nature immunology

A multiply redundant genetic switch 'locks in' the transcriptional signature of regulatory T cells

Wenxian Fu[1], Ayla Ergun[1,2,8], Ting Lu[2,3,8], Jonathan A Hill[1,7], Sokol Haxhinasto[1,7], Marlys S Fassett[1], Roi Gazit[4], Stanley Adoro[5], Laurie Glimcher[5], Susan Chan[6], Philippe Kastner[6], Derrick Rossi[4], James J Collins[2], Diane Mathis[1] & Christophe Benoist[1]

The transcription factor Foxp3 participates dominantly in the specification and function of Foxp3$^+$CD4$^+$ regulatory T cells (T_{reg} cells) but is neither strictly necessary nor sufficient to determine the characteristic T_{reg} cell signature. Here we used computational network inference and experimental testing to assess the contribution of other transcription factors to this. Enforced expression of Helios or Xbp1 elicited distinct signatures, but Eos, IRF4, Satb1, Lef1 and GATA-1 elicited exactly the same outcome, acting in synergy with Foxp3 to activate expression of most of the T_{reg} cell signature, including key transcription factors, and enhancing occupancy by Foxp3 at its genomic targets. Conversely, the T_{reg} cell signature was robust after inactivation of any single cofactor. A redundant genetic switch thus 'locked in' the T_{reg} cell phenotype, a model that would account for several aspects of T_{reg} cell physiology, differentiation and stability.

Regulatory T cells (T_{reg} cells) have a key role in immunological homeostasis, control autoimmune deviation, prevent runaway responses to microbes or allergens and regulate certain nonimmunological functions[1,2]. Most T_{reg} cells differentiate in the thymus as a rescue pathway for cells that express a self-reactive T cell antigen receptor (TCR)[3], but some also differentiate in peripheral organs in response to chronic challenges such as commensal bacteria[4]. Phenotypic stability is an important consideration for T_{reg} cells, as the self-reactivity of their TCRs makes it important for their suppressive phenotype to be stable, lest they convert into aggressive effector cells. Support for the idea of T_{reg} cell instability, and for the proposal that T_{reg} cells turned into aggressive effector cells by the loss of the Forkhead-family transcription factor Foxp3 have a role in autoimmune diseases, stemmed from results obtained by transfer of T_{reg} cells into alymphoid hosts[5–7] and from lineage-tracing experiments that relied on continuously active *Foxp3*-driven transgenes encoding Cre recombinase[8]. However, those results have been largely refuted by the observation that T_{reg} cells transferred into normal hosts are stable for long periods of time and by lineage-tracing experiments done in pulse-chase mode with a tamoxifen-controlled Cre recombinase system[9]. Thus, with the exception of effector cells that transiently express Foxp3 after activation[10], the phenotype of committed T_{reg} cells seems very stable over time[9].

T_{reg} cell function is underwritten by a canonical 'T_{reg} cell signature', a set of transcripts that are over- or underexpressed in T_{reg} cells relative to their expression in the CD4$^+$ conventional T cell (T_{conv} cell) counterparts of the T_{reg} cells[11,12]. This signature is established very early during T_{reg} cell differentiation[11], and the genes encode proteins with a range of cellular locations and several molecular mediators of T_{reg} cell action[13]. Foxp3 is essential for the specification and maintenance of T_{reg} cells, as shown by the lethal lymphoproliferation and multiorgan autoimmunity that develop in its absence in scurfy mutant mice or in patients with the immunological disorder IPEX syndrome (immune dysregulation, polyendocrinopathy, enteropathy, X-linked)[14], and plays an important part in determining the T_{reg} cell signature[11,15,16]. Foxp3 was initially considered the 'master regulator' of T_{reg} cells, but a more nuanced view has since emerged. Cells with many T_{reg} cell characteristics, including a transcriptionally active *Foxp3* locus (T_{reg} cell–like cells), can differentiate in the absence of Foxp3, albeit in lower numbers and with less stability[17,18], and perhaps some patients with IPEX[19]. A segment of the T_{reg} cell signature can also be elicited in transforming growth factor-β (TGF-β)-induced T_{reg} cells derived from CD4$^+$ T cells of scurfy mice[11]. Conversely, the transduction of *FOXP3* or its induction by TGF-β are not sufficient to elicit the full T_{reg} cell signature[11,20].

Many other transcription factors have been reported to interact with Foxp3 and to promote T_{reg} cell function. These include factors from a variety of families, and physical or functional interactions have been demonstrated for Runx1, NFAT, Eos, phosphorylated STAT3,

[1]Division of Immunology, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA. [2]Howard Hughes Medical Institute, Department of Biomedical Engineering, and Center for BioDynamics, Boston University, Boston, Massachusetts, USA. [3]Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA. [4]Immune Disease Institute, Program in Cellular and Molecular Medicine, Children's Hospital Boston, Boston, Massachusetts, USA. [5]Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, USA. [6]Institut de Génétique et de Biologie Moléculaire et Cellulaire, IIIkirch, France. [7]Present addresses: Tempero Pharmaceuticals, Cambridge, Massachusetts, USA (J.A.H.), and Boehringer Ingelheim Pharmaceuticals, Ridgefield, Connecticut, USA (S.H.). [8]These authors contributed equally to this work. Correspondence should be addressed to D.M. (cbdm@hms.harvard.edu) or C.B. (cbdm@hms.harvard.edu).

Received 13 February; accepted 13 August; published online 9 September 2012; doi:10.1038/ni.2420

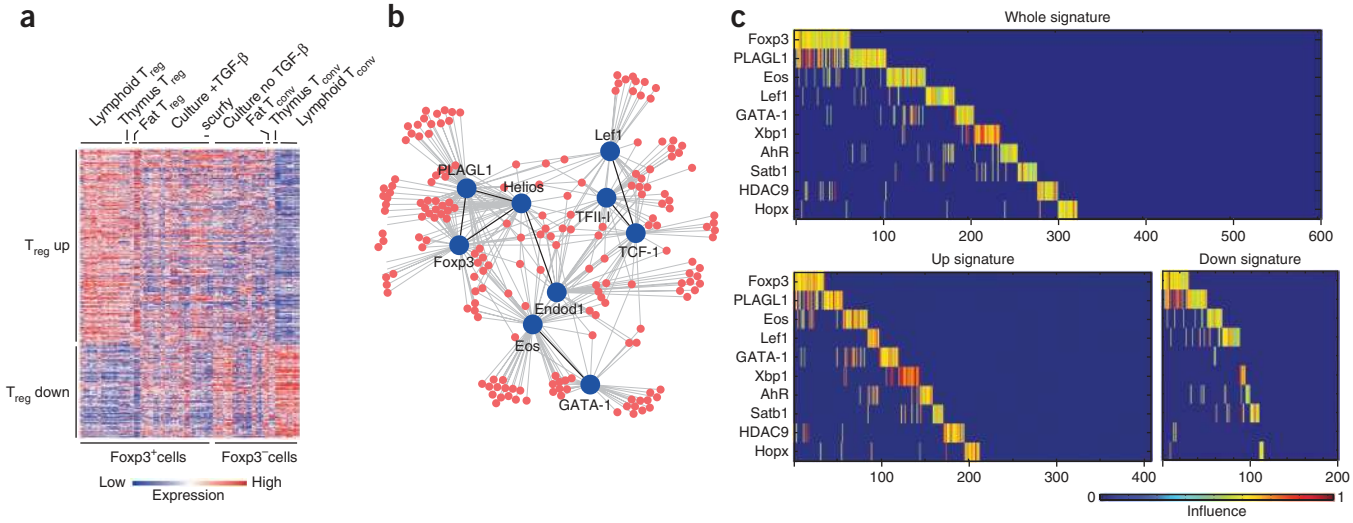

Figure 1 Computational prediction of transcription factors in control of the T_{reg} cell signature. (**a**) Expression profiles of genes (rows) from matched pairs of Foxp3$^+$ and Foxp3$^-$ cells (below plot) and TGF-β-treated cultures of CD4$^+$ T cells from Foxp3-null scurfy mice (scurfy; columns (**Supplementary Table 1**)); these results were used in the computational reconstruction in **b**. (**b**) Transcription factors (blue) most connected to genes of the T_{reg} cell signature (red), as predicted by the CLR algorithm. (**c**) Result of mathematical optimization of the CLR scores in **a**, with combinations of transcription factors selected to maximize the portion of the T_{reg} cell signature accounted for. Colors indicate intensity of influence, from blue (background; no effect) to green, yellow, red (increasing effect). Up, upregulation; Down, downregulation. Data are representative of three experiments.

IRF4, T-bet, GATA-3, RORγt, RORα, Foxo1 and Foxo3, Satb1 and HIF-1α[21–31]. Several of these are important for full T_{reg} cell function. In addition, different T_{reg} cell subphenotypes control various facets of effector T cells, and these are themselves dependent on distinct transcription factors such as T-bet, IRF4 or STAT3 (refs. 24,27,30).

How the contributions of those various cofactors' activities are orchestrated is unknown. A plausible hypothesis is that each cofactor might 'condition', alone or in combination with Foxp3, a segment of the genes of the T_{reg} cell signature, and the full signature and functional activity would thus result from the cumulative effects of these transcription factors. To test this hypothesis, we first used a computational approach to 'reverse engineer' the transcriptional regulatory network of T_{reg} cells, a successful strategy in simpler regulatory systems[32]. We then tested the computational predictions in loss- and gain-of-function experiments. Our results led to a rather different perspective in which the T_{reg} cell phenotype is controlled by a highly redundant 'genetic switch'.

