**Article**

# Quantitative Proteomics Reveals the Dynamic Protein Landscape during Initiation of Human Th17 Cell Polarization



1. Characterization of the proteome during Th17 polarization

2. Biological context — Signaling by TGF-beta Receptor Complex — Jak-STAT signaling

3. Expression changes between Th17 cells and Th0 cells show high degree of correlation in proteomics and transcriptomics

4. Differentially expressed proteins between Th17 and Th0 in human and mouse have poor overlap

5. New regulators of Th17 differentiation in human identified

Subhash K. Tripathi, Tommi Välikangas, Ankitha Shetty, ..., Sanjeev Galande, Laura L. Elo, Riitta Lahesmaa

laura.elo@utu.fi (L.L.E.)
rilahes@utu.fi (R.L.)

**HIGHLIGHTS**

Quantitative proteomics analysis of early human Th17 cell polarization

The proteome and transcriptome highly correlate during early Th17 polarization

Poor overlap of proteome profiles of human and mouse during early Th17 polarization

The results underline the importance of human studies for translational research

**DATA AND SOFTWARE AVAILABILITY**
GSE118974

# iScience
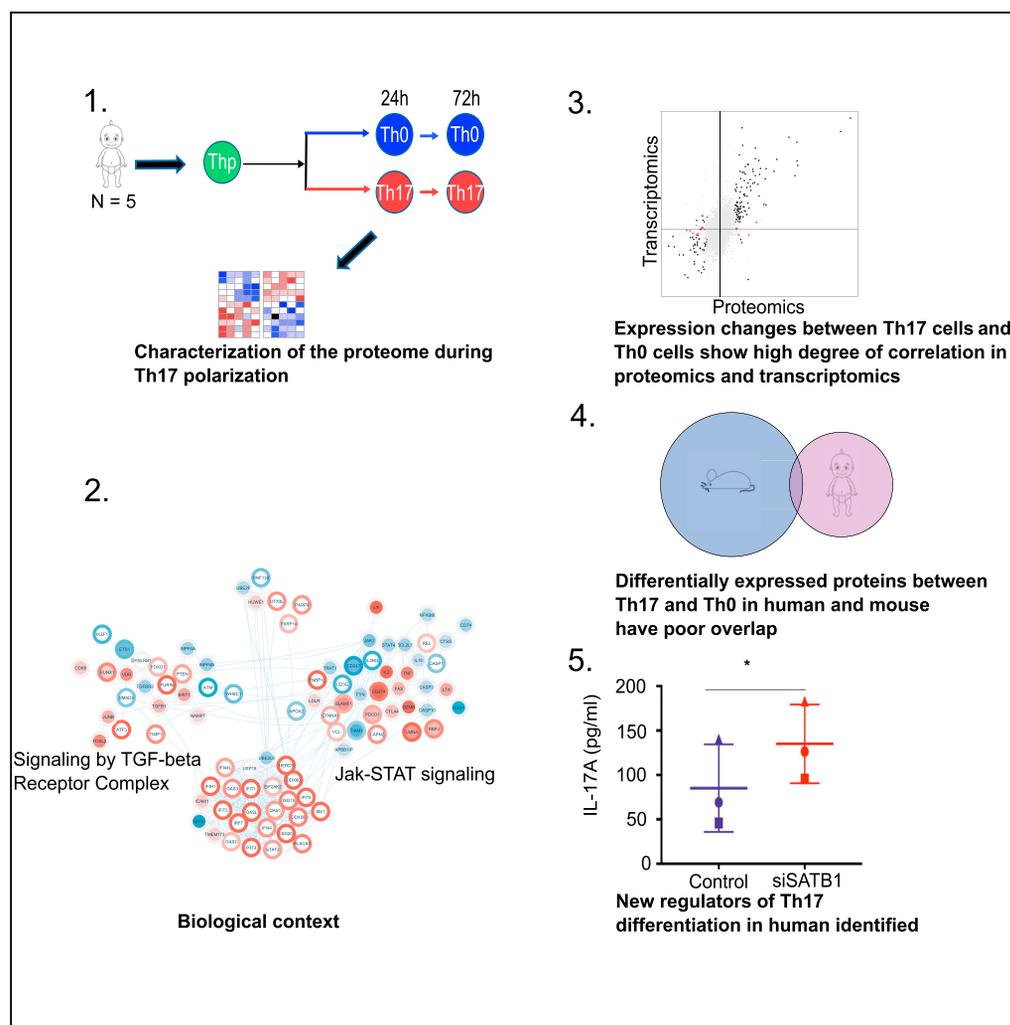
## Article

# Quantitative Proteomics Reveals the Dynamic Protein Landscape during Initiation of Human Th17 Cell Polarization

Subhash K. Tripathi,[1,5] Tommi Välikangas,[1,2,5] Ankitha Shetty,[1,4,5] Mohd Moin Khan,[1,3] Robert Moulder,[1] Santosh D. Bhosale,[1,3] Elina Komsi,[1] Verna Salo,[1,3] Rafael Sales De Albuquerque,[1] Omid Rasool,[1] Sanjeev Galande,[4] Laura L. Elo,[1,5,*] and Riitta Lahesmaa[1,5,6,*]

## SUMMARY

**Th17 cells contribute to the pathogenesis of inflammatory and autoimmune diseases and cancer. To reveal the Th17 cell-specific proteomic signature regulating Th17 cell differentiation and function in humans, we used a label-free mass spectrometry-based approach. Furthermore, a comprehensive analysis of the proteome and transcriptome of cells during human Th17 differentiation revealed a high degree of overlap between the datasets. However, when compared with corresponding published mouse data, we found very limited overlap between the proteins differentially regulated in response to Th17 differentiation. Validations were made for a panel of selected proteins with known and unknown functions. Finally, using RNA interference, we showed that SATB1 negatively regulates human Th17 cell differentiation. Overall, the current study illustrates a comprehensive picture of the global protein landscape during early human Th17 cell differentiation. Poor overlap with mouse data underlines the importance of human studies for translational research.**

## INTRODUCTION

CD4+ T cells are key players of the adaptive immune system. Upon antigenic stimulation, the naive CD4+ T (T helper precursor [Thp]) cells polarize into distinct T helper ([Th], i.e., Th1, Th2, Th9, and Th17) and regulatory T (Treg) cells (O'Shea and Paul, 2010). Although their primary function is to provide protective immunity against various intracellular and extracellular pathogens, they can also exhibit inappropriate responses in inflammatory and autoimmune diseases (McKinstry et al., 2010; Ghoreschi et al., 2011; Cosmi et al., 2014). Th17 cells are a subset of Th cells that are characterized by the expression of their key transcription factors, STAT3 and RORC and the chemokine receptor CCR6 as well as by the secretion of signature cytokines interleukin (IL)-17A and IL-17F. They are critical in combating fungal infections and contribute to the pathogenesis of several inflammatory and autoimmune diseases and various cancers (Hernández-Santos and Gaffen, 2012; Bailey et al., 2014; Burkett et al., 2015). The characterization of the molecular mechanisms regulating the differentiation and function of the Th17 cells is therefore of great interest for research into the etiology and treatment of these diseases.

Until now, our understanding of the proteins that are most important in Th17 differentiation and function in both human and mouse has originated from transcriptional profiling analyses (Ciofani et al., 2012; Tuomela et al., 2012; Yosef et al., 2013; Gaublomme et al., 2015). These analyses have identified a number of genes that now serve as key markers for Th17 cells. However, this approach is limited as some transcriptional changes are not necessarily reflected at the proteome level (Vogel and Marcotte, 2012). For example, several post-transcriptional and post-translational mechanisms modulate the stability and activity of many of the proteins involved (Liu et al., 2016). Characterization of the cellular proteome of Th17 cells enables the discovery of unique protein signatures and regulated cellular pathways that drive Th17 cell development and function. A better understanding of the proteome during the differentiation process facilitates the rational design of drugs targeting Th17-mediated inflammatory and autoimmune diseases.

Mass-spectrometry (MS)-based proteomic analysis is a powerful tool for comprehensively profiling the proteome in different cellular systems, including T cells (Cox and Mann, 2011; Howden et al., 2013). Previously, MS-based proteomic analysis of differentiating T cells focused mainly on classical Th1 and Th2 cells using in vitro differentiation systems (Loyet et al., 2005; Rautajoki et al., 2007). In addition, addressing

[1]Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Tykistökatu 6, FI-20520 Turku, Finland

[2]Doctoral Programme in Mathematics and Computer Sciences (MATTI), University of Turku, University Hill, FI-20014 Turku, Finland

[3]Turku Doctoral Programme of Molecular Medicine (TuDMM), University of Turku, Tykistökatu 6, FI-20520 Turku, Finland

[4]Centre of Excellence in Epigenetics, Department of Biology, Indian Institute of Science Education and Research (IISER), Pune 411008, India

[5]These authors contributed equally

[6]Lead Contact

*Correspondence: laura.elo@utu.fi (L.L.E.), rilahes@utu.fi (R.L.) https://doi.org/10.1016/j.isci.2018.12.020

disease-related traits, the proteomic profiles were compared for *in vivo* differentiated Th1 and Th1/Th17 cell clones isolated from biopsies of gut samples from patients with Crohn disease (Riaz et al., 2016). Recently, a number of studies identified a distinct set of differentially regulated proteins when comparing the proteomes of CD4+CD25+ Foxp3 expressing natural Treg cells and induced Treg (iTreg) with CD4+ conventional T cells both in human and mouse (Kubach et al., 2007; Duguet et al., 2017; Cuadrado et al., 2018; Schmidt et al., 2018). Most recently, a study reported Th17 proteome profiles in mouse (Mohammad et al., 2018). Although studies of the molecular profiles and mechanisms governing different Th and Treg cell differentiation have been mostly performed in mouse, previous reports that have compared the transcriptomic profiles of human and mouse have revealed significant differences between the two species (Schwanhüusser et al., 2011; Vogel and Marcotte, 2012). As the findings from studies based on mouse disease models often cannot be replicated in humans, studies in humans are critical (Mestas and Hughes, 2004; Mak et al., 2014).

In the current study, we utilized a label-free MS-based approach to build a quantitative dataset on the cellular proteome of naive CD4+ human T cells, CD3/CD28 activated T (Th0) cells, and Th17 cells at 24 and 72 hr after the initiation of polarization. Statistical analysis of the data revealed a Th17-cell-specific proteome signature with a number of proteins regulated during Th17 cell differentiation already at the early stage of the differentiation process. Moreover, selected proteins with previously known and unknown Th17-related functions were validated in additional samples by distinct methods to confirm the results. Furthermore, the proteomics and transcriptomics data generated in this study were compared to determine the degree of concordance between these two. Notably, a comparison of our human Th17-regulated proteome with the mouse Th17 proteome demonstrated poor overlap between the two species. Last, using the RNA interference (RNAi) approach, we demonstrated SATB1 as a negative regulator of human Th17 cell differentiation process in contrast to mouse, where it positively regulates Th17 cell differentiation (Ciofani et al., 2012). This study illustrates the global protein landscape and the mRNA-protein associations during early human Th17 cell differentiation. This dataset provides a valuable resource of candidate proteins potentially regulating the differentiation and functions of Th17 cells in humans. Further investigation on these candidate proteins might lead to the rational design of targets with therapeutic potential for modulating Th17-mediated immune responses in humans.
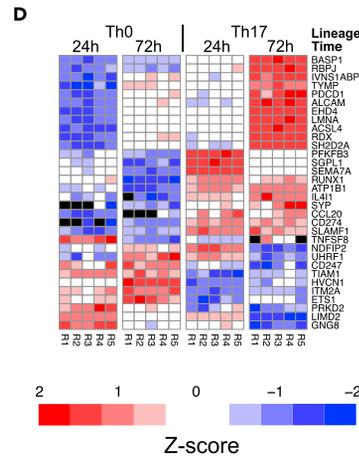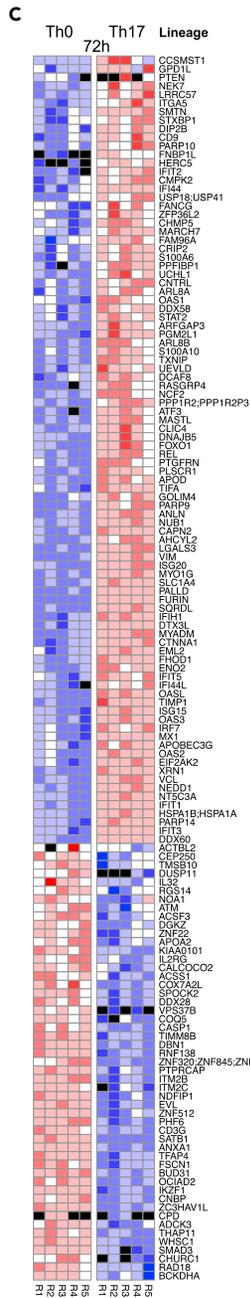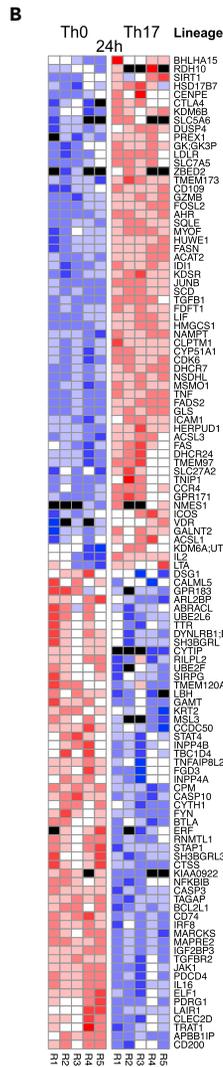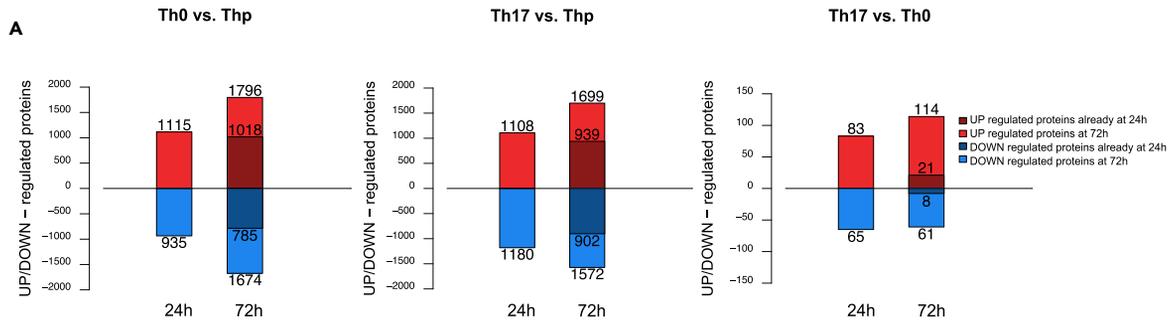
## RESULTS

### Quantitative Proteomic Analysis during Initiation of Human Th17 Differentiation

We investigated the quantitative changes in the cellular proteome of human naive Th cells during the early stages of human Th17 cell differentiation using shotgun label-free quantitative (LFQ) proteomics. Naive CD4+ T cells isolated from the human umbilical cord blood were either activated by T cell receptor (TCR) cross-linking with CD3 and CD28 antibodies (Th0 cells) or polarized with a cytokine cocktail of IL-1β, TGF-β, and IL-6 in combination with TCR/CD28 cross-linking to initiate Th17 cell differentiation. Polarization was monitored by the expression of Th17 cell marker genes, including CCR6, IL-17A, and IL-17F, and the master transcription factor RORC (Figures S1A–S1D). The Thp cell sample and the Th0 and Th17 cell samples at 24- and 72-hr time points post-initiation of polarization were collected from five individual donors. To rule out the possibility of polarization toward interferon (IFN)-γ-expressing pathogenic Th17 cells, expression of the IFN-γ cytokine was monitored in three separate Th17 cultures prepared in a similar manner (Figures S1E and S1F). The samples were prepared using the filter-aided sample preparation protocol and analyzed in triplicate by liquid chromatography-tandem MS (LC-MS/MS) (Figure S1G).

Using the label-free MS approach, we identified more than 5,600 unique proteins among the Thp, Th0, and Th17 cell subsets (Table S1). Samples in the normalized data clustered according to the biological replicates and cell lineages indicating successful normalization, good reproducibility, and general good quality of the data (Figure S2). The proportion of missing values in the samples was low in general (<7%), being largest in the Thp cells (∼4.4%–6.1%).

Furthermore, a considerable number of proteins were detected as differentially expressed (DE) between the different cell subsets. On comparing Th0 and Thp cells, 2,050 and 3,470 proteins were detected as DE (false discovery rate [FDR] ≤ 0.05) at 24 and 72 hr after the onset of the polarization, respectively (Figure 1A; Tables S1 and S2). Similarly, on comparing Th17 and Thp cells, 2,288 proteins were DE at 24 hr and 3,271 at 72 hr, respectively. Among the proteins DE in response to TCR activation (Th0 versus Thp), 935 and 1,674 proteins were downregulated and 1,115 and 1,796 proteins were upregulated in Th0 at 24 and 72 hr,

**A**



Th0 vs. Thp

Th17 vs. Thp

Th17 vs. Th0

**B** Th0 Th17 24h Lineage

**C** Th0 Th17 72h Lineage

**D** Th0 Th17 Lineage Time

**Figure 1. The Differentially Expressed (DE) Proteins in the Proteomics Data**

(A–D) (A) The number of up- and downregulated proteins between Th0 and Thp cells, Th17 and Thp cells, and Th17 and Th0 cells at 24 and 72 hr. At 72 hr, the number of proteins detected up- or downregulated already at 24 hr is shown in darker color. The Z score standardized expressions of the DE proteins between Th17 and Th0 unique to 24 hr (B), unique to 72 hr (C), and common to both time points (24 and 72 hr) (D). Black color in the heatmaps stands for undetected/missing value.

respectively. Likewise, in response to combined TCR activation and Th17-polarizing cytokines (Th17 versus Thp), 1,180 and 1,572 proteins were downregulated and 1,108 and 1,699 proteins were upregulated in Th17 cells at 24 and 72 hr, respectively. As expected, owing to the similarity of the cells at the early stage of Th17 polarization, the differential expression analysis between Th17 and Th0 cells revealed only 148 and 175 proteins differentially regulated at 24 and 72 hr, respectively (Figure 1A; Tables S1 and S2). Among the proteins that were DE in the Th17-polarizing conditions in a time-point-specific fashion (i.e., detected as DE only at either time points), 60 and 92 proteins were upregulated and 56 and 51 proteins were downregulated at 24 and 72 hr, respectively (Figures 1B and 1C). Interestingly, only a small number of DE proteins were found to be common at both time points (Figure 1D), suggesting that the differentially regulated proteins have a stage-specific function in driving Th17 differentiation.

