Cell

A Validated Regulatory Network for Th17 Cell Specification

Maria Ciofani,[1,10] Aviv Madar,[3,4,10] Carolina Galan,[1] MacLean Sellars,[1] Kieran Mace,[3] Florencia Pauli,[5] Ashish Agarwal,[3] Wendy Huang,[1] Christopher N. Parkurst,[1] Michael Muratet,[5] Kim M. Newberry,[5] Sarah Meadows,[5] Alex Greenfield,[2] Yi Yang,[1] Preti Jain,[5] Francis K. Kirigin,[2] Carmen Birchmeier,[6] Erwin F. Wagner,[7] Kenneth M. Murphy,[8,9] Richard M. Myers,[5] Richard Bonneau,[3,4,*] and Dan R. Littman[1,9,*]

[1]Molecular Pathogenesis Program, The Kimmel Center for Biology and Medicine of the Skirball Institute
[2]Computational Biology Program, The Sackler Institute
New York University School of Medicine, New York, NY 10016, USA
[3]Department of Biology, Center for Genomics and Systems Biology
[4]Computer Science Department, Courant Institute of Mathematical Sciences
New York University, New York, NY, 10003 USA
[5]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA
[6]Developmental Biology, Max Delbruck for Molecular Medicine, 13125 Berlin, Germany
[7]Cancer Cell Biology Programme, Spanish National Cancer Research Centre (CNIO), E-28029 Madrid, Spain
[8]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63108, USA
[9]The Howard Hughes Medical Institute
[10]These authors contributed equally to this work
*Correspondence: bonneau@nyu.edu (R.B.), dan.littman@med.nyu.edu (D.R.L.)
http://dx.doi.org/10.1016/j.cell.2012.09.016

SUMMARY

Th17 cells have critical roles in mucosal defense and are major contributors to inflammatory disease. Their differentiation requires the nuclear hormone receptor RORγt working with multiple other essential transcription factors (TFs). We have used an iterative systems approach, combining genome-wide TF occupancy, expression profiling of TF mutants, and expression time series to delineate the Th17 global transcriptional regulatory network. We find that cooperatively bound BATF and IRF4 contribute to initial chromatin accessibility and, with STAT3, initiate a transcriptional program that is then globally tuned by the lineage-specifying TF RORγt, which plays a focal deterministic role at key loci. Integration of multiple data sets allowed inference of an accurate predictive model that we computationally and experimentally validated, identifying multiple new Th17 regulators, including Fosl2, a key determinant of cellular plasticity. This interconnected network can be used to investigate new therapeutic approaches to manipulate Th17 functions in the setting of inflammatory disease.

INTRODUCTION

The vertebrate immune system, composed of numerous phenotypically well-defined cell types, is ideally suited for studying the combinatorial action of transcription factors (TFs) and epigenetic regulators whose target gene products confer unique cellular functions. TFs that are selectively expressed in subsets of myeloid and lymphoid lineage cells have been designated "master regulators" if they are both essential and sufficient to induce defined cell fates. It is becoming clear, however, that networks of multiple TFs are required to achieve the full differentiation programs (Mattick et al., 2010; Novershtern et al., 2011). How such factors cooperate to determine specific programs remains poorly understood.

CD4-expressing T lymphocytes are among the best-characterized immune system cells (Zhu et al., 2010). They develop in the thymus and acquire the potential to become T helper cells that guide B lymphocytes to produce distinct classes of antibody and carry out multiple other effector functions; or they upregulate the TF Foxp3 and become anti-inflammatory regulatory T cells (Treg). T helper cells differentiate further in the periphery following induction or activation of TFs in response to signals from the T cell antigen receptor (TCR), cytokines, and other ligands in the microenvironment. T helper effector subsets include Th1 cells, which produce interferon-γ and control infections with intracellular microbes; Th2 cells, which secrete IL-4, IL-5, and IL-13 and are required for clearance of helminthes; and Th17 cells, producers of IL-17A, IL-17F, and IL-22, which protect mucosa from bacterial and fungal infection (Korn et al., 2009). In addition, follicular helper T cells (T_{FH}) provide B cells with signals for immunoglobulin class switching and affinity maturation (Crotty, 2011). $CD4^+$ T cell subsets exhibit plasticity but are considered distinct lineages based on expression of TFs with properties of "master regulators." Th1 cells are defined by their expression of T-bet (Tbx21), Th2 cells by GATA3, Th17 cells by RORγt, and T_{FH} cells by Bcl6. Effector T cells expressing distinct subset-specific cytokines are common in vivo, although cells with combinations of such cytokines are often observed. Differentiation of naive $CD4^+$ T cells into Th1, Th2, Th17, or

Treg cells can be mimicked in vitro by TCR stimulation and combinations of defined cytokines. Genome-wide histone modifications, chromatin accessibility, and occupancy by lineage-specifying TFs have been studied in such models (Durant et al., 2010; Kwon et al., 2009; Wei et al., 2009).

Th17 cells have critical functions in many autoimmune diseases and in cancer (Korn et al., 2009). The orphan nuclear receptor RORγt is required for the differentiation of Th17 cells and for inflammatory diseases in mice. Its forced expression in mouse and human T cells induces transcripts present in Th17 cells, including those coding for the key cytokines, for the IL-23 receptor, and for the chemokine receptor CCR6 (Ivanov et al., 2006; Manel et al., 2008). However, RORγt is not sufficient to specify the full Th17 program, and other TFs, including STAT3, IRF4, BATF, and IκBζ, are required for induction of RORγt and IL-17A in vivo and upon polarization in vitro by IL-6 and TGF-β, with or without IL-1β and IL-23 (Brüstle et al., 2007; Okamoto et al., 2010; Schraml et al., 2009; Yang et al., 2007). Multiple other TFs are also involved in Th17 cell differentiation, including c-Maf, Runx1, and Ahr (Bauquet et al., 2009; Veldhoen et al., 2008; Zhang et al., 2008). RORα, which is closely related to RORγt, can also contribute to IL-17 expression in the absence of RORγt (Yang et al., 2008).

Investigation of TF functions in Th17 cell differentiation has been limited to how single factors affect expression of a limited number of targets (e.g., IL-17A). However, the Th17 differentiation program extends beyond functions of individual cytokines, as highlighted by studies showing Th17-mediated pathogenesis in the absence of IL-17A and IL-17F (Codarri et al., 2011; Leppkes et al., 2009). We therefore wished to examine how multiple TFs regulate each other and their targets in order to model a transcriptional network and identify critical factors in Th17 cell differentiation. New regulatory interactions can be found using genome-wide methods to learn networks from time series and genetic and environmental perturbations (Bonneau et al., 2007; Ernst et al., 2007; Faith et al., 2007; Greenfield et al., 2010). Integrating multiple data types and analytical methods for network inference using meta-analyses (averaging over several data types and computational approaches) can leverage the complementary weaknesses and strengths of each component to produce higher accuracy networks (Marbach et al., 2012b). We have applied an integrative approach, with meta-analysis of genome occupancy of multiple TFs, RNA-seq of TF-deficient T cells, and immune cell transcriptome data, to build a network model for Th17 cells. We find in Th17 cell differentiation early cooperative binding of BATF and IRF4 that governs chromatin accessibility and subsequent recruitment of RORγt to regulate a select set of Th17-relevant genes. We used the network model to identify additional candidate genes that were in turn incorporated through an iterative process and validated several genes critical in Th17 cell differentiation, including TFs that influence the expression of > 2,000 genes. We found that the AP-1 family member Fosl2 (confidently predicted to be a core Th17 factor) has a key role in a mouse model of autoimmune disease, limiting plasticity of T helper cell differentiation. In addition, loci that are implicated in genome-wide association studies (GWAS) to have roles in autoimmune disease were enriched in the Th17 network. This analysis can therefore identify candidate genes that serve as cogs in functional specialization of Th17 cells and that have the potential to be valuable for new therapeutic approaches.

RESULTS

TF Co-Occupancy Enriches for Functional *cis* Regulatory Modules

We studied early Th17 cell specification events in a largely synchronized population of naive $CD4^+$ T cells that produce IL-17 upon stimulation of the TCR in the presence of IL-6 and low levels of TGF-β (Figure S1A available online). In this model, cells receiving TCR stimulation alone serve as a nonpolarized (Th0) control.

To assemble a high-resolution map of TF-DNA interactions in Th17 differentiation, ChIP-seq was performed with antibodies directed against STAT3, IRF4, BATF, c-Maf, RORγ, and p300 with cells cultured in Th0 and Th17 conditions for 48 hr, when Th17-specifying TFs are simultaneously expressed or active (Figure S1B). Many high-confidence bound regions were observed for each TF in Th17 conditions and for BATF and IRF4 in Th0 cells with the cognate consensus binding motif recovered for each (Figure S1C). TFs binding at key lineage-associated loci (*Il17a*, *Il17f*, *Il12rb1*, *Il1r1*, *Rorc*) revealed a high degree of colocalization in Th17 cells (Figures 1A and S1D), indicating that these TFs occupy common *cis* regulatory regions and highlighting their roles in integrating cytokine and TCR-derived signals. In contrast, insulator-binding factor CTCF displayed a distinct occupancy pattern.

To examine genome-wide TF binding patterns in Th0 and Th17 cells, we merged TF binding peaks with summits that clustered in close proximity to each other (within 100 bp) to define putative *cis* regulatory modules (pCRM; Figure 1B). In this manner, 162,113 significant TF peaks clustered into 83,138 unique pCRMs, with distributions provided in Figure 1C. The clustered heat map in Figure 1B displays TF binding significance for TSS-proximal (± 5 kb) pCRMs and the associated fold change in gene expression in Th17 versus Th0 cells for the nearest gene (Figure 1B, right).

The pCRM clustering revealed that the most prominent signature was co-occupancy by all five TFs in Th17 cells and by IRF4 and BATF in Th0 cells (clusters 1 and 2, Figure 1B). This was not a result of TF binding enrichment expected near the TSSs, as ~70% of the five-TF pCRMs localized to distal sites (>5 kb from TSS; Figure S1E). Of note, BATF and IRF4 showed a striking binding overlap regardless of occupancy by other Th17 TFs (clusters 3–6, Figure 1B). TF average binding significance increased with pCRM order and was the highest at five-TF pCRMs (Figure 1C), possibly reflecting enhanced accessibility and/or cooperativity between factors at these regions. Moreover, strong RORγt binding was almost exclusively restricted to five-factor pCRMs (clusters 1 and 2), suggesting that these elements represent important regulatory domains for the integration of specification signals. Accordingly, these pCMRs were proximal to loci showing differential expression in Th17 versus Th0 cells (Figure 1B, right). Nearly all (>99%) five-TF pCRMs colocalized with the histone acetyltransferase p300

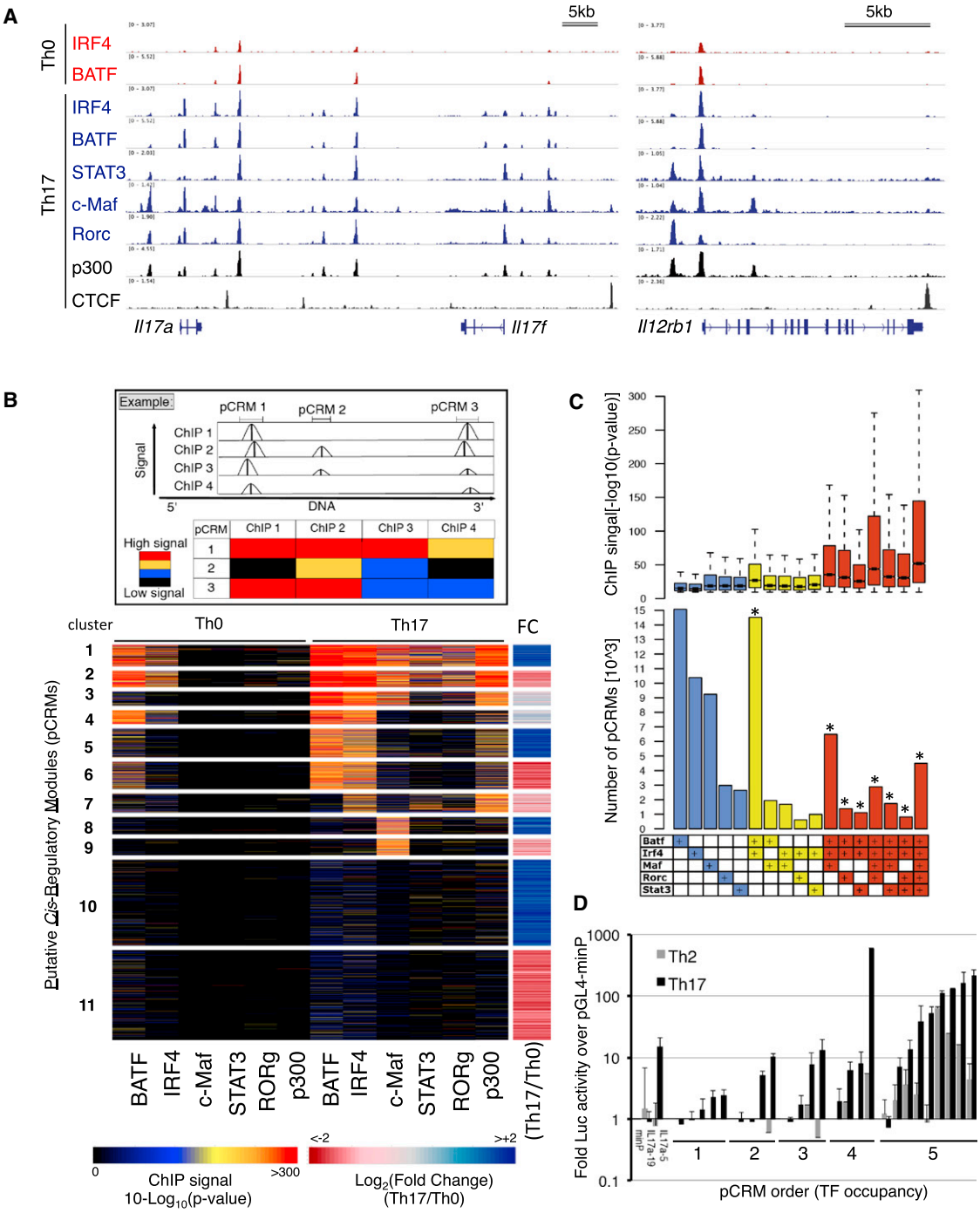


Figure 1. Genome-wide Co-Occupancy of Th17 Lineage TFs

(A) ChIP-seq binding tracks for core TFs, p300, and CTCF at selected Th17 loci in Th0 and Th17 cells at 48 hr (visualized with IGV, Broad Institute).

(B) A clustered heat map of pCRM regions (rows) based on TF ChIP signals and the associated gene expression fold changes (FC) in Th17 versus Th0 cells. A schematic illustration of the clustering approach is shown (top).

(C) Numbers of pCRMs with > 500 occurrences in Th17 cells (bottom) with associated distributions of TF ChIP p values (top). Box plots (top) show median (line), 25th–75th percentile (box) ± 1.5 interquartile range. Asterisk denotes that we observed significantly more pCRMs than expected by chance based on 10,000 simulations ($p < 0.001$).

(D) Luciferase reporter assay of enhancer activity for selected pCRM DNA regions in CD4 T cells cultured under Th2 and Th17 conditions. (y axis) Luciferase activity relative to that of pGL4-minP with a minimal promoter. (x axis) TF occupancy order of tested pCRMs (1 to 5 TFs); Il17a-5 and Il17a-19 serve as positive and negative controls, respectively. Error bars represent SD of two experiments.

See also Figure S1 and Table S1.

(Figure 1B), a factor for which occupancy can be predictive of tissue-specific regulatory activity (Visel et al., 2009). Of note, p300 binding in Th17 cells, though induced relative to Th0, was not restricted to upregulated loci, which may reflect the contribution of inhibitory gene regulation by distal pCRMs.

To relate pCRM occupancy to *cis* regulatory function, genomic regions corresponding to pCRMs occupied by different numbers of TFs were cloned upstream of a minimal promoter driving a luciferase reporter and assayed for activity (Figure 1D). The most active regions coincided with high-order occupancy, which was also significantly more active in Th17 versus control Th2 cells. Thus, *cis* regions that integrate Th17 signals (4 and 5 TF pCRMs) display lineage selectivity, and their combinatorial occupancy increases the likelihood of activity. Together, these findings support the view that core Th17 TFs function synergistically and that shared regulatory targets of STAT3, IRF4, BATF, c-Maf, and RORγt are likely enriched with key targets for Th17 cell specification.

Cooperative Binding of BATF/IRF4 Complexes Prepatterns Chromatin for Specification

The strong association between IRF4 and BATF occupancy in both Th17 and Th0 cells suggested a regulatory interaction between these factors. Indeed, in five-TF pCRMs, the binding summits for IRF4 and BATF were more proximal than for any other pair of Th17 TFs (Figure 2A), and BATF was uniquely coimmunoprecipitated with IRF4 in a DNA-dependent manner in Th17-polarized cells (Figure 2B), indicating that they form a complex on DNA. Consistent with this, pCRMs co-occupied by BATF and IRF4 lacked the interferon-stimulated response element (ISRE) consensus but were enriched with two types of dominant AP1-ISRE composite elements with AP-1 motifs adjacent to ISRE half sites (Figure 2C). As a similar motif structure underlies cooperative binding of IRF4 and PU.1 (Eisenbeis et al., 1995), we tested functional cooperativity of BATF and IRF4 binding in a cellular context. We thus performed ChIP-seq for each TF in Th0- and Th17-polarized cells genetically deficient for the other TF. IRF4 and BATF occupancy was markedly reduced in *Batf* and *Irf4* mutant cells, respectively. The effect was most significant at pCRMs occupied by IRF4 and BATF in combination or with additional TFs when compared to regions harboring single IRF4 or BATF peaks, where cooperativity is not expected (Figure 2D). The dependence between IRF4 and BATF was stronger at distal pCRMs, suggesting that additional factors binding near promoters compensate for the loss of either TF. Similar results were obtained in Th0 cells in the absence of cytokine-induced factors (Figure S2A). Several examples of this mutual binding dependency are shown (Figures 2E and S2B).

