Risk variants disrupting enhancers of T$_H$1 and T$_{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 Tsalikian[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

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$_H$1 and T$_{REG}$ cells isolated from healthy and T1D subjects. We uncover a large number of deregulated enhancers and altered transcriptional circuitries in both cell types of T1D patients. We identify four SNPs (rs10772119, rs10772120, rs3176792, 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 α (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.

enhancer | noncoding variant | diabetes | GWAS | epigenomics

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$_H$1) cells play a critical role in the pathogenesis of T1D by producing IFN-γ. The pathogenic role of T$_H$1 cells is demonstrated by the transfer of disease through transplantation of pathogenic T$_H$1 cells (1). T$_H$1-initiated destruction of beta cells can be modulated by Foxp3$^+$ regulatory T cells (T$_{REG}$), 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 β (TGF-β) and interleukin 10 (IL-10). The protective role of T$_{REG}$ cells is impaired in susceptible population, and studies have demonstrated a deficiency in number and function of T$_{REG}$ 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 autoimmunity-associated genes *DEX1* (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$_H$1 and T$_{REG}$ cells from T1D

patients, despite of the pivotal roles of T$_H$1 and T$_{REG}$ cells in the pathogenesis of T1D. In this study, we conducted epigenomic and transcriptomic profiling of T$_H$1 and T$_{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$_H$1 and T$_{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$_H$1 and T$_{REG}$ enhancers. We validated the functional roles of four candidate T$_{REG}$ SNPs using a combination of luciferase reporter assay, genome-editing, transcription factor chromatin immunoprecipitation (ChIP), and chromosome conformation capture (3C) assays.

Results

Transcriptome Changes in T$_H$1 and T$_{REG}$ Cells of T1D Patients. Using a panel of established cell surface markers, we purified effector memory T$_{REG}$ cells (CD3$^+$ CD4$^+$ CD25$^+$ CD127$^{dim/-}$ CD45RO$^+$) (9, 10) and effector memory T$_H$1 cells (CD3$^+$ CD4$^+$ CXCR3$^+$ CCR6$^-$ CCR7$^-$ 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).

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 epigenomics 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.

SYSTEMS BIOLOGY

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 T_H1 and T_{REG} cells, respectively [false-discovery rate (FDR) < 0.1, Fig. 1*A* and *SI Appendix*, Table S4]. Consistent with the pathogenic role of T_H1 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 T_{REG} 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. 1*C*).

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 T_H1 cells, and tyrosine kinase 2 (*TYK2*), *IL10*, and major histocompatibility complex, class II, DQ β1 (*HLA-DQB1*) in T_{REG} cells (Fig. 1*B*). 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 chemokinase receptor 4 (*CCR4*), and signal transducer and activator of transcription 5A/5B (*STAT5A/B*) in T_H1, and interleukin-5 (*IL5*), IFN regulatory factor 4 (*IRF4*), and signal transducer and activator of transcription 2 (*STAT2*) in T_{REG}. *STAT5A* and *STAT5B* are reported to have a role in T_H1 cell differentiation (11). *SMAD3*, which is down-regulated in the case group, has a critical role in T_H1 inhibition and immune tolerance (12, 13). Together with *FOXP3* and other transcription factors (TFs), *IRF4* can form a transcriptional network that governs T_{REG} cell differentiation (14). *IL5* is reported to promote induction of antigen-specific T_{REG} cells that suppress autoimmunity and reduced expression of *IL5* will disrupt the immune balance (15) (Fig. 1*B* and *SI Appendix*, Table S4).

Deregulated Enhancers in T_H1 and T_{REG} 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 T_H1 cells and T_{REG} 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 (T_H1, T_{REG}), and histone marks (H3K4me1 and H3K27ac) (*SI Appendix*, Fig. S3A). These correlations are higher than a similar study performed using T_{REG} 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 T_H1 cell, we predicted 13,017 and 12,145 enhancers in case and control groups, respectively. For T_{REG} cell, we predicted 11,915 and 14,860 enhancers in case and control groups, respectively (Fig. 2*A* and *SI Appendix*, Table S6). As expected, 83% of T_H1 enhancers and 74% of T_{REG} enhancers are shared between the two cell types, given their common origin from naive $CD4^+$ T cells (Fig. 2*B*). However, within each cell type, a large fraction of the enhancers has altered activities between the case and control groups (21% for T_H1 and 25% for T_{REG}) (Fig. 2 *A* and *D*). The percentages of group-specific enhancers were significantly higher than expected (null distribution computed by
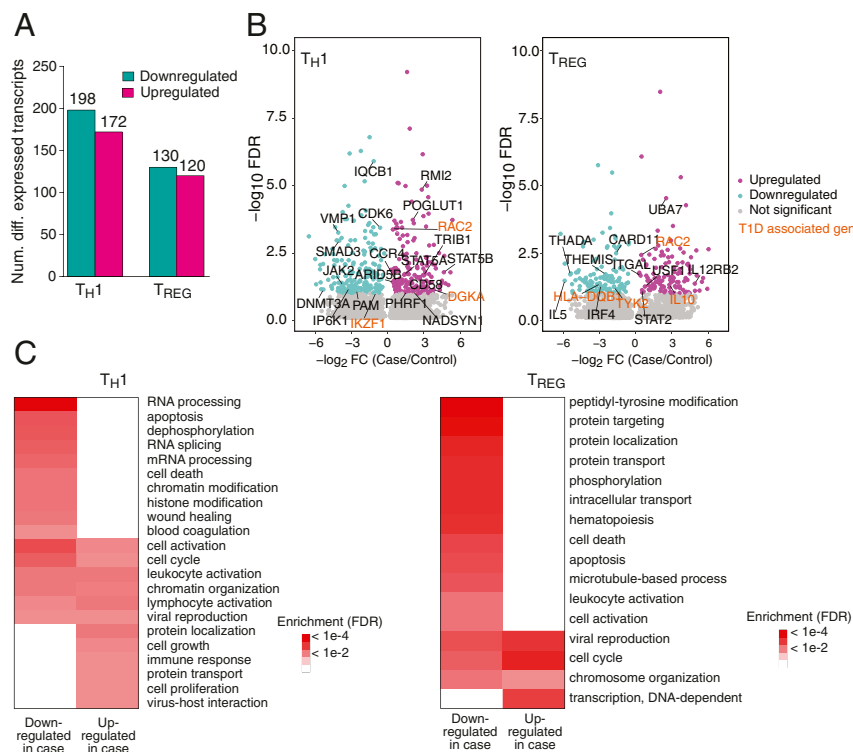