Among the proteins detected as DE in this data, several are previously known to be preferentially expressed and regulated during Th17 differentiation. These included AHR (Veldhoen et al., 2008), JUNB (Carr et al., 2017), FOSL2 (Ciofani et al., 2012), RBPJ (Meyer zu Horste et al., 2016), SIRT1 (Lim et al., 2015), REL (Chen et al., 2011), CCL20 (Eriksson et al., 2015), and TNFSF8 (Sun et al., 2010). Notably, our analysis also identified several other proteins (e.g., UHRF1, GK, ICAM1, KDSR, ATF3, IL4IT, APOD, VIM, PALLD, PTGFRN, and KDM6A) differentially upregulated with unexplored/unknown functions in Th17 cells (Figures 1B and 1C). Interestingly, proteins related to antiviral pathways (e.g., IRF7, MX1, OAS, OAS31, and OASL) were upregulated, suggesting the role of IFN signaling in driving Th17 differentiation. We observed the upregulation of the ADP ribosyl transferase poly (ADP-ribose) polymerase (PARP) superfamily proteins, including PARP-9, PARP-10, and PAPR-14, together with the RNA helicase DEAD-box proteins, such as DDX58 and DDX60. PARP-14 is highly expressed and positively regulated during the development of Th17 and T helper follicular cells in mouse by modulating the phosphorylation of STAT3, a key transcriptional regulator of both (Mehrotra et al., 2015). Several negative regulators of Th17 cell development (Ciofani et al., 2012; Yosef et al., 2013), including IRF8, ETS1, STAT4, FYN, CASP1, and SMAD3, were significantly downregulated in differentiating Th17 cells. Interestingly, among the downregulated proteins, we also observed several proteins with unexplored function in Th17 biology including ELF1, IL16, BTLA, LAIR1, UHRF1, ANXA1, NDFIP1, and NDFIP2.

## Functional Analysis of the Th17 DE Proteins

We examined the enrichment of Gene Ontology (GO) annotations for biological processes among the DE proteins over both time points. The GO analysis revealed significant (FDR ≤0.05) enrichment of approximately 300 biological processes (Table S3). The enriched processes included a number of important immunological processes, such as regulation of immune system development, immune system processes, response to external stimulus, cell surface receptor signaling pathways, cytokine production, cytokine-mediated signaling pathways, positive regulation of cellular metabolic processes, T cell activation, and lymphocyte differentiation (Figure 2A; Table S3).

We observed that proteins related to lipid metabolic process were highly regulated in Th17 cells, which is in agreement with the recent studies showing the role of fatty acid metabolism in regulating Th17 cell differentiation and function (Figure 2B) (Berod et al., 2014; Santori et al., 2015). Known proteins of interest, such as VDR (Chang et al., 2010), are involved in lipid metabolism and have previously been linked to Th17 differentiation. MSMO1 and CYP51A1 are both also involved in lipid metabolism and have been reported to regulate RORC expression and the pathogenic activity of Th17 cells (Santori et al., 2015). Our analysis additionally identified a number of other proteins related to lipid metabolism signaling whose functions in regulating Th17 differentiation have not yet been defined (Table S3). As expected, we observed a number of proteins that were common to the immune system process and response to cytokines, including CD274 (PD-L1), CCL20, IFIT1, IRF7, and ETS1 (Figure 2B). Notably, proteins induced in response to IFN and antiviral pathways were highly enriched for the GO term "response to cytokine," including OASL, ISG20, IRF7, and IFN-induced protein with tetratricopeptide repeat (IFIT) family proteins.

The Ingenuity Pathways Analysis (IPA; https://www.qiagenbioinformatics.com) was used to assign the molecular functional types and cellular locations of the proteins DE between Th17 and Th0 cells (Figures S3A
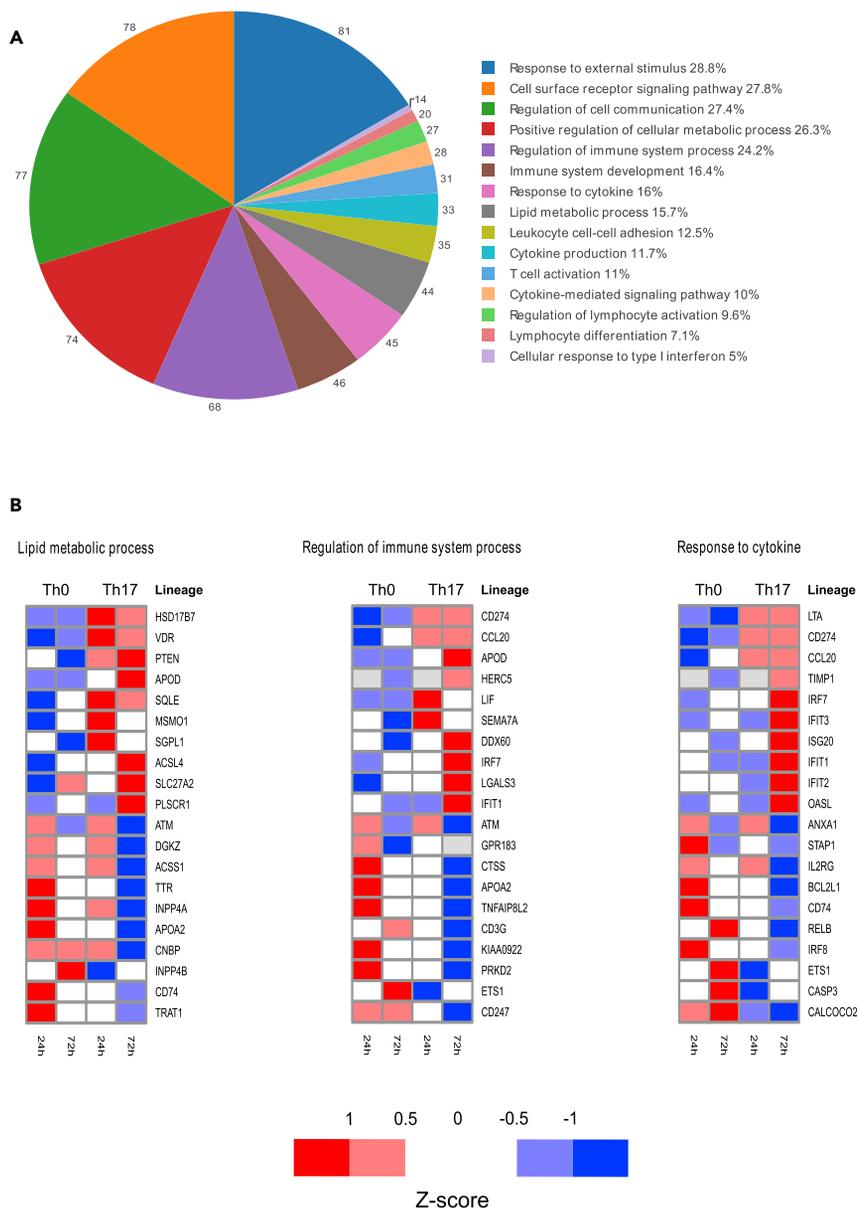
**Figure 2. Enriched GO Biological Processes for the Proteins Differentially Expressed between Th17 and Th0 in the Proteomics Data**

(A) The proportions of DE proteins for the 15 chosen representative GO biological processes out of all the annotated DE proteins.

(B) The averaged standardized expressions of 10 of the most upregulated and 10 of the most downregulated proteins in each of the three chosen representative GO biological processes. The GO analysis was performed using DAVID over both time points. The detected proteome was used as a reference background.

and S3B; Table S3). The IPA analysis revealed distinct molecular function types for the Th17-regulated proteins, including enzymes (~23%), transcription regulators (~12%), transporters and kinases (5% each), transmembrane receptors (4.6%), peptidases and cytokines (3.5% each), phosphatases (2.5%), and G-protein receptors (1.1%). In addition, a small fraction of the Th17 DE proteins was recognized as growth factors, ligand-dependent nuclear receptors, ion channels, and translation regulators (<1% each). A major fraction of the Th17 DE proteins (>37%) were classified as others owing to their undefined functions (Figure S3A). Moreover, the IPA analysis for cellular compartment revealed four major cell locations for the proteins DE between Th17 and Th0 cells, including the cytoplasm (45.4%), nucleus (23.8%), plasma membrane

(19.9%), and extracellular space (8.2%), with only 2.8% of the DE proteins belonging to an undefined (other) cell location (Figure S3B; Table S3). Thus the enrichment of proteins assigned to various molecular functional types and their location at different cellular compartments indicate the involvement of distinct signaling pathways and transcriptional mechanisms during differentiation of Th17 cells. In conclusion, our analysis identified proteins from almost all the functional types and cellular compartments, providing a valuable resource for further investigation of candidate proteins potentially regulating the differentiation and function of Th17 cells in humans.
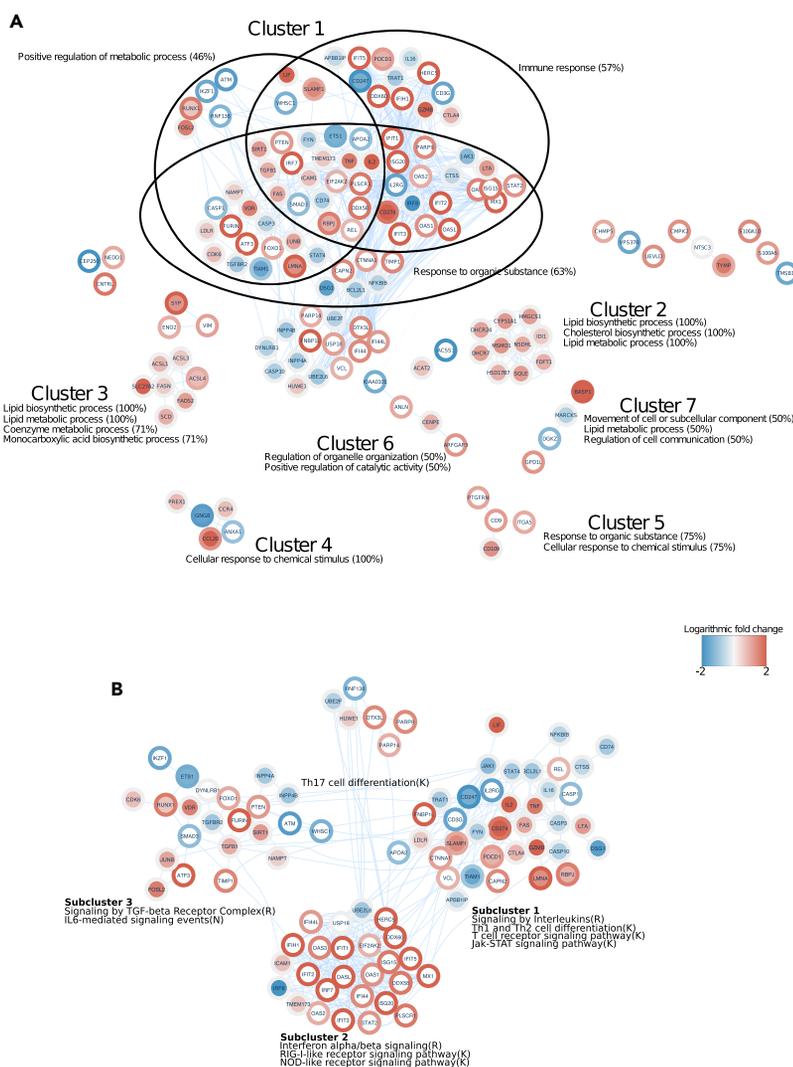
## An Interaction Network for the DE Proteins during Early Th17 Cell Polarization

To understand the relationships between proteins identified in our data, we performed a protein-protein interaction network analysis for the proteins DE in Th17 cells at both time points by mining the STRING database (Szklarczyk et al., 2017). Of the 291 unique proteins detected as DE between Th17 and Th0, 146 were found to have known experimental or predicted interactions (Figure 3). Altogether 443 interactions between the 146 proteins were detected. To further examine the possible functional elements among the interacting proteins, we identified clusters based on the interactions in the resulting network (Table S4). This yielded altogether 15 clusters, with seven of the clusters having four or more interacting proteins. The average interaction score between the cluster members in the identified clusters was high (0.77–0.93, range 0–1). To infer the functional relevance of the identified clusters, we examined the prevalence of the enriched GO biological processes within the seven clusters (Table S3). To further complement our network analysis, we included the changes in expression between Th17 and Th0 for the interacting proteins. The functional interaction analysis identified cluster 1 as the largest cluster with altogether 78 interconnected proteins related to distinct GO biological processes, including immune responses, positive regulation of cellular metabolic processes, and response to organic substances as the top three representative GO biological processes (Figure 3A). A large proportion of the proteins included in cluster 1 were either upregulated at 24 or 72 hr or downregulated at 24 hr, whereas only a small portion seemed to be downregulated at 72 hr. In addition, all the proteins of clusters 2 and 3 were related to lipid metabolic processes (Figure 3A). Furthermore, all proteins included in clusters 2 and 3 were found to be differentially regulated only at 24 hr (except for ACSL4, which was upregulated also at 72 hr).

Furthermore, to gain insight into the biological implication of expression changes in cluster 1, we analyzed the proteins in the context of biological pathways. A close examination identified three distinct functional sub-clusters within cluster 1 that were further called for functional enrichment of biological pathways (Figure 3B). The proteins in sub-clusters 1 and 3 were enriched for biological pathways involving signaling by TGF-β receptor complex, IL-6-mediated signaling events, signaling by interleukins, Th1 and Th2 cell differentiation, and the JAK-STAT signaling pathway. Interestingly, we observed a number of proteins in these two sub-clusters that are known modulators of Th17 cell differentiation, including CDK6, RUNX1, JUNB, SIRT1, RBPJ, TNF, SLAMF1, and FAS (positive regulators of Th17 cell differentiation) and FOSL2, SMAD3, ETS1, STAT4, and FYN (negative regulators of Th17 cell differentiation). Noticeably, the majority of proteins in sub-cluster 2 were upregulated in Th17 condition at 72 hr and enriched for antiviral immune pathways including IFN-α/β signaling, RIG-I-like receptor signaling pathway, and NOD-like receptor pathway (Figure 3B). The enrichment of antiviral proteins in the network at this early stage of differentiation is interesting because they are known to negatively modulate Th17 differentiation (Ramgolam et al., 2009; Ye et al., 2017). However, the expression of IRF-8 was downregulated in Th17 cells at 24 hr. IRF-8 is a known suppressor of Th17 cell differentiation in mouse (Ouyang et al., 2011). Further studies are needed to characterize the role of proteins related to antiviral pathway in human Th17 cell differentiation. To summarize, we identified protein-protein interaction networks of the proteins related to distinct biological processes and cellular pathways indicating their coordinated regulation during early Th17 cell differentiation program.

## Validation of Mass-Spectrometry-Identified Th17 DE Proteins

MS-based global proteomics approaches for protein identification rely on probability algorithms and benefit from further validation by targeted MS or other methods, such as immunoblotting or flow cytometry. We used all these three means to validate the expression profiles of selected candidate proteins. Results using western blot and flow cytometry analysis are shown in Figures 4A, 4B, and S4). We have previously shown the changes in protein expression of BASP1 (Figure 4A), LMNA (Figure 4A), and ATP1B1 (in Figure 4D) in Th17 cells, and these were used here as marker genes that further confirmed consistency between our current and previous results (Tuomela et al., 2012). Western blot analysis for the

**Figure 3. The Interaction Network among the DE Proteins between Th17 and Th0**

The network is based on all high-confidence experimental and predicted interactions from the STRING database.

(A) The network of protein-protein interactions from the DE proteins between Th17 and Th0 over both time points. The representative significant GO biological processes and cluster identities are shown only for clusters 1–7 (clusters sized >3).

(B) Cluster 1 divided into sub-clusters and the enriched pathways identified using the ReactomeFIViz Cytoscape plug-in. The letter in parentheses after each pathway name denotes the source of the annotation: R, Reactome; K, KEGG; N, NCI PID. The edges in the networks were bundled for a clearer representation. The logarithmic fold change between Th17 and Th0 at 24 hr (node inner circles) and at 72 hr (node outer circles) was implemented as continuous color mapping. For a clearer representation, the logarithmic fold changes are truncated to a range of −2 to 2.

selected Th17 upregulated proteins (JUNB, SIRT1, CTNNA1, FAS, VIM, and IRF7) and downregulated proteins (SATB1, SMAD3, STAT4, and ETS1) showed consistent results with the proteomic analysis (Figures 4A, 4B, and S4). Using flow cytometry, we also validated the expression of OASL and ATF3 at 72 after the initiation of Th17 cell polarization. Again, the results were in agreement with the proteome data showing Th17-specific increased expression of OASL and ATF3 (Figure 4C). The function of BASP1, LMNA, CTNNA1, OASL, ATF3, and VIM in modulating Th17 cell differentiation and functions remains to be studied.

In addition, we confirmed the expression of five additional candidates with unknown function in Th17 biology, namely, PALLD, ACSL4, SMTN, RDX, and FHOD1 in addition to ATP1B1 using a targeted MS-based selected reaction monitoring analysis (Figure 4D). Long-chain acyl-CoA synthetase 4 (ACSL4)
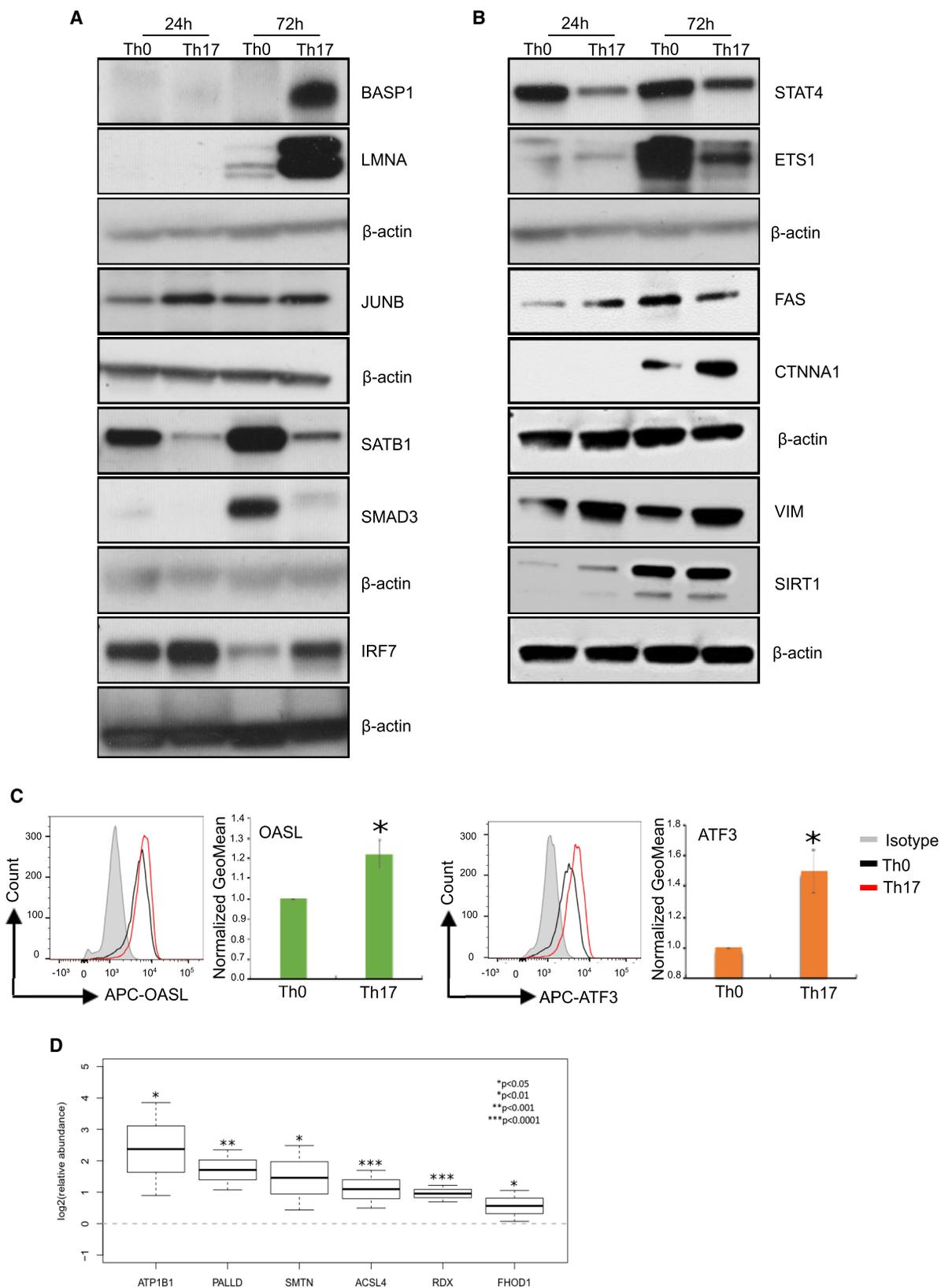
**Figure 4. Validation of the Mass-Spectrometry (MS)-Identified Proteins Using Immunoblotting, Flow Cytometry, and Targeted Mass-Spectrometry-Based Selected Reaction Monitoring (SRM) Analysis**

(A and B) Immunoblot validations of MS identified up- or downregulated proteins with known and unknown Th17-related function. Blots show the protein extracts from the Th0 and Th17 cells at 24 and 72 hr. A representative blot from three biological replicates is shown.

