HHS Public Access

Author manuscript

Nat Immunol. Author manuscript; available in PMC 2017 September 27.

Published in final edited form as:

Nat Immunol. 2017 May ; 18(5): 563–572. doi:10.1038/ni.3714.

Quality of TCR signaling encoded by differential enhancer affinities for the composite BATF-IRF4 transcription factor complex

Arifumi Iwata[1], **Vivek Durai**[1], **Roxane Tussiwand**[2], **Carlos G. Briseño**[1], **Xiaodi Wu**[1], **Gary E. Grajales-Reyes**[1], **Takeshi Egawa**[1], **Theresa L. Murphy**[1], and **Kenneth M. Murphy**[1,3]

[1]Department of Pathology and Immunology, Washington University in St. Louis, School of Medicine, St. Louis, MO 63110, USA [2]Department of Biomedicine, University of Basel, Mattenstrasse 28, CH - 4058 Basel Switzerland [3]Howard Hughes Medical Institute, Washington University in St. Louis, School of Medicine, St. Louis, MO 63110, USA

Abstract

Variable strengths of T cell receptor (TCR) signaling can produce divergent outcomes, but the mechanism remains obscure. The abundance of the transcription factor IRF4 increases with TCR signal strength, but how this would induce distinct types of responses is unclear. We compared T_H2 gene expression with BATF/IRF4 enhancer occupancy at varying strengths of TCR stimulation. BATF/IRF4-dependent genes clustered into distinct TCR-sensitivities. Enhancers exhibited a spectrum of occupancy by BATF/IRF4 ternary complex that correlated with TCR-sensitivity of gene expression. DNA sequences immediately flanking the previously defined AICE motif controlled the affinity for BATF/IRF4 for direct binding to DNA. ChIP-exo analysis allowed identification of a novel high-affinity AICE2 motif at a human SNP of *CTLA4* associated with resistance to autoimmunity. Thus, the affinity of different enhancers for the BATF-IRF4 complex may underlie divergent signaling outcomes in response to various strengths of TCR signaling.

Introduction

Strength of T cell receptor (TCR) signaling can influence thymocyte fate choice[1] and T cell effector outcome[2], but how signal strength controls different gene programs remains unclear[3]. Antigen dose can alter T_H1/T_H2 balance[4,5], T_H1/T_{FH} balance[6] and interleukin 10

Correspondence should be addressed to K.M.M. (kmurphy@wustl.edu). Phone 314-362-2009, Fax 314-747-4888.

Accession codes. GEO: microarray data, GSE85173; ChIP-seq and ChIP-exo data, GSE85172.

Author Contributions.

A.I., T.L.M. and K.M.M. designed the study; A.I., C.G.B, performed experiments related to cell sorting, culture and flow cytometry with advice from V.D., G.E.G.-R., R.T., and T.E.; A.I., and C.G.B. performed microarray experiments with advice from X.W., T.L.M. ; A.I. performed and analyzed ChIP-Seq experiments with advice from V.D. and T.E.; A.I., T.L.M. and K.M.M. wrote the manuscript with contributions from all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

(IL-10) production by T_H1 cells[7]. BCL-6 and BLIMP-1, supporting T_{FH} or T_H1 development, do show graded abundance at different TCR signal strengths but cannot explain all graded T cell responses[2], and how different strengths of TCR signaling regulates their differential abundance is unknown.

The transcription factor IRF4 may mediate some aspects of variable TCR signaling[3], including BLIMP-1 abundance[8]. IRF4 abundance increases in proportion to TCR-signal strength and correlates with T cell expansion and gene expression for metabolic and biosynthetic pathways[8–10]. IRF4 is required for T cell effector function[11], and for B cell development, class switch recombination and plasma cell differentiation[12]. IRF4 binds the DNA sequence GAAA but requires heterodimerization with other factors for high affinity binding. In B cells and T cells, IRF4 forms a complex with a heterodimer composed of basic leucine zipper ATF-like transcription factor (BATF) and Jun, which binds DNA at a specific sequence motif, the activation protein 1(AP-1)/IRF composite element, or AICE[13–16]. IRF4 is also recruited to ETS/IRF composite elements (EICE) through interactions with the ETS PU.1 and SpiB in B cells and DCs, but not T cells due to their low ETS abundance[11,17]. In plasma cells, which express low amounts of BATF, IRF4 is expressed at high levels and binds to interferon-sensitive response elements (ISREs)[18].

The BATF subfamily of AP-1factors includes BATF, BATF2 and BATF3, which all bind DNA as heterodimers with Jun factors[17]. BATF expression is restricted to the immune system and is required for differentiation T_H9, T_H17, T_{FH} cells[17,19] and for effector CD8 T cell differentiation and expansion[20]. *Batf*$^{-/-}$CD8$^+$ T cells produce increased interferon-γ (IFN-γ) and granzyme B, implying that BATF also regulates activation[20]. BATF is also required for the germinal center reaction and class switch recombination in B cells[21]. Another BATF member, BATF3, is expressed in DCs and is required for CD24$^+$ DC development[22]. BATF and BATF3 are expressed in different cell types but can compensate for each other when expressed in the same cells[13,17]. BATFs enable IRF4/8 dependent transcription by binding cooperatively to two variants of AICEs, AICE1 (TTTCnnnnTGASTCA) and AICE2 (GAAATGASTCA)[14,15,17]. BATF and IRF4 are both induced within four hours of TCR stimulation and so may initiate expression of many genes associated with activation and differentiation[23].

The role of BATF in T_H2 differentiation was unclear due to differing results[13,24,25]. We found BATF was not required for T_H2 development[13,21,25], but others reported impaired T_H2 development using distinct *Batf*$^{-/-}$ mice[24]. We later reported that *Batf3* expression compensated for loss of BATF in T_H2 development, maintaining IL-4 and IL-10, but not CTLA-4, expression, and that *Batf*$^{-/-}$*Batf3*$^{-/-}$ (*Batf1/3* DKO) T cells lacked IL-4, IL-10 and CTLA-4 expression[13]. This suggested that *Batf*-target genes included some that were sensitive to compensation by endogenous BATF3, but that varying conditions of activation[24] might influence the amount BATF3 or the sensitivity of target genes to compensation by BATF3. In either case, the basis for such differential sensitivity remained unclear.

Here, we first document clearly distinct sensitivities of several genes to compensation by BATF3 in *Batf*$^{-/-}$$T_H2$ cells. We show that enhancers controlling BATF-dependent TCR-inducible genes respond to different levels of BATF/IRF4. Genes highly sensitive to low

BATF levels can be compensated by endogenous BATF3 in *Batf*$^{-/-}$ T_H2 cells, while genes require higher levels of BATF are not. Using ChIP-seq and ChIP-exo analysis, we find that sensitivity of enhancers in these genes to BATF can be regulated by sequences surrounding AICEs that influence the affinity for the BATF/IRF4 ternary complex. ChIP-exo analysis helped to identify a novel AICE2 motif conferring high affinity for BATF/IRF4 that may control a single nucleotide polymorphism in the *CTLA4* locus known to decrease the incidence of autoimmune diseases[26,27].

RESULTS

GATA-3 and CTLA-4 respond to distinct signal strengths

We examined expression of BATF, IRF4, GATA-3 and CTLA-4 in T_H2 cells activated with a graded level of TCR signaling. IRF4 was induced in a gradual and uniform manner in proportion to TCR signal strength (Fig. 1a), consistent with previous reports[8–10]. BATF expression was also graded and had similar dose responses to TCR signaling as IRF4 (Fig. 1a). *Gata3* and *Ctla4* are known target genes of IRF4 and BATF[13,14,28], but were induced at different strengths of TCR signaling (Fig. 1b). GATA-3 was induced at low signal strength, while CTLA-4 was induced at higher signal strength (Fig. 1c). During secondary stimulation, a similar graded expression in response to TCR stimulation was observed for BATF and IRF4. GATA-3 expression remained high even at low TCR signal strength during secondary stimulation but was more sensitive to TCR signal strength compared to CTLA-4 (Supplementary Fig. 1a–c). Likewise, BATF expression was induced in a graded manner in proportion to peptide dose, with GATA-3 induction occurring at low peptide dose and CTLA-4 induction occurring only at higher peptide dose (Supplementary Fig. 1d–f).

The kinetics of gene expression showed similar patterns to the TCR dose response. BATF and IRF4 were induced on day 1 and accumulated over the next three days (Supplementary Fig. 2a). In contrast, GATA-3 was expressed at nearly maximal levels on day 2, but CTLA-4 was not fully expressed until day 3, correlating with high levels of BATF and IRF4 (Supplementary Fig. 2b–f). Thus, BATF and IRF4 expression showed a graded response to both the strength of TCR stimulation and length of activation.

Batf targets in T_H2 cells show differential compensation by *Batf3*

Expression of CTLA-4 in T_H2 cells was partially dependent on *Batf*, but was completely lacking in *Batf*$^{-/-}$*Batf3*$^{-/-}$ T_H2 cells (*Batf1/3* DKO T_H2 cells), even at a high dose of anti-CD3 in a secondary stimulation, as we reported[13] (Fig. 2a,b). In contrast, expression of GATA-3 in T_H2 cells was *Batf*-independent, and significantly reduced only in *Batf1/3* DKO T_H2 cells, similar to IL-10 (Fig. 2a,b). Thus, endogenous BATF3 could compensate for BATF for expression of GATA-3 and IL-10, but not for CTLA-4. This suggests that *Batf* target genes in T_H2 cells can have distinct sensitivities to the combined levels of BATF and BATF3.

BATF-dependent gene induction relies on interactions of IRF4 with several amino acid residues of BATF (H55, K63, 77) that face out from the leucine zipper (LZ)[13]. Mutations in three BATF amino acids, H55Q, K63D, and E77K (*Batf*-HKE) eliminates BATF

transcriptional activity, including IL-17 expression in T_H17 cells, class switch recombination in B cells, and $CD24^+$ DC development[13] (Supplementary Fig. 3a–c). We tested the activity of *Batf*-HKE for GATA-3, IL-10 and CTLA-4 expression in T_H2 cells. Retroviral expression of *Batf* or *Batf3* into *Batf1/3* DKO T_H2 cells fully restored GATA-3, IL-10 and CTLA-4 expression to similar levels as in wild-type T_H2 cells (Fig. 2c), as expected[13]. By contrast, expression of *Batf*-HKE into *Batf1/3* DKO T_H2 cells induced GATA-3 and IL-10 to intermediate levels and failed to induce CTLA-4 (Fig. 2c).