Fig. 1. Differentially expressed genes between case and control subjects in T_H1 and T_{REG} cells. (*A*) Number of differentially expressed transcripts. FDR < 0.1, computed using EBSeq and corrected for multiple testing using the Benjamini–Hochberg (BH) method. Down-regulated, Expression is lower in the case group. Up-regulated, Expression is higher in the case group. (*B*) Volcano plot showing differentially expressed transcripts. Turquoise, Down-regulated in case group; magenta, up-regulated in case group. Genes associated with autoimmune disease are labeled. Genes associated with T1D are highlighted in orange. (*C*) Enriched GO terms among sets of differentially expressed genes in T_H1 (*Left*) and T_{REG} cells (adjusted hypergeometric test, *P* < 0.05, corrected for multiple testing using the BH method).
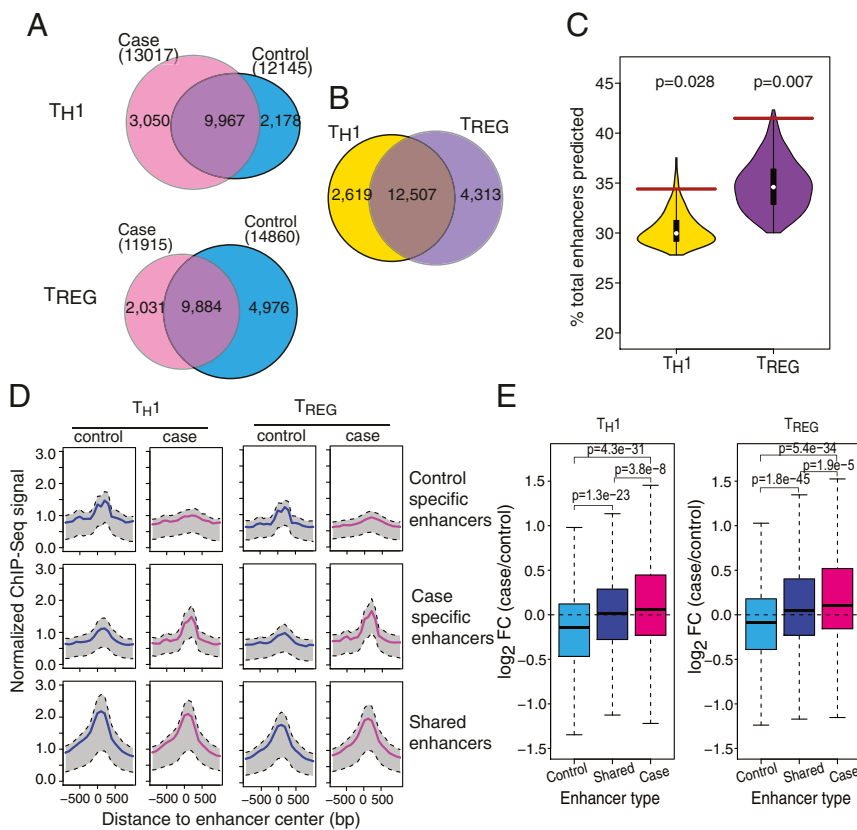
Fig. 2. Enhancer repertoires of T_H1 and T_{REG} cells are significantly altered in T1D patients. (*A*) Overlap of enhancers between case and control subjects. Enhancers overlapping by >500 bp are considered to be overlapped. (*B*) Overlap of enhancers between T_H1 and T_{REG} cells. For each cell type, the total number of enhancers in case and control subjects is used. (*C*) *P* values for observed numbers of group-specific enhancers. Violin plots show the background distribution based on permutated ChIP-seq data. The brown horizontal lines show the observed percentage of group-specific enhancers. (*D*) Pile-up plots of histone mark H3K27ac signals associated with case-specific, control-specific, and shared enhancers. Signals are normalized by library size and input. (*E*) Enhancer activity profile is correlated with target gene expression profile (*P* values are calculated using two-sided Wilcoxon rank-sum test).

permuting the sample labels in the dataset; Fig. 2*C*). 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*A*). 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. 2*E*), 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 *RASGRP1* 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 T-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. 3*A*).