(C) Flow cytometry validations of OASL and ATF3 detected as differentially upregulated in Th17 cells by MS analysis. A representative histogram image from Th17 (red), Th0 (black), and isotype control (gray) as well as the mean of four biological replicates are shown. Error bars represent standard deviation values across biological replicates (n=4; *p<0.05 for Th0 versus Th17; two-tailed t-test analysis).
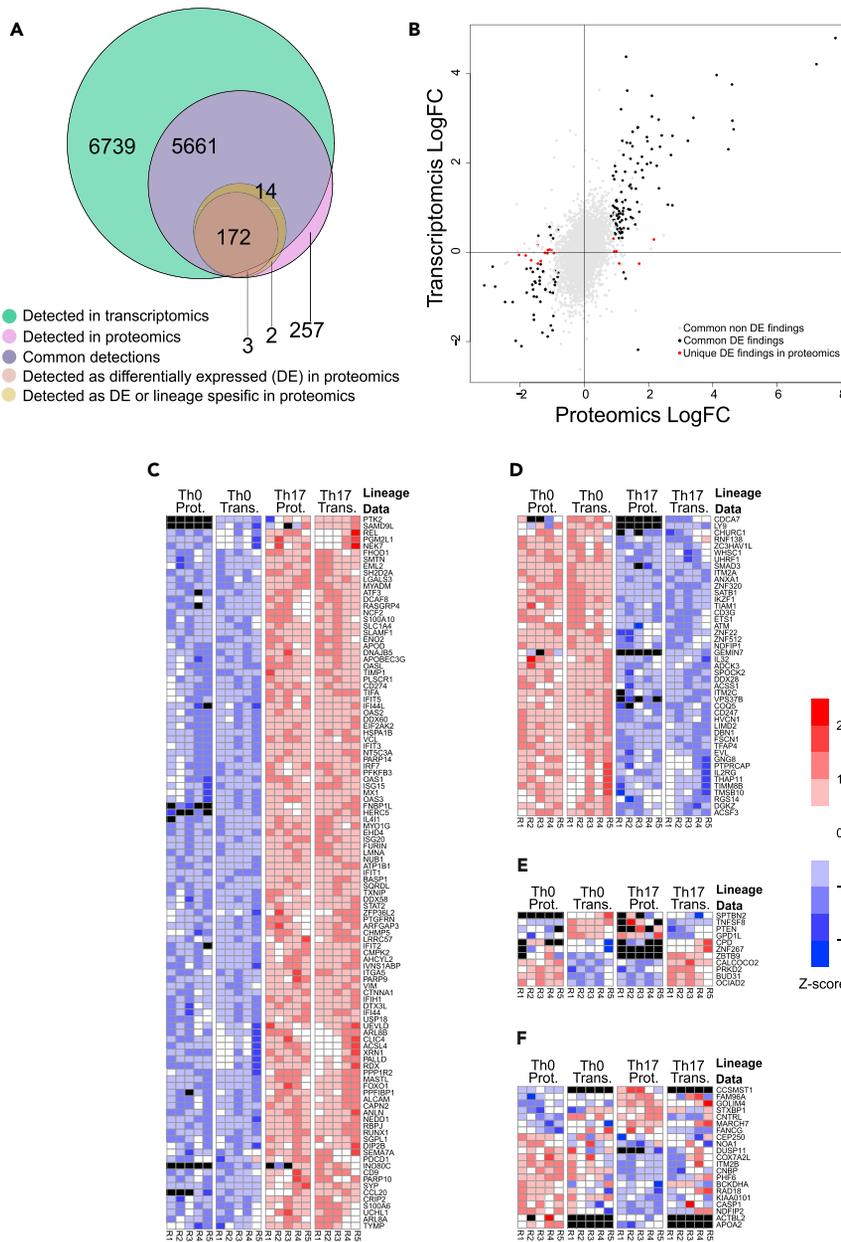
(D) Confirmation of the protein expression of six proteomics candidates, ATP1B1, PALLD, ACSL4, FHOD1, SMTN, and RDX, using targeted SRM analysis. The SRM data is from four biological replicates and represented as the mean values +/- 95% confidence interval. (n = 4; *p < 0.05, **p < 0.01, ***p < 0.0001, for Th17 versus Th0).

belongs to the class of enzymes that catalyses addition of a coenzyme-A (CoA) group to free long-chain fatty acids, specifically arachidonic acid and eicosapentaenoic acid, and converts them into acyl-CoA forms (Kang et al., 1997). As recent studies have advocated the role of lipid metabolism in regulating Th17 cell differentiation and function (Berod et al., 2014; Wang et al., 2015; Sun et al., 2017; Young et al., 2017), it will be of interest to investigate how ACSL4 regulates Th17 cell function. The formin-homology-domain-containing protein FHOD, radixin (RDX), and smoothelin (SMTN) are cytoskeleton proteins that regulate cell migration and contraction (Niessen et al., 2005; Schulze et al., 2014). Radixin is a part of membrane-associated ERM (ezrin-radixin-moesin) family proteins that plays a key role in activation and T cell homeostasis by regulating cell adhesion, migration, and mobility and promotes T cell-antigen-presenting cell interaction (Faure et al., 2004; Müller et al., 2013). The role of these proteins in regulating Th17 cell functions remains to be established. To summarize, we confirmed the protein expression changes for the selected proteins detected by proteomics analysis by alternative methods in separate, independent biological replicates. These candidates provide interesting targets for further functional studies to reveal unknown mechanisms regulating Th17 cell differentiation and function.

## Comparison of the Transcriptome and Proteome of the Differentiating Th17 Cells

To find out the degree of concordance between protein and corresponding mRNA expression, we compared proteome and RNA sequencing (RNA-seq) datasets from the same human Th0 and Th17 cells at the 72-hr time point. Altogether 12,400 transcripts were detected from RNA-seq data (Figure 5A; Table S5). Proteomics and transcriptomics showed high degree of overlap as there were transcripts corresponding to 95.5% (5,661) of the 5,923 proteins detected by LC-MS (Table S5). Approximately half (54.3%, 6,739) of the protein-coding genes detected in RNA-seq data was not detected in the corresponding proteomics data. This is consistent with recent studies on other types of human T cells showing that approximately 35%–60% of the protein-coding genes detected by RNA-seq data were not detected in the corresponding proteomic data (Cuadrado et al., 2018; Schmidt et al., 2018). To gain further insight into the Th17-cell-specific protein expression changes and their relative mRNA expression changes, we examined the overlap of differentially regulated proteins with the mRNA transcriptome data. Of total 191 differentially regulated proteins, 186 were detected in transcriptomics data, among which 172 were DE (Th17 versus Th0). Altogether 262 proteins were uniquely detected in proteomics data, of which 257 proteins were not differentially regulated between Th0 and Th17 cells. Three proteins were DE in proteomics but not detected in transcriptomics data, and two proteins were detected only in one cell type in proteomics data, but not detected in transcriptomics data (Figure 5A; Table S5). The Pearson correlation coefficient analysis of the logarithmic fold changes showed a good correlation for all the common detections between proteomics and transcriptomics (0.519, p value < 0.001, n = 5,661). Importantly, an improved correlation was observed for the common DE proteins and DE transcripts (0.825, p < 0.001, n = 154). As anticipated the DE findings unique to proteomics among the common detection showed a poor correlation (0.233, p = 0.35, n = 18) (Figure 5B, Table S5).

For a more comprehensive comparison of the proteomics and transcriptomics data, we focused our analyses on the common differentially regulated proteins (163 proteins) either detected only in one condition (nine proteins; detected either in Th0 or Th17 condition) or DE proteins (154 proteins) having corresponding DE transcripts between Th17 and Th0 (Figures 5C–5E). Noticeably, among the 163 differentially regulated proteins with corresponding detected transcripts, a large majority (93.2%) was changed in a similar fashion between mRNA and protein levels (Figures 5C and 5D; Table S5). Among these, 107 proteins and their corresponding mRNAs were consistently upregulated, including RBPJ, RUNX1, CCL20, BASP1, ATP1B1, and LMNA, and 45 proteins and their corresponding mRNAs were consistently downregulated, such as SMAD3, SATB1, ETS1, IL2RG, and Nedd4-family-interacting protein 1 (Ndfip1). However,

**Figure 5. Correspondence of the Differentially Regulated Proteins between Th17 and Th0 and the Transcriptomics Data at 72 hr**

(A) Venn diagram demonstrating the total number of common and unique detections and the total number of differentially regulated and differentially expressed (DE) proteins detected uniquely in the proteomics and commonly with the transcriptomics data. Differentially regulated proteins refer to those proteins detected as differentially expressed or detected only in one condition.

(B–E) (B) The logarithmic fold changes (logFC) of the all the common detections in proteomics and transcriptomics. The differentially expressed (DE) proteins detected as DE also in transcriptomics are marked with black and those uniquely DE in proteomics are marked with red. The Pearson correlation coefficient over the logFCs over all the common detections was 0.519 (p value < 0.001, n = 5,661), over the common DE findings 0.825 (p<0.001, n = 154), and over the DE findings unique to proteomics among the common detections 0.233 (p = 0.35, n = 18). The $Z$ score standardized expression levels of the DE proteins and the corresponding DE transcripts between Th17 and Th0 (C) upregulated in Th17 in both datasets, (D) downregulated in Th17 in both datasets, and (E) divergently regulated between datasets.

(F) The $Z$ score standardized expression levels of the DE proteins not detected as DE in transcriptomics. Black color in the heatmaps stands for undetected/missing value.

11 molecules showed opposite expression changes between mRNA and protein expression for both up- and downregulated genes at both time points (Figure 5E). This anti-correlation profile of proteins and mRNA observed in our data was also seen in another study comparing human iTreg transcriptome and proteome (Schmidt et al., 2018). Of note, 21 proteins DE between Th17 and Th0 cells were not detected as DE in the transcriptome data (Figure 5F and Table S5), suggesting that they may be post-transcriptionally regulated. Furthermore, the transcripts corresponding to CCSMST1, ACTBL2, and APOA2 proteins were not detected in transcriptomics data across all replicates. Further validation and evaluation of the differences in the expression profiles of these proteins with available good-quality antibodies such as GOLIM4, CNTRL, CASP1, RAD18, CNBP, NDFIP2, and ITM2B are required to understand the mechanisms responsible for their regulation. Thus our comparative analysis of the transcriptome and the proteome showed mostly similar, and also opposite expression, changes in the mRNA-protein levels, consistent with dynamic regulation of mRNA and protein expression. Moreover, our analysis highlighted the importance of proteomic analysis in complementing measurements on the transcriptome level.
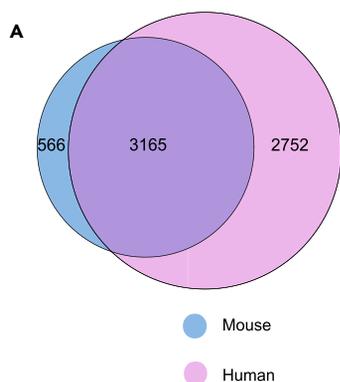
## Comparison of the Human and Mouse Proteomes during Early Th17 Cell Polarization

Mouse models have served as a useful tool for characterizing the biochemical and *in vivo* functions of genes. However, when used to model human diseases, the extent to which findings from mouse genetic models translate to humans has gained limited success. Importantly, using a systems biology approach, previous studies have revealed differences between mouse and human (Mestas and Hughes, 2004; Shay et al., 2013; Pishesha et al., 2014). Our recent study identified conserved and diversified gene signatures from the comparison of human and mouse transcriptomes during Th17 cell polarization (Tuomela et al., 2016). To further establish the level of dissimilarity, we made direct comparison with data from a recently published proteomics dataset from mouse Th17 cells cultured for 72 hr (Mohammad et al., 2018). The proteome profiles of CD4+ T cells 72 hr after TCR activation and Th17 polarization were compared using the aforementioned mouse data and the human data from the current study.

Comparison of the proteins detected in human (5,917 proteins) and mouse (3,731 proteins) revealed high degree of overall overlap as homologs for almost 85% of the mouse proteins were also detected in human (Figure 6A, Table S6). To make the comparison of the differentially regulated proteins more comprehensive, we used a threshold of FDR 0.1 (used in Mohammad et al., 2018) in both datasets to define the DE proteins. In agreement with our previous study (Tuomela et al., 2016), there was very limited overlap between the proteins differentially regulated between Th17 and Th0 cells, respectively, in human and mouse. Of the 758 and 397 proteins differentially regulated at 72 hr between Th17 and Th0 in mouse and in human, respectively, only 33 proteins were detected as differentially regulated in a similar fashion in human and mouse (Figure 6B; Table S6). Moreover, 18 proteins showed regulation in expression the opposite direction in human and mouse.

Of the 33 proteins whose expression was regulated similarly in human and mouse (Figure 6B), 15 and 18 proteins were upregulated and downregulated, respectively, in Th17 conditions at 72 hr (Figure 6C). Further analysis of the 18 proteins that showed opposite expression patterns between Th17 and Th0 in the two species at 72 hr revealed eight proteins (including DDX58, CD44, REL, and PRDX4) with upregulation in human and downregulation in mouse. The 10 proteins differentially upregulated in mouse and downregulated in human Th17 conditions included IL16, SATB1, LEF1, and LPXN (Figure 6D). The relative changes in expression of CD44 and SATB1 were validated both in human and mouse with western blot. RORγt and BATF proteins were used as polarization markers for mouse Th17 cell differentiation (Figure S5A). The immunoblot results demonstrated that the expression of these proteins was consistent with comparative analysis of quantitative proteomic data (Figures 6E and S5B).

SATB1, a chromatin organizer and transcription factor, is known to control the expression of a large number of genes involved in T cell development, activation, and differentiation both in mouse and human (Alvarez et al., 2000; Ahlfors et al., 2010; Notani et al., 2010; Burute et al., 2012). Both in mouse and human, SATB1 regulates Th2 cell differentiation by directly controlling the expression of its target genes and mediates topological looping of the transcriptionally active chromatin to the Th2 cytokine locus (Cai et al., 2006; Notani et al., 2010). FOXP3-mediated inhibition of SATB1 expression in human Treg cells is critical for their suppressive function as well as repression of genes for effector T cell differentiation (Beyer et al., 2011). Although the expression of SATB1 is conserved between Th2 and Treg cells, its opposite expression profile in human and mouse in Th17 cells suggests a different role in driving Th17 cell differentiation.
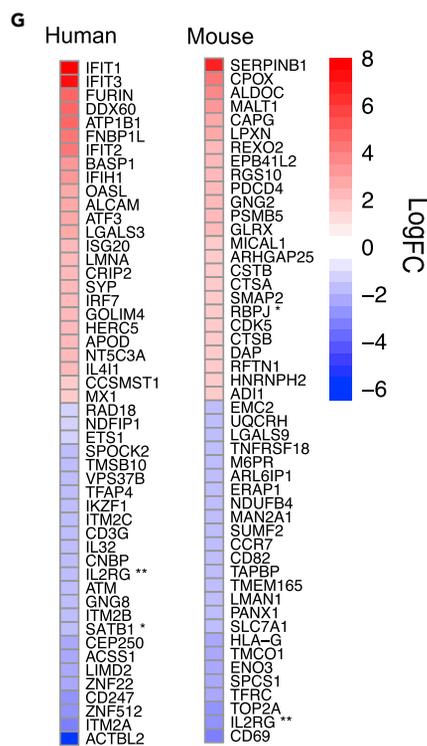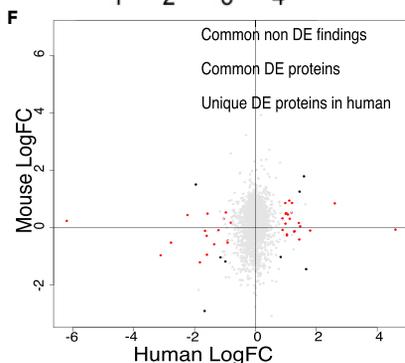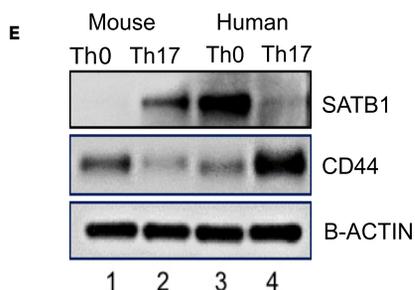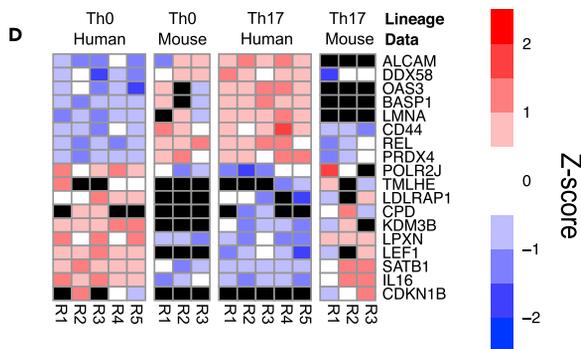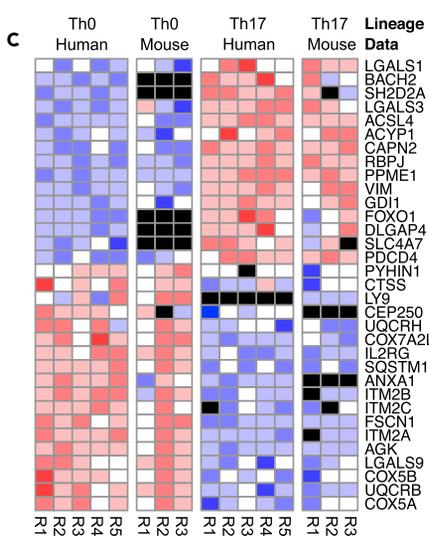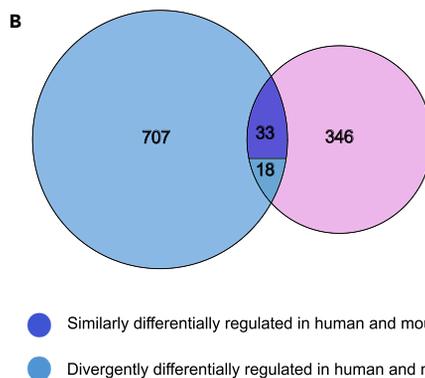
**Figure 6. Comparison of Differential Regulation between Th17 and Th0 at 72 hr in Human and Mouse**

(A–D) Venn diagrams demonstrating (A) the total number of common and unique detections in human and mouse proteomics data and (B) the number of common and unique differentially regulated proteins detected. The numbers of differentially regulated proteins behaving in the similar or opposite fashion is shown in the overlap. Differentially regulated proteins refer to those proteins detected as DE or only in one condition. The $Z$ score standardized expression levels of the commonly detected differentially regulated proteins regulated (C) similarly or (D) in an opposite fashion in human and mouse.