RESULTS

Bioinformatics prediction of T_{reg} cell transcriptional control

Transcriptional regulation is governed by extensive and interconnected networks of regulator proteins and their target genes. This complexity can be tackled by computational methods that start from a large number of gene-expression data sets and reconstruct plausible regulatory models, then infer and rank potential connections between target genes and a set of putative regulators[33,34]. Such algorithms, which are typically based on multiple regression or related approaches, analyze the pairwise variation between regulator(s) and their target(s) across a large number of related data sets in response to a range of perturbations (differentiation or genetic or chemical perturbations). Here, we used 129 gene-expression profiles previously generated on a microarray platform with samples from the following various CD4$^+$ T cells: primary T_{reg} cells and T_{conv} cells from various anatomical locations and of different surface phenotypes; TGF-β-induced Foxp3$^+$ cells; cells from mutant mice (with deficiency in RARα or Foxp3); and cells transfected to express the kinase Akt or various transcription factors (**Supplementary Table 1**). We selected as potential regulators 2,021 transcription-control factors from an annotation of the Gene Ontology project (conventional transcription factors as well as chromatin modifiers) and selected 603 target genes that compose the canonical T_{reg} cell signature[11] (407 and 196 genes over- and underexpressed, respectively, in T_{reg} cells; **Fig. 1a**). We used the context likelihood of relatedness (CLR) algorithm[35], a relevance network-reconstruction method that operates by combining the relative strength of coexpression of a regulator and potential targets (regulators with the highest scores, **Table 1** and **Fig. 1b**; results, **Supplementary Table 2**). The top predicted regulators included Foxp3 and other factors associated with T_{reg} cell function, such as Eos (encoded by *Ikzf4*) and Helios (encoded by *Ikzf2*)[23,36], but they also included some additional transcription factors not discovered before with T_{reg} cells, such as Lef1 and GATA-1. Genes encoding some of the predicted regulators themselves had different expression in T_{reg} cells versus T_{conv} cells (*Ikzf2*, *Ikzf4* and *Lef1*), whereas others had only modestly different expression (*Gata1*).

Many target genes were predicted to be influenced by several transcription factors (**Fig. 1b**), which made it difficult to infer the regulators of the T_{reg} cell signature. In keeping with our hypothesis of additive transcriptional control by a panel of transcription factors, we started from those predicted regulators and formulated an optimization process with the ILOG Cplex mathematical programming solver to identify a set of transcription factors that would, in combination, account for the greatest fraction of the T_{reg} cell signature. With this model, we identified ten transcription factors that could account for 330 of the 603 genes of the T_{reg} cell signature (**Fig. 1c**). Foxp3 led the list, covering 10.8% of the signature, which was lower than but in the same range as estimates obtained in transduction experiments[11]. Most of the transcription factors were predicted to be both stimulatory and repressive, depending on the target, although some seemed to be only activating (GATA-1 and HDAC9).

Loss-of-function confirmation of computational predictions

We undertook a set of complementary gain- and loss- of function experiments to determine whether the computational predictions had

Table 1 Top regulators and their strongest connections based on CLR inference

Target gene	Expression in T_{reg} cells	Regulator								
		Foxp3	TFII-I	PLAGL1	TCF-I	Endod1	Helios	Eos	GATA-1	Lef1
Mfsd6	Down	**6.88**	3.28	**4.14**	2.48	2.39	3.49	**5.61**	3.17	0.47
Lypd6b	Down	3.96	**6.90**	3.20	**5.85**	4.03	2.79	3.59	1.87	3.30
2610019F03Rik	Down	2.41	**5.20**	2.79	**9.06**	3.59	2.36	**4.14**	0.95	**5.45**
Ahr	Up	2.37	1.32	2.01	1.90	**6.30**	3.06	2.58	1.32	3.06
Atp8b4	Down	**5.29**	2.79	**7.89**	1.54	1.28	2.68	3.70	–0.09	1.12
Casp7	Up	2.38	3.66	1.26	2.70	**6.65**	0.99	**4.27**	**5.47**	1.87
Cd83	Up	**4.42**	0.29	**6.87**	–0.54	2.57	**4.97**	2.17	0.39	1.56
Cyfip1	Up	1.99	**5.58**	1.61	2.78	**6.45**	2.83	1.84	1.81	2.81
Dennd2d	Down	0.61	**4.78**	0.86	**7.17**	1.37	0.46	2.53	0.77	2.41
Dst	Up	3.10	**6.44**	3.46	2.40	3.40	1.59	**4.10**	2.17	2.02
Themis	Down	**5.13**	1.38	**6.92**	1.43	1.36	**4.39**	2.18	0.96	2.98
Enc1	Down	**4.60**	0.43	2.42	2.13	**4.23**	**6.18**	3.78	2.27	3.71
Entpd1	Up	1.92	2.67	2.11	**6.51**	**5.05**	3.34	2.98	0.71	3.56
Gpr83	Up	**5.51**	2.41	4.03	2.82	1.52	3.97	**6.46**	3.80	0.65
Il1rl2	Down	3.92	0.55	3.91	1.68	3.78	**7.05**	2.95	–0.61	1.80
Il2ra	Up	1.82	3.48	1.03	3.36	2.80	1.42	**5.20**	**4.72**	0.39
Nlk	Down	0.28	**5.90**	1.54	3.29	1.31	0.52	1.57	0.87	3.38
Pde3b	Down	**8.31**	1.65	**6.46**	0.23	3.02	**5.58**	1.49	0.38	1.35
Lycat	Up	**4.95**	0.47	**5.96**	0.79	0.99	**6.02**	0.63	–0.62	2.15
Mctp2	Down	**6.52**	**5.76**	**6.26**	2.86	0.44	2.84	**5.01**	1.15	1.36
Niban	Up	4.01	2.95	3.60	2.74	**4.99**	**6.24**	1.70	1.75	2.62
Nrp1	Up	**5.81**	2.00	**5.15**	1.34	3.78	**4.45**	2.59	1.07	1.80
Rcn1	Up	0.17	1.36	0.20	1.16	1.19	–0.09	2.29	**8.22**	0.91
Slamf1	Up	3.72	2.32	**7.31**	0.72	0.82	2.49	–0.01	0.18	1.96
Socs2	Up	1.43	2.58	0.76	2.30	3.14	0.74	3.17	**6.31**	1.14
Stx11	Up	0.92	1.81	3.05	3.16	**4.90**	**4.29**	1.15	0.63	**6.13**

CLR scores of regulators with the most connections. Bolding indicates predicted interactions between regulators and genes of the T_{reg} cell signature with the highest scores kept in the network.

actual biological relevance. We chose a subset of transcription factors based on the availability of mice deficient in those factors and/or the availability of vectors for enforced expression of those factors. First, we analyzed the predicted targets of Foxp3 by comparing TGF-β-induced cultures of $CD4^+$ T cells from wild-type and Foxp3-deficient scurfy mice to identify those transcripts strictly dependent on Foxp3 (ref. 11). Predicted Foxp3 targets were skewed to the extremes of the distribution of this analysis, more so than the T_{reg} cell signature as a whole was (**Fig. 2a**), which indicated that the computational prediction of Foxp3 targets was valid.

We then analyzed the transcriptomes of T_{reg} cells in the following four available mouse strains deficient in a subset of those predicted cofactors: mice with fdf complete knockout of *Ikzf4* (which encodes Eos), which were viable and fertile with no noted abnormal autoimmune phenotype and had normal T_{reg} cell numbers and phenotypes (**Supplementary Fig. 1** and R.G. and D.R., data not shown), perhaps contrary to expectations[23]; mice with deletion of the *Gata1* promoter[37], in which T_{reg} cells and other T cells seemed normal (**Supplementary Fig. 1** and J.H., data not shown) but other known GATA-1 target genes were affected; mice with conditional knockout of *Xbp1* (ref. 38), in which the T_{reg} cell populations in lymphoid organs were again normal (**Supplementary Fig. 1** and S.A. and L.G., data not shown); and mice with knockout of *Helios*, in which T_{reg} cells seem normal[39]. We generated gene-expression profiles from purified splenic $CD4^+CD25^{hi}$ T_{reg} cells of

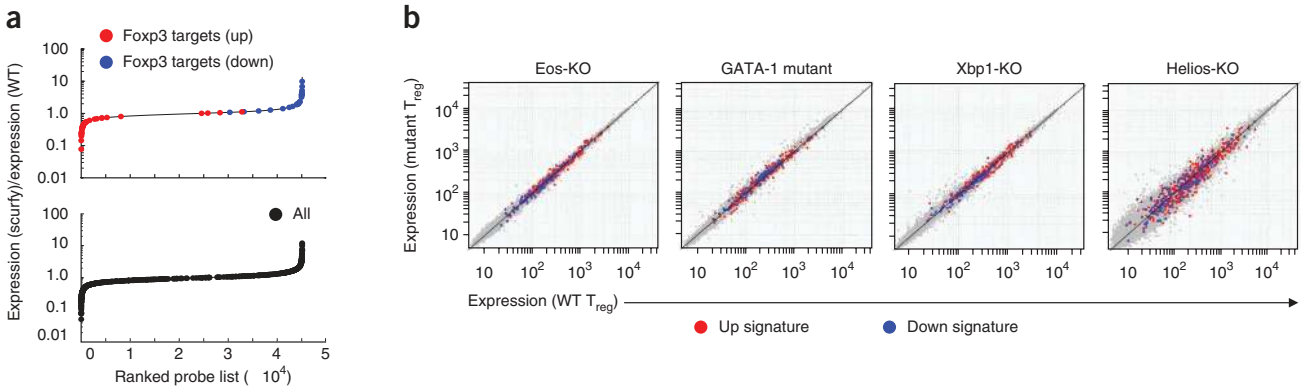

Figure 2 Validity of the predicted Foxp3 targets. (**a**) Ratio of the expression of CLR-predicted targets (top) or the entire T_{reg} cell signature (bottom) in TGF-β induced T cells from scurfy mice relative to that in wild-type (WT) mice[11], ranked by genome-wide change for all probes in scurfy mice versus wild-type mice. $P = 6.9 \times 10^{-6}$ (Kolmogorov-Smirnov test). (**b**) Profiling of the expression of CLR-predicted transcription factors in T_{reg} cells deficient in (-KO) Eos, Xbp1 or Helios or with mutant GATA-1 (Mutant T_{reg}), compared with that of T_{reg} cells from their wild-type littermates (WT T_{reg}). Data are representative of three experiments (**a**) or two experiments (**b**; average of duplicates).

