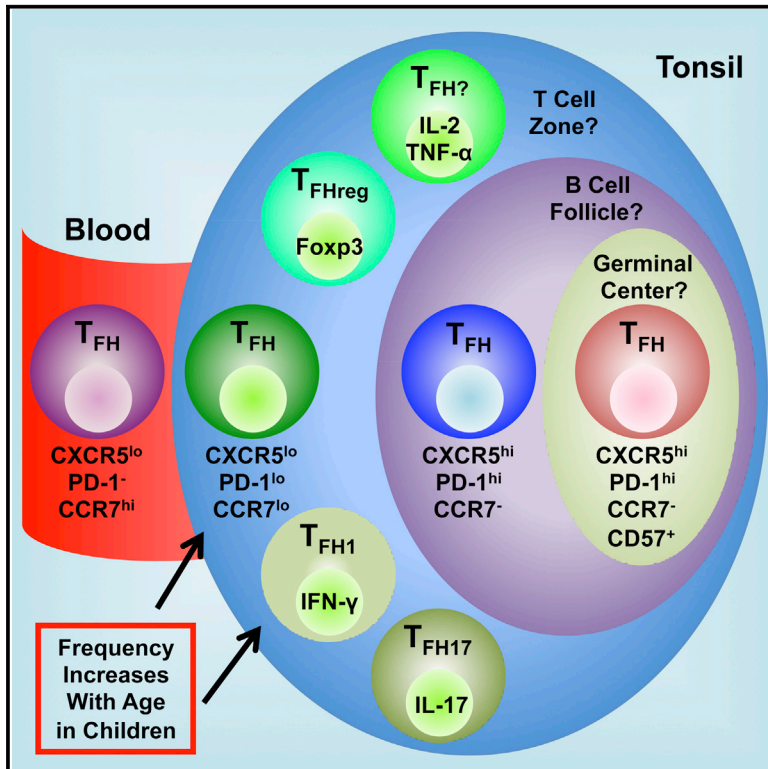


# Cell Reports

## Mapping the Diversity of Follicular Helper T Cells in Human Blood and Tonsils Using High-Dimensional Mass Cytometry Analysis

### Graphical Abstract



### Authors

Michael T. Wong, Jinmiao Chen, Sriram Narayanan, ..., Maria Alicia Curotto De Lafaille, Michael Poidinger, Evan W. Newell

### Correspondence

evan\_newell@immunol.a-star.edu.sg

### In Brief

In this study, Wong et al. use mass cytometry to comprehensively probe the phenotypic and functional profiles of human blood- and tonsil-derived CD4<sup>+</sup> T cells. This analysis uncovers multiple subtypes of T<sub>FH</sub> cells spanning blood and tonsils. Regression analysis reveals that during childhood, memory T<sub>FH</sub> cells accumulate with age.

### Highlights

- High-dimensional analysis of relationships between T cell phenotype and function
- Identification of IFN-γ<sup>+</sup> T<sub>H</sub>1<sup>-</sup>, IL-17<sup>+</sup> T<sub>H</sub>17<sup>-</sup>, and Foxp3<sup>+</sup> T<sub>reg</sub><sup>-</sup> like T<sub>FH</sub> cells
- Functional diversity of T<sub>FH</sub> cells concentrated at extrafollicular stages
- Memory T<sub>FH</sub> cell accumulation in tonsils during childhood



# Mapping the Diversity of Follicular Helper T Cells in Human Blood and Tonsils Using High-Dimensional Mass Cytometry Analysis

Michael T. Wong,<sup>1,3</sup> Jinmiao Chen,<sup>1,3</sup> Sriram Narayanan,<sup>1</sup> Wenyu Lin,<sup>1</sup> Rosslyn Anicete,<sup>2</sup> Henry Tan Kun Kiaang,<sup>2</sup> Maria Alicia Curotto De Lafaille,<sup>1</sup> Michael Poidinger,<sup>1</sup> and Evan W. Newell<sup>1,\*</sup>

<sup>1</sup>Agency for Science, Technology and Research (A\*STAR), Singapore Immunology Network (SIgN), Singapore 138648, Singapore

<sup>2</sup>KK Women's and Children's Hospital, Department of Otolaryngology, Singapore 229899, Singapore

<sup>3</sup>Co-first author

\*Correspondence: [evan\\_newell@immunol.a-star.edu.sg](mailto:evan_newell@immunol.a-star.edu.sg)

<http://dx.doi.org/10.1016/j.celrep.2015.05.022>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## SUMMARY

Single-cell analysis technologies such as mass cytometry allow for measurements of cellular heterogeneity with unprecedented dimensionality. Here, we applied dimensionality reduction and automated clustering methods on human T helper (T<sub>H</sub>) cells derived from peripheral blood and tonsils, which showed differential cell composition and extensive T<sub>H</sub> cell heterogeneity. Notably, this analysis revealed numerous subtypes of follicular helper T (T<sub>FH</sub>) cells that followed a continuum spanning both blood and tonsils. Furthermore, we identified tonsillar CXCR5<sup>lo</sup>PD-1<sup>lo</sup>CCR7<sup>lo</sup> T<sub>FH</sub> cells expressing interferon- $\gamma$  (IFN- $\gamma$ ), interleukin-17 (IL-17), or Foxp3, indicating that T<sub>FH</sub> cells exhibit diverse functional capacities within extrafollicular stages. Regression analysis demonstrated that CXCR5<sup>lo</sup>PD-1<sup>-</sup> and CXCR5<sup>lo</sup>PD-1<sup>lo</sup> cells accumulate during childhood in secondary lymphoid organs, supporting previous findings that these subsets represent memory T<sub>FH</sub> cells. This study provides an in-depth comparison of human blood and tonsillar T<sub>FH</sub> cells and outlines a general approach for subset discovery and hypothesizing of cellular progressions.

## INTRODUCTION

CD4<sup>+</sup> T helper (T<sub>H</sub>) cells orchestrate highly diverse immune responses depending on the context of infection and the cellular microenvironment. Upon stimulation by antigen-presenting cells, naive CD4<sup>+</sup> T cells can differentiate into distinct subsets of T<sub>H</sub> cells such as T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>17, T regulatory (T<sub>reg</sub>), and T follicular helper (T<sub>FH</sub>) cells (Yamane and Paul, 2013). These subsets are often defined by their expression of transcription factors that determine lineage commitment, trafficking receptors involved in tissue localization, and cytokines that mediate cellular activation and recruitment (Griffith et al., 2014; Zhu and Paul, 2010).

T<sub>FH</sub> cells are distinct from other T<sub>H</sub> subsets based on their expression of the transcription factor, Bcl-6. The chemokine receptor CXCR5 is expressed at high levels on the surface of T<sub>FH</sub> cells, which facilitates their migration into B cell follicles formed within secondary lymphoid organs such as tonsils (Cannons et al., 2013). In the follicles, these cells secrete interleukin-21 (IL-21) along with other T<sub>H</sub>-associated cytokines to aid in B cell proliferation, affinity maturation, antibody secretion, and class-switching, resulting in the formation of germinal centers and ultimately the humoral immune response. The expression of CD40 ligand (CD40L) on T<sub>FH</sub> cells also sends signals to the co-stimulatory molecule CD40 expressed on B cells, resulting in B cell activation.

Adding to the complexity of T<sub>H</sub> cell biology and differentiation, T<sub>FH</sub> cells mediate diverse responses in a variety of immunological settings. Within the T<sub>FH</sub> lineage, numerous subsets of blood T<sub>FH</sub> cells have been described based on expression of distinct surface markers such as CXCR3, CCR6, PD-1, and ICOS (Schmitt et al., 2014). These subsets are able to induce differing antibody responses when co-cultured with B cells and appear to function similarly to T<sub>FH</sub> cells found in the tonsils. However, the cytokine-secreting capacity of these subsets was not assessed at the single-cell level. Understanding the relatedness between these subsets in the blood and tonsils requires methods for simultaneously assessing the surface and intracellular markers expressed on T<sub>FH</sub> cells, along with ways to synthesize and make sense of these relationships.

The ability to analyze protein expression at the single-cell level has become invaluable for identifying numerous cellular subsets. Mass cytometry (also known as cytometry by time-of-flight, or CyTOF) combines the advantages of both flow cytometry and mass spectrometry by utilizing antibodies conjugated to metal isotopes (Bendall et al., 2012; Newell and Davis, 2014). Cellular protein expression can be measured by the unique time-of-flight pattern of the metal isotopes, which allows for the simultaneous analysis of more than 40 parameters with minimal signal overlap between channels. Recently, we utilized mass cytometry to investigate the phenotype and function of human peripheral blood CD8<sup>+</sup> T cells (Newell et al., 2012). Several studies have characterized the heterogeneity of other major cell types, including bone marrow hematopoietic cells, natural killer cells,

and myeloid cells (Becher et al., 2014; Bendall et al., 2011; Horowitz et al., 2013).

One of the current bottlenecks in mass cytometry data analysis lies in the limited number of methods that can cope with the high dimensionality of the data. To overcome this, several methods for dimensionality reduction and clustering have been used to aid in mass cytometry analysis (Amir et al., 2013; Qiu et al., 2011; Shekhar et al., 2014). We previously used principal component analysis (PCA) to study the heterogeneity of CD8<sup>+</sup> T cells, which distills the multidimensional data into a manageable number of summary variables (Newell et al., 2012). PCA is able to segregate major populations of cells and even infer the inter-relatedness between subsets. However, because PCA is a linear transformation, it is unable to segregate rare or subtly different populations of cells.

In addition to PCA, a number of alternative dimensionality reduction methods have been developed. Isometric feature mapping (ISOMAP) maps a nonlinear manifold's intrinsic geometry by measuring geodesic distances between cells, which allows for learning the data's global geometry while retaining local metric information (Amir et al., 2013; Tenenbaum et al., 2000). Working on different principles, t-distributed stochastic non-linear embedding (t-SNE) maps closely related objects to nearby points (van der Maaten and Hinton, 2008) and has been used to visually delineate cell subsets (Amir et al., 2013). Adding automation to the identification of cell subsets, automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE) performs density-based partitioning and automatically identifies cellular subpopulations on data analyzed with t-SNE (Shekhar et al., 2014). We have recently used a modified form of ACCENSE, DensVM, to analyze the mouse myeloid system (Becher et al., 2014).