(E) Immunoblot validations of SATB1 and CD44 expression in human and mouse during Th17 cell polarization. Blots show the protein extracts from the Th0 and Th17 cells at 72 hr. A representative blot from three biological replicates is shown.

(F) The logarithmic fold changes (logFC) of all the common detections in human and mouse. The DE proteins detected as DE also in mouse are marked with black, and those uniquely DE in humans are marked with red. The Pearson correlation coefficient for the logFCs of all the common detections was 0.102 (p value < 0.001, n = 3,143), 0.304 (p value = 0.46, n = 8) with the common DE proteins detected, and 0.34 (p = 0.04, n = 73) for the DE findings unique to humans among the common detections.

(G) The logFCs and identities of the 25 most up- and downregulated DE proteins in humans and mouse. The common detection is marked with **. Proteins detected as DE in the other species but not among the 25 most up- and downregulated are marked with *. Mouse genes corresponding to mouse proteins are transformed as orthologous human genes for comparison using Ensembl BioMart. Black color in the heatmaps stands for undetected/missing value.

CD44 is a transmembrane glycoprotein with diverse functions in a variety of cell types (Baaten et al., 2010b). In T cells, CD44 deficiency in Th1 cells leads to compromised cell survival and abrogated generation of cellular memory to viral infection (Baaten et al., 2010a). CD44 is also required for the suppressive function of Treg cells by promoting persistent expression of FOXP3 and suppressive cytokines TGF-α1 and IL-10 (Bollyky et al., 2009). Mice lacking *Cd44* gene have an increased secretion of IL-17 and severe experimental autoimmune encephalomyelitis (EAE) disease score (Flynn et al., 2013). However, a contrasting study revealed CD44 deletion to downregulate Th1/Th17 differentiation, promote Th2 differentiation, and ameliorate clinical EAE disease score (Guan et al., 2011). Because, the majority of the understanding on the regulation and functions of CD44 in T cells have come from studies in mouse, the opposite expression profile of CD44 in human and mouse Th17 cells deserves further studies to clarify its function in humans. Our results suggest these proteins to have different roles in polarizing Th17 cells in human and mouse. Moreover, the vast majority of the proteins detected as DE between Th17 and Th0 at 72 hr after the initiation of Th17 polarization in human were not detected as DE in mouse, and vice versa. Furthermore, we detected a number of proteins in one species, but not in the other. For instance, the proteins CEP250, ANXA1, ALCAM, OAS3, and BASP1 were not detected at all in Th17 cells in mouse, whereas LY9 and CDKN1B were not detected in human Th17 cells. Moreover, the proteins, BACH2, SH2D2A, FOXO1, DLGAP4, SLC4A7, TMLHE, LDLRAP1, CPD, and KDM3B were not detected in mouse Th0 condition.

To address the level of conservation (similarity) of protein signatures between human and mouse, Pearson correlation coefficients were calculated for logarithmic fold changes (Th17 versus Th0) of the proteins detected in human and mouse. Overall, the analysis revealed very little correlation ($R^2$ = 0.102, p value < 0.001, n = 3,143), although it was somewhat better for the proteins that were DE in both human and mouse ($R^2$ = 0.304, p value = 0.46, n = 8) and for the DE proteins unique to human among the common detections ($R^2$ = 0.34, p = 0.04, n = 73) (Figure 6F, Table S6). Upon examination of 25 most up- and downregulated DE proteins in human and mouse, we found only IL2RG detected in both species. In addition, SATB1 (human) and RBPJ (mouse) were also detected as DE in both species, but not among the 25 most up- and downregulated in one of the species (Figure 6G, Table S6). Overall, our analysis identified a very low degree of similarities in DE protein signatures between human and mouse, as well as a number of proteins even with opposite expression patterns, underscoring why studies using mouse genetic models have limited success in recapitulating the results in humans.

## Repression of SATB1 Expression Contributes to Human Th17 Cell Differentiation

The expression of SATB1 in distinct Th and Treg cells is shown to be conserved between human and mouse (Ribeiro de Almeida et al., 2009; Ahlfors et al., 2010; Beyer et al., 2011). SATB1 expression has been previously reported to be upregulated in mouse Th17 cells (Ciofani et al., 2012; Mohammad et al., 2018). However, we found SATB1 transcription to be downregulated during human Th17 cell differentiation (Tuomela et al., 2012, 2016). This observation was further complemented by the findings, both at transcript and protein levels from our current transcriptomics and proteomics data and further validated by real-time quantitative PCR analysis and western blot analysis (Figures 4A, 6D, 6E, 7A, and 7B; Table S4). To examine the role of SATB1 in human Th17 cells, RNAi was used to silence its expression. Introduction of SATB1-specific small interfering RNA (siRNA) (siSATB1) in differentiating Th17 cells resulted in a significant decrease in the expression of SATB1 both at mRNA and protein levels when compared with non-targeting scramble
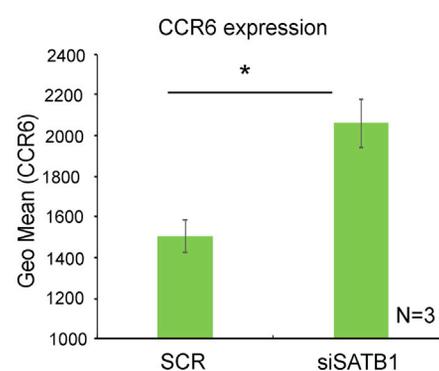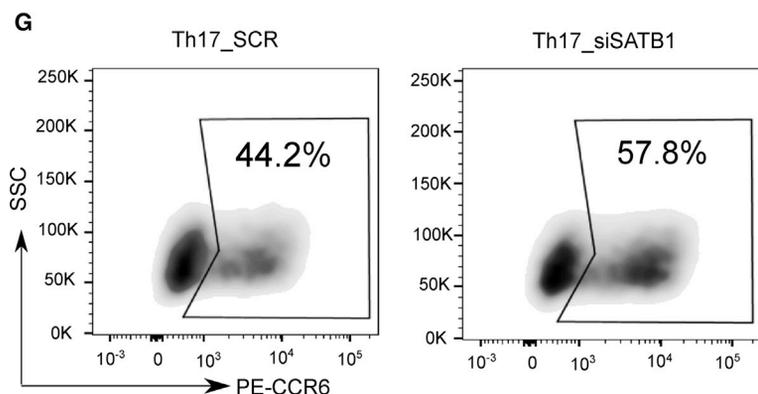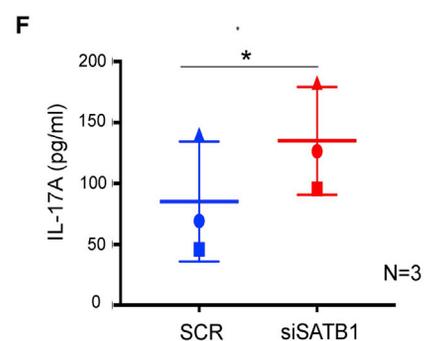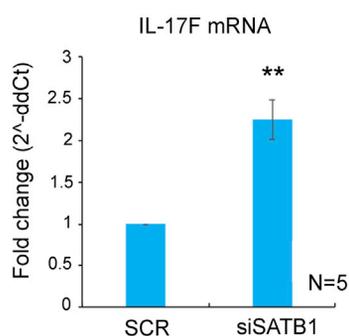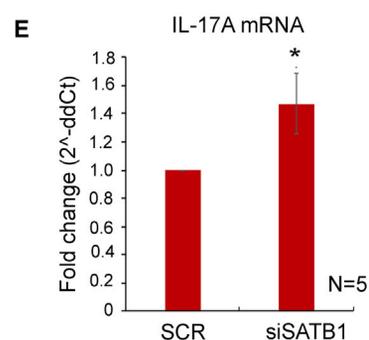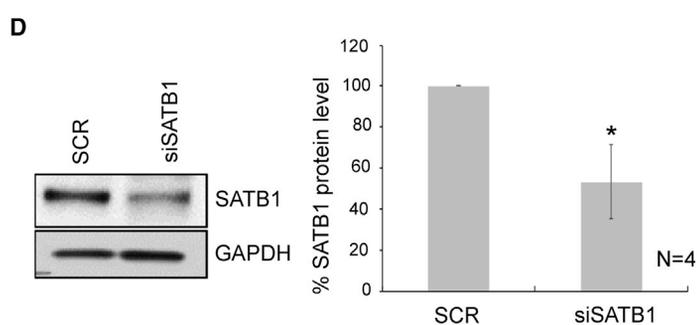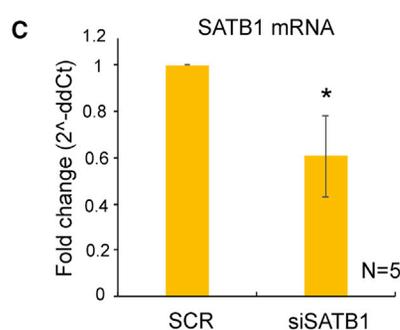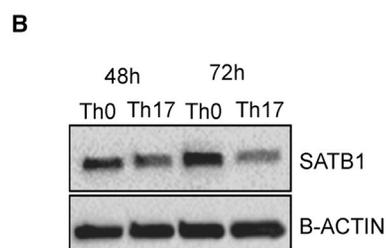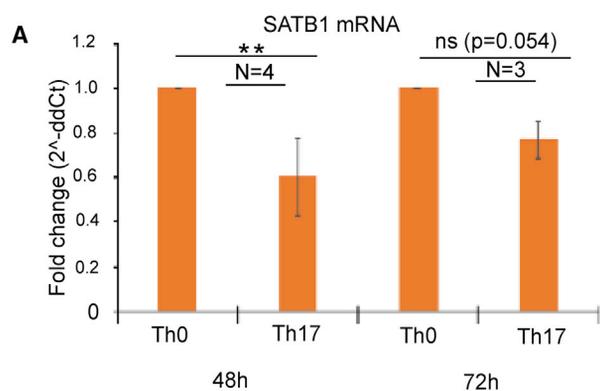
**Figure 7. SATB1 Is a Negative Regulator of Human Th17 Cell Differentiation**

The mRNA expression profile of SATB1 in Th0 and Th17 cells at 48 and 72 hr after cell activation (A). Western blot analysis of SATB1 expression in Th0 and Th17 cells at 48 and 72hr after cell activation (B). The RNA interference (RNAi)-mediated knockdown of SATB1 at 24 hr post-Th17 cell polarization using specific siRNAs. Scramble siRNA (SCR) was used as a negative control. Silencing of SATB1 at mRNA (C) and protein (D) level was measured. Effect of SATB1 knockdown on IL-17A and IL-17F mRNA levels (E) by real-time TaqMan PCR analysis and on IL-17A protein secretion (F) was quantitated by ELISA assay from SATB1 siRNA nucleofected Th17 cells relative to NT-siRNA at 72 hr post-initiation of Th17 polarization. CCR6 surface expression detected by flow cytometry was measured (G) from SATB1 siRNA nucleofected Th17 cells relative to NT-siRNA at 72 hr post-initiation of Th17 polarization. Data are representative of 3–5 independent experiments. Error bars represent standard deviation values across biological replicates (n = 3–5; *p < 0.05, **p < 0.01 for Th0 versus Th17 and SCR versus siSATB1; two-tailed t-test analysis).

siRNA (Figures 7C, 7D, and S5C). SATB1 silencing in Th17 cells led to increased transcription of marker cytokines *IL-17A* and *IL-17F* genes as well as the secretion of IL-17A cytokine (Figures 7E and 7F). In addition, depletion of SATB1 resulted in the increased expression of Th17-specific chemokine receptor, CCR6 (Figure 7G). Overall, depletion of SATB1 results in increased expression of the Th17 cell signature genes suggesting it acts as a negative regulator of Th17 cell differentiation in humans.

## DISCUSSION

The role of Th17 cells has been established in the pathogenesis of inflammatory and autoimmune diseases and cancer (Ghoreschi et al., 2011; Bailey et al., 2014). Since their discovery a decade ago, most studies have used mouse model to identify key regulators and molecular mechanisms driving Th17 cell development and function. However, the molecular mechanisms and signaling pathways that regulate human Th17 cell differentiation remain largely unclear. Our understanding of the human Th17 cell development and functions rely on studies of individual gene function or studies utilizing genome-wide transcriptional profiling to gain a holistic systems view (Tuomela et al., 2012, 2016; Gaublomme et al., 2015). However, genome-wide transcriptome analysis only describes the general characteristics of molecular changes (Stubbington et al., 2015) and does not necessarily reflect the specific phenotypic changes, which are controlled at the protein level.

In the recent years, global proteomic approaches have been adapted for the global analysis of complex proteomes in several distinct human immune cell types, including B and plasma cells (Salonen et al., 2013), macrophages and dendritic cells (Becker et al., 2012; Eligini et al., 2015; Worah et al., 2016), neutrophils (Tomazella et al., 2009), resting T cells (Howden et al., 2013) cytotoxic T cells (CTL) (Hukelmann et al., 2016; van Aalderen et al., 2017), natural killer T cells (Scheiter et al., 2013), Th1 and Th2 cells (Filén et al., 2009), and Treg cells (Kubach et al., 2007; Duguet et al., 2017; Cuadrado et al., 2018; Schmidt et al., 2018). In addition, a recent study describing the proteomes of Th1/Th17 clones derived from the gut biopsies of the patient with Crohn disease identified protein signatures of "mixed" Th1/Th17 phenotypes using MS-based protein identification and quantification (Riaz et al., 2016).

In the present study, we have used a label-free MS approach to reveal qualitative and quantitative proteomic differences between Th0 and Th17 cells during early stages of cellular stimulation and differentiation from naive CD4+ T cell subsets isolated from the umbilical cord blood. To understand the dynamics of protein expression changes during early stages of Th17 cell differentiation, we characterized proteome changes at 24 and 72 hr after initiation of the Th17 cell polarization. Despite the fact that the vast majority of the proteome is shared between TCR-activated cells (Th0) and cells primed for Th17 cell polarization at this early stage, our analyses demonstrated significant temporal and lineage-specific changes in the proteome. Overall, we observed clear differences in protein abundance associated with Th17 cell polarization.

Our analysis of the cellular proteome during early Th17 cell differentiation identified the differential expression of 291 proteins. The detected DE proteins included key protein markers used to characterize early differentiating human Th17 cells, such as cell surface markers (CCR4, CTLA-4, ICOS, and PDCD1), cytokines (CCL20 and TNF-α), and nuclear proteins (RUNX1, SIRT1, AHR, RBPJ, JUNB, REL, and FOSL2). These findings support the overall strength of the proteomic data generated and the MS-based approach implemented in this study. In addition, we discovered a number of additional proteins with significant differential expression between Th17 and Th0 cells that might have functional effects on Th17 development and function. Such examples are two glycosylphosphatidylinositol (GPI)-anchored cell surface glycoproteins, namely, semaphorin 7A (SEMA7A; also known as CD108) and CD109. SEMA7A is a negative regulator of T cell responses (Czopik et al., 2006). It has important roles in inflammatory and autoimmune disorders, including rheumatoid arthritis, multiple sclerosis, and colitis (Suzuki et al., 2007; Jongbloets et al., 2013;

Gutiérrez-Franco et al., 2016). Yet, the molecular mechanism of how SEMA7A regulates differentiation and functioning of Th17 cells is not known. Likewise, CD109 is a co-receptor for TGF-β that negatively regulates TGF-β signaling. CD109 does this by coupling with caveolin-1, which mediates internalization and degradation of TGF-β receptors via receptor-mediated endocytosis (Bizet et al., 2011) or via regulating SMAD7/Smurf2-mediated degradation of TGFBR1 (Bizet et al., 2012). The regulation of TGF-β signaling contributes to the development of Th17 cells. At lower concentrations, TGF-β supports Th17 differentiation, and at higher concentrations, it promotes the differentiation of Treg cells. Besides its role in controlling TGF-β signaling, in human keratinocytes CD109 amplifies the activation of STAT3 (Litvinov et al., 2011), a transcription factor also required for Th17 development.

Interestingly, FURIN, found to be upregulated in Th17 cells in our analysis, is a pro-protein convertase enzyme involved in the processing of CD109 into the 180-kDa secreted form and 25-kDa GPI-anchored form, suggesting that FURIN participates in regulating CD109-mediated molecular functions. In mouse, the deficiency of FURIN in T cells causes T cell expansion/activation, impaired production of TGF-β1, and breakdown of the peripheral tolerance (Pesu et al., 2008). However, the molecular mechanisms by which FURIN regulates Th17 cell development are still unclear. Consistent with these findings, Th17-specific upregulation of FURIN and CD109 in our analysis suggests that they may participate in controlling the development and function of Th17 cells via an overlapping pathway.

