cell transcriptomes during an experimental malaria infection, and computationally reconstructing the course of events, we have resolved the bifurcation of naïve CD4$^+$ T cells into Th1 and Tfh cells at a molecular level. Importantly, GPfates modelling of scRNA-seq data is not limited to immune cells or single bifurcation events. This model can also be combined with existing computational workflows, including alternative methods to estimate pseudo-temporal dynamics (19, 36) (see section 5.2 of the Computational Supplement). The GPfates approach permits analysis of cellular differentiation towards two fates (Comp. Supp. Figure 11), and in principle, towards multiple fates (Comp. Supp. Figure 12). However, GPfates exhibits some limitations. Most notably, the ability to identify and time bifurcation events is linked to changes in the

transcriptome that reflect these cellular decisions. In particular, because scRNA-seq profiles are subject to high levels of noise, this means that changes will only be detectable with some lag time (Computational Supplement). The processed expression data and the GPfates model presented in this study can be accessed at www.plasmoTH.org, where users can visualise their genes of interest.

Our data provide the framework for revealing molecular insights into the early stages of Th cell differentiation, and describe the sequence of transcriptional events before and after the bifurcation of Th1 and Tfh fates. Transcriptomic profiling previously suggested developmental similarities between Tfh and Th1 cells (37). However, highly immunogenic viral or bacterial infections induced $CD4^+$ T cells to segregate into $Bcl6^+$ (Tfh) or $Blimp\text{-}1^+$ (Th1) subpopulations within two days (38, 39). In our parasitic model, single $CD4^+$ T cell transcriptomes remained remarkably similar until four days of infection. Although it is difficult to directly compare infection models, we speculate that *Plasmodium* infection in mice may not drive Th bifurcation as early as observed with highly immunogenic viruses or bacteria, particularly given evidence of immune-suppression (40).

IL-10-producing Tr1 cells can suppress immune responses, which could aid the treatment of immune-mediated disorders (41), or be detrimental for chronic infections (42). Despite this, their relationship to Th1 cells is not clear (43). In our model, Tr1 cells emerged from the Th1 trajectory. This observation, coupled with similar transcriptomes for Th1 and Tr1 cells provides evidence that Tr1 cells are highly related to, and derive directly from Th1 cells in this model. Thus, our modelling of scRNA-seq data revealed molecular relationships between Th1, Tr1 and Tfh cells, and showed that a single naïve $CD4^+$ T cell can simultaneously give rise to more than one cell fate during experimental malaria.

Activated $CD4^+$ T cells may experience different microenvironments within secondary lymphoid tissue. The observation that bifurcation towards Th1 and Tfh fates was preceded by upregulation of chemokine receptors prompted us to investigate possible cell-cell interactions with chemokine-expressing myeloid cells. Previous studies have highlighted the potential for cDCs in lymph nodes to produce Th1-associated chemokines (10). Our study, which focused on the spleen further implicated inflammatory monocytes in Th1 support. However, since our transgenic approach for depleting monocytes also removed a small portion of splenic red pulp macrophages, we cannot fully discount the possibility that they may partly contribute to a Th1 fate. Nevertheless, we propose that splenic monocytes/macrophages influence bifurcation by supporting a Th1 fate during *Plasmodium* infection. Our studies emphasise that although cDCs are key for initiating $CD4^+$T cell activation in the spleen, other myeloid cells can also promote a Th1 fate in the presence of cDCs. In contrast, given that CXCR5 was the only chemokine receptor significantly associated with bifurcation towards a Tfh fate, cellular interaction with B-cell follicles may be the primary mechanism for supporting a Tfh fate. Our model proposes that activated, uncommitted $CD4^+$ T cells become receptive to competing chemoattractant signals from different zones of the spleen, and suggests intercellular communication as a major driver of bifurcation. However, upstream of these processes, internal stochasticity in uncommitted $CD4^+$ T cells may control the balance of chemokine receptor expression. Future experiments combining our integrated single-cell genomics and computational modelling with *in vivo* positional and trafficking

Europe PMC Funders Author Manuscripts

data may reveal molecular relationships between internal stochasticity, migratory behaviour and Th cell fate.

Materials and Methods

Study design

The goal of this study was to use scRNA-seq to capture the transcriptomes of individual splenic PbTII cells at various time points during the first week of a blood-stage *Pc*AS infection. Multiple mice were used for most time points to test for possible batch effects, with an independent experimental repeat performed on a different scRNA-seq platform. scRNA-seq data was modelled using Gaussian processes, with statistical testing for significance of both genes and cells associated with the Gaussian processes.

Experimental mice and infections

Wild-type and transgenic inbred mouse strains were housed, and employed in blood-stage *Plasmodium* infections, as described in the supplementary materials and methods.

Flow Cytometry

Splenocytes were isolated and assessed by flow cytometry as described in the supplementary materials and methods.

Single-cell mRNA sequencing

Single-cell capture and processing, as well as quality control analysis of scRNA-seq data was performed as described in the supplementary material and methods.

Statistics

Statistical analyses were conducted using R, or Python, or GraphPad prism. The types of statistical tests and significance levels are described in respective figure legends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Wellcome Trust Sanger Institute Sequencing Facility for performing Illumina sequencing, Wellcome Trust Sanger Institute Single-cell Genomics Core Facility for single-cell sample processing and the Wellcome Trust Sanger Institute Research Support Facility for care of the mice used in these studies. We thank QIMR Berghofer Flow Cytometry and Animal Facilities for expert advice and care of wild-type and transgenic mice. We wish to acknowledge Stephan Lorenz, Joanna Cartwright and Tom Metcalf for expert technical assistance. We thank Guy Emerton for constructing the database and the interface for accessing the data. We thank Michel Raymond for his work in defining cytokines and cell-surface receptors. Susanna Ng is acknowledged for assistance with graphic design.

Funding

This work was supported by Wellcome Trust (no. WT098051), European Research Council grant ThSWITCH (no. 260507), Australian National Health and Medical Research Council Project grant (number 1028641) and Career Development Fellowship (no. 1028643), University of Queensland, Australian Infectious Disease Research Centre grants and the Lister Institute for Preventative Medicine. KRJ was supported by grants from EMBL Australia and

OzEMalaR. FOB was supported by the Lundbeck Foundation. MZ was supported by the Marie Curie ITN grant "Machine Learning for Personalized Medicine" (EU FP7-PEOPLE Project Ref 316861, MLPM2012).