A similar pattern of dependence was observed when BATF levels were examined across a spectrum comparing WT, heterozygous and knockout alleles. GATA-3 and IL-10 expression in $Batf^{+/-}Batf3^{-/-}$ T_H2 cells, containing one functional *Batf* allele, was equal to that in $Batf^{+/+}Batf3^{-/-}$ T_H2 cells, containing two *Batf* alleles (Fig. 2d). In contrast, CTLA-4 expression was significantly lower in $Batf^{+/-}Batf3^{-/-}$ T_H2 cells compared to $Batf^{+/+}Batf3^{-/-}$ T_H2 cells (Fig. 2d). Finally, retroviral expression of *Batf* in wild-type T_H2 cells shifted the TCR dose response such that both GATA-3 and CTLA-4 were induced at a lower dose compared to control vector (Fig. 2e). In summary, these results show that GATA-3 and IL-10 are highly responsive to BATF, requiring only a low amount of total BATF and responding to weak BATF-IRF4 interaction provided by *Batf*-HKE. In contrast, CTLA-4 is less sensitive to BATF, requiring high amount of total BATF and responding only with strong a BATF-IRF4 interaction provided by wild-type *Batf*, but not *Batf*-HKE.

Batf targets show varying sensitivity to TCR signal strength

We compared global gene expression in wild-type and *Batf1/3* DKO T_H2 cells under weak TCR stimulation (10 ng/ml of anti-CD3ε on day 2, or 2 ng/ml of anti-CD3ε on day 4), or under strong TCR stimulation (10 ng/ml of anti-CD3ε on day 4). *Batf1/3* DKO T_H2 cells did not express *Ifng, Il4*, or *Il17a* at any dose of anti-CD3ε, indicating that they did not take on a T_H1, or T_H17 phenotype under T_H2 conditions (Supplementary Fig. 3d). With weak TCR stimulation, there were 30 or 68 genes expressed at least 3-fold more highly in WT compared to *Batf1/3* DKO; with strong TCR stimulation, this increased to 207 genes (Fig. 3a, red circles). Spearman's rank-order correlation of gene expression was greatest between samples of similar strength of TCR stimulation, and there was strong positive correlation between the two weak TCR stimulation conditions (Fig. 3b). These results indicate that distinct sets of genes are induced by low and high levels of TCR stimulation.

We carried out hierarchical clustering of genes based on the fold change of expression between wild-type and *Batf1/3* DKO T_H2 cells (Fig. 3c,d). Cluster I and cluster II contained genes with large expression changes between wild-type and *Batf1/3* DKO T_H2 cells with weak TCR stimulation, while cluster III contained genes that differed in expression only with strong TCR stimulation. Cluster IV contained genes that were increased in activated *Batf1/3* DKO T_H2 cells, including *Foxp3* (Supplementary Table 1). Cluster I and cluster II contained genes induced in wild-type T_H2 cells by weak TCR stimulation, including *Il10*, *Ccr4*, *Maf*, and *Prdm1*. Cluster III contained genes that were induced only by strong TCR stimulation in wild-type T_H2 cells, including *Ahr*, *Vdr*, *Ctla4, Hif1a* and *Socs3*. Within clusters, a spectrum of sensitivity to strength of activation was evident (Fig. 3d). In

summary, BATF-dependent genes induced T_H2 cells can exhibit a wide range of sensitivity to CR signal strengths.

BATF-IRF4 enhancer occupancy correlates with TCR sensitivity

We performed ChIP-seq for BATF in wild-type T_H2 cells and for IRF4 in wild-type and *Batf1/3* DKO T_H2 cells. To mimic low BATF-IRF4 interaction, we did ChIP-seq for IRF4 in *Batf1/3* DKO T_H2 cells stably reconstituted with *Batf*-HKE. BATF and IRF4 binding peaks were co-localized and had similar tag counts within selected genes from microarray clusters I and II, such as *Maf* and *Il10*, and within selected genes from cluster III, such as *Ahr* and *Vdr* (Fig. 4a). The co-localization is consistent with BATF and IRF4 binding as a complex, as we previously described for BATF/IRF4 ChIP-seq in T_H17 cells[13–16]. IRF4 did not bind DNA in *Batf1/3* DKO T_H2 cells, showing that IRF4 binding is BATF-dependent for these loci in T_H2 cells. Some IRF4 peaks were maintained in the presence of *Batf*-HKE in the *Maf* and *Il10* loci, suggesting these are high-affinity IRF4 binding sites not requiring strong BATF-IRF4 interaction. However, no IRF4 peaks were maintained with *Batf*-HKE in the *Ahr* and *Vdr* loci, suggesting that these binding sites are of low-affinity and require strong BATF-IRF4 interaction.

We performed additional ChIP-seq for IRF4 using the same three TCR stimulation conditions as used for global gene expression analysis. We merged all IRF4 binding peaks observed with any conditions of primary stimulation (see Methods) and performed a Spearman's rank correlation analysis for BATF and IRF4 binding intensities (tag counts per peak) for all seven ChIP-seq experiments. BATF and IRF4 binding were highly correlated in wild-type cells ($r=0.77$). IRF4 binding intensity was well correlated ($r>0.6$) between the twice-activated wild-type T_H2 cells and strong TCR stimulation conditions, and between the two weakly stimulated samples (Fig. 4b). However IRF4 binding intensities were more weakly correlated ($r<0.6$) between strong and weak TCR stimulation conditions (Fig. 4b). When experimental conditions were ordered by decreasing BATF expression, Spearman's rank correlation coefficient for IRF4 tag counts were highest between samples of similar BATF expression (Fig. 4b). Globally, IRF4 binding to enhancers correlates with sensitivity to TCR stimulation. With strong TCR stimulation, IRF4 tag counts of peaks in genes from clusters I and II (TSS ± 50 kb) were equal to those from cluster III (Fig. 4c). But with weak TCR stimulation, IRF4 tag counts of peak in genes from cluster III were lower than those from clusters I and II (Fig. 4c), in agreement with gene-specific differences in IRF4 binding for *Maf* and *Il10* (clusters I and II) and *Ahr* and *Vdr* (cluster III) (Fig. 4a). Specifically, IRF4 binding was similar between strong and weak TCR stimulation for the peaks in *Maf* and *Il10* (Fig. 4d). In contrast, IRF4 binding to peaks in *Ahr* and *Vdr* are present with strong TCR stimulation, but absent with weak TCR stimulation (Fig. 4d). The T_H2-related genes, *Il4*, *Il5*, *Il13*, and *Ctla4* fall into cluster III, and have IRF4 ChIP-seq binding consistent with low sensitivity to TCR stimulation (Supplementary Fig. 4). These results suggest that genes with high-affinity BATF/IRF4 enhancers can be induced at low TCR signal strength and tolerate weaker interactions between BATF and IRF4. In contrast, genes with low-affinity enhancers are induced at higher TCR signal strength and require full strong interaction between the BATF leucine zipper and IRF4.

To test whether compensation for BATF by BATF3 was through direct binding of BATF3, we performed BATF3 and IRF4 ChIP-seq in *Batf*$^{-/-}$ cells using conditions of strong TCR stimulation. BATF3 and IRF4 binding sites in *Batf*$^{-/-}$ cells coincided with BATF and IRF4 binding sites in wild-type cells in Cluster I/II and Cluster III genes, as well as T_H2 related genes (Supplementary Fig. 4c,d). Thus, BATF3 is sufficient to allow IRF4 binding on many sites in *Batf*$^{-/-}$ cells, on cluster I and II genes, and some sites in cluster III genes, in agreement with compensation for GATA-3 and IL10 expression by *Batf3* in *Batf*$^{-/-}$ T_H2 cells. Some BATF/IRF4 binding peaks in the *Ctla4* locus in WT cells were not bound by BATF3 and IRF4 in *Batf*$^{-/-}$ cells (Supplementary Fig. 4d,) perhaps explaining why *Batf3* does not compensate for *Batf* in CTLA4 expression. In summary, some BATF-dependent genes with high-affinity BATF/IRF4 enhancers can be expressed in *Batf*$^{-/-}$ T_H2 cells because they bind BATF3 even though it is expressed in lower amounts than BATF.

ChIP-exo reveals a high-sensitivity AICE motif

Some IRF4 peaks were present only with strong TCR stimulation (e.g., *Ahr* −108 kb peak), while others (e.g., *Ahr* −90 kb peak) were also present with weak TCR stimulation. To identify sequences that might discriminate high and low affinity peaks, we first classified peaks as high-, intermediate-, or low-affinity based on the number of conditions of TCR stimulation in which they occur (Supplementary Fig. 5a). Genes from gene clusters I and II contain a greater number of high-affinity peaks and fewer low-affinity peaks compared with genes from gene cluster III (Supplementary Fig. 5b). However, *de novo* motif analysis of IRF4 peaks of all categories identified enrichment for AP1, AICE1, AICE2 and ETS motifs, but found no significant differences in sequence motifs between categories (Supplementary Fig. 5c).

We directly tested whether nucleotides flanking an AICE could influence BATF-IRF4 binding affinity. We used four pairs of electrophoretic mobility shift assay (EMSA) probes with identical AICE sequences but different flanking regions (Supplementary Fig. 6a–d). We generated AICE probes that had identical core AICE sequences from four high-affinity peaks (AICE1: *Enpp6* -45 kb and *Bcor* +65 kb; AICE2: *Prdm1* +14 kb and *Ccr4* +8 kb) and from intermediate-affinity or low-affinity peaks (AICE1: *Ptchd3* -26 kb and *Mzt1* +230 kb; AICE2: *Ctla4* -33 kb and *Snrpe* +38 kb). These probes differed in the flanking genomic sequences surrounding identical AICE motifs. Probes from all high-affinity peak produced a strong BATF/IRF4 complex in EMSA, while probes from intermediate- or low-affinity peaks produced a substantially weaker complex (Supplementary Fig. 6e,f). Thus, the affinity of an AICE for BATF/IRF4 can be affected by DNA sequences outside the core AICE motif, but *de novo* motif analysis of ChIP-seq data was unable to resolve the motif further.

Often, more than one AICE motif was found within one IRF4 ChIP-seq peak (Fig. 5a, Supplementary Fig. 7a,b). Conceivably, a peak could arise on the basis of only one AICE actually being a true BATF/IRF4 binding site, with the other(s) being included on the basis of their proximity to this site. To optimize motif analysis to distinguish sequence requirements of high-affinity versus a low-affinity AICEs, we needed to identify which motif(s) within a peak were functional for BATF/IRF4 binding. Thus, we performed ChIP-exo sequencing[29] in WT T_H2 cells. As an example, a peak within intron 2 of *Rbm47* has two

AICE motif sites. However, ChIP-exo revealed that only one site is actually occupied (Fig. 5a–c). In other peaks, both AICEs identified from motif analysis of ChIP-seq were also found to bind in ChIP-exo analysis (Supplementary Fig. 7a–c).