The cooperativity of IRF4 and BATF, paired with our finding that *cis* regions bound by both TFs in TCR stimulation conditions (Th0) exclusively acquire additional strong binding by STAT3, RORγt, c-Maf, and p300 in Th17 cells (Figure 1B), suggests that IRF4 and BATF function as pioneer factors in nucleating binding of Th17 TFs upon cytokine-stimulated differentiation. Consistent with this, occupancy of IRF4 and BATF is enriched in the center of five-TF pCRMs (Figure S3A). To investigate this hypothesis directly, we examined chromatin accessibility at regions co-occupied by all five TFs, using formaldehyde-assisted isolation of regulatory elements sequencing (FAIRE-seq). Deletion of IRF4 or BATF in Th0 or Th17 cells had little effect on regions already accessible in naive cells, but most regions with inducible FAIRE-seq signal exhibited marked reductions in *Irf4*$^{-/-}$ and *Batf*$^{-/-}$ compared to WT cells in both Th17 and Th0 conditions (Figure S3B). Thus, in the absence of IRF4 and BATF, regions normally bound by all five Th17 TFs are less accessible, providing further evidence that IRF4 and BATF remodel the chromatin landscape, potentially facilitating subsequent recruitment of TFs involved in regulating expression of adjacent genes.

Consistent with IRF4 and BATF mediating accessibility, these TFs also globally affect p300 occupancy, which was reduced in IRF4- or BATF null Th17 cells (Figure S3C). STAT3 deficiency also reduced p300 binding, but loss of RORγt resulted in a much smaller genome-wide effect (Figure S3C). The focal influence of RORγt was also reflected in occupancy of IRF4 and STAT3 and the presence of H3K4me2 and H3K4me3, histone marks associated with active transcription, in RORγt-deficient Th17 cells (Figure S3D). Strikingly, few genes were dependent on RORγt, as measured by more than 2-fold reduction (and p value < 0.01) in both H3K4me3 at locus-linked pCRMs and expression of the respective genes, namely *Il17a*, *Il17f*, *Il23r*, *Ccl20*, *Il1r1*, *Ltb4r1*, which were known and novel targets (*2310007L2Rik*, *Furin*, *Fam124b*, *Tmem176a*, and *Tmem176b*). These findings are consistent with RORγt having a highly specific regulatory footprint relative to initiator TFs that establish broader changes in chromatin remodeling.

The Th17 Network Reveals Lineage Specification by Combinatorial Regulation

Although it provides mechanistic insight, TF occupancy does not sufficiently explain target gene regulation. We thus complemented TF ChIP-seq (which finds direct targets) with RNA-seq of TF wild-type versus knock-out (KO) Th17 cells (which finds functional targets). ChIP p values for peaks falling within 10 kb of the gene body were consolidated to a gene-wide p value, and KO-RNA-seq p values were calculated based on target gene differential expression. For each data set, p values were mapped to rank-based scores from 0 (least significant) to 1 (most significant) and combined such that the integrated ChIP + KO scores ranged from 0 to 2, with the highest-scoring genes likely to be direct and functional targets. Rank-based approaches have proved successful in similar settings (Madar et al., 2010; Marbach et al., 2012a, 2012b; Prill et al., 2010) (see Extended Experimental Procedures). High-confidence TF-target interactions (score > 1.5; FDR < 10%) are visualized in Cytoscape.

The resulting causal network captures known relationships between core TFs, including *Rorc* activation by STAT3, IRF4, and BATF (Brüstle et al., 2007; Schraml et al., 2009; Yang et al., 2007) and further reveals feed-forward loops that reinforce *Rorc* expression in response to TCR and cytokine signals (Figure 3A, box). Of note, there is high interconnectivity among TFs, including positive feedback loops reinforcing expression of initiator TFs BATF, IRF4, and STAT3, and a negative feedback loop (c-Maf to BATF) that limits response. Conversely, RORγt

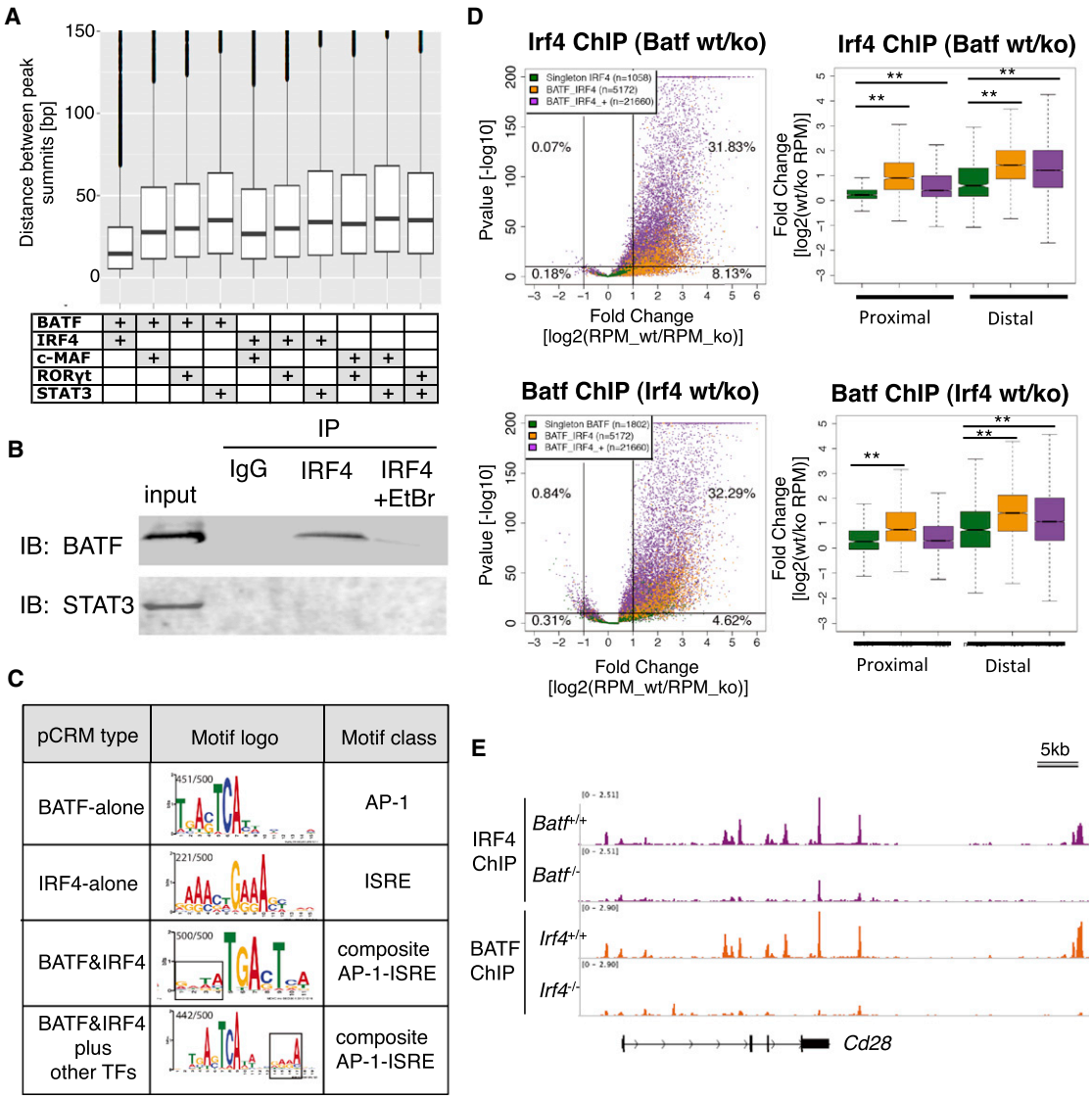


Figure 2. Cooperative Occupancy by BATF and IRF4

(A) Proximal binding of BATF and IRF4. Distribution of the distance between ChIP peak summits for pairs of TFs in 5-TF pCRMs. Box plot is as shown in Figure 1; points represent outliers.

(B) Coimmunoprecipitation of BATF, but not STAT3, with IRF4 in 48 hr Th17-polarized cultures. Ethidium bromide (EtBr) disrupts DNA-protein interactions. IP, immunoprecipitating antibody; IB, immunoblotting antibody. Representative of two experiments.

(C) MEME-ChIP motifs identified in four subtypes of pCRMs as indicated. The AP-1 and ISRE consensus is recovered in regions singly occupied by BATF and IRF4, respectively. A new AP-1-ISRE composite motif comprising an AP-1 site (TGA(C/G)TCA) adjacent to an ISRE half-site (GAAA; boxed region) is only recovered at pCRMs occupied by BATF and IRF4. ISRE half-site orientation differs according to whether or not there is a 4 bp interval. The fraction of pCRMs for which the motif is found is indicated.

(D) Genome-wide interdependence of IRF4 and BATF co-occupancy in Th17-polarized cells. Scatter plots display the fold change in ChIP-seq reads versus significance. (Top) IRF4 in *Batf* WT versus KO. (Bottom) BATF in *Irf4* WT versus KO. Differences in ChIP-seq reads displayed for three relevant pCRM subtypes: (1) BATF or IRF4 alone (green); (2) BATF and IRF4 (orange); and (3) BATF, IRF4, plus additional TFs (purple). Distribution of fold changes of WT versus KO occupancy are displayed for proximal and distal pCRMs. Box plots are as shown in Figure 1C. RPM, reads per million. $^{**}p < 0.001$, Kolmogorov-Smirnov test.

(E) Reciprocally reduced occupancy of BATF and IRF4 at the *Cd28* locus in Th17-polarized cells deficient for IRF4 and BATF, respectively. ChIP-seq tracks were normalized for library size.

See also Figures S2 and S3.

does not participate in stabilizing positive feedback relationships with inducing TFs, thus rendering its expression sensitive to changing environmental signals. This is consistent with the need for continuous STAT3 activation (McGeachy et al., 2009) and the plasticity of the Th17 subset when cytokine conditions are altered (Hirota et al., 2011; Lee et al., 2009).

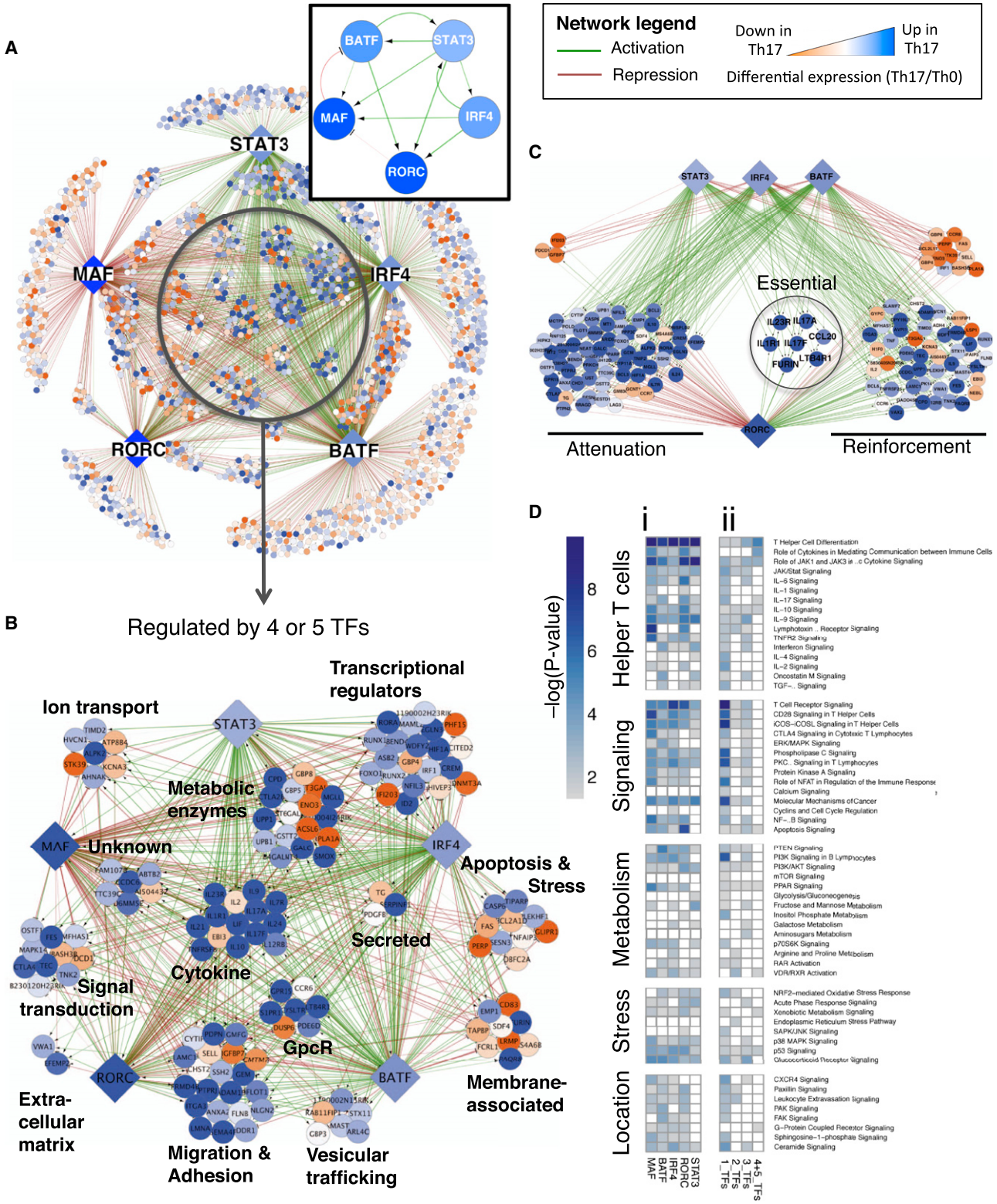


Figure 3. TF Combinatorial Interactions Specify the Th17 Lineage

(A) High-confidence regulatory edges (FDR < 10%; based on 10,000 simulations) focused on five core TFs identify direct (ChIP-seq) and functional (KO RNA-seq) regulatory targets (visualized using Cytoscape). Boxed inset displays the regulatory interactions between core Th17 TFs. See network legend for visualization scheme.

Highly regulated nodes, defined by combinatorial regulation by 4 or 5 core TFs, comprise many genes with critical lineage modulatory and effector activity. These include key signature genes (*Il17a*, *Il17f*, and *IL23r*), other relevant cytokines and receptors (*Il2*, *Il9*, *Lif*, *Il10*, *Il1r1*, *Il21*, *Il12rb1*, *Ebi3* [*IL-27*] *Ltb4r1*, and *Ccr6*), and TFs (*Rora*, *Hif1a*, *Runx1*, and *Foxo1*) (Korn et al., 2009). This indicates that other highly regulated genes in diverse categories (e.g., ion transport, migration, metabolism, and stress response) likely include novel Th17 regulators or effectors (Figure 3B).

To assess the regulatory relationships between Th17 TFs, we summarized the activating and repressing regulatory inputs for each network gene in a clustered heat map (Figure S4A). Of note, initiator TFs BATF, IRF4, and STAT3 regulate the largest number of genes and impose complementary control of shared targets, particularly in activation. This is most striking for highly regulated Th17 genes (Figures 3C and S4B) and is mirrored by the regulation of similar pathways (Figure 3D, i), including helper T cell differentiation and activation, cytokine signaling, metabolism, and oxidative/xenobiotic stress response (some previously attributed to these TFs [Durant et al., 2010; Kwon et al., 2009; Schraml et al., 2009]). Thus, together, initiator TFs establish a broad and coherent transcriptional program in Th17 cells.

c-Maf is generally appreciated as an activator of cytokine loci (Ho et al., 1999). Unexpectedly, in Th17 cells, it functions mainly as a negative regulator (Figures 3A and S4A), attenuating the expression of proinflammatory loci (e.g., *Rora*, *Runx1*, *Il1r1*, *Ccr6*, and *Tnf*) and globally repressing genes in pathways regulated by other core TFs (Figure 3D, i). Of note, c-Maf does positively regulate a few loci, several linked to attenuating inflammation (e.g., *Il9*, *Il10*, *Lif*, and *Ctla4*). Together with the recent description of c-Maf as an *Il22* repressor (Rutz et al., 2011), the global c-Maf target repertoire identifies an underappreciated general anti-inflammatory role for this TF.

RORγt, the key lineage specifier, functions as an activator and a repressor within the network (Figure S4A). Of note, it either reinforces or antagonizes the coherent activation program initiated by IRF4, BATF, and STAT3 (Figures 3C and S4B). Though RORγt positively regulates many genes, it has a strong role at only a small number of key Th17 loci (as defined above; Figures 3C and S3D; see Figure S4C for the magnitude of RORγt dependency). As a repressor, RORγt limits target expression, including regulators of metabolism and quiescence (e.g., *Il10*, *Hif1a*, *Egln3*, *Foxo1*, and *Il7r*) and alternative lineage fates (*Il4ra* and *Il12rb2*) promoted by initiator TFs. In this regard, RORγt acts as a modulator; its repressive activity is poorly correlated with actual expression changes, with its repressed targets often upregulated in Th17 relative to Th0 cells (Figures 3C and S4D). Thus, RORγt is essential in licensing the expression of a select few loci, and elsewhere it functions as a rheostat to tune mRNA levels to those of a Th17-specifying program.