A number of proteins involved in the regulation of nuclear structure organization or function were found to be differentially upregulated (e.g., LMNA, MYOF, BASP1, IRF7, BHLHA15, PARP9, PARP10, PARP14, and CDK6) or downregulated (e.g., ELF1, ACTBL2, CHURC, CNBP, SATB1, IRF8, and IKZF1) in Th17 cells. These observations suggest an important role of these proteins in influencing nuclear architecture and function, which are active parts of regulating chromatin structure or transcription. In mouse, IRF8 has been shown to inhibit Th17 differentiation where IRF8 physically interacts with RORγt, and suppresses *Il-17a* transcription (Ouyang et al., 2011). The reduced expression of IRF8 in human Th17 cells when compared with Th0 cells further supports its role as a negative regulator of Th17 cell differentiation. The expression of IKZF (Ikaros) was downregulated in Th17 cells across all replicates. However, a study of Ikaros knockout mouse has shown that IKZF1 positively regulates the development and function of Th17 cells (Wong et al., 2013). The role of IKZF1 in the human system has not been reported. Further studies are needed to uncover the mechanisms and functions of these proteins in the regulation of Th17 cell development.

The mapping of protein-protein interactions is important in understanding the molecular mechanisms regulating cellular processes. The STRING analysis showed a highly cohesive protein-protein network and revealed abundant functional interactions between the proteins involving distinct GO biological processes. The majority of the protein-protein interactions in the network were related to the regulation of cellular metabolic processes with a number of the interactions involving proteins of lipid metabolic process. The role of a few proteins (e.g., FADS2 and FASN, ACC1, AHR, and HSD17B7) of the lipid metabolism pathway from our network has been recently reported to regulate Th17 function (Berod et al., 2014; Hu et al., 2015; Santori et al., 2015; Wang et al., 2015; Sun et al., 2017).

Among the validated Th17-regulated proteins, JUNB, SATB1, SMAD3, STAT4, ATF3, and ETS1 have previously been shown to have functional roles either in the context of Th17 cells or other Th cell subsets. In addition, RUNX1, ATP1B1, KDSR, LMNA, VDR, and BASP1 were validated at the protein level also in our previous report (Tuomela et al., 2012). The STAT3-dependent expression of JUNB and its role in positively regulating Th17 cell differentiation have been recently described in mouse model (Carr et al., 2017). Furthermore, SMAD3 (Martinez et al., 2009) and ETS1 (Moisan et al., 2007) negatively regulate the murine Th17 cell function. STAT4 is a major regulator of Th1 cell differentiation and function, and is a critical component in maintaining the balance between Th17 and Treg cells under disease states (Xu et al., 2011). Importantly, some of the validated proteins, including BASP1, LMNA, IRF7, ATF3, ACSL4, FHOD1, SMTN, ATP1B1, RDX, PALLD, and OASL have previously been poorly characterized or not reported to regulate Th17 cell differentiation and function and are hence candidates for functional follow-up studies. For example, we have previously reported that ATF3 directly regulates the human IFN-γ expression and promotes Th1 differentiation (Filén et al., 2010). However, the role of ATF3 in regulating the differentiation and function of human Th17 cells remains to be studied. In a mouse model, IRF7 silencing in resting CD4+ cells resulted in the increased secretion of Th17-associated cytokines (e.g., IL-17A, IL-17F, IL-21, IL-22, and IL-9), suggesting that IRF7 is a negative regulator of the Th17 response

(Tao et al., 2014). However, the molecular mechanisms responsible for the expression of IRF7 and how it regulates human Th17 cell polarization are still unclear.

Recent studies have addressed the role of lipid metabolism in regulating Th17 cell differentiation and function (Berod et al., 2014; Wang et al., 2015; Sun et al., 2017; Young et al., 2017). As an enzyme involved in lipid metabolic pathways, ACSL4 is an interesting candidate for further studies on regulation of Th17 cell differentiation. Other candidates of interest, although not related to lipid metabolism, are FHOD1, RDX, and SMTN. These structural proteins regulate cell migration and contraction (Niessen et al., 2005; Schulze et al., 2014). However, their role in Th17 cell differentiation and function remains undetermined.

A comparison between the proteomic and transcriptomic profiles can provide valuable insight into the relationship between mRNA and protein expression that may not be feasible to achieve by separate analysis of either dataset alone (Nagaraj et al., 2011; Worah et al., 2016). To that end, we compared the current proteomics data with RNA-seq data from the same samples processed for MS-based proteome analysis. Almost 95% of the proteins detected by LFQ analysis had transcripts for the corresponding protein-coding genes and displayed good correlation in their relative abundance versus Th0 (0.519) (Figure 5B) consistent with the recent studies on other cells including human Treg cells (Weekes et al., 2014; Cuadrado et al., 2018; Schmidt et al., 2018). The DE proteins showed even improved correlation (0.819). However, we detected 21 proteins DE in a Th17-specific manner that were not observed at the transcriptome level. However, of these 21 DE proteins, only two proteins ITM2B and COX7A2L were also found to be DE and downregulated in Th17 condition in our reanalysis of the mouse proteomics data from Mohammad et al., 2018. Moreover, they reported that the expression of ITM2B was DE in mouse Th17 transcriptomics analysis with opposite expression profile compared with their protein data (i.e., downregulated at the RNA level and upregulated at the protein level when Th17 polarized cells were compared with activated cells). COX7A2L was found to be downregulated in mouse both at the transcript and protein levels. The expression of CEP250 was observed as DE and downregulated in human proteomics data, whereas it was detected only in the Th0 condition in mouse proteomics data and not detected as DE in mouse transcriptomics data. Thus based on our analysis it appears that the majority of these 21 DE proteins are uniquely post-transcriptionally regulated in humans. To conclude, our study provides several candidate proteins whose roles in the regulation of Th17 differentiation and function are yet to be determined.

Comparative analysis of transcriptional changes in human and mouse during immune cell differentiation, including Th17 cell specification, revealed a conserved signature between these species (Shay et al., 2013; Tuomela et al., 2016). This includes a panel of potential therapeutic targets that are regulated in a similar fashion in human and mouse and thus suitable for follow-up studies in experimental *in vivo* mouse models. In this study, we complemented our previous study by comparing human and mouse proteomes, which demonstrated remarkable differences in the regulation of protein expression between human and mouse during Th17 cell priming. Interestingly, only 33 proteins were differentially regulated in a similar fashion, whereas 18 showed regulation to the opposite direction in human and mouse. Although consistent with our previous results comparing DE transcripts from human and mouse during Th17 cell differentiation (Tuomela et al., 2016), the lack of overlap at the protein level is even more striking. Moreover, our results are in line with other recent reports, including those from the ImmGen consortium, suggesting inter-species differences in the cellular machinery due to the operations of conservation and diversification of the *trans-* and *cis*-regulatory programs (Mestas and Hughes, 2004; Shay et al., 2013). Although the proteins regulated in a similar manner in human and mouse are good candidates for follow-up functional *in vivo* studies in experimental mouse models, further investigation on the proteins regulated in opposite manner among common DE proteins in human and mouse along with human-specific DE proteins is critical for understanding their role in regulating Th17 cell differentiation and function in humans.

The expression of SATB1 was downregulated in human Th17 cells when compared with Th0 cells, suggesting its role as a negative regulator of human Th17 cell differentiation. However, a previous mouse study reported that SATB1 positively regulates Th17 function by controlling the expression of key Th17-related genes including *Il-22, Il-17A* and *Il-17F, Il21, Il1R1,* and *Ccl20* (Ciofani et al., 2012). Opposite regulation of this key chromatin organizer (i.e., SATB1) during early specification of human and mouse Th17 cells suggests fundamental differences in the process between the two species. Moreover, recently, conditional deletion of SATB1 in mouse showed impaired production of IL-17- and IFN-γ in pathogenic T cells, and the mouse were resistant to the induction of EAE, suggesting a role of SATB1 for the generation of pathogenic Th cells (Akiba et al., 2018). Using the RNAi approach, we showed that SATB1 knockdown induced

an elevated expression of marker cytokines IL-17A and IL-17F and CCR6, a chemokine expressed on the surface of Th17 cells. This suggests SATB1 to be a negative regulator of Th17 cell priming. Further studies are required to gain insight into the molecular mechanisms through which SATB1 controls Th17 differentiation and function in humans.

In summary, our results revealed differences in the proteome of the cell types and time points studied and shed light on various perspectives of the early regulation of human Th17 cell priming. Beyond previous transcriptomics applications describing the regulatory programs that govern Th17 cell development and function in human and mouse, our study reports comprehensive proteomics data describing the initiation of human Th17 cell priming of naive CD4+ T cells derived from the umbilical cord blood. Hence, this study complements previous investigations to understand human Th17 cell priming. This study has resulted in the identification of a panel of proteins with unknown or poorly characterized functions in the context of Th17 development and function. We provide several previously unknown candidate proteins for further studies to reveal their influence on human Th17 specification and immunomodulatory and therapeutic potential for Th17-mediated immune diseases.

## Limitations of the Study

This study describes the global cellular proteomic landscape during early human Th17 cell differentiation. However, the current data only represents unfractionated cellular lysates. The fractionation of the cellular compartments could have provided further insight into the differentiation process. Furthermore, mechanistic follow-up studies on key findings reported here are required and are interesting targets for future studies. Our comparison of human to mouse Th17 proteomics data showed very limited overlap. It would be important to consider the kinetics of the process in the mouse—here we compared our data with data from a reported parallel study in mouse, where only one time point had been analyzed. It will be important to functionally validate in the mouse system the few findings with consistent results in human. Understanding the molecular mechanisms of proteins regulated in a similar fashion in human and mouse is crucial for translational research.

## Ethical Aspects

The usage of blood of unknown donors was approved by the Ethics Committee of the Hospital District of Southwest Finland. Experiments using animals' cells were in accordance with the relevant University Animal Welfare guidelines.

## METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

## DATA AND SOFTWARE AVAILABILITY

The PRIDE accession number (Vizcaíno et al., 2016) for the mass spectrometry proteomic profiling data presented in this paper is PXD008973. The selected reaction monitoring (SRM) validation data are available through PASSEL with the dataset identifier PASS01204. The RNAseq data from this study is submitted to the Gene Expression Omnibus (GEO) with identifier GSE118974.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, five figures, and seven tables and can be found with this article online at https://doi.org/10.1016/j.isci.2018.12.020.

## ACKNOWLEDGMENTS

## AUTHOR CONTRIBUTIONS

## DECLARATION OF INTERESTS

## REFERENCES

van Aalderen, M.C., van den Biggelaar, M., Remmerswaal, E.B.M., van Alphen, F.P.J., Meijer, A.B., Ten Berge, I.J.M., and van Lier, R.A.W. (2017). Label-free analysis of CD8+T cell subset proteomes supports a progressive differentiation model of human-virus-specific T cells. Cell Rep. *19*, 1068–1079.

Ahlfors, H., Limaye, A., Elo, L.L., Tuomela, S., Burute, M., Gottimukkala, K.V., Notani, D., Rasool, O., Galande, S., and Lahesmaa, R. (2010). SATB1 dictates expression of multiple genes including IL-5 involved in human T helper cell differentiation. Blood *116*, 1443–1453.

Akiba, Y., Kuwabara, T., Mukozu, T., Mikami, T., and Kondo, M. (2018). SATB1 is required for the development of experimental autoimmune encephalomyelitis to maintain T cell receptor responsiveness. Microbiol. Immunol. *62*, 255–268.

Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev. *14*, 521–535.

Baaten, B.J.G., Li, C.R., Deiro, M.F., Lin, M.M., Linton, P.J., and Bradley, L.M. (2010a). CD44 regulates survival and memory development in Th1 cells. Immunity *32*, 104–115.

Baaten, B.J.G., Li, C.R., and Bradley, L.M. (2010b). Multifaceted regulation of T cells by CD44. Commun. Integr. Biol. *3*, 508–512.

Bailey, S.R., Nelson, M.H., Himes, R.A., Li, Z., Mehrotra, S., and Paulos, C.M. (2014). Th17 cells in cancer: the ultimate identity crisis. Front. Immunol. *5*, 276.

Becker, L., Liu, N.C., Averill, M.M., Yuan, W., Pamir, N., Peng, Y., Irwin, A.D., Fu, X., Bornfeldt, K.E., and Heinecke, J.W. (2012). Unique proteomic signatures distinguish macrophages and dendritic cells. PLoS One *7*, e33297.

Berod, L., Friedrich, C., Nandan, A., Freitag, J., Hagemann, S., Harmrolfs, K., Sandouk, A., Hesse, C., Castro, C.N., Bähre, H., et al. (2014). De novo fatty acid synthesis controls the fate between regulatory T and T helper 17 cells. Nat. Med. *20*, 1327–1333.

Beyer, M., Thabet, Y., Müller, R.U., Sadlon, T., Classen, S., Lahl, K., Basu, S., Zhou, X., Bailey-Bucktrout, S.L., Krebs, W., et al. (2011). Repression of the genome organizer SATB1 in regulatory T cells is required for suppressive function and inhibition of effector differentiation. Nat. Immunol. *12*, 898–907.

Bizet, A.A., Liu, K., Tran-Khanh, N., Saksena, A., Vorstenbosch, J., Finnson, K.W., Buschmann, M.D., and Philip, A. (2011). The TGF-β co-receptor, CD109, promotes internalization and degradation of TGF-β receptors. Biochim. Biophys. Acta *1813*, 742–753.

Bizet, A.A., Tran-Khanh, N., Saksena, A., Liu, K., Buschmann, M.D., and Philip, A. (2012). CD109-mediated degradation of TGF-?? receptors and inhibition of TGF-?? responses involve regulation of SMAD7 and Smurf2 localization and function. J. Cell Biochem. *113*, 238–246.

Bollyky, P.L., Falk, B.A., Long, S.A., Preisinger, A., Braun, K.R., Wu, R.P., Evanko, S.P., Buckner, J.H., Wight, T.N., and Nepom, G.T. (2009). CD44 costimulation promotes FoxP3+ regulatory T cell persistence and function via production of IL-2, IL-10, and TGF-beta. J. Immunol. *183*, 2232–2241.

Burkett, P.R., Zu Horste, G.M., and Kuchroo, V.K. (2015). Pouring fuel on the fire: Th17 cells, the environment, and autoimmunity. J. Clin. Invest. *125*, 2211–2219.

Burute, M., Gottimukkala, K., and Galande, S. (2012). Chromatin organizer SATB1 is an important determinant of T-cell differentiation. Immunol. Cell Biol. *90*, 852–859.

Cai, S., Lee, C.C., and Kohwi-Shigematsu, T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat. Genet. *38*, 1278–1288.

Carr, T.M., Wheaton, J.D., Houtz, G.M., and Ciofani, M. (2017). JunB promotes Th17 cell identity and restrains alternative CD4+T-cell programs during inflammation. Nat. Commun. *8*, 301.

Chang, S.H., Chung, Y., and Dong, C. (2010). Vitamin D suppresses Th17 cytokine production by inducing C/EBP Homologous Protein (CHOP) expression. J. Biol. Chem. *285*, 38751–38755.

Chen, G., Hardy, K., Pagler, E., Ma, L., Lee, S., Gerondakis, S., Daley, S., and Shannon, M.F. (2011). The NF-κB transcription factor c-Rel is required for Th17 effector cell development in experimental autoimmune encephalomyelitis. J. Immunol. *187*, 4483–4491.

Ciofani, M., Madar, A., Galan, C., Sellars, M., Mace, K., Pauli, F., Agarwal, A., Huang, W., Parkhurst, C.N., Muratet, M., et al. (2012). A validated regulatory network for Th17 cell specification. Cell *151*, 289–303.

Cosmi, L., Maggi, L., Santarlasci, V., Liotta, F., and Annunziato, F. (2014). T helper cells plasticity in inflammation. Cytometry A *85*, 36–42.

Cox, J., and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. Annu. Rev. Biochem. *80*, 273–299.

Cuadrado, E., van den Biggelaar, M., de Kivit, S., Chen, Y.-Y., Slot, M., Doubal, I., Meijer, A., van Lier, R.A.W., Borst, J., and Amsen, D. (2018). Proteomic analyses of human regulatory T cells reveal adaptations in signaling pathways that protect cellular identity. Immunity *48*, 1046–1059.e6.

Czopik, A.K., Bynoe, M.S., Palm, N., Raine, C.S., and Medzhitov, R. (2006). Semaphorin 7A is a negative regulator of T cell responses. Immunity *24*, 591–600.

Duguet, F., Locard-Paulet, M., Marcellin, M., Chaoui, K., Bernard, I., Andreoletti, O., Lesourne, R., Burlet-Schiltz, O., Gonzalez de Peredo, A., Saoudi, A., et al. (2017). Proteomic analysis of regulatory T cells reveals the importance of Themis1 in the control of their suppressive function. Mol. Cell. Proteomics *16*, 1416–1432.

Eligini, S., Brioschi, M., Fiorelli, S., Tremoli, E., Banfi, C., and Colli, S. (2015). Human monocyte-derived macrophages are heterogenous: proteomic profile of different phenotypes. J. Proteomics *124*, 112–123.

Eriksson, P., Andersson, C., Cassel, P., Nyström, S., and Ernerudh, J. (2015). Increase in Th17-associated CCL20 and decrease in Th2-associated CCL22 plasma chemokines in active ANCA-associated vasculitis. Scand. J. Rheumatol. *44*, 80–83.

Faure, S., Salazar-Fontana, L.I., Semichon, M., Tybulewicz, V.L., Bismuth, G., Trautmann, A., Germain, R.N., and Delon, J. (2004). ERM proteins regulate cytoskeleton relaxation promoting T cell-APC conjugation. Nat. Immunol. *5*, 272–279.

Filén, J.-J., Filén, S., Moulder, R., Tuomela, S., Ahlfors, H., West, A., Kouvonen, P., Kantola, S., Björkman, M., Katajamaa, M., et al. (2009). Quantitative proteomics reveals GIMAP family proteins 1 and 4 to be differentially regulated during human T helper cell differentiation. Mol. Cell. Proteomics *8*, 32–44.

Filén, S., Ylikoski, E., Tripathi, S., West, A., Björkman, M., Nyström, J., Ahlfors, H., Coffey, E., Rao, K.V., Rasool, O., and Lahesmaa, R. (2010). Activating transcription factor 3 is a positive regulator of human IFNG gene expression. J. Immunol. *184*, 4990–4999.

Flynn, K.M., Michaud, M., and Madri, J.A. (2013). CD44 deficiency contributes to enhanced experimental autoimmune encephalomyelitis: a role in immune cells and vascular cells of the blood-brain barrier. Am. J. Pathol. *182*, 1322–1336.

Gaublomme, J.T., Yosef, N., Lee, Y., Gertner, R.S., Yang, L.V., Wu, C., Pandolfi, P.P., Mak, T., Satija, R., and Shalek, A.K. (2015). Single-cell genomics unveils critical regulators of Th17 cell pathogenicity. Cell *163*, 1400–1412.

Ghoreschi, K., Laurence, A., Yang, X.P., Hirahara, K., and O'Shea, J.J. (2011). T helper 17 cell heterogeneity and pathogenicity in autoimmune disease. Trends Immunol. *32*, 395–401.