To determine which predicted motifs contained ChIP-exo binding, we analyzed the mean ChIP-exo tag counts ±50 bp of all AICE1, AICE2 and ETS motifs that were predicted by *de novo* motif analysis or identified as consensus motifs. BATF and IRF4 binding by ChIP-exo was not enriched at ETS motif sites within IRF4 peaks, suggesting that BATF and IRF4 do not bind directly to ETS motifs (Supplementary Fig. 8a,b). However, in AICE1 and AICE2 sites, IRF4 ChIP-exo tag counts were significantly enriched at a position just upstream of the IRF binding consensus (TTTC/GAAA) and BATF ChIP-exo tag counts were enriched upstream of the AP-1 consensus (TGASTCA) as well as upstream of the IRF binding consensus (Fig. 5d, Supplementary Fig. 8a). Next, we filtered AICE1 and AICE2 motifs, either predicted by *de novo* motif analysis or by consensus motifs, to identify sites that had significantly higher ChIP-exo tag counts upstream of IRF4 and BATF binding sites than on flanking control bases. This analysis defined 1960 sites with ChIP-exo binding for AICE1 and 1506 sites for AICE2.

The AICE1 motif identified from the sites with ChIP-exo binding was nearly identical to the AICE1 motif determined by *de novo* motif analysis of ChIP-seq (Supplementary Fig. 8c,d). In contrast, the AICE2 motif identified from sites with ChIP-exo binding was enriched for T at a position four bp upstream of the IRF binding consensus compared with the AICE2 motif from ChIP-seq (Fig. 6 a,b). This T-containing AICE2 motif was enriched in genes from cluster I and II compared with genes from cluster III, consistent with its association with high-affinity peaks (Fig. 6 c,d). These results suggest that a T located at -4 bp in the AICE2 motif may increase the affinity of BATF-IRF4 DNA binding and could tune the sensitivity of target genes for activation by BATF-IRF4.

A novel AICE consensus motif has high affinity

We tested the functional effect of a T nucleotide at −4 bp in an AICE2 by analyzing three EMSA probes derived from peaks in *Ctla4* and *Bcl11b* that were IRF4 binding in ChIP-exo analysis (Fig. 7a). A probe derived from the low affinity peak in the -33 kb region of *Ctla4* was an AICE2 with C nucleotide at -4 bp position in the native genome; the native probe did not form a BATF-IRF4 complex in EMSA even at high concentrations of IRF4 (Fig. 7b). However, changing the -4 bp nucleotide from C to T led to formation of a strong BATF-IRF4 complex (Fig. 7b). Two other probes derived from high affinity peaks in the +30 kb region of *Ctla4* and the +33 kb region of *Bcl11b*, contain T at the -4 bp position in the native genome (Fig. 7a). Both probes formed a strong BATF-IRF4 complex by EMSA. However, changing the -4 bp T > C led to loss of the complex in all concentrations of IRF4 (Fig. 7b). These results show that the affinity of AICE2 motifs can be tuned by the DNA sequence at the -4 bp position, which lies outside the previously recognized AICE consensus.

Interestingly, an example of this high affinity AICE sequence is present within single nucleotide polymorphism (SNP) in the human *CTLA4* locus, in which a G > T change -38 kb, rs231735, was reported as a protective SNP for rheumatoid arthritis[26] and granulomatosis with polyangiitis[27]. Since these are T_H17-related diseases[30,31], we analyzed

CTLA-4 expression in murine T_H17, finding that it also BATF-dependent in these cells (Fig. 8a,b). The *CTLA4* SNP is located at -4 bp relative to an AICE2 (Fig. 8c). Conceivably, this SNP could increase the sensitivity of CTLA-4 expression to TCR stimulation by increasing BATF-IR$ binding. An EMSA probe based on the major G-allele of rs231735 did not form a BATF-IRF4 complex even with high concentrations of IRF4 (Fig. 8d). However, a probe based on the T allele of rs231735 formed a BATF-IRF4 complex, similar to T-AICE probes above.

We tested this SNP in a functional assay *in vivo* in T cells. We used a retroviral-based reporter[32] containing the *Ctla4* minimal promoter with or without insertion of an upstream 36 bp region containing either the G- or T-allele of rs231735-T. We tested the activity of these reporters after stable integration into T_H2 cells by activation of increasing doses of anti-CD3ε (Fig. 8f). The promoter lacking either region showed basal levels of stimulation. Promoter activity was increased somewhat by the rs231735-G allele, but was significantly increased further at all doses of anti-CD3ε by the rs231735-T allele enhancer (Fig. 8f). Notably, with the intermediate level of TCR stimulation (4.6 ng/ml of anti-CD3ε), activity was observed only with the T-allele. These results highlight the importance of a single nucleotide polymorphism for transcription factor binding to a site and for optimal, finely tuned gene expression.

DISCUSSION

Our results resolve a previous discrepancy related to the role of BATF in T_H2 development[24,25]. We reported a requirement for BATF in T_H17 and T_{FH} development, but not in T_H1 or T_H2 development[21,25], using *Batf*$^{-/-}$ mice targeting exons 1 and 2, but an independent study of *Batf*$^{-/-}$ mice targeting exon 3 reported an additional role in T_H2 development[24]. This discrepancy could result from differences in genetic backgrounds, since we had used 129SvEv mice[21,25] while the other study used C57BL/6 mice[24]. However, we found that T_H2 development could be abrogated in *Batf*$^{-/-}$ mice in which *Batf3* was also deleted[13]. This result showed that T_H2 development depends on BATF family activity and suggesting the discrepancy could result from differences in the amount of *Batf3* compensation between the studies. Here, we confirm this interpretation by showing directly that *Batf3* can induce T_H2 development when expressed in *Batf1/3* DKO T cells. Further, *Gata3* is a BATF-dependent gene induced by low TCR signal strength and is fully induced at 50% levels of BATF in a *Batf3*$^{-/-}$ background, and its expression can be compensated by *Batf*-HKE. For these reasons, GATA-3 appears to be selectively compensated by low levels of BATF3 expressed in T cells. Thus, differences in strength of TCR stimulation and the genetic backgrounds used could contribute to the apparent discrepancy in T_H2 development in *Batf*$^{-/-}$mice[24,25]. However, we have not addressed whether these findings also explain the published claims regarding antigen does and T_H1/T_H2 balance[4,5] or T_H1/T_{FH} balance[6].

Second, our results address the actions of graded IRF4 expression in CD8 T cell activation[9,33]. CD8 T cell expansion requires IRF4, which is induced in an graded manner in response to different strengths of TCR signaling[9]. We show that BATF is also induced in an graded manner coordinately with IRF4. At increasing levels of BATF/IRF4 expression, we identified a hierarchy of induced genes, with cluster I and II representing highly sensitive

responder targets and cluster III being low sensitivity targets. We also identified a hierarchy of IRF4 binding sites by ChIP-seq in response to increasing levels of BATF/IRF4 expression. Notably, global levels of IRF4 binding are different for these clusters at low, but not high, levels of signaling or at early times after activation. Further, the genes in cluster I and II contain more high-affinity IRF4 binding sites compared with genes in cluster III. A recent study showed that H3K27 acetylation at BATF binding sites[15] was increased as the strength of TCR signaling was increased[34]. Our results extend this correlation by showing that levels of IRF4 binding to BATF binding sites are also increased in proportions with the strength of TCR stimulation. However, we show that IRF4 binding does not increase uniformly across the genome, but occurs preferentially on genes that have high sensitivity to BATF. In summary, we show that analog expression of BATF and IRF4 increasing in proportion to the strength of TCR stimulation can induce a hierarchy of gene expression based on differing affinities of enhancer binding sties for the BATF/IRF4 complex.

Third, our results demonstrate how enhancers with AICEs can respond to different levels of BATF/IRF4. In essence, we show that flanking sequences surrounding a recognized AICE motif strongly influences affinity for the BATF/IRF4 complex in a chromatin-independent manner. Varying the affinity of enhancers for transcription factors is a recognized mechanism for generating graded responses to varying strength of signaling, as for example in morphogen-dependent expression of target genes controlled by motifs with different affinities for Dorsal[35,36]. Similarly, we identified enhancers with AICE1 or AICE2 of identical core sequence that have variable *in vivo* affinity by ChIP-seq analysis. When we compared these regions by EMSA, we found that flanking sequences determined overall binding in EMSA in the same pattern as observed *in vivo* by ChIP-seq. For example, the *Prdm1* and *Ctla4* genes each have an AICE2 motif identified by ChIP-seq with identical core sequences, GAAATGAGTCT. The site from *Prdm1* is high affinity, based on occupancy at low TCR signal strength, while the site in *Ctla4* is occupied only at high TCR signal strength. Notably, the EMSA complex formed by the *Prdm1* AICE2 region was much stronger than that of the *Ctla4* region, despite having identical core motifs, suggesting that *in vitro* binding may reflect *in vivo* affinity. As we previously provided data that BATF and IRF4 bind to AICE that function in T_H17 development[16], the ability of flanking sequences to modulate AICE affinity for BATF/IRF4 provides a mechanism for controlling the sensitivity of target genes to TCR signal strength. Flanking regions are known to regulate target transcription factor binding through DNA structure and GC content[37]. Immune dysregulation can be observed in CTLA-4 heterozygous germline mutations[38], and the human *CTLA4* SNP examined, which is associated with protection for autoimmunity[39]. The T located at -4 bp in the AICE2 motif of this SNP increased *in vitro* binding in EMSA and increased enhancer activity *in vivo*, so that perhaps an increase in CTLA-4 expression could act to represses autoimmunity[40], although further work is clearly needed.

METHODS

Mice

Wild-type, *Batf*$^{-/-}$, and *Batf*$^{-/-}$ *Batf3*$^{-/-}$ mice[22,25] of 129S6/SvEvTac background, DO11.10 and BALB/c mice were housed in our specific pathogen-free animal facility according to

institutional guidelines. All experiments were performed with sex-matched mice 6–12 weeks of age without randomization or blinding.

Antibodies and flow cytometry

Cells were stained at 4 °C in the presence of Fc Block (2.4G2; BioXcell) in flow cytometry buffer (0.5% BSA in PBS). The following antibodies were purchased from Becton Dickinson (BD): PE-conjugated anti-CTLA-4 (UC10-4F10-11); PE-Cy7–conjugated anti-CD25 (PC81); biotin-conjugated anti-CD8b (53-5.8), from BioLegend: Pacific blue-conjugated anti-CD4 (RM4-5); PerCP/Cy5.5–conjugated anti-Thy1.1 (OX-7); Brilliant Violet421–conjugated anti-human CD4 (OKT4); APC-conjugated anti-human CD4 (RPA-T4), biotin-conjugated anti CD45R/B220 (RA3-6B2), from eBioscience: APC-conjugated IL-10 (JESS-16E3); efluor660-conjugated anti-GATA-3 (TWAJ); PE-conjugated anti-IRF4 (3E4), biotin-conjugated anti-CD49b (DX5), from Tonbo Biosciences: FITC conjugated anti-CD4 (GK1.5), from Invitrogen, R-PE-conjugated anti-human CD4 (S3.5). For IL-10 and CTLA-4 staining, cells were fixed in 2% paraformaldehyde for 15 min at RT and permeabilized with 0.5% Saponin before staining. For BATF, GATA-3 and IRF4, cells were fixed and permeabilized with Foxp3 Staining Buffer Set (eBioscience) following manufacturers' instructions. Cells were analyzed on a FACSCanto II or FACS Aria Fusion and data were analyzed with FlowJo software (TreeStar).