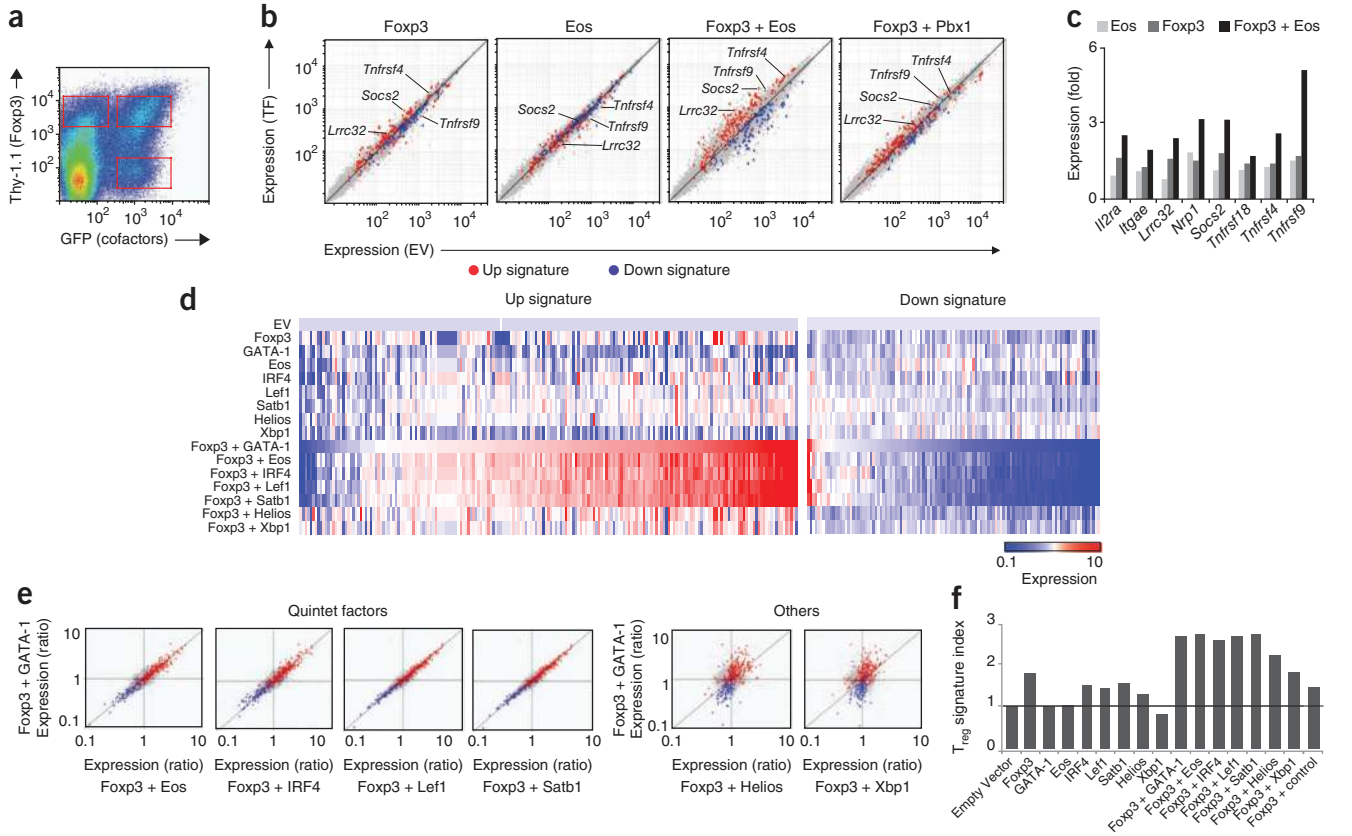

Figure 3 Transcriptional induction of genes of the T_{reg} cell signature by Foxp3 and other transcription factors. (**a**) Flow cytometry profile of purified mouse T_{conv} cells activated and retrovirally transduced with expression vectors encoding Foxp3 (with a Thy-1.1 reporter) and various cofactors (with a green fluorescent protein (GFP) reporter) and sorted after 3 d of culture. Outlined areas indicate cells expressing Foxp3 alone (top left), Foxp3 and the cofactors (top right) or the cofactors alone (bottom right). (**b**) Expression profiles of T_{conv} cells transduced to express Foxp3 and Eos alone or together, or Foxp3 plus the control transcription factor Pbx1 (TF; vertical axis), compared with those of cells transduced with empty vector (EV; horizontal axis). (**c**) RT-PCR analysis of the expression of a subset of genes of the T_{reg} cell signature in cells transduced as in **b**, presented relative to that of cells transduced with empty vector. (**d**) Genes of the T_{reg} cell signature upregulated or downregulated after transduction of vector encoding each candidate transcription factor alone (top group) or together with vector encoding Foxp3 (bottom group). (**e**) Change in the expression of genes of the T_{reg} cell signature elicited by Foxp3 plus various cofactors (horizontal axes) compared with changes elicited by Foxp3 plus GATA-1 (vertical axis); ratios are relative to expression in cells transduced with empty vector (ratio). (**f**) Transition toward the T_{reg} cell phenotype, assessed by a cumulative T_{reg} cell signature index. Data are representative of two independent experiments.

these mice and their wild-type littermates (**Fig. 2b**). We detected no bias in any of the mutant T_{reg} cells, whether for the T_{reg} cell signature as a whole (**Fig. 2b**) or for the computationally predicted targets of each transcription factor (**Supplementary Fig. 2**). Thus, whereas each of these transcription factors may have affected the T_{reg} cell signature when they varied naturally in the diverse cell types used for the computational analysis, the T_{reg} cell signature was robust after the complete elimination of any one of them.

Gain-of-function confirmation of computational predictions

We then did gain-of-function experiments by retrovirally transducing activated $CD4^+$ T_{conv} cells with cDNA encoding each of the various candidate transcription factors, alone or together with retrovirus containing cDNA encoding human Foxp3 (which has a transcriptional signature similar to that of mouse Foxp3 (ref. 11); this allowed us to distinguish the ectopic expression from the endogenous transcripts). These manipulations resulted in expression of Foxp3 and other transcription factors in the same general range as that in normal *ex vivo* T cells (**Supplementary Fig. 3**). We sorted cells expressing each transcription factor, alone or together with Foxp3, 3 d after transduction and profiled their gene expression (**Fig. 3a**). Consistent with published findings[11,20], Foxp3 alone influenced only a limited number of genes of the T_{reg} cell signature (**Fig. 3b**). Enforced expression of each cofactor alone had even less effect, but we observed robust induction of upregulated genes of the T_{reg} cell signature and repression of downregulated genes of the T_{reg} cell signature when Foxp3 and cofactors were both present (**Fig. 3b**). We confirmed those results by RT-PCR analysis of a subset of genes from independent samples (**Fig. 3c**). We noted that synergistic outcome with each of the seven candidate transcription factors tested (**Supplementary Fig. 4**), but five of them (Eos, IRF4, GATA-1, Lef1 and Satb1; called the 'quintet' here) had a notable effect, acting together with Foxp3 to similarly shift most of the T_{reg} cell signature (**Fig. 3d** and **Supplementary Table 3**). Indeed, each of this quintet of transcription factors, together with Foxp3, regulated the same set of genes (direct comparison, **Fig. 3e**). This synergy was not an artifact of the dual transduction, as results obtained with cells transduced to express Foxp3 and a control transcription factor (Pbx1) were similar to those of cells expressing Foxp3 alone (**Fig. 3b**,**f**). Rates of cell division were also identical in singly and doubly

