ARTICLES
https://doi.org/10.1038/s41588-019-0524-6
nature genetics

The impact of proinflammatory cytokines on the β-cell regulatory landscape provides insights into the genetics of type 1 diabetes

Mireia Ramos-Rodríguez[1], Helena Raurell-Vila[1], Maikel L. Colli[2], Maria Inês Alvelos[2], Marc Subirana-Granés[1], Jonàs Juan-Mateu[2], Richard Norris[1], Jean-Valery Turatsinze[2], Ernesto S. Nakayasu[3], Bobbie-Jo M. Webb-Robertson[3], Jamie R. J. Inshaw[4], Piero Marchetti[5], Lorenzo Piemonti[6], Manel Esteller[7,8,9,10], John A. Todd[4], Thomas O. Metz[3], Décio L. Eizirik[2] and Lorenzo Pasquali[1,7,11]*

The early stages of type 1 diabetes (T1D) are characterized by local autoimmune inflammation and progressive loss of insulin-producing pancreatic β cells. Here we show that exposure to proinflammatory cytokines reveals a marked plasticity of the β-cell regulatory landscape. We expand the repertoire of human islet regulatory elements by mapping stimulus-responsive enhancers linked to changes in the β-cell transcriptome, proteome and three-dimensional chromatin structure. Our data indicate that the β-cell response to cytokines is mediated by the induction of new regulatory regions as well as the activation of primed regulatory elements prebound by islet-specific transcription factors. We find that T1D-associated loci are enriched with newly mapped *cis*-regulatory regions and identify T1D-associated variants disrupting cytokine-responsive enhancer activity in human β cells. Our study illustrates how β cells respond to a proinflammatory environment and implicate a role for stimulus response islet enhancers in T1D.

In T1D, early inflammation of the pancreatic islets (insulitis) by T and B cells contributes to both the primary induction and secondary amplification of the immune assault, with inflammatory mediators such as the cytokines interleukin-1β (IL-1β) and interferon-γ (IFN-γ) contributing to the functional suppression and apoptosis of β cells[1–3].

Genome-wide association studies (GWAS) have made a substantial contribution to the knowledge of T1D genetic architecture, uncovering >60 regions containing thousands of associated genetic variants. Nevertheless, translating variants to function is a main challenge for T1D and other complex diseases. Most of the associated variants do not reside in coding regions[4], suggesting that they may influence transcript regulation rather than altering protein coding sequences. Recent studies showed a primary enrichment of T1D association signals in T- and B-cell enhancers[4,5]. A secondary[5], or a lack of enrichment, was instead observed in islet regulatory regions. While such observation points to a major role of the immune system, we hypothesize that a subset of T1D variants may also act at the β-cell level but only manifest on islet cell perturbation and are thus not captured by the current maps of islet regulatory elements.

We have now mapped inflammation-induced *cis*-regulatory networks, transcripts, proteins and three-dimensional (3D) chromatin structure changes in human β cells (Fig. 1a). We leverage these data to reveal functional T1D genetic variants as well as key candidate genes and regulatory pathways contributing to β-cell autoimmune destruction. Such analyses permit elucidation of the role of gene regulation and its interaction with T1D genetics in the context of the autoimmune reaction that drives β-cell death.

Results

Proinflammatory cytokines impact the β-cell chromatin landscape. To characterize the effect of proinflammatory cytokines on the β-cell regulatory landscape, we first mapped all accessible or open chromatin sites in human pancreatic islets exposed or unexposed to IFN-γ and IL-1β. We assayed chromatin accessibility by assay for transposase-accessible chromatin using sequencing (ATAC-seq) and, to focus on the β-cell fraction and decrease interindividual variability, in parallel with human pancreatic islet assays, we performed ATAC-seq in the clonal human β-cell line EndoC-βH1 (EC)[6], exposed or unexposed to the proinflammatory cytokines (overall number of peaks identified in human islets: 92,610–229,588; in EC cells: 52,735–110,715; see Extended Data Fig. 1a). Such experiments revealed an important remodeling of β-cell chromatin resulting in approximately 12,500 highly confident chromatin sites that gained accessibility (false discovery rate (FDR)-adjusted $P < 0.05$; $\log_2$ fold change > 1; Extended Data Fig. 1b) on exposure to proinflammatory cytokines. Importantly, the changes

[1]Endocrine Regulatory Genomics Laboratory, Germans Trias i Pujol University Hospital and Research Institute, Badalona, Spain. [2]Center for Diabetes Research and Welbio, Medical Faculty, Université Libre de Bruxelles, Brussels, Belgium. [3]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA. [4]JDRF/Wellcome Diabetes and Inflammation Laboratory, Wellcome Centre for Human Genetics, Nuffield Department of Medicine, NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, UK. [5]Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy. [6]Diabetes Research Institute, San Raffaele Scientific Institute, Milan, Italy. [7]Josep Carreras Leukaemia Research Institute, Barcelona, Spain. [8]Centro de Investigación Biomédica en Red Cáncer, Madrid, Spain. [9]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. [10]Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain. [11]CIBER de Diabetes y Enfermedades Metabólicas Asociadas, Barcelona, Spain. *e-mail: lpasquali@igtp.cat