References and Notes

1. World malaria report 2016. World Health Organization; Geneva Switzerland: 2016.
2. Pinzon-Charry A, et al. Low doses of killed parasite in CpG elicit vigorous CD4+ T cell responses against blood-stage malaria in mice. J Clin Invest. 2010; 120:2967–2978. [PubMed: 20628205]
3. Boyle MJ, et al. Human antibodies fix complement to inhibit Plasmodium falciparum invasion of erythrocytes and are associated with protection against malaria. Immunity. 2015; 42:580–590. [PubMed: 25786180]
4. Obeng-Adjei N, et al. Circulating Th1-Cell-type Tfh cells that exhibit impaired B cell help are preferentially activated during acute malaria in children. Cell Rep. 2015; 13:425–439. [PubMed: 26440897]
5. Perez-Mazliah D, Langhorne J. CD4 T-cell subsets in malaria: TH1/TH2 revisited. Front Immunol. 2015; 5:671. [PubMed: 25628621]
6. Tubo NJ, et al. Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection. Cell. 2013; 153:785–796. [PubMed: 23663778]
7. deWalick S, et al. Cutting edge: conventional dendritic cells are the critical APC required for the induction of experimental cerebral malaria. J Immunol. 2007; 178:6033–6037. [PubMed: 17475826]
8. Celli S, Lemaître F, Bousso P. Real-time manipulation of T cell-dendritic cell interactions in vivo reveals the importance of prolonged contacts for CD4+ T cell activation. Immunity. 2007; 27:625–634. [PubMed: 17950004]
9. Bajénoff M, Wurtz O, Guerder S. Repeated antigen exposure is necessary for the differentiation, but not the initial proliferation, of naive CD4(+) T cells. J Immunol. 2002; 168:1723–1729. [PubMed: 11823503]
10. Groom JR, et al. CXCR3 chemokine receptor-ligand interactions in the lymph node optimize CD4+ T helper 1 cell differentiation. Immunity. 2012; 37:1091–1103. [PubMed: 23123063]
11. Hohl TM, et al. Inflammatory monocytes facilitate adaptive CD4 T cell responses during respiratory fungal infection. Cell Host Microbe. 2009; 6:470–481. [PubMed: 19917501]
12. Nakano H, et al. Blood-derived inflammatory dendritic cells in lymph nodes stimulate acute T helper type 1 immune responses. Nat Immunol. 2009; 10:394–402. [PubMed: 19252492]
13. Johnston RJ, et al. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. Science. 2009; 325:1006–1010. [PubMed: 19608860]
14. Szabo SJ, et al. A novel transcription factor, T-bet, directs Th1 lineage commitment. Cell. 2000; 100:655–669. [PubMed: 10761931]
15. Hale JS, et al. Distinct memory CD4+ T cells with commitment to T follicular helper- and T helper 1-cell lineages are generated after acute viral infection. Immunity. 2013; 38:805–817. [PubMed: 23583644]
16. Lawrence ND. The Gaussian Process Latent Variable Model. Technical Report. 2006
17. Lazaro-Gredilla M, Van Vaerenbergh S, Lawrence ND. Overlapping Mixtures of Gaussian Processes for the data association problem. Pattern Recognition. 2012; 45:1386–1395.
18. Marco E, et al. Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. Proc Natl Acad Sci U S A. 2014; 111:E5643–5650. [PubMed: 25512504]
19. Trapnell C, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014; 32:381–386. [PubMed: 24658644]
20. Chen J, Schlitzer A, Chakarov S, Ginhoux F, Poidinger M. Mpath maps multi-branching single-cell trajectories revealing progenitor cell progression during development. Nat Commun. 2016; 7:11988. [PubMed: 27356503]
21. Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods. 2016; 13:845–848. [PubMed: 27571553]

22. Setty M, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat Biotechnol. 2016; 34:637–645. [PubMed: 27136076]
23. Treutlein B, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature. 2014; 509:371–375. [PubMed: 24739965]
24. Guo F, et al. The transcriptome and DNA methylome landscapes of human primordial germ cells. Cell. 2015; 161:1437–1452. [PubMed: 26046443]
25. Choi YS, et al. LEF-1 and TCF-1 orchestrate T(FH) differentiation by regulating differentiation circuits upstream of the transcriptional repressor Bcl6. Nat Immunol. 2015; 16:980–990. [PubMed: 26214741]
26. Xu L, et al. The transcription factor TCF-1 initiates the differentiation of T(FH) cells during acute viral infection. Nat Immunol. 2015; 16:991–999. [PubMed: 26214740]
27. Shaw LA, et al. Id2 reinforces TH1 differentiation and inhibits E2A to repress TFH differentiation. Nat Immunol. 2016; 17:834–843. [PubMed: 27213691]
28. Hardtke S, Ohl L, Förster R. Balanced expression of CXCR5 and CCR7 on follicular T helper cells determines their transient positioning to lymph node follicles and is essential for efficient B-cell help. Blood. 2005; 106:1924–1931. [PubMed: 15899919]
29. Poncini CV, et al. Trypanosoma cruzi Infection Imparts a regulatory program in dendritic cells and T cells via galectin-1-dependent mechanisms. J Immunol. 2015; 195:3311–3324. [PubMed: 26324777]
30. Garín MI, et al. Galectin-1: a key effector of regulation mediated by CD4+CD25+ T cells. Blood. 2006; 109:2058–2065. [PubMed: 17110462]
31. Whitfield ML, George LK, Grant GD, Perou CM. Common markers of proliferation. Nat Rev Cancer. 2006; 6:99–106. [PubMed: 16491069]
32. Macaulay IC, et al. Single-cell RNA-sequencing reveals a continuous spectrum of differentiation in hematopoietic cells. Cell Rep. 2016; 14:966–977. [PubMed: 26804912]
33. Breitfeld D, et al. Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production. J Exp Med. 2000; 192:1545–1552. [PubMed: 11104797]
34. Schaerli P, et al. CXC chemokine receptor 5 expression defines follicular homing T cells with B cell helper function. J Exp Med. 2000; 192:1553–1562. [PubMed: 11104798]
35. Jaitin DA, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 2014; 343:776–779. [PubMed: 24531970]
36. Bendall SC, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell. 2014; 157:714–725. [PubMed: 24766814]
37. Liu X, et al. Bcl6 expression specifies the T follicular helper cell program in vivo. J Exp Med. 2012; 209:1841–1852. S1841–1824. [PubMed: 22987803]
38. Choi YS, et al. ICOS receptor instructs T follicular helper cell versus effector cell differentiation via induction of the transcriptional repressor Bcl6. Immunity. 2011; 34:932–946. [PubMed: 21636296]
39. Pepper M, Pagán AJ, Igyártó BZ, Taylor JJ, Jenkins MK. Opposing signals from the Bcl6 transcription factor and the interleukin-2 receptor generate T helper 1 central and effector memory cells. Immunity. 2011; 35:583–595. [PubMed: 22018468]
40. Haque A, et al. Type I IFN signaling in CD8- DCs impairs Th1-dependent malaria immunity. J Clin Invest. 2014; 124:2483–2496. [PubMed: 24789914]
41. Clemente-Casares X, et al. Expanding antigen-specific regulatory networks to treat autoimmunity. Nature. 2016; 530:434–440. [PubMed: 26886799]
42. Couper KN, et al. IL-10 from CD4CD25Foxp3CD127 adaptive regulatory T cells modulates parasite clearance and pathology during malaria infection. PLoS Pathog. 2008; 4:e1000004. [PubMed: 18401464]
43. Apetoh L, et al. The aryl hydrocarbon receptor interacts with c-Maf to promote the differentiation of type 1 regulatory T cells induced by IL-27. Nat Immunol. 2010; 11:854–861. [PubMed: 20676095]