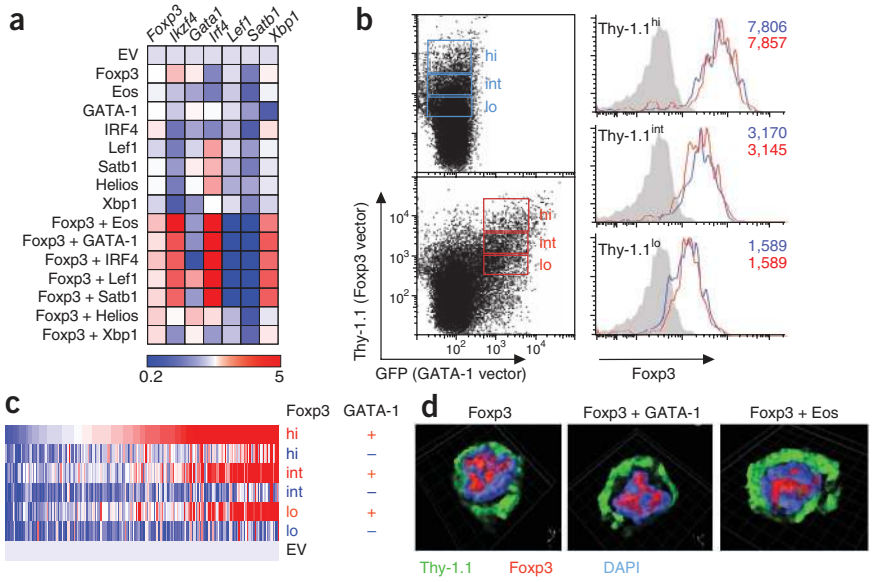

Figure 4 Mechanistic effect of Foxp3 cofactors. (**a**) Expression of endogenous transcripts encoding Foxp3 and cofactors in cells transduced to express various factors (as in **Fig. 3d**). (**b**) Sorting of CD4$^+$ T_{conv} cells transduced to express Foxp3 (blue) or Foxp3 plus GATA-1 (red) into matching bins of Thy-1.1 (Foxp3) reporter intensity (left; intensity indicated by labels adjacent to outlined areas), and intracellular staining of Foxp3 in the sorted cells (right). Numbers in plots at right indicate the mean fluorescence intensity of Foxp3 protein (key colors match line colors). Gray shaded curves, isotype-matched control antibody. (**c**) Expression profiles of genes of the T_{reg} cell signature in cells prepared as in **b**. (**d**) Confocal microscopy of CD4$^+$ cells transduced to express Foxp3 alone or together with Eos or GATA-1, stained for Foxp3 and Thy-1.1 and with the DNA-intercalating dye DAPI. Original magnification, ×100. Data are representative of two experiments (**a**), three experiments (**b**,**d**) or one experiment (**c**).

transduced cells, as measured by dilution of the cytosolic dye CFSE (**Supplementary Fig. 5**). The synergistic outcome was different from that prompted by Helios or Xbp1, although the latter also acted in synergy with Foxp3, as shown by the integrated T_{reg} cell signature index (**Fig. 3f**). We then sought to determine whether a combination of two transcription factors of the quintet would induce the T_{reg} cell signature without Foxp3. Indeed, the combination of Eos and Lef1 or of GATA-1 and Satb1 had a partial effect, including a modest induction of *Foxp3* (**Supplementary Fig. 6**).

We used microarrays with features that span the length of the transcripts, which allowed us to parse signals from the transfected genes and their endogenous homologs (**Fig. 4a** and **Supplementary Table 4**). Foxp3 plus any factor of the quintet modified the expression of endogenous transcripts encoding transcription factors, inducing *Foxp3*, *Ikzf4*, *Irf4* and *Xbp1* and repressing *Lef1* and *Satb1*. Thus, each factor of the quintet acted in synergy with Foxp3 to induce widespread reassortment of the cell's balance of regulatory transcription factors in an autoassembly of the T_{reg} cell profile. This involved the induction of transcription factors that were normally overexpressed (Eos and IRF4) and the repression of those that were underexpressed (Lef1 and Satb1) in T_{reg} cells (**Supplementary Fig. 7**). Thus, these results indicated a synergistic effect between Foxp3 and cofactors that propagated to other transcription factors and 'locked in' the T_{reg} cell signature. Accordingly, the genes affected here included genes of the T_{reg} cell signature found to be Foxp3 independent in published studies[11].

How the cofactors operate

We then sought to determine how the cofactors of the quintet might elicit such a transition. This did not occur through the stabilization of Foxp3 protein, whose abundance (as measured by intracellular staining) was identical whether or not the cells were cotransduced to express a factor of the quintet, across a range of expression values of the cotranscribed reporter (**Fig. 4b**). It was possible that factors of the quintet had a quantitative influence on the activity of Foxp3 by simply displacing a threshold of activity below which Foxp3 would be ineffective. To test this hypothesis, we sorted and profiled matched bins of cells transduced to express various amounts of Foxp3, alone or together with GATA-1, chosen as a representative of the factors of the quintet (**Fig. 4c**). As might be expected, increasing amounts of Foxp3 did have a more substantial effect on transcription. However, even the greatest amount of Foxp3 alone was unable to match the induction of genes of the T_{reg} cell signature obtained by Foxp3 with GATA-1. The cooperative effect of GATA-1 was apparent at all amounts of Foxp3. Thus, the factors of the quintet did not merely provide a quantitative 'boost' to Foxp3 but instead enhanced its transcriptional activity. The nuclear/cytoplasmic distribution of Foxp3 was unaltered by the experimental conditions, as it was almost exclusively nuclear whether transduced alone or together with a quintet factor (**Fig. 4d**).

The effects reported above also suggested that Foxp3 interacted molecularly with the factors of the quintet in nuclear complexes. Such interactions have already been demonstrated for IRF4 and Eos[23,24], so we tested the other three factors. Indeed, reciprocal coimmunoprecipitation of proteins from transfected cells identified interactions between Foxp3 and GATA-1, Satb1 and Lef1 but not between Foxp3 and the control transcription factor Pbx1 (**Supplementary Fig. 8**).

The synergizing activity of the cofactors, which was most notable with factors of the quintet and was not accounted for by quantitative effects on Foxp3 or its global cellular localization, could have the following two interpretations: through cooperative binding, the cofactors might recruit Foxp3 to genomic locations near T_{reg} cell signature genes; or the cofactors might enhance the activity of Foxp3 molecules already bound to DNA. To distinguish between those scenarios, we used chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq) to assess how factors of the quintet affected the genome-wide localization of Foxp3. We prepared chromatin from primary CD4$^+$ T_{conv} cells transduced to express Flag-tagged Foxp3 alone or together with GATA-1 (the preparation of doubly transduced cells in the numbers needed for ChIP-Seq was technically very demanding, so we chose GATA-1 as a representative factor), immunoprecipitated proteins with anti-Flag and assessed the bound DNA by deep sequencing (ChIP-Seq statistics, **Supplementary Table 5**). Immunoprecipitation with antibody to RNA polymerase II or immunoprecipitation of whole-cell extracts provided a genome-wide control for transcriptional start sites (TSSs) or for sequencing nonhomogeneity, respectively. Overall, summing of the genome-wide signals relative to TSS locations showed that Foxp3 localized mainly in the vicinity of known TSSs, as expected (**Fig. 5a**); other experiments of ChIP-Seq with irrelevant control antibody have shown a paucity of signal around the TSS[40], which substantiates the importance of