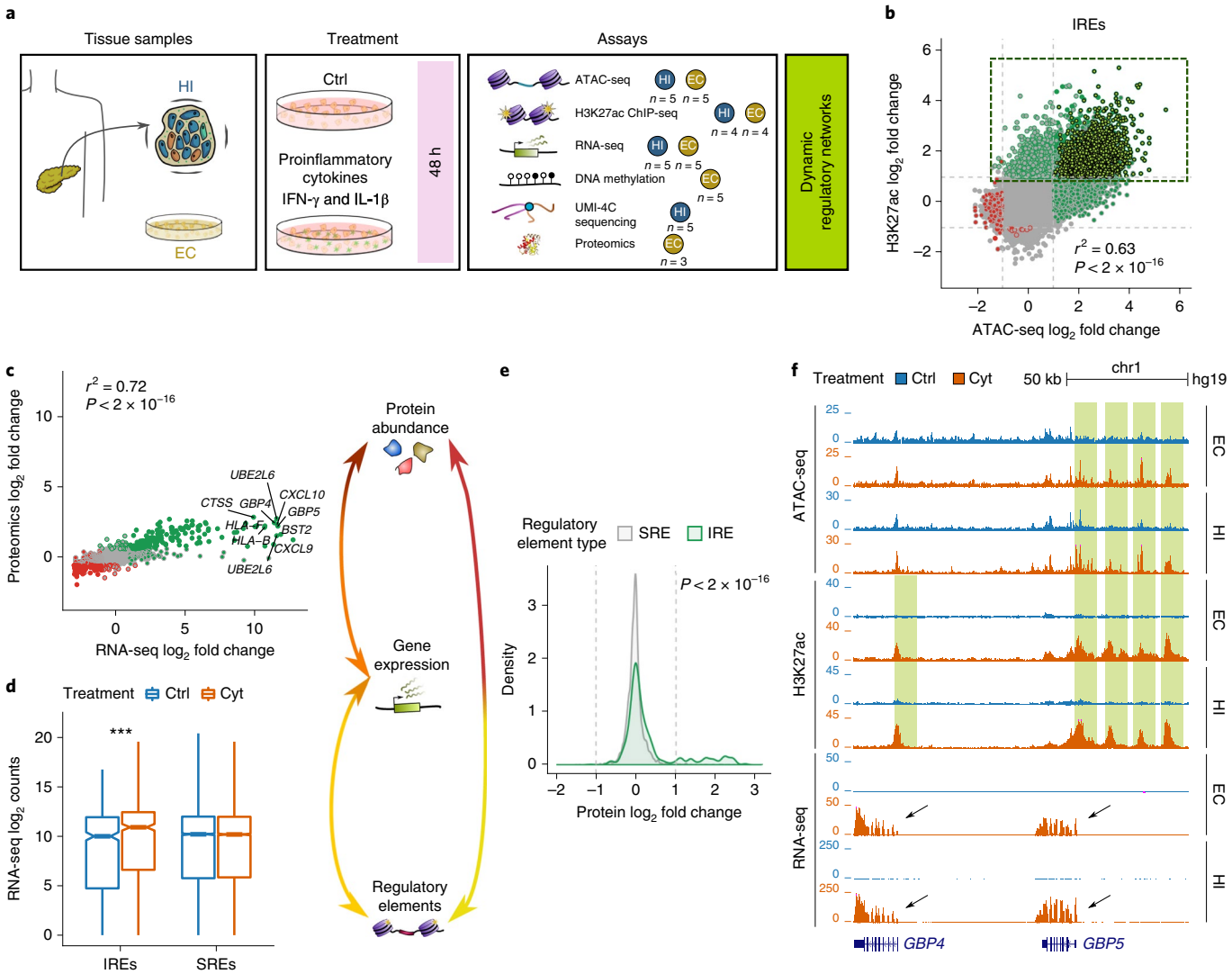


Fig. 1 | Proinflammatory cytokine exposure causes profound remodeling of the β-cell regulatory landscape. a, Summary of the experimental design. The number of EC and human pancreatic islet (HI) samples used in different assays is shown. **b**, Correlation between chromatin accessibility and H3K27ac deposition; each dot corresponds to a chromatin site. The point fill refers to the ATAC-seq; the border refers to the H3K27ac classification (green = gained; red = lost; gray = stable). The dashed box depicts the regulatory elements (IREs) and the lighter shade of green depicts a subtype named neo-IREs (see text). **c**, Correlation between changes in RNA expression and protein abundance in EC cells. The point fill and border indicate the classification of RNA-seq and proteins, respectively (upregulated = green; downregulated = red; equally regulated = gray). **d**, Genes proximal to the IREs (see Methods) show cytokine-induced expression in EC cells exposed or unexposed to proinflammatory treatment. CYT = cytokine-exposed. Two-sided Wilcoxon test ***$P < 0.001$. The boxplot limits show the upper and lower quartiles; the whiskers extend to 1.5× the interquartile range. **e**, Translation of proteins encoded by IRE-associated genes is induced by cytokine exposure in EC cells. This is shown by the significantly different (two-sided Wilcoxon test $P < 2 \times 10^{-16}$) $\log_2$ fold change distribution of protein abundance obtained after cytokine exposure for proteins encoded by genes associated with IREs or SREs. **f**, Representative view of the *GBP4* and *GBP5* genes, encoding the IFN-inducible guanylate binding proteins, illustrating their upregulation on cytokine exposure and the nearby induction of IREs characterized by gains in chromatin accessibility and enrichment in H3K27ac (green boxes).

observed in the human β-cell line were concordant with those observed in the human islet preparations (Extended Data Fig. 1c).

We reasoned that changes in chromatin accessibility may reflect the activation of noncoding *cis*-regulatory elements. Thus, we used chromatin immunoprecipitation with sequencing (ChIP-seq) to map cytokine-induced changes in H3K27ac (Extended Data Fig. 1a), a key histone modification associated with active *cis*-regulatory elements shown to be dynamically regulated in response to acute stimulation[7]. We observed genome-wide deposition of the active histone modification mark on exposure to proinflammatory cytokines in both EC and human pancreatic islets (Extended Data Fig. 1b,c).

Integrative analysis of ATAC-seq and ChIP-seq indicates that changes in chromatin accessibility are strongly correlated with deposition of H3K27ac ($P < 2 \times 10^{-16}$, $r^2 = 0.63$) allowing the identification of approximately 3,800 open chromatin regions that gained H3K27ac (FDR-adjusted $P < 0.05$; $\log_2$ fold change >1) on exposure to proinflammatory cytokines (Fig. 1b and Extended Data Fig. 1d). We found that this subset of open chromatin regions is preferentially located distally to gene transcription start sites (TSS) (Extended Data Fig. 1e), and their sequence is evolutionarily conserved (Extended Data Fig. 1f) and enriched for specific transcription factor binding sites (Extended Data Fig. 1g). We named