Guan, H., Nagarkatti, P.S., and Nagarkatti, M. (2011). CD44 reciprocally regulates the differentiation of encephalitogenic Th1/Th17 and Th2/regulatory T cells through epigenetic modulation involving DNA methylation of cytokine gene promoters, thereby controlling the development of experimental autoimmune encephalomyelitis. J. Immunol. *186*, 6955–6964.

Gutiérrez-Franco, A., Costa, C., Eixarch, H., Castillo, M., Medina-Rodríguez, E.M., Bribián, A., de Castro, F., Montalban, X., and Espejo, C. (2016). Differential expression of sema3A and sema7A in a murine model of multiple sclerosis: Implications for a therapeutic design. Clin. Immunol. *163*, 22–33.

Hernández-Santos, N., and Gaffen, S.L. (2012). Th17 cells in immunity to *Candida albicans*. Cell Host Microbe *11*, 425–435.

Howden, A.J.M., Geoghegan, V., Katsch, K., Efstathiou, G., Bhushan, B., Boutureira, O., Thomas, B., Trudgian, D.C., Kessler, B.M., Dieterich, D.C., et al. (2013). QuaNCAT: Quantitating proteome dynamics in primary cells. Nat. Methods *10*, 343–346.

Hu, X., Wang, Y., Hao, L.Y., Liu, X., Lesch, C.A., Sanchez, B.M., Wendling, J.M., Morgan, R.W., Aicher, T.D., Carter, L.L., et al. (2015). Sterol metabolism controls TH17 differentiation by generating endogenous RORγ agonists. Nat. Chem. Biol. *11*, 141–147.

Hukelmann, J.L., Anderson, K.E., Sinclair, L.V., Grzes, K.M., Murillo, A.B., Hawkins, P.T., Lamond, A.I., and Cantrell, D.A. (2016). The cytotoxic T cell proteome and its shaping by the kinase mTOR. Nat. Immunol. *17*, 104–112.

Jongbloets, B.C., Ramakers, G.M.J., and Pasterkamp, R.J. (2013). Semaphorin7A and its receptors: Pleiotropic regulators of immune cell function, bone homeostasis, and neural development. Semin. Cell Dev. Biol. *24*, 129–138.

Kang, M.J., Fujino, T., Sasano, H., Minekura, H., Yabuki, N., Nagura, H., Iijima, H., and Yamamoto, T.T. (1997). A novel arachidonate-preferring acyl-CoA synthetase is present in steroidogenic cells of the rat adrenal, ovary, and testis. Proc. Natl. Acad. Sci. U S A *94*, 2880–2884.

Kubach, J., Lutter, P., Bopp, T., Stoll, S., Becker, C., Huter, E., Richter, C., Weingarten, P., Warger, T., Knop, J., et al. (2007). Human CD4+CD25+ regulatory T cells: proteome analysis identifies galectin-10 as a novel marker essential for their anergy and suppressive function. Blood *110*, 1550–1558.

Lim, H.W., Kang, S.G., Ryu, J.K., Schilling, B., Fei, M., Lee, I.S., Kehasse, A., Shirakawa, K., Yokoyama, M., Schnölzer, M., et al. (2015). SIRT1 deacetylates ROR t and enhances Th17 cell generation. J. Exp. Med. *212*, 607–617.

Litvinov, I.V., Bizet, A.A., Binamer, Y., Jones, D.A., Sasseville, D., and Philip, A. (2011). CD109 release from the cell surface in human keratinocytes regulates TGF-beta receptor expression, TGF-beta signalling and STAT3 activation: relevance to psoriasis. Exp. Dermatol. *20*, 627–632.

Liu, Y., Beyer, A., and Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. Cell *165*, 535–550.

Loyet, K.M., Ouyang, W., Eaton, D.L., and Stults, J.T. (2005). Proteomic profiling of surface proteins on Th1 and Th2 cells. J. Proteome Res. *4*, 400–409.

Mak, I.W.Y., Evaniew, N., and Ghert, M. (2014). Lost in translation: Animal models and clinical trials in cancer treatment. Am. J. Transl. Res. *6*, 114–118.

Martinez, G.J., Zhang, Z., Chung, Y., Reynolds, J.M., Lin, X., Jetten, A.M., Feng, X.H., and Dong, C. (2009). Smad3 differentially regulates the induction of regulatory and inflammatory T cell differentiation. J. Biol. Chem. *284*, 35283–35286.

McKinstry, K.K., Strutt, T.M., and Swain, S.L. (2010). Regulation of CD4+ T-cell contraction during pathogen challenge. Immunol. Rev. *236*, 110–124.

Mehrotra, P., Krishnamurthy, P., Sun, J., Goenka, S., and Kaplan, M.H. (2015). Poly-ADP-ribosyl polymerase-14 promotes T helper 17 and follicular T helper development. Immunology *146*, 537–546.

Mestas, J., and Hughes, C.C.W. (2004). Of mice and not men: differences between mouse and human immunology. J. Immunol. *172*, 2731–2738.

Meyer zu Horste, G., Wu, C., Wang, C., Cong, L., Pawlak, M., Lee, Y., Elyaman, W., Xiao, S., Regev, A., and Kuchroo, V.K. (2016). RBPJ controls development of pathogenic Th17 cells by regulating IL-23 receptor expression. Cell Rep. *16*, 392–404.

Mohammad, I., Nousiainen, K., Bhosale, S.D., Starskaia, I., Moulder, R., Rokka, A., Cheng, F., Mohanasundaram, P., Eriksson, J.E., Goodlett, D.R., et al. (2018). Quantitative proteomic characterization and comparison of T helper 17 and induced regulatory T cells. PLoS Biol. *16*, e2004194.

Moisan, J., Grenningloh, R., Bettelli, E., Oukka, M., and Ho, I.C. (2007). Ets-1 is a negative regulator of Th17 differentiation. J. Exp. Med. *204*, 2825–2835.

Müller, N., Fischer, H.J., Tischner, D., van den Brandt, J., and Reichardt, H.M. (2013). Glucocorticoids induce effector T cell depolarization via ERM proteins, thereby impeding migration and APC conjugation. J. Immunol. *190*, 4360–4370.

Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. *7*, 548.

Niessen, P., Rensen, S., van Deursen, J., De Man, J., De Laet, A., Vanderwinden, J.M., Wedel, T., Baker, D., Doevendans, P., Hofker, M., et al. (2005). Smoothelin-A is essential for functional intestinal smooth muscle contractility in mice. Gastroenterology *129*, 1592–1601.

Notani, D., Gottimukkala, K.P., Jayani, R.S., Limaye, A.S., Damle, M.V., Mehta, S., Purbey, P.K., Joseph, J., and Galande, S. (2010). Global regulator SATB1 recruits β-catenin and regulates TH2 differentiation in Wnt-dependent manner. PLoS Biol. *8*, e1000296.

O'Shea, J., and Paul, W.E. (2010). Mechanisms underlying lineage commitment and plasticity of helper CD4 + T cells. Science *327*, 1098–1102.

Ouyang, X., Zhang, R., Yang, J., Li, Q., Qin, L., Zhu, C., Liu, J., Ning, H., Shin, M.S., Gupta, M., et al. (2011). Transcription factor IRF8 directs a silencing programme for TH17 cell differentiation. Nat. Commun. *2*, 314.

Pesu, M., Watford, W.T., Wei, L., Xu, L., Fuss, I., Strober, W., Andersson, J., Shevach, E.M., Quezado, M., Bouladoux, N., et al. (2008). T-cell-expressed proprotein convertase furin is essential for maintenance of peripheral immune tolerance. Nature *455*, 246–250.

Pishesha, N., Thiru, P., Shi, J., Eng, J.C., Sankaran, V.G., and Lodish, H.F. (2014). Transcriptional divergence and conservation of human and mouse erythropoiesis. Proc. Natl. Acad. Sci. U S A *111*, 4103–4108.

Ramgolam, V.S., Sha, Y., Jin, J., Zhng, X., and Markovic-Plese, S. (2009). IFN-beta inhibits human Th17 cell differentiation. J. Immunol. *183*, 5418–5427.

Rautajoki, K.J., Marttila, E.M., Nyman, T.A., and Lahesmaa, R. (2007). Interleukin-4 inhibits caspase-3 by regulating several proteins in the Fas pathway during initial stages of human T helper 2 cell differentiation. Mol. Cell. Proteomics *6*, 238–251.

Riaz, T., Sollid, L.M., Olsen, I., and de Souza, G.A. (2016). Quantitative proteomics of gut-derived Th1 and Th1/Th17 clones reveal the presence of CD28+ NKG2D- Th1 cytotoxic CD4+ T cells. Mol. Cell. Proteomics *15*, 1007–1016.

Ribeiro de Almeida, C., Heath, H., Krpic, S., Dingjan, G.M., van Hamburg, J.P., Bergen, I., van de Nobelen, S., Sleutels, F., Grosveld, F., Galjart, N., and Hendriks, R.W. (2009). Critical role for the transcription regulator CCCTC-binding factor in the control of Th2 cytokine expression. J. Immunol. *182*, 999–1010.

Salonen, J., Rönnholm, G., Kalkkinen, N., and Vihinen, M. (2013). Proteomic changes during B cell maturation: 2D-DIGE approach. PLoS One *8*, e77894.

Santori, F.R., Huang, P., van de Pavert, S.A., Douglass, E.F., Jr., Leaver, D.J., Haubrich, B.A., Keber, R., Lorbek, G., Konijn, T., and Rosales, B.N. (2015). Identification of natural RORγ ligands that regulate the development of lymphoid cells. Cell Metab. *21*, 286–297.

Scheiter, M., Lau, U., van Ham, M., Bulitta, B., Gröbe, L., Garritsen, H., Klawonn, F., König, S., and Jänsch, L. (2013). Proteome analysis of distinct developmental stages of human natural killer (NK) cells. Mol. Cell. Proteomics *12*, 1099–1114.

Schmidt, A., Marabita, F., Kiani, N.A., Gross, C.C., Johansson, H.J., Éliás, S., Rautio, S., Eriksson, M., Fernandes, S.J., Silberberg, G., et al. (2018). Time-resolved transcriptome and proteome landscape of human regulatory T cell (Treg) differentiation reveals novel regulators of FOXP3. BMC Biol. *16*, 47.

Schulze, N., Graessl, M., Blancke Soares, A., Geyer, M., Dehmelt, L., and Nalbant, P. (2014). FHOD1 regulates stress fiber organization by controlling the dynamics of transverse arcs and dorsal fibers. J. Cell Sci. *127*, 1379–1393.

Schwanhüusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Shay, T., Jojic, V., Zuk, O., Rothamel, K., Puyraimond-Zemmour, D., Feng, T., Wakamatsu, E., Benoist, C., Koller, D., and Regev, A.; ImmGen Consortium (2013). Conservation and divergence in the transcriptional programs of the human and mouse immune systems. Proc. Natl. Acad. Sci. U S A *110*, 2946–2951.

Stubbington, M.J.T., Mahata, B., Svensson, V., Deonarine, A., Nissen, J.K., Betz, A.G., and Teichmann, S.A. (2015). An atlas of mouse CD4+ T cell transcriptomes. Biol. Direct *10*, 14.

Sun, L., Fu, J., and Zhou, Y. (2017). Metabolism controls the balance of Th17/T-regulatory cells. Front. Immunol. *8*, 1632.

Sun, X., Yamada, H., Shibata, K., Muta, H., Tani, K., Podack, E.R., and Yoshikai, Y. (2010). CD30 ligand/CD30 plays a critical role in Th17 differentiation in mice. J. Immunol. *185*, 2222–2230.

Suzuki, K., Okuno, T., Yamamoto, M., Pasterkamp, R.J., Takegahara, N., Takamatsu, H., Kitao, T., Takagi, J., Rennert, P.D., Kolodkin, A.L., et al. (2007). Semaphorin 7A initiates T-cell-mediated inflammatory responses through α1β1 integrin. Nature *446*, 680–684.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. *45*, D362–D368.

Tao, Y., Zhang, X., Chopra, M., Kim, M.J., Buch, K.R., Kong, D., Jin, J., Tang, Y., Zhu, H., Jewells, V., and Markovic-Plese, S. (2014). The role of endogenous IFN-beta in the regulation of Th17 responses in patients with relapsing-remitting multiple sclerosis. J. Immunol. *192*, 5610–5617.

Tomazella, G.G., da Silva, I., Laure, H.J., Rosa, J.C., Chammas, R., Wiker, H.G., de Souza, G.A., and Greene, L.J. (2009). Proteomic analysis of total cellular proteins of human neutrophils. Proteome Sci. *7*, 32.

Tuomela, S., Salo, V., Tripathi, S.K., Chen, Z., Laurila, K., Gupta, B., Äijö, T., Oikari, L., Stockinger, B., Lähdesmäki, H., and Lahesmaa, R. (2012). Identification of early gene expression changes during human Th17 cell differentiation. Blood *119*, e151–e160.

Tuomela, S., Rautio, S., Ahlfors, H., Öling, V., Salo, V., Ullah, U., Chen, Z., Hämälistö, S., Tripathi, S.K., Äijö, T., et al. (2016). Comparative analysis of human and mouse transcriptomes of Th17 cell priming. Oncotarget *7*, 13416–13428.

Veldhoen, M., Hirota, K., Westendorf, A.M., Buer, J., Dumoutier, L., Renauld, J.C., and Stockinger, B. (2008). The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. Nature *453*, 106–109.

Vizcaíno, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. *44*, D447–D456.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. *13*, 227–232.

Wang, C., Yosef, N., Gaublomme, J., Wu, C., Lee, Y., Clish, C.B., Kaminski, J., Xiao, S., Meyer Zu Horste, G., Pawlak, M., et al. (2015). CD5L/AIM regulates lipid biosynthesis and restrains Th17 cell pathogenicity. Cell *163*, 1413–1427.

Weekes, M.P., Tomasec, P., Huttlin, E.L., Fielding, C.A., Nusinow, D., Stanton, R.J., Wang, E.C., Aicheler, R., Murrell, I., Wilkinson, G.W., et al. (2014). Quantitative temporal viromics: an approach to investigate host-pathogen interaction. Cell *157*, 1460–1472.

Wong, L.Y., Hatfield, J.K., and Brown, M.A. (2013). Ikaros sets the potential for Th17 lineage gene expression through effects on chromatin state in early T Cell Development. J. Biol. Chem. *288*, 35170–35179.

Worah, K., Mathan, T.S.M., Vu Manh, T.P., Keerthikumar, S., Schreibelt, G., Tel, J., Duiveman-de Boer, T., Sköld, A.E., van Spriel, A.B., de Vries, I.J.M., et al. (2016). Proteomics of human dendritic cell subsets reveals subset-

specific surface markers and differential inflammasome function. Cell Rep. *16*, 2953–2966.

Xu, J., Yang, Y., Qiu, G., Lal, G., Yin, N., Wu, Z., Bromberg, J.S., and Ding, Y. (2011). Stat4 is critical for the balance between Th17 cells and regulatory T cells in colitis. J. Immunol. *186*, 6597–6606.

Ye, J., Wang, Y., Liu, X., Li, L., Opejin, A., Hsueh, E.C., Luo, H., Wang, T., Hawiger, D., and Peng, G. (2017). TLR7 signaling regulates Th17 cells and autoimmunity: novel potential for autoimmune therapy. J. Immunol. *199*, 941–954.

Yosef, N., Shalek, A.K., Gaublomme, J.T., Jin, H., Lee, Y., Awasthi, A., Wu, C., Karwacz, K., Xiao, S., Jorgolli, M., et al. (2013). Dynamic regulatory

network controlling TH17 cell differentiation. Nature *496*, 461–468.

Young, K.E., Flaherty, S., Woodman, K.M., Sharma-Walia, N., and Reynolds, J.M. (2017). Fatty acid synthase regulates the pathogenicity of Th17 cells. J. Leukoc. Biol. *102*, 1229–1235.

**Supplemental Information**

**Quantitative Proteomics Reveals the Dynamic**

**Protein Landscape during Initiation of Human Th17**

**Cell Polarization**

Subhash K. Tripathi, Tommi Välikangas, Ankitha Shetty, Mohd Moin Khan, Robert Moulder, Santosh D. Bhosale, Elina Komsi, Verna Salo, Rafael Sales De Albuquerque, Omid Rasool, Sanjeev Galande, Laura L. Elo, and Riitta Lahesmaa

# Supplemental Information

**Title: Quantitative proteomics reveals the dynamic protein landscape during initiation of human Th17 cell polarization**

**One Sentence Summary**: Proteome analysis of early human Th17 cell differentiation

**Authors:** Subhash K. Tripathi[1±], Tommi Välikangas[1,2±], Ankitha Shetty[1,4±], Mohd Moin Khan[1,3], Santosh D Bhosale[1,3], Robert Moulder[1], Elina Komsi[1], Verna Salo[1,3], Rafael Sales De Albuquerque[1], Omid Rasool[1], Sanjeev Galande[4], Laura L. Elo[*1±] and Riitta Lahesmaa[*1±]

# Supplementary figures and legends

## Supplementary Figure 1 (related to figure 1)
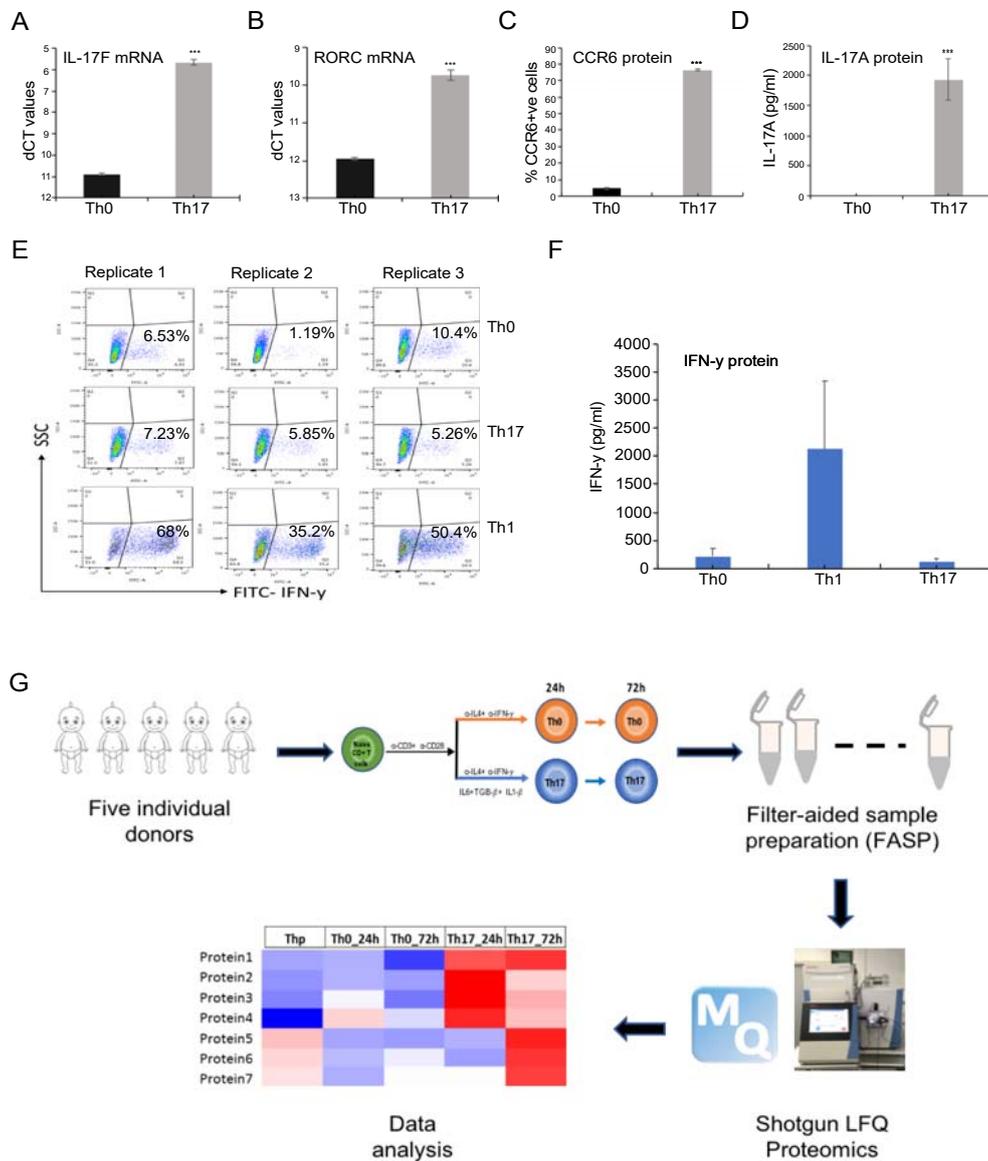


**Figure S1.** The expression of Th17 markers in Th17 cultures and a schematic representation of the study setting. (A) IL17F mRNA expression, (B) RORC mRNA expression, (C) CCR6 protein expression, and (D) IL17A protein secretion from the cultures at 72h post initiation of Th17 polarization. The data is from five individual donors. Error bars show the standard deviation. P-values were calculated using the two-tailed Student's t-test. (E) The expression of IFN-γ was measured by intracellular staining using flow cytometry and (F) the secretion of IFN-γ with Luminex analysis from three biological replicates. (G) A schematic representation of the study. Error bars represent standard deviation across biological replicates (n = 3-5; P <0.001(**) for Th0 vs Th17; paired t-test analysis).