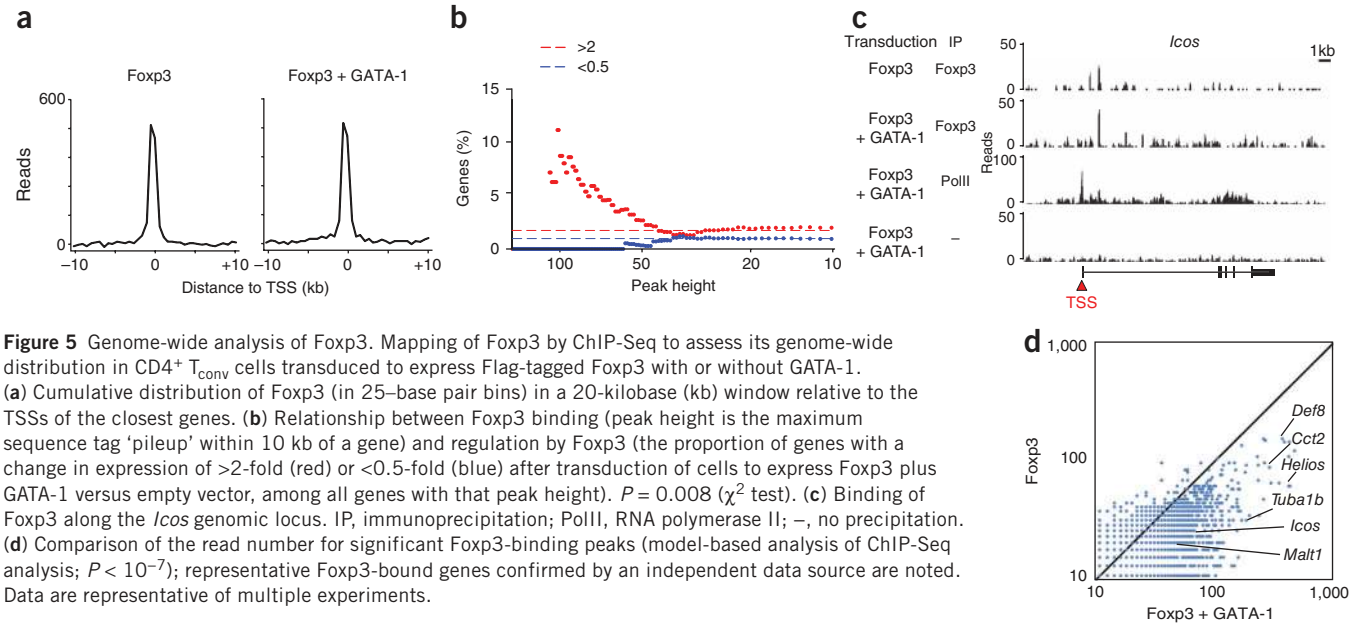

Figure 5 Genome-wide analysis of Foxp3. Mapping of Foxp3 by ChIP-Seq to assess its genome-wide distribution in CD4$^+$ T_{conv} cells transduced to express Flag-tagged Foxp3 with or without GATA-1. (**a**) Cumulative distribution of Foxp3 (in 25–base pair bins) in a 20-kilobase (kb) window relative to the TSSs of the closest genes. (**b**) Relationship between Foxp3 binding (peak height is the maximum sequence tag 'pileup' within 10 kb of a gene) and regulation by Foxp3 (the proportion of genes with a change in expression of >2-fold (red) or <0.5-fold (blue) after transduction of cells to express Foxp3 plus GATA-1 versus empty vector, among all genes with that peak height). $P = 0.008$ (χ^2 test). (**c**) Binding of Foxp3 along the *Icos* genomic locus. IP, immunoprecipitation; PolII, RNA polymerase II; –, no precipitation. (**d**) Comparison of the read number for significant Foxp3-binding peaks (model-based analysis of ChIP-Seq analysis; $P < 10^{-7}$); representative Foxp3-bound genes confirmed by an independent data source are noted. Data are representative of multiple experiments.

the signals observed here. The data allowed statistically robust detection of more than 5,000 Foxp3-binding sites (model-based analysis of ChIP-Seq analysis; $P < 10^{-7}$; **Supplementary Table 6**). We confirmed binding signals on many of these sites by comparison with similar data from *ex vivo* T_{reg} cells (R. Samstein and A. Rudensky, personal communication). To further confirm these data, we computed the distribution of genes whose expression was affected by transduction of retrovirus expressing Foxp3 and cofactors, for the range of genes showing significant peaks of Foxp3 binding. As might be expected, the group of genes with the highest peaks of Foxp3 binding was enriched for those genes activated by Foxp3 in the microarray data (of the 57 genes with a Foxp3 binding peak height of >75 'reads', 12.2% had a change in expression of >1.6-fold after transduction, compared with 4.7% for the entire data set; **Fig. 5b**). This was not absolutely true, however, and many sites of strong Foxp3 binding did not correspond to a significant transcriptional change in expression, as is often noted for ChIP-Seq data. In addition, transcripts repressed by Foxp3 were not over-represented among those with the highest Foxp3 binding.

In cells doubly transduced to express Foxp3 and GATA-1, we did not observe additional sites of significant binding of Foxp3. Instead, there was quantitatively enhanced occupancy by Foxp3 at the same locations as in the singly transduced cells, as shown for the Foxp3-binding site in the first intron of the representative gene *Icos* (**Fig. 5c**), or when we quantified genome-wide binding in parallel (**Fig. 5d**). Thus, the factors of the quintet did not spread the recruitment of Foxp3 to different genomic locations but seemed to functionally stabilize Foxp3 and enhance the activity of Foxp3 independently bound to its target sites.

Signature 'lock' by feedback loops: computational modeling

The results presented above (**Fig. 3d**) raised the following question: how could five distinct transcription factors, of widely different structure, DNA sequence specificity and functional activity, elicit the same transcriptional outcome? This was particularly paradoxical for Lef1 and Satb1, which are repressed in T_{reg} cells[26] (**Supplementary Fig. 7**). However, the retroviral vectors used contained only the coding regions of *Satb1* or *Lef1* and lacked the 3′ and 5′ untranslated regions, which have been shown to be involved in the regulation of endogenous Satb1 expression[26]; this probably led to 'constitutive' expression of exogenous Lef1 and Satb1 during the culture period. The following plausible interpretation was suggested by the effect on endogenous expression of transcription factors (**Fig. 4a**): the T_{reg} cell signature, along with the regulatory factors it includes, is organized with regulatory feedback loops, both positive and negative, such that it has the ability to autoassemble and 'lock in' once the expression of Foxp3 and some cofactors is 'pushed' beyond that of T_{conv} cells. Intuitively, such 'locking in' could be achieved not only by positive feedback but also by double-negative inhibition of repressive factors.

To assess the plausibility of that proposal, we used computational simulation to determine whether such a self-reinforcing system that incorporated repressed regulators could actually function. We developed a mathematical model to simulate the dynamics of such a system (**Fig. 6** and **Supplementary Note**). The model consisted of the main regulator Foxp3 (F), with its active conformation F* (the transition from F to F* can mean quantitative induction, post-translational modification (such as acetylation) or complex formation, as suggested by coimmunoprecipitation assays, that potentiate or stabilize F), and a set of downstream regulatory factors of the T_{reg} cell signature whose genes were either upregulated (U_i) or downregulated (D_i) by the main regulator F (**Fig. 6a**; details, **Supplementary Note**). The molecules encoded by those signature genes themselves controlled smaller subnetworks of factors, some mere effectors (U_4 and D_3) and others able to regulate F* (such as U_1–U_3, D_1 and D_2). The output of these subnetworks, which themselves could operate with 'and' or 'or' logic, then controlled a larger set of signature genes but most importantly influenced the conversion from F to F*.

To make the model computationally manageable, we bypassed subnetwork calculations and omitted cross-regulatory influences between cofactors, which were likely to occur. Differential equations paired up with Hill functions described the biochemical kinetics engaged in the model. After a reasonable parameter set was fixed, this minimal model successfully reproduced the bi-stable features of the T_{reg} cell program and the outcome of the experimental perturbations (**Figs. 2** and **3**). After *in silico* transduction (**Fig. 6b**), expression of genes of the T_{reg} cell signature remained off when Foxp3 or any of the cofactors was expressed alone, but all signature genes transitioned to the T_{reg} cell state and were 'locked in' when Foxp3 was overexpressed together

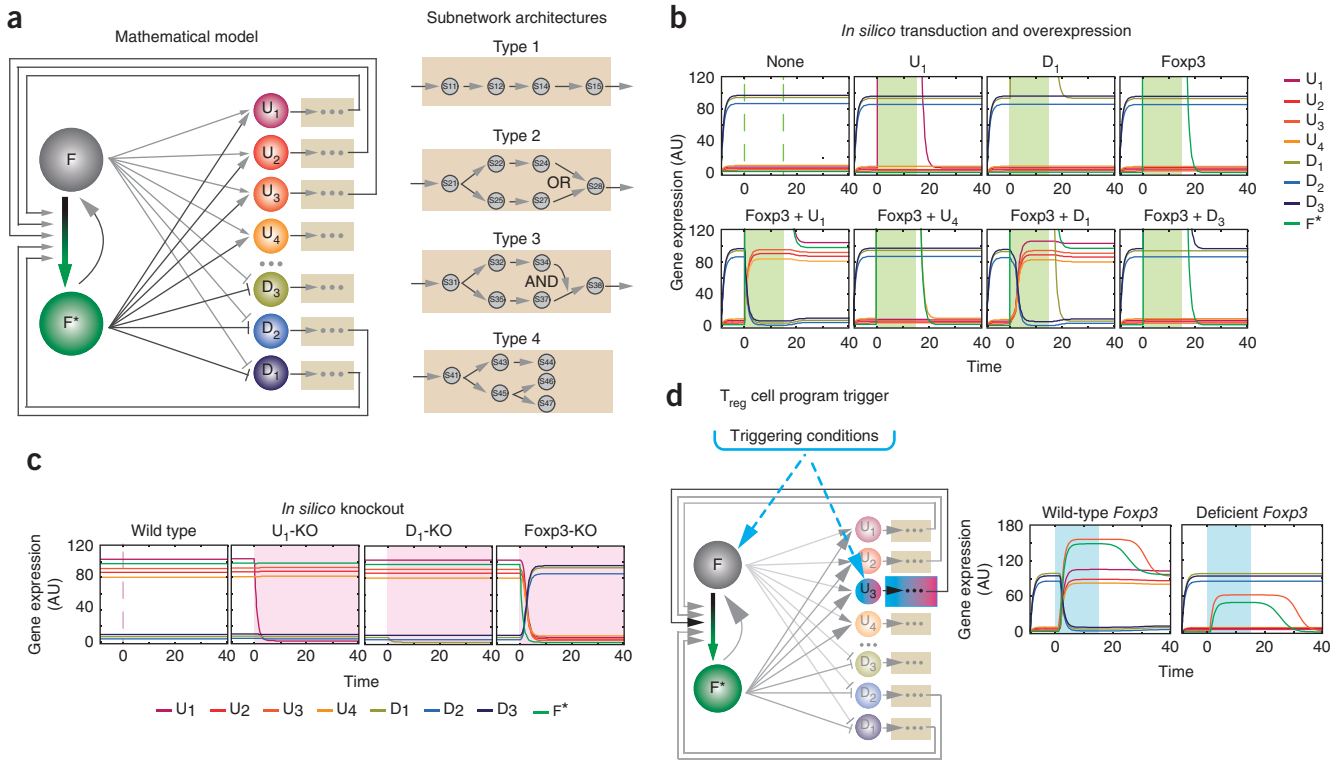

Figure 6 Mathematical modeling of a 'self-locking' network. (**a**) Mathematical model consisting of the main regulator Foxp3 (F), with its active conformation F* (where this transition can represent transcriptional or post-transcriptional activation) and a set of downstream regulatory factors of the T_{reg} cell signature whose genes are upregulated (U_1–U_4) or downregulated (D_1–D_3). Subsets of molecules encoded by the signature genes (U_1–U_3, D_1 and D_2) positively activate the transition from F to F*, directly or through the subnetworks they control. (**b**) *In silico* simulation of transduction and overexpression experiments: lines indicate expression of transcription factors (key), presented in arbitrary units (AU); green shading indicates the time window of factor overexpression. (**c**) Simulated knockout experiments (presented as in **b**); pink shading corresponds to the time frame after factor elimination. (**d**) Activation of the T_{reg} cell program; blue shading (bottom) indicates the time window of the inducing conditions.