Here, we apply the aforementioned dimensionality reduction and automated clustering methods to study the diversity of CD4<sup>+</sup> T<sub>H</sub> cells. In particular, we broadly compare and quantify the phenotype and function of T<sub>FH</sub> cells derived from human peripheral blood mononuclear cells (PBMCs) and tonsils. Application of t-SNE together with automated clustering led to identification of 15 subtypes of T<sub>FH</sub> cells, each composed of cells derived from blood, tonsils, or a combination of both tissues. Using ISOMAP, we hypothesize phenotypic progressions spanning both blood and tonsils. Quantifying the frequency of the identified T<sub>FH</sub> subtypes, we find that memory T<sub>FH</sub> cells accumulate with age in the tonsils during pre-adulthood.

## RESULTS

### t-SNE Segregates Known Populations of T<sub>H</sub> Cells

CD4<sup>+</sup> T<sub>H</sub> cells are extremely diverse, consisting of multiple subtypes that are derived from naive precursors. In order to capture the breadth and complexity of T<sub>H</sub> cell diversity, we stimulated cells isolated from four PBMCs and three tonsil donors with phorbol-12-myristate-13-acetate (PMA) and ionomycin in the presence of brefeldin A (BFA) and monensin for 4 hr, which would prevent the cytokine-producing potential of different T<sub>H</sub> cells. Cells were stained with a unique panel of 40 markers primarily found on T<sub>H</sub> cells, including antibodies that measured differentiation/activation status, trafficking receptors, and function (Table S1).

The staining quality of each marker was assessed using biaxial plots (Figure S1A). While we did not observe detectable IL-9 and IL-10 expression from these PBMC and tonsil samples, we confirmed IL-9 and IL-10 staining using CD4<sup>+</sup> memory T cells cultured under different T<sub>H</sub>-polarizing conditions as performed previously (Kattah et al., 2008; Wong et al., 2010) (Figure S1A). In order to detect surface markers that downregulate during stimulation, we stained select markers, primarily chemokine receptors, at 37° for 30 min prior to stimulation (see Experimental Procedures and Table S1). Using this method, we found that the frequency of these surface markers does not change substantially with stimulation, regardless of using freshly isolated or cryopreserved/thawed cells (Figure S2A).

After gating on CD4<sup>+</sup> cells and merging these cells from all PBMC and tonsil donors, we applied t-SNE on the CD4<sup>+</sup> dataset. We gated and overlaid naive (CD45RA<sup>+</sup>CCR7<sup>+</sup>CD45RO<sup>-</sup>), T<sub>H</sub>1 (interferon-γ [IFN-γ]<sup>+</sup>, T<sub>H</sub>17 (IL-17A<sup>+</sup>), and T<sub>FH</sub> (CXCR5<sup>hi</sup>PD-1<sup>hi</sup>) cells onto the first two dimensions of t-SNE (Figures 1A and S2B). Each subset occupied distinct regions in the t-SNE plot, indicating that t-SNE is able to segregate known populations of T<sub>H</sub> cells. However, the overlaid cells did not account for all of the cells in the t-SNE analysis.

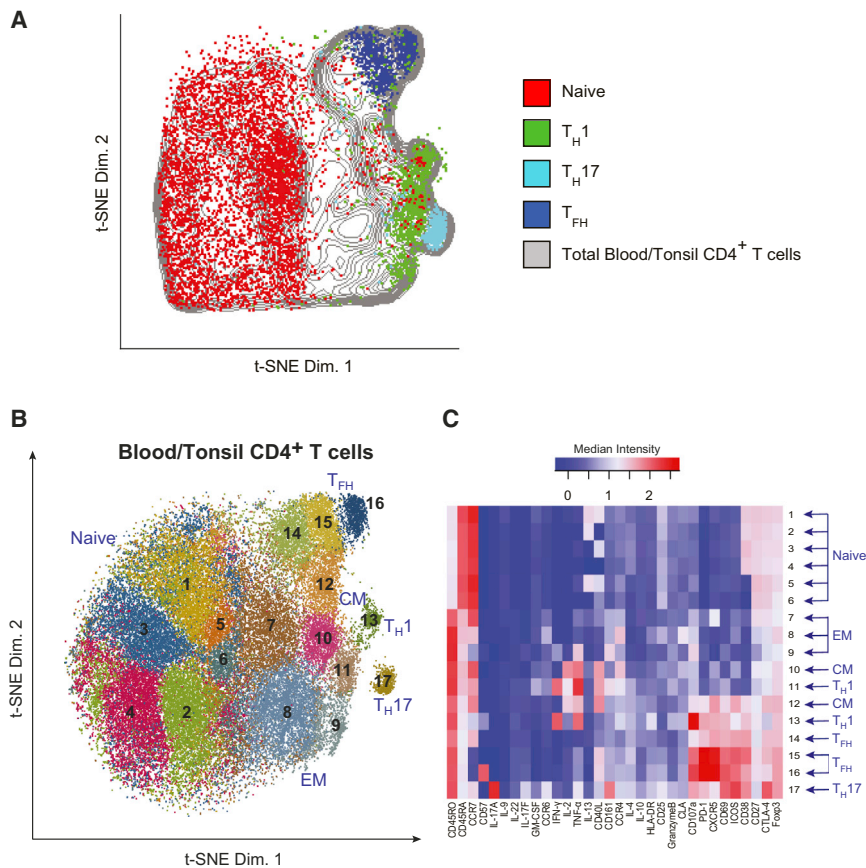
### Human CD4<sup>+</sup> T<sub>H</sub> Cells Are Extremely Diverse across Blood and Tonsils

In order to automatically detect the unidentified t-SNE clusters in an unbiased fashion, we utilized the clustering algorithm DensVM on the CD4<sup>+</sup> dataset (Figure 1B). DensVM is a modified version of ACCENSE (Shekhar et al., 2014) (Figure S3A), which we used previously for automated clustering and population boundary detection of mouse myeloid cells (Becher et al., 2014). After identification of clusters by DensVM, we used heatmaps to access the phenotypic profile of each cluster, as well as annotated the clusters based on reported subsets of T<sub>H</sub> cells (Figure 1C).

DensVM classified six clusters (clusters 1–6) of naive T cells that expressed CD45RA and CCR7 (Figure 1B). These six clusters showed variations in the expression of CD40L and CD38 (Figure 1C). Although naive T cells are commonly considered as a homogenous population, these data supported previous studies reporting naive cell heterogeneity (Alves et al., 2007; Bains et al., 2013; Jameson and Fulton, 2011). In addition to segregating T<sub>CM</sub> (clusters 10 and 12) and T<sub>EM</sub> (clusters 7–9) cells, DensVM also identified one cluster consisting of T<sub>H</sub>17 and T<sub>reg</sub> cells (cluster 17), three clusters of T<sub>FH</sub> cells (clusters 14–16), and two clusters of T<sub>H</sub>1 cells (clusters 11 and 13) (Figure 1C). These data indicated a remarkable amount of CD4<sup>+</sup> T cell diversity across blood and tonsils.

### Comparative Analysis of PBMC- and Tonsil-Derived CD4<sup>+</sup> T<sub>H</sub> Cells

Several studies have characterized the phenotype and function of CD4<sup>+</sup> T<sub>H</sub> cells from human peripheral blood and tonsils, including T<sub>H</sub>17, T<sub>reg</sub>, and CXCR5<sup>+</sup> T<sub>FH</sub> cells (Lim et al., 2008; Schmitt et al., 2014; Voo et al., 2009). To further investigate the relationships between CD4<sup>+</sup> T<sub>H</sub> cells in the blood and tonsils, we applied t-SNE and DensVM on the merged dataset containing four PBMC and three tonsil samples, wherein each sample



### Figure 1. Human CD4<sup>+</sup> T<sub>H</sub> Cells Are Highly Diverse across Peripheral Blood and Tonsils

Cells isolated from four PBMC and three tonsil donors were stimulated with PMA and ionomycin in the presence of brefeldin A and monensin. After 4 hr, cells were stained with a panel of 40 surface or intracellular markers (see Table S1).

(A) t-SNE was performed after pooling the PBMC and tonsil samples and gating on CD4<sup>+</sup> cells. The indicated cellular subsets (colored legends) were manually gated and overlaid onto the total CD4<sup>+</sup> population (gray) (see Figures S1 and S2).

(B) The numbered and colored clusters, corresponding to clusters identified by DensVM (see Figure S3), were overlaid onto the t-SNE analysis. Where indicated, clusters were annotated with the subset name.

(C) Heatmap shows hierarchical clustering of median surface and intracellular marker expression levels of the indicated clusters identified by DensVM. Scale values of median intensity indicate the number of ion counts after logicle transformation, and range from low (blue) to high (red) expression. Each cluster is annotated with the indicated subset names (blue arrows). CM, central memory; EM, effector memory.