Individual Th17 TFs regulate broad cellular functions (Figure 3D, i). When network targets were subdivided according to the complexity of inputs (either 1, 2, 3, or 4/5 TF edges), pathway analysis revealed a selective enrichment for genes involved in helper T cell differentiation with increasing number of TF inputs (Figure 3D, ii). Accordingly, nodes regulated by 4 or 5 TFs were most enriched for genes highly differentially expressed in Th17 cells (Figure 3A, compare node color in center versus periphery). Thus, although lineage TFs orchestrate the expression of genes in similar pathways, lineage specificity is a product of high-order combinatorial regulation.

Data Integration Allows for Discovery of Th17-Relevant Genes

The ChIP- and KO-based network is accurate but lacks regulatory information for other TFs with roles in Th17 cells. To learn interactions for those TFs, we integrated into our pipeline two other data sets (Figure 4A). The four data types are designated: [C] ChIP-seq; [K] RNA-seq KO; [R] 155 helper T cell RNA-seq experiments; and [I] public microarray data spanning 167 immune cell types and conditions (Immunological Genome Project; ImmGen) (Heng and Painter, 2008). In addition to providing regulatory information for new TFs, R and I can provide further support for interactions identified by the other two data types (C and K). Z scores for putative TF target regulatory interactions were individually assigned for R and I using the Inferelator (Greenfield et al., 2010) and were converted to rank-based scores, as above (ranging from 0 to 1), to facilitate their integration with K and C (Extended Experimental Procedures).

To validate our approach and evaluate complementarity of data sets, we estimated the predictive power of the resulting Th17 networks at recovering 74 literature-curated Th17-relevant genes (Table S1). As key Th17 genes were strongly regulated by multiple core TFs (Figure 3B), we reasoned that ranking genes based on their summed scores over all five core TFs would enrich for relevant target genes. We used two performance metrics, the areas under precision recall (aucPR) and under receiver operator curves (aucROC) that together provide a balance between estimating the sensitivity (aucROC) and accuracy (aucPR) of top-ranked predictions. Regardless of the metric used, for each data combination, summed TF target predictions (Figure 4B, bar) better identified Th17-relevant genes than targets for individual TFs (Figure 4B, points in bar), highlighting the predictive power of leveraging combinatorial regulation (also Figure S5A). Moreover, combining across data types

(B) Expanded view of highly regulated nodes with four to five core regulatory inputs, grouped based on general functional categories.
(C) Regulatory interactions shared by STAT3, IRF4, BATF, and RORγt highlighting different aspects of RORγt transcriptional function. (Attenuation) RORγt repression targets that are upregulated in Th17 cells; (reinforcement): activation targets that are upregulated in Th17 cells; (essential) targets with a 2-fold change in RORγt KO differential expression and KO H3K4me3 ChIP.
(D) Targets for single TFs are enriched for pathways in multiple functional categories (i). Targets of multiple TFs (increasingly regulated by 2, 3, or 4+5 TFs) are selectively enriched for pathways related to T helper differentiation and effector function (ii). Analysis performed using the ingenuity analysis tool (IPA) is presented as a heat map of enrichment p values.
See also Figure S4 and Table S4.

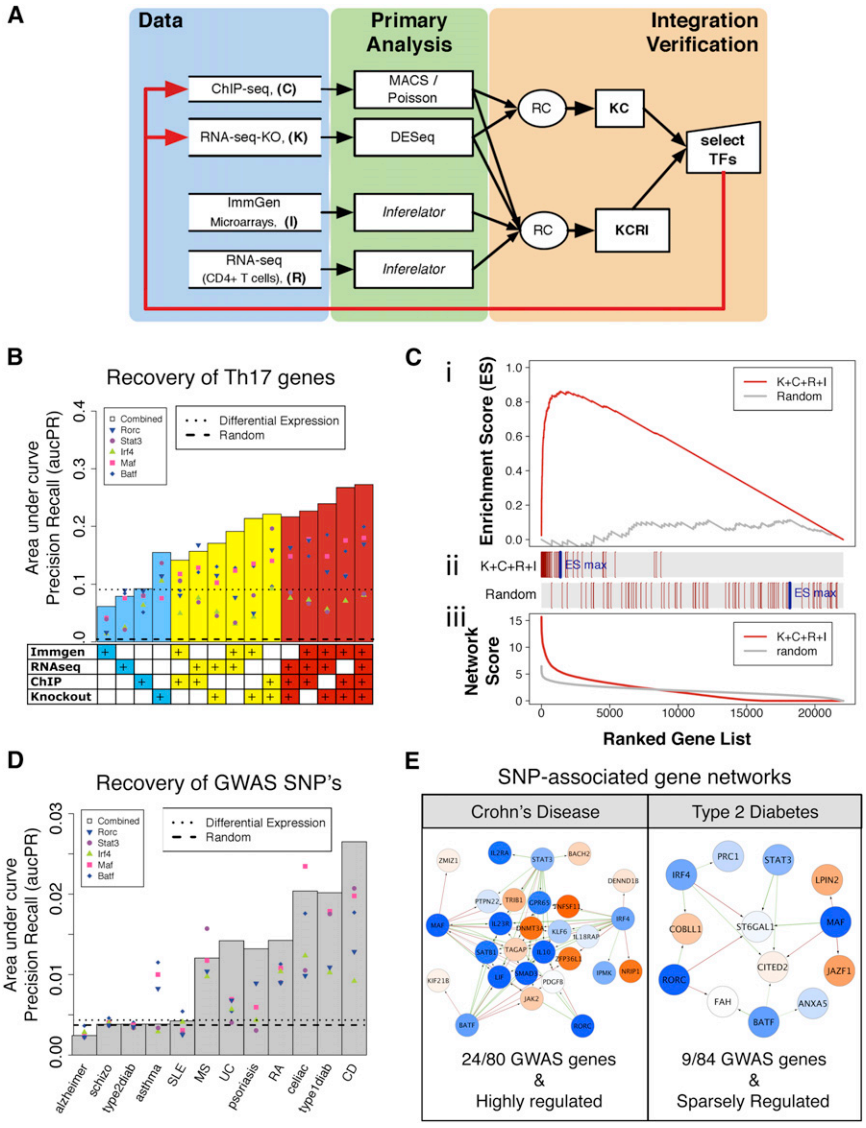


Figure 4. Network Model Performance and Validation

(A) Schematic for integration of four genomics and systems data types (K, C, R, and I) using a rank combined (RC) approach, resulting in the KC and KCRI networks.

(B) Performance measured as aucPR values indicating enrichment of literature-curated Th17 genes in networks derived from all possible data combinations. Points indicate single TF predictions (e.g., *BATF → target*), and bars indicate TF sum predictions (i.e., *[BATF+IRF4+STAT3+MAF+RORγt]→target*). (Dotted line) Reference performance for targets prioritized by differential expression (Th17 versus Th0); (dashed line) for random.

(C) Gene set enrichment analysis (GSEA, Broad Institute) for the top-performing KCRI network. (i) The ranked list of TF sum targets recovers Th17-relevant genes with a maximal enrichment score (ES) of 0.86 out of 1 (red line); random (gray line). (ii) Vertical red lines indicate where in the ranked list literature-derived genes were recovered. (iii) Summed TF score distributions for KCRI (red line) and random (gray line).

(D) The KCRI network selectively recovers GWAS SNP-linked genes for Th17-implicated inflammatory diseases. Recovery of SNP-associated disease genes is measured in terms of aucPR for TF sum predictions (gray bars) and single TF predictions (points).

(E) Core TF networks for genes associated with GWAS of Crohn's disease and type 2 diabetes. The KCRI network recovered 24 out of 80 Crohn's disease genes ($p = 10^{-7}$) and 9 out of 84 type 2 diabetes genes ($p = 0.29$). Network display is as in Figure 3. p values calculated by Fisher's exact test.

See also Figure S5 and Tables S1 and S4.

resulted in higher quality networks (note the performance increase with data integration; Figures 4B and S5A). This is true for both K + C, which produces a smaller (core-TF-centered) but accurate network, and for R + I, which generates a more comprehensive (more than 170 differentially expressed TFs) but less accurate model. Thus, combining data as in K + C + R + I provides additional network information without compromising the K + C network accuracy, leading to a 3-fold performance boost at finding relevant Th17 genes over the conventional differential expression list (Figure 4B, dotted line). Moreover, the nonparametric rank-based scheme is better suited for combining data than an alternative parametric Fisher method for combining p values (Figures S5C and S5D).

We further characterized the predictions of the top-performing KCRI network using gene set enrichment analysis (GSEA; Figure 4C). Literature-curated genes were highly enriched as top predictions (60 of 74 total), ranking many more putative Th17-relevant genes with comparable KCRI scores (1,328 of ~22,000 total).

The wealth of GWAS information for diseases in which Th17 cells are implicated allowed us to judge the relevance of our mouse Th17 network for human disease (Hindorff et al., 2012). Loci linked to SNPs associated with ulcerative colitis (UC), Crohn's disease (CD), multiple sclerosis, psoriasis, and rheumatoid arthritis were significantly enriched in the core (KCRI) Th17 network (Figures 4D, 4E, and S5B). This is in striking contrast to diseases with enrichment scores expected by chance (Alzheimer's disease and schizophrenia) and other inflammatory diseases in which Th17 cell involvement is not established (e.g., type 2 diabetes and systemic lupus erythematosus [SLE]). Of note, for those diseases with strong links to Th17 function, GWAS-associated genes are more highly regulated by the core Th17 TFs (Figure 4E), and TF sum scores (Figure 4D, bars) better recovered GWAS loci relative to individual TFs (Figure 4D, points). This strongly supports the notion that synergy between the five core Th17 TFs is also relevant for function of the Th17 lineage in human disease.

Network Analysis Identifies Modulators of the Th17 Program

Our validations revealed that highly regulated targets of the KCRI network can be exploited to identify Th17-relevant effector genes. However, early response regulators that function upstream of or in parallel with core TFs may not be captured (Figure 3A). To address this, we used an independent Inferelator-derived ImmGen network (I) to predict new TFs that demonstrate significant target overlap with the five core TFs (KC network; see Extended Experimental Procedures); a similar method was proven successful (Carro et al., 2010). Accordingly, 26 candidate Th17 TFs were prioritized that either: (1) were top-scoring KCRI-network-regulated genes (90th percentile) or (2) showed significant overlap between their predicted targets and targets of the five core TFs. The latter TF enrichment analysis identified all core TFs as top hits and several known Th17 regulators (see full list in Table S2), including RORα, AHR, RBPJ, and TBX21 (Alam et al., 2010; Veldhoen et al., 2008; Yang et al., 2008).

To assess the effects of prioritized TFs on Th17 cell differentiation, we performed gain- and loss-of-function experiments. Among the 16 TFs overexpressed in $CD4^+$ T cells by retroviral transduction, several had a significant effect on the percentage of IL-$17A^+$ cells generated, including Ets factor Etv6, Ncoa2, Smad3, Hif1a, Skil, and Trib3 (Figures S6A and S6B). Most striking was the AP-1 family member Fosl2, the top-predicted factor in the TF enrichment analysis (Table S2), whose overexpression significantly reduced the number of IL-17A-producing cells.

In a complementary approach, we performed siRNA-mediated knockdown (KD) experiments for 14 candidate TFs with siRNA pools electroporated into activated $CD4^+$ T cells subjected to Th17-polarizing conditions. Reductions in target TF mRNAs were similar to those observed with siRorc, which effectively reduced RORγt (Figures S6C and S6D), its targets (*Il17a/f*, *Il23r*, and *Il22*), and Th17 differentiation (Figures 5A, 5B, and S6C). Strikingly, six TF KDs significantly altered IL-17A production relative to a nontargeting control, without affecting the concomitant generation of $Foxp3^+$ iTreg cells. These included Etv6, Nfatc2, Bcl11b, Crem, and regulators of chromatin remodeling (Satb1 and Kdm6b [Jmjd3]) (Figures 5A and 5B). Thus, the network method prioritized and made accurate predictions about genes that influence expression of IL-17A, a key component of the Th17 phenotype.

To assess factor influence on the broad Th17 program, we performed RNA-seq of TF KD cultures. Global pathway analysis for TF-dependent genes identified distinct factor-related pathways yet showed a striking convergence in enrichment for genes involved in T helper cell differentiation/function for all factors except Sirt2 and a non-TF control, Ccr6 (Figure 5C). Indeed, TF-dependent loci comprise a variety of helper T cell effector or lineage-specializing genes (*Il4*, *Ifng*, *Gata3*, *Foxp3*, *Tbx21*, *Il22*, *Il1r1*, *Il23r*, *Il10*, *Il24*, *Il9*, and *Ccl20*) (Figure 5D). Of particular interest, Bcl11b influenced the expression of a broad set of helper T cell modulatory genes, suggesting a key regulatory role in subset diversification. Similarly, Jmjd3, a lysine K27 demethylase and known T-bet partner (Miller et al., 2010), displayed a marked influence on the activation of multiple Th17-expressed cytokines, suggesting that it may partner with core Th17 TFs. Indeed, Jmjd3 shares many direct targets (KD + C) with RORγt and STAT3 (Figure S6E). Taken together, the results show that the network model identifies many candidate modulators of the Th17 program.

Fosl2 Is a Core Component of the Th17 Specification Program

The AP-1 family TF Fosl2 was the highest-ranking candidate to coregulate targets with core TFs (Table S2). We interrogated its role in helper T cell differentiation and function using mice with conditional deletion of Fosl2 in T cells ($Fosl2^{fl/fl}$ CD4-Cre). Fosl2-deficient $CD4^+$ T cells could be polarized in vitro into Th1, Th2, Th17, or iTreg cells, but notably, cytokine production was dysregulated (Figures 6A and S7A). Fosl2 null Th17 cultures were markedly increased for IL-17A-producing and atypical $Foxp3^+$IL-$17A^+$ cells. There was also low-level derepression of Il17a in mutant Th1 and Th2 cells, consistent with Fosl2 function as an Il17a repressor (Figures S6A, S6B, and S7A). Fosl2 deficiency also enabled IFNγ production in Th17 and Th2 cultures, particularly when Th17 cells were subsequently exposed to Th1-skewing conditions (Figures S7A and S7B).

We next examined the role of Fosl2 in the Th17-dependent disease model experimental autoimmune encephalomyelitis (EAE), which mimics the CNS pathology in multiple sclerosis. $Fosl2^{fl/fl}$ CD4-Cre mice had significantly attenuated disease severity compared to wild-type controls (Figure 6B). Analysis of spinal cord infiltrates at 21 days postimmunization revealed reduced $CD4^+$ T cells but similar percentages of IL-17A, IFNγ, and GM-CSF producers in mutant mice. Strikingly, the Fosl2-deficient cytokine-producing T helper cells also expressed the TF Foxp3, which specifies the Treg program (Figures 6B and 6C). Consistent with our in vitro observations (Figure 6A), these findings suggest Fosl2 as a key regulator of T helper lineage plasticity, particularly under inflammatory conditions.

To gain insight into Fosl2 function, we identified its direct targets using K + C analysis. Consistent with previous studies, Fosl2 both activates and represses target loci (Wagner and Eferl, 2005). It attenuates expression of Th17 signature genes in addition to Il17a (*Il17f*, *Ccl20*, *Ccr6*, *Il1r1*, and *Batf*) and of Th1 regulatory loci (*Tbx21*, *Il18r1*, *Il18rap*, and *Il2*), suggesting a role in controlling inflammatory responses and preventing Th1 specification (Figure 6D). Conversely, Fosl2 also promotes the expression of genes that drive Th17 maintenance and survival (*Il6ra*, *Il-23r*, *Il12rb1*, *il7r*, and *Il21*) and helper T cell diversification or function (*Il4ra*, *Il12rb2*, *Il2ra*, *Il10ra*, *Ltb4r1*, *Smad3*, and *Hif1a*) (Figure 6D). These loci are also targeted by the five core TFs, indicating that Fosl2 modulates the lineage identity and functional programs regulated by core Th17 TFs.