Isolation and culture of $CD4^+$ T cells

All cells are cultured in IMDM supplemented with 10% fetal calf serum, 2 mM glutamine, 100 IU/ml penicillin, 100 μg/ml streptomycin, 1 mM sodium pyruvate, non-essential amino acids, and 55 μM β-mercaptoethanol. Spleen and lymph node cells were harvested, treated with ACK lysis buffer and passed through a 70 μm nylon filter.

For time course and TCR dose titration experiments (FACS, microarray and ChIP) naïve $CD4^+$ T cells were sorted as $CD4^+CD25^-CD44^-$ cells using spleen and lymph node cells after negative selection with biotinylated antibodies against B220, DX5 and CD8, streptavidin-nanobeads and MojoSort Magnetic Cell Separation system (Biolegend). Naïve $CD4^+$ T cells were activated with soluble anti-CD3ε (145-2C11, 1 μg/ml, BioXCell) at various concentrations cross linked by plate-bound anti-hamster IgG (MP biomedicals, 1:50) and anti-CD28 (37.51, BioXCell) (4 μg/ml) under T_H2 conditions (anti-IFN-γ 10 μg/ml (XMG1.2, BioXCell), anti-IL-12 10 μg/ml (Tosh), IL-4 10 ng/ml (Peprotech)) and were analyzed on the indicated day after primary activation.

For peptide dose titration, naïve $KJ126^+$ cells from DO11.10 mice were purified and activated with MACS purified $CD11c^+$ dendritic cells from spleens of BALB/c mice in the presence of ovalbumin (323–339) peptide under T_H2 conditions. Cells were expanded three-fold on day 3 and analyzed by FACS on day 4.

For experiments using secondary stimulations, $CD4^+$ T cells were isolated using Dynabeads FlowComp Mouse CD4 kit (Invitrogen) and were activated on anti-CD3ε (500A2 ascites diluted 1:400 to achieve maximal activation) and anti-CD28 (37.51, 4 μg/ml, BioXCell) coated plates under T_H2 conditions. On day 3, cells were diluted three-fold in fresh media on uncoated plates. On day 7, cells were re-stimulated under the same conditions or were re-

stimulated using soluble anti-CD3ε (145-2C11, 1 μg/ml, BioXCell) at various concentrations cross linked by plate-bound anti-hamster IgG (MP biomedicals, 1:50) and anti-CD28 (37.51, BioXCell) (4 μg/ml) under T_H2 conditions. On day 4 of the 2nd stimulation, cells were analyzed by flow cytometry for expression of GATA-3 and CTLA-4, or were activated by PMA and ionomycin in the presence of brefeldin A for 5 h for analysis of IL-10 expression, or for 2 h for ChIP.

EMSA

Oligonucleotide pairs were annealed to generate probes that were labeled with ^{32}P-dCTP using Klenow polymerase (Supplementary Table 2). HEK293FT cells were transiently transfected with retroviral vectors for *Batf*, *JunB* or *Irf4* using TransIT-LT1. After 48 h, cells were lysed with buffer A (10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl) containing 0.2% NP40 and protease inhibitors. Nuclei were pelleted, resuspended in buffer C (20 mM HEPES-KOH pH 7.9, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2mM EDTA, 25% glycerol), and centrifuged to obtain nuclear extracts[41]. EMSA was essentially as described[42] using combinations of nuclear extracts from cells transfected with *Batf, JunB* and *Irf4* (up to 1.5 μg total), 0.25 μg poly dI-dC (Sigma) and ^{32}P-labeled probes in 10 μl binding reactions for 20 min at 4 °C. Reactions were separated on 4–7%T 3.3%C polyacrylamide mini-gels in 0.4×TBE for 50 min at 250 V and 4 °C and analyzed by autoradiography.

Retroviral analysis

Retroviral vectors were transfected into Plat-E cells with TransIT-LT1 (Mirus Bio) and viral supernatants were collected 2 days later. On day 1 after activation of $CD4^+$ T cells, culture media was replaced with supernatants of transfected packaging cells containing 6 μg/ml polybrene. Cells were transduced by centrifuging at room temperature for 90 min at 1170*g*. Viral supernatant was replaced by T_H2 culture media.

Enhancer elements were cloned into a retroviral reporter with additional insertion of polyA sequence upstream of enhancer sites (hCD4 pA GFP RV)[32] (Supplementary Table 3). For analysis, we used integrated MFI, which combines the metrics of frequency and MFI as a measure of total functional response[43,44].

For ChIP-seq analysis of retrovirally transduced cells, infected cells were sorted on day 6 after primary activation, re-stimulated on day 7 and harvested on day 11 following PMA/ionomycin activation for 2 h.

Expression microarray analysis

Total RNA was extracted using RNAqueous-Micro Kit (Ambion) and was amplified with the Ovation Pico WTA Sytem (NuGEN) and hybridized to GeneChip Mouse Gene 1.0 ST microarrays (Affymetrix). Data were normalized by robust multiarray average summarization and quartile normalization with ArrayStar software (DNASTAR). The log_2-transformed data were imported into the software of the R project (version 3.2.3). Differential expression analyses were performed using limma package of R and *P*-value were corrected by the Benjamini-Hochberg procedure[45]. Spearman's rank correlation

coefficient plot was generated by R. Hierarchical clustering was performed with Euclidean distance and Ward clustering using of R.

ChIP-seq

ChIP was performed as described[32] with minor modifications. Briefly, 10^7 activated $CD4^+$ T cells were collected, crosslinked with 1% formaldehyde at room temperature for 8 min, quenched with 1.25 M glycine and washed twice with PBS. Pellets were 'flash frozen' for storage at −80 °C. Chromatin was sonicated for 24 cycles of 20s on and 50s off per cycle with a Vivra-Cell VCX130PB and CV188 sonicators (Sonics & Material) in lysis buffer (10 mM Tris pH?, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Sodium Deoxycholate, 0.5% N-lauroylsarcosine). Chromatin was immunoprecipitated overnight at 4 °C with Dynabeads Protein A or G (Invitrogen), that had been pre-incubated with 5 μg of antibody: anti-IRF4 (sc-6059X; Santa Cruz Biotechnology), rabbit anti-BATF[25], or rabbit anti-BATF3[13]. Beads containing protein-DNA complexes were washed two times with RIPA buffer (10 mM Tris pH?, 140 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton-X), two times with RIPA buffer + 0.3 M NaCl, two times with LiCl buffer (0.21 M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate), two times with Tris-EDTA/Triton-X buffer (10 mM Tris pH?, 1 mM EDTA, 0.2% Triton-X), and twice with Tris-EDTA. DNA fragments were eluted and reverse-crosslinked by incubation for 5 h at 65 °C in Tris-EDTA pH? with 0.3% SDS, 1 mg/ml Proteinase K (New England Biolabs). DNA was purified by phenol-chloroform extraction followed by ethanol precipitation. Libraries for ChIP-seq were prepared from 1 ng of ChIPed DNA with a ThruPLEX-FD kit (Rubicon Genomics) and were sequenced with an Illumina HiSeq 2500 as single 'reads' extending 50 bases.

Computational analysis for ChIP-seq

ChIP-seq data sets were aligned to the mouse genome (NCBI37/mm9 assembly) by Bowtie software (version 1.1.1)[46] with the following parameters: –sam –best -p4 -m 1 –chunkmbs 8000. Uniquely mapped reads were masked with Samtools[47] with blacklist of the ENCODE project[48] and the RepeatMasker program (which screens for interspersed repeats and low complexity) in the UCSC Genome Browser. Duplicated reads are discarded using 'make tag directory' of Homer software package[49] with the parameter -tbp 1. Data were visualized with the 'makeUCSCfile' of Homer. Peaks from individual conditions were identified with 'findPeaks' of Homer with a 300 bp window and Poisson *P*-value < 1×10^{-10} and with normalized tag counts four-fold more than control (Input sample). IRF4 peaks from individual of ChIP-seq experiments at primary stimulation conditions (10 ng/ml and 2 ng/ml on day 4 and 10 ng/ml on day 2) were merged using 'mergePeaks' of Homer and were centered and trimmed to 300 bp. Tag counts per peak were calculated with 'annotatePeaks.pl' of Homer. Spearman's rank correlation coefficient of tag counts on merged IRF4 peaks between each ChIP-seq experiment were performed by R. Peak related indicated genes were extracted from merged IRF4 peaks within TSS ± 50 kb using 'intersectBed' of BedTools package[50]. High-affinity peaks were defined as merged IRF4 peaks with tag counts with more than a threshold of Poisson *P*-value 1×10^{-10} in all three experiments (10 ng/ml and 2 ng/ml on day 4 and 10 ng/ml on day 2). Intermediate-affinity peaks were defined as merged IRF4 peaks with tag counts more than threshold in two of

three experiments, and low-affinity peaks have tag counts more than threshold in one of three experiments. The heatmap of binding intensity was generated using 'annotatePeaks.pl' of Homer and R. *De novo* motif analysis was performed from the top 3000 peaks, which were ranked by sum of tag counts per peak of the three experiments (10 ng/ml and 2 ng/ml on day 4 and 10 ng/ml on day 2), using 'findMotifsGenome.pl' of Homer with 150 bp window. Putative motif loci of motifs from each category (high-affinity, intermediate-affinity and low-affinity) were extracted from merged IRF4 peaks with *de novo* motifs using 'annotatePeaks.pl', length of motifs were adjusted and motifs were merged to one bed file using 'intersectBed'. Motif logos were generated by the 'seqLogo' package of R.

ChIP-exo

ChIP-exo was performed as described[51] with minor modifications. Briefly, ChIPed DNA-antibody-beads complexes from 30×10^6 cells per experiment were washed with RIPA buffer six times, and Tris-EDTA pH? twice. The ChIP-exo libraries were made by the following enzymatic reactions with four washes between each reaction; End polishing by T4 DNA polymerase, Klenow DNA polymerase, T4 polynucleotide kinase, ligation of P7 exo-adaptor by T4 DNA ligase, nick repair by Phi29 DNA polymerase, exonuclease reaction by lambda exonuclease and cleaning-up single strand DNA by RecJf exonuclease. DNA was eluted and reverse crosslinked as above. The libraries were generated by P7 primer extension with Phi29 DNA polymerase, ligation of P5 exo-adaptor with T4 DNA ligase, PCR amplification with Phusion polymerase for 12 cycles and were cleaned up by AMPure XP.