44. Stubbington MJ, et al. An atlas of mouse CD4(+) T cell transcriptomes. Biol Direct. 2015; 10:14. [PubMed: 25886751]
45. Marshall HD, et al. Differential expression of Ly6C and T-bet distinguish effector and memory Th1 CD4(+) cell properties during viral infection. Immunity. 2011; 35:633–646. [PubMed: 22018471]
46. Santos A, Wernersson R, Jensen LJ. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. Nucleic Acids Res. 2015; 43:D1140–1144. [PubMed: 25378319]
47. Stubbington MJ, et al. T cell fate and clonality inference from single-cell transcriptomes. Nat Methods. 2016; 13:329–332. [PubMed: 26950746]
48. Flores-Langarica A, et al. T-zone localized monocyte-derived dendritic cells promote Th1 priming to Salmonella. Eur J Immunol. 2011; 41:2654–2665. [PubMed: 21630252]
49. Crawford A, et al. Molecular and transcriptional basis of $CD4^+$ T cell dysfunction during chronic infection. Immunity. 2014; 40:289–302. [PubMed: 24530057]
50. Williams, CKI., Rasmussen, CE. Gaussian Processes for machine learning. Vol. 2. MIT Press; 2006. p. 4
51. Titsias, MK., Lawrence, ND. International Conference on Artificial Intelligence and Statistics; 2010.
52. Reid JE, Wernisch L. Pseudotime estimation: deconfounding single cell time series. Bioinformatics. 2016; 32:2973–2980. [PubMed: 27318198]
53. Hensman J, Rattray M, Lawrence ND. Fast Nonparametric Clustering of Structured Time-Series. IEEE Trans Pattern Anal Mach Intell. 2015; 37:383–393. [PubMed: 26353249]
54. Hensman, J., Rattray, M., Lawrence, ND. Advances in Neural Information Processing Systems 25. Pereira, F.Burges, CJC.Bottou, L., Weinberger, KQ., editors. Curran Associates, Inc.; 2012. p. 2888-2896.
55. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. Nat Methods. 2014; 11:740–742. [PubMed: 24836921]
56. Haghverdi L, Buettner F, Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics. 2015; 31:2989–2998. [PubMed: 26002886]
57. Campbell K, Yau C. Bayesian Gaussian Process Latent Variable Models for psuedotime. 2015
58. Diaconis P, Goel S, Holmes S. Horseshoes in multidimensional scaling and local kernel methods. 2008
59. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008; 9:2579–2605.
60. Stegle O, et al. A robust Bayesian two-sample test for detecting intervals of differential gene expression in microarray time series. J Comput Biol. 2010; 17:355–367. [PubMed: 20377450]
61. Yang J, Penfold CA, Grant MR, Rattray M. Inferring the perturbation time from biological time course data. Bioinformatics. 2016; 32:2956–2964. [PubMed: 27288495]
62. Niu AL, et al. Rapid evolution and copy number variation of primate RHOXF2, an X-linked homeobox gene involved in male reproduction and possibly brain function. BMC Evol Biol. 2011; 11:298. [PubMed: 21988730]
63. Sasaki T, Shiohama A, Minoshima S, Shimizu N. Identification of eight members of the Argonaute family in the human genome. Genomics. 2003; 82:323–330. [PubMed: 12906857]
64. Owens ND, et al. Measuring absolute RNA copy numbers at high temporal resolution reveals transcriptome kinetics in development. Cell Rep. 2016; 14:632–647. [PubMed: 26774488]

Europe PMC Funders Author Manuscripts

One Sentence Summary

Using single-cell RNA sequencing and computational modelling of temporal mixtures we resolve the developmental trajectories of Th1 versus Tfh cells during experimental malaria, and identify novel T cell-intrinsic and extrinsic factors that influence fate.