In light of the dominant regulation by AP-1 factors in the Th17 lineage, it was interesting to observe a high degree of overlap in occupancy by Fosl2 and BATF (Figure 6E), suggestive of an antagonistic relationship between them. This may be mediated by direct competition for the same binding sites (Figure S7C) and enhanced by direct transcriptional repression of *Batf* by Fosl2 (Figure S7D). Thus, as predicted, Fosl2 is a highly interconnected component of the core Th17 specification program. This is in contrast to Hif1α, a recently identified regulator of Th17 cells (Dang et al., 2011), which was not predicted to share a significant

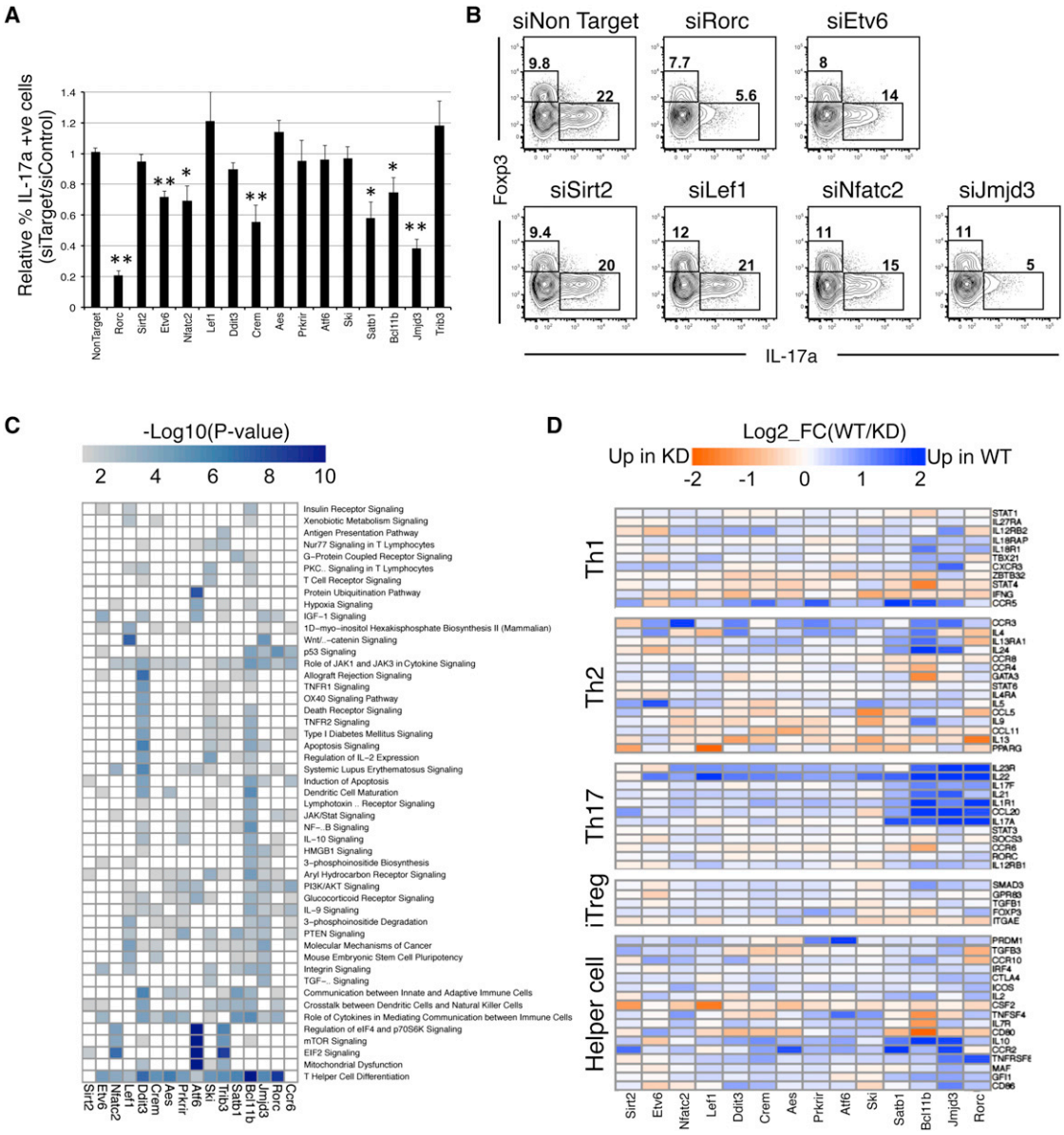


Figure 5. Identification of Th17 Regulators

(A) siRNA knockdown screen of candidate TF function in Th17 differentiation. Percent of IL17A-producing cells relative to the control siRNA condition for knockdown cultures analyzed at 24 hr of Th17 polarization. Error bars represent SD of two experiments conducted in triplicate. *$p < 0.05$ and **$p < 0.01$, t test.

(B) Flow cytometric analysis for Th17-polarized cells transfected with siRNAs for the indicated gene targets.

(C) Shared and unique functions of Th17 TF regulators. Heat map of ingenuity pathway enrichment (IPA, $p < 0.01$) for candidate TF-dependent genes.

(D) TF candidates influence the expression of immune-modulatory genes. Heat map of log2 fold change in expression of T helper signature genes in the siRNA knockdowns relative to nontargeting control for 24 hr Th17 polarization cultures.

See also Figure S6 and Table S2.

number of targets with core TFs (Table S3) and is not as interconnected as Fosl2 (Figure 6D).

Visualization and Exploration of the Extended KCRI Th17 Network

The final KCRI network comprises 173 differentially expressed TFs controlling 3,679 genes with ~19,000 interactions. The subset of this network surrounding regulators based on ChIP-seq and KO data is more accurate but is limited (this network has seven TFs controlling 2,218 genes, with 4,237 edges). To facilitate interrogation of the Th17 transcriptional network by the scientific community, we provide access to the primary data (see Table S3 and Supplemental Information), the networks (KC and KCRI), and analysis tools at http://th17.bio.nyu.edu.

DISCUSSION

Th17 cells exert critical functions in immune defense at mucosal barriers and are implicated as contributors to multiple autoimmune diseases (Korn et al., 2009). Since the discovery of Th17 cells, multiple TFs involved in the production of IL-17A and in inflammation have been described, but little was known of how they collaborate in the global transcriptional program governing Th17 specification and function. Here, we aimed to accurately define how Th17 TFs integrate functionally to execute this program and created a useful network model that can be exploited to uncover lineage regulators, effectors, and potential therapeutic targets (Figure 7). To achieve this, we used a culture system for Th17 differentiation to build a transcriptional network model based on combinations of data sets and analytical approaches and determined its performance using both computational and experimental validations, testing the role of predicted regulators in vitro and in a disease model of inflammation. Our combined computational and experimental approach allowed for iteration between the generation of a data integrative network and follow-up investigation of individual genes, and, hence, for continuous network refinement (Figure 4A). This work provides a clear experimental design and analysis framework that can be adopted for other cell lineages in the immune system and elsewhere.

Dynamics of TF Function in the Specification of CD4$^+$ T Cells

Transcription initiation at sites occluded by nucleosomes and high-order chromatin structure requires mechanisms for making specific regions accessible to appropriate regulators (Zaret and Carroll, 2011). In TCR-activated CD4$^+$ T cells, BATF and IRF4 bind cooperatively to sites throughout the genome. In the presence of Th17-polarizing cytokines, STAT3, c-Maf, and RORγt are recruited to many of the same sites. Chromatin accessibility analysis suggests that BATF/IRF4 complexes pioneer the access of other TFs that further specify functional subsets. Indeed, BATF and IRF4 have critical roles in multiple Th cells (Brüstle et al., 2007; Ise et al., 2011; Rengarajan et al., 2002; Schraml et al., 2009). As these TFs are upregulated in Th0 cells, it is interesting to speculate that pioneering function provides the T cell with plasticity to differentiate in multiple directions, depending on the cytokine environment. Thus, whereas TGF-β and STAT3-activating signals would recruit STAT3/RORγt to a subset of BATF/IRF4 binding sites, Th1 or Th2 signals may recruit STAT1/T-bet or STAT6/GATA-3 to others. It will be of interest to compare the global distribution of BATF, IRF4, and lineage-specifying TFs in Th1 and Th2 cells.

Fosl2 is a negative regulator of IL-17A. Thus, the finding that Fosl2-deficient mice had a reduced inflammatory response in the EAE model was unexpected. This may reflect the requirement for Fosl2 for expression of key loci supporting Th17 cell maintenance. The result may also be explained by derepression of Foxp3 in inflammatory T cells producing IL-17A, IFNγ, and GM-CSF, which may be mediated, in part, by reduced expression of Hif1α—a Foxp3 inhibitor—in Fosl2-deficient T cells (Dang et al., 2011). Foxp3$^+$ T helper cells that produce effector cytokines have been described in humans and have been shown to have regulatory activity (Voo et al., 2009). Hence, reduced disease scores in Fosl2 deficiency may be due to increased activity of Treg-like cells infiltrating the CNS. Fosl2-deficient T cells also derepress T-bet and IFNγ, suggesting that Fosl2 serves as a brake, binding to sites otherwise occupied by BATF and IRF4 to prevent expression. Indeed, Fosl2 occupancy overlaps with that of BATF and IRF4 in Th0 and Th17 cells, where it likely competes for AP-1 sites (Figure 7). Thus, Fosl2 is a highly integrated regulator of T helper cell lineage identity, functioning to limit plasticity of Th17 cells by repressing Th1 and Treg transcriptional programs potentially by balancing the activity of BATF/IRF4 at key loci. Our analysis also highlights that regulation of a single cytokine, i.e., IL-17A, does not reflect broad functions of the controlling TFs. A global perspective in the context of a multi-TF causal regulatory network aids in deciphering the role of individual factors.

RORγt has been described as a "master regulator" for the Th17 program, yet it has a surprisingly small regulatory footprint. RORγt deficiency had limited effects on p300 recruitment and H3K4 methylation, suggesting that it lacks a major role in remodeling its regulated loci. However, a handful of loci were highly dependent on RORγt for these early inductive events (Figure 7); what distinguishes this selectivity remains to be uncovered. This focal mode of regulation, coupled with a generalized program upon which RORγt functions to tune expression, is consistent with the plasticity of Th17 cells, suggesting that expression and chromatin state at key Th17 loci might be amenable to rapid change depending on cytokine environment. The lack of stabilizing positive feedback of RORγt to initiators may permit such T helper program switching. Moreover, though RORγt attenuates the expression of regulators of alternative Th subsets (*il4ra*, *il12rb*, and *Tbx21*), these loci are nevertheless expressed. Thus, RORγt is not a prototypical "master" regulator that functions to "lock in" lineage programs. This renders RORγt an exceptional drug target, as therapeutic intervention would not be expected to perturb the generic regulatory programs shared by other cell types.

A Highly Predictive Th17 Cell Network Model

The iterative approach applied here was successful in uncovering important aspects of Th17 biology, generating a model that captures most of the previously identified Th17-relevant genes among the top candidates and predicting many more with equal confidence. Among these, several TFs and chromatin modifiers were shown to affect Th17 differentiation or the expression of immune-modulatory genes, including *Bcl11b*, *Etv6*, and *Jmjd3*. Although we focused on the top-predicted TF *Fosl2*, we expect that many more candidates will be pertinent to Th17 biology. We anticipate the network model presented herein to be a highly useful tool for exploration and for generation of new hypotheses.

Although the Th17 network largely models in vitro differentiation, it is nonetheless likely to be relevant for in vivo Th17 cell functions. Indeed, *Fosl2* mutant T cells were compromised in effector function in an autoimmunity model, and similar phenotypes were reported in mice deficient for many top-scoring network genes. Moreover, the network is selectively enriched for genes with orthologs that harbor SNPs associated with human inflammatory diseases linked to Th17 cell-mediated pathology, such as Crohn's disease and psoriasis. GWAS

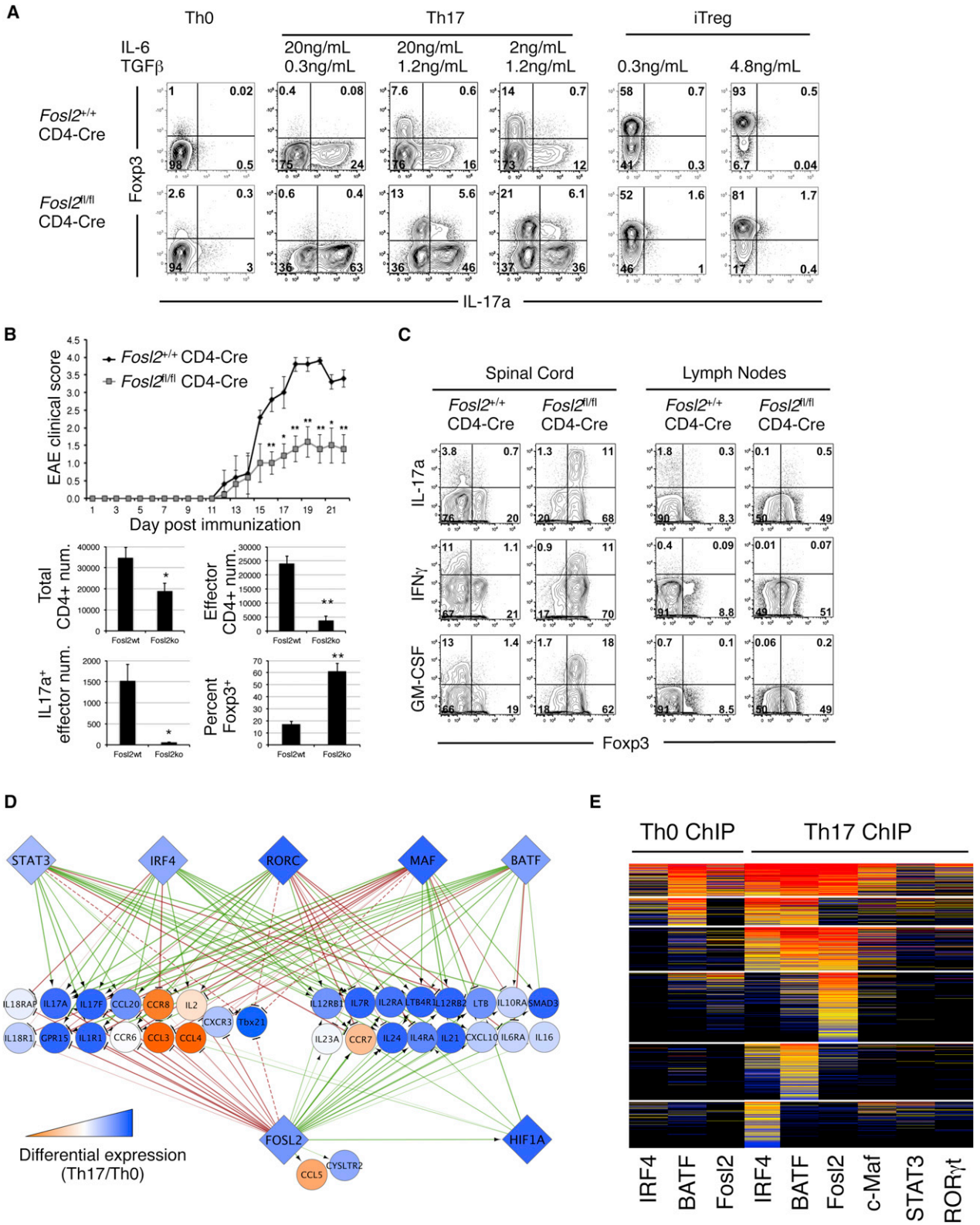


Figure 6. Fosl2 Regulates Loci Critical to Lineage Identity and Function

(A) Fosl2 negatively regulates IL-17A expression. Flow cytometric analysis for naive *Fosl2* WT and KO CD4 T cells cultured as indicated for 3 days. Representative of four experiments.

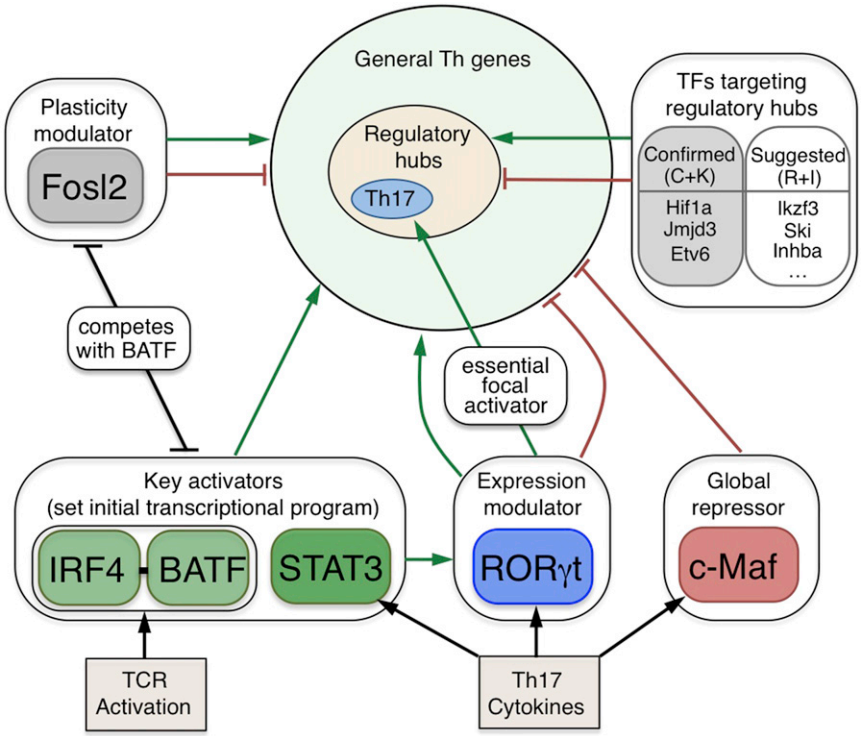


Figure 7. Model for Th17 TF Functions during Lineage Specification

Functions of the core TFs and a selection of newly identified TFs in regulating expression of general T-helper-cell- and Th17-cell-associated genes. BATF/IRF4 complexes, transcriptionally induced following TCR signaling, mutually activate the expression of a large set of target genes, together with STAT3. RORγt drives expression of a small subset of key Th17 genes and modulates the expression of genes activated by initiator TFs, BATF/IRF4/STAT3. Fosl2 restricts the expression of genes required for alternate CD4$^+$ differentiation programs. c-Maf functions as a general repressor. Regulatory hubs include loci that receive a high level of input from Th17 TFs and are enriched for genes that are critical for Th17 differentiation and function.

studies, while facilitating the identification of genes involved in complex diseases involving multiple cell types, are often difficult to translate into biological hypotheses amenable to investigation. However, our analysis identified several GWAS-implicated genes as candidate Th17-specific mediators of pathogenesis (e.g., *PTPN22*, *LIF*, and *KLF6*) and may be used to implicate Th17 cells in the etiology of particular conditions. Deconvoluting GWAS data by leveraging the information from accurate and comprehensive transcriptional regulatory networks to provide cellular context, reveal functional epistasis, and prioritize genes of potential medical importance will likely prove to be a powerful approach in uncovering disease mechanisms and developing new diagnostic and therapeutic tools (Califano et al., 2012). Taken together, this body of work is an excellent example of how the power of systems biology can be harnessed to answer a specific large-scale biological question, thus providing a validated paradigm for similar undertakings.