Computional analysis for ChIP-exo

ChIP-exo data sets were aligned and masked as above. We kept duplicated tags and shrunk 50 bp of reads to first 1 bp of 5′ position for further analysis. For visualizing, four replicates of BATF and two replicates of IRF4 were combined and normalized to 10 million tags. Bedgraphs were generated by 'genomeCoverageBed' of BedTools. ChIP-exo tag counts per base of motif ±50 bp were measured by 'coverageBed' of BedTools and were visualized by R after normalization to 10 million tags per experiments. We chose putative narrow exonuclease stopped position (exo binding region) around motifs and non-binding region (control region) based on the mean ChIP-exo tag counts around consensus motifs and motifs predicted by ChIP-seq (Fig. 5d, Supplementary Fig. 6a). We applied 20 bp of exo-binding region and control regions for BATF and 10 bp for IRF4. We defined the motif sites with BATF and IRF4 ChIP-exo binding that satisfied the following criteria; threshold: total tag counts on target region of four experiments for BATF or two experiments for IRF4 were more than the threshold (Poisson *P*-value 1×10^{-6}); significance between exo binding region and control region: differential analysis between $\log_2$-transformed ChIP-exo tag counts on exo-binding region (8 regions for BATF, 4 regions for IRF4) and ChIP-exo tag counts on control regions (16 regions for BATF, 8 regions for IRF4) in each experiment were performed by Welch's *t*-test with Storey's correction[52] (*P*-value < 0.05); fold change: the mean of tag counts on exo-binding region were more than two fold higher than the mean of tag counts on control region. Motif logos were generated by the 'seqLogo' package of R.

Statistical analysis

All statistical analyses were performed using Prism (GraphPad Software) or R.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Alvin J. Siteman Cancer Center at Washington University School of Medicine for use of the Center for Biomedical Informatics and Multiplex Gene Analysis Genechip Core Facility. This work was supported by the Howard Hughes Medical Institute (K.M.M.), the US National Institutes of Health (RO1 AI097244-01A1 to T.E., F30DK108498 to V.D., 1F31CA189491-01 to G.E.G.-R.).

References

1. Li MO, Rudensky AY. T cell receptor signalling in the control of regulatory T cell differentiation and function. Nat Rev Immunol. 2016; 16:220–233. [PubMed: 27026074]
2. Tubo NJ, Jenkins MK. TCR signal quantity and quality in CD4+ T cell differentiation. Trends Immunol. 2014; 35:591–596. [PubMed: 25457838]
3. Mayya V, Dustin ML. What Scales the T Cell Response? Trends Immunol. 2016 vol? Pages??
4. Constant S, et al. Extent of T cell receptor ligation can determine the functional differentiation of naive CD4+ T cells. J Exp Med. 1995; 182:1591–1596. [PubMed: 7595230]
5. Hosken NA, et al. The effect of antigen dose on CD4+ T helper cell phenotype development in a T cell receptor-alpha beta-transgenic model. Journal of Experimental Medicine. 1995; 182:1579–1584. [PubMed: 7595228]
6. Tubo NJ, et al. Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection. Cell. 2013; 153:785–796. [PubMed: 23663778]
7. Saraiva M, et al. Interleukin-10 production by Th1 cells requires interleukin-12-induced STAT4 transcription factor and ERK MAP kinase activation by high antigen dose. Immunity. 2009; 31:209–219. [PubMed: 19646904]
8. Man K, et al. The transcription factor IRF4 is essential for TCR affinity-mediated metabolic programming and clonal expansion of T cells. Nat Immunol. 2013; 14:1155–1165. [PubMed: 24056747]
9. Nayar R, et al. Graded levels of IRF4 regulate CD8+ T cell differentiation and expansion, but not attrition, in response to acute virus infection. The Journal of Immunology. 2014; 192:5881–5893. [PubMed: 24835398]
10. Yao S, et al. Interferon regulatory factor 4 sustains CD8(+) T cell expansion and effector differentiation. Immunity. 2013; 39:833–845. [PubMed: 24211184]
11. Huber M, Lohoff M. IRF4 at the crossroads of effector T-cell fate decision. Eur J Immunol. 2014; 44:1886–1895. [PubMed: 24782159]
12. De Silva NS, et al. The diverse roles of IRF4 in late germinal center B-cell differentiation. Immunol Rev. 2012; 247:73–92. [PubMed: 22500833]
13. Tussiwand R, et al. Compensatory dendritic cell development mediated by BATF-IRF interactions. Nature. 2012; 490:502–507. [PubMed: 22992524]
14. Glasmacher E, et al. A Genomic Regulatory Element That Directs Assembly and Function of Immune-Specific AP-1-IRF Complexes. Science. 2012; 338:975–980. [PubMed: 22983707]
15. Li P, et al. BATF-JUN is critical for IRF4-mediated transcription in T cells. Nature. 2012; 490:543–546. [PubMed: 22992523]
16. Ciofani M, et al. A validated regulatory network for th17 cell specification. Cell. 2012; 151:289–303. [PubMed: 23021777]
17. Murphy TL, Tussiwand R, Murphy KM. Specificity through cooperation: BATF-IRF interactions control immune-regulatory networks. Nat Rev Immunol. 2013; 13:499–509. [PubMed: 23787991]
18. Ochiai K, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. Immunity. 2013; 38:918–929. [PubMed: 23684984]

19. Jabeen R, et al. Th9 cell development requires a BATF-regulated transcriptional network. J Clin Invest. 2013; 123:4641–4653. [PubMed: 24216482]
20. Kurachi M, et al. The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells. Nat Immunol. 2014; 15:373–383. [PubMed: 24584090]
21. Ise W, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. Nat Immunol. 2011; 12:536–543. [PubMed: 21572431]
22. Hildner K, et al. Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. Science. 2008; 322:1097–1100. [PubMed: 19008445]
23. Yosef N, et al. Dynamic regulatory network controlling TH17 cell differentiation. Nature. 2013; 496:461–468. [PubMed: 23467089]
24. Betz BC, et al. Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. Journal of Experimental Medicine. 2010; 207:933–942. [PubMed: 20421391]
25. Schraml BU, et al. The AP-1 transcription factor Batf controls T(H)17 differentiation. Nature. 2009; 460:405–409. [PubMed: 19578362]
26. Gregersen PK, et al. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat Genet. 2009; 41:820–823. [PubMed: 19503088]
27. Chung SA, et al. Meta-analysis of genetic polymorphisms in granulomatosis with polyangiitis (Wegener's) reveals shared susceptibility loci with rheumatoid arthritis. Arthritis Rheum. 2012; 64:3463–3471. [PubMed: 22508400]
28. Lohoff M, et al. Dysregulated T helper cell differentiation in the absence of interferon regulatory factor 4. Proc Natl Acad Sci USA. 2002; 99:11808–11812. [PubMed: 12189207]
29. Rhee HS, Pugh BF. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell. 2011; 147:1408–1419. [PubMed: 22153082]
30. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. N Engl J Med. 2011; 365:2205–2219. [PubMed: 22150039]
31. Chen M, Kallenberg CG. ANCA-associated vasculitides–advances in pathogenesis and treatment. Nat Rev Rheumatol. 2010; 6:653–664. [PubMed: 20924413]
32. Grajales-Reyes GE, et al. Batf3 maintains autoactivation of Irf8 for commitment of a CD8alpha(+) conventional DC clonogenic progenitor. Nat Immunol. 2015; 16:708–717. [PubMed: 26054719]
33. Nayar R, et al. IRF4 Regulates the Ratio of T-Bet to Eomesodermin in CD8+ T Cells Responding to Persistent LCMV Infection. PLoS One. 2015; 10:e0144826. [PubMed: 26714260]
34. Allison KA, et al. Affinity and dose of TCR engagement yield proportional enhancer and gene activity in CD4+ T cells. Elife. 2016; 5 doi or article identifier.
35. Papatsenko D, Levine M. Quantitative analysis of binding motifs mediating diverse spatial readouts of the Dorsal gradient in the Drosophila embryo. Proc Natl Acad Sci USA. 2005; 102:4966–4971. [PubMed: 15795372]
36. Ashe HL, Briscoe J. The interpretation of morphogen gradients. Development. 2006; 133:385–394. [PubMed: 16410409]
37. Levo M, Segal E. In pursuit of design principles of regulatory sequences. Nat Rev Genet. 2014; 15:453–468. [PubMed: 24913666]
38. Kuehn HS, et al. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. Science. 2014; 345:1623–1627. [PubMed: 25213377]
39. Romo-Tena J, Gomez-Martin D, Alcocer-Varela J. CTLA-4 and autoimmunity: new insights into the dual regulator of tolerance. Autoimmun Rev. 2013; 12:1171–1176. [PubMed: 23851140]

Methods References

40. Lenardo M, Lo B, Lucas CL. Genomics of Immune Diseases and New Therapies. Annu Rev Immunol. 2016; 34:121–149. [PubMed: 26735698]
41. Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 1983; 11:1475–1489. [PubMed: 6828386]

42. Szabo SJ, et al. Identification of cis-acting regulatory elements controlling interleukin-4 gene expression in T cells: roles for NF-Y and NF-ATc [published erratum appears in Mol Cell Biol 1993 Sep;13(9):5928]. Molecular & Cellular Biology. 1993; 13:4793–4805. [PubMed: 8336717]
43. Darrah PA, et al. Multifunctional TH1 cells define a correlate of vaccine-mediated protection against Leishmania major. Nat Med. 2007; 13:843–850. [PubMed: 17558415]
44. Shooshtari P, et al. Correlation analysis of intracellular and secreted cytokines via the generalized integrated mean fluorescence intensity. Cytometry A. 2010; 77:873–880. [PubMed: 20629196]
45. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47–??. [PubMed: 25605792]
46. Langmead B, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25–??. [PubMed: 19261174]
47. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]
48. ENCODE Project Consortium An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
49. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010; 38:576–589. [PubMed: 20513432]
50. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]
51. Serandour AA, et al. Development of an Illumina-based ChIP-exonuclease method provides insight into FoxA1-DNA binding properties. Genome Biol. 2013; 14:R147–??. [PubMed: 24373287]
52. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci USA. 2003; 100:9440–9445. [PubMed: 12883005]

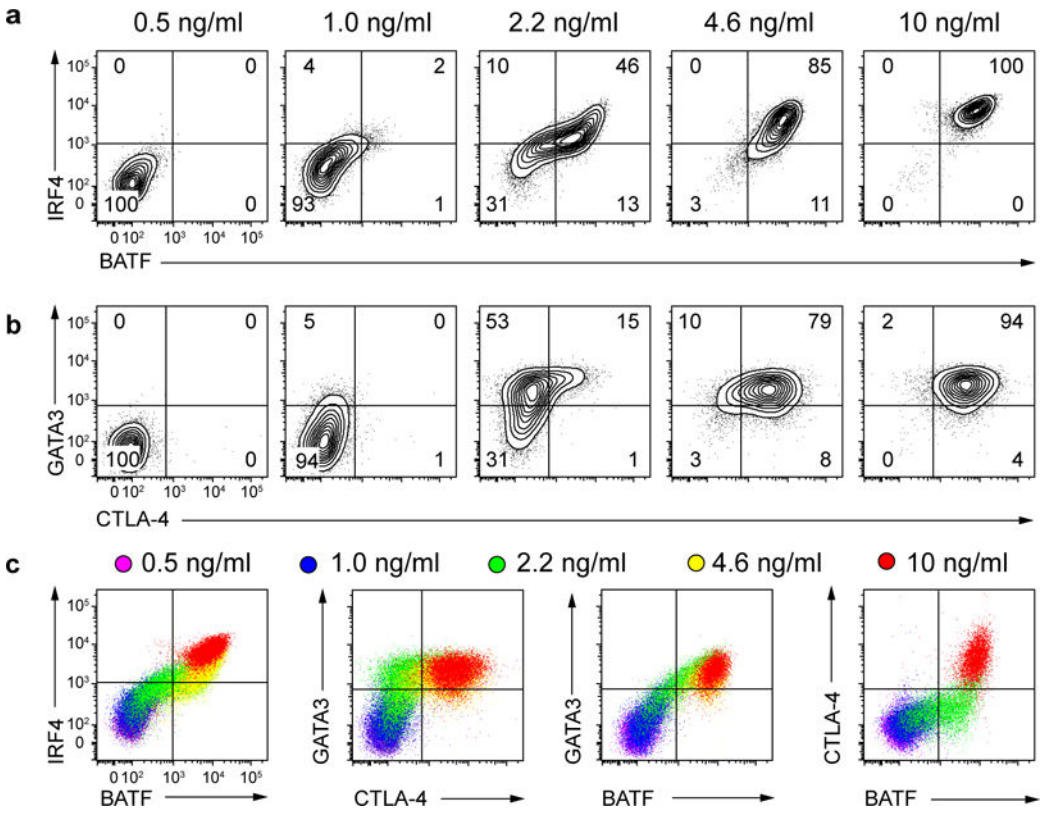

Figure 1.

