Risk variants disrupting enhancers of \(T_H1\) and \(T_{\text{REG}}\) cells in type 1 diabetes

Peng Gao\(^{a,b,c,1}\), Yasin Uzun\(^{a,b,c,1}\), Bing He\(^{a,b,c}\), Sarah E. Salamati\(^d\), Julie K. M. Coffey\(^d\), Eva Tsaklikan\(^d\), and Kai Tan\(^{a,b,c,e,f,g,2}\)

\(^{a}\)Division of Oncology, Children’s Hospital of Philadelphia, Philadelphia, PA 19104; \(^{b}\)Center for Childhood Cancer Research, Children’s Hospital of Philadelphia, Philadelphia, PA 19104; \(^{c}\)Department of Biomedical and Health Informatics, Children’s Hospital of Philadelphia, Philadelphia, PA 19104; \(^{d}\)Stead Family Department of Pediatrics, University of Iowa Carver College of Medicine, Iowa City, IA 52242; \(^{e}\)Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; \(^{f}\)Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; and \(^{g}\)Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

Edited by Wing Hung Wong, Stanford University, Stanford, CA, and approved March 1, 2019 (received for review September 5, 2018)

Genome-wide association studies (GWASs) have revealed 59 genomic loci associated with type 1 diabetes (T1D). Functional interpretation of the SNPs located in the noncoding region of these loci remains challenging. We perform epigenomic profiling of two enhancer marks, H3K4me1 and H3K27ac, using primary \(T_H1\) and \(T_{\text{REG}}\) cells isolated from healthy and T1D subjects. We uncover a large number of deregulated enhancers and altered transcriptional circuits in both cell types of T1D patients. We identify four SNPs (rs10772119, rs10772120, rs3177692, rs883868) in linkage disequilibrium (LD) with T1D-associated GWAS lead SNPs that alter enhancer activity and expression of immune genes. Among them, rs10772119 and rs883868 disrupt the binding of retinoic acid receptor \(\alpha\) (RARA) and Yin and Yang 1 (YY1), respectively. Loss of binding by YY1 also results in the loss of long-range enhancer-promoter interaction. These findings provide insights into how noncoding variants affect the transcriptomes of two T-cell subtypes that play critical roles in T1D pathogenesis.

Significance

Functional interpretation of noncoding genetic variants identified by genome-wide association studies is a major challenge in human genetics and gene regulation. We generated epigenomic data using primary cells from type 1 diabetes patients. Using these data, we identified and validated multiple novel risk SNPs for this disease. In addition, our ranked list of candidate risk SNPs represents the most comprehensive annotation based on T1D-specific T-cell data. Because many autoimmune diseases share some genetic underpinnings, our dataset may be used to understand causal noncoding mutations in related autoimmune diseases.

Author contributions: P.G. and K.T. designed research; P.G. and K.T. performed research; Y.U., B.H., S.E.S., J.K.M.C., and E.T. contributed new reagents/analytic tools; Y.U. analyzed data; and P.G., Y.U., and K.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE112342).

1\(^{p}\) G. and Y. U. contributed equally to this work.
2\(^{2}\) To whom correspondence should be addressed. Email: tank1@email.chop.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1815336116/-/DCSupplemental.

Published online March 25, 2019.

www.pnas.org/cgi/doi/10.1073/pnas.1815336116

PNAS | April 9, 2019 | vol. 116 | no. 15 | 7581–7590

---

**Results**

Transcriptome Changes in \(T_H1\) and \(T_{\text{REG}}\) Cells of T1D Patients. Using a panel of established cell surface markers, we purified effector memory \(T_{\text{REG}}\) cells (CD3\(^+\) CD4\(^+\) CD25\(^+\) CD127\(^{dim/−}\)) (9, 10) and effector memory \(T_H1\) cells (CD3\(^+\) CD4\(^−\) CXCR3\(^+\) CXCR5\(^−\) CD45RO\(^−\)) (9) from the peripheral blood of 11 subjects, including 6 T1D patients and 5 age-matched healthy controls (SI Appendix, Fig. S1 and Tables S1 and S2).

---

**Type 1 diabetes (T1D)** is an autoimmune disease caused by immune-mediated destruction of the insulin-producing pancreatic beta cells. There is overwhelming evidence that imbalance between effector and regulatory T cells is a major cause of autoimmunity. Of all effector T cells, T-helper 1 (\(T_H1\)) cells play a critical role in the pathogenesis of T1D by producing IFN-\(\gamma\). The pathogenic role of \(T_H1\) cells is demonstrated by the transfer of disease through transplantation of pathogenic \(T_H1\) cells (1). \(T_H1\)-initiated destruction of beta cells can be modulated by Foxp3\(^+\) regulatory T cells (\(T_{\text{REG}}\)), which play an important role in the maintenance and regulation of immune tolerance and prevention of autoimmunity. They can modulate T-cell activation and promote immune tolerance by direct cell–cell interactions and production of immune modulatory cytokines such as transforming growth factor \(\beta\) (TGF-\(\beta\)) and interleukin 10 (IL-10). The protective role of \(T_{\text{REG}}\) cells is impaired in susceptible populations, and studies have demonstrated a deficiency in number and function of \(T_{\text{REG}}\) cells in T1D (2).

Despite years of study, the molecular mechanisms responsible for the loss of immune tolerance remain to be fully elucidated. Genome-wide association studies (GWASs) have revealed 119 lead SNPs (59 genomic loci) that are associated with T1D (ImmunoBase; https://www.immunobase.org), but very few etiologic SNPs have been demonstrated. One challenge is that many associated SNPs are located in noncoding regions, which represent 98.5% of the human genome. Although several risk noncoding SNPs have been reported for the autoimmune-associated genes DEXI (3), GLUT1 (4), and IL6 (5), our knowledge of risk noncoding SNPs for autoimmune diseases remains limited.