### ISOMAP Reproducibly Represents T<sub>H</sub>1 and T<sub>FH</sub> Cells in a Phenotypic Continuum across Blood and Tonsils

Despite the power of using t-SNE and DensVM in identifying novel and known

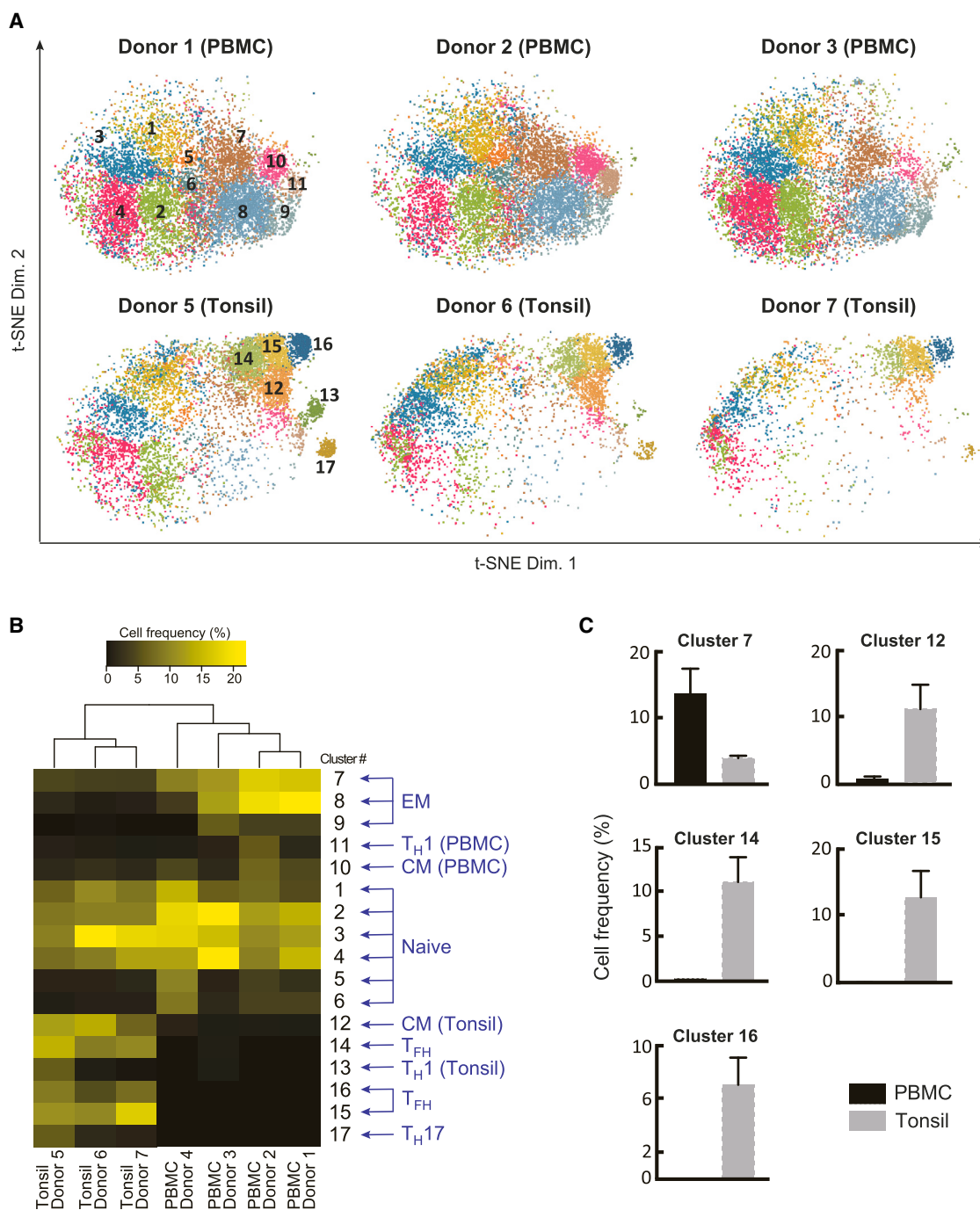
populations of CD4<sup>+</sup> T cells, these analysis tools were inadequate for mapping potential relationships between cell subsets. ISOMAP may be more appropriate for this task because it takes into account local distances for similar cells while retaining the global geometry between different cell types (Tenenbaum et al., 2000). When visualizing either T<sub>FH</sub>-related marker expression (CXCR5, PD-1, and CD57) or cytokine expression (tumor necrosis factor  $\alpha$  [TNF- $\alpha$ ], IL-2, and IFN- $\gamma$ ) on the ISOMAP generated for blood and tonsillar T<sub>H</sub> cells, we observed gradients of expression for each marker along both dimensions (Figure 3A).

We then visualized the clusters identified by t-SNE/DensVM in the ISOMAP and numbered each cluster according to Figure 1 (Figure 3B). Using the expression of the six markers from Figure 3A, we drew three hypothesized paths of cellular progressions, each seemingly derived from naive T cells (clusters 1–6) (Figure 3B). One progression, consisting of clusters 14, 17, 15, and 16, resolved different stages of progression for tonsillar T<sub>FH</sub> cells. In this progression, CXCR5<sup>lo</sup>PD-1<sup>lo</sup>CD57<sup>-</sup> (clusters 14 and 17) cells were connected with CXCR5<sup>hi</sup>PD-1<sup>hi</sup>CD57<sup>-</sup> (cluster 15) cells, which were connected to CXCR5<sup>hi</sup>PD-1<sup>hi</sup>CD57<sup>+</sup> T<sub>FH</sub> cells (cluster 16). The final stage of the ISOMAP progression was the expression of CD57 by CXCR5<sup>hi</sup>PD-1<sup>hi</sup> T<sub>FH</sub> cells, which have been reported to be exclusively localized to the germinal centers (Kim et al., 2001). Another continuum of cells connecting clusters 8, 9, 10, and 11 described various subtypes of T<sub>H</sub>1-like cells found in the peripheral blood (Figure 3B). Along this phenotypic progression,

populations of CD4<sup>+</sup> T cells, these analysis tools were inadequate for mapping potential relationships between cell subsets. ISOMAP may be more appropriate for this task because it takes into account local distances for similar cells while retaining the global geometry between different cell types (Tenenbaum et al., 2000). When visualizing either T<sub>FH</sub>-related marker expression (CXCR5, PD-1, and CD57) or cytokine expression (tumor necrosis factor  $\alpha$  [TNF- $\alpha$ ], IL-2, and IFN- $\gamma$ ) on the ISOMAP generated for blood and tonsillar T<sub>H</sub> cells, we observed gradients of expression for each marker along both dimensions (Figure 3A).

We then visualized the clusters identified by t-SNE/DensVM in the ISOMAP and numbered each cluster according to Figure 1 (Figure 3B). Using the expression of the six markers from Figure 3A, we drew three hypothesized paths of cellular progressions, each seemingly derived from naive T cells (clusters 1–6) (Figure 3B). One progression, consisting of clusters 14, 17, 15, and 16, resolved different stages of progression for tonsillar T<sub>FH</sub> cells. In this progression, CXCR5<sup>lo</sup>PD-1<sup>lo</sup>CD57<sup>-</sup> (clusters 14 and 17) cells were connected with CXCR5<sup>hi</sup>PD-1<sup>hi</sup>CD57<sup>-</sup> (cluster 15) cells, which were connected to CXCR5<sup>hi</sup>PD-1<sup>hi</sup>CD57<sup>+</sup> T<sub>FH</sub> cells (cluster 16). The final stage of the ISOMAP progression was the expression of CD57 by CXCR5<sup>hi</sup>PD-1<sup>hi</sup> T<sub>FH</sub> cells, which have been reported to be exclusively localized to the germinal centers (Kim et al., 2001).

Another continuum of cells connecting clusters 8, 9, 10, and 11 described various subtypes of T<sub>H</sub>1-like cells found in the peripheral blood (Figure 3B). Along this phenotypic progression,



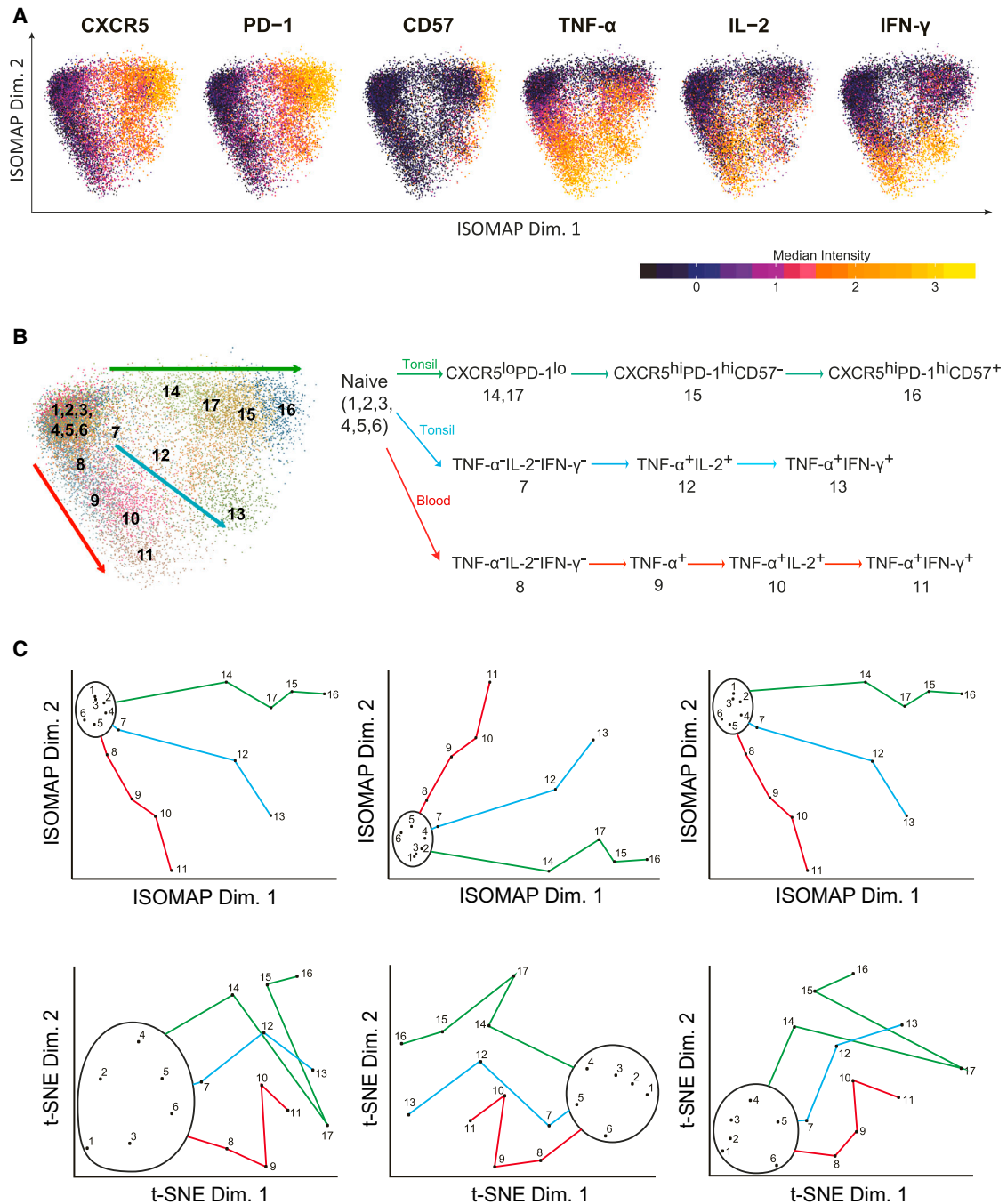
**Figure 2. Comparative Analysis of Human CD4<sup>+</sup> T Cells from Peripheral Blood and Tonsils**

t-SNE was performed on CD4<sup>+</sup> T cells merged from four PBMC and three tonsil donors.