Europe PMC Funders Author Manuscripts

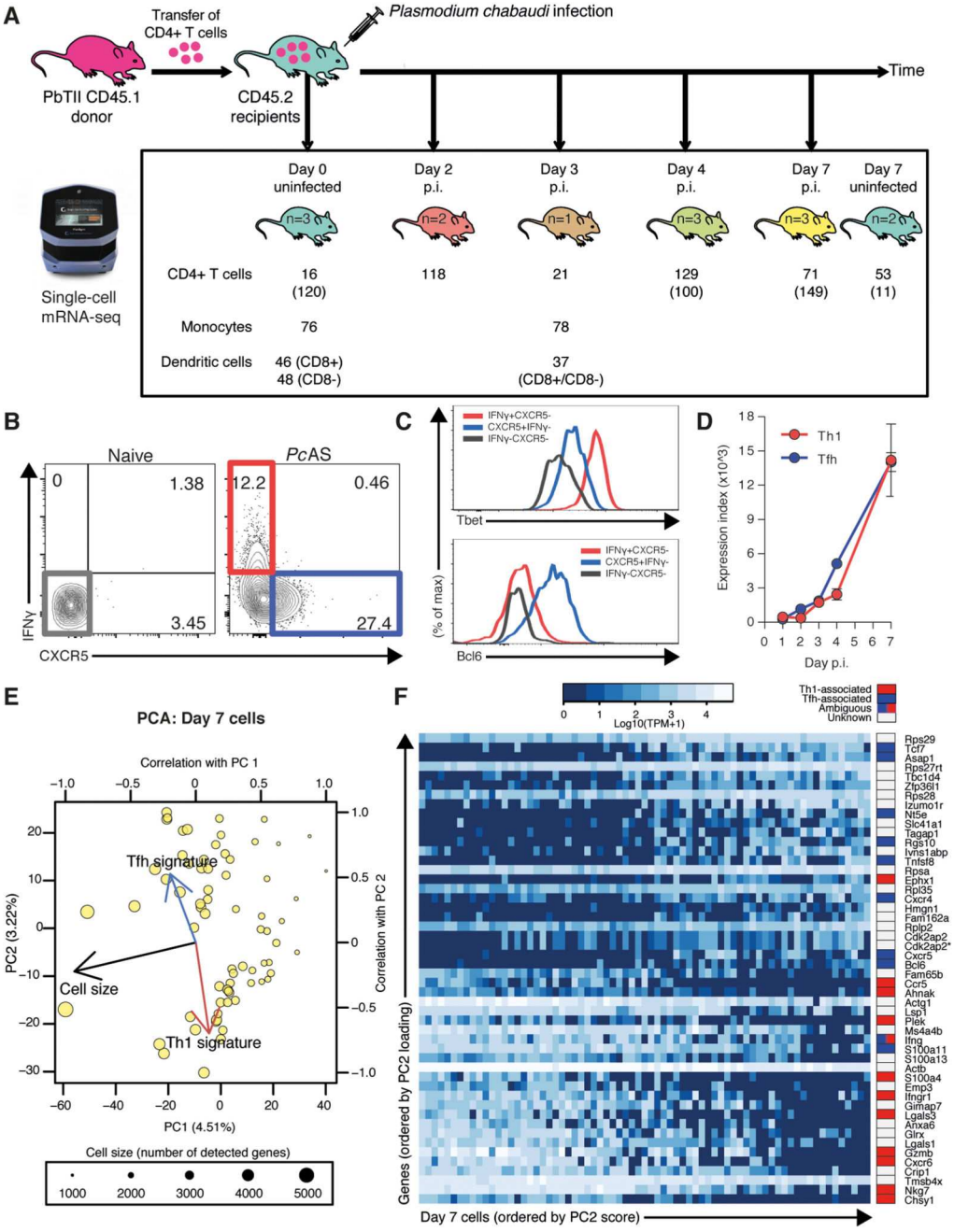

Figure 1. Single-cell mRNA-sequencing of PbTII cells.

(**A**) PbTII cells were transferred from a single donor to multiple recipients. The numbers denote single cells from which mRNA-seq data was successfully recorded. Numbers in parentheses refer to the replicate experiment presented in Fig. S12.

(**B-C**) Representative FACS plots showing bifurcation of splenic Th1 (T-bet$^+$IFNγ^+) and Tfh (Bcl6$^+$CXCR5$^+$) PbTII CD4$^+$ T cells at day 7 post-infection with *Pc*AS.

(**D**) Flow cytometry data indicate concurrent differentiation of Th1 (IFNγ^+) and Tfh (CXCR5$^+$) PbTII CD4$^+$T cells within the spleen of *Pc*AS-infected mice (n=4). Index

expression is the product of MFI (mean fluorescence intensity) and proportion IFNγ^+ or CXCR5$^+$. Data are representative of two independent experiments.

(**E**) PCA of single PbTII cells at day 7 post-infection with *Pc*AS. The arrows represent Pearson correlation with PC1 and PC2. Cell size refers to the number of detected genes. "Th1 signature" and "Tfh signature" refer to cumulative expression of genes associated with Th1 or Tfh phenotypes (total TPM of all genes in the set) (15).

(F) Expression levels of the leading 50 genes with the largest PC2 loadings at day 7 (D). Genes were annotated either as Th1- or Tfh-associated based on public datasets (15, 37, 44, 45). **Cdk2ap2* appears twice because two alternative genomic annotations exist.