Although multiple studies have demonstrated enrichment of T1D GWAS variants at T-cell–specific transcription enhancers in healthy donors (6, 7), to date, no study has examined the enhancer repertoire in primary \(T_H1\) and \(T_{\text{REG}}\) cells from T1D patients, despite of the pivotal roles of \(T_H1\) and \(T_{\text{REG}}\) cells in the pathogenesis of T1D. In this study, we conducted epigenomic and transcriptomic profiling of \(T_H1\) and \(T_{\text{REG}}\) cells isolated from a cohort of five healthy donors and six newly diagnosed T1D patients. Our data (8) reveal significant alteration in the enhancer repertoire and transcriptional regulatory circuitry in \(T_H1\) and \(T_{\text{REG}}\) cells of T1D patients. Intersecting our epigenomic data with a catalog of SNPs located in previously reported T1D-associated genomic loci, we identified several novel risk SNPs located in \(T_H1\) and \(T_{\text{REG}}\) enhancers. We validated the functional roles of four candidate \(T_{\text{REG}}\) SNPs using a combination of luciferase reporter assay, genome-editing, transcription factor chromatin immunoprecipitation (ChIP), and chromosome conformation capture (3C) assays.

---

**Significance**

Functional interpretation of noncoding genetic variants identified by genome-wide association studies is a major challenge in human genetics and gene regulation. We generated epigenomic data using primary cells from type 1 diabetes patients. Using these data, we identified and validated multiple novel risk variants for this disease. In addition, our ranked list of candidate risk SNPs represents the most comprehensive annotation based on T1D-specific T-cell data. Because many autoimmune diseases share some genetic underpinnings, our dataset may be used to understand causal noncoding mutations in related autoimmune diseases.

---

**Data and Materials**

Functional interpretation of noncoding genetic variants identified by genome-wide association studies is a major challenge in human genetics and gene regulation. We generated epigenomic data using primary cells from type 1 diabetes patients. Using these data, we identified and validated multiple novel risk variants for this disease. In addition, our ranked list of candidate risk SNPs represents the most comprehensive annotation based on T1D-specific T-cell data. Because many autoimmune diseases share some genetic underpinnings, our dataset may be used to understand causal noncoding mutations in related autoimmune diseases.

---

**Author Contributions**

P.G. and K.T. designed research; P.G. and K.T. performed research; Y.U., B.H., S.E.S., J.K.M.C., and E.T. contributed new reagents/analytic tools; Y.U. analyzed data; and P.G., Y.U., and K.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE112342).

---

**Acknowledgments**

P.G. and Y.U. contributed equally to this work. To whom correspondence should be addressed. Email: tank1@email.chop.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1815336116/-/DCSupplemental.

Published online March 25, 2019.
We profiled the transcriptome of patients and healthy donor subjects using RNA-seq (SI Appendix, Table S3) and found 370 and 250 differentially expressed transcripts between case and control groups for T1H1 and TREG cells, respectively [false-discovery rate (FDR) < 0.1, Fig. L4 and SI Appendix, Table S4]. Consistent with the pathogenic role of T1H1 in T1D, up-regulated genes are enriched for functions such as lymphocyte activation, cell growth, cell proliferation, and immune response, whereas down-regulated genes are enriched for functions such as cell death and apoptosis.

For TREG cells, the up-regulated genes are involved in transcription, cell cycle, and chromosome organization, whereas down-regulated genes are involved in apoptosis, protein modification, targeting, and transport (Fig. 1C).

Several T1D-associated genes (obtained from ImmunoBase) are differentially expressed between case and control groups, including Rac family small GTPase 2 (RAC2) in both cell types, Ikaros 1 (IKZF1) and diacylglycerol kinase α (DGKA) in T1H1 cells, and tyrosine kinase 2 (TYK2), IL10, and major histocompatibility complex, class II, DO β1 (HLA-DQB1) in TREG cells (Fig. 1B). Besides those known T1D-associated genes, several genes associated with other autoimmune diseases are also differentially expressed, including cyclin-dependent kinase 6 (CDK6), Janus kinase 2 (JAK2), C-C motif chemokine receptor 4 (CCR4), and signal transducer and activator of transcription 5A/5B (STAT5A/B) in T1H1, and interleukin-5 (IL5), IFN regulatory factor 4 (IRF4), and signal transducer and activator of transcription 2 (STAT2) in TREG. STAT5A and STAT5B are reported to have a role in T1H1 cell differentiation (11). SMAD3, which is down-regulated in the case group, has a critical role in T1H1 inhibition and immune tolerance (12, 13). Together with FOXP3 and other transcription factors (TFs), IRF4 can form a transcriptional network that governs TREG cell differentiation (14). IL5 is reported to promote induction of antigen-specific TREG cells that suppress autoreactivity and reduced expression of IL5 will disrupt the immune balance (15) (Fig. 1B and SI Appendix, Table S4).

Deregulated Enhancers in T1H1 and TREG Cells of T1D Patients. Traditional ChIP followed by high-throughput sequencing (ChIP-seq) protocols require millions of cells. From a typical blood draw of 50 mL, it is only possible to purify approximately half a million T1H1 cells and TREG cells, which is not sufficient for profiling multiple histone marks and the transcriptome. We thus developed a low-input ChIP-seq protocol using as low as 20,000 cells. We validated the protocol using the human lymphoblastoid cell line, GM12878. Data generated using our low-input protocol have an excellent agreement with data generated with conventional protocol using millions of cells (SI Appendix, Fig. S2).

Using our low-input ChIP-seq protocol, we profiled two histone modification marks, H3K4me1 and H3K27ac, the combination of which marks active transcriptional enhancers (SI Appendix, Table S5). By using normalized ChIP-seq signals in 1,000-bp windows genome-wide, we obtained a median genome-wide interindividual correlation ranging between 0.84 and 0.87 considering subject groups (control, case), cells (T1H1, TREG), and histone marks (H3K4me1 and H3K27ac) (SI Appendix, Fig. S3A). These correlations are higher than a similar study performed using TREG cells from 11 healthy donors (16) (median, 0.70) (SI Appendix, Fig. S3C).