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. S4*B*). 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. 3*B*), 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
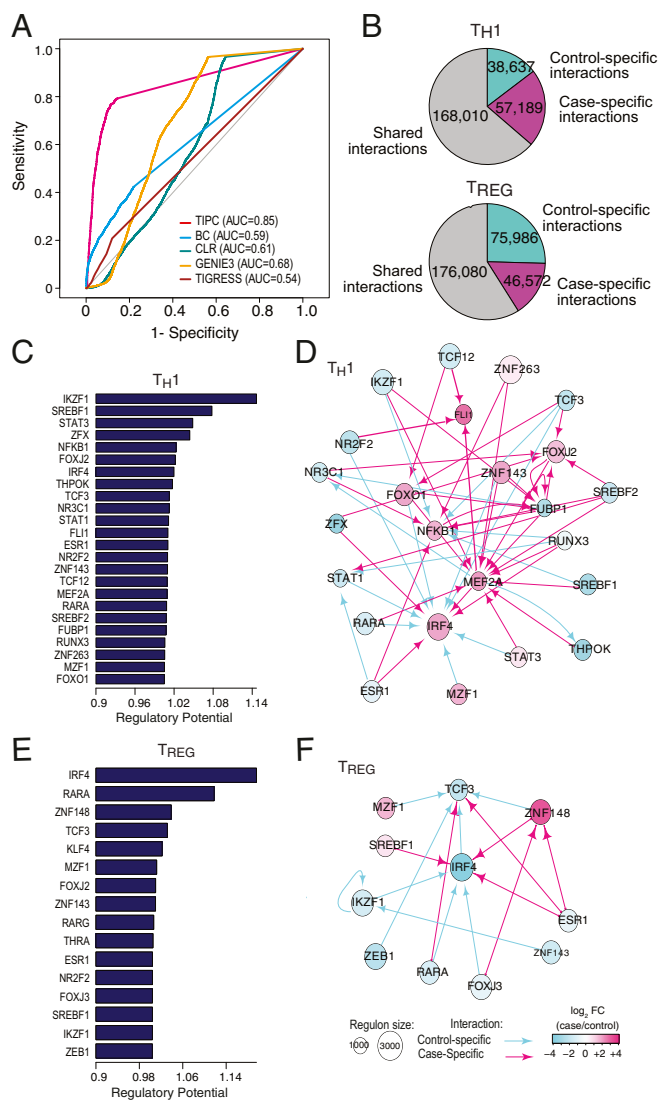
Fig. 3. Transcriptional regulatory networks (TRNs) and key transcription factors. (*A*) Benchmarking of our method for constructing TRN. BC, Pearson correlation between the gene pairs. Other methods are CLR, GENIE3, and TIGRESS (*Methods*). (*B*) Distribution of condition-specific and shared TF–target interactions for the TRNs constructed. (*C* and *E*) Predicted key TFs that play critical roles in T1D onset for T_H1 (*C*) and T_{REG} (*E*). Bars show the median proximity (1/distance) of TF to differentially expressed targets. (*D* and *F*) Condition-specific interactions between the key TFs for T_H1 (*D*) and T_{REG} (*F*). Control-specific interactions are denoted with turquoise, whereas case-specific interactions are shown with magenta. Shared interactions between case and control are not shown for better readability. The arrows show the direction of interaction (from TF to target). Node sizes are proportional to size of the TF regulon. Node colors show expression change of TFs between control and case.

differential expression into a distance measure such that the distance between two highly differentially expressed targets is very short. As a result, TFs that have shorter median distance to the set of differentially expressed targets are ranked higher (see *Methods* for details).

We identified 24 and 16 key TFs in T_H1 and T_{REG} cells, respectively (Fig. 3 *C* and *E*). Thirty-seven percent of T_H1 TFs and 56% T_{REG} TFs are shared between the two cell types. Consistent with their role as key regulators, these TFs target more other TFs than nonkey TFs ($P < 0.001$, *t* test; *SI Appendix*, Fig. S5). Most of the identified TFs have a reported role in either the biology of

T_H1 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 T_H1 and *IRF4* and *RARA* for T_{REG}. *IKZF1* plays a role in the T_H1 versus T_H2 polarization (25, 26). A SNP (rs10272724) in the 3′-UTR of *IZKF1* 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 T_H1 cells. All-*trans* retinoic acid (ATRA), an endogenous ligand of RARA, can prevent human natural T_{REG} cells from converting to T_H1/T_H17 cells and sustains their suppressive function in inflammatory environments (30). Likewise, RARA can sustain T_H1 cell lineage stability and prevents transition to a T_H17 cell program in vivo (31). Consistent with the role of *RARA* in T_H1 and T_{REG} cells, *RARA* level decreases in both cell types (fold changes are 0.55 and 0.08 for T_H1 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 T_H1 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 T_H1 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, that 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 T_H1- 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 T_H1 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. 4*A*, *Left*). Since hypergeometric test does not take into account linkage disequilibrium, we used an alternative approach, linkage disequilibrium score regression (36, 37), to evaluate the statistical significance. We again observed that T_{REG} and T_H1 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. 4*A*, *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 T_H1 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.
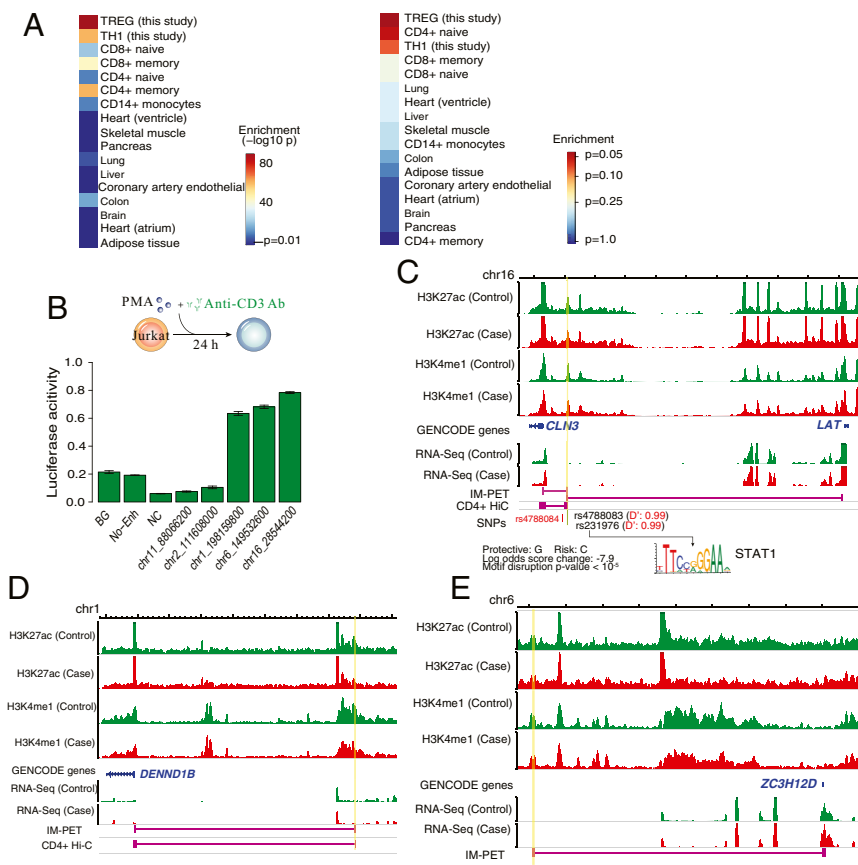
Fig. 4. T_H1 enhancers validated by luciferase reporter assay. (*A*) T_H1 and T_{REG} enhancers have higher degree of overlap with T1D-associated SNPs. (*Left*) Overlap *P* values were computed using hypergeometric test. (*Right*) Overlap *P* values were computed using linkage disequilibrium score regression. (*B*, *Top*) Schematic for the protocol for differentiating T_H1 from Jurkat cells. (*B*, *Bottom*) Luciferase reporter activity of candidate enhancers. Enhancers are labeled by the genomic coordinate of the center of the enhancer. Enhancer activity is relative to cotransfected *Renilla* control. BG, No DNA vector control; NC, negative control, genomic region without H3K4me1 and H3K27ac signals. Data shown are means ± SEM of six replicates. (*C*) Genome browser tracks for the 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 interactions predicted by IM-PET algorithm, supporting capture-Hi-C data [Javierre et al. (38)] for the enhancer–promoter interaction (if any), and SNP locations. In addition, TF motif and binding sites (if any) that are significantly affected by the SNP(s) are also shown. There are two SNPs, rs4788083 and rs231976, in this enhancer. rs231976 perturbs the binding site of STAT1 (BH-adjusted empirical value of $P < 10^{-5}$; *Methods*). D′, Measurement of linkage disequilibrium between an enhancer SNP and the lead T1D GWAS SNP rs4708084 (shown in red). (*D*) Enhancer located at chr1: 198158800–198160800 targeting the *DENND1B* gene. (*E*) Enhancer located at chr6:149531600–149533600 targeting the *ZC3H12D* gene.