(A) Individual donor CD4<sup>+</sup> T cells from PBMCs or tonsils were separated using barcodes and visualized using the t-SNE map generated from the merged dataset. Three of four PBMC donors are shown. The numbered and colored clusters correspond to cell populations identified by DensVM in Figure 1.

(B) Heatmap shows hierarchical clustering of the indicated cluster frequencies within CD4<sup>+</sup> cells from the individual PBMC and tonsil donors. Scale values indicate cell frequency from low (black) to high (yellow).

(C) Bar graphs show average frequency of the indicated clusters from PBMC (black bar) and tonsil (gray bar) CD4<sup>+</sup> cells. All cluster frequencies shown are significantly different ( $p < 0.05$ ) between PBMCs and tonsils.



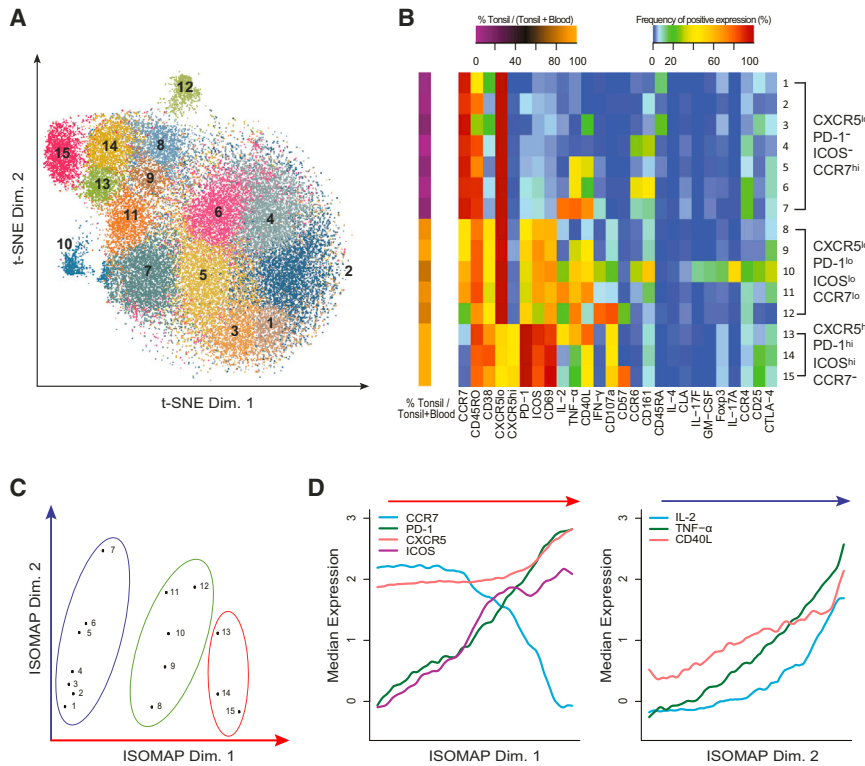
### Figure 3. ISOMAP Analysis Suggests Phenotypic Progression of Human CD4<sup>+</sup> T Cells in Peripheral Blood and Tonsils

ISOMAP and t-SNE were performed on CD4<sup>+</sup> T cells merged from four PBMC and three tonsil donors (see Figure S3).

(A) Plots show median expression of the indicated surface or intracellular markers in the ISOMAP analysis. Scale values of median intensity indicate the number of ion counts after logicle transformation, and range from low (violet) to high (yellow).

(B) Each cluster identified by DensVM in Figure 1 is numbered and overlaid onto the ISOMAP analysis (left panel). A schematic (right panel) is shown annotating each cluster with expression of the indicated markers. Green (tonsil-derived), blue (tonsil-derived), and red (blood-derived) lines represent three distinct hypothesized paths of phenotypic progression. CD45RA<sup>+</sup>CCR7<sup>+</sup>CD45RO<sup>-</sup> naive cells (clusters 1–6) represent the beginning of each pathway.

(C) Shown is three replicate ISOMAP (top three panels) and t-SNE (bottom three panels) analysis plots. The median coordinates of each cluster identified by DensVM are represented as numbered dots in the ISOMAP and t-SNE plots. Naive cells (circled clusters 1–6) represent the beginning of each pathway.



**Figure 4. Analysis of CXCR5<sup>lo</sup> and CXCR5<sup>hi</sup> Cells Reveals Distinct Subtypes of Peripheral Blood and Tonsillar T<sub>FH</sub> Cells**

t-SNE and ISOMAP were performed on CD4<sup>+</sup> CXCR5<sup>lo</sup> and CD4<sup>+</sup> CXCR5<sup>hi</sup> T cells combined from four PBMC and three tonsil donors.

(A) The numbered and colored clusters, corresponding to clusters identified by DensVM, were overlaid onto the t-SNE analysis.

(B) Heatmap shows hierarchical clustering of the indicated marker frequencies in each of the 15 clusters identified by DensVM (see Figure S4), with three major subtypes of T<sub>FH</sub> cells annotated. Violet to orange scale represents the ratio of cluster frequency in tonsil to the sum of the cluster frequency in tonsil and blood (see Figure S5). Rainbow scale indicates percent range from low to high.

(C) The medians of each cluster identified by DensVM are represented as numbered dots in the ISOMAP analysis, with the circled clusters representing the three major subtypes of T<sub>FH</sub> cells annotated in (B).

(D) Median expression levels (colored lines) of the indicated surface markers along the first ISOMAP dimension (left panel) and the indicated intracellular markers along the second ISOMAP dimension (right panel) are shown.

TNF- $\alpha$ <sup>-</sup>IL-2<sup>-</sup>IFN- $\gamma$ <sup>-</sup> cells (cluster 8) or cells producing only TNF- $\alpha$  (cluster 9) were closely associated with naive cells, which were followed by cells that produce TNF- $\alpha$  and IL-2 (cluster 10). The end of this continuum contained cells capable of secreting TNF- $\alpha$ , IFN- $\gamma$ , and reduced levels of IL-2 (cluster 11). Comparable to T<sub>H</sub>1 cells in blood, a third pathway found primarily in the tonsils revealed a transition from TNF- $\alpha$ <sup>-</sup>IL-2<sup>-</sup>IFN- $\gamma$ <sup>-</sup> (cluster 7) to TNF- $\alpha$ <sup>+</sup>IL-2<sup>+</sup> (cluster 12) to TNF- $\alpha$ <sup>+</sup>IFN- $\gamma$ <sup>+</sup> (cluster 13) cells, implying the progression of tonsillar CXCR5<sup>lo</sup> T<sub>H</sub>1 cells.

We assessed the reproducibility of our analysis by running ISOMAP on three independently subsampled subsets of the same data, which resulted in reliable maps that preserved spatial connectivity and accurately represented the phenotypic progressions between different clusters as described above (Figure 3C). In contrast, t-SNE's geometry was not consistent across the three subsamples (Figure 3C). This is consistent with ISOMAP's primary focus on modeling large geodesic distances rather than small ones (Tenenbaum et al., 2000), while t-SNE focuses on preserving local geometry (van der Maaten and Hinton, 2008). These data not only demonstrate the individual strengths and weaknesses of using t-SNE or ISOMAP for high-dimensional analysis but also illustrate the utility of combining these techniques for subset discovery and generating hypotheses about relationships between different cell types.