these newly mapped regions induced regulatory elements (IREs) (Supplementary Table 1 and Supplementary Table 2).

Chromatin changes link to variation in transcription and translation. We next explored whether the newly identified IREs were associated with changes in gene expression and protein translation. To identify β-cell transcripts and proteins induced by the proinflammatory cytokines, we assayed gene expression by RNA sequencing (RNA-seq; five replicates in EC and five replicates in human pancreatic islets[8]; see Extended Data Fig. 1a) and collected multiplex proteomics data for three EC replicates after exposure or lack of exposure to proinflammatory cytokines.

In line with the chromatin assays, which indicated extensive gene regulatory activation, we unraveled cytokine-induced transcriptional activation resulting in approximately 1,200 upregulated genes (FDR-adjusted $P<0.05$; $\log_2$ fold change >1; Extended Data Fig. 2a,b). By multiplex proteomics, after rigorous filtering, a subset of 10,166 proteins was confidently quantified and retained for significance testing. A total of 348 proteins displayed significant changes in abundance (FDR/$Q<0.15$ and absolute fold change >1.5; absolute $\log_2$ fold change > 0.58). Of the overall detected proteins, 2.19% were upregulated (Extended Data Fig. 2c), 76% of which had induced messenger RNA levels at 48 h, confirming consistency between RNA-seq and protein changes ($r^2=0.72$, $P<2\times10^{-16}$; Fig. 1c). Protein–protein interactions inferred from β-cell cytokine-induced proteins resulted in a network more connected than expected by chance ($P<10^{-3}$) and significantly enriched for Molecular Signatures Database (http://software.broadinstitute.org/gsea/msigdb/) pathways including IFN-γ signaling, antigen processing and presentation, apoptosis and T1D (Kyoto Encyclopedia of Genes and Genomes T1D $P=7.9\times10^{-8}$; Extended Data Fig. 2d).

As expected, we found that IREs were linked to the upregulation of the nearby gene(s) as well as to an induced abundance of the corresponding protein (Fig. 1d,e and Extended Data Fig. 2e). Moreover, gene induction was highly correlated with the number of associated IREs, suggesting a cumulative effect of IREs on cytokine-induced changes in gene expression (Extended Data Fig. 2f).

Taken together, these findings reveal that the pancreatic β-cell response to proinflammatory cytokines is dynamic, involving extensive chromatin remodeling and profound changes in the regulatory landscape (Fig. 1f and Extended Data Fig. 2g). Such changes are associated with induction of transcription and protein translation including pathways implicated in the pathogenesis of T1D. Newly defined regulatory maps can be visualized online along with other islet regulatory annotations at www.isletregulome.org.