with any of the cofactors, including those repressed in T_{reg} cells (for example, D_1). The model showed no effect of the single *in silico* deletion of any of the cofactors (**Fig. 6c**), consistent with the experimental results, but the T_{reg} cell signature 'shut off' with extinction of Foxp3. That last finding deviated somewhat from the experimental results, as thymuses of mice with inactivated Foxp3 protein do contain cells with some T_{reg} cell features, including partial activation of the T_{reg} cell signature and a transcriptionally active *Foxp3* locus[17,18]. The discrepancy could be resolved in the model by the postulation that the differentiation of T_{reg} cells triggers, directly or indirectly, transient expression of Foxp3 and one of its cofactors (such as U_3 in **Fig. 6d**, top). In the simulation, transient activation of F and U_3 resulted in activation of the whole network (**Fig. 6d**, bottom left), but only if the external inducing conditions were modeled to activate both F and U3. This activation was then stable after removal of the inductive stimulus. With inactive Foxp3, however, the cells showed only partial T_{reg} cell features, which reverted to the T_{conv} cell state some time after withdrawal of the differentiating stimulus (**Fig. 6d**, bottom right), a scenario consistent with the unstable T_{reg} cell–like cells mentioned above. Thus, the simulation arrived at a model of T_{reg} cell differentiation compatible with most experimental outcomes and with several aspects of T_{reg} cell physiology.

DISCUSSION

Our work here arrived at a conceptual framework different from its origin. The intent was to use computational network inference to predict the panel of transcription factors that act together with Foxp3 to determine the canonical T_{reg} cell signature. We expected that experimental confirmation by loss- or gain-of-function experiments would define discrete gene modules affected by each of the cofactors, probably with some degree of synergy. Instead, we arrived at a very different perspective in which the T_{reg} cell signature involves a very high degree of positive and negative feedback, such that the signature autoassembles and reaches the same state in response to different triggers. Accordingly, the T_{reg} cell signature proved impervious to the removal of any one of the factors, with the exception of Foxp3.

Although it does so with much more complex determinism, the control of the T_{reg} cell signature acts much like a classic genetic switch. A genetic switch describes stable and inheritable changes in the phenotypic state of a genetic system that are conserved after termination of the initiating signal. First shown to explain the stable lysogenic state of the bacteriophage λ driven by the cI repressor[41], genetic switches based on the reciprocal action of transcription factors have been demonstrated in diverse phenomena such as long-term memory potentiation[42], cell transformation[43] or the maintenance of pluripotency in embryonic stem cells[44]. Positive feedback loops combined with suppression, either direct or indirect, are inherent to the operation of a switch and to the bi-stable states achieved. Much as neural memory must persist over time, the self-reactivity of TCRs expressed by T_{reg} cells makes it important for their suppressive phenotype to be stable over the course of an infection and to prevent autoimmunity[9]. For T_{reg} cells, modifications of the methylation status

of the *Foxp3* locus also contribute to this stability[45]. Genetic switches also ensure that a state outlives the conditions that set it: lysogeny by bacteriophage λ is self-perpetuating once established; for T_{reg} cells, the TCR ligand and the cytokine milieu that led to the establishment of the T_{reg} cells need not be maintained. This remanence could be important for the thymic induction of T_{reg} cell differentiation by self antigens, which may not be encountered in the same processed form in the periphery, or for T_{reg} cells induced by gut commensal bacteria, which are cells that should persist even with fluctuating microbial flora.

Unlike the minimalist simplicity of the bacteriophage λ switch, the T_{reg} cell switch was very complex. First, several factors participated in a synergistic manner, and the factors of the quintet had to activate several distinct pathways and loops. Second, two inputs were necessary for the establishment of the T_{reg} cell state. Neither Foxp3 alone nor any of the cofactors were sufficient to 'lock in' the T_{reg} cell signature. There are distinct advantages to 'two-key' control systems, which diminish the risk of the long-term consequences that would result from erroneous activation, in this case from 'noisy' gene expression such as the transient induction of *Foxp3* after the activation of $CD4^+$ effector T cells. Signaling along TCR and interleukin 2 receptor (IL-2R)–STAT5 pathways that promote commitment to the T_{reg} cell fate might achieve this duality (for example, by activating the transcription factors NF-κB and Foxp3, respectively). The scenario modeled by the computational simulation is consistent with a two-step process of T_{reg} cell differentiation, which goes through a $Foxp3^-CD25^{hi}$ intermediate cell that secondarily converts to a $Foxp3^+$ cell under the influence of IL-2 or other trophic cytokines[46,47]. In addition, the model probably accounts for the somewhat divergent results obtained by different groups after transduction for Foxp3 expression[11,15,16]. Although we observed only very limited effects, even after Foxp3 expression in amounts similar to those of *ex vivo* T_{reg} cells, other studies have reported much more functional activity of Foxp3. The precise conditions of culture and of cell activation for retroviral transduction, such as supplementation with IL-2, may have induced in some experimental settings one of the cofactors needed to act in synergy with Foxp3 and activate the switch.

Finally, there was multiple redundancy in the T_{reg} cell switch, as exemplified by the actions of the factors of the quintet. This ability to flip the switch was not universal (Helios and Xbp1 did not have that potential), but five of the seven transcription factors tested here had it, and there is no reason to think that the list is complete. Such redundancy not only ensures additional stability, as exemplified by the data obtained with knockout mice, but also allows several different physiological pathways to arrive at the same state. This may be relevant to the different thymic and extra-thymic contexts of T_{reg} cell differentiation[4]. Lymphopenic conditions or chronic exposure to antigen or molecules produced by gut microbes may each induce one or another cofactor able to 'lock in' the T_{reg} cell transcriptional network.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GEO: microarray and ChIP-Seq data, GSE7460, GSE7596, GSE7852, GSE13306 and GSE40278.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank R. Samstein and A. Rudensky for the unpublished ChIP-seq data; S. Smale (University of California, Los Angeles) for mouse cDNA encoding Helios; M. Calderwood and the Center for Cancer Systems Biology for expression cDNA; P. Rahl for advice on ChIP-Seq; J. Ericson, S. Davis, H. Paik and R. Cruse for genomic data analysis; H. Chen and Q. Cai for experimental support; and J. LaVecchio and G. Buruzala for sorting. This work benefited from public data generated by the Immunological Genome Project consortium. Supported by the US National Institutes of Health (AI051530 to C.B. and D.M.; AI072073 to C.B., D.M. and J.C.; training grant T32 DK7260 for support of M.S.F.; and 3R24AI072073-03S1 for support of A.E.), GlaxoSmithKline, the Damon Runyon Cancer Research Foundation (S.H.), the American Diabetes Association (7-07-BETA-14 to W.F.) and the Canadian Institutes of Health Research (J.H.).