We predicted active enhancers for each cell type and cohort (case and control) separately using H3K4me1 and H3K27ac ChIP-seq data and the Chromatin Signature Identification by Artificial Neural Network (CSI-ANN) algorithm (17). For T1H1 cell, we predicted 13,017 and 12,145 enhancers in case and control groups, respectively. For TREG cell, we predicted 11,915 and 14,860 enhancers in case and control groups, respectively (Fig. 2A and SI Appendix, Table S6). As expected, 83% of T1H1 enhancers and 74% of TREG enhancers are shared between the two cell types, given their common origin from naïve CD4+ T cells (Fig. 2B). However, within each cell type, a large fraction of the enhancers has altered activities between the case and control groups (21% for T1H1 and 25% for TREG) (Fig. 2A and D). The percentages of group-specific enhancers were significantly higher than expected (null distribution computed by
permuting the sample labels in the dataset; Fig. 2C). **T**h1 cells in T1D patients appear to have an overall gain of active enhancers, whereas **T**reg cells in T1D patients appear to have an overall loss of active enhancers. Taken together, these data suggest a considerable change in the transcriptional regulatory networks (TRNs) of both cell types in T1D patients.

To understand the impact of case-specific enhancers on the transcriptomes, we need to know their target genes. We recently developed the Integrated Method for Predicting Enhancer Targets (IM-PET) algorithm (18). It predicts enhancer–promoter interactions by integrating four statistical features derived by integrating transcriptome, epigenome, and genome sequence data. Using IM-PET, on average, each gene is predicted to be targeted by 1.5 and 1.6 enhancers in **T**h1 and **T**reg cells, respectively. We compared our EP predictions with a recently published Capture-Hi-C data on CD4+ T cells (SI Appendix, Fig. S4). Approximately 55% of our predictions are supported by Capture-Hi-C data, suggesting high quality of our predictions. Finally, for both cell types, there is a positive correlation between enhancer activity and gene expression level across the case and control groups (Fig. 2E); providing further support to the accuracy of our target gene prediction.

We next focused on target genes of case/control-specific enhancers (SI Appendix, Table S7). Many of the condition-specific enhancers are linked to genes already implicated in T1D, such as **IKZF1**, **CCR5**, **CLEC16A**, **IL2RA**, and **UBASH3A** for **T**h1 and **IKZF1**, **IKZF4**, **RAC2**, and **RASGRF1** for **T**reg (7, 19). Gene ontology analysis of the enhancer targets suggests deregulation of specific pathways in **T**h1 and **T**reg cells of T1D patients, such as cell-cell activation, lymphocyte activation, leukocyte activation, innate immune response, and cellular response to organic substances (SI Appendix, Table S8).

### Key Transcription Factors Mediating Transcriptome Changes in **T**h1 and **T**reg Cells of T1D Patients

Our knowledge about the TRNs in **T**h1 and **T**reg cells of T1D patients is rather limited (20). Motivated by our knowledge of transcriptional regulation, we developed a method, target inference via physical connection (TIPC), to infer condition-specific TRNs. TIPC computes probability scores for three key components of transcriptional regulation, including probability of a DNA sequence being an enhancer, probability of a TF binding to an enhancer given the TF motif model and the enhancer sequence, and probability of enhancer–promoter interaction. The overall score for a TF regulating a target gene is the product of the three component probabilities and the expression level of the TF (see Methods for details). We evaluated the performance of TIPC using two approaches. First, using a set of gold-standard TF-target pairs in embryonic stem cells, we found that TIPC outperforms four state-of-the-art methods based on Pearson correlation (BC), mutual information [context likelihood ratio (CLR) (21)], decision trees [gene network inference with ensemble of trees (GENIE3) (22)], and regression [trustful inference of gene regulation with stability selection (TIGRESS) (23)] for predicting TF–target interactions (Fig. 3A). We further evaluated the predicted TF–target interactions using publicly available TF ChIP-seq data (24) for T cells. We found that our predicted TF–target interactions have significant overlap with ChIP-seq data (SI Appendix, Fig. S4B). We predicted 263,836 and 298,638 TF–target pairs in **T**h1 and **T**reg cells, respectively. Thirty-six percent and 41% of the TF–target pairs are either control- or case-specific in **T**h1 and **T**reg cells, respectively (Fig. 3B), suggesting a significant degree of rewiring of the TRN in both cell types of T1D patients.

Next, we developed a method to systematically identify key TFs that play critical roles in the state transition of a TRN between two conditions. Our method is based on the assumption that a key TF can influence a larger set of differentially expressed targets compared with a nonkey TFs. To capture such signal, for two adjacent genes in the TRN. We convert their P values of
T cells, in-...<0.001, t test; SI Appendix, Fig. S5). Most of the identified TFs have a reported role in either the biology of T1H and T REG cells, or the pathogenesis of either T1D or other autoimmune diseases (SI Appendix, Table S9).

When ranked according to their regulatory potential (the inverse of median distance of TF to differentially expressed targets), the top key TFs are IKZF1 and SREBF1 for T1H and IRF4 and RARA for T REG. IKZF1 plays a role in the T1H versus T1H2 polarization (25, 26). A SNP (rs10272724) in the 3′-UTR of IKZF1 has been shown to be protective from T1D (27). SREBF1 plays an important role in inducing genes that encode numerous genes in the lipid biosynthesis pathway, which controls full activation, proliferation, and differentiation of CD4+ T cells, including T REG cells (28, 29).