Primed and neo-regulatory elements mediate cytokine response. We next sought to gain an insight into the dynamic activation of IREs. The relationship between chromatin openness and H3K27ac deposition on exposure to proinflammatory cytokines allows the distinction of two classes of IREs (Fig. 1b and Fig. 2a–c): opening IREs ($n=2,436$), which gain both chromatin accessibility ($\log_2$ fold change > 1) and H3K27ac ($\log_2$ fold change > 1); and primed IREs ($n=1,362$), which are already accessible chromatin sites before treatment (ATAC-seq $\log_2$ fold change < 1) and gain H3K27ac ($\log_2$ fold change > 1) on exposure to the stimulus. Primed and opening IREs are both associated with gene expression induction (Extended Data Fig. 3a), are phylogenetically conserved (Extended Data Fig. 3b) and preferentially map distally relative to a gene's TSS (Extended Data Fig. 3c). We further revealed that 70% of opening IREs ($n=1,716$), before cytokine exposure, are inactive and inaccessible (that is, undetectable by ATAC-seq under basal conditions; see Methods). We named the latter neo-IREs. Neo-IREs represent 45% of all IREs and may mirror a class of regulatory elements identified on stimulation of mouse macrophages and named 'latent enhancers'[7].

Because chromatin openness, the feature distinguishing the two classes of IREs, is believed to reflect transcription factor occupancy, we analyzed their sequence composition in search of recognition sequences of key transcription factors orchestrating the β-cell response to proinflammatory cytokines. Even though IREs are mostly distal to TSS (Extended Data Fig. 3c), to reduce sequence bias, we excluded all annotated promoters from this analysis. The two classes of distal IREs predominantly mapped to the enhancer chromatin state (Extended Data Fig. 3d) and showed clear differences in sequence composition. Newly induced enhancers were enriched for the binding motifs of inflammatory response transcription factors including IFN-sensitive response element (ISRE), signal transducer and activator of transcription (STAT) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) (Extended Data Fig. 3e). Instead, primed enhancers were enriched for binding motifs of inflammatory response transcription factors (ISRE, STAT) and, unexpectedly, islet-specific transcription factors (HNF1A/B, NEUROD1, PDX1, MAFB, NKX6.1; Extended Data Fig. 3f). Importantly, we found that in primed enhancers, inflammatory response and islet-specific transcription factor binding motifs mapped to the same genomic regions, suggesting cobinding and possibly cooperation of the two classes of transcription factors (Extended Data Fig. 3g,h).

Sequence composition bias per se does not imply transcription factor occupancy. Thus, we took advantage of published ChIP-seq datasets of islet-specific transcription factors (MAFB, PDX1, FOXA2, NKX6.1 and NKX2.2) mapped in unstimulated human pancreatic islets[9] to measure transcription factor occupancy in primed and neo-enhancers before the proinflammatory stimulus. As expected from the sequence composition analysis, primed enhancers (unlike neo-enhancers) are bound by tissue-specific transcription factors even before their activation (Fig. 2d and Extended Data Fig. 3i). Transcription factor occupancy can also be indirectly assessed by ATAC-seq, which assays the protection of the bound sequence to transposase cleavage (footprint). Footprint analysis is effective for transcription factors with a long residence time[10], such as the IFN regulatory factor (IRF) and STAT transcription factor families. Our analyses revealed the emergence of footprint marks on proinflammatory treatment in correspondence to ISRE motifs in both primed and neo-enhancers (Fig. 2e), indicating cytokine-induced transcription factor occupancy of IREs.

Gene regulation is orchestrated by different epigenetic mechanisms. DNA methylation is a relatively stable epigenetic mark contributing to maintenance of cellular identity[11,12]. Moreover, high-resolution DNA methylation maps, obtained from multiple tissues, suggested that the vast majority of tissue-specific, differentially methylated regions are located at distal, mostly noncoding regulatory sites[13]. Consequently, characterization of the DNA methylome in the context of relevant stimuli is important for understanding the functional mechanisms of tissue-specific responses in human disease[14]. Thus, we explored if cytokine-induced chromatin remodeling is associated with changes in DNA methylation. We quantified DNA methylation changes by performing dense methylation arrays in EC cells exposed or unexposed to IFN-γ and IL-1β. The Infinium MethylationEPIC array was designed to interrogate with high precision and coverage >850,000 CpG sites (approximately 3% of all sites in the genome) selected primarily because of their location close to gene promoters and CpG island regions. By focusing on the 1,230 IRE enhancers harboring one or more CpG sites interrogated by the array, we observed that primed enhancers overlap lowly methylated CpGs (median $\beta=0.12\pm0.08$), which did not vary significantly on cytokine exposure. Such observation is in sharp contrast with neo-enhancers, which were highly methylated under control conditions (median $\beta=0.77\pm0.10$), but underwent a significant loss of DNA methylation (two-sided Wilcoxon test, $P=4.13\times10^{-4}$) on treatment (Fig. 2f). While we did not observe cytokine-induced methylation,