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 mark signals between case and control in T_H1 and T_{REG} 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 (*SI Appendix*, Figs. S7 and S8 and 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 T_H1 cells (39) and a Foxp3-expressing Jurkat cell line as the model for T_{REG} 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 T_H1 enhancers, three of them showed significant activity compared with the negative controls (Fig. 4*B*, $P < 0.01$, 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. 4*C*) 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 $P < 10^{-5}$, corrected for multiple testing with Benjamini–Hochberg (BH) method], a key TF involved in the IL-27–mediated early commitment to the T_H1 lineage (45).

The second validated enhancer (chr1:198158800–198160800) is predicted to target the gene DENN domain containing 1B (*DENND1B*) (Fig. 4*D*). 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 (*ZC3H12D*), and there is significant reduction in H3K4me1 level in T1D patients compared with healthy individuals ($P < 0.01$; Fig. 4*E*). *ZC3H12D* is an RNase that destabilizes a set of mRNAs. *ZC3H12D*$^{-/-}$ mice exhibit multiorgan inflammation (48) and autoimmune disease (49). A set of genes, including *IL6*, *IL12*, *REL*, *TNFRSF4*, and *IL2*, which are important for the activation of T_H1 cells, are targets of *ZC3H12D* (49).

For T_{REG} cells, all five enhancers show significant activity compared with the negative controls (Fig. 5*A*; $P < 0.01$, Wilcoxon rank-sum test). The first enhancer (chr12: 9908600–9911600) targets *CD69*, *CLEC2B*, *KLRB1*, and *EIF2S3L*, 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,
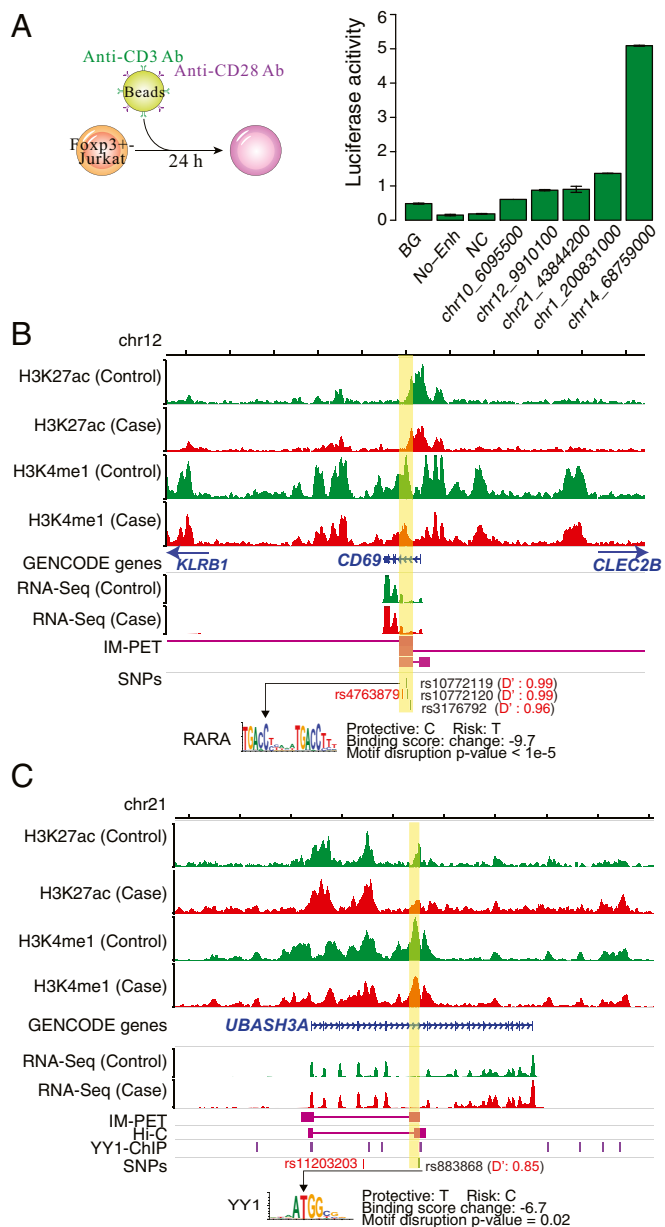
Fig. 5. T_{REG} enhancers validated by reporter assay. (*A, Left*) Schematic for the protocol for differentiating T_{REG} from Foxp3⁺-Jurkat cells. (*A, Right*) Luciferase reporter activity of candidate enhancers. Enhancers are labeled by the genomic coordinate of the center of the enhancer. Enhancer activity is relative to cotransfected *Renilla* control. BG, No DNA vector control; NC, negative control, genomic region without H3K4me1 and H3K27ac signals. Data shown are means ± SEM of six replicates. (*B*) Genome browser tracks for the enhancer located at chr12: 9908600–9911600 (highlighted with yellow). The SNP rs10772119 affects the binding site of RARA (BH-adjusted empirical value of $P < 10^{-5}$). D′, Measurement of linkage disequilibrium between an enhancer SNP and the lead T1D GWAS SNP rs4763879 (shown in red). (*C*) Enhancer located at chr21: 43843200–43845200. The SNP rs883868 affects the binding site of YY1 (BH-adjusted empirical P value, 1.8×10^{-2}). D′, Measurement of linkage disequilibrium between the enhancer SNP rs883868 and the lead T1D GWAS SNP rs11203203 (shown in red).