Europe PMC Funders Author Manuscripts

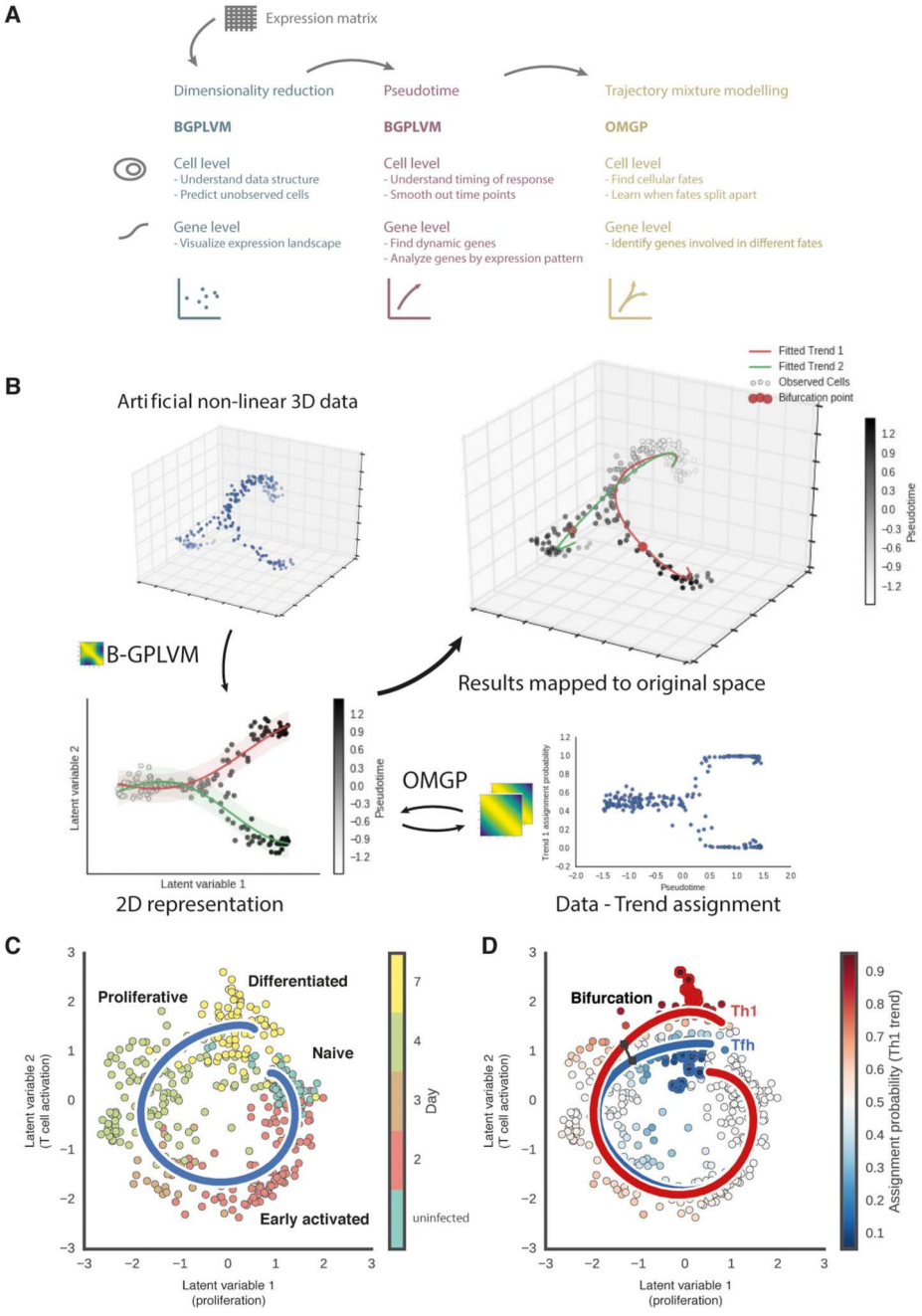

Figure 2. GPfates modelling of bifurcation processes using scRNA-seq data.

(A) Overview of the analysis workflow that underlies GPfates, consisting of dimensionality reduction of high-dimensional single-cell transcriptomes (left), inference of a pseudo temporal ordering of the cells (middle) and the reconstruction of trajectories using temporal mixture modelling (right). These individual steps build on models derived using the Gaussian process framework. Once fitted, GPfates enables for different downstream analyses, including cell orderings, bifurcation time point estimates and inference of the genes that drive bifurcation events.

(B) Illustration of intermediate results obtained from GPfates. Left: a low-dimensional representation as well as a pseudo temporal ordering of the cells is inferred using a non-linear dimensionality reduction (Gaussian process latent variable model). Temporal trajectories and bifurcations are then reconstructed using a temporal mixture model (Overlapping Mixture of Gaussian Processes), with data-trend assignments per cell. **(C)** The low-dimensional representation (2D) of the complete datasets (408 single-cell transcriptomes). The blue line depicts the inferred progression of pseudotime. Text labels illustrate features typical features of cells in the corresponding pseudotime region. **(D)** Inference of two simultaneous trends based on the pseudotime using the temporal mixture model.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

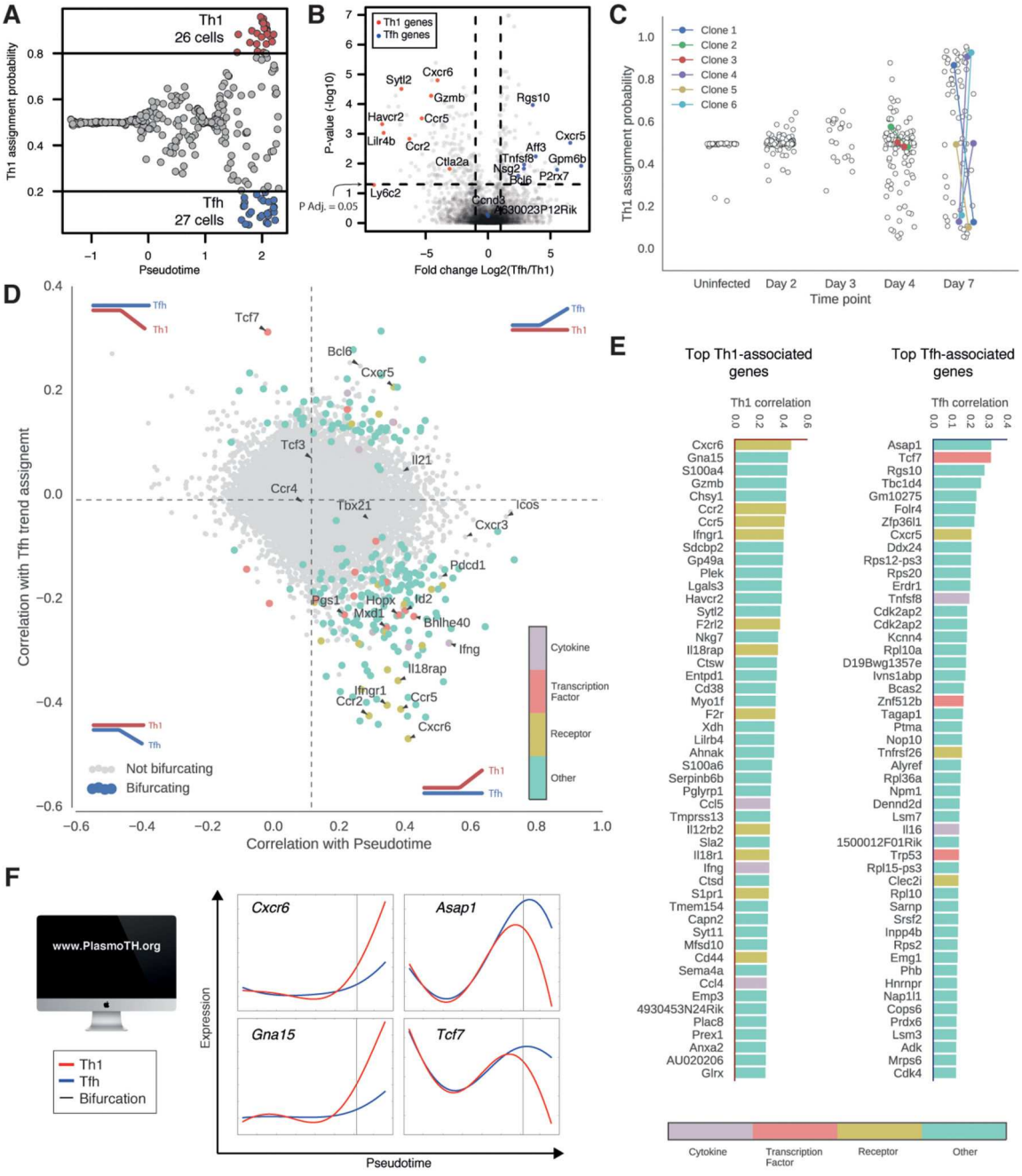

Figure 3. The relationship of known Th1 and Tfh transcriptomics signatures and the GPfates trajectories.