Supplementary Figure 2: Related to Figure 2.

A



B



**Figure S2.** Quality of the proteomics data. (A) The proportion of missing values in each sample. (B) The similarity and clustering of the samples. Pearson correlation coefficient was used as a similarity measure and the hierarchical clustering with complete linkage was used for the clustering.
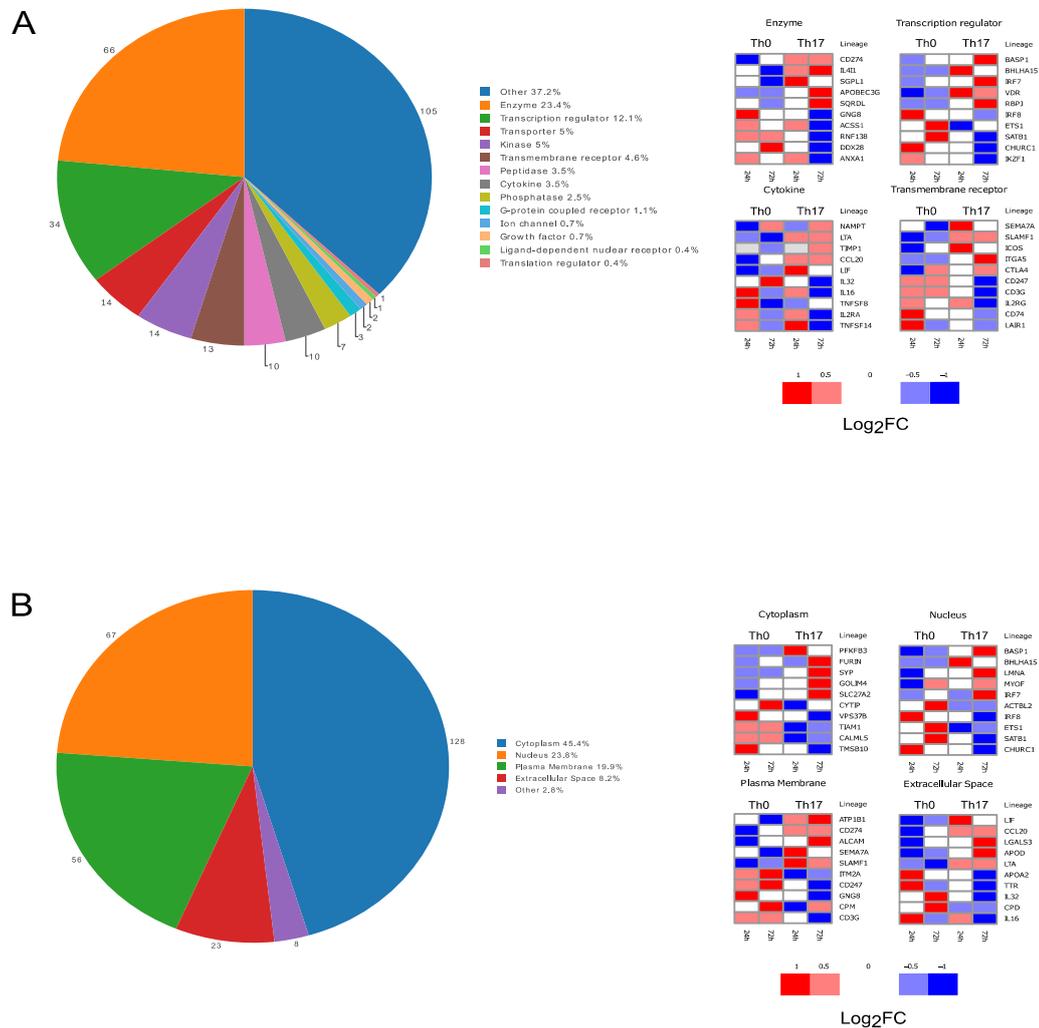
**Figure S3.** The functional types, and cellular locations among the differentially expressed proteins between Th17 and Th0. (A) Proportions and counts of the IPA functional types associated with the differentially expressed proteins over both time-points and the standardized averaged expression of the most up- and down-regulated proteins based on average logarithmic fold changeover both time-points for the selected types. (B) Proportions and counts of the IPA cellular locations associated with the DE proteins over both timepoints and the standardized averaged expression of the most up- and down-regulated proteins based on average logarithmic fold change over both timepoints for the selected locations.
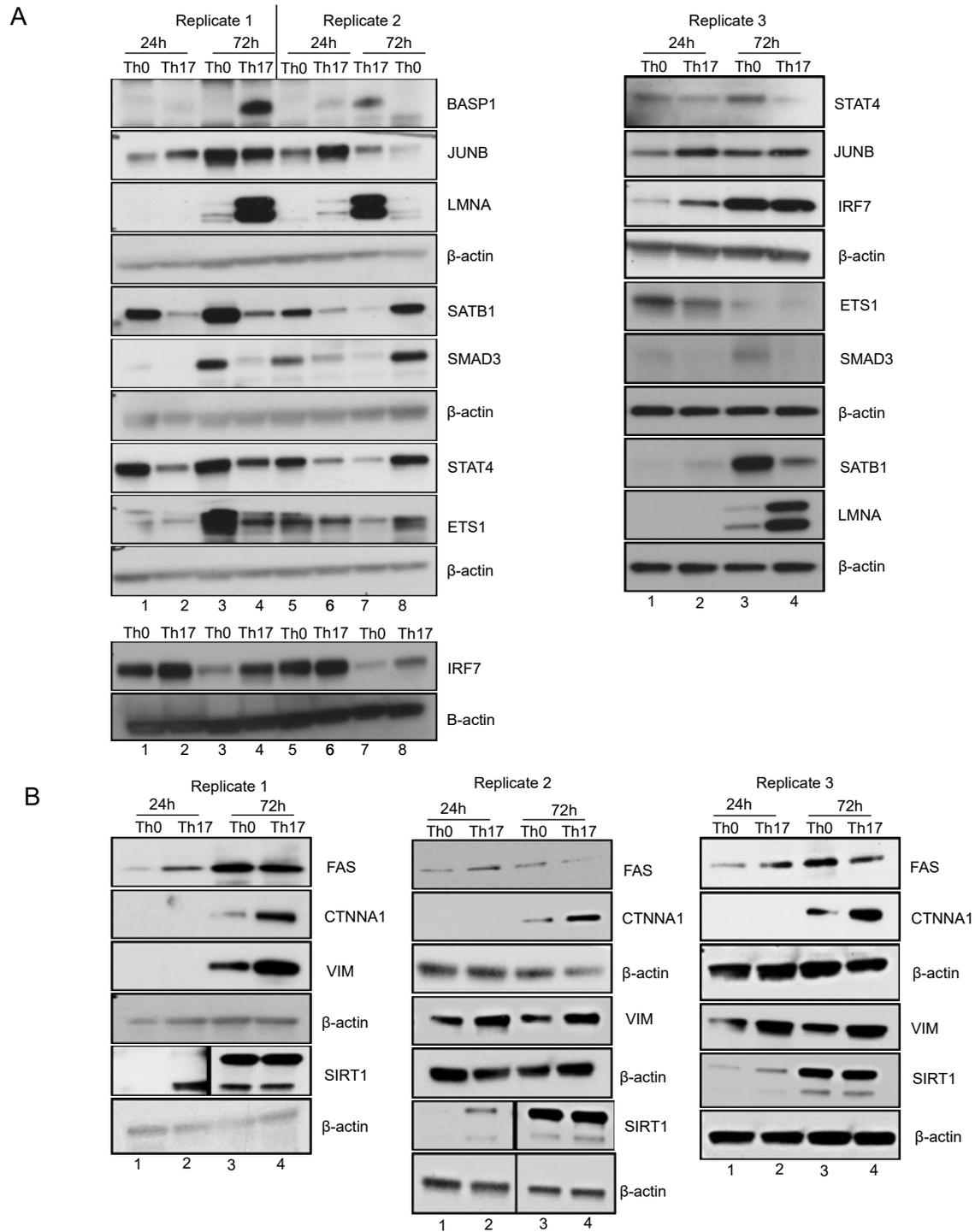
**Figure S4.** Validation of the Mass-spectrometry (MS) identified proteins using Immunoblotting. (A and B), Immunoblot validations of MS identified up- or downregulated proteins with known and unknown Th17-related function. Blots show the protein extracts from the Th0 and Th17 cells at 24h and 72h.

## Supplementary Figure 5 (related to Figure 6 and 7)

A   Related to figure 6

RORyt and Batf proteins expression in Mouse Th17 cells
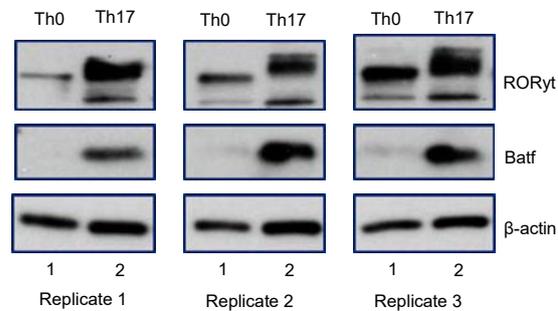


B   Related to figure 6
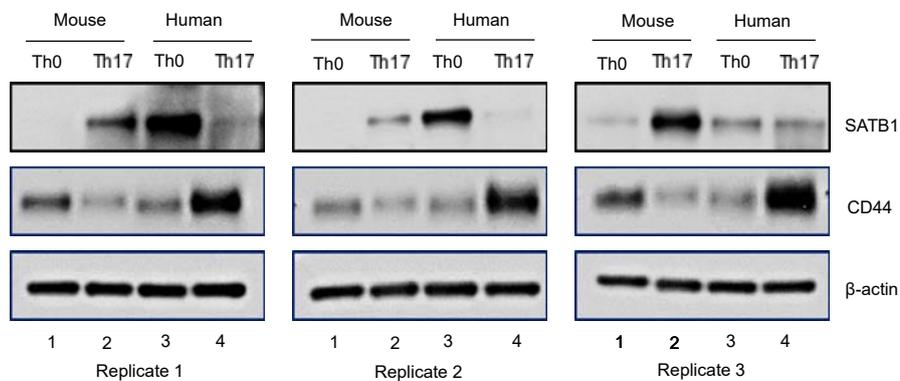


C   Related to figure 7
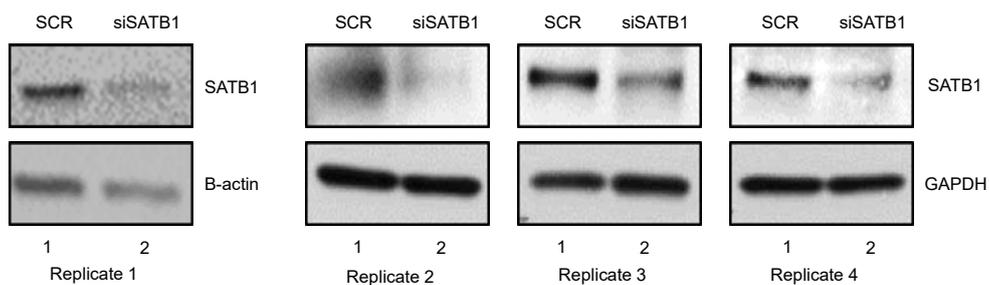


**Figure S5.** Validation of opposite expression of SATB1 and CD44 in human and mouse during Th17 polarization and SATB1 expression upon SATB1 RNAi silencing. (A) Immunoblot analysis of Th17 marker transcription factors, RORyt and BATF expression during mouse Th17 cell polarization, (B) Immunoblot analysis of SATB1 and CD44 expression in human and mouse from the Th0 and Th17 cells at 72h, and (C) Immunoblot analysis of SATB1 expression upon SATB1 knockdown from four biological replicates.

# Inventory of Supplementary Table Legends

All the supplementary tables are provided as separate Excel spreadsheet and can be found in online version.

**Table S1.** The normalized data, missing value proportions of the samples, Pearson correlations between the samples and the differential expression results from all the comparisons. (related to Figure1 and Supplementary Figure 2)

**Table S2.** The up- and downregulated proteins between all cell types (Thp, Th0 and Th17) and the z-score standardized expression values of the differentially expressed proteins between Th17 and Th0. Supplementary Table information related to Figure 1.

**Table S3.** The GO biological process enrichment information relating to Figure 2, and the IPA type and location information relating to Supplementary Figure 3.

**Table S4:** The nodes, edges, clusters and enrichment information of the protein-protein interaction networks relating to Figure 4.

**Table S5:** The common and unique genes and proteins between the proteomics and the transcriptomics datasets at 24h, 72h, and overall and the averaged z-score standardized expression values for the common genes in proteomics and transcriptomics. Supplementary Table information related to Figure 5.

**Supplementary Table 6:** Supplementary Table information related to Figure 6.

**Supplementary Table 7:** Supplementary table information for antibodies, quantitative PCR primer and probes and SATB1 siRNA sequence.

## Transparent Methods:

### Primary human CD4+ T-cell isolation and Th17 cell culture

To obtain naive CD4+ T cells, human peripheral blood mononuclear cells (PBMCs) were isolated from the umbilical cord blood of healthy neonates (Turku University Central Hospital, Turku, Finland) by by the Ficoll- Paque density gradient method (Ficoll-Paque PLUS; GE Healthcare). Naive CD4+ cells were further purified using CD4+ Dynal positive selection beads (Dynal CD4 Positive Isolation Kit; Invitrogen), and after the isolation, cells from individual donors were either activated directly or pooled (in case of validation experiments) before activation. CD4+ T cells were stimulated with plate-bound αCD3 (3750 ng/6-well culture plate well; Immunotech) and soluble αCD28 (1 μg/mL; Immunotech) in a density of $2.5 \times 10^6$ cells/mL of X-vivo 20 serum-free medium (Lonza). The X-vivo 20 medium was supplemented with L-glutamine (2 mM, Sigma-Aldrich), and antibiotics (50 U/mL penicillin and 50 μg/mL streptomycin; Sigma-Aldrich). Th17 cell polarization was initiated with a cytokine cocktail of IL-6 (20 ng/mL; Roche), IL-1β (10 ng/mL) and TGF-β (10 ng/mL) in the presence of neutralizing anti-IFNγ (1 μg/mL) and anti-IL-4 (1 μg/mL) to block Th1 and Th2 differentiation, respectively. For the control cells (Th0), CD4+ T cells were TCR stimulated with αCD3 and αCD28 in the presence of neutralizing antibodies without differentiating cytokines and cultured in parallel. All cytokines and neutralizing antibodies used in the study were purchased from R&D Systems unless otherwise stated. All cell cultures were maintained at 37°C in a humidified atmosphere of 5% (v/v) $CO_2$/air.

### Isolation of mouse cells and in vitro cell culture

BALB/c mice were purchased from the University of Turku animal facility. Animals were handled and housed in accordance with the University of Turku animal welfare guidelines. Cells were obtained from spleens of 8- to 10-week-old mice. Spleens were first macerated

using a cell strainer and syringe plunger to make a single cell suspension, red blood cells were removed using ACK lysis buffer (Gibco by life technology, cat# A10492-01). Cells were then isolated by positive selection using monoclonal antibodies to CD4+CD62L+ coupled with magnetic beads (MACS Miltenyi Biotec; cat# 130-106-643) using a MACS preparation MACS LS/MS column (MACS Miltenyi Biotec).

Cell cultures were performed in IMDM (Gibco) media supplemented with 5% fetal calf serum, 2 mM L-glutamine (Sigma-Aldrich), and 100 U/mL penicillin and 100 µg/mL streptomycin (Sigma-Aldrich) and 50 µM β-mercaptoethanol (Gibco). Cells were activated by plate-bound α-CD3 (1 µg/mL;  BD PharMingen,  cat# 553238) and soluble α-CD28 (2 µg/mL; BD PharMingen, cat# 557393) for 3 days (unless otherwise indicated) and cultured either under neutral conditions (Th0, TCR control) or Th17 differentiation conditions, which were induced by culturing the cells in the presence of TGFβ (1 ng/ml; R&D, cat# 240-B), IL6 (20 ng/ ml; R&D, cat# 406-ML), and IL-1β (10 ng/ml; R&D, cat# 201-LB). Neutralising antibodies anti-IFN-y (cat# 557530), and anti-IL-4(cat# 559062) (both at 10 µg/mL, BD PharMingen) were added to both neutral and Th17 differentiation conditions.