respectively). The lead SNP is located in the *CD69* gene body (50) (Fig. 5*B*). 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. 5*C*).

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 T_{REG} cells. *CD69* is required for the repressive function of T_{REG} cells (51). Diminished *CD69* expression is associated with compromised function of T_{REG} 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 T_{REG} have not yet been investigated. Nevertheless, *KLRB1* could be of potential interest because it defines a subset of T_{REG} 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* has been reported to negatively regulate the function of T_{REG} 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 T_{REG} 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 T_{REG} and results in a significant reduction of T_{REG} cells in mouse (59).

Two Enhancer SNPs Affecting RARA and YY1 Binding at T_{REG} Enhancers Regulating Immune Response Genes. There are 25 SNPs in the five validated T_{REG} 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 T_{REG} 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. 6*A*).

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. 6*B*). 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. 6*C*). Taken together, these results are consistent with the association between the genotype and gene expressions we found in our cohort.
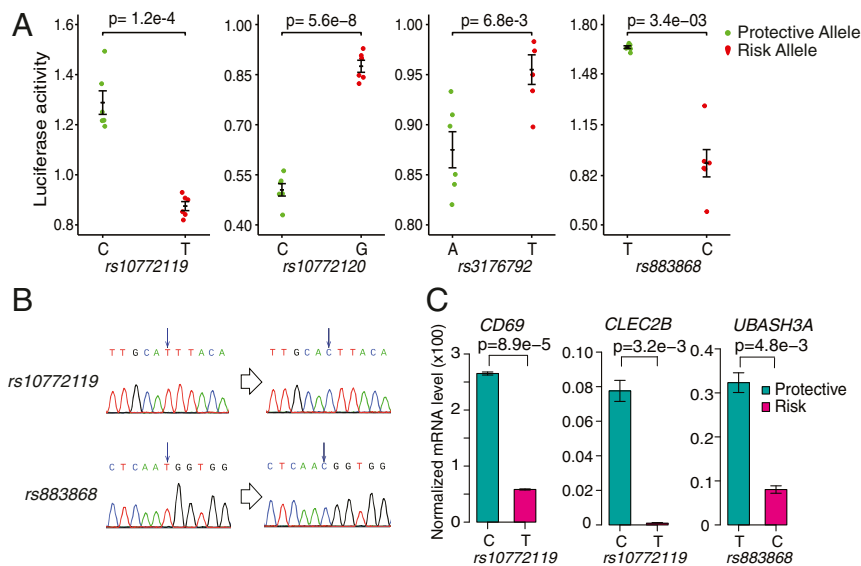
Fig. 6. Four risk SNPs affect enhancer activity and target gene expression. (*A*) Luciferase reporter activity for enhancers containing risk SNPs with protective or risk allele. The green dots are replicates of protective genotype. The red dots represent risk genotype. Data shown are means ± SEM of six replicates. *P* values are computed using Wilcoxon rank-sum test, two-sided. (*B*) Sanger sequencing confirmation of genome-edited cells. (*C*) Effect of genotypes on the expression of target genes in T_{REG} cells measured by RT-qPCR. CRISPR-mediated mutations of three SNPs in T_{REG} cells affect the expression of target genes. The turquoise bars represent protective genotypes, and the magenta bars represent risk genotypes. mRNA level of each gene was normalized to ACTB gene, and relative levels were calculated using the ΔΔCt method. Data shown are means ± SEM of two biological replicates. *P* values are computed using Wilcoxon rank-sum test, two-sided.

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 T_{REG} 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. 7*A*). 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. 5*B*) and rs883868 exhibits allelic-specific binding by YY1 (Fig. 5*C*). For both TFs, we performed ChIP-qPCR to confirm the binding of the TF to the predicted binding sites (Fig. 7*B*). 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. 7*C*), 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. 7*A*).