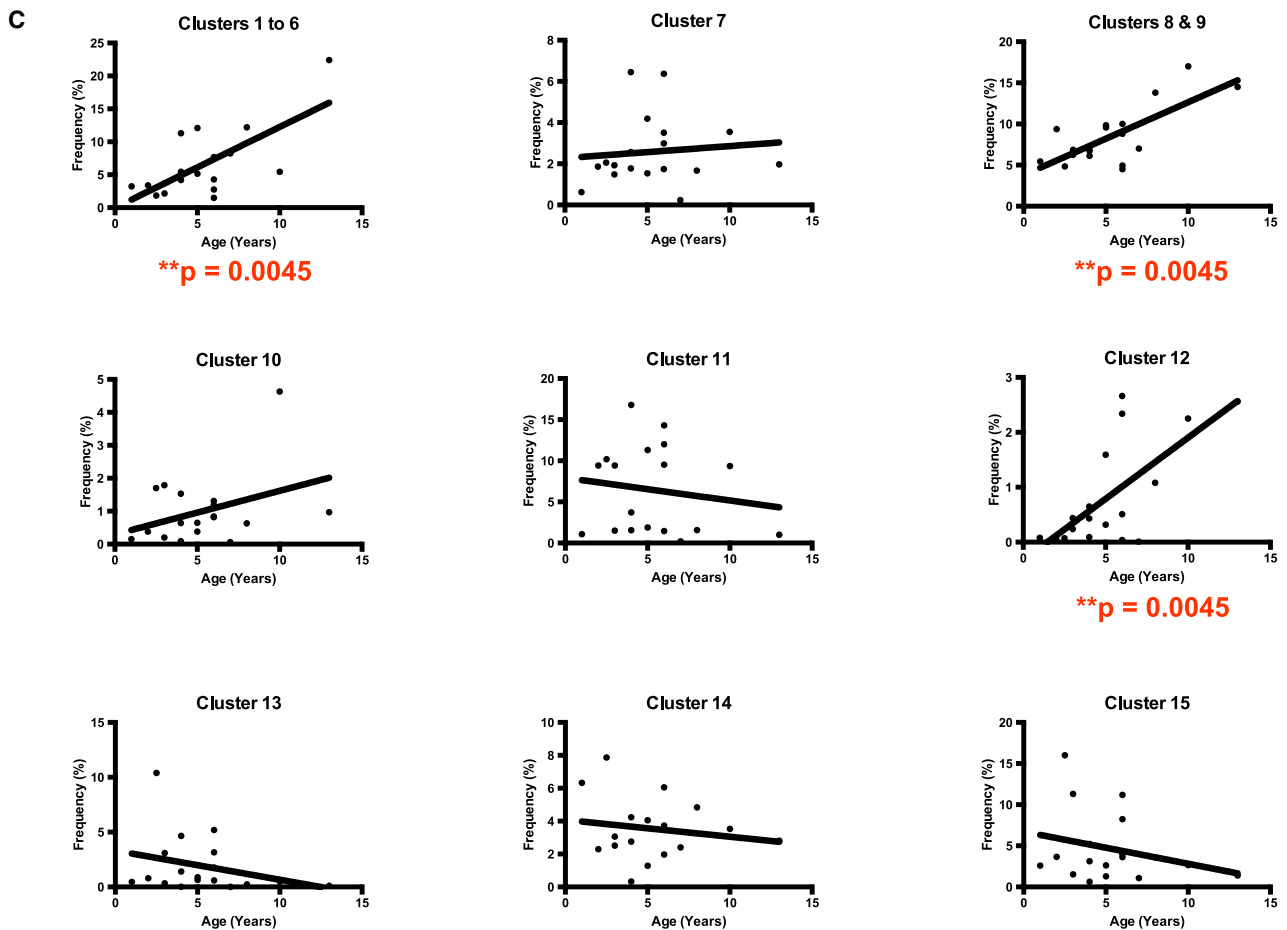
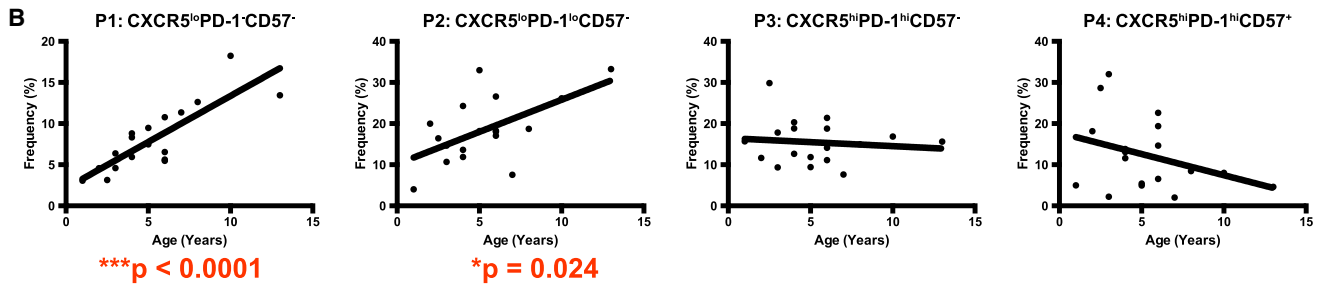
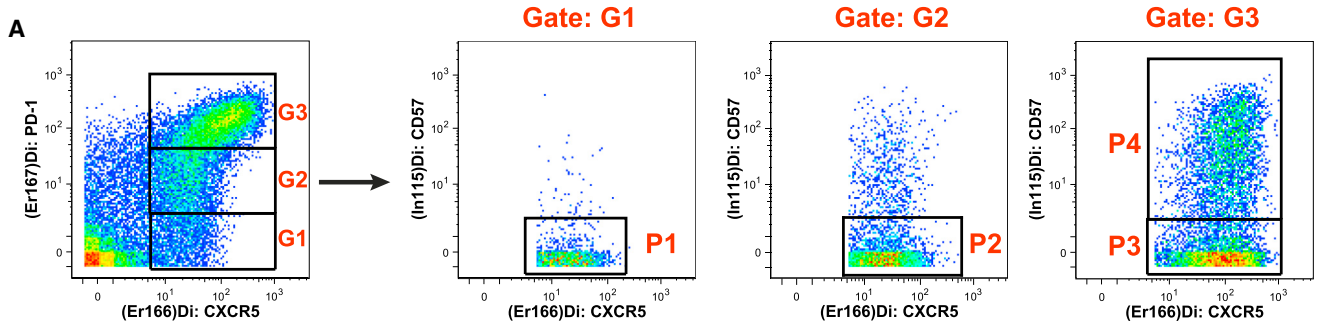
### High-Dimensional Analysis of CXCR5<sup>lo</sup> and CXCR5<sup>hi</sup> Cells Reveals Three Major Subtypes of T<sub>FH</sub> Cells Spanning Blood and Tonsils

Because few studies have simultaneously compared the phenotype and relationship between T<sub>FH</sub> cells from peripheral blood

and secondary lymphoid tissue, we applied the aforementioned analysis methods (Figure S3B) on CD4<sup>+</sup> CXCR5<sup>+</sup> T<sub>FH</sub>-like cells combined from the four PBMC and three tonsil donors described above. After gating on all CXCR5<sup>lo</sup> and CXCR5<sup>hi</sup> cells (Figure S4), DensVM identified 15 distinct clusters of T<sub>FH</sub>-like cells within the t-SNE map (Figure 4A). Using manual gating on biaxial plots, we then generated a heatmap that quantified the frequency of positively expressed markers within each DensVM cluster (Figures 4B and S4). Hierarchical clustering revealed three major types of T<sub>FH</sub> cells: (1) CXCR5<sup>lo</sup>PD-1<sup>-</sup>ICOS<sup>-</sup>CCR7<sup>hi</sup> (clusters 1–7), (2) CXCR5<sup>lo</sup>PD-1<sup>lo</sup>ICOS<sup>lo</sup>CCR7<sup>lo</sup> (clusters 8–12), and (3) CXCR5<sup>hi</sup>PD-1<sup>hi</sup>ICOS<sup>hi</sup>CCR7<sup>-</sup> (clusters 13–15).

The first population type was primarily found in the peripheral blood (Figure S5) and was likely segregated based on surface expression of CCR4, CCR6, and CD161, as well as functional expression of TNF- $\alpha$ , IL-2, and CD40L (Figure 4B). Notably, clusters 4, 6, and 7 contained cells that were CCR4<sup>+</sup>CCR6<sup>+</sup>CD161<sup>+</sup>, which are surface markers typically expressed on T<sub>H</sub>17 cells (Acosta-Rodriguez et al., 2007; Cosmi et al., 2008). However, in our data, these cells were not capable of producing IL-17 upon ex vivo stimulation with PMA and ionomycin. This illustrated that the use of chemokine receptor expression as proxies for functional differentiation is not always accurate.

The second major type of T<sub>FH</sub> cells was found primarily in tonsils (Figure S5) and represented an effector memory phenotype based on reduced expression of CCR7 (Figure 4B). Within this population, we identified bona fide T<sub>H</sub>1-like (cluster 12) and T<sub>H</sub>17-like (cluster 10) T<sub>FH</sub> cells that expressed IFN- $\gamma$  and IL-17A/F, respectively. The T<sub>H</sub>17-like T<sub>FH</sub> cells (cluster 10) also expressed the canonical T<sub>H</sub>17 markers CCR4, CCR6 and CD161.



(legend on next page)



Furthermore, cluster 10 contained a small population of Foxp3<sup>+</sup> T<sub>reg</sub> cells, suggesting that T<sub>H</sub>17<sup>-</sup> and T<sub>reg</sub>-like T<sub>FH</sub> cells are phenotypically similar based on their close proximity in the t-SNE map. This is consistent with T<sub>H</sub>17 cells being developmentally related to T<sub>reg</sub> cells (Zhou et al., 2008).

The third major type of cells represented bona fide T<sub>FH</sub> cells based on the highest expression of CXCR5, PD-1, and ICOS, and this group was derived exclusively from the tonsils (Figure S5). Based on the absence of CCR7 and expression of CD40L, these cells were likely capable of helping B cells in the follicles (Cannons et al., 2013; Schmitt et al., 2014). One cluster (cluster 15) within this group represented a population of CD57<sup>+</sup>CD107a<sup>+</sup> late effector/killer-like T<sub>FH</sub> cells. These data collectively demonstrate the abundant diversity of CXCR5<sup>+</sup> T<sub>FH</sub> cells in both the peripheral blood and secondary lymphoid organs.

In order to investigate the relatedness between the numerous subtypes of T<sub>FH</sub> cells, we applied ISOMAP to CXCR5<sup>+</sup> cells merged from the aforementioned four PBMC and three tonsil donors. We then annotated the center of each DensVM cluster on the first two dimensions of ISOMAP and circled the three populations of T<sub>FH</sub> cells found with hierarchical clustering (Figure 4C). Along the first ISOMAP dimension, we observed that as the expression of CXCR5, PD-1, and ICOS simultaneously increases, there is a concurrent decrease in CCR7 expression (Figure 4D). These results suggested that circulating CXCR5<sup>+</sup> cells might be closely related to bona fide CXCR5<sup>hi</sup> T<sub>FH</sub> cells in the secondary lymphoid organs, perhaps representing their counterparts.

Along the second ISOMAP dimension, we observed a gradual increase in functional capacity based on the expression of IL-2, TNF- $\alpha$ , and CD40L (Figure 4D). In the peripheral blood, the functional capacity of T<sub>FH</sub> cells rose with increasing expression of CXCR5, PD-1, and ICOS. However, tonsillar T<sub>FH</sub> cells lost functional capacity to produce these cytokines in conjunction with increasing expression of CXCR5, PD-1, and ICOS. Overall, these data revealed an unprecedented amount of diversity within T<sub>FH</sub> cells.

### Memory CXCR5<sup>+</sup> T<sub>FH</sub> Cells Accumulate with Age in Tonsils

The majority of tonsillectomies are performed prior to adulthood, allowing us to assess T<sub>FH</sub> diversity from childhood to pre-adulthood. Using the data generated by t-SNE/DensVM/ISOMAP, we analyzed the accumulation of different T<sub>FH</sub> populations in 18 tonsil sample donors that required tonsillectomies due to tonsillar hypertrophy or sleep apnea (Table S2). Included in these 18 tonsil samples were the three samples used in the aforementioned t-SNE analysis, as well as 15 additional samples that were stained in separate experiments. Because batch effects and minor modifications to the staining panels made it impossible to perform t-SNE on all the samples simultaneously, we used

manual or Boolean gating on all 18 tonsil samples to identify the T<sub>FH</sub> populations from the aforementioned t-SNE analysis.