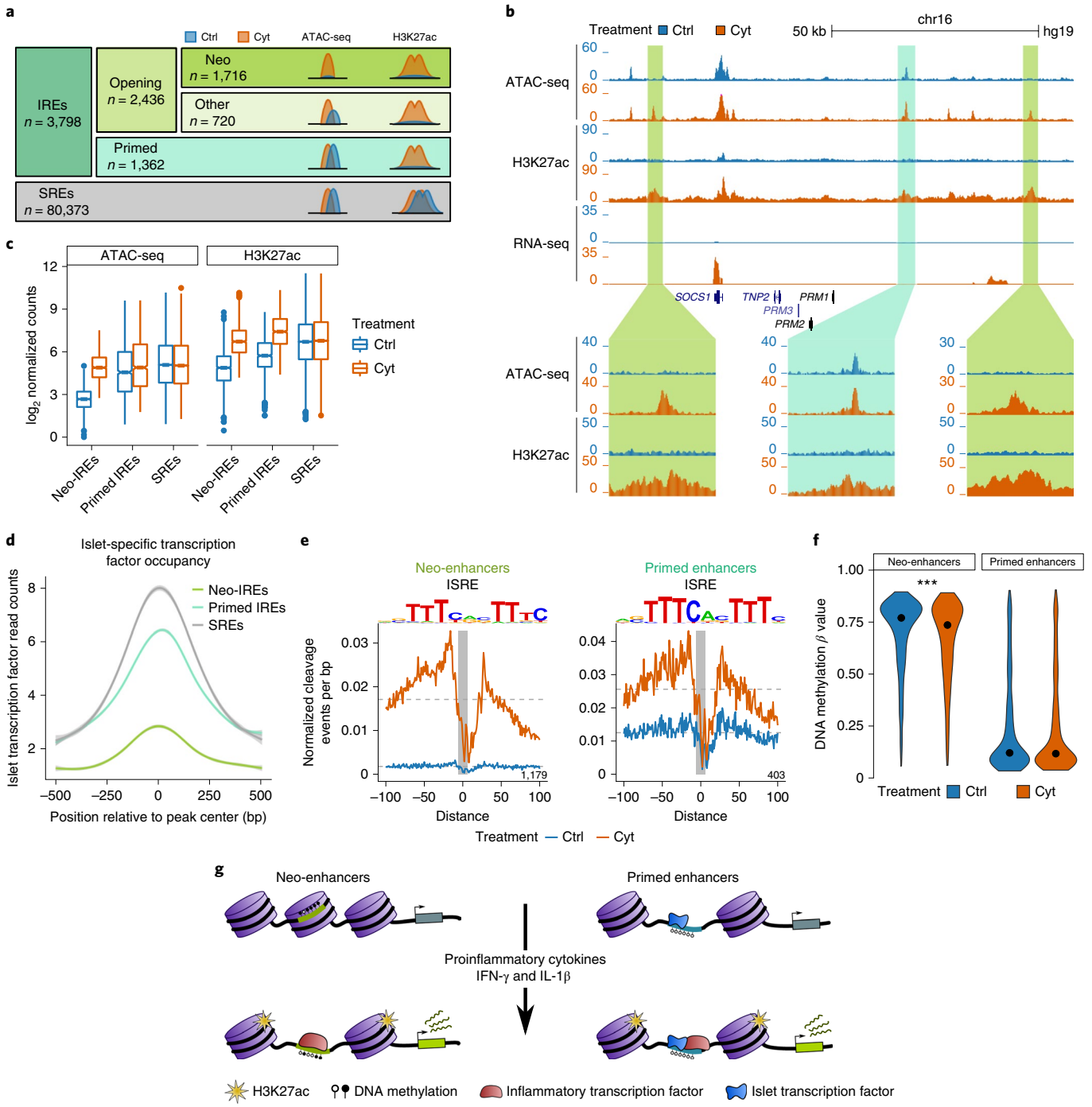


Fig. 2 | The β-cell response to proinflammatory cytokines unveils neo and primed IREs. a, Classification of ATAC-seq open chromatin sites on exposure of human β cells to IFN-γ and IL-1β. **b**, View of the *SOCS1* locus, a gene strongly induced on proinflammatory cytokine exposure. Representative examples of primed (blue box) and neo-IREs (green boxes) are shown. **c**, Boxplot distribution of ATAC-seq and H3K27ac normalized tag counts at different classes of IREs. The boxplot limits show the upper and lower quartiles; the whiskers extend to 1.5× the interquartile range; the individual data points represent the outliers; the notch represents the confidence interval around the median. **d**, Islet-specific transcription factor occupancy at neo, primed and stable regulatory elements. Read density for PDX1, NKX2.2, FOXA2, NKX6.1 and MAFB was calculated in 10 bp bins in 1 kb windows centered on the regulatory element. The lines represent the mean, while the gray shading depicts the s.d. **e**, Footprint analysis of ISRE motifs in neo (left) and primed regulatory elements (right) in cells exposed or unexposed to IFN-γ and IL-1β (blue = Ctrl; orange = cytokines). **f**, Violin plots showing the distribution of DNA methylation β values in neo and primed enhancers, exposed or unexposed to proinflammatory cytokines. Two-sided Wilcoxon test $^{***}P < 0.001$. **g**, Model showing two types of IREs driving the response to proinflammatory cytokines in human β cells.

we found that approximately 70% of significantly demethylated probes (FDR-adjusted $P \leq 0.05$; β-value differences between cytokine-exposed and untreated cells ($\beta_{Cyt} - \beta_{Ctrl}$) < -0.20) mapping to IREs were located at neo-enhancers (Extended Data Fig. 3j,k).

These results suggest that neo-enhancers are enriched for methylated CpGs that undergo preferential demethylation on cytokine treatment, whereas primed enhancers are enriched for unmethylated CpGs that do not change their methylation status on exposure to cytokines.

Taken together these analyses lead to a model, where proinflammatory cytokines elicit a regulatory response in β cells characterized by: (1) new induction of distal regulatory elements coupled with reduction of DNA methylation and binding of inflammatory response transcription factors; and (2) activation of regulatory elements prebound by islet-specific transcription factors and induced by inflammatory response transcription factors (Fig. 2g).

Collectively, these results allow the reconstruction of *cis*-regulatory networks activated in human pancreatic β cells on exposure to the proinflammatory cytokines IFN-γ and IL-1β (Extended Data Fig. 4a–c and Supplementary Table 1).

Changes in islet 3D chromatin structure. Regulatory regions can exert control over genes at megabase distances through the formation of DNA loops. These loops are often confined within structures known as topologically associating domains[15–17]. Topologically associating domains are largely conserved on evolution, are invariant in different cell types and have their boundaries defined by the regulatory scope of tissue-specific enhancers[18,19]. Our knowledge regarding the general characteristics and mechanisms of loops is improving[20–23], but much less is known regarding the mechanisms and functional significance of dynamic looping events during biological processes.