IRF4, the top key regulator in T REG, is reported to interact with FOXP3 and promote T REG function (14). RARA is critical for the normal differentiation and functions of both T REG and T1H cells. All-trans retinoid acid (ATRA), an endogenous ligand of RARA, can prevent human natural T REG cells from converting to T1H/T1H17 cells and sustains their suppressive function in inflammatory environments (30). Likewise, RARA can sustain T1H1 cell lineage stability and prevents transition to a T17 cell program in vivo (31). Consistent with the role of RARA in T1H1 and T REG cells, RARA level decreases in both cell types (fold changes are 0.55 and 0.08 for T1H1 and T REG, respectively; SI Appendix, Table S4) in T1D patients.

Fig. 3 D and F show the case/control-specific regulatory interactions among the key TFs. It highlights the complex and dynamic interactions among the key TFs during T1D development. For most of the key TFs, there are several case and control interactions. Moreover, although ~47% of the TFs are shared between T1H1 and T REG cells, the interaction partners of many of them are different in the two cell types, such as IKZF1, TCF3, and RARA.

Candidate Noncoding Risk Variants Affecting Enhancers in T1H and T REG Cells. A recent study reported that T1D-associated SNPs are enriched in enhancers active in thymus, T and B cells, and CD34+ hematopoietic cells (7). However, this study is based on lymphocyte data from healthy individuals and thus may not fully capture the set of deregulated enhancers in T1D patients. We computed the overlap between sets of tissue- or cell-specific enhancers (including our T1H-, and T REG-specific enhancers) and the set of T1D-associated SNPs, which consists of 119 lead GWAS SNPs (32–35) and 3,844 additional SNPs that are in the same linkage disequilibrium blocks with the lead SNPs. Consistent with the previous study, we found enrichment of T1D-associated SNPs in subsets of CD4+ and CD8+ T cells, especially memory T cells, using hypergeometric test. However, we found T REG and T1H enhancers from this study have much higher enrichment for T1D-associated SNPs compared with enhancers identified in all subsets of CD4+ and CD8+ T cells used here from healthy individuals (Fig. 4A, Left). Since hypergeometric test does not take into account linkage disequilibrium, we used an alternative approach, linkage disequilibrium score regression (11), to evaluate the statistical significance. We again observed that T REG and T1H enhancers from this study tend to have higher statistical significance of overlapping GWAS SNPs than enhancers identified in all subsets of CD4+ and CD8+ T cells from healthy individuals (Fig. 4A, Right). We found no significant difference in overlapping GWAS SNPs between case and control-specific enhancers. This is not unexpected since enhancer SNPs may exert their effect on the phenotype via both loss or gain of function. This result provides additional support for the direct involvement of T REG and T1H cells in T1D pathogenesis. It also highlights the importance of using primary cells from patients and control subjects for finding disease-associated genetic variants. The set of enhancers that overlap with T1D-associated SNPs is provided in SI Appendix, Table S10.
As the first step to identify noncoding risk SNPs that alter enhancer activity, we compared the normalized H3K4me1 and H3K27ac signals of each enhancer between case and control groups. We identified enhancers with significantly altered histone signals between case and control in T1H1 and TREG cells, respectively (P < 0.1, Wilcoxon rank-sum test, two-sided, SI Appendix, Table S11). From this set of enhancers, using a combination of multiple orthogonal supporting evidence, including overlapping with super enhancers, expression quantitative trait loci (eQTL) and perturbation of TF binding sites (Table S12), we selected five enhancers for each cell type to test their activity using luciferase reporter assay. We differentiated a Jurkat cell line as the model for T1H1 cells (39) and a Foxp3-expressing Jurkat cell line as the model for TREG cells (40–42), hereby termed Foxp3+–Jurkat cells (Methods). Both cell models were further validated by checking the expression pattern of known signature genes for each cell type (SI Appendix, Fig. S9).

Of the five candidate T1H1 enhancers, three of them showed significant activity compared with the negative controls (Fig. 4B, P < 0.01, Wilcoxon rank-sum test, one-sided). The first enhancer (chr16: 28543200–28545200) has a significant reduction in H3K4me1 signal (P < 0.01, Wilcoxon rank-sum test, two-sided; Fig. 4C) in T1D patients compared with healthy individuals. The predicted target genes of the enhancer are ceroid-lipofuscinosis, neuronal 3 (CLN3), and linker for activation of T cells (LAT). Both predictions are supported by published Capture-Hi-C data in CD4+ T cells (38). CLN3 is important for the normal function of lysosomes. Impaired lysosome function caused by mitochondrial respiration deficiency subverts T-cell differentiation toward proinflammatory subsets and exacerbates the in vivo inflammatory response (43). LAT is required for T-cell receptor-mediated signaling both in mature T cells and during T-cell development. Mice in which tyrosine 136 of LAT is constitutively mutated accumulate CD4+ T cells that trigger autoimmunity and inflammation (44). This enhancer harbors two T1D-associated SNPs, rs4788083 and rs231976. Using randomly selected SNPs from 1000 Genomes Project as the background null distribution, we found that rs231976 perturbs the binding site of STAT1 [empirical P < 10−5; Methods D; Measurement of linkage disequilibrium between an enhancer SNP and the lead T1D GWAS SNP rs4708084 (shown in red)]. (2) Enhancer located at chr16: 28543200–28545200 (highlighted in yellow), including normalized histone modification ChIP-seq signals, RNA-seq signals of target genes of the enhancer, enhancer-promoter interaction (if any), and SNP locations. In addition, TF motif and binding sites (if any) that are significantly affected by the SNPs are also shown. There are two SNPs, rs4788083 and rs231976, in this enhancer. rs231976 perturbs the binding site of STAT1 [empirical P < 10−5; corrected for multiple testing with Benjamini–Hochberg (BH) method], a key TF involved in the IL-27–mediated early commitment to the T1H1 lineage (45).