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 T_H1 and T_{REG} 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 T_{REG} cell-dependent immune tolerance by suppressing both CD4⁺ and CD8⁺ T effector cells, while promoting T_{REG} cell expansion (67). Supporting this mechanism, ATRA cannot exert the protective role in mice with abrogated T_{REG} 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*, *CELE2B*, *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 T_{REG} cell lines engineered to have a C
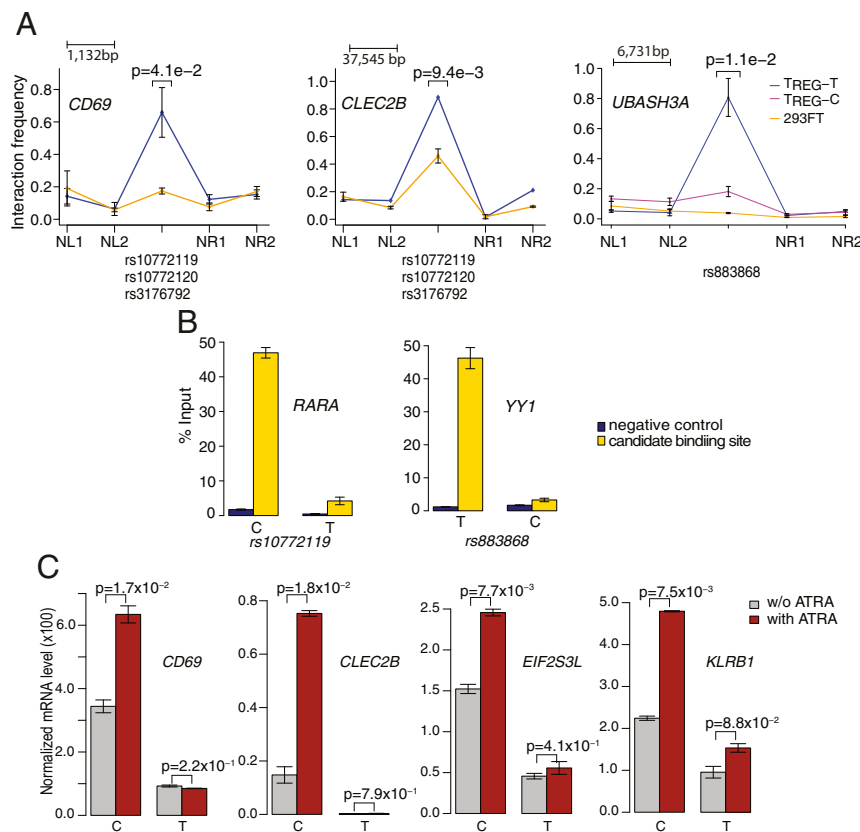
Fig. 7. Two risk SNPs disrupt the binding of RARA and YY1 to T_{REG} enhancers. (*A*) Chromatin interaction between SNP-bearing enhancers and promoters of predicted target genes. 3C-qPCR experiment was performed for each enhancer–promoter pair in Foxp3$^+$-Jurkat cells. 293FT cells were used as negative control. For each pair of enhancer and promoter, six primers were designed, including two primers for the promoter and enhancer and four primers surrounding the enhancer (NL1, NL2, NR1, and NR2). Each consecutive primer pairs have roughly equal distance between them. Data shown are means ± SEM of two biological replicates. For enhancer bearing the YY1 binding site and the SNP rs883868, 3C-qPCR was performed using genome-edited cells having homozygous "T" (T_{REG}-T) or "C" (T_{REG}-C) allele at rs883868 to examine the effect of the SNP on enhancer–promoter interaction. Data shown are means ± SEM of two replicates. All *P* values are calculated with one-sided *t* test. (*B*) Change in TF binding at SNPs with homozygous alleles created by CRISPR-based genome editing of Foxp3$^+$-Jurkat cells. Shown are normalized ChIP-qPCR value expressed as percent of input. Negative control site, Genomic region without any histone modification signal. Data shown are means ± SEM of four replicates. (*C*) Target genes of RARA enhancer are responsive to treatment of ATRA in Foxp3$^+$-Jurkat cells. Genome-edited cells bearing homozygous alleles at rs10772119 were treated with ATRA and target gene expression was measured using RT-qPCR. Data shown are means ± SEM of two replicates. All *P* values are calculated with one-sided *t* test.

allele at rs10772119 (Fig. 7*C*). Taken together, our finding provides a mechanistic explanation for which rs10772119 promotes the onset of T1D via ATRA signaling in T_{REG}.

YY1 is a zinc finger protein that functions either as a transcriptional activator or repressor. Research in rats shows that YY1 may be associated with decreased incidence of T1D (71, 72). We found that the risk SNP rs883868 can disrupt the binding of YY1 in T_{REG} cells, which leads to lower expression of the target gene *UBASH3A*. In support of the regulation of *UBASH3A* by YY1, YY1 knockout in pro-B cells can decrease the expression of *UBASH3A* (73). 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 T_{REG} 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 T_{REG} and T_H1 Cells from Peripheral Blood. Fifty milliliters of peripheral blood were collected from each subject and processed within 0.5 h

of 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 swinging-bucket rotor without brake. The PBMC layer was transferred to a new 50-mL 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 1 × 10⁷ cells per 50 μL. After adding 5 μL of CD4⁺ T cell biotin–antibody mixture per 10⁷ cells, cells were mixed and incubated for 5 min at 4 °C. Thirty-five microliters of 1× staining buffer per 10⁷ cells and 10 μL of CD4⁺ T Cell MicroBead mixture per 10⁷ cells were added and mixed. After incubation for 10 min at 4 °C, cells were diluted to 4 mL and applied to autoMACS Separator for negative selection of CD4⁺ cells using the "Deplete" program. In total, 0.2 × 10⁶ cells were taken out as unstained cells, and the remaining cells were pelleted by centrifugation at 300 × g, 20 °C for 7 min. Supernatant was discarded and cells were resuspended in 1× staining buffer 1 to 2 × 10⁷ cells per mL. Cells were incubated with an antibody mixture (anti-CD3, anti-CD4, anti-CD25, anti-CD45RO, anti-CD127, anti-CCR6, anti-CCR7, anti-CXCR3) [1:400 (vol/vol)] (see list of antibodies in *SI Appendix*, Table S14) at 20 °C in the dark for 30 min and washed with 1× staining buffer twice. Cells were resuspended in 1× staining buffer 1 to a final concentration of 1 × 10⁷ cells per mL and sorted on an Aria flow cytometer (BD Biosciences) to obtain two cell populations: effector memory T_REG cells (CD3⁺CD4⁺CD25⁺CD127^dim/−CD45RO⁺) and effector memory T_H1 cells (CD3⁺CD4⁺CXCR3⁺CCR6⁻CCR7⁻CD45RO⁺). Cells were collected in 0.5 mL of 1× staining buffer 2 [1× DPBS plus 1% human AB serum (Sigma)] within 1.5-mL siliconized low-retention microcentrifuge tubes (Fisher) coated with human AB serum before use. The purities of sorted cells were over 99% (*SI Appendix*, Fig. S1). Sorted cells were used for ChIP-seq and RNA-seq.