We took advantage of promoter capture Hi-C performed in human pancreatic islets[24] to explore the long-range interactions between gene promoters and cytokine-induced and invariant distant regulatory elements. Interestingly, we observed that the interaction confidence scores captured between IRE enhancers and gene promoters in untreated islets were significantly reduced compared with distal stable regulatory element (SRE) ($P=1.8\times10^{-11}$; Extended Data Fig. 5a). Since this finding pointed to potential dynamic properties of the interaction maps, we next sought to investigate if cytokine-induced regulatory changes are linked to modification of 3D chromatin structure and if induction of β-cell cytokine-responsive regulatory elements is coupled with the new formation of DNA looping interactions.

Hi-C profiles are limited in sequencing coverage and library complexity, resulting in maps of reduced resolution relative to regulatory maps of functional elements. On the other hand, circular chromosome conformation capture (4C) approaches are difficult to interpret quantitatively mainly due to potential amplification bias. Thus, we applied targeted chromosome capture with unique molecular identifiers (UMI-4C), a recently developed method[25], to quantitatively measure interaction intensities in human islets before and after exposure to proinflammatory cytokines. We centered the conformation capture viewpoint at the promoter of 13 genes (*TNFSF10*, *GBP1* and *CIITA*, among others) whose expression was strongly induced by cytokine exposure.

UMI-4C showed marked changes in the 3D chromatin structure at the analyzed loci. Promoters of the induced genes gained chromatin interactions, with distal genomic regions reflecting the formation of new DNA looping events (Fig. 3a,b and Extended Data Fig. 5b–d). Importantly, such new contacts were preferentially engaged with newly mapped human islet cytokine-responsive IREs (Fig. 3c).

These results demonstrate that cytokine exposure induces changes in human islet 3D chromatin conformation including the formation of new enhancer–promoter interactions. Such changes allow the newly activated distal IREs to contact their target gene promoters.

Islet cytokine enhancers are implicated in T1D genetic susceptibility. GWAS have identified approximately 60 chromosome regions associated with T1D[26], with many of the association signals having been assigned to candidate genes with immunological functions. Consistent with this notion, several studies reported a primary enrichment of T1D risk variants in T- and B-cell regulatory elements[4,5]. Furthermore, there is a substantial lack of statistically significant overlap of T1D-associated variants in islet enhancers, while such regulatory elements are instead enriched for GWAS signals for type 2 diabetes (T2D) and fasting glucose[9,27]. Nonetheless, the molecular mechanisms linking T1D association signals to cellular functions are poorly described for most of the regions of association identified.