The second validated enhancer (chr1:198158800–198160800) is predicted to target the gene DENN domain containing 1B (DENND1B) (Fig. 4D). The enhancer–promoter interaction is also supported by Capture-Hi-C data in CD4+ T cells (38). The H3K4me1 signal at the enhancer is significantly reduced in T1D patients compared with healthy individuals (P < 0.01, Wilcoxon rank-sum test, two-sided). Polymorphisms in DENND1B are associated with asthma and other immune disorders (46, 47).

The third validated enhancer (chr6: 149531600–149533600) targets the gene zinc finger CCCH-type containing 12D (ZCH12D), and there is significant reduction in H3K4me1 level in T1D patients compared with healthy individuals (P < 0.01; Fig. 4E). ZCH12D is an RNase that destabilizes a set of mRNAs. ZCH12D−/− mice exhibit multiorgan inflammation (48) and autoimmune disease (49). A set of genes, including IL6, IL12, REL, TNRFSF4, and IL2, which are important for the activation of T1H1 cells, are targets of ZCH12D (49).

For TREG cells, all five enhancers show significant activity compared with the negative controls (Fig. 5A, P < 0.01, Wilcoxon rank-sum test). The first enhancer (chr12: 9808600–9911600) targets CD69, CLEC2B, KLRB1, and E1F23L, and harbors three SNPs (rs10772119, rs10772120, and rs3176792), which are in strong linkage disequilibrium with the T1D-associated GWAS lead SNP rs4763879 from ImmunoBase (D′ = 0.99, 0.99, and 0.96, respectively).
respectively). The lead SNP is located in the CD69 gene body (50) (Fig. 5B). The second enhancer (chr21: 43843200–43845200) targets UBASH3A and harbors the SNP rs883868, which is in strong linkage disequilibrium with the T1D-associated GWAS lead SNP rs11203203. The lead SNP is located in the UBASHA gene body (50) (Fig. 5C).

Of the five target genes by the first two enhancers, CD69 and UBASH3A are known T1D-associated genes (19), and CD69 is well studied for its role in T1D and TREG cells. CD69 is required for the repressive function of TREG cells (51). Diminished CD69 expression is associated with compromised function of TREG cells in systemic sclerosis (52). UBASH3A is primarily expressed in T cells and can negatively regulate T-cell signaling (53). Knockout of UBASH3A up-regulates inflammatory cytokine production and increase susceptibility to autoimmunity in a mouse model of multiple sclerosis (54). Roles of the other three genes in T1D and/or TREG have not yet been investigated. Nevertheless, KLRB1 could be of potential interest because it defines a subset of TREG cells capable of producing proinflammatory cytokines (55).

The third enhancer (chr10: 6094400–6096600) harbors the SNP rs10795763 and targets the gene IL2RA, which is associated with T1D, rheumatoid arthritis, and multiple sclerosis (19). Polymorphisms in IL2RA have been reported to negatively regulate the function of TREG cells (56). The SNP affects the binding of MBD2-interacting zinc finger protein (MIZF) (SI Appendix, Fig. S10). Methyl CpG binding protein 2 (MBD2) promotes demethylation of FOXP3 promoter and TREG function (57). The fourth enhancer (chr14: 68757800–68760200) harbors the SNP rs35763290 and targets the gene Actinin Alpha 1 (ACTN1). ACTN1 is an actin-binding protein and involved in the formation of immunological synapse, which is critical for T-cell activation, migration, and effector function at the interface between a T cell and its cognate antigen-presenting cell or target cell (58). The SNP affects the binding of nuclear respiratory factor 1 (NRF1). The fifth enhancer (chr1: 200830000–200832000) targets the PHLDA3 gene but does not harbor any SNP that disrupts any known TF binding motif. PHLDA3 is a p53-regulated repressor of Akt signaling, inhibition of which can suppress the activation and proliferation of TREG and results in a significant reduction of TREG cells in mice (59).

Two Enhancer SNPs Affecting RARA and YY1 Binding at TREG Enhancers Regulating Immune Response Genes. There are 25 SNPs in the five validated TREG enhancers described in the previous section. They are either a GWAS lead SNPs documented in the ImmunoBase or SNPs that are in linkage disequilibrium with a GWAS lead SNPs. We genotyped all 25 SNPs in the 11 study subjects using Sanger sequencing (SI Appendix, Table S13). Of these 25 SNPs, 9 have case-specific genotypes (the genotypes that are exclusively observed in T1D patients but not in healthy control), suggesting they are potential candidate risk SNPs. Among them, four SNPs (rs10772119, rs10772120, rs3176792, and rs883868), which are located in two enhancers, are associated with significant changes in histone modification and expression of the target genes of the host enhancer (Wilcoxon rank-sum test, P < 0.1; SI Appendix, Fig. S11). We therefore focused on these two enhancers and the four SNPs to further investigate their role in transcriptional regulation in TREG cells.

To test whether the SNPs can alter enhancer activity, we first mutated all four SNPs by site-directed mutagenesis and performed luciferase reporter assay in Foxp3+Jurkat cells. We found that all four mutations cause significant change in the enhancer activity (P < 0.01, two-sided Student’s t test; Fig. 6A).