After manually gating on four different populations of CXCR5<sup>+</sup> T<sub>FH</sub> cells in tonsils (Figure 5A), we quantified the frequencies of these populations with respect to donor age (Figure 5B). Linear regression analysis revealed that CXCR5<sup>lo</sup>PD-1<sup>-</sup>CD57<sup>-</sup> and CXCR5<sup>lo</sup>PD-1<sup>lo</sup>CD57<sup>-</sup> cells, likely representing early-stage memory T<sub>FH</sub> cells residing in extrafollicular regions, increased significantly with age ( $p < 0.0001$  and  $p = 0.024$ , respectively). In contrast, the frequency of bona fide CXCR5<sup>hi</sup>PD-1<sup>hi</sup> T<sub>FH</sub> cells did not correlate significantly with age, regardless of CD57 expression. We also did not find any significant correlations with these T<sub>FH</sub> populations and the other clinical parameters listed in Table S2 (data not shown).

In order to further confirm these findings, we quantified cells representing the 15 clusters identified by t-SNE/DensVM (Figure 4B) using a Boolean gating strategy, with the defining markers for each cluster listed in Table S3. Due to the similarity of marker expression in clusters 1–6, as well as clusters 8 and 9, we treated these clusters as a single population during the Boolean gating analysis. Using this approach, we found that clusters 1–6, representing CXCR5<sup>lo</sup>PD-1<sup>-</sup>CD57<sup>-</sup> cells, and clusters 8 and 9, representing CXCR5<sup>lo</sup>PD-1<sup>lo</sup>CD57<sup>-</sup> cells, increased significantly with age ( $p = 0.0045$ ) (Figure 5C). These data were similar to the regression analysis performed using manually gated populations of these T<sub>FH</sub> subsets (Figure 5B). Of the remaining clusters, only cluster 12, representing T<sub>H</sub>1-like T<sub>FH</sub> cells expressing IFN- $\gamma$ , correlated positively with age ( $p = 0.0045$ ). Collectively, these data suggest that CXCR5<sup>lo</sup>PD-1<sup>-</sup> and CXCR5<sup>lo</sup>PD-1<sup>lo</sup> memory T<sub>FH</sub> cells accumulate in secondary lymphoid organs during childhood, likely to mount a rapid secondary immune response upon re-challenge with antigen.

## DISCUSSION

CD4<sup>+</sup> T<sub>H</sub> cells are extremely heterogeneous, possessing the ability to differentiate into numerous subsets depending on the cellular microenvironment (Yamane and Paul, 2013; Zhu and Paul, 2010). While several major subsets have been characterized, the diversity within these subsets has been underappreciated. Here, we utilize the power of mass cytometry to evaluate human T<sub>FH</sub> cells at the single-cell level, employing several-dimensionality reduction and automated clustering methods to identify various subtypes of T<sub>FH</sub> cells and infer their relatedness. This study compares T<sub>FH</sub> cells in blood versus tonsils, wherein we hypothesize phenotypic progressions of T<sub>FH</sub> cells spanning both tissues. Our approach was distinct from a recent report using mass cytometry to evaluate intracellular signaling in tonsillar T cells (Sen et al., 2014), which did not evaluate T<sub>FH</sub> cells in their mass cytometry staining panel.

Circulating T<sub>FH</sub> cells are hypothesized to be counterparts of tonsillar T<sub>FH</sub> cells (Morita et al., 2011; Schmitt et al., 2014),

### Figure 5. Linear Regression Analysis of Tonsillar T<sub>FH</sub> Cells Shows Increasing Frequency of Extrafollicular Memory T<sub>FH</sub> Cells with Age

Mass cytometry staining was performed on cells derived from 18 tonsil donors (see Table S2).

(A) Representative gating scheme is shown for identifying four subtypes of tonsillar CXCR5<sup>+</sup> T<sub>FH</sub> cells (P1–P4).

(B and C) Regression analysis shows the relationship between the frequency of (B) the indicated T<sub>FH</sub> subtypes or (C) the indicated clusters (identified by DensVM in Figure 4) with age (see Table S3). All frequencies are calculated as a percentage of CD4<sup>+</sup> cells. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

but the phenotypic relationship between these cells remains elusive. Using dimensionality reduction and automated clustering methods, we highlight three major populations of  $T_{FH}$  cells across blood and tonsils.  $CXCR5^{lo}PD-1^{lo}ICOS^{lo}CCR7^{lo}$  cells were found in both tissues, suggesting these cells as phenotypic intermediates for  $T_{FH}$  cells trafficking from blood into tonsils. In an unbiased fashion, ISOMAP mapped these populations in a phenotypic progression, where CCR7 expression negatively correlated with CXCR5, PD-1, and ICOS. This continuum likely represented cells migrating out of blood into T cell zones and finally into B cell follicles, since downregulation of CCR7 and concomitant upregulation of CXCR5 is required for this migration to occur (He et al., 2013; Schmitt et al., 2014). These data, along with the aforementioned studies, supported the hypothesis that circulating  $T_{FH}$  cells might be the memory cell counterparts of bona fide  $T_{FH}$  cells in secondary lymphoid organs.

Previous studies have identified various subsets within the circulating  $CD4^{+}CXCR5^{+}$  compartment (Bentebibel et al., 2013; He et al., 2013; Locci et al., 2013; Morita et al., 2011), implying that  $T_{FH}$  cells adopt different phenotypes depending on the immune response. One report described  $T_{H1}$ - ( $T_{FH1}$ ),  $T_{H2}$ - ( $T_{FH2}$ ), and  $T_{H17}$ -like ( $T_{FH17}$ )  $T_{FH}$  cells in the blood based on differential expression of CXCR3 and CCR6, though cytokine secretion was not analyzed at the single-cell level (Morita et al., 2011). While we were unable to stain for CXCR3 due to downregulation upon stimulation, we identified tonsillar  $CXCR5^{lo}PD-1^{lo}$   $T_{FH}$  cells that expressed IFN- $\gamma$  and IL-17, perhaps representing counterparts of the circulating  $T_{FH1}$  and  $T_{FH17}$  cells. In addition, tonsillar-derived  $Foxp3^{+}$   $T_{FH}$  cells were found in the same DensVM cluster as  $T_{FH17}$  cells, supporting previous murine studies describing that  $T_{FH}$  cells could adopt a regulatory phenotype (Chung et al., 2011; Wollenberg et al., 2011). Thus, our study provides further evidence that  $T_{FH}$  cells are plastic but that heterogeneity in terms of cytokine production is mostly restricted to  $CXCR5^{lo}PD-1^{lo}$  subsets that may represent intermediately differentiated  $T_{FH}$  cells. Furthermore, once  $T_{FH}$  cells enter the follicles, represented by  $CXCR5^{hi}PD-1^{hi}$  cells, they seem to lose this plasticity based on our data.

In addition to a continuum implying  $T_{FH}$  cell trafficking, we identified a  $T_{FH}$  progression representing an increase of functional capacity with IL-2, TNF- $\alpha$ , and CD40L. While functional capacity positively correlated with PD-1 and ICOS expression in  $CXCR5^{lo}$   $T_{FH}$  cells,  $CXCR5^{hi}$  cells in the tonsils demonstrated decreasing functional capacity with increasing PD-1 and ICOS expression. This is consistent with observations of germinal center  $T_{FH}$  cells being poor cytokine producers (Rasheed et al., 2006). The expression of CD57 on  $CXCR5^{hi}$   $T_{FH}$  cells represented the last stage of cellular progression, with cells expressing the lowest levels of IL-2, TNF- $\alpha$ , and CD40L. Interestingly, CD57 expression is often associated with T cell exhaustion and senescence (Brenchley et al., 2003; Wood et al., 2009), but two studies found that  $CD57^{+}$   $T_{FH}$  cells were highly capable of helping B cells (Kim et al., 2001, 2005; Rasheed et al., 2006). On the contrary, another study found that  $T_{FH}$  cells were able to help B cells independent of CD57 expression (Rasheed et al., 2006). Regardless, these  $CD57^{+}$   $T_{FH}$  cells have been found to localize exclusively to germinal centers (Kim et al., 2001), further supporting that the progression inferred by ISOMAP cor-

responds to  $T_{FH}$  trafficking from blood to the germinal centers of secondary lymphoid organs.

Human memory  $CD4^{+}$  T cells have been shown to increase in frequency over time in both blood and tissues due to constant antigen exposure (Farber et al., 2014). Our findings that  $CXCR5^{lo}PD-1^{-}$  and  $CXCR5^{lo}PD-1^{lo}$  populations accumulate in the tonsils with age during pre-adulthood supports the notion that these populations represent memory  $T_{FH}$  cells (Locci et al., 2013). In contrast, the frequency of bona fide  $CXCR5^{hi}PD-1^{hi}$   $T_{FH}$  cells did not correlate with age, suggesting that the frequency of these cells are more dependent on the context of inflammation. These results indicate that the pool of  $CXCR5^{lo}PD-1^{-/lo}$  cells increase over time in the tonsils to support a secondary challenge with antigen. While obtaining a large number of tonsils from adults would be difficult, it would be interesting to determine whether the  $T_{FH}$  clusters identified in this study correlate with age, given that the germinal center response generally declines during late adulthood (Linterman, 2014).

Our analysis of  $CD4^{+}$   $T_H$  cells highlights the strengths and weaknesses of various dimensionality reduction methods, and we propose integrating some of these methods to exploit their individual strengths. When compared to PCA and ISOMAP, t-SNE best segregated known subsets of  $T_H$  cells in blood and tonsils. One limitation of t-SNE is that it is unable to consistently segregate extremely rare populations of cells. For example, we were unable to observe distinct clusters for  $T_{H2}$  and  $T_{H22}$  cells. This limitation may be alleviated by performing down-sampling functions to give more equal weight to rare populations, which is incorporated into some analysis programs such as SPADE (Qiu et al., 2011). Another caveat is that t-SNE can often detect subtle changes in marker expression, often resulting in excessive cluster identification by ACCENSE/DensVM. For example, while we detected six clusters of naive T cells in blood and tonsils, closer analysis likely indicates that there should only be two to three clusters. Thus, it is important to validate the clusters identified by these methods before proceeding to downstream analysis, either through manual gating or other approaches.