AUTHOR CONTRIBUTIONS

W.F., A.E., T.L., M.S.F., J.A.H., S.A., J.J.C., D.M. and C.B. designed experiments; W.F., J.A.H., S.H., M.S.F. and R.G. did experiments; A.E. and T.L. did computation; L.G., S.C., P.K. and D.R. provided mice and advice; and W.F., A.E., T.L., J.A.H., R.G., M.S.F., J.J.C., D.M. and C.B. analyzed data and wrote manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ni.2420.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T cells and immune tolerance. *Cell* **133**, 775–787 (2008).
2. Feuerer, M. *et al.* Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. *Nat. Med.* **15**, 930–939 (2009).
3. Josefowicz, S.Z., Lu, L.F., Rudensky, A.Y. & Regulatory, T. Cells: Mechanisms of differentiation and function. *Annu. Rev. Immunol.* **30**, 531–564 (2012).
4. Curotto de Lafaille, M.A. & Lafaille, J.J. Natural and adaptive $foxp3^+$ regulatory T cells: more of the same or a division of labor? *Immunity* **30**, 626–635 (2009).
5. Duarte, J.H. *et al.* Natural Treg cells spontaneously differentiate into pathogenic helper cells in lymphopenic conditions. *Eur. J. Immunol.* **39**, 948–955 (2009).
6. Tsuji, M. *et al.* Preferential generation of follicular B helper T cells from $Foxp3^+$ T cells in gut Peyer's patches. *Science* **323**, 1488–1492 (2009).
7. Murai, M. *et al.* Interleukin 10 acts on regulatory T cells to maintain expression of the transcription factor Foxp3 and suppressive function in mice with colitis. *Nat. Immunol.* **10**, 1178–1184 (2009).
8. Zhou, X. *et al.* Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells *in vivo*. *Nat. Immunol.* **10**, 1000–1007 (2009).
9. Rubtsov, Y.P. *et al.* Stability of the regulatory T cell lineage *in vivo*. *Science* **329**, 1667–1671 (2010).
10. Komatsu, N. *et al.* Heterogeneity of natural $Foxp3^+$ T cells: a committed regulatory T-cell lineage and an uncommitted minor population retaining plasticity. *Proc. Natl. Acad. Sci. USA* **106**, 1903–1908 (2009).
11. Hill, J.A. *et al.* Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity* **27**, 786–800 (2007).
12. Feuerer, M., Hill, J.A., Mathis, D. & Benoist, C. $Foxp3^+$ regulatory T cells: differentiation, specification, subphenotypes. *Nat. Immunol.* **10**, 689–695 (2009).
13. Vignali, D.A., Collison, L.W. & Workman, C.J. How regulatory T cells work. *Nat. Rev. Immunol.* **8**, 523–532 (2008).
14. Ziegler, S.F. FOXP3: of mice and men. *Annu. Rev. Immunol.* **24**, 209–226 (2006).
15. Hori, S., Nomura, T. & Sakaguchi, S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**, 1057–1061 (2003).
16. Fontenot, J.D. *et al.* Regulatory T cell lineage specification by the forkhead transcription factor foxp3. *Immunity* **22**, 329–341 (2005).
17. Gavin, M.A. *et al.* Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**, 771–775 (2007).
18. Lin, W. *et al.* Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol.* **8**, 359–368 (2007).
19. Otsubo, K. *et al.* Identification of FOXP3-negative regulatory T-like ($CD4^+CD25^+CD127^{low}$) cells in patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome. *Clin. Immunol.* **141**, 111–120 (2011).
20. Sugimoto, N. *et al.* Foxp3-dependent and -independent molecules specific for $CD25^+CD4^+$ natural regulatory T cells revealed by DNA microarray analysis. *Int. Immunol.* **18**, 1197–1209 (2006).
21. Ono, M. *et al.* Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**, 685–689 (2007).
22. Wu, Y. *et al.* FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* **126**, 375–387 (2006).
23. Pan, F. *et al.* Eos mediates Foxp3-dependent gene silencing in $CD4^+$ regulatory T cells. *Science* **325**, 1142–1146 (2009).
24. Zheng, Y. *et al.* Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control T_H2 responses. *Nature* **458**, 351–356 (2009).
25. Zhou, L. *et al.* TGF-β-induced Foxp3 inhibits T_H17 cell differentiation by antagonizing RORγt function. *Nature* **453**, 236–240 (2008).

26. Beyer, M. *et al.* Repression of the genome organizer SATB1 in regulatory T cells is required for suppressive function and inhibition of effector differentiation. *Nat. Immunol.* **12**, 898–907 (2011).
27. Chaudhry, A. *et al.* CD4$^+$ regulatory T cells control TH17 responses in a Stat3-dependent manner. *Science* **326**, 986–991 (2009).
28. Shi, L.Z. *et al.* HIF1α-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells. *J. Exp. Med.* **208**, 1367–1376 (2011).
29. Dang, E.V. *et al.* Control of T_H17/T_{reg} balance by hypoxia-inducible factor 1. *Cell* **146**, 772–784 (2011).
30. Koch, M.A. *et al.* The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nat. Immunol.* **10**, 595–602 (2009).
31. Wang, Y., Souabni, A., Flavell, R.A. & Wan, Y.Y. An intrinsic mechanism predisposes Foxp3-expressing regulatory T cells to Th2 conversion *in vivo*. *J. Immunol.* **185**, 5983–5992 (2010).
32. Gardner, T.S. & Faith, J.J. Reverse-engineering transcription control networks. *Phys. Life Rev.* **2**, 65–88 (2010).
33. Basso, K. *et al.* Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37**, 382–390 (2005).
34. Battle, A., Segal, E. & Koller, D. Probabilistic discovery of overlapping cellular processes and their regulation. *J. Comput. Biol.* **12**, 909–927 (2005).
35. Faith, J.J. *et al.* Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5**, e8 (2007).
36. Thornton, A.M. *et al.* Expression of Helios, an Ikaros transcription factor family member, differentiates thymic-derived from peripherally induced Foxp3$^+$ T regulatory cells. *J. Immunol.* **184**, 3433–3441 (2010).
37. Yu, C. *et al.* Targeted deletion of a high-affinity GATA-binding site in the GATA-1 promoter leads to selective loss of the eosinophil lineage *in vivo*. *J. Exp. Med.* **195**, 1387–1395 (2002).
38. Lee, A.H., Scapa, E.F., Cohen, D.E. & Glimcher, L.H. Regulation of hepatic lipogenesis by the transcription factor XBP1. *Science* **320**, 1492–1496 (2008).
39. Cai, Q. *et al.* Helios deficiency has minimal impact on T cell development and function. *J. Immunol.* **183**, 2303–2311 (2009).
40. Giraud, M. *et al.* Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. *Proc. Natl. Acad. Sci. USA* **109**, 535–540 (2012).
41. Johnson, A.D. *et al.* λ Repressor and cro—components of an efficient molecular switch. *Nature* **294**, 217–223 (1981).
42. Pittenger, C. & Kandel, E. A genetic switch for long-term memory. *C.R. Acad. Sci. III* **321**, 91–96 (1998).
43. Iliopoulos, D., Hirsch, H.A. & Struhl, K. An epigenetic switch involving NF-κB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. *Cell* **139**, 693–706 (2009).
44. Young, R.A. Control of the embryonic stem cell state. *Cell* **144**, 940–954 (2011).
45. Huehn, J., Polansky, J.K. & Hamann, A. Epigenetic control of FOXP3 expression: the key to a stable regulatory T-cell lineage? *Nat. Rev. Immunol.* **9**, 83–89 (2009).
46. Lio, C.W. & Hsieh, C.S. A two-step process for thymic regulatory T cell development. *Immunity* **28**, 100–111 (2008).
47. Burchill, M.A. *et al.* Linked T cell receptor and cytokine signaling govern the development of the regulatory T cell repertoire. *Immunity* **28**, 112–121 (2008).

ONLINE METHODS

Mice. C57BL/6 mice were from the Jackson Laboratory. Mice with a mutated *Gata1* promoter on the BALB/c background[37] (with a deletion in the double-GATA site 21 base pairs upstream of the first hematopoietic exon) were from The Jackson Laboratory. Mice with conditional knockout of *Xbp1* (with *lox*P-flanked *Xbp1* alleles deleted by Cre expressed from the poly(I:C)-inducible gene *Mx1*) on the C57BL/6 background have been described[38]; 5- to 6-week-old mice were injected intraperitoneally three times with 250 μg poly(I:C) (invivoGen) with 2-day intervals for the induction of *Cre* expression; mice were used for experiments 6 weeks after the final injection of poly(I:C). Mice deficient in Helios *(Ikzf2)* have been described[39]. For the generation of mice constitutively deficient in Eos (*Ikzf4*), mice with insertion of *lox*P sites flanking exons 1–4 of *Ikzf4* were crossed with mice that have germline expression of Cre. Targeting of the genomic locus was confirmed by Southern blot analysis with 5′ and 3′ external probes, deletion of exons 1–4 was confirmed by quantitative RT-PCR. Homozygous Eos-deficient mice are viable and fertile with no apparent abnormality. All mice cared for in accordance with the ethical guidelines of the Center for Animal Resources and Comparative Medicine at Harvard Medical School under procedures approved by the Institutional Animal Care and Use Committee (protocol 02954).

CLR network regulatory prediction. For this analysis, we compiled 129 published gene-expression data sets obtained with purified $CD4^+$ T cell populations in several experimental contexts: *ex vivo* T_{conv} cells or T_{reg} cells from various anatomical locations; cultured T_{reg} cells; TGFβ-induced $Foxp3^+$ cells; and retinoic acid–treated cultures[11,12,48]. The Affymetrix M430v2 microarray raw data were preprocessed by the robust multiarray average algorithm in the GenePattern genomic analysis platform[49], and averaged expression values were used for analysis.

For a robust definition of the transcriptional signature of mature T_{reg} cells, results from several independent experiments had been brought together. The consensus peripheral T_{reg} cell signature had been defined by calculation of the ratio of the expression in T_{reg} cells to that in T_{conv} cells, with retention of only those genes with consistent 1.5-fold overexpression or underexpression in T_{reg} cells in all four data sets, with Affymetrix M430v2 arrays. This resulted in a total of 603 genes (407 overexpressed and 196 underexpressed, respectively, in T_{reg} cells)[11].

To predict regulatory connections, we used the CLR algorithm[35], which operates by combining the relative strength of coexpression of a regulator and potential targets. The CLR algorithm builds on the relevance network strategies by applying a background-correction step. First we reconstructed a relevance network of 2,021 transcriptional regulators and 603 T_{reg} cell signature genes[11]. After computing mutual information (MI) values for all pairwise transcription factor–T_{reg} cell signature gene pairs, the algorithm compares the MI of a transcription factor–gene pair to the background distribution of MI scores of all genes associated with the transcription factor or all transcription factors associated with the gene of interest. After this background correction, the most probable interactions are those whose combined scores were significantly above the background distribution of MI scores. The background corrected CLR scores were filtered at a false-discovery rate of 0.005 (computed with the Bonferroni correction) to generate the CLR network of T_{reg} cells. MATLAB software (MathWorks) was used for CLR computations.