To investigate the role of the risk SNPs in target gene expression, we used CRISPR-Cas9–based genome editing to engineer Foxp3+Jurkat cell lines with homozygous alleles at the SNPs (Fig. 6B). Using these cell lines, we found that cells with a homozygous T allele (risk allele) at rs10772119 have significantly reduced expression levels of CD69 and CLEC2B. Likewise, cells with a homozygous allele C (risk allele) at rs883868 have significantly reduced UBASH3A expression (Fig. 6C). Taken together, these results are consistent with the association between the genotype and gene expressions we found in our cohort.
To further establish the enhancer–target gene relationship, we tested the predicted enhancer–promoter interactions in Foxp3+–Jurkat cells using 3C coupled with quantitative PCR (3C-qPCR). As a negative control for the enhancer–promoter interaction, we used the human embryonic kidney 293FT (HEK 293FT) cell line. Compared with TREG cells, 293FT has very low expression of CD69 and UBASH3A genes, but the mRNA levels of CLEC2B in both cell types are comparable. Compared with negative control regions, the enhancer at CD69 locus significantly interacts with the promoters of CD69 and CLEC2B (P < 0.01, one-sided t test). The enhancer at UBASH3A locus interacts with the promoter of UBASH3A (Fig. 7A). Consistent with mRNA expression, we only found significant but weaker interaction between the enhancer with the CLEC2B promoter in HEK 293FT cells.

To identify TFs whose binding is affected by the two risk SNPs, we performed TF motif analysis using a collection of 1,772 motifs from the CIS-BP database (60). We found that rs10772119 exhibits allelic-specific binding by RARA (Fig. 5B) and rs883868 exhibits allelic-specific binding by YY1 (Fig. 5C). For both TFs, we performed ChIP-qPCR to confirm the binding of the TF to the predicted binding sites (Fig. 7B). Furthermore, ChIP-qPCR using cells with homozygous alleles at the SNP showed that the sites bearing the risk alleles has significantly lower binding affinity, confirming the result of our computational analysis.

To further test the role of the RARA binding site at rs10772119 in target gene expression, we treated Foxp3+–Jurkat cells with ATRA and measured the expression of the target genes using RT-qPCR. We found that only in cells with homozygous C alleles, the target genes are significantly induced by ATRA treatment (Fig. 7C), providing additional support to the role of rs10772119 in RARA binding and target gene expression.

YY1 has been reported to mediate enhancer–promoter interactions in a number of mammalian cell types (61). We hypothesized that reduced YY1 binding to the enhancer harboring rs883868 can result in reduced enhancer–promoter interaction. We therefore performed 3C-qPCR using Fox3+–Jurkat cells engineered to have homozygous T and C alleles at the SNP. Consistent with our hypothesis, we found significantly reduced enhancer–promoter interaction in cells with homozygous C allele at the SNP rs883868 (Fig. 7A).

In summary, the series of experiments above uncovered four candidate risk SNPs (rs10772119, rs10772120, rs3176792, and rs883868) that are in linkage disequilibrium with known T1D GWAS SNPs and alter enhancer activity and expression of immune response genes (Figs. 4–7). Among them, rs10772119 and rs883868 disrupt the binding of RARA and YY1, respectively. Moreover, allelic-specific binding by YY1 results in the allelic-specific long-range enhancer–promoter interaction involving UBASH3A (Fig. 7).

**Discussion**

GWASs have revealed 59 high-confidence genomic loci associated with T1D, which harbor 119 lead SNPs and 3,844 additional SNPs that are in the same linkage disequilibrium blocks with the lead SNPs. Among those SNPs, functional interpretation of noncoding SNPs remains particularly challenging. Toward this goal, we conducted an epigenomics-based fine map to identify candidate risk noncoding SNPs. We mapped the genome-wide distribution of two enhancer-associated histone modifications (H3K4me1 and H3K27ac) in T1H and TREG cells from five healthy and six T1D patient subjects using a low-input ChIP-seq protocol.

We identified four new risk SNPs for T1D and two of them exhibit allele-specific binding by two TFs, RARA and YY1. ATRA induces the normal development of pancreas and affects the function of beta cells. RA signaling has a strong association with the onset of T1D (62). T1D patients (63, 64) are known to have vitamin A deficiency. Dietary RA reduces diabetes in diabetic-prone BB/Wor rats (65, 66). ATRA treatment significantly reduces diabetes incidence and delays the onset of diabetes transferred from NOD mice to NOD/SCID recipient mice. In summary, these findings suggest a protective role of ATRA against T1D. This protective role is due to induced TREG cell-dependent immune tolerance by suppressing both CD4+ and CD8+ T effector cells, while promoting TREG cell expansion (67). Supporting this mechanism, ATRA cannot exert the protective role in mice with abrogated TREG cells (68).

Here, we found that rs10772119 can disrupt the binding of RARA to the enhancer located at chr12: 9908600–9911600. Subjects with this risk SNP have lower expression of several target genes, including CD69, CLE2B, KLRB1, and EIF2S3L. CD69 is an early marker of T-cell activation. ATRA can increase the expression of CD69 in a dose-dependent fashion (69), and this effect is aborted if RARA was knockout (70). Consistent with this finding, we found that RA can induce the expression of CD69 and CLEC2B in TREG cell lines engineered to have a C
allele at rs10772119 (Fig. 7C). Taken together, our finding provides a mechanistic explanation for which rs10772119 promotes the onset of T1D via ATRA signaling in TREG.

YY1 is a zinc finger protein that functions either as a transcriptional activator or repressor. Research in rats shows that YY1 is one of a few proteins that has a reported role in mediating long-range chromatin interactions. Consistent with its role in chromatin looping, here we showed that a single SNP can disrupt YY1 binding and consequently leads to the loss of enhancer–promoter looping mediated by YY1. Our finding adds to increasing reports of genetic variants that can disrupt three-dimensional genome organization and gene expression (74–76).

None of the four risk SNPs identified in this study are lead GWAS SNPs. Instead, they are in linkage disequilibrium with lead SNPs. This result highlights the challenge of finding risk variants in the presence of linkage disequilibrium. Two of the four enhancer SNPs, rs10772120 and rs3176792, do not appear to overlap with any known TF binding motif, although mutating the SNPs showed an effect on enhancer activity and target gene expression. Additional studies are needed to uncover the mechanism by which the two SNPs affect gene expression in TREG cells.