From our analysis, ISOMAP geometrically preserved cellular progressions better than t-SNE. This was best illustrated in our analysis of  $CXCR5^{+}$   $T_{FH}$  cells, as running ISOMAP on three separate occasions resulted in visually consistent cellular progressions compared to t-SNE. These data supported the notion that individual cells measured by mass cytometry lie on a high-dimensional nonlinear manifold space, where ISOMAP naturally learns the underlying global geometry of the data while retaining local metric information (Tenenbaum et al., 2000). One alternative method for analyzing cellular progressions, Wanderlust, is able to map developmental trajectories along a one-dimensional path (Bendall et al., 2014), which would be appropriate if pathways are assumed to be non-branching. In cases where cellular progressions are assumed to involve multiple pathways, ISOMAP may be suitable for this analysis.

In summary, we present a broad phenotypic analysis of  $T_{FH}$  cells across blood and tonsils, employing a range of methods each with unique advantages. By combining ISOMAP with t-SNE and DensVM, we were able to simultaneously visualize high-resolution diversity and cellular inter-relatedness with  $CXCR5^{+}$   $T_{FH}$  cells. These data provide further evidence that

blood  $T_{FH}$  cells are the counterparts of  $T_{FH}$  cells found in secondary lymphoid organs, and map out a hypothetical pathway by which circulating  $T_{FH}$  cells develop and traffic to the germinal centers, which should be tested and validated with in vivo mouse models. Future studies should also address the functional relevance of the numerous subtypes of  $T_{FH}$  cells, as well as determine their association with clinical parameters such as age or response to vaccination similar to previous studies (Bentebibel et al., 2013; He et al., 2013; Locci et al., 2013).

## EXPERIMENTAL PROCEDURES

### PBMC and Tonsil Cell Isolation

Whole blood and leukocyte reduction system (LRS) cone blood from healthy donors were obtained from the Singapore Immunology Network and the Health Sciences Authority of Singapore, respectively, according to institutional review board (IRB) protocol. PBMCs were isolated by density gradient centrifugation (Ficoll-Paque PLUS, GE Healthcare) as previously described (Wong et al., 2010). Tonsil cells were obtained from KK Women's and Children's Hospital (Singapore) according to IRB protocol. Tonsil tissue was cut into smaller pieces and mashed through a 100- $\mu$ m nylon mesh placed on a six-well plate (BD Falcon) containing PBS to create a single-cell suspension. The cells were washed and resuspended in PBS before filtering once more through a nylon mesh. PBMCs and tonsil cells were cryopreserved in 90% fetal calf serum + 10% DMSO.

### Stimulation, Staining, and CyTOF Data Acquisition

Cryopreserved PBMCs or tonsil cells were thawed and washed with pre-warmed complete RPMI media (cRPMI - 10% fetal calf serum [FCS], 1 $\times$  penicillin/streptomycin/L-glutamine, HEPES, and 1 $\times$   $\beta$ -mercaptoethanol) and rested overnight at 37°C in 24-well tissue culture plates (BD Falcon). Cells were washed and stained at 37°C in 96-well non-tissue culture treated round bottom plates (BD Falcon) with the indicated antibodies (Table S1) for 30 min prior to stimulation. The addition of these antibodies did not stimulate the cells based on comparing to the condition with no antibodies added (data not shown). Cells were stimulated with 150 ng/ml PMA and 1  $\mu$ M ionomycin in the presence of 1 $\times$  BFA (eBioscience) and 1 $\times$  monensin (eBioscience).

After 4 hr of stimulation, cells were washed twice in cold CyFACS buffer (PBS + 2% FCS + 2 mM EDTA + 0.05% sodium azide) and incubated for 5 min on ice with 200  $\mu$ M cisplatin (Sigma), a viability marker. Cells were washed with CyFACS, stained with primary antibody on ice for 30 min, washed twice with CyFACS, and then stained with antibody cocktail (Table S1). After 30 min, cells were washed twice with CyFACS, once with PBS, and then fixed in PBS with 2% paraformaldehyde (Electron Microscopy Sciences) overnight at 4°C. For PBMC versus tonsil comparison studies, cells were first resuspended in Foxp3 fixation/permeabilization buffer (eBioscience) for 10 min on ice, washed twice with PBS, and fixed in PBS with 2% paraformaldehyde overnight at 4°C.

On the next day, cells were washed three times with intracellular staining permeabilization (perm) buffer (BioLegend) and stained with primary antibody (Table S1) diluted in perm buffer for 45 min at room temperature. After two washes in perm buffer, cells were stained at room temperature with an intracellular antibody cocktail (Table S1) diluted in perm buffer. After 45 min, cells were washed once with perm buffer and once with PBS prior to barcoding.

For barcoding, 2 mM bromoacetamidobenzyl-EDTA (BABE, Dojindo) with 0.5 mM PdCl<sub>2</sub> dissolved in HEPES buffer or DOTA-maleimide (DM, Macrocyclics) with 50 mM RhCl<sub>3</sub> (Sigma) or LnCl<sub>3</sub> (Trace Sciences) dissolved in L buffer (DVS Sciences) was used. For each sample, a unique dual combination of each of the following barcodes was added to cells for 30 min on ice: DM-Rh-103, BABE-Pd-104, BABE-Pd-106, BABE-Pd-108, BABE-Pd-110, and DM-Ln-113. Cells were subsequently washed with perm buffer, incubated with CyFACS for 5 min on ice, washed once more with CyFACS, and labeled at room temperature with 250 nM iridium interchelator (DVS Sciences) diluted in PBS with 2% paraformaldehyde. After 20 min, cells were washed twice with CyFACS and twice in distilled water before diluting to 5  $\times$  10<sup>5</sup> cells/ml

in distilled water. Cells were acquired on CyTOF using dual-count mode and analyzed with noise-reduction turned off. Calibration of the instrument with tuning solution was performed as instructed by DVS Sciences.

### Antibody Conjugation

Purified antibodies lacking carrier proteins were purchased from the companies listed in Table S1. Antibody conjugation was performed according to the protocol provided by DVS Sciences.

### Data Analysis

We analyzed the data using the following steps (Figure S3B), which have been implemented into an R-package. The R-package and all data used for this study are publicly available for download at <https://github.com/signbioinfo/sidap>. An expanded version of the data analysis is provided in Supplemental Experimental Procedures.

#### Pre-processing

FCS files were imported into the R environment via the read.FCS function in the flowCore package. Intensity values of marker expression were then logicle-transformed, and markers as listed in Figure 1 were extracted for downstream analysis. Because the number of CD4<sup>+</sup> T cell events varied dramatically between blood and tonsil donors, we randomly sampled up to 10,000 CD4<sup>+</sup> T cell events per donor to partially normalize the contribution of each donor.

#### Dimensionality Reduction Using t-SNE

We utilized bh\_tsne, an efficient implementation of t-SNE via Barnes-Hut approximations (van der Maaten and Hinton, 2008). bh\_tsne was implemented and compiled in C++. We then implemented an interface function to execute bh\_tsne from R.

#### Cluster Analysis Using DensVM

We used density-based clustering aided by support vector machine (DensVM) to automate subset detection from the t-SNE map (Figure S3A), which has been previously described (Becher et al., 2014).

#### Cluster Annotation Using Heatmaps

We annotated individual clusters by using heatmaps. We grouped cell events by clusters and calculated the median intensity values per cluster for every marker. Heatmaps visualizing the median expression of every marker in every cluster were generated with no scaling on the row or column direction. Hierarchical clustering was generated using Euclidean distance and complete agglomeration method. We used the heatmaps to interrogate marker expression to identify each cluster's defining markers. Based on this, we designated the individual clusters as one of previously described or unknown populations.

#### Comparative Analysis and Statistical Tests

Because independent analysis of two similar samples will result in very different maps, we performed the analysis on cells combined from the four PBMC and three tonsil donors. A trellis visualization of the t-SNE map was then generated to visually identify the difference between donors. The frequencies of DensVM clusters were calculated on a per donor basis and a heatmap together with a dendrogram was plotted to illustrate the differences of cell subset frequencies. The grouping of donors was shown by the clustering dendrogram on donors. Based on the DensVM analysis, we ran t test and BH correction on cluster frequencies to identify which clusters have significantly different frequencies between blood and tonsils.

#### Construction of Subset Transition Graph Using ISOMAP

We first downsampled the data by randomly selecting a comparable number of cell events from each of the clusters that were identified by the t-SNE and DensVM analysis. The sampled cell events were pooled and subjected to ISOMAP dimensionality reduction. On the first two ISOMAP dimensions, we placed nodes at the centroid of each cluster and used the inter-cluster continuum to draw edges connecting proximate clusters. Along the first and second ISOMAP dimensions, we generated 100 bins of equal intervals and calculated the median intensities of markers expressed by cells within each bin. Smoothed curves were then plotted using the R package LOWESS to show the progressive phenotypic change.

#### Post-processing

We coded the DensVM cluster assignment of each cell into a two-dimensional coordinate system that was then inverse-logicle transformed. Similarly, we inverse-logicle transformed the coordinates of the t-SNE map and ISOMAP. The

cluster coordinates, together with the t-SNE and ISOMAP coordinates, were added to the FCS files as additional parameters. Subsequently, we manually gated on populations of interest and overlaid the gated subsets on the t-SNE or ISOMAP plots using the FlowJo software.

### Linear Regression Analysis

Linear regression analysis was performed to correlate age with frequency of the indicated subsets. Quantification of subsets was performed using a Boolean gating strategy, with the markers used to define each subset listed in Table S3. Adjusted *p* values are shown after multiple test correction (Figure 5C).

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2015.05.022>.

### AUTHOR CONTRIBUTIONS

M.T.W., W.L., and S.N. performed the experiments. J.C., M.T.W., and E.W.N. analyzed the data and generated the figures with guidance from M.P., M.A.C.D., H.T.K.K., and R.A. provided tissues and/or critical analytical support. M.T.W., J.C., and E.W.N. wrote the manuscript. The study was financially supported by E.W.N., M.P., and M.A.C.D.

### ACKNOWLEDGMENTS

The authors thank the SigN community and all members of the Newell and Poidinger labs for helpful discussion and technical support. In particular, we thank Chen Hao for developing the graphical user interface of our R package. Some antibodies used for generating the mass cytometry panels were generously provided by Robert Balderas and Anne Tiong (Becton Dickinson). This study was funded by A-STAR/SigN core funding (E.W.N., M.P., and M.A.C.D.).