In the second phase, starting from the CLR scores, we formulated an optimization problem whose objective was to identify the transcription factors that together influenced the T_{reg} cell signature the most and accounted for the most genes. This is a mixed-integer optimization problem that we solved with the CPLEX9.0 optimization package (ILOG tool; IBM) for the AMPL modeling language and tool set.

Retroviral transduction. For enforced expression of Foxp3 and other transcription factors, the retroviral expression vector MSCV-IRES-Thy-1.1/GFP[11] was used throughout. The cDNA encoding human Foxp3, GATA-1, Eos, IRF4, Lef1, Satb1 or Xbp1 was from human ORFeome. Mouse cDNA encoding Helios was provided by S. Smale. Plat-E cells were plated 1 d before transfection with those plasmids, together with the packaging construct pCL-Eco through the use of Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen).

For the transduction of $CD4^+$ T cells, cell suspensions were prepared from spleens and lymph nodes of 6- to 8-week-old C57BL/6 mice by physical dissociation and red blood cells were lysed in 0.8% ammonium chloride lysis buffer. $CD4^+$ T cells were negatively purified by magnetic selection (with labeling with phycoerythrin-conjugated anti-CD11b (M1/70), anti-CD11c (N418), anti-CD19 (6D5), anti-CD8α (53-6.7), anti-CD25 (PC61) and anti-NK1.1 (PK136; all from BioLegend). After cells were washed, anti-phycoerythrin beads (130-048-801; Miltenyi Biotec) were added to the cell suspension and then $CD4^+$ T_{conv} cells were purified with MACS LD columns (Miltenyi) to a purity of >95%. Cells were then activated with beads coated with anti-CD3 and anti-CD28 (Invitrogen) at a ratio of one bead per cell, with addition of 20 U/ml recombinant human IL-2 (Proleukin; Chiron) in complete culture medium (RPMI-1640 medium supplemented with 10% FCS, 2 mM L-glutamine, 100 U/ml penicillin-streptomycin and 50 μM 2-mercaptoethanol). T cells were cultured for 24 h and then were spin-infected (for 2 h at 2,000 r.p.m. and 32 °C) with retroviral supernatants. Cells were then cultured for an additional 72 h. Infected cells were sorted by flow cytometry as $CD4^+$ cells further gated on Thy-1.1 and GFP that report expression of Foxp3 and of the other cofactors, respectively. For the experiments in **Figure 4b**,**c**, infected T cells were sorted into fractions on the basis of the intensity of Thy-1.1 expression (high, intermediate and low) for microarray profiling and analysis of Foxp3 protein.

For confirmation and quantification of Foxp3 expression in the transduced cells, sorted cells were fixed and permeabilized for intracellular staining with anti-Foxp3 (FJK-16s; eBioscience) according to the manufacturer's instructions, and were analyzed by flow cytometry (LSR II; BD) for quantitative analysis of Foxp3 protein expression or by confocal microscopy (Axiovert 200M; Zeiss) for analysis of the localization of Foxp3 protein.

Gene-expression profiling. For analysis of gene expression in knockout mice or after retroviral transduction, sorted cell populations were lysed in TRIzol reagent and RNA was prepared according to the manufacturer's instructions (Invitrogen). RNA amplification was conducted for two rounds with the MessageAmp aRNA kit (Ambion), followed by biotin labeling with the BioArray high yield RNA transcription labeling kit (Enzo Life Sciences) and purification with the RNeasy mini kit (Qiagen). The resulting cRNA was hybridized to Mouse Gene 1.0 ST arrays (Affymetrix). These steps followed the Immunological Genome Project pipeline and were done at Expression Analysis. Data were normalized with the robust multiarray average algorithm implemented in Affymetrix Power Tools after prefiltration to remove unannotated probes and visualized on GenePattern Multiplot module.

We developed a T_{reg} cell signature index to estimate the global expression of genes of the T_{reg} cell signature in the transduced cells (**Fig. 3f**). First, we calculated the fraction of signature genes upregulated under various conditions (*F*); then we calculated the median value of the change in expression (-fold) relative to the expression in control transduced cells for all upregulated (Up) signature genes (*M*), and the T_{reg} cell Up signature index was established as follows: $I_{Up} = F^*M^*2$. As expected, $I_{Up} = 1$ for control samples. Similar calculation was done for downregulated (Down) genes of the T_{reg} cell signature, and a composite T_{reg} cell signature index was calculated as follows: $I = [I_{Up}+I_{Down}]/2$.

To distinguish the expression of transduced transcription factors from that of their endogenous counterparts, we used feature-level analysis of the 1.0 ST microarray data. The Affymetrix Mouse Gene 1.0 ST Array offers whole-transcript coverage, as each of the 28,853 genes is represented on the array by approximately 27 oligonucleotides ('features') spread across the full length of the gene. This characteristic allowed us to distinguish the expression of endogenous mouse transcription factors from the ectopic transcription factors, which were of human origin. First, nucleotide sequences of all the features (25-residue oligonucleotides) for one particular gene (such as *Foxp3*) were retrieved, and the sequence similarities between mouse and human were analyzed with the basic local alignment search tool of the National Center for Biotechnology Information. Features with substantial dissimilarity between mouse and human (more than 5 mismatches among 25 nucleotides) were considered mouse-specific probes, and their expression values were averaged and normalized to arrive at the values presented (**Fig. 4a** and **Supplementary Table 4**).

doi:10.1038/ni.2420

Immunoprecipitation and immunoblot analysis. Human embryonic kidney (HEK293) cells were doubly transfected with vectors for Flag-tagged Foxp3 plus another transcription factors (Satb1, Lef1, GATA-1 or Pbx1), were lysed on ice with hypotonic solution (10 mM HEPES, 1.5 mM MgCl2, 10 mM KCl and 0.05% NP-40–like IgePal Ca-630) supplemented with EDTA-free complete protease inhibitors (Roche). Nuclear pellets were subsequently treated with nuclear lysis buffer (20 mM HEPES, 300 mM NaCl, 20 mM KCl and EDTA-free complete protease inhibitor 'cocktail') and MNase (nuclease S7; Roche). Chromatin digestion was stopped by the addition of EDTA to a concentration of 5 mM, and post-nuclear supernatants were incubated overnight at 4 °C with Protein G Sepharose beads coupled to antibodies for immunoprecipitation (anti-Flag (M2; Sigma), anti-Foxp3 (FJK-16s; eBioscience), anti-GATA-1 (Ab28839; Abcam), anti-Lef1 (Ab124271; Abcam), anti-Satb1 (611182; BD) or control immunoglobulin G (rat IgG2a (eBR2a; eBioscience), mouse IgG2a (553454; BD) and rabbit IgG (sc628; Santa Cruz))), with constant rotation. Bound proteins were eluted by boiling, were separated by SDS-PAGE and were electrotransferred to PVDF membranes. After blockade of nonspecific binding (2 h in 5% milk and 0.02% Tween-20 in 1x PBS), blots were probed for 1 h at room temperature with antibodies for immunoblot analysis (anti-Foxp3 (FJK-16s; eBioscience), anti-GATA-1 (Ab28839; Abcam), anti-Lef1 (Ab124271; Abcam) and anti-Satb1 (611182; BD)).

ChIP-Seq. Mouse primary $CD4^+$ T cells, transduced and sorted as described above, were used in this assay. ChIP was done as described[50]. Cells ($\sim 1 \times 10^7$) were crosslinked with formaldehyde (11%). Cell lysates were sonicated on ice (eight cycles of 30 s at intervals of 60 s; Misonix), then were incubated with 10 μg antibody (antibody to total RNA polymerase II (sc-899; Santa Cruz), anti-Flag (M2; Sigma) or anti-GATA-1 (ab28839; Abcam) prebound to protein G–conjugated Dynal beads (Invitrogen). Immunoprecipitated DNA was purified and used for library construction with a ChIP-Seq DNA Sample Prep Kit for Illumina sequencing[50]. Sequences were aligned to the genome with Bowtie software (version 0.12.7) to National Center for Biotechnology Information Build 36 (UCSC mm9) of the mouse genome. Peaks of binding were 'called' with model-based analysis of ChIP-Seq analysis software (1.4.0rc2). The number of reads in each tag 'pileup' were first normalized to the total number of reads in the sample. For accurate comparison of the local tag densities in peak regions of the different samples (particularly for Foxp3 binding in samples transduced to express either Foxp3 alone or Foxp3 plus GATA-1), values were rescaled by a constant calculated from the integrated values of the noise in regions devoid of any Foxp3-binding peaks (seven regions ranging from 60 kb to 650 kb). This correction stemmed from the assumption that the experimental noise should be constant even when true signal (and hence the total number of reads) might be expected to vary between parallel samples and that a normalization factor calculated from the genome background amount allows appropriate compensation for variability in amplification during the construction of sequencing libraries.

Mathematical modeling. Information on mathematical modeling is available in the **Supplementary Note**.

Statistical analysis. Statistical significance was determined with the Kolmogorov-Smirnov test or the χ^2 test.

48. Hill, J.A. *et al.* Retinoic acid enhances Foxp3 induction indirectly by relieving inhibition from CD4+CD44hi Cells. *Immunity* **29**, 758–770 (2008).
49. Reich, M. *et al.* GenePattern 2.0. *Nat. Genet.* **38**, 500–501 (2006).
50. Rahl, P.B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).