EXPERIMENTAL PROCEDURES

For full details of experimental and computational approaches, see the Extended Experimental Procedures and Extended Computational Methods. In brief, naive CD4$^+$ T cells were purified from lymph nodes and spleen of wild-type mice or TF KO mice. The cells were cultured under Th0 and Th17 conditions and were processed for ChIP-, RNA- and FAIRE-seq. The analysis pipeline included dedicated methods for learning from individual data sources.

ACCESSION NUMBERS

The GEO accession number for the full data set (ChIP-seq and RNA-seq) is GSE40918.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, Extended Computational Methods, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2012.09.016.

ACKNOWLEDGMENTS

We are grateful to Andrew Sczesnak for generating TDF files and for multiple other contributions to data analysis during the course of the study. We thank R. Dalla Favera for providing IRF4 mutant mice; O. Uluckan and S. Wurm for coordinating shipping of AP-1 mutant mice; K. Ward for system administration; and K. Gunsalus and M. Dustin for helpful discussions. Support was provided

(B) Conditional deficiency of *Fosl2* in CD4 T cells reduces the severity of EAE. (Top) Clinical scores (mean ± SEM) for CD4-cre mice with WT or conditional *Fosl2* alleles (*$p < 0.05$ and **$p < 0.01$, t test). (Bottom) Total numbers or percentages of CD4$^+$ T cell populations isolated from spinal cord (mean ± SEM). Effector CD4$^+$ T cells are defined as Foxp3$^-$. Representative of two experiments.

(C) Predominance of cytokine-producing Foxp3$^+$ cells in mice during EAE. Flow cytometric analysis of restimulated CD4$^+$ T cells isolated from the spinal cord and lymph nodes of mice on day 22 post-EAE induction. Plots were gated on CD4$^+$CD45$^+$ cells.

(D) Combinatorial core TF targets involved in T cell specification are highly regulated by Fosl2. Network depiction is as in Figure 3. (Dashed lines) Edges added manually to account for Tbx21 with ChIP peaks outside of the set boundaries (10 kb flanking the transcribed region).

(E) The clustered heat map of Fosl2 and core TF occupancy reveals that the binding domain of Fosl2 largely overlaps with that of BATF.

See also Figure S7.

by NIH grants RC1 AI087266 and RC4 AI092765 (A.A., D.R.L, R.B, and R.M.M.), PN2 EY016586 (A.G., A.M, and R.B.), and IU54CA143907-01 and EY016586-06 (A.G. and R.B.); NSF grant IOS-1126971 (R.B.); fellowships from the Leukemia and Lymphoma Society and Crohn's and Colitis Foundation of America (M.C.), the National Arthritis Research Foundation (Y.Y.); and Irvington Institute Fellowships from the Cancer Research Institute (M.S. and W.H.).

Received: August 23, 2012
Revised: September 12, 2012
Accepted: September 17, 2012
Published online: September 25, 2012

REFERENCES

Alam, M.S., Maekawa, Y., Kitamura, A., Tanigaki, K., Yoshimoto, T., Kishihara, K., and Yasutomo, K. (2010). Notch signaling drives IL-22 secretion in CD4+ T cells by stimulating the aryl hydrocarbon receptor. Proc. Natl. Acad. Sci. USA *107*, 5943–5948.

Bauquet, A.T., Jin, H., Paterson, A.M., Mitsdoerffer, M., Ho, I.C., Sharpe, A.H., and Kuchroo, V.K. (2009). The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. Nat. Immunol. *10*, 167–175.

Bonneau, R., Facciotti, M.T., Reiss, D.J., Schmid, A.K., Pan, M., Kaur, A., Thorsson, V., Shannon, P., Johnson, M.H., Bare, J.C., et al. (2007). A predictive model for transcriptional control of physiology in a free living cell. Cell *131*, 1354–1365.

Brüstle, A., Heink, S., Huber, M., Rosenplänter, C., Stadelmann, C., Yu, P., Arpaia, E., Mak, T.W., Kamradt, T., and Lohoff, M. (2007). The development of inflammatory T(H)-17 cells requires interferon-regulatory factor 4. Nat. Immunol. *8*, 958–966.

Califano, A., Butte, A.J., Friend, S., Ideker, T., and Schadt, E. (2012). Leveraging models of cell regulation and GWAS data in integrative network-based association studies. Nat. Genet. *44*, 841–847.

Carro, M.S., Lim, W.K., Alvarez, M.J., Bollo, R.J., Zhao, X., Snyder, E.Y., Sulman, E.P., Anne, S.L., Doetsch, F., Colman, H., et al. (2010). The transcriptional network for mesenchymal transformation of brain tumours. Nature *463*, 318–325.

Codarri, L., Gyülvészi, G., Tosevski, V., Hesske, L., Fontana, A., Magnenat, L., Suter, T., and Becher, B. (2011). RORγt drives production of the cytokine GM-CSF in helper T cells, which is essential for the effector phase of autoimmune neuroinflammation. Nat. Immunol. *12*, 560–567.

Crotty, S. (2011). Follicular helper CD4 T cells (TFH). Annu. Rev. Immunol. *29*, 621–663.

Dang, E.V., Barbi, J., Yang, H.Y., Jinasena, D., Yu, H., Zheng, Y., Bordman, Z., Fu, J., Kim, Y., Yen, H.R., et al. (2011). Control of T(H)17/T(reg) balance by hypoxia-inducible factor 1. Cell *146*, 772–784.

Durant, L., Watford, W.T., Ramos, H.L., Laurence, A., Vahedi, G., Wei, L., Takahashi, H., Sun, H.W., Kanno, Y., Powrie, F., and O'Shea, J.J. (2010). Diverse targets of the transcription factor STAT3 contribute to T cell pathogenicity and homeostasis. Immunity *32*, 605–615.

Eisenbeis, C.F., Singh, H., and Storb, U. (1995). Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. Genes Dev. *9*, 1377–1387.

Ernst, J., Vainas, O., Harbison, C.T., Simon, I., and Bar-Joseph, Z. (2007). Reconstructing dynamic regulatory maps. Mol. Syst. Biol. *3*, 74.

Faith, J.J., Hayete, B., Thaden, J.T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J.J., and Gardner, T.S. (2007). Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol. *5*, e8.

Greenfield, A., Madar, A., Ostrer, H., and Bonneau, R. (2010). DREAM4: Combining genetic and dynamic information to identify biological networks and dynamical models. PLoS ONE *5*, e13397.

Heng, T.S., and Painter, M.W.; Immunological Genome Project Consortium. (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094.

Hindorff, L.A., MacArthur, J., Wise, A., Junkins, H.A., Hall, P.N., Klemm, A.K., and Manolio, T.A. (2012). A Catalog of Published Genome-Wide Association Studies. (http://www.genome.gov/gwastudies).

Hirota, K., Duarte, J.H., Veldhoen, M., Hornsby, E., Li, Y., Cua, D.J., Ahlfors, H., Wilhelm, C., Tolaini, M., Menzel, U., et al. (2011). Fate mapping of IL-17-producing T cells in inflammatory responses. Nat. Immunol. *12*, 255–263.

Ho, I.C., Kim, J.I., Szabo, S.J., and Glimcher, L.H. (1999). Tissue-specific regulation of cytokine gene expression. Cold Spring Harb. Symp. Quant. Biol. *64*, 573–584.

Ise, W., Kohyama, M., Schraml, B.U., Zhang, T., Schwer, B., Basu, U., Alt, F.W., Tang, J., Oltz, E.M., Murphy, T.L., and Murphy, K.M. (2011). The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. Nat. Immunol. *12*, 536–543.

Ivanov, I.I., McKenzie, B.S., Zhou, L., Tadokoro, C.E., Lepelley, A., Lafaille, J.J., Cua, D.J., and Littman, D.R. (2006). The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. Cell *126*, 1121–1133.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V.K. (2009). IL-17 and Th17 cells. Annu. Rev. Immunol. *27*, 485–517.

Kwon, H., Thierry-Mieg, D., Thierry-Mieg, J., Kim, H.P., Oh, J., Tunyaplin, C., Carotta, S., Donovan, C.E., Goldman, M.L., Tailor, P., et al. (2009). Analysis of interleukin-21-induced Prdm1 gene regulation reveals functional cooperation of STAT3 and IRF4 transcription factors. Immunity *31*, 941–952.

Lee, Y.K., Turner, H., Maynard, C.L., Oliver, J.R., Chen, D., Elson, C.O., and Weaver, C.T. (2009). Late developmental plasticity in the T helper 17 lineage. Immunity *30*, 92–107.

Leppkes, M., Becker, C., Ivanov, I.I., Hirth, S., Wirtz, S., Neufert, C., Pouly, S., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., et al. (2009). RORgamma-expressing Th17 cells induce murine chronic intestinal inflammation via redundant effects of IL-17A and IL-17F. Gastroenterology *136*, 257–267.

Madar, A., Greenfield, A., Vanden-Eijnden, E., and Bonneau, R. (2010). DREAM3: network inference using dynamic context likelihood of relatedness and the inferelator. PLoS ONE *5*, e9803.

Manel, N., Unutmaz, D., and Littman, D.R. (2008). The differentiation of human T(H)-17 cells requires transforming growth factor-beta and induction of the nuclear receptor RORgammat. Nat. Immunol. *9*, 641–649.

Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., Kellis, M., Collins, J.J., and Stolovitzky, G.; The DREAM5 Consortium. (2012a). Wisdom of crowds for robust gene network inference. Nat. Methods *9*, 796–804.

Marbach, D., Roy, S., Ay, F., Meyer, P.E., Candeias, R., Kahveci, T., Bristow, C.A., and Kellis, M. (2012b). Predictive regulatory models in Drosophila melanogaster by integrative inference of transcriptional networks. Genome Res. *22*, 1334–1349.

Mattick, J.S., Taft, R.J., and Faulkner, G.J. (2010). A global view of genomic information—moving beyond the gene and the master regulator. Trends Genet. *26*, 21–28.

McGeachy, M.J., Chen, Y., Tato, C.M., Laurence, A., Joyce-Shaikh, B., Blumenschein, W.M., McClanahan, T.K., O'Shea, J.J., and Cua, D.J. (2009). The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. Nat. Immunol. *10*, 314–324.

Miller, S.A., Mohn, S.E., and Weinmann, A.S. (2010). Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. Mol. Cell *40*, 594–605.

Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T., et al. (2011). Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell *144*, 296–309.

Okamoto, K., Iwai, Y., Oh-Hora, M., Yamamoto, M., Morio, T., Aoki, K., Ohya, K., Jetten, A.M., Akira, S., Muta, T., and Takayanagi, H. (2010). IkappaBzeta

regulates T(H)17 development by cooperating with ROR nuclear receptors. Nature *464*, 1381–1385.

Prill, R.J., Marbach, D., Saez-Rodriguez, J., Sorger, P.K., Alexopoulos, L.G., Xue, X., Clarke, N.D., Altan-Bonnet, G., and Stolovitzky, G. (2010). Towards a rigorous assessment of systems biology models: the DREAM3 challenges. PLoS ONE *5*, e9202.

Rengarajan, J., Mowen, K.A., McBride, K.D., Smith, E.D., Singh, H., and Glimcher, L.H. (2002). Interferon regulatory factor 4 (IRF4) interacts with NFATc2 to modulate interleukin 4 gene expression. J. Exp. Med. *195*, 1003–1012.

Rutz, S., Noubade, R., Eidenschenk, C., Ota, N., Zeng, W., Zheng, Y., Hackney, J., Ding, J., Singh, H., and Ouyang, W. (2011). Transcription factor c-Maf mediates the TGF-β-dependent suppression of IL-22 production in T(H)17 cells. Nat. Immunol. *12*, 1238–1245.

Schraml, B.U., Hildner, K., Ise, W., Lee, W.L., Smith, W.A., Solomon, B., Sahota, G., Sim, J., Mukasa, R., Cemerski, S., et al. (2009). The AP-1 transcription factor Batf controls T(H)17 differentiation. Nature *460*, 405–409.

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.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854–858.

Voo, K.S., Wang, Y.H., Santori, F.R., Boggiano, C., Wang, Y.H., Arima, K., Bover, L., Hanabuchi, S., Khalili, J., Marinova, E., et al. (2009). Identification of IL-17-producing FOXP3+ regulatory T cells in humans. Proc. Natl. Acad. Sci. USA *106*, 4793–4798.

Wagner, E.F., and Eferl, R. (2005). Fos/AP-1 proteins in bone and the immune system. Immunol. Rev. *208*, 126–140.

Wei, G., Wei, L., Zhu, J., Zang, C., Hu-Li, J., Yao, Z., Cui, K., Kanno, Y., Roh, T.Y., Watford, W.T., et al. (2009). Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. Immunity *30*, 155–167.

Yang, X.O., Panopoulos, A.D., Nurieva, R., Chang, S.H., Wang, D., Watowich, S.S., and Dong, C. (2007). STAT3 regulates cytokine-mediated generation of inflammatory helper T cells. J. Biol. Chem. *282*, 9358–9363.

Yang, X.O., Pappu, B.P., Nurieva, R., Akimzhanov, A., Kang, H.S., Chung, Y., Ma, L., Shah, B., Panopoulos, A.D., Schluns, K.S., et al. (2008). T helper 17 lineage differentiation is programmed by orphan nuclear receptors ROR alpha and ROR gamma. Immunity *28*, 29–39.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev. *25*, 2227–2241.

Zhang, F., Meng, G., and Strober, W. (2008). Interactions among the transcription factors Runx1, RORgammat and Foxp3 regulate the differentiation of interleukin 17-producing T cells. Nat. Immunol. *9*, 1297–1306.

Zhu, J., Yamane, H., and Paul, W.E. (2010). Differentiation of effector CD4 T cell populations (*). Annu. Rev. Immunol. *28*, 445–489.

EXTENDED EXPERIMENTAL PROCEDURES

Mice

Mice were bred and maintained in the animal facility of the Skirball Institute (Langone Medical Center, NYU) in specific pathogen-free conditions. C57Bl/6, and *Hif1a*$^{fl/fl}$ (Ryan et al., 2000) mice were obtained from Jackson laboratories. *Rorc(t)* knock-out mice harboring a GFP reporter cDNA at the translation initiation site have been described (Eberl et al., 2004). Mutant strains were kindly provided by the following researchers: *Stat3*$^{fl/fl}$ (Lee et al., 2002), D. Levy (NYU); *Irf4*$^{fl/fl}$,(Klein et al., 2006) R. Dalla-Favera (Columbia University); *Maf* $^{fl/fl}$ (Wende et al., 2012), C. Birchmeier (MDC, Germany); *Batf* $^{fl/fl}$ (Schraml et al., 2009), K.M. Murphy (Washington University); and *Fosl2*$^{fl/fl}$ (Karreth et al., 2004), E. Wagner (CNIO, Spain). *Irf4*$^{fl/fl}$ mice were mated with EIIa-Cre transgenic mice to obtain fully IRF4 null animals. All animal procedures in accordance with protocols approved by the Institutional Animal Care and Usage Committee of New York University.

Cell Culture

Naive CD4$^+$ T cells were purified by cell sorting from spleen and lymph nodes as previously described (Ivanov et al., 2006) using the Aria II (BD). Briefly, red blood cells were cleared from organ cell suspensions using ACK lysis buffer (Lonza). The resulting leukocytes were depleted of B220$^+$ and CD8$^+$ cells by magnetic-activated cell sorting (MACS, Miltenyi) according to the product protocol. The negative fraction was cell surface stained using antibodies specific for CD4, CD25, CD44, and CD62L, and CD4$^+$CD25$^-$CD62L$^+$CD44$^{lo/-}$ naive CD4$^+$ T cells were isolated by cell sorting using the Aria II to greater than 98% purity based on post-sort analysis. Naive CD4+ T cells were cultured in 48-, 24-, or 12- well plates coated with an anti-hamster IgG secondary antibody (MP Biomedicals), in complete IMDM media (containing 10% FCS) containing soluble anti-CD3ε (0.25 μg/ml) and anti-CD28 (1 μg/ml) for TCR stimulation. Cultures were supplemented as follows, or as indicated in figures: with anti-IL4 (2 μg/ml) and anti-IFNγ (2 μg/ml) for Th0 conditions and additionally with IL-6 (20ng/ml; eBioscience) and TGFβ (0.3ng/mL; PeproTech) for Th17 conditions; or 5ng/mL TGFβ for iTreg conditions. For Th1 differentiation, IL-4 and anti-IFNγ (2 μg/ml) were added; for Th2 differentiation, IL-12 and anti-IL-4 (2 μg/ml) were added; cytokine concentrations as labeled in figures. Unless otherwise indicated antibodies were purchased from eBioscience.

Antibodies, Surface, and Intracellular Staining

For analysis of cytokine production, cells were incubated for 4-5 hr with phorbol 12-myristate 13-acetate (50ng/ml; Sigma), ionomycin (500ng/ml; Sigma), and GolgiStop (BD) at 37°C in a tissue culture incubator. Surface cell staining was carried out with fluorescence-labeled antibodies in PBS containing 0.5% BSA and 2mM EDTA at 4°C for 20 min. For live cell analysis or sorting, cells were washed once in staining buffer and resuspended in 200ng/ml of DAPI in staining buffer to exclude dead cells. For intracellular staining, cells were first stained with the fixable Aqua dead cell exclusion kit (Invitrogen), washed twice with PBS, and resuspended in Fixation-Permeabilization solution (Cytofix/Cytoperm kit; BD Biosciences or eBioscience) and intracellular staining was carried out according to the manufacturer's protocol. All fluorescence-labeled antibodies were purchased from eBioscience. An LSR II (BD Biosciences) was used for flow cytometric acquisition, followed by analysis with FlowJo software (Tree Star). All analysis plots are gated to exclude dead cells.