(**A**) Th1 and Tfh assignment probabilities of individual cells. For differential expression analysis (B), Th1 and Tfh were defined as cells with assignment probability of ≥0.8 for the respective trend.

(B) Differential expression patterns between cells assigned to Th1 and Tfh states. Shown are fold differences (x-axis) and the corresponding adjusted pvalue (y-axis) of differential expression for expressed genes (in at least 20% of cells). Statistical significance was

determined using Wilcoxon rank sum test, with Benjamini & Hochberg correction for multiple testing. The horizontal and vertical dashed lines denote adjusted p-value of 0.05 and twofold change, respectively.

(**C**) Parallel Th1 and Tfh differentiation within cells of a single $CD4^+$ T cell clone. Colors correspond to individual clones determined by sequence analysis of endogenous T cell receptor genes (Supplementary Tables 2 and 3).

(**D**) Identification of genes associated with Th1 and Tfh trajectories. For each gene, shown is the expression correlation with pseudotime (x-axis) *versus* the correlation with the Tfh trend assignment (y-axis). Gene relevance was determined using the bifurcation statistic (methods, Figure S9C). The top 248 bifurcating genes, with bifurcation statistic > 49, are represented in colors according to the functional classification of the genes (Supplementary Methods and Supplementary Table 4).

(**E**) Genes with the strongest association with Th1 or Tfh differentiation, filtered using the bifurcation score as in (D). The genes are ranked in descending order of association with the respective trend. *Cdk2ap2* appears twice because of alternative genomic annotations.

(**F**) Web address for GPfates database, where the expression kinetics of genes of interest can be visualized. Examples illustrate the top-ranking bifurcating genes from (E).

Europe PMC Funders Author Manuscripts

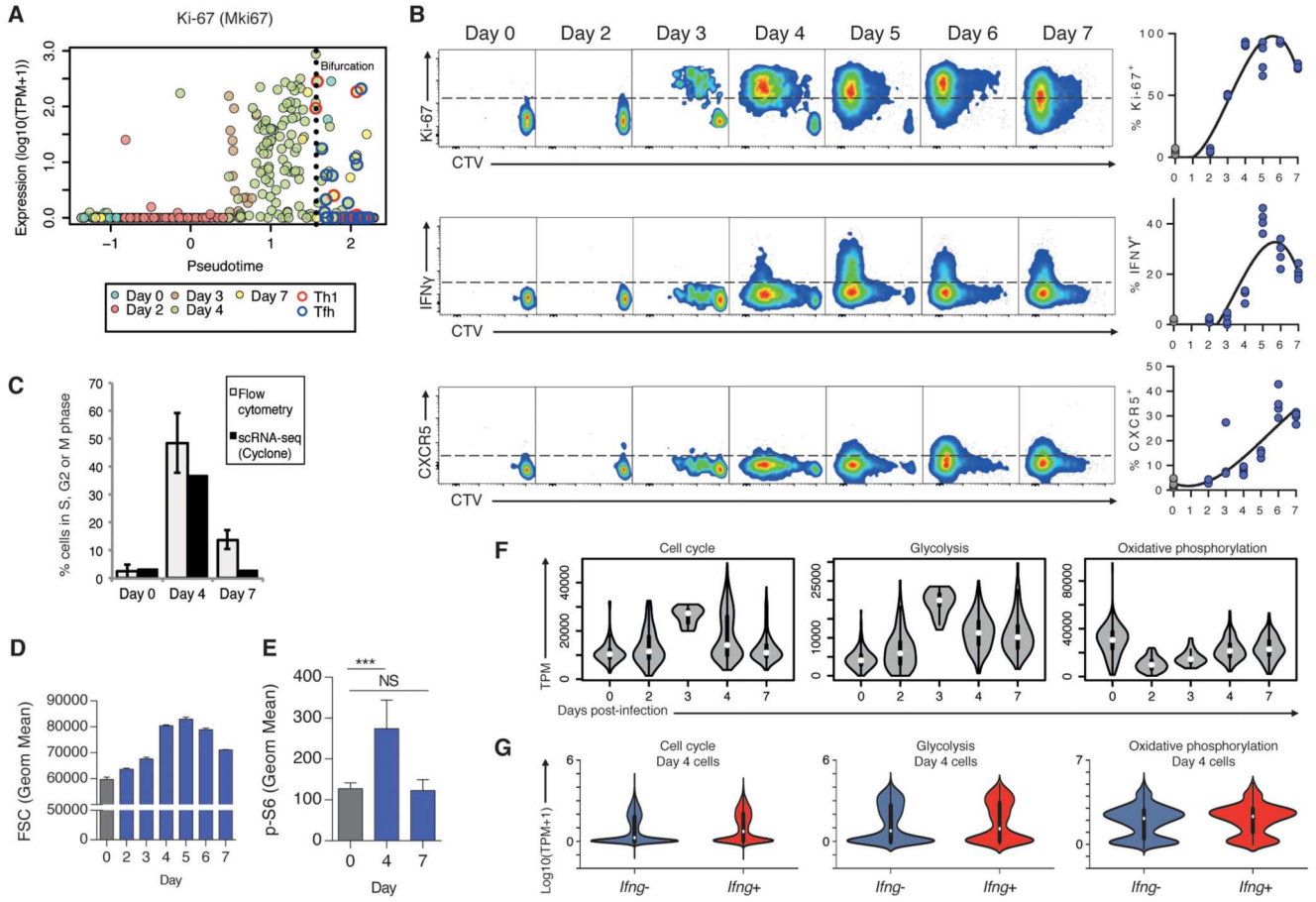

Figure 4. The bifurcation of T cell fates is accompanied by changes in transcription, proliferation and metabolism.

(**A**) The expression kinetics of *Mki67*, encoding the proliferative marker Ki-67, as a function of pseudotime.