**Mass-spectrometry sample preparation, preprocessing and preliminary analysis**

*Cell lysis*: Proteins were extracted from the cell pellet using a lysis buffer (4% SDS, 0.1 M DTT, 0.1 M Tris-HCl, pH 7.6), heated at 95ºC for 5 min. The lysate was then sonicated at high voltage with a setting of 5 cycles for 30 seconds and 30 seconds rest between cycles. The cell debris were cleared by centrifugation at 16000g for 20 min, and a DC Protein Assay (#5000116, BioRad) was used to estimate protein amounts.

*Filter Aided Sample Preparation (FASP) Method*: Briefly, an aliquot corresponding to 50 µg of protein from each biological replicate corresponding to different time points (i.e., 24h and 72h) (n=5 for both Th17 and corresponding Th0 controls) were mixed with FASP urea buffer

(8 M urea in 0.1 M Tris-HCl, pH 8.5) in a 30-kDa filter tube (Millipore) to eliminate the SDS. The proteins were reduced with dithiothretiol (DTT) and alkylated with iodoacetamide in dark for 20 min. Finally, they were digested with sequencing grade modified trypsin in 1:30 (protein:protease) ratio overnight at 37°C. The digested peptides were then acidified and desalted using Sep-Pak C18 cartridges (WAT054955 Vac 1 cc 50 mg, Waters). The desalted samples were dried using a centrifugal evaporator (Thermo Scientific) and stored at -80°C until further LC-MS/MS analysis.

*Mass spectrometry analysis*: The dried peptides were reconstituted in formic acid/acetonitrile mixture, and an amount corresponding to 400 ng was analysed using Easy-nLC 1200 coupled to Q Exactive HF mass spectrometer (Thermo Scientific). The peptides were separated on 75 µm ID X 40-cm HPLC column, packed in-house with 1.9 µm Reprosil C18 particle (Dr Maisch GmbH). The peptides were eluted with a gradient from 7 to 25% B phase in 75 min then to 90% B in 15 min, at flow rate of 300 nL/min. The mobile phase compositions were, water with 0.1 % formic acid (A) and 80% acetonitrile 0.1% formic acid (B). The temperature of the column was maintained at 60°C using a column oven. The tandem mass spectra were acquired with higher-energy C-trap dissociation (HCD) of the 10 most intense ions (m/z 300–2000, charge states > 1+). The MS1 resolution was set to 120,000, with $3 \times 10^6$ AGC target value and a maximal injection time of 100 ms. MS/MS spectra were acquired in the Orbitrap with a resolution of 15,000 (at m/z 400), a target value of 50,000 ions, a maximum injection time of 250 ms. Dynamic exclusion was set to 30 s. Triplicate analysis were performed for all samples in randomized batches.

**Peptide and protein identification and quantification**

The mass spectrometry raw files were processed using MaxQuant software version 1.5.5.1 (Cox and Mann, 2008). Uniprot human database (May 2017) was used to search the peptide

data using Andromeda (Cox et al., 2011) as a search algorithm. The search parameters specified trypsin digestion with a maximum of two missed cleavages, carbamidomethylation of cysteine as fixed term modification and N-terminal acetylation and methionine oxidation as variable modifications. The peptide and protein level false discovery rates (FDR) were set to 0.01. The match between the runs option was enabled to transform the identifications across the mass spectrometric measurements. The label free quantification method (MaxLFQ) was used to determine the relative intensity values of proteins and to normalize the protein intensities between the samples (Cox et al., 2014). Prior to the downstream data-analysis, data was filtered to remove proteins with less than two unique peptides. Contaminants and reverse hits were also removed (Table S1). The proteomic mass spectrometry data presented in this paper were submitted to PRIDE (Vizcaíno et al., 2016) and have the accession number PXD008973.

**Proteomics data analysis**

All data analyses were performed using the R statistical programming software environment version 3.4.3 (R Core Team, 2015).

**Exploratory data analysis**

To explore the similarity of the samples and the grouping of the biological replicates in the LFQ-normalized data, the R-package pheatmap (Kolde, 2015) was used (Supplementary Figure 2). Pearson correlation coefficient was used as a similarity measure and hierarchical clustering with complete linkage for clustering the samples.

**Differential expression analysis**

The Reproducibility Optimized Test Statistic (ROTS) (Elo et al., 2008; Suomi et al., 2017) was used to detect the DE proteins between the conditions. Differential expression was examined separately for each comparison and time point. The examined comparisons were Th0 – Thp at 24h, Th0 – Thp at 72h, Th17 – Thp at 24h, Th17 – Thp at 72h, Th17 – Th0 at 24h and Th17 – Th0 at 72h. Technical replicates for a biological replicate were averaged and the data was log2-transformed prior to the differential expression analysis. FDR of 0.05 was used as a threshold to define the DE proteins. Differentially expressed proteins whose logarithmic fold change (LogFC) was > 0, were considered as up-regulated and proteins whose LogFC was < 0 were considered as down-regulated. Z-score standardization of the DE proteins in the compared samples was used for visualizing the changes in expression with heatmaps.

**Enrichment analysis**

The enriched GO biological processes were identified using the Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8 (Huang et al., 2009a, 2009b). The GO FAT terms, which filter out the broadest terms, were considered. The enrichment analysis was performed using the DE proteins over both time points as the input and the whole detected and filtered proteome as the background reference. A biological process was considered enriched if it had FDR ≤ 0.05.

The enrichment of molecular types and cellular locations was further examined using the Ingenuity Pathway Analysis (IPA) (QIAGEN Inc.) (Krämer et al., 2014). The enrichment of cellular locations and protein types between Th17 and Th0 was examined at 24h and 72h using the time point specific DE proteins as input for IPA and the whole detected and filtered proteome as a background reference. All the resulting IPA location and type information was collected.

**Protein-protein interactions**

To investigate protein-protein interactions, the DE proteins over both time points (24h and 72h) were entered into the STRING functional protein association networks database (Szklarczyk et al., 2017). Both predicted and known high confidence (interaction score ≥ 0.7) interactions were considered. The resulting interaction network was imported into Cytoscape version 3.6.0 (Shannon et al., 2003) for visualization. Clusters in the interaction data were identified using the Cytoscape plug-in clusterMaker v2 (Morris et al., 2011) with the granularity parameter (inflation value) set to 1.8 as suggested by (Brohée and van Helden, 2006). The GO biological process enrichment results over both time points were used to calculate the most frequent terms for each cluster, from which the most representative biological processes for each cluster were selected. The proteins in the main cluster were grouped according to the representative processes.

To further examine functionality in the identified network, enriched biological pathways in the main cluster (cluster 1) was examined using the Cytoscape plug-in ReactomeFIViz version 7.0.1 (Wu et al., 2014). The ReactomeFIViz-plugin was used to discover subclusters in cluster 1 and to identify the enriched biological pathways in the discovered subclusters. ReactomeFIVIz examines enriched pathways in a gene/protein set against multiple pathway-databases (Wu et al., 2014)]. An FDR of 0.01 was used as a threshold for biological pathways in the subclusters to discover only the most important functionalities.

**Targeted Proteomics Validation**

Selected reaction monitoring (SRM) mass spectrometry was used to validate the relative abundance of ATP1B1, PALLD, ACSL4, FHOD1, SMTN and RDX in the Th17 cells at 72 h. Heavy-labeled synthetic peptides (lysine $^{13}C_6$ $^{15}N_2$ and arginine $^{13}C_6$ $^{15}N_4$) peptides were obtained for the targets of interest (Thermo Fischer Scientific) and were selected on the

basis of their stability, consistency and intensity in the discovery data.  For these validations, four additional cultures were prepared from the cord blood of four donors.

Skyline software (MacLean et al., 2010) was used to evaluate the top five most intense transitions from the MS/MS spectra of the heavy labelled standard peptides and assess the relative performance of the native peptides in the spiked validation samples.

The samples were prepared using the same FASP digestion and desalting protocols used for discovery. These were then spiked with synthetic heavy labelled analogues of the peptide targets and a retention time standard (MSRT1, Sigma) for scheduled selected reaction monitoring.  The LC-MS/MS analyses were conducted using Easy-nLC 1000 liquid chromatograph (Thermo Scientific) coupled to a TSQ Vantage Triple Quadrupole Mass Spectrometer (Thermo Scientific).  The column configuration included a 20 x 0.1 mm i.d. pre-column in conjunction with a 150 mm x 75 μm i.d. analytical column, both packed with 5 μm Reprosil $C_{18}$-bonded silica (Dr Maisch GmbH). A separation gradient from 8% to 43% B in 27 min, then to 100% B in 3 min, was used at a flow rate of 300 nl/min (the mobile phase compositions are as indicate above). The raw SRM data are available through PASSEL (Farrah et al., 2012) with the dataset identifier PASS01204

Skyline was used to select the transition used for the assays subsequently process the data generated. The MSStats (3.8.4) plugin included in the Skyline software was used for the group comparison between cases and controls. The summed intensities of GAPDH peptides were used as a global standard to normalize the data from each analysis and Tukey's median polish method was used as the summary method.

**Transcriptomics data analysis**

**RNAseq sample preparation**

RNA samples from five biological replicates derived from Th0 and Th17 cultures of five individual donors were collected at 72h time point. RNA was isolated (RNeasy Mini Kit,

QIAGEN) and DNase treated (RNase-Free Dnase Set; QIAGEN). Library preparation was performed according to Illumina TruSeq® Stranded mRNA Sample Preparation Guide (part # 15031047). RNA-seq with 50 nt read length was performed at the Finnish Functional Genomics Centre (FFGC) with HiSeq 3000 instrument using TruSeq chemistry and the raw data was base called with CASAVA1.8.

**Preprocessing of the raw data**

The RNA-seq raw reads were mapped to the Ensembl human reference genome GRCh38 (Genome Reference Consortium Human Build 38) (Zerbino et al., 2018) using the STAR aligner (Dobin et al., 2013) version 2.5.2. The read counts were generated using the featureCounts tool (Liao et al., 2014) in the Subread software package (Liao et al., 2013) version 1.5.1. Uniquely mapped reads were used for further analysis. The uniquely mapped reads were filtered for lowly expressed genes (genes with counts per million (cpm) >1 in at least 5 replicate samples were retained) and used for further analysis. The filtered gene counts were normalized using the trimmed mean of M-values (TMM) normalization from the Bioconductor package edgeR (McCarthy et al., 2012; Robinson et al., 2010) after which the data was transformed to counts per million (cpm), offsetted by 1 and log2-transformed. Differential expression analysis was performed similarly to proteomics using ROTS (Elo et al., 2008; Suomi et al., 2017). An FDR of 0.05 was used as a threshold to define the DE genes. The RNAseq data presented in this study was submitted to the Gene Expression Omnibus (GEO) and has the Series record GSE118974.

**Comparison between human and mouse proteomics data**

To compare differentially regulated proteins (differentially expressed proteins and proteins detected in only one condition) between Th17 and Th0 in human and mouse at 72h during Th17 polarization, we used the published mouse proteomics raw data (Mohammad et al., 2018)  and pre-processed it similarly to the human data using MaxQuant (version 1.5.5.1)

with LFQ-normalization, filtering out proteins with less than two unique peptides, and removing contaminants and reverse hits. Similarly, as with the human proteomics data, the differential expression analysis was performed using ROTS. To make the comparison of the differentially regulated proteins more comprehensive, we used a threshold of FDR 0.1 in both datasets to define the DE proteins. Mouse genes related to proteins were mapped to orthologous human genes for the comparison using Ensembl BioMart (Zerbino et al., 2018). All the orthologous mouse genes from Ensembl 92 database to the human reference genome GRCh38 were considered. If multiple orthologous human genes existed for a given mouse gene, the most similar human orthologous gene according to the Ensembl database was selected.

**Immunoblot Analysis**

For the western blot analysis, cell samples were lysed in either RIPA (Pierce, #89901) or Triton-X buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.5% Triton-X-100; 5% glycerol; 1% SDS), supplemented with proteinase (Roche) and phosphate inhibitors (Roche) on ice with vortex every 5 min. Lysed samples were sonicated for 5 min under ice cold conditions (Bioruptor UCD-200; Diagenode), followed by centrifugation at maximum speed for 20 min at 4°C. Supernatants containing proteins were collected and quantified (DC Protein Assay; Bio-Rad). The samples containing 15–30 $\mu$g of protein was boiled with 6x sample loading dye (330 mM Tris-HCl, pH 6.8; 330 mM SDS; 6% $\beta$-ME; 170 mM bromophenol blue; 30% glycerol). Samples were loaded on 4–20% gradient SDS-PAGE gels (Biorad), transferred to nitrocellulose membranes (Bio-Rad), and probed with the antibodies listed in Supplemental Table 7. During the course of this study, immuno-blotting membranes were either cut into sections and each section was probed with separate antibodies, or the blots were striped with striping buffer (25mM Glycine and 1%SDS; pH 2.5) and re-probed successively with different antibodies recognizing proteins with different molecular mass.

**IL-17A secretion**

IL-17A production is a characteristic of CD4+ T cells differentiated toward the Th17 phenotype and can be detected in cell-culture supernatant at 72h using either the Milliplex MAP human IL-17A kit (Merck Millipore; HCYTOMAG-60K-01), Bioplex Human IL-17A Cytokine/Chemokine 96-Well Plate Assay (Bio Rad; Cat. no. 171B5014M, 171304090M) or Human IL-17A Duoset ELISA kit (R&D Biosystems DY317-05, DY008). The amount of IL-17A secreted by Th17 cells was normalized with the number of living cells determined based on forward and side scattering in flow cytometric analysis (LSRII flow cytometer; BD Biosciences).

**IFN-γ secretion**

Culture supernatants from CD4+ T cells differentiated toward the Th17 phenotype at 72h were harvested and assayed by ELISA for IFN-γ secretion (Milliplex MAP human IL-17A+IFN-γ kit; Cat.no. HCYTOMAG-60K-02), according to the manufacturer's protocols. For appropriate analysis, all values below the detectable range were considered zero.

**Flow cytometry**

Flow cytometry analysis of cell-surface receptor CCR6 detection was done at 72h after initiation of Th17 cell priming. In short, cells were washed twice with PBS, and staining was done in staining buffer (0.5% FBS/0.1% Na-azide/PBS) for 15 min at 4°C and data is either acquired in flow cytometer on same day in staining buffer or cells were fixed with 1% formalin. LSRII flow cytometer (BD Biosciences) was used in data acquisition. Living cells were gated for analysis based on forward and side scattering. Detection of IFN-γ producing cells was determined by intracellular cytokine staining with anti-IFN-γ-FITC (BD Biosciences). Briefly, cells were stimulated with phorbol 12-myristate 13-acetate and ionomycin (PMA) for 5 h. GolgiStop (BD Biosciences) was added after 2 h, and cells were fixed in 4% paraformaldehyde solution. Fixed cells were stained with fluorescent antibodies

in 0.1% saponin permeabilization buffer and analysed on a LSRII (BD Biosciences). For OASL and ATF3 staining, cells were first fixed with 4% paraformaldehyde and permeabilized with Perm III buffer (BD Biosciences) before analysis. Cells were incubated with primary antibodies, followed by 30 min incubations with the secondary antibody when needed. The staining was controlled by using isotype specific antibodies or only with the secondary antibodies. The information regarding the antibodies used for staining is listed in Supplemental Table 7.

**Quantitative Real-time PCR**

Total RNA was isolated using RNeasy kit (Qiagen, Valencia, CA). cDNA was synthesized with Reverse Transcription kit (Applied Biosystems, Foster City, CA) using oligo dT primers according to the manufacturer's instruction. TaqMan primers and probes for IL-17A, IL-17F, RORC and SATB1 were designed with Universal ProbeLibrary Assay Design Centre (Roche), in Absolute QPCR ROX Mix (Thermo Scientific). EF1a gene was used as endogenous control. The qPCR runs were analysed using the 7900HT Fast Real-Time PCR System (Applied Biosystems).

**Cell transfection with small interfering RNA (siRNA)**

Freshly isolated CD4+ T cells from umbilical cord blood were suspended in Optimem I (Invitrogen) and transfected with SATB1 targeting siRNA oligonucleotide (Sigma) (Table S7) using the nucleofection technique (Lonza). Four million cells were transfected with 5ug of siRNA after which the cells were rested at $37^0$ C for 24h in RPMI 1640 medium (Sigma-Aldrich) supplemented with pen/strep, 2 mM L-glutamine and 10% FCS ($2x10^6$ cells/ml) and subsequently activated and cultured under Th17 culturing condition as described above.

**Statistical analysis**

A two-tailed student's t-test was used for determining the statistical significance of IL17A and IFN-y secretion and percentage of CCR6 and IFN-y expressing cells at 72h of culture

from three to five independent cultures. Statistical analysis of the mass spectrometry data is described in the respective method.

## Supplemental References

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotech *26*, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V, and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Cox, J., Hein, M.Y., Luber, C. a, and Paron, I. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. … *13*, 2513–25

Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. *44*, D447–D456.

R Core Team (2015). R: A Language and Environment for Statistical Computing. R Found. Stat. Comput. Vienna Austria.

Kolde, R. (2015). pheatmap : Pretty Heatmaps. R Packag. Version 1.0.8.

Elo, L.L., Filén, S., Lahesmaa, R., and Aittokallio, T. (2008). Reproducibility-optimized test statistic for ranking genes in microarray studies. In IEEE/ACM Transactions on Computational Biology and Bioinformatics, p.

Suomi, T., Seyednasrollah, F., Jaakkola, M.K., Faux, T., and Elo, L.L. (2017). ROTS: An R package for reproducibility-optimized statistical testing. PLoS Comput. Biol. *13*.

Huang, D.W., Sherman, B.T., Lempicki, R.A., Huang, D., Sherman, B., Lempicki, R., Huang, D., Sherman, B., Lempicki, R., Maere, S., et al. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc.

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity pathway analysis. Bioinformatics.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. *45*, D362–D368.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res.

Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D., and Ferrin, T.E. (2011). ClusterMaker: A multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics.

Brohée, S., and van Helden, J. (2006). Evaluation of clustering algorithms for protein-protein interaction networks. BMC Bioinformatics.

Wu, G., Dawson, E., Duong, A., Haw, R., and Stein, L. (2014). ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. F1000Research.

MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: An open source document editor for creating and analyzing targeted proteomics experiments.

Bioinformatics

Farrah, T., Deutsch, E.W., Kreisberg, R., Sun, Z., Campbell, D.S., Mendoza, L., Kusebauch, U., Brusniak, M.Y., Hüttenhain, R., Schiess, R., et al. (2012). PASSEL: The PeptideAtlas SRMexperiment library. Proteomics.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., et al. (2018). Ensembl 2018. Nucleic Acids Res

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res.

Liao, Y., Smyth, G.K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics.

Mohammad, I., Nousiainen, K., Bhosale, S.D., Starskaia, I., Moulder, R., Rokka, A., Cheng, F., Mohanasundaram, P., Eriksson, J.E., Goodlett, D.R., et al. (2018). Quantitative proteomic characterization and comparison of T helper 17 and induced regulatory T cells. PLOS Biol.