Chromatin Immunoprecipitation

TF ChIP-Seq was performed in biological duplicate as described (Johnson et al., 2007) with the following modifications. For each ChIP, 20-80 million cells were cross-linked with paraformaldehyde; chromatin was isolated and fragmented with a Vibra-Cell VCX130PB (Sonics & Materials). Following immunoprecipitation, the protein-DNA crosslinks were reversed and DNA was purified. DNA from control samples was prepared similarly but without immunoprecipitation. Histone ChIP of native chromatin was performed as previously described (Kirigin et al., 2012). Sequencing libraries were made from the resulting DNA fragments for both ChIP and controls as described (Reddy et al., 2012). The ChIP-seq libraries were sequenced with single-end 36 bp reads on an Illumina GAIIx or single-end 50 bp reads on an Illumina HiSeq 2000.

Commercial antibodies used for ChIP for each protein were as follows: IRF4 (IRF-4 M-17; Santa Cruz Biotech, sc-6059), BATF (BATF; Santa Cruz Biotech, sc-100974), STAT3 (Stat3 C-20; Santa Cruz Biotech, sc-482), p300 (p300 C-20; Santa Cruz Biotech, sc-585), Maf (Bethyl Laboratories, A300-613A), FOSL2 (Fra-2 Q-20; Santa Cruz Biotech, sc-604), HIF1α (Novus, NB100-105), ETV6 (TEL; Santa Cruz Biotech, sc-8546), JMJD6 (abcam, ab64575), NRF2 (H-300, Santa Cruz Biotech, sc-13032), H3K4me2 (Millipore, 07-030), and H3K4me3 (Millipore, 05-745R). The anti-RORγ rabbit polyclonal antibody was raised against amino acids 79-301 (Covance) and affinity purified antibody was isolated from serum using the same immunogen. The JMJD3 affinity purified antibody was kindly provided by G. Natoli (IFOM-IEO, Italy). The specificity of each transcription factor antibody was validated by immunoblot or conventional ChIP assay comparing wild-type to factor-deficient (or knock-down) Th17 subset polarized cells. In addition, ChIP-Seq was performed in knock-out cells for the core TFs to provide an additional negative control for each ChIP-seq.

FAIRE-Seq

FAIRE was performed as previously described (Simon et al., 2012). FAIRE reads were mapped using Bowtie (-k 1 –best) (Langmead et al., 2009) on the Galaxy platform (Goecks et al., 2010). For visualization of FAIRE signal around pCRMs, normalized alignment files were prepared using HOMER (Heinz et al., 2010), and heatmaps were made using SEQMINER (Ye et al., 2011).

Coimmunoprecipitations

Naive CD4 positive T cells were sorted and cultured under Th17 polarizing conditions for 48h prior to assay. Whole cells lysates were prepared with high salt buffer (10mM Tris, 420mM NaCl, 0.5% NP40, 1mM EDTA), sonicated, spun to remove insoluble particles, and diluted to a final concentration of 150mM NaCl for co-immunoprecipitation. Endogenous IRF4 was immunoprecipitated using anti-IRF4 antibody in the presence or absence of 50ug/ml ethidium bromide. Co-IP pulled-downs were resolved by SDS electrophoresis and anti-BATF (Santa Cruz Biotech) and STAT3 (Cell Signaling) antibodies were used for Western blot detection.

Luciferase Assay

pCRM activity was assessed using luciferase reporter assays by cloning the ChIP-defined genomic region (average of approx. 750bp) upstream of a minimal promoter driving a luciferase gene (pGL4.23[luc2/minP]; Promega). Importantly, pCRMs were selected in a non-biased manner based on ranked average binding ChIP p-values for occupying TFs. Genomic coordinates are available in Table S5.

pCRM activity was assessed using luciferase reporter assays by cloning the ChIP-defined genomic region upstream of a minimal promoter driving a luciferase gene (pGL4.23[luc2/minP]; Promega). Naive $CD4^+$ T cells were sorted and cultured under Th2, or Th17 polarizing conditions for 48h prior to being harvested for electroporation. Briefly, 5 million cells were pre-incubated with 10 μg of pCRM-pGL4minP or empty pGL4minP construct and 2 μg of renilla luciferase plasmid in 500 μl of RPMI on ice. Cells were electroporated using a BioRad Electroporator at 300V and 750 μF. After 10 min of recovery on ice, cells were placed into pre-warmed polarizing culture medium under TCR and cytokine stimulation conditions (Th2 or Th17). 24h post electroporation, cells were collected and luciferase assays were performed using the Dual Luciferase Reagents (Promega). Firefly luciferase values were normalized to renilla luciferase values for each sample and expressed as fold change over empty pGL4-minP. pGL4-minP harboring regions from the *Il17a* locus: Il17a-5 (a known enhancer) (Wang et al., 2012) and Il17a-19 (a non-TF occupied conserved region 19kb upstream of TSS) served as positive and negative controls, respectively.

Preparation of RNA-Seq Libraries

mRNA was prepared from total RNA by poly-A selection and cDNA synthesis was carried out as described (Mortazavi et al., 2008). The resulting dsDNA was prepared for sequencing by ligation of Illumina sequencing adapters, selection of 225 bp fragments from a 2% agarose SizeSelect E-Gel (Invitrogen), and amplification with 15 cycles of PCR using Illumina paired-end primers. Alternatively, some libraries were made using the Nextera tagmentation protocol described (Gertz et al., 2012). The RNA-seq libraries were sequenced with single-end 36 bp reads on an Illumina GAIIx or single-end 50 bp reads on an Illumina HiSeq 2000. Biological duplicates were carried out for each experiment. Sequence reads were mapped to the *Mus musculus* genome (version mm9) with Bowtie (version 0.12.7) and with the following settings: -k 1–best. The–phred33-quals or–phred64-quals parameter was set as needed depending on the format of the input fastq file. Anywhere between 8.5M and 78.7M reads aligned per library. Read counts for annotated genomic features were computed using the htseq-count script from the HTSeq (version 0.5.3p3) software suite with parameters:–stranded = no–mode = union.

siRNA Knockdowns

For knockdown of network genes in T cell polarization cultures, naive C57/Bl6 CD4 T cells were sort purified and cultured for 16-18h in RPMI/Th0 conditions. 2 million stimulated cells were transfected with 300pmol of control siRNAs for Ccr6, Rorc, and a non-Targeting pool (pool #2; SMARTpool siRNA; Dharmacon), in addition to SMARTpool siRNAs for network target genes (Dharmacon). Transfections were performed using the Amaxa Mouse T cell Nucleofector Kit with the X-001 program (Amaxa) according the manufacture's protocol. After a 4h recovery at 37°C, cells were stimulated in Th17 conditions in RPMI media. RNA was prepared from cells collected at 24h post polarization to assess knockdown efficiency and for RNA-Seq. Flow cytometric analysis for IL-17A and Foxp3 24h post polarization; viability was assessed by Aqua exclusion (Invitrogen) and cell counts by Accucount particles.

Collection of GWAS and SNP Data for Network Validation

Disease-associated SNPs compiled from the National Human Genome Research GWAS Catalog (available at http://www.genome.gov/gwastudies; accessed Feb 29, 2012). For each condition, gene lists were produced by selecting catalog-annotated human genes within 100 kb of associated SNPs. In cases where a SNP falls between two loci, the closest gene was chosen for association. Gene lists were used with no regard to human-mouse synteny.

Retroviral Gene Transfer

Retroviral constructs were generated by subcloning of the cDNA of interest into MSCV-Thy1.1 5′ of the internal ribosomal entry site, permitting the bicistronic expression of candidate TFs and cell surface Thy1.1. Retrovirus was generated by transfection of retroviral constructs into the PlatE producer cell line (Morita et al., 2000); viral supernatants were used at 48h post transfection. FACS sorted naive CD4 T cells were stimulated under Th0 conditions for 20-24h prior to retroviral transduction. For gene transfer, cells were spin transduced for 2 hr at 2500rpm with viral supernatants in the presence of 6.7 ug/mL of polybrene (hexadimethrine bromide, Sigma), and media was replaced with T cell polarization media for differentiation to Th17 and control Th1. Cells were harvested after 48h (Th17) and 5 days (Th1) for flow cyotmetric analysis of cytokine production.

EAE Induction

For induction of EAE, mice were immunized subcutaneously on day 0 with 200 μg/mouse MOG 35-55 peptide (UCLA peptide synthesis facility), emulsified in CFA (CFA supplemented with 2 mg/ml Mycobacterium tuberculosis), and injected intravenously on days 0 and 2 with 200 ng/mouse of pertussis toxin (Sigma Aldrich). The following scoring system used was 0—no disease, 1—limp tail, 2—weak/partially paralyzed hind legs, 3—completely paralyzed hind legs, 4—complete hind and partial front leg paralysis, 5—complete paralysis/death. Mice with disease levels 4 and 5 were considered moribund and were euthanized.

Isolation of Mononuclear Cells from Spinal Cords

Before spinal cord (SC) dissection, mice were perfused with 30 ml of cold Ca^{2+}/Mg^{2+}-free PBS. The spinal columns were dissected, cut open, and intact SCs separated carefully from the vertebrae. The SCs were cut into several small pieces and placed in 2 ml digestion solution containing 10 mg/ml Collagenase D (Roche) in PBS with 5% FCS. Digestion was performed for 30 min at 37°C. Digestion was terminated by the addition of EDTA to a final concentration of 12.5 mM and an additional 5 min incubation. The resulting digested tissue was passed through a 70 um cell screen. The cells were washed once in PBS, placed in 10 ml of 38% Percoll solution, and pelleted for 30 min at 2000 rpm with no brake. Cells pellets were washed once in PBS, re-suspended in FACS buffer or T cell medium and stimulated for assessment of cytokine production and Foxp3 expression as described above.

SUPPLEMENTAL REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bonneau, R., Facciotti, M.T., Reiss, D.J., Schmid, A.K., Pan, M., Kaur, A., Thorsson, V., Shannon, P., Johnson, M.H., Bare, J.C., et al. (2007). A predictive model for transcriptional control of physiology in a free living cell. Cell *131*, 1354–1365.

Bonneau, R., Reiss, D.J., Shannon, P., Facciotti, M., Hood, L., Baliga, N.S., and Thorsson, V. (2006). The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. Genome Biol. *7*, R36.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Carro, M.S., Lim, W.K., Alvarez, M.J., Bollo, R.J., Zhao, X., Snyder, E.Y., Sulman, E.P., Anne, S.L., Doetsch, F., Colman, H., et al. (2010). The transcriptional network for mesenchymal transformation of brain tumours. Nature *463*, 318–325.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Eberl, G., Marmon, S., Sunshine, M.J., Rennert, P.D., Choi, Y., and Littman, D.R. (2004). An essential function for the nuclear receptor RORgamma(t) in the generation of fetal lymphoid tissue inducer cells. Nat. Immunol. *5*, 64–73.

Faith, J.J., Hayete, B., Thaden, J.T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J.J., and Gardner, T.S. (2007). Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol. *5*, e8.

Fisher, R.A. (1925). Statistical Methods for Research Workers (Edinburgh: Oliver and Boyd).

Gertz, J., Varley, K.E., Davis, N.S., Baas, B.J., Goryshin, I.Y., Vaidyanathan, R., Kuersten, S., and Myers, R.M. (2012). Transposase mediated construction of RNA-seq libraries. Genome Res. *22*, 134–141.

Gilchrist, M., Thorsson, V., Li, B., Rust, A.G., Korb, M., Roach, J.C., Kennedy, K., Hai, T., Bolouri, H., and Aderem, A. (2006). Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. Nature *441*, 173–178.

Goecks, J., Nekrutenko, A., and Taylor, J.; Galaxy Team (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. *11*, R86.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Heng, T.S., and Painter, M.W.; Immunological Genome Project Consortium (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol. *9*, 1091–1094.

Ivanov, I.I., McKenzie, B.S., Zhou, L., Tadokoro, C.E., Lepelley, A., Lafaille, J.J., Cua, D.J., and Littman, D.R. (2006). The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. Cell *126*, 1121–1133.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497–1502.

Karreth, F., Hoebertz, A., Scheuch, H., Eferl, R., and Wagner, E.F. (2004). The AP1 transcription factor Fra2 is required for efficient cartilage development. Development *131*, 5717–5725.

Kirigin, F.F., Lindstedt, K., Sellars, M., Ciofani, M., Low, S.L., Jones, L., Bell, F., Pauli, F., Bonneau, R., Myers, R.M., et al. (2012). Dynamic microRNA gene transcription and processing during T cell development. J. Immunol. *188*, 3257–3267.

Klein, U., Casola, S., Cattoretti, G., Shen, Q., Lia, M., Mo, T., Ludwig, T., Rajewsky, K., and Dalla-Favera, R. (2006). Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. Nat. Immunol. *7*, 773–782.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lee, C.K., Raz, R., Gimeno, R., Gertner, R., Wistinghausen, B., Takeshita, K., DePinho, R.A., and Levy, D.E. (2002). STAT3 is a negative regulator of granulopoiesis but is not required for G-CSF-dependent differentiation. Immunity *17*, 63–72.

Lefebvre, C., Rajbhandari, P., Alvarez, M.J., Bandaru, P., Lim, W.K., Sato, M., Wang, K., Sumazin, P., Kustagi, M., Bisikirska, B.C., et al. (2010). A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol. Syst. Biol. *6*, 377.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics *27*, 1696–1697.

Madar, A., Greenfield, A., Ostrer, H., Vanden-Eijnden, E., and Bonneau, R. (2009). The Inferelator 2.0: a scalable framework for reconstruction of dynamic regulatory network models. Conf. Proc. IEEE Eng. Med. Biol. Soc. *2009*, 5448–5451.

Madar, A., Greenfield, A., Vanden-Eijnden, E., and Bonneau, R. (2010). DREAM3: network inference using dynamic context likelihood of relatedness and the inferelator. PLoS ONE *5*, e9803.

Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., Kellis, M., Collins, J.J., and Stolovitzky, G.; The DREAM5 Consortium (2012a). Wisdom of crowds for robust gene network inference. Nat. Methods *9*, 796–804.

Marbach, D., Roy, S., Ay, F., Meyer, P.E., Candeias, R., Kahveci, T., Bristow, C.A., and Kellis, M. (2012b). Predictive regulatory models in Drosophila melanogaster by integrative inference of transcriptional networks. Genome Res. *22*, 1334–1349.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521–533.

Morita, S., Kojima, T., and Kitamura, T. (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther. *7*, 1063–1066.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods *5*, 621–628.

Ouyang, Z., Zhou, Q., and Wong, W.H. (2009). ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. Proc. Natl. Acad. Sci. USA *106*, 21521–21526.

Park, H., Li, Z., Yang, X.O., Chang, S.H., Nurieva, R., Wang, Y.H., Wang, Y., Hood, L., Zhu, Z., Tian, Q., and Dong, C. (2005). A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. Nat. Immunol. *6*, 1133–1141.

Prill, R.J., Marbach, D., Saez-Rodriguez, J., Sorger, P.K., Alexopoulos, L.G., Xue, X., Clarke, N.D., Altan-Bonnet, G., and Stolovitzky, G. (2010). Towards a rigorous assessment of systems biology models: the DREAM3 challenges. PLoS ONE *5*, e9202.

Reddy, T.E., Gertz, J., Pauli, F., Kucera, K.S., Varley, K.E., Newberry, K.M., Marinov, G.K., Mortazavi, A., Williams, B.A., Song, L., et al. (2012). Effects of sequence variation on differential allelic transcription factor occupancy and gene expression. Genome Res. *22*, 860–869.

Ryan, H.E., Poloni, M., McNulty, W., Elson, D., Gassmann, M., Arbeit, J.M., and Johnson, R.S. (2000). Hypoxia-inducible factor-1alpha is a positive factor in solid tumor growth. Cancer Res. *60*, 4010–4015.

Schraml, B.U., Hildner, K., Ise, W., Lee, W.L., Smith, W.A., Solomon, B., Sahota, G., Sim, J., Mukasa, R., Cemerski, S., et al. (2009). The AP-1 transcription factor Batf controls T(H)17 differentiation. Nature *460*, 405–409.

Simon, J.M., Giresi, P.G., Davis, I.J., and Lieb, J.D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. Nat. Protoc. *7*, 256–267.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Wang, X., Zhang, Y., Yang, X.O., Nurieva, R.I., Chang, S.H., Ojeda, S.S., Kang, H.S., Schluns, K.S., Gui, J., Jetten, A.M., and Dong, C. (2012). Transcription of Il17 and Il17f is controlled by conserved noncoding sequence 2. Immunity *36*, 23–31.

Wende, H., Lechner, S.G., Cheret, C., Bourane, S., Kolanczyk, M.E., Pattyn, A., Reuter, K., Munier, F.L., Carroll, P., Lewin, G.R., and Birchmeier, C. (2012). The transcription factor c-Maf controls touch receptor development and function. Science *335*, 1373–1376.

Ye, T., Krebs, A.R., Choukrallah, M.A., Keime, C., Plewniak, F., Davidson, I., and Tora, L. (2011). seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. *39*, e35.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhou, X., Sumazin, P., Rajbhandari, P., and Califano, A. (2010). A systems biology approach to transcription factor binding site prediction. PLoS ONE *5*, e9878.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. J. R. Statist. Soc. B *67*, 301–320.