(**B**) Representative FACS plots showing kinetics of CellTrace™ Violet (CTV) dilution and Ki67, IFNγ or CXCR5 expression, with summary graphs showing % of PbTII cells expressing these (after 10^6 PbTII cells transferred) in uninfected (Day 0) and *Pc*AS-infected mice at indicated days post-infection (n=4 mice/time point, with data from individual mice shown in summary graphs; solid line in summary graphs indicates temporal trends fit using a third order polynominal regression). Data are representative of two independent experiments.

(**C**) Relative cell-cycle speed of PbTII cells, determined by measuring the fraction of cells in S, G2, or M phases. Shown are results when allocating cells to cell cycle phases using flow cytometry (Figure S16C) or computational assignments based on the scRNA-seq data.

(**D**) Cell size estimation using FSC (Forward Scatter) measurements of PbTII cells.

(**E**) Cellular metabolic activity of PbTII cells in naive mice (n=3) and at days 4 and 7 post-infection (n=6) as determined by flow cytometric assessment of ribosomal protein S6 phosphorylation (p-S6). Histogram and proportions are representative of two independent experiments. Statistics are one-way ANOVA and Tukey's multiple comparisons tests ***p<0.001.

Europe PMC Funders Author Manuscripts

(F) Expression kinetics of genes associated with the cell cycle (251 genes derived from Cyclebase (46), glycolysis (41 genes, GO:0006096) and oxidative phosphorylation (30 genes, GO:0006119) during *Pc*AS infection. Shown are cumulative expression levels of genes in the respective categories per single cell. Data from all cells and mice (Fig. 1A) were pooled.

(G) Differential Expression analysis comparing the experiment-corrected expression of genes associated with cell cycle (p-value < 10^{-103}), glycolysis (p-value < 10^{-4}) and oxidative phosphorylation (p-value < 10^{-5}) (F) in *Ifng* positive (≥10 TPM) and *Ifng* negative cells (<10 TPM) at day 4 post infection with *Pc*AS from both experiments combined.

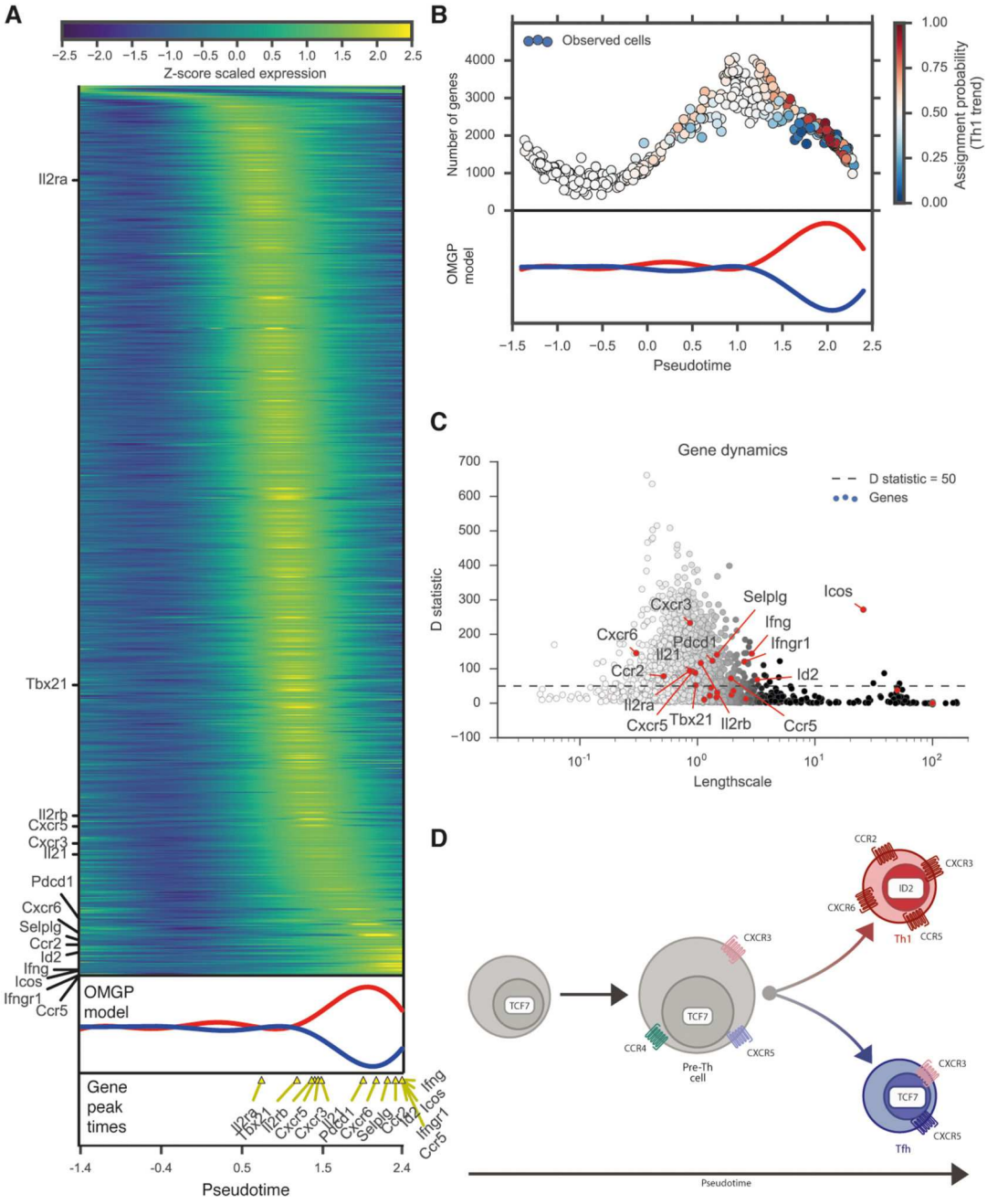

Figure 5. Temporal gene expression dynamics during *Pc*AS infection.

(A) Expression patterns over pseudotime shown for top 2061 dynamic genes, (defined by D-statistic >50). Genes are ordered per peak expression time. Th1 and Tfh probability trajectories from the GPfates OMGP model presented at the bottom to depict bifurcation and provide temporal context between the gene expressions and cellular fates. Various dynamically expressed immune receptors, transcription factors and secreted molecules are annotated.