GATA-3 and CTLA-4 are differentially sensitive to graded expression of BATF and IRF4 in T_H2 cells following increasing strength of TCR stimulation. (**a**) Flow cytometry analyzing IRF4 and BATF expression on day 4 of primary activation in WT $CD4^+$ T cells cultured under T_H2 conditions (anti-IFN-γ, anti-IL-12 and IL-4) with anti-CD28 and the indicated concentration of anti-CD3ε crosslinked by plate-bound anti-hamster IgG. Numbers indicate the percentage of live $CD4^+$ cells in each quadrant. (**b**) Flow cytometry analyzing GATA-3 and CTLA-4 expression on day 4 in WT $CD4^+$ T cells cultured as in (**a**). (**c**) Overlays of flow cytometry data from (**a**) and (**b**) showing expression of the indicated proteins at various doses of plate bound anti-CD3ε. Data are representative of two experiments using five biological replicates (**a**–**c**).

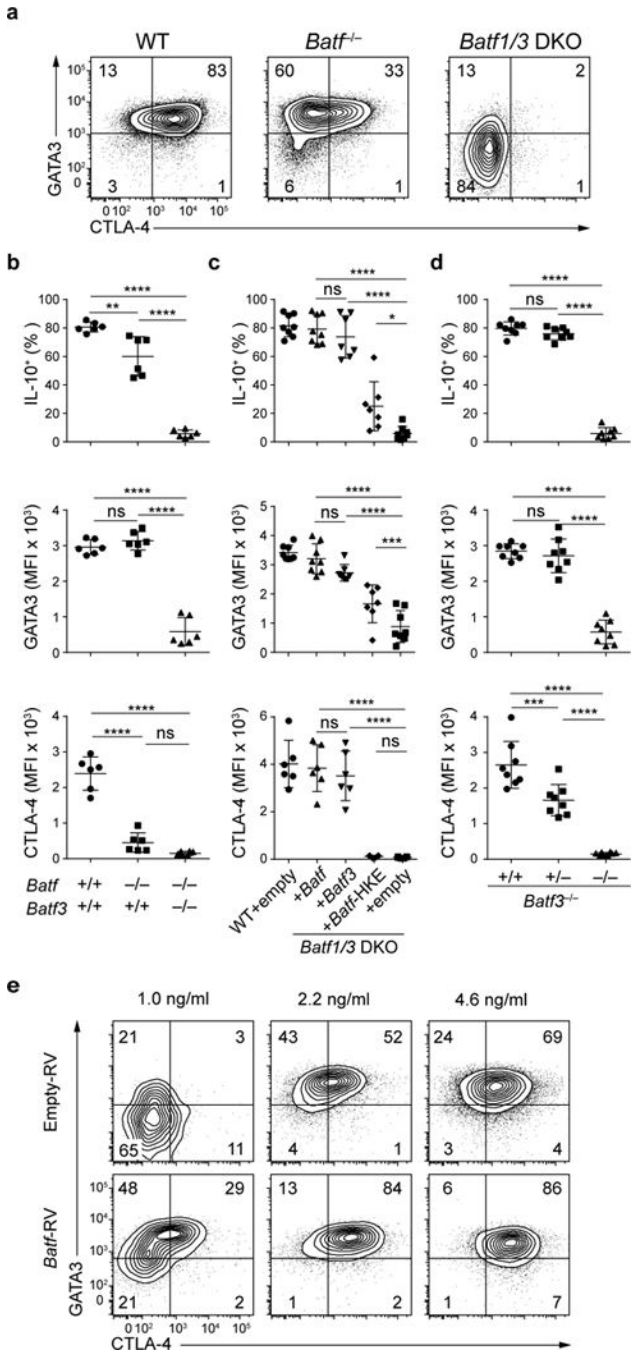

Figure 2.

GATA3, IL-10 and CTLA-4 are differentially sensitive to the level of BATF. (**a**) Flow cytometry analyzing GATA-3 and CTLA-4 expression in T_H2 cells from WT, *Batf*$^{-/-}$, and *Batf*$^{-/-}$ *Batf3*$^{-/-}$ double knock out (*Batf1/3* DKO) mice on day 4 after secondary stimulation with 500A2 ascites (1:400) and anti-CD28 under T_H2 conditions. Numbers indicate the percentage of live CD4$^+$ cells in each quadrant. (**b**) Percentage of IL-10 positive cells and MFI of GATA-3 and CTLA-4 in cells from the indicated genotype cultured as in (**a**). Data are pooled from two experiments with n=6 for each genotype. (**c**) Percentage of IL-10

positive cells and MFI of GATA-3 and CTLA-4 in T_H2 cells from WT or *Batf1/3* DKO mice infected with empty retrovirus (empty) or retrovirus expressing *Batf*, *Batf3*, or *Batf*-HKE. Cells were analyzed on day 4 after secondary stimulation as in (**a**) and gated on human CD4 as the marker of retroviral infection. n=6 (CTLA-4), n=7(DKO+*Batf3*, DKO+*Batf*-HKE in IL-10 or GATA3), n=8 (WT+empty, DKO+empty, DKO+*Batf* in IL-10 or GATA3), pooled from two (CTLA-4) or four (IL-10, GATA3) experiments. (**d**) Percentage of IL-10 positive cells and MFI of GATA-3 and CTLA-4 in T_H2 cells from *Batf3*$^{-/-}$ mice that were *Batf*$^{+/+}$, *Batf*$^{+/-}$ or *Batf*$^{-/-}$. Cells were analyzed on day 4 after secondary stimulation as in (**a**). n=8 in each genotype. Data are pooled from three experiments. One-way ANOVA with Tukey's (**b**,**d**) or Sidak's (**c**) multiple comparison test. **p < 0.05; **p < 0.01; ***p < 0.001; ****p<0.00001; ns, not significant. (**e**) Flow cytometry analyzing GATA-3 and CTLA-4 expression on day 4 of primary activation of WT $CD4^+$ T cells cultured under T_H2 conditions with anti-CD28 and the indicated concentration of anti-CD3ε crosslinked by plate-bound anti-hamster IgG. Cells were infected on day 1 with empty retrovirus or retrovirus expressing *Batf*. Numbers indicate percentage of infected T cells in each quadrant. Data are representative of two experiments.

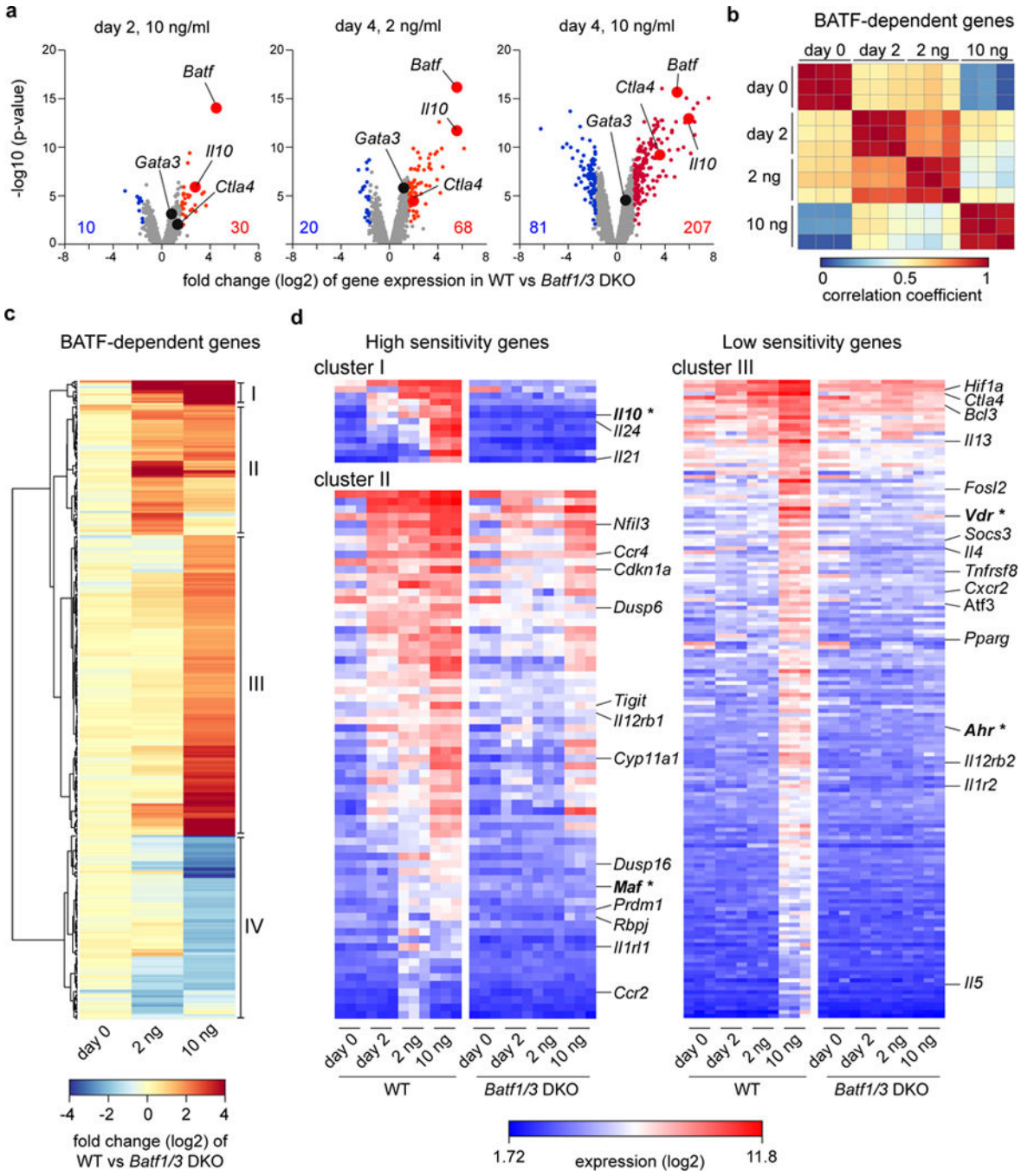

Figure 3.