Low-Input ChIP-Seq. Purified cells were crosslinked with 1% formaldehyde (Thermo Scientific) in 1× fixing buffer for 5 min at room temperature according to the vendor's manual (Covaris). Cross-linked cells were then resuspended in 1× shearing buffer and sonicated using Covaris E220 for 780 s using the following settings: duty factor, 5%; peak incident power, 105; cycles per burst, 200. Five percent of sheared chromatin was used as the input, and the remaining chromatin was used for immunoprecipitation (IP). IP was performed using the ChIP-IT high-sensitivity kit (Active Motif). IP and input samples were treated with RNase A followed by proteinase K treatment. Cross-linking was reversed by incubation overnight at 65 °C. Reverse cross-linked DNA was purified using a MinElute PCR purification kit (Qiagen). All ChIPed DNA and 1 ng of input DNA were used for sequencing library preparation using the ThruPLEX-FD Prep Kit (Rubicon Genomics). Ten to 12 PCR cycles were used for IP DNA and 9 PCR cycles were used for input DNA at step 5 of the library preparation protocol. Libraries were sequenced on an Illumina HiSeq 2500 sequencer in single-end mode with a read length of 50 nt.

CRISPR/Cas9-Mediated Genome Editing. Single-guide RNAs targeting the SNPs (rs10772119 and rs883868) were designed using CRISPOR (crispor.tefor.net) and cloned into the CRISPR vector pX459 (Addgene plasmid no. 448139). The 91-nt single-stranded DNA oligonucleotide (ssODN) repair templates were designed with homology arms centered around the SNPs (*SI Appendix*, Table S15). CRISPR/Cas9-mediated genome editing of Foxp3⁺-Jurkat cells was performed as described (77) with some modifications. Briefly, plasmids and ssODN templates were cotransfected to Foxp3⁺-Jurkat cells using TransIT-LT1 (Mirus Bio). Twenty-four hours posttransfection, cells were selected by puromycin for 72 h. Dead cells were removed by centrifugation at 300 × g, at room temperature for 5 min and discarding supernatant. Cell concentration and viability were measured using Countess II (Life Technologies). Live cells were resuspended at a concentration of five cells per mL. One hundred microliters of cell suspension was transferred to a 96-well plate and cultured for 2–3 wk for clonal cell line isolation. Genomic DNA was isolated using Quick-DNA 96 kit (ZYMO) and used for screening for desired mutations using allele-specific PCR and fluorescence melting curves as described (78). Genomic DNA of positive clones were further confirmed by Sanger sequencing.

Single-Nucleotide Variants Associated with T1D. The list of T1D-associated SNPs was obtained from the ImmunoBase. The database reports 59 genomic loci containing 119 lead SNPs that are associated with T1D from GWAS studies. We identified linkage disequilibrium blocks using the Haploview software (79) and the genotyping data from 1000 Genomes project (80). We used the default parameter setting except for *blockout*, which was set to "ALL." All SNPs that fall within the same linkage disequilibrium block as the T1D-associated GWAS lead SNP were also considered as candidate risk SNPs.

Code Availability. All published software used for data processing are open source. Additional custom Perl, Python, R, and Unix shell scripts are developed for data processing and analysis. All custom scripts are available upon request.

Data Availability. The data reported in this article have been deposited in the Gene Expression Omnibus database (accession no. GSE112342).

1. Katz JD, Benoist C, Mathis D (1995) T helper cell subsets in insulin-dependent diabetes. *Science* 268:1185–1188.
2. Brusko TM, Wasserfall CH, Clare-Salzler MJ, Schatz DA, Atkinson MA (2005) Functional defects and the influence of age on the frequency of CD4⁺ CD25⁺ T-cells in type 1 diabetes. *Diabetes* 54:1407–1414.
3. Leikfoss IS, et al. (2013) Multiple sclerosis-associated single-nucleotide polymorphisms in CLEC16A correlate with reduced SOCS1 and DEXI expression in the thymus. *Genes Immun* 14:62–66.
4. Ng DP, et al. (2002) Minor effect of GLUT1 polymorphisms on susceptibility to diabetic nephropathy in type 1 diabetes. *Diabetes* 51:2264–2269.
5. Fishman D, et al. (1998) The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. *J Clin Invest* 102:1369–1376.
6. Farh KK, et al. (2015) Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 518:337–343.
7. Onengut-Gumuscu S, et al.; Type 1 Diabetes Genetics Consortium (2015) Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet* 47:381–386.
8. Peng G, et al. (2018) Novel risk variants affecting enhancers of TH1 and TREG cells in type 1 diabetes. Gene Expression Omnibus. Available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112342. Deposited March 26, 2018.
9. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E (2013) The who's who of T-cell differentiation: Human memory T-cell subsets. *Eur J Immunol* 43:2797–2809.
10. Rosenblum MD, Way SS, Abbas AK (2016) Regulatory T cell memory. *Nat Rev Immunol* 16:90–101.
11. Murphy KM (2008) Permission to proceed: Jak3 and STAT5 signaling molecules give the green light for T helper 1 cell differentiation. *Immunity* 28:725–727.
12. Giroux M, et al. (2011) SMAD3 prevents graft-versus-host disease by restraining Th1 differentiation and granulocyte-mediated tissue damage. *Blood* 117:1734–1744.
13. Takimoto T, et al. (2010) Smad2 and Smad3 are redundantly essential for the TGF-beta-mediated regulation of regulatory T plasticity and Th1 development. *J Immunol* 185:842–855, and erratum (2011) 186:636.
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⁺ T regulatory cells that suppress autoimmunity. *Blood* 119:4441–4450.
16. Arvey A, et al. (2015) Genetic and epigenetic variation in the lineage specification of regulatory T cells. *eLife* 4:e07571.
17. Firpi 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 (2014) Global view of enhancer-promoter interactome 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 (Basel)* 4:499–521.
20. Fang D, Zhu J (2017) Dynamic balance between master transcription factors determines the fates and functions of CD4 T cell and innate lymphoid cell subsets. *J Exp Med* 214:1861–1876.
21. Haned H, Benschop CCG, 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. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P (2010) 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* 116:1327–1336.
26. Thomas RM, et al. (2010) Ikaros silences T-bet expression and interferon-gamma production during T 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* 60:1041–1044.
28. Angela M, et al. (2016) Fatty acid metabolic reprogramming via mTOR-mediated inductions of PPARγ directs early activation of T cells. *Nat Commun* 7:13683.