(B) The relationship of transcriptional activity and the divergence of Th1 and Tfh fates. The number of detected genes per cell is shown across pseudotime. The color of the data points represents trend assignment probability (Fig. 2). Th1 and Tfh trajectories from the GPfates OMGP model presented to depict relation to the bifurcating behaviour.
(C) Gene expression dynamics assessed using D-statistic and optimal Squared Exponential kernel lengthscales. Genes with a D-statistic >50 selected as displaying non-linear expression patterns over pseudotime. The optimal lengthscales of the squared exponential kernels of the Gaussian Processes plotted on x-axis, where small values indicate that some rapid change in expression over pseudotime occur.
(D) A model summarizing the expression patterns of key chemokine receptors and the transcription factors *Id2* and *Tcf7* during Th1-Tfh cell fate determination. The size of the cell represents proliferative capacity (Fig. 4A-F).

Europe PMC Funders Author Manuscripts

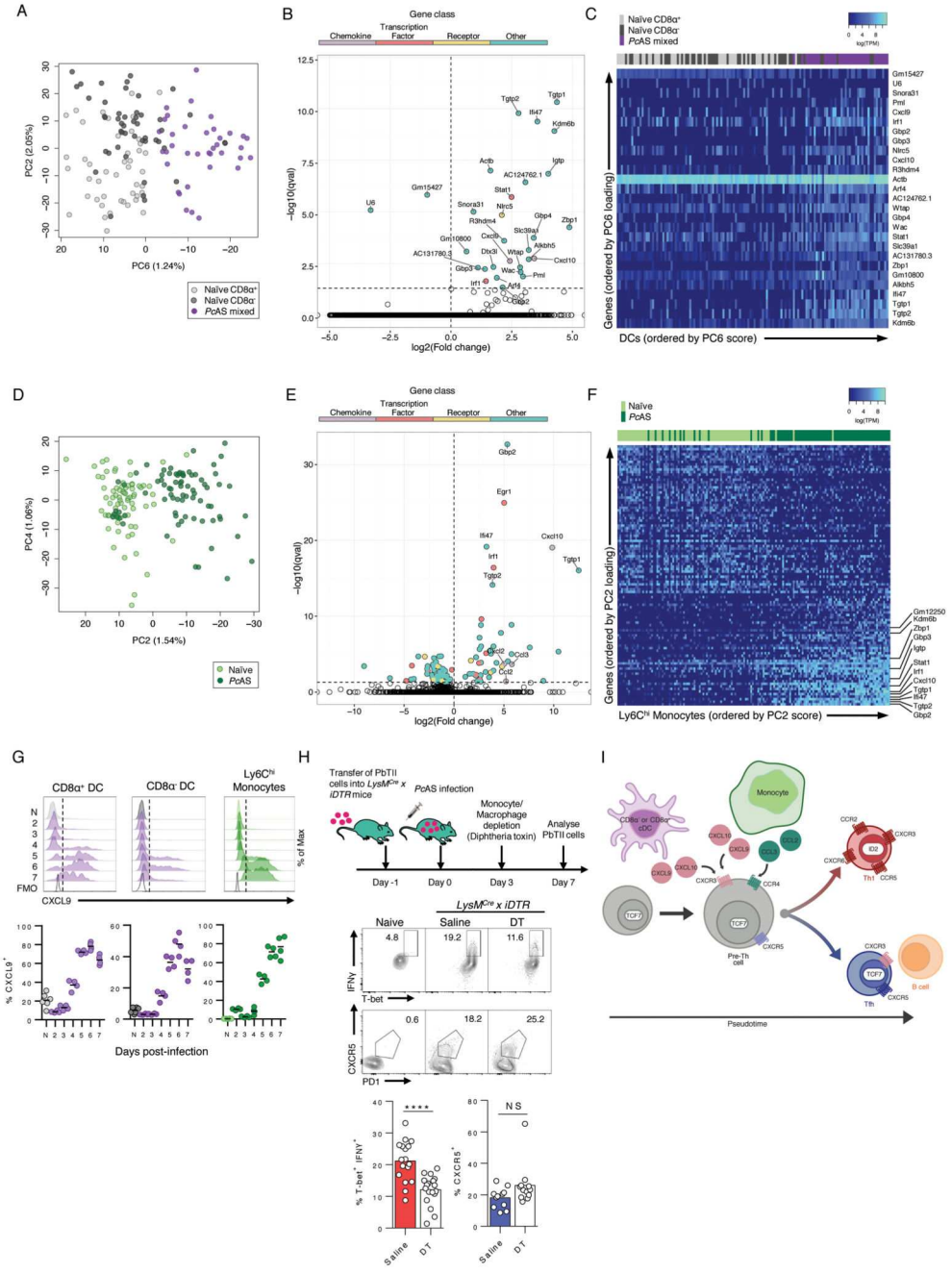

Figure 6. Myeloid cells influence Th bifurcation in uncommitted PbTII cells.

(**A-C**) Splenic CD8α$^+$ and CD11b$^+$ CD8α$^-$ cDCs from a naïve mouse, mixed cDCs from an infected mouse, and **(D-F)** Ly6C^{hi} monocytes from naïve and infected mice were analysed by scRNA-seq, with mRNA reads filtered by minimum expression of 100 TPM in at least 2 cells. (**A**& **D)** PCA showing clustering of (A) cDCs or (D) monocytes. (**B** & **E**) Fold-change and confidence for differentially expressed genes (19) between infected and naive (B) cDCs or (E) monocytes; genes filtered on expression in >10 cells; genes satisfying qval < 0.05 are colored per function. **(C**& **F)** Differentially expressed genes (qval<0.05) in (C) cDCs and (F)

monocytes, between naive and infected mice: cells and genes are ordered according to PC score and loading respectively. Common genes between heatmaps are annotated in (F). (**G**) Representative FACS histograms and proportions of splenic CD8α^{+} cDCs, CD8α^{-} cDCs and Ly6C^{hi} monocytes expressing CXCL9 in naive and infected mice; data shows individual mice with line at mean, and is representative of two independent experiments (n=4 mice/time point/experiment). (**H**) PbTII cells were transferred into *LysM^{Cre}* x *iDTR* mice 1 day prior to infection. At 3 days p.i., mice were treated with diphtheria toxin (DT) or saline. Proportions of Th1 (T-bet^{hi}IFNγ^{+}) and Tfh (CXCR5^{+}) PbTII cells at 7 days p.i.; data pooled from three independent experiments (n=5-6/experiment). Statistics: Mann-Whitney U Test. ****p<0.0001; NS, not significant. (**I**) Summary model proposing chemokine interactions between non-bifurcated PbTII cells and myeloid cells support a Th1 fate, while B cells support a Tfh fate.

Europe PMC Funders Author Manuscripts