We hypothesized that a subset of T1D genetic signals may reflect an altered capacity of β cells to react to an inflammatory environment. Thus, we sought to explore to what extent genetic signals underlying T1D susceptibility act through pancreatic islet regulatory response to proinflammatory cytokines.

Causal *cis* variants are expected to be found in sequences that act as regulatory regions in state-specific and disease-relevant tissues. Thus, we examined nonshared loci with genome-wide significant association to T2D and T1D in European populations and considered all variants in high linkage disequilibrium (1000 Genomes Project, phase 3 European population (EUR), $R^2>0.8$) with a lead SNP reported in the National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) GWAS catalog[26]. In line with previous observations[4,9], we found that T2D but not T1D risk variants overlap human islet noncytokine-responsive regulatory elements (that is, SREs) more than expected by chance (T2D SNPs in SREs $P<2\times10^{-16}$, $z=5.47$). In contrast, we found that human islet IREs are enriched for T1D but not T2D risk variants (T1D SNPs in IREs $P=3\times10^{-6}$, $z=4.61$) (Fig. 4a). This result was reproduced when using regulatory elements detected in EC cells (Extended Data Fig. 6a). Such findings revealed 9 T1D-associated regions (13% of the total) containing at least 1 islet cytokine-induced regulatory element directly overlapping a T1D-associated variant (Supplementary Table 3 and Extended Data Fig. 6b–f).

We noted that the two T1D lead SNPs at the 1q24.3 and 16q13.13 loci (rs78037977 (refs. [28,29]) and rs193778 (ref. [4]), respectively) were directly overlapping IREs in the islets. We used GWAS genotyping data from a cohort of 14,575 individuals (5,909 T1D cases and 8,721 controls (Ctrls); see Methods) to confirm their association with T1D. Both variants were included in the 99% credible set of their respective locus and displayed strong association *P* values (rs78037977, $P=6.94\times10^{-10}$; rs193778, $P=1.33\times10^{-7}$; see Supplementary Table 4 for the posterior probability of association and variant ranking in the credible set), indicating that they could potentially be causal.

At the 1q24.3 locus, rs78037977 (NC_000001.10:g.172715702 A>G) overlaps an islet cytokine-induced chromatin site (Fig. 4b), which is prebound by islet-specific transcription factors and is a predicted enhancer in other cell types (Extended Data Fig. 6g). We created allele-specific luciferase reporter constructs and measured enhancer activity in the EC cell line before and after cytokine exposure. The sequence exerts enhancer activity exclusively after cytokine exposure, which is disrupted by the rs78037977 T1D-associated G allele (one-way analysis of variance (ANOVA), $F=26$, $P=4.34\times10^{-5}$; Fig. 4c and Extended Data Fig. 6h). This is consistent with a causal role of the variant at this locus. To identify the gene target of this T1D-susceptible enhancer, we reconstructed the 3D chromatin structure using chromatin capture experiments. UMI-4C in human islets identified a cytokine-induced interaction of the enhancer with *TNFSF18*, a gene activated in islets on cytokine exposure (Fig. 4d,e). *TNFSF18* encodes a cytokine, glucocorticoid-induced tumor necrosis factor receptor-related protein (GITR; also known as TNFRSF18), which modulates the inflammatory reaction and regulation of autoimmune responses[30]. Interestingly, we noted that cytokine exposure results in upregulation of *TNFSF18* in human islets but not in the EC β-cell line, suggesting differences in gene regulatory dynamics in primary tissue or the activation of an islet cell subpopulation.

At the 16q13.13 locus, rs193778 (NC_000016.9:g.11351211 A>G) maps to a phylogenetically conserved, cytokine-responsive