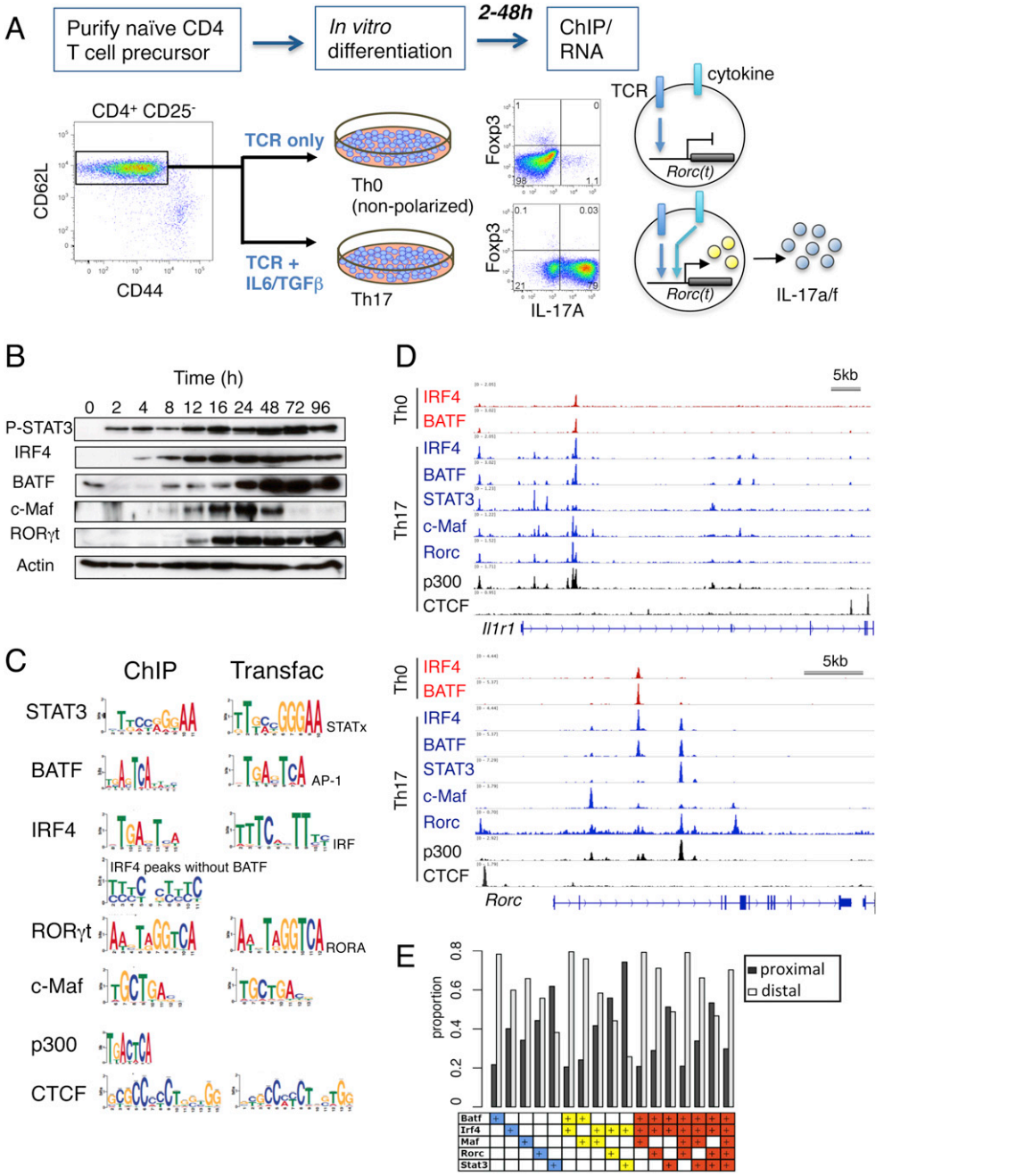


Figure S1. Genome-wide Co-Occupancy of Th17 Lineage TFs, Related to Figure 1

(A) In vitro model system for Th17 cell differentiation from naive CD4$^+$ T cell precursors and schematic for experimental protocol. Th0 cultures provide control cells that receive TCR activation in the absence of exogenous polarizing cytokines (IL-6 + TGFβ).

(B) Western blot analysis of Th17 polarization time series starting from FACS-purified naive CD4 T cells (time = 0).

(C) Recovery of cognate consensus motifs from TF-ChIP-Seq.

(D) High degree of co-occupancy among Th17 lineage TFs. ChIP-seq binding tracks are displayed for core TFs, CTCF, and p300 at selected Th17 loci in both non-polarized Th0 and Th17 conditions. Visualized using the Integrative Genomics Viewer (IGV; Broad Institute).

(E) High-order pCRMs are not correlated with proximity to TSS. Bar chart of proportion of proximal versus distal pCRMs with respect to increasing order of occupancy.

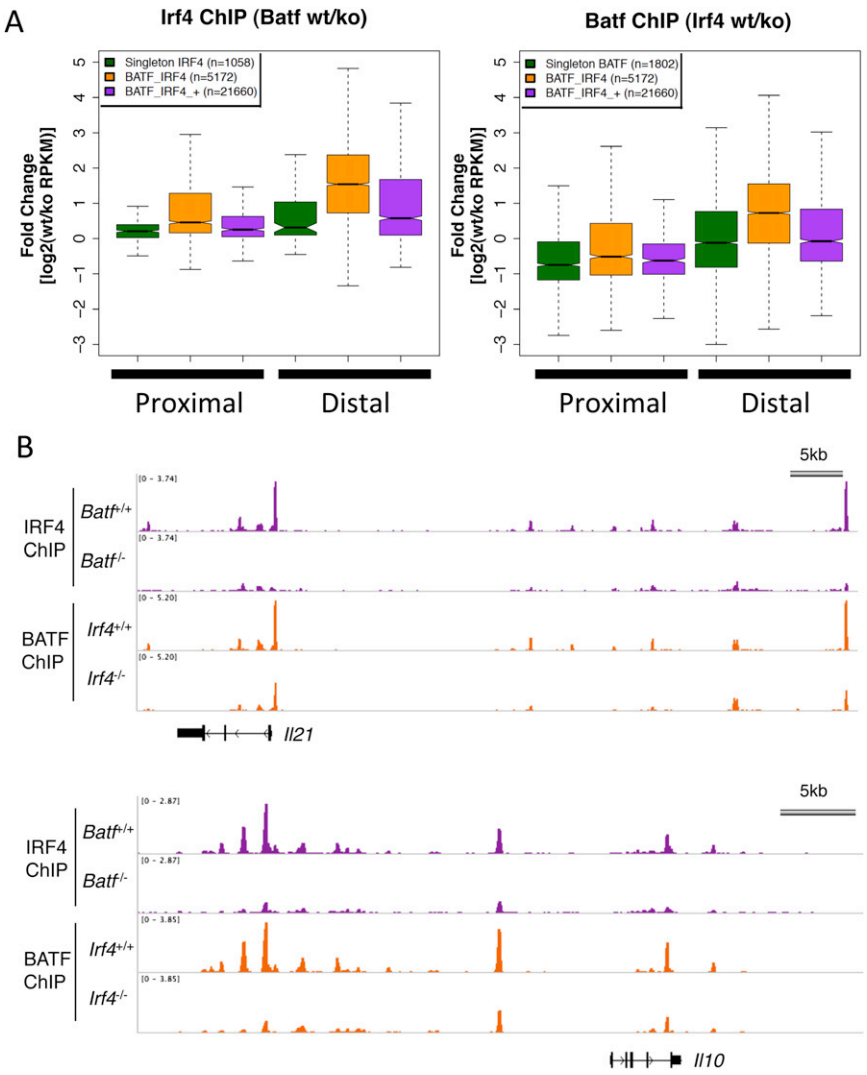


Figure S2. Cooperative Occupancy by BATF and IRF4, Related to Figure 2

(A) Genome-wide interdependence of IRF4 and BATF co-occupancy in Th0 cells. Box plots displaying the fold change in ChIP-seq reads for IRF4 in *Batf* wild-type (wt) versus knockout (ko) and for BATF in *Irf4* (wt/ko) for 48h Th0 cultured cells. Differences in ChIP-seq reads are assayed within relevant pCRM regions. Three sub-types of pCRMs were interrogated: BATF or IRF4 alone; BATF and IRF4 alone; and BATF, IRF4, plus additional TFs (+) as indicated by color-coding. Displayed is the data distribution: median (line), 25th to 75th percentile (box) +/− 1.5 Interquartile range (whiskers). To compute fold change in ChIP values, reads localized to a given pCRM were normalized by library size (i.e., reads per million; RPM) prior to calculations.

(B) Interdependent binding of IRF4 and BATF at selected loci in Th17 cells.

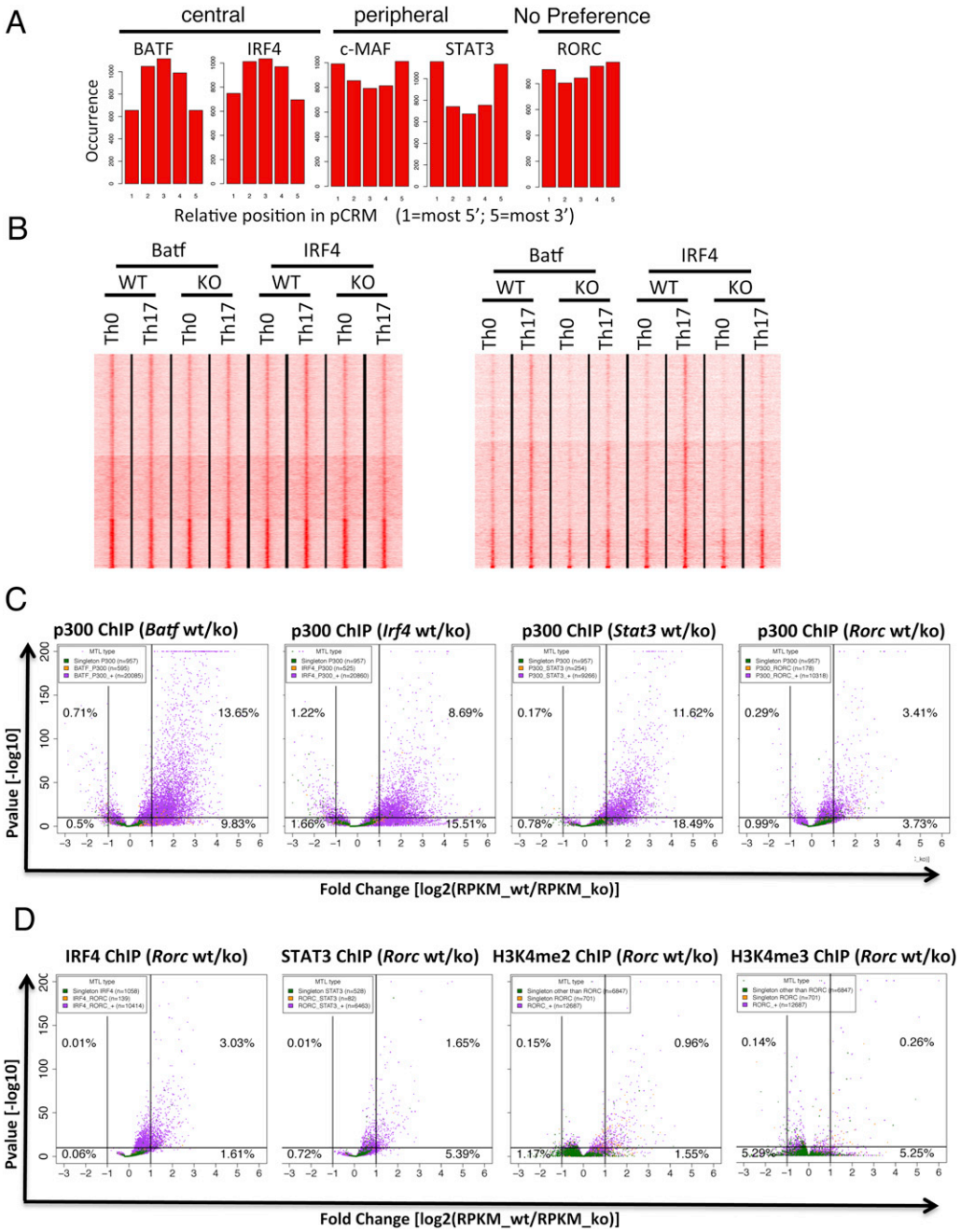


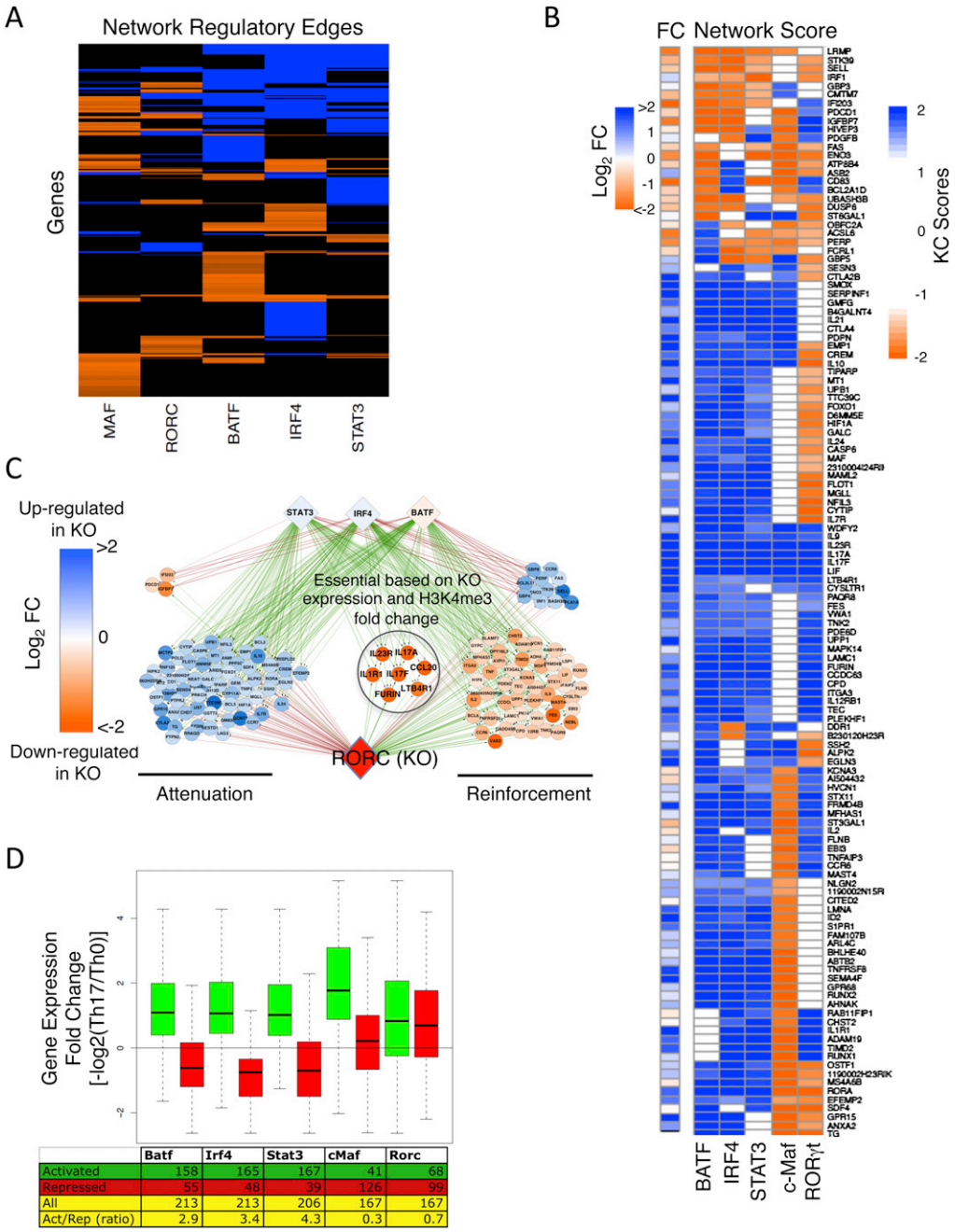
Figure S3. Genome-wide Requirement for Th17 TFs for Accessibility and TF Occupancy, Related to Figure 2

(A) Spatial correlations between Th17 TFs within pCRMs occupied by all five TFs. Bar chart plots the occurrence with which the summit of a given TFs occupies relative position 1 through 5 when the summits of all five are ordered from 5′ to 3′.

(B) IRF4 and BATF regulate chromatin accessibility at TCR-induced *cis* regions that are co-occupied by Th17 TFs. FAIRE signal at Th17 5TF+p300 pCRMs was compared between WT and *Irf4*$^{-/-}$ and *Batf*$^{-/-}$ Th0 and Th17 polarized T cells. pCRMs were divided according to their accessibility status in naive CD4^{+} T cells: constitutive pCRMs are accessible in naive cells (2,930 regions, left panel), while induced pCRMs are not (1,575 regions, right). Biological replicate samples were averaged, and normalized FAIRE reads were aligned around the median summit position of overlapping TF binding peaks, +/− 2,000bp.

(C) Limited requirement for RORγt for p300 occupancy as compared to IRF4, BATF, and STAT3. Differential occupancy of p300 in TF wild-type versus deficient 48h Th17 polarization cultures is displayed as scatter plots of fold change versus significance. Various pCRM subtypes are compared as indicated in the figure. Percentage of pCRMs with differential ChIP are indicated in plot.

(D) Limited requirement for RORγt for IRF4 and STAT3 occupancy and for presence of H3K4me2 and H3K4me3 modifications. Scatter plots as in (C).