SYSTEMS BIOLOGY

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 all-trans retinoic acid in stabilizing human natural regulatory T cells under inflammatory conditions. *Proc Natl Acad Sci USA* 111: E3432–E3440.
31. Brown 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: 6338.
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. Bulik-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 ppromoters. *Cell* 167:1369–1384.
39. Smeets RL, et al. (2012) Molecular pathway profiling of T lymphocyte signal transduction pathways; Th1 and Th2 genomic 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. Baixauli 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-bet through activation of STAT1 during initial Th1 commitment. *J Immunol* 170: 4886–4890.
46. Fiuza BSD, et al. (2017) Polymorphisms in DENND1B gene are associated with asthma and atopy phenotypes in Brazilian children. *Mol Immunol* 90:33–41.
47. Nakamura M (2012) Analysis of disease-pathway by identifying susceptible genes to primary biliary cirrhosis. *Nihon Rinsho Meneki Gakkai Kaishi* 35:503–510, and erratum (2013) 36:1.
48. Miao R, et al. (2013) Targeted disruption of MCPIP1/Zc3h12a results in fatal inflammatory disease. *Immunol Cell Biol* 91:368–376.
49. Uehata T, et al. (2013) Malt1-induced cleavage of regnase-1 in CD4+ helper T cells regulates immune activation. *Cell* 153:1036–1049.
50. MacArthur J, et al. (2017) The new NHGRI-EBI catalog of published genome-wide association studies (GWAS catalog). *Nucleic Acids Res* 45:D896–D901.
51. Cortés JR, et al. (2014) Maintenance of immune tolerance by Foxp3+ regulatory cells requires CD69 expression. *J Autoimmun* 55:51–62.
52. Radstake TR, et al. (2009) Increased frequency and compromised function of T regulatory cells in systemic sclerosis (SSc) is related to a diminished CD69 and TGFbeta expression. *PLoS One* 4:e5981.
53. Ge Y, Paisie TK, Newman JRB, McIntyre LM, Concannon P (2017) UBASH3A mediates risk for type 1 diabetes through inhibition of T-cell receptor-induced NF-κB signaling. *Diabetes* 66:2033–2043.
54. Carpino N, et al. (2004) Regulation of ZAP-70 activation and TCR signaling by two related proteins, Sts-1 and Sts-2. *Immunity* 20:37–46.
55. Pesenacker AM, et al. (2013) CD161 defines the subset of FoxP3+ T cells capable of producing proinflammatory cytokines. *Blood* 121:2647–2658.
56. Garg G, et al. (2012) Type 1 diabetes-associated IL2RA variation lowers IL-2 signaling and contributes to diminished CD4+CD25+ regulatory T cell function. *J Immunol* 188: 4644–4653.
57. Wang L, et al. (2013) Mbd2 promotes foxp3 demethylation and T-regulatory-cell function. *Mol Cell Biol* 33:4106–4115.
58. Ritter AT, Angus KL, Griffiths GM (2013) The role of the cytoskeleton at the immunological synapse. *Immunol Rev* 256:107–117.
59. Abu-Eid R, et al. (2014) Selective inhibition of regulatory T cells by targeting the PI3K-Akt pathway. *Cancer Immunol Res* 2:1080–1089.
60. Weirauch MT, et al. (2014) Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158:1431–1443.
61. Weintraub AS, et al. (2017) YY1 is a structural regulator of enhancer-promoter loops. *Cell* 171:1573–1588.
62. Rhee EJ, Plutzky J (2012) Retinoid metabolism and diabetes mellitus. *Diabetes Metab J* 36:167–180.
63. Baena RM, et al. (2002) Vitamin A, retinol binding protein and lipids in type 1 diabetes mellitus. *Eur J Clin Nutr* 56:44–50.
64. Forga L, et al. (2016) Low serum levels of prealbumin, retinol binding protein, and retinol are frequent in adult type 1 diabetic patients. *J Diabetes Res* 2016:2532108.
65. Driscoll HK, et al. (1996) Vitamin A status affects the development of diabetes and insulitis in BB rats. *Metabolism* 45:248–253.
66. Kadison A, et al. (2001) Retinoid signaling directs secondary lineage selection in pancreatic organogenesis. *J Pediatr Surg* 36:1150–1156.
67. Van YH, et al. (2009) All-trans retinoic acid inhibits type 1 diabetes by T regulatory (Treg)-dependent suppression of interferon-gamma-producing T-cells without affecting Th17 cells. *Diabetes* 58:146–155.
68. Stosić-Grujicić S, Cvjetićanin T, Stojanović I (2009) Retinoids differentially regulate the progression of autoimmune diabetes in three preclinical models in mice. *Mol Immunol* 47:79–86.
69. Dawson HD, Collins G, Pyle R, Key M, Taub DD (2008) The retinoic acid receptor-alpha mediates human T-cell activation and Th2 cytokine and chemokine production. *BMC Immunol* 9:16.
70. Hall JA, et al. (2011) Essential role for retinoic acid in the promotion of CD4+ T cell effector responses via retinoic acid receptor alpha. *Immunity* 34:435–447.
71. Klöting N, Klöting 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.
72. Klöting N, Klöting 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.
73. Kleiman E, Jia H, Loguercio S, Su AI, Feeney AJ (2016) YY1 plays an essential role at all stages of B-cell differentiation. *Proc Natl Acad Sci USA* 113:E3911–E3920.
74. Hnisz D, et al. (2016) Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351:1454–1458.
75. Smemo S, et al. (2014) Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 507:371–375.
76. Northcott PA, et al. (2014) Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* 511:428–434.
77. Ran FA, et al. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8: 2281–2308.
78. 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.
79. Barrett JC (2009) Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc* 2009:pdb.ip71.
80. Auton A, et al.; 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526:68–74.