Batf-dependent genes display a spectrum of sensitivity to the strength of TCR signaling. (**a**) WT and *Batf1/3* DKO T cells activated under T_H2-inducing conditions using the indicated concentration of anti-CD3ε antibody were harvested on day 2 or 4 after primary activation. P-values were corrected by Benjamini-Hochberg procedure. Colors indicate p-value < 0.01 and expression more than threefold higher (red) or lower (blue) in WT than *Batf1/3* DKO T_H2 cells. (**b**) Spearman's rank correlation coefficient for gene expression (log_2) of BATF-dependent genes identified in (**a**) in unstimulated WT $CD4^+$ T cells (day 0) and WT $CD4^+$ T

cells activated as in (**a**). (**c**) Hierarchical clustering of BATF-dependent genes identified in (**a**) using fold change of unstimulated WT $CD4^+$ T cells (day 0) and WT $CD4^+$ T cells activated by the indicated concentration of anti-CD3ε for 4 days (primary activation). (**d**) Expression (log_2) in WT and *Batf1/3* DKO T_H2 of genes from clusters I, II and III determined in (**c**).

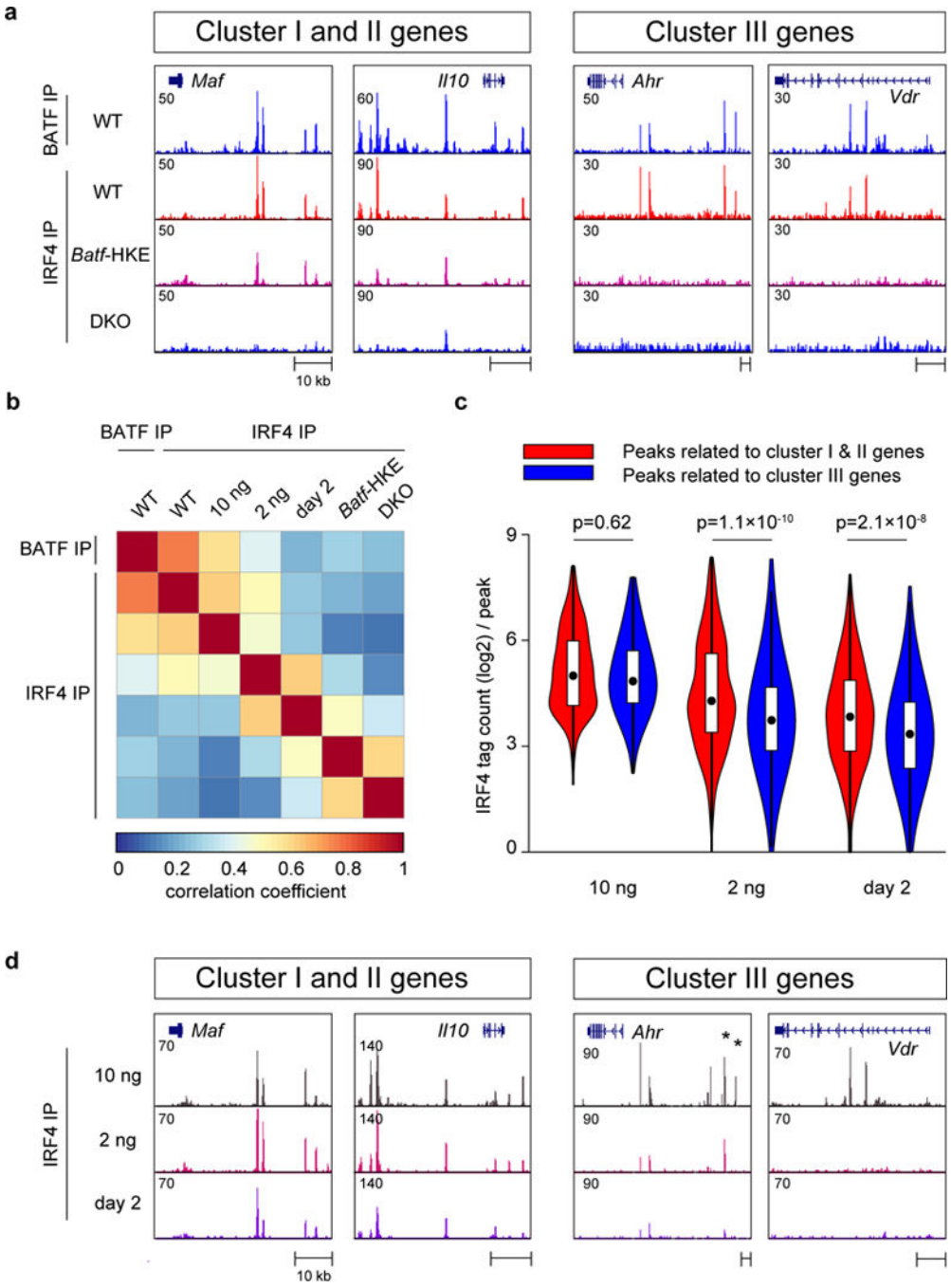

Figure 4.

BATF and IRF4 ChIP-seq binding correlates with target sensitivity to TCR signaling. (**a**) ChIP-seq for BATF in WT T_H2 cells and IRF4 in WT and *Batf1/3* DKO T_H2 cells (DKO) and in *Batf1/3* DKO T_H2 cells reconstituted with retroviral *Batf*-HKE (*Batf*-HKE). Cells were prepared for ChIP after PMA/ionomycin stimulation on day 4 of secondary stimulation. (**b**) Spearman's rank correlation coefficient analysis of ChIP-seq tag counts (log_2) obtained from merged IRF4 peaks after primary stimulation at various doses of anti-TCR on day 2 and day 4 (10 ng/ml day 4, 2.2 ng/ml day 4, 10 ng/ml day 2) comparing ChIP-

seq experiments from (**a)** (BATF IP WT, IRF4 IP WT, IRF4 IP *Batf*-HKE, and IRF4 IP DKO) and from primary stimulations (IRF4 IP 10 ng, IRF4 IP 2 ng, and IRF4 IP day 2). (**c**) Violin plot for IRF4 tag counts (log_2) per merged IRF4 peak ±50 kb from the transcription start site of cluster I & II genes (red) or cluster III genes (blue) in the indicated ChIP-seq experiments. Mann-Whitney U test with Bonferroni correction. (**d**) IRF4 ChIP-seq in WT T_H2 cells cultured with 10 ng/ml of anti-CD3ε for 4 days, 2.2 ng/ml of anti-CD3ε for 4 days, or 10 ng/ml of anti-CD3ε for 2 days.

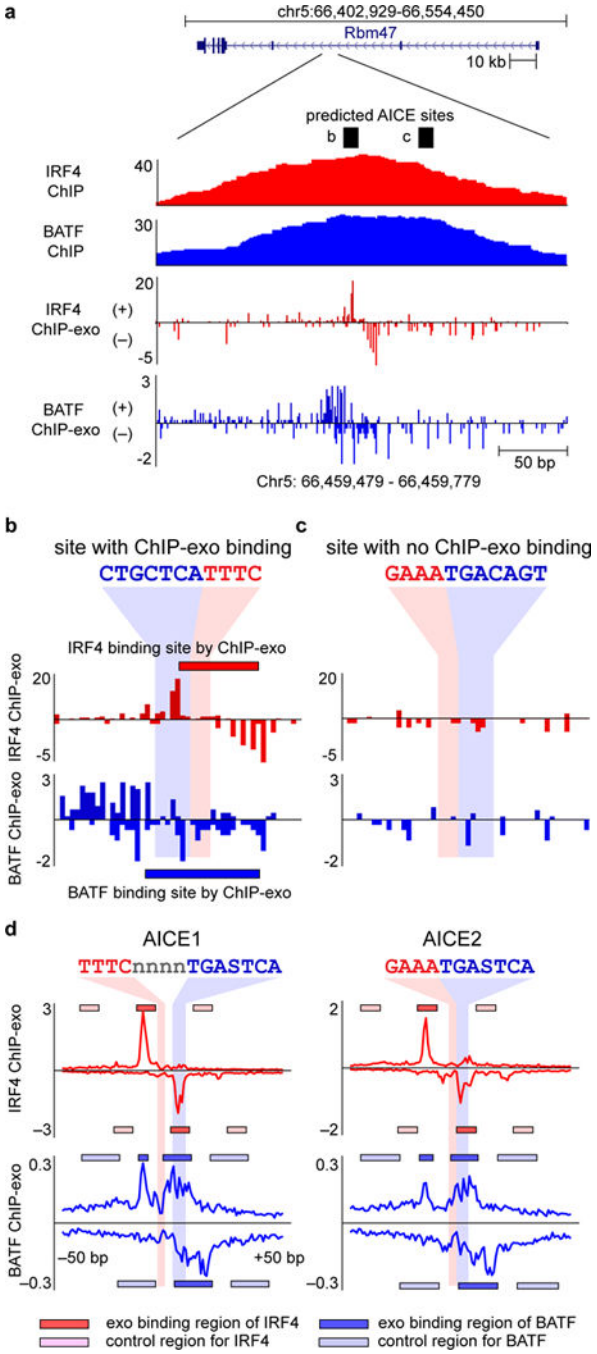

Figure 5.

ChIP-exo reveals narrow BATF and IRF4 binding in the sites with AICE motifs. (**a**) ChIP-seq analysis and ChIP-exo analysis of BATF and IRF4 binding in WT T_H2 cells; black bars: predicted AICE sites from *de novo* motif analysis. Shown is the genomic region containing *Rbm47*. Genomic coordinates refer to NCBI37/mm9. (**b**) Example of a site from (**a**) with ChIP-exo binding. (**c**) Example of a site from (**a**) with no ChIP-exo binding. (**d**) Mean ChIP-exo tag counts of BATF and IRF4 binding on consensus AICE1 and AICE2 motifs within IRF4 ChIP-seq peaks showing ±50 bp flanking regions; red box: exo binding region of

IRF4, light red box: control region for IRF4, blue box: exo binding region of BATF, light blue box: control region for BATF.