	Batf	Irf4	Stat3	cMaf	Rorc
Activated	158	165	167	41	68
Repressed	55	48	39	126	99
All	213	213	206	167	167
Act/Rep (ratio)	2.9	3.4	4.3	0.3	0.7

Figure S4. Relationship between Core TF Regulation and Expression of Target Genes in the Th17 Network, Related to Figure 3

(A) Heat map summarizing genome-wide regulatory inputs for the core TF network displayed in Figure 3A. Orange represents repression and blue represents activation. Rows are target genes.

(B) Heat map of activation and repression inputs for highly regulated genes (4 or 5 inputs) by core TFs. Orange represents repression and blue represents activation. View is limited to genes with KC scores > 1.5 for a given TF-target regulatory interaction. Also displayed is the fold change (FC) in expression observed in Th17 relative to Th0 cells.

(C) Regulatory interactions shared by STAT3, IRF4, BATF, and RORγt (as in Figure 3C). Node color depicts the extent to which genes are differentially expressed in RORγt deficient Th17 cells when compared to wild-type Th17 cells. Different modes of RORγt regulation are as indicated.

(D) Effect of individual core TFs on target gene transcription. Positive and negative regulation by STAT3, IRF4, and BATF is well correlated with expression changes associated with Th17 differentiation. In contrast, the regulatory effect of RORγt is consistent with a modulatory role. Box plots display the fold change in expression of both TF activation (green) and repression (red) targets for Th17 relative to Th0 culture conditions. Displayed is the data distribution: median (line), 25th to 75th percentile (box) +/− 1.5 Interquartile range (whiskers). Genes co-regulated by either 4 or 5 of the TFs are considered in this analysis.

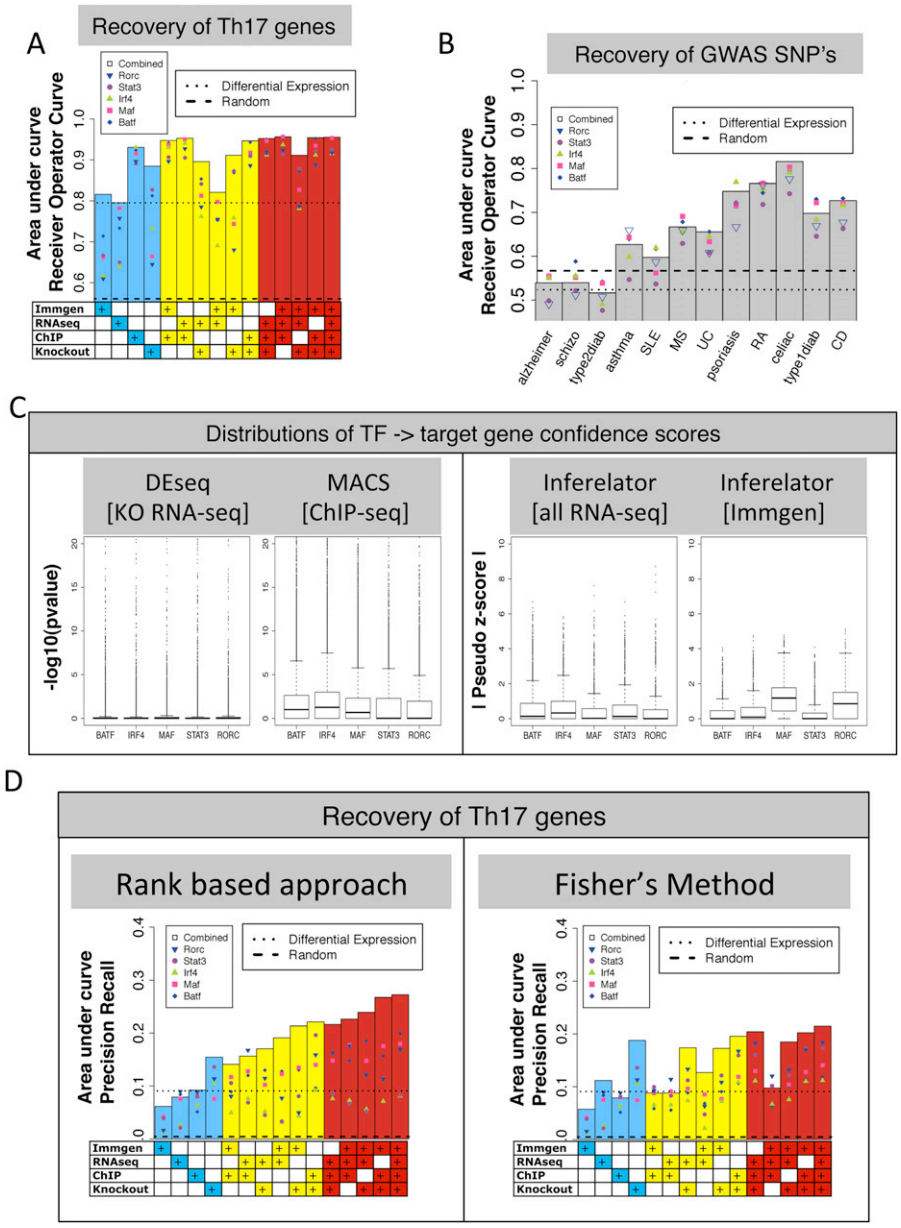


Figure S5. aucROC Performance and Comparison of Two Meta-Analysis Strategies: Rank-Based and Fisher's Method, Related to Figure 4

(A) Recovery of validated biologically relevant Th17 targets based on integration of regulatory models from multiple functionally relevant TFs and of multiple data types. The graph summarizes area under curve (auc) of receiver operator curve (ROC) plot results indicating the degree to which 74 literature-based Th17-relevant genes are enriched as top network predictions under different data combinations. Individual TFs (scatter plot) versus combined TFs (bar plots) are compared for each combination of data types. As a reference, differential expression in Th17 versus Th0 (dotted line) and random performance (dashed line; based on 200 simulations) are provided.

(B) The KCRI network selectively recovers genes linked to SNPs that are associated with Th17-implicated inflammatory disease in GWAS studies. Gray bars and scatter plots in each column correspond to the recovery of SNP-associated disease-relevant targets as top predictions within the ranked list of the KCRI network scores, using the aucROC analysis. Gene lists of disease-associated SNPs were compiled from the National Human Genome Research GWAS Catalog.

(C) Distributions of genome-wide TF - > target gene scores (-log10 p-values or absolute z-scores) for five TFs based on method [data-source] used.

(D) Area under curve of Precision Recall for the Rank based approach used in this work and Fisher's method for combining p-values (pseudo z-scores were transformed to p-values before combining assuming a normal distribution). Performance measured as aucPR values indicating enrichment of literature-curated Th17 genes in networks derived from all possible data combinations. Points indicate single TF predictions (e.g., *BATF → target*) and bars indicate TF sum predictions (i.e., *[BATF+IRF4+STAT3+MAF+RORγt]→target*). Dotted line, reference performance for targets prioritized by differential expression (Th17 versus Th0) and, dashed line, for random. Fisher's method achieves better performance in combining TFs - > target p-values over single data sets (e.g., compare blue bars for knock out data), but does not perform well when incorporating p-values from distinct data sources (compare yellow and red bars) due to inherent differences in p-value distributions derived from different data types and methods.

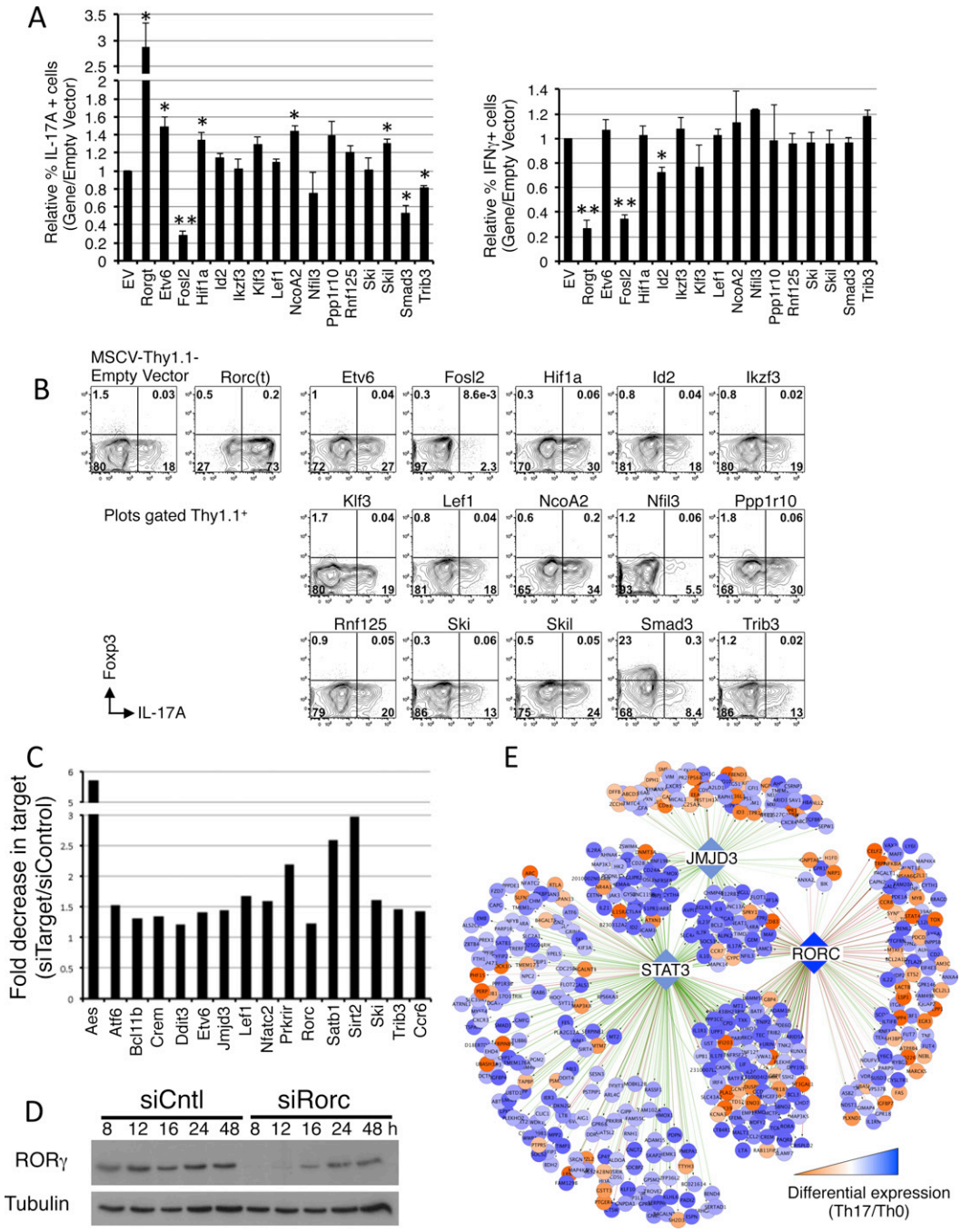


Figure S6. Gain- and Loss-of-Function Screens Identify Regulators of Th17 Specification, Related to Figure 5

(A) Overexpression screen of network TF candidates as putative Th17 subset regulators. Bar charts show the percent of IL-17A-producing or IFNγ-producing cells relative to the control empty vector after transduction of retroviruses encoding candidate factors and Th17 polarization for 48h, or Th1 polarization for 5 days, respectively. Results are mean ± s.e.m. for four biological replicates, each conducted in duplicate. Significance at *$p < 0.05$ and **$p < 0.01$ by t test.

(B) Representative flow cytometric analysis for IL-17A and Foxp3 expression in Th17-polarized cultures transduced with the indicated cDNAs in a retroviral vector also encoding the Thy1.1 reporter. Cells were gated for Thy1.1 expression to analyze proportions that were IL-17A$^+$.

(C) Knock-down efficiency of target mRNAs in the siRNA screen. Data represent reads per kilobase million (RPKM) expression values for the target TF in siRNA knock-down Th17 cultures relative to a non-targeting control. Analysis is at 24h post Th17 differentiation.

(D) Western-blot analysis of RORγt protein levels for *Rorc* knock-down at indicated times post initiation of Th17 polarization.

(E) JMJD3 regulates the expression of many RORγt and STAT3 targets in Th17 cells. Network representation is as in Figure 3. Due to space constraints, the display is limited to genes that have differential expression in Th17 relative to Th0 cells (z-score > 2.5, < −2.5 based on statistical analysis of microarray for 8 independent experiments).

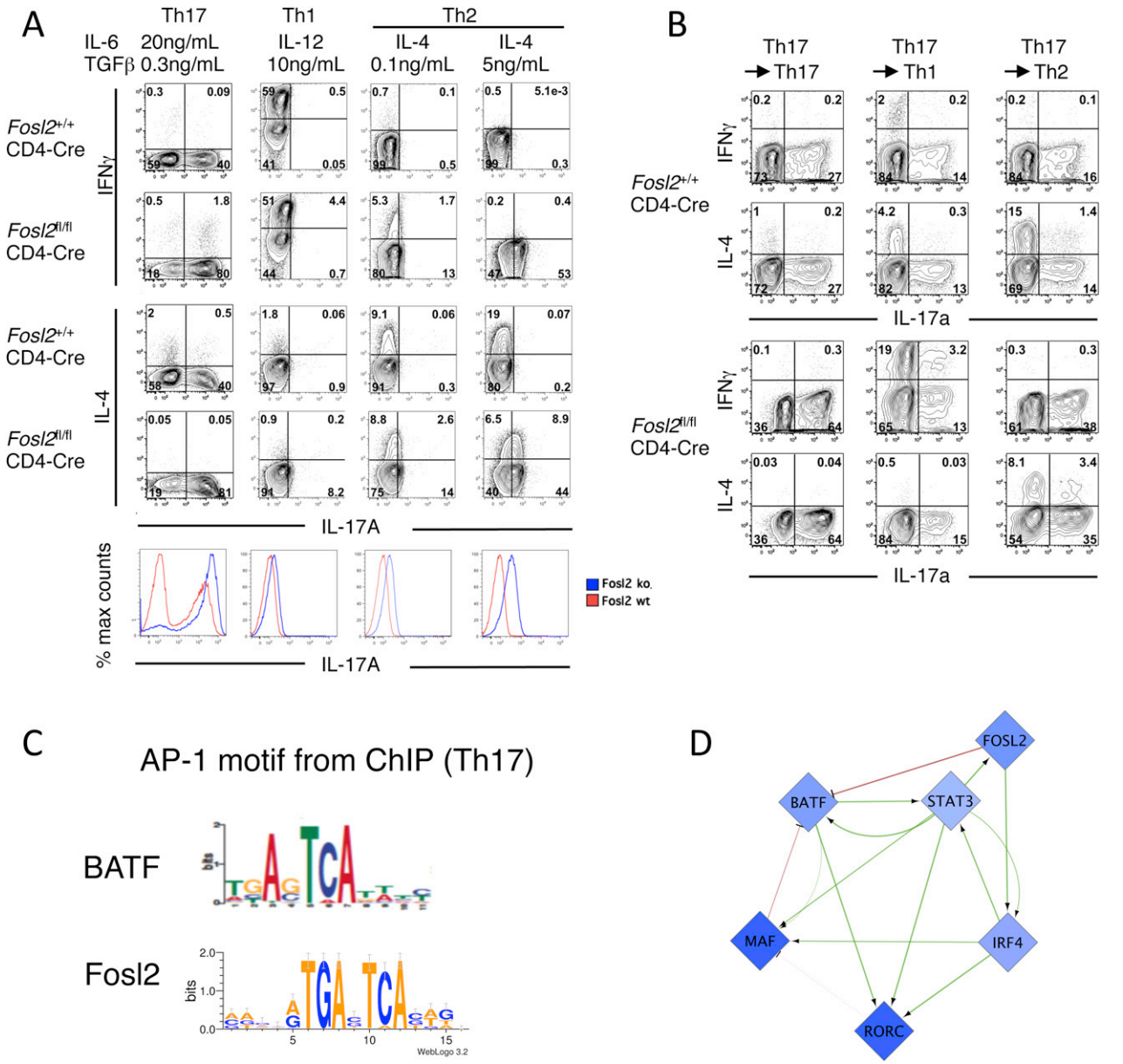


Figure S7. Fosl2 Restricts the Plasticity of Th17 Subset Cells, Related to Figure 6

(A) Dysregulated cytokine production in the absence of Fosl2. Fosl2 wild-type and deficient naive CD4$^+$ T cells were polarized under Th1, Th2, and Th17 conditions for 6 days. Flow cytometric analysis was then performed for IL-17A, IL-4, and IFNγ.

(B) Fosl2 restricts IFNγ production among IL-17A-producing cells. Fosl2 wild-type and deficient naive CD4$^+$ T cells were polarized under Th17 conditions (20ng/mL IL-6 and 0.3ng/mL TGFβ + blocking antibodies for IFNγ and IL-4) for three days. Thereafter, the media was replaced with cytokines for either (a) Th17-; (b) Th1- (10ng/mL IL-12); or (c) Th2- (2ng/mL IL-4) promoting conditions for an additional three days prior to analysis.

(C) De novo motif analysis of high confidence binding regions for BATF and Fosl2 ChIP-seq experiments of Th17 cells showing that the AP-1 consensus motif is recovered in both instances.

(D) Regulation of core Th17 TFs by Fosl2. Edges represent integration of data from ChIP-seq and KO RNA-seq differential expression; line weight is relative to network score; FDR < 5%. Nodes are colored to indicate the differential expression in Th17 relative to Th0 (blue = upregulated, orange = downregulated in Th17 cells).