**Methods**

**Study Subjects.** Eleven subjects were recruited at The University of Iowa Children’s Hospital. Recent-onset T1D subjects (within 1 y from the first day of diagnosis; n = 6) were recruited through the Division of Pediatric Endocrinology and Diabetes at The University of Iowa Children’s Hospital. Diabetes was defined according to World Health Organization criteria and included blood glucose levels of 200 mg/dL with symptoms confirmed by a physician. Healthy subjects (n = 5) were recruited by posting flyers at The University of Iowa Children's Hospital. The control criteria comprised fasting blood glucose of 100 mg/dL, no familial history of any autoimmune disorder, and lack of islet autoantibodies. All study subjects were free of known infection at the time of sample collection. For T1D subjects, no use of corticosteroids or glucocorticoids within the prior 6 mo was required. For healthy control subjects, no medications (especially steroids) and nonrelatives of T1D patients were required. At the time of each visit, the following clinical measurements were taken: HbA1c, autoantibodies, weight, and body mass index (subject information is shown in SI Appendix, Tables S1 and S2).

The research protocol was approved by the IRB of the University of Iowa, and participants and/or their parents (guardians) provided written informed consent.

**Purification of TREG and T Cell from Peripheral Blood.** Fifty milliliters of peripheral blood were collected from each subject and processed within 0.5 h.
collection. Blood was diluted with 1× DPBS (Gibco) and subjected to Ficoll-Hypaque centrifugation at 700 × g at 20 °C for 20 min in a swing-out bucket rotor without a brake. A 2-mL aliquot of the middle BMC layer was transferred to a new 50-μL conical tube and cells were washed twice with 1× staining buffer 1 (1× DPBS plus 1% FBS). Cells were resuspended with 1× staining buffer 1 to 10 × 10^6 cells per 50 μL. After adding 5 μL of CD4^+ T cell biotin–antibody mixture per 10^6 cells, cells were mixed and incubated for 5 min at 4 °C. Thirty-five microliters of 1× staining buffer per 10^6 cells and 10 μL of CD4^+ T Cell MicroBead mixture per 10^6 cells were added and mixed. After incubation for 10 min at 4 °C, cells were Giroux M, et al. (2011) SMAD3 prevents graft-versus-host disease by restraining T helper 1 cell differentiation.

3. Leikfoss IS, et al. (2013) Multiple sclerosis-associated single-nucleotide polymorphisms in CLEC16A correlate with reduced SOCS1 and DEXI expression in the thymus. Nat Rev Immunol 13:972–980.

4. Ng DP, et al. (2002) Minor effect of GLUT1 polymorphisms on susceptibility to diabetic childhood acute lymphoblastic leukemia protects against type 1 diabetes. J Immunol 168:5798–5804.

5. Tran GT, et al. (2012) IL-5 promotes induction of antigen-specific CD4^+CD25^+Foxp3^+ regulatory T cells. Proc Natl Acad Sci USA 109:12672–12677.

6. Sánchez-Castillo M, et al. (2015) CODEX: A next-generation sequencing experiment to study gene REgulation using stability selection. BMC Syst Biol 9:49.

7. Gregory GD, Raju SS, Winandy S, Brown MA (2006) Mast cell IL-4 expression is regulated by Ikaros and influences encephalitogenic Th1 responses in EAE. Nat Immunol 7:1368–1374.

8. Murphy KM (2008) Permission to proceed: Jak3 and STATS signaling molecules give the green light for T helper 1 cell differentiation. Immunity 28:725–727.

9. Giroux M, et al. (2011) SMAD3 prevents graft-versus-host disease by restraining Th1 differentiation and granulocyte-macrophage–mediated tissue damage. Blood 117:1734–1744.

10. Takimoto T, et al. (2010) Smad2 and Smad3 are redundantly essential for the TGF-β pathway in T cells and T1D. Immunity 33:806–817.

11. Sánchez-Castillo M, et al. (2015) CODEX: A next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. Nucleic Acids Res 43:D1117–D1123.

12. Giroux M, et al. (2011) SMAD3 prevents graft-versus-host disease by restraining Th1 differentiation and granulocyte-macrophage–mediated tissue damage. Blood 117:1734–1744.

13. Takimoto T, et al. (2010) Smad2 and Smad3 are redundantly essential for the TGF-β pathway in T cells and T1D. Immunity 33:806–817.

14. Fu W, et al. (2012) A multiply redundant genetic switch “locks in” the transcriptional signature of regulatory T cells. Nat Immunol 13:972–980.

15. Tran GT, et al. (2012) IL-5 promotes induction of antigen-specific CD4^+CD25^+ FOXP3^+ regulatory T cells that suppress autoimmunity. Blood 119:4441–4450.

16. Arvey A, et al. (2013) Genomic and epigenetic variation in the lineage specification of regulatory T cells. eLife 3:e04751.

17. Feri HA, Ucar D, Tan K (2010) Discover regulatory DNA elements using chromatin signatures and artificial neural network. Bioinformatics 26:1579–1586.

18. He B, Chen C, Teng L, Tan K (2010) Global view of enhancer-promoter-interactor in human cells. Proc Natl Acad Sci USA 111:E2191–E2199.

19. Bakay M, Pandey R, Hakonarson H (2013) Genes involved in type 1 diabetes: An update. Genes (Base) 4:499–521.

20. Fang D, Zhu J (2017) Dynamic balance between master transcription factor determinates the fates and functions of CD4 T cell and innate lymphoid cell subsets. J Exp Med 214:1861–1876.

21. Haned H, Benschop CGG, Gill PD, Sijen T (2015) Complex DNA mixture analysis in a forensic context: Evaluating the probative value using a likelihood ratio model. Forensic Sci Int Genet 16:17–25.

22. Hughes-Thru XL, Ibarra A, Whitehead L, Geurts P (2011) Inferring regulatory networks from expression data using tree-based methods. PLoS One 5:e12776.