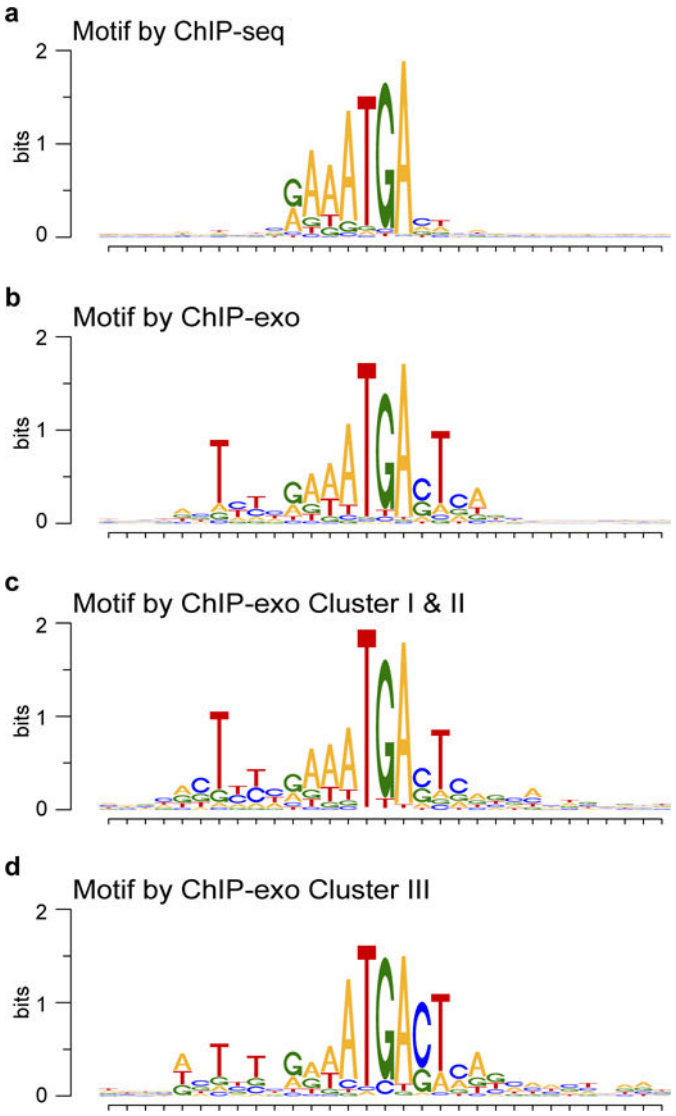

Figure 6.

ChIP-exo reveals a novel AICE2 (**a**) AICE2 motif predicted by *de novo* motif analysis of IRF4 ChIP-seq. (**b**) Novel AICE2 motif identified from sites with ChIP-exo binding. (**c**) AICE2 motif determined by ChIP-exo in peaks ±50 kb from TSS of Cluster I and II genes. (**d**) AICE2 motif determined by ChIP-exo in peaks ±50 kb from TSS of Cluster III genes.

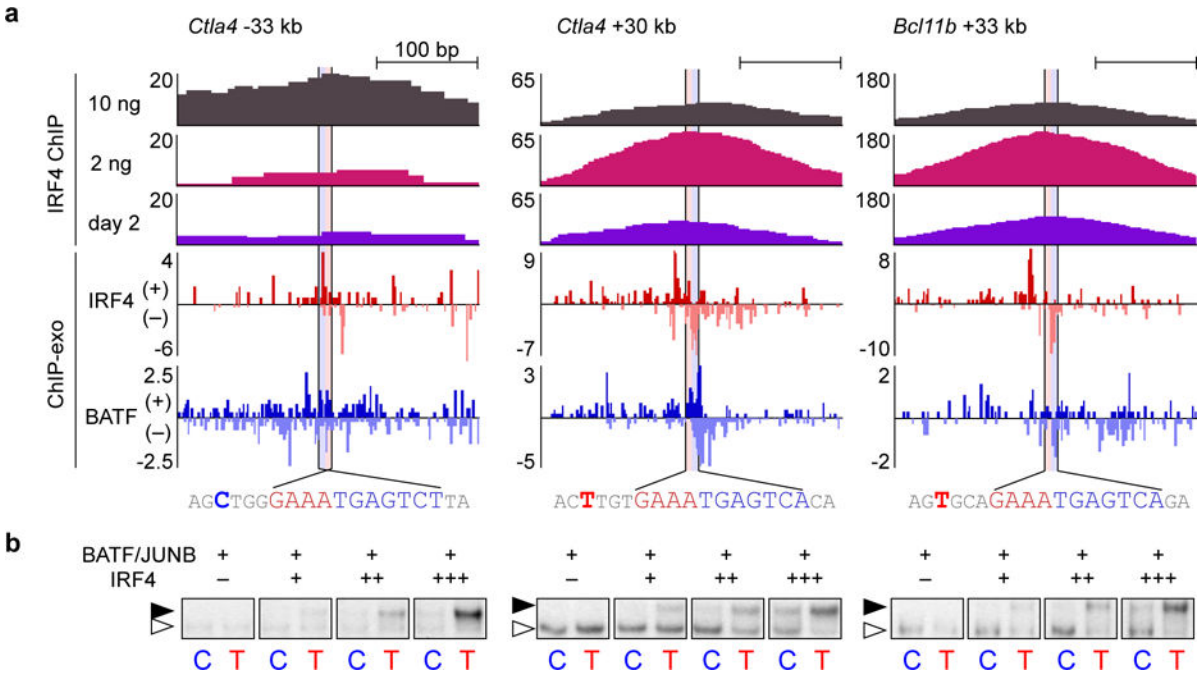

Figure 7.

DNA sequences flanking identical AICE motifs regulate BATF-IRF4 DNA binding affinity and enhancer activity. (**a**) ChIP-seq for IRF4 using the indicated primary activation condition of mouse T_H2 cells and ChIP-exo for IRF4 and BATF showing selected AICE2-containing peaks containing an upstream C (*Ctla4* -33 kb) or an upstream T (*Ctla4* +30 kb and *Bcl11b* +33 kb). (**b**) EMSA using nuclear extracts of HEK293 FT cells expressing BATF, JUNB and varying amounts of IRF4 using probes based on the AICE2 motifs in (**a**) containing the native sequences, a C > T mutation (*Ctla4* -33 kb) or T > C mutations (*Ctla4* +30 kb and *Bcl11b* +33 kb) at the -4 bp position. Open triangle: BATF/JUNB. Solid triangle: BATF/JUNB/IRF4. Data were representative of two experiments (**b**). Full sequence of probes are shown in Supplementary Table 2 (**b**).

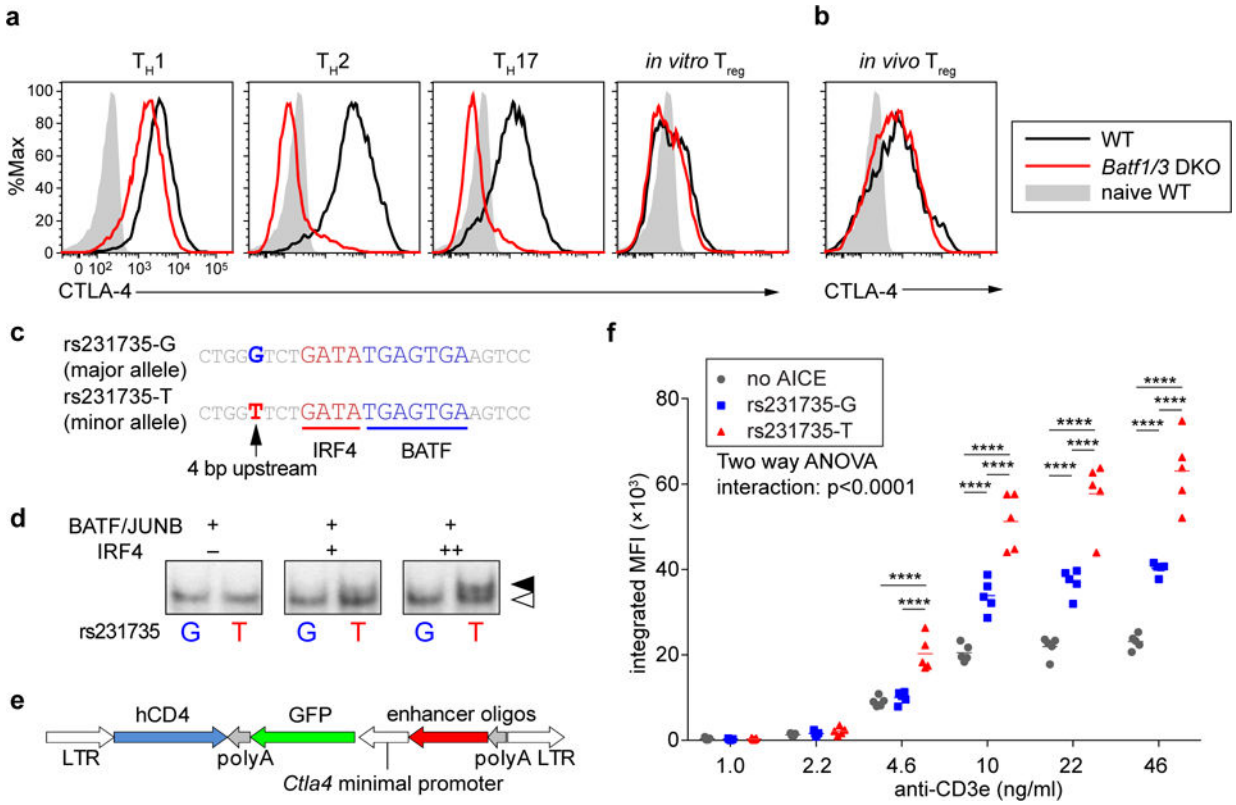

Figure 8.

Human *CTLA4* SNP affects BATF-IRF4 DNA binding affinity and enhancer activity. (**a**) Flow cytometry analyzing CTLA-4 expression in indicated helper T cell subsets cultured on crosslinked anti-CD3ε and anti-CD28 for 4 days under T_H1 (anti-IL-4, IL-12 and IFN-γ), T_H2 (anti-IL-12, anti-IFN-γ and IL-4), T_H17 (anti-IFN-γ, anti-IL-12, anti-IL-4, IL-6, TGF-β and IL-1β), Treg (anti-IFN-γ, anti-IL-12, anti-IL-4, TGF-β) conditions. (**b**) Flow cytometry analyzing CTLA-4 expression in $CD3^+CD4^+$ $Foxp3^+$ T cells in mesenteric lymph node. Data are representative of two replicates. (**c**) Sequences of human SNP rs231735 (*CTLA4* -38kb). (**d**) EMSA using nuclear extract of HEK293 FT cells expressing BATF, JUNB and varying amounts of IRF4 with probes based on of rs231735-G and rs231735-T. (**e**) Structure of reporter retrovirus for (**f**). (**f**) Integrated MFI analysis of reporter activity in T_H2 cells expressing retroviral reporter construct containing no AICE, rs231735-G, or rs231735-T. Two-way ANOVA with Tukey's multiple comparison; **** P<0.0001. n =5, pooled two experiments (**f**). Data were representative of two experiments (**a**). Full sequence of probes are shown in Supplementary Table 2.