Received: October 21, 2014

Revised: April 13, 2015

Accepted: May 8, 2015

Published: June 11, 2015

### REFERENCES

Acosta-Rodriguez, E.V., Rivino, L., Geginat, J., Jarrossay, D., Gattorno, M., Lanzavecchia, A., Sallusto, F., and Napolitani, G. (2007). Surface phenotype and antigenic specificity of human interleukin 17-producing T helper memory cells. *Nat. Immunol.* **8**, 639–646.

Alves, N.L., van Leeuwen, E.M., Remmerswaal, E.B., Vrisekoop, N., Tesselaar, K., Roosnek, E., ten Berge, I.J., and van Lier, R.A. (2007). A new subset of human naive CD8+ T cells defined by low expression of IL-7R alpha. *J. Immunol.* **179**, 221–228.

Amir el, A.D., Davis, K.L., Tadmor, M.D., Simonds, E.F., Levine, J.H., Bendall, S.C., Shenfeld, D.K., Krishnaswamy, S., Nolan, G.P., and Pe'er, D. (2013). viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* **31**, 545–552.

Bains, I., Yates, A.J., and Callard, R.E. (2013). Heterogeneity in thymic emigrants: implications for thymectomy and immunosenescence. *PLoS ONE* **8**, e49554.

Becher, B., Schlitzer, A., Chen, J., Mair, F., Sumatoh, H.R., Teng, K.W., Low, D., Ruedl, C., Riccardi-Castagnoli, P., Poidinger, M., et al. (2014). High-dimensional analysis of the murine myeloid cell system. *Nat. Immunol.* **15**, 1181–1189.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696.

Bendall, S.C., Nolan, G.P., Roederer, M., and Chattopadhyay, P.K. (2012). A deep profiler's guide to cytometry. *Trends Immunol.* **33**, 323–332.

Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Simonds, E.F., Chen, T.J., Shenfeld, D.K., Nolan, G.P., and Pe'er, D. (2014). Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell* **157**, 714–725.

Bentebibel, S.E., Lopez, S., Obermoser, G., Schmitt, N., Mueller, C., Harrod, C., Flano, E., Mejias, A., Albrecht, R.A., Blankenship, D., et al. (2013). Induction of ICOS+CXCR3+CXCR5+ TH cells correlates with antibody responses to influenza vaccination. *Science translational medicine* **5**, 176ra132.

Bodenmiller, B., Zunder, E.R., Finck, R., Chen, T.J., Savig, E.S., Bruggner, R.V., Simonds, E.F., Bendall, S.C., Sachs, K., Krutzik, P.O., and Nolan, G.P. (2012). Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867.

Brenchley, J.M., Karandikar, N.J., Betts, M.R., Ambrozak, D.R., Hill, B.J., Crotty, L.E., Casazza, J.P., Kuruppu, J., Migueles, S.A., Connors, M., et al. (2003). Expression of CD57 defines replicative senescence and antigen-induced apoptotic death of CD8+ T cells. *Blood* **101**, 2711–2720.

Cannons, J.L., Lu, K.T., and Schwartzberg, P.L. (2013). T follicular helper cell diversity and plasticity. *Trends Immunol.* **34**, 200–207.

Chung, Y., Tanaka, S., Chu, F., Nurieva, R.I., Martinez, G.J., Rawal, S., Wang, Y.H., Lim, H., Reynolds, J.M., Zhou, X.H., et al. (2011). Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nat. Med.* **17**, 983–988.

Cosmi, L., De Palma, R., Santarlasci, V., Maggi, L., Capone, M., Frosali, F., Rodolico, G., Querci, V., Abbate, G., Angeli, R., et al. (2008). Human interleukin 17-producing cells originate from a CD161+CD4+ T cell precursor. *J. Exp. Med.* **205**, 1903–1916.

Farber, D.L., Yudanin, N.A., and Restifo, N.P. (2014). Human memory T cells: generation, compartmentalization and homeostasis. *Nat. Rev. Immunol.* **14**, 24–35.

Griffith, J.W., Sokol, C.L., and Luster, A.D. (2014). Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu. Rev. Immunol.* **32**, 659–702.

He, J., Tsai, L.M., Leong, Y.A., Hu, X., Ma, C.S., Chevalier, N., Sun, X., Vandenberg, K., Rockman, S., Ding, Y., et al. (2013). Circulating precursor CCR7(lo) PD-1(hi) CXCR5+ CD4s+ T cells indicate Tfh cell activity and promote antibody responses upon antigen reexposure. *Immunity* **39**, 770–781.

Horowitz, A., Strauss-Albee, D.M., Leipold, M., Kubo, J., Nemat-Gorgani, N., Dogan, O.C., Dekker, C.L., Mackey, S., Maecker, H., Swan, G.E., et al. (2013). Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry. *Science translational medicine* **5**, 208ra145.

Jameson, S.C., and Fulton, R.B. (2011). Not all naïve CD8 T cells are created equal. *Immunol. Cell Biol.* **89**, 576–577.

Kattah, M.G., Wong, M.T., Yocum, M.D., and Utz, P.J. (2008). Cytokines secreted in response to Toll-like receptor ligand stimulation modulate differentiation of human Th17 cells. *Arthritis Rheum.* **58**, 1619–1629.

Kim, C.H., Rott, L.S., Clark-Lewis, I., Campbell, D.J., Wu, L., and Butcher, E.C. (2001). Subspecialization of CXCR5+ T cells: B helper activity is focused in a germinal center-localized subset of CXCR5+ T cells. *J. Exp. Med.* **193**, 1373–1381.

Kim, J.R., Lim, H.W., Kang, S.G., Hillsamer, P., and Kim, C.H. (2005). Human CD57+ germinal center-T cells are the major helpers for GC-B cells and induce class switch recombination. *BMC Immunol.* **6**, 3.

Lim, H.W., Lee, J., Hillsamer, P., and Kim, C.H. (2008). Human Th17 cells share major trafficking receptors with both polarized effector T cells and FOXP3+ regulatory T cells. *J. Immunol.* **180**, 122–129.

Linterman, M.A. (2014). How T follicular helper cells and the germinal centre response change with age. *Immunol. Cell Biol.* **92**, 72–79.

Locci, M., Havenar-Daughton, C., Landais, E., Wu, J., Kroenke, M.A., Arlehamn, C.L., Su, L.F., Cubas, R., Davis, M.M., Sette, A., et al.; International AIDS Vaccine Initiative Protocol C Principal Investigators (2013). Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and

- correlate with broadly neutralizing HIV antibody responses. *Immunity* 39, 758–769.
- Morita, R., Schmitt, N., Bentebibel, S.E., Ranganathan, R., Bourdery, L., Zurawski, G., Foucat, E., Dullaers, M., Oh, S., Sabzghabaei, N., et al. (2011). Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity* 34, 108–121.
- Newell, E.W., and Davis, M.M. (2014). Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. *Nat. Biotechnol.* 32, 149–157.
- Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., and Davis, M.M. (2012). Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity* 36, 142–152.
- Qiu, P., Simonds, E.F., Bendall, S.C., Gibbs, K.D., Jr., Bruggner, R.V., Linderman, M.D., Sachs, K., Nolan, G.P., and Plevritis, S.K. (2011). Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat. Biotechnol.* 29, 886–891.
- Rasheed, A.U., Rahn, H.P., Sallusto, F., Lipp, M., and Müller, G. (2006). Follicular B helper T cell activity is confined to CXCR5(hi)ICOS(hi) CD4 T cells and is independent of CD57 expression. *Eur. J. Immunol.* 36, 1892–1903.
- Schmitt, N., Bentebibel, S.E., and Ueno, H. (2014). Phenotype and functions of memory Tfh cells in human blood. *Trends Immunol.* 35, 436–442.
- Sen, N., Mukherjee, G., Sen, A., Bendall, S.C., Sung, P., Nolan, G.P., and Arvin, A.M. (2014). Single-cell mass cytometry analysis of human tonsil T cell remodeling by varicella zoster virus. *Cell Rep.* 8, 633–645.
- Shekhar, K., Brodin, P., Davis, M.M., and Chakraborty, A.K. (2014). Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding (ACCENSE). *Proc. Natl. Acad. Sci. USA* 111, 202–207.
- Tenenbaum, J.B., de Silva, V., and Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 2319–2323.
- van der Maaten, L., and Hinton, G. (2008). Visualizing High-Dimensional Data Using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
- Voo, K.S., Wang, Y.H., Santori, F.R., Boggiano, C., Wang, Y.H., Arima, K., Bover, L., Hanabuchi, S., Khalili, J., Marinova, E., et al. (2009). Identification of IL-17-producing FOXP3+ regulatory T cells in humans. *Proc. Natl. Acad. Sci. USA* 106, 4793–4798.
- Wollenberg, I., Agua-Doce, A., Hernández, A., Almeida, C., Oliveira, V.G., Faro, J., and Graca, L. (2011). Regulation of the germinal center reaction by Foxp3+ follicular regulatory T cells. *J. Immunol.* 187, 4553–4560.
- Wong, M.T., Ye, J.J., Alonso, M.N., Landrigan, A., Cheung, R.K., Engleman, E., and Utz, P.J. (2010). Regulation of human Th9 differentiation by type I interferons and IL-21. *Immunol. Cell Biol.* 88, 624–631.
- Wood, K.L., Twigg, H.L., 3rd, and Doseff, A.I. (2009). Dysregulation of CD8+ lymphocyte apoptosis, chronic disease, and immune regulation. *Front Biosci (Landmark Ed)* 14, 3771–3781.
- Yamane, H., and Paul, W.E. (2013). Early signaling events that underlie fate decisions of naive CD4(+) T cells toward distinct T-helper cell subsets. *Immunol. Rev.* 252, 12–23.
- Zhou, L., Lopes, J.E., Chong, M.M., Ivanov, I.I., Min, R., Victora, G.D., Shen, Y., Du, J., Rubtsov, Y.P., Rudensky, A.Y., et al. (2008). TGF-beta-induced Foxp3 inhibits T(H)17 cell differentiation by antagonizing RORgamma function. *Nature* 453, 236–240.
- Zhu, J., and Paul, W.E. (2010). Heterogeneity and plasticity of T helper cells. *Cell Res.* 20, 4–12.