23. Haury AC, Mordelet F, Vera-Licona P, Vert JP (2012) TIGRESS: Trustful inference of gene regulation using stability selection. BMC Syst Biol 6:145.

24. Sánchez-Castillo M, et al. (2015) CODEX: A next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. Nucleic Acids Res 43:D1117–D1123.

25. Gregory GD, Raju SS, Winandy S, Brown MA (2006) Mast cell IL-4 expression is regulated by Ikaros and influences encephalitogenic Th1 responses in EAE. J Clin Invest 117:1330–1336.

26. Thomas RM, et al. (2010) Ikaros silences T-bet expression and interferon-gamma production during Th helper 2 differentiation. J Biol Chem 285:2545–2553.

27. Swafford AD, et al. (2011) An allele of IKZF1 (Ikaros) conferring susceptibility to childhood acute lymphoblastic leukemia protects against type 1 diabetes. Diabetes 60:1044–1044.

28. Angela M, et al. (2016) Fatty acid metabolic reprogramming via mTOR-mediated inclusions of PPARy directs early activation of T cells. Nat Commun 7:13683.
29. Lochner M, Berod L, Sparwasser T (2015) Fatty acid metabolism in the regulation of T cell function. Trends Immunol 36:81-91.
30. Lu L, et al. (2014) Critical role of an all-trans retinoic acid in stabilizing human natural regulatory T cells under inflammatory conditions. Proc Natl Acad Sci USA 111: E3432-E3440.
31. Stein CC, et al. (2015) Retinoic acid is essential for Th1 cell lineage stability and prevents transition to a Th17 cell program. Immunity 42:499-511.
32. Burren OS, et al. (2011) T1DBase: Update 2011, organization and presentation of large-scale data sets for type 1 diabetes research. Nucleic Acids Res 39:D997-D1001.
33. Hulbert EM, et al. (2007) T1DBase: Integration and presentation of complex data for type 1 diabetes research. Nucleic Acids Res 35:D742-D746, and erratum (2007) 35: E638.
34. Smink LJ, et al. (2005) T1DBase, a community web-based resource for type 1 diabetes research. Nucleic Acids Res 33:D544-D549.
35. Barrett JC, et al.: Type 1 Diabetes Genetics Consortium (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 41:703-707.
36. Finucane HK, et al.: ReproGen Consortium; Schizophrenia Working Group of the Psychiatric Genomics Consortium; RACI Consortium (2015) Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 47:1228-1235.
37. Buehl-Sullivan BK, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium (2015) LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 47:291-295.
38. Javierre BM, et al. (2016) Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell 167:1369-1384.
39. Smets RL, et al. (2012) Molecular pathway profiling of T lymphocyte signal transduction pathways; Th1 and Th2 genetic fingerprints are defined by TCR and CD28-mediated signaling. BMC Immunol 13:12.
40. Allan SE, et al. (2008) Generation of potent and stable human CD4+ T regulatory cells by activation-independent expression of FOXP3. Mol Ther 16:194-202.
41. Kim JY, et al. (2007) Functional and genomic analyses of FOXP3-transduced Jurkat T cells as regulatory T (Treg)-like cells. Biochem Biophys Res Commun 362:44-50.
42. Nie H, et al. (2013) Phosphorylation of FOXP3 controls regulatory T cell function and is inhibited by TNF-α in rheumatoid arthritis. Nat Med 19:322-328.
43. Baxaui F, et al. (2015) Mitochondrial respiration controls lysosomal function during inflammatory T cell responses. Cell Metab 22:485-498.
44. Mingueneau M, et al. (2009) Loss of the LAT adaptor converts antigen-responsive T cells into pathogenic effectors that function independently of the T cell receptor. Immunity 31:197-208.
45. Takeda A, et al. (2003) Cutting edge: Role of IL-27/WSX-1 signaling for induction of T regulatory (Treg)-dependent suppression of interferon-gamma-producing T-cells without affecting Th17 cells. Diabetes 58:146-155.
46. Stosis-Grujić S, Cvjeticanić T, Stojanovic I (2009) Retinoids differentially regulate the progression of autoimmune diabetes in three preclinical models in mice. Mol Immunol 47:79-86.
47. Dawson HD, Collins G, Pyle R, Key M, Taub DD (2008) The retinoic acid receptor-alpha (RaA) and type 1 diabetes mellitus in the BB rat. Proc Natl Acad Sci USA 105:20479-20484.
48. Kloting N, Kloting I (2004) Congenic mapping of type 1 diabetes—protective gene(s) in an interval of 4 Mb on rat chromosome 6q32. Biochem Biophys Res Commun 323: 388-394.
49. Kloting N, Kloting I (2004) Genetic variation in the multifunctional transcription factor Yy1 and type 1 diabetes mellitus in the BB rat. Mol Genet Metab 82:255-259.
50. Kleinman J, Jia H, Loguercio S, Su AI, Feeny AJ (2016) YY1 plays an essential role at all stages of B-cell differentiation. Proc Natl Acad Sci USA 113:E3911-E3920.
51. Hinaz D, et al. (2016) Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science 351:1454-1458.
52. Smemo S, et al. (2014) Obesity-associated variants within FTO form long-range regulatory neighborhoods. Nature 507:371-375.
53. Northcott PA, et al. (2014) Enhancer hijacking activates Gfi1 family oncogenes in medulloblastoma. Nature 511:428-434.
54. Cappuccio P, et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8: 2281-2308.
55. Papp AC, Pinsonneault JK, Cooke G, Sadée W (2003) Single nucleotide polymorphism genotyping using allele-specific PCR and fluorescence melting curves. Biotechniques 34:1068-1072.
56. Barrett JC (2009) Haploviev: Visualization and analysis of SNP genotype data. Cold Spring Harb Protoc 2009: pdb.ip71.
57. Auton A, et al.; 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. Nature 526:68-74.