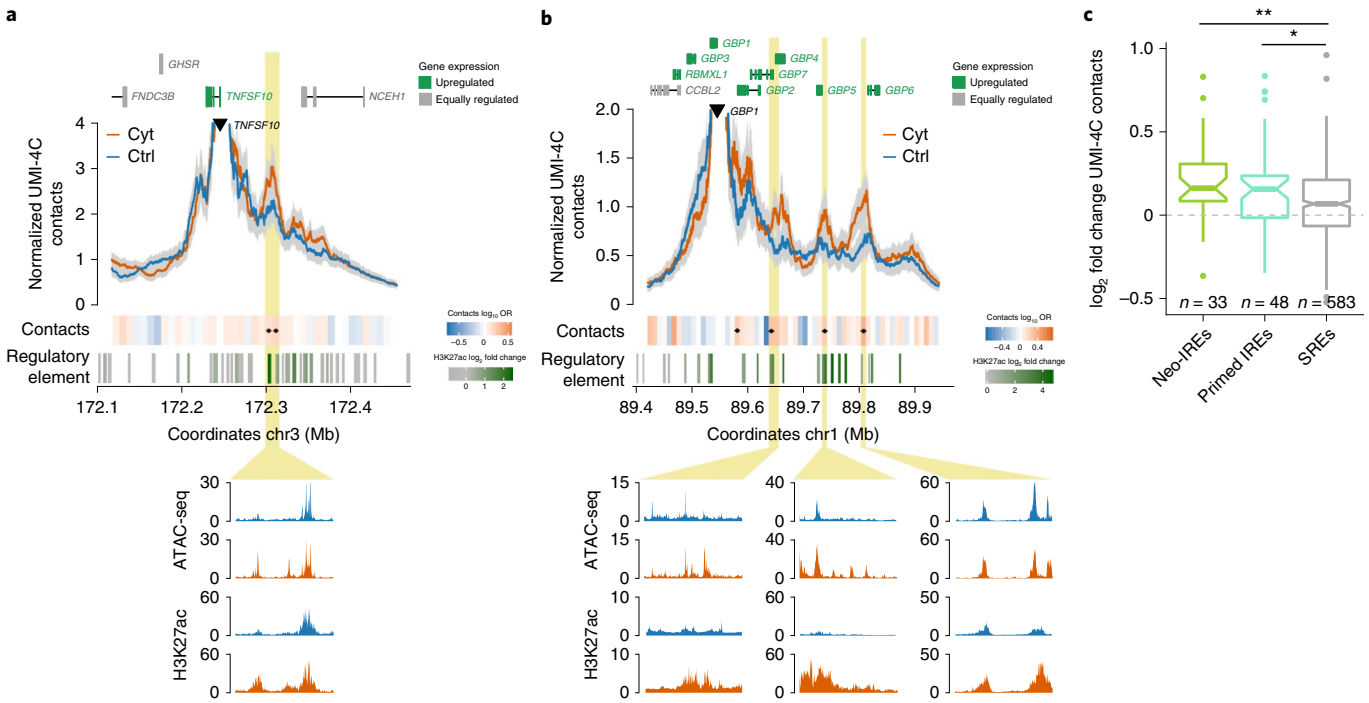


Fig. 3 | Cytokine exposure induces changes in human islet 3D chromatin structure. a,**b**, View of the UMI-4C chromatin contacts of the *TNFSF10* (**a**) and *GBP1* (**b**) promoters, before and after exposure to proinflammatory cytokines. The yellow boxes indicate IREs that gained contacts with the upregulated gene promoters. A heatmap under the 4C track represents the $\log_{10}$ odds ratio of the UMI-4C contact difference in cytokine-exposed (Cyt) versus untreated cells (Ctrl). The small black diamonds over the contact heatmap indicate a significant difference between cytokine-treated and untreated 3D chromatin contacts (chi-squared test $P < 0.05$). ATAC-seq peaks are represented by rectangles shaded from gray to green proportionally to the cytokine-induced H2K27ac $\log_2$ fold change observed at that site. **c**, Distribution of the UMI-4C contact $\log_2$ fold changes (Cyt versus Ctrl) at the different types of islet open chromatin sites classified in Fig. 2a. The data, obtained by analyzing viewpoints centered at the promoter of cytokine-induced genes, show that the chromatin structural changes are preferentially happening at IREs. The boxplot limits show the upper and lower quartiles; the whiskers extend to 1.5× the interquartile range; the individual data points represent outliers; and the notch represents the confidence interval around the median. Two-sided Wilcoxon test $^{*}P < 0.05$, $^{**}P < 0.01$.

regulatory element (Fig. 4f). This sequence displays enhancer activity in both treated and untreated β cells. However, exclusively in cytokine-exposed β cells, the T1D-associated G allele exerts significantly higher enhancer activity than the protective variant (one-way ANOVA, $F = 12.34$, $P = 1.23 \times 10^{-3}$; Fig. 4g and Extended Data Fig. 6i). The locus includes several upregulated genes (*SOCS1*, *DEXI*, *CIITA*, *RMI2*) that could represent potential targets of this IRE. Recent research points to *DEXI* as a T1D candidate gene in immune cells and β cells[31,32]. By performing UMI-4C experiments in human islets, we observed a strong chromatin contact between the promoter of *DEXI* and the regulatory element bearing the rs193778 T1D-associated variant (Fig. 4h). Such data points to *DEXI* as a potential causal gene in pancreatic islets.

Altogether, these results illustrate how unraveling cytokine-induced chromatin dynamics in human islets can guide the identification of *cis*-regulatory variants that are strong candidates in driving T1D-association signals.

Discussion

Our work illustrates the human pancreatic β-cell chromatin dynamics in response to an external stimulus that may be relevant in the context of T1D. We show that exposure to proinflammatory cytokines causes profound remodeling of the β-cell regulatory landscape coupled with changes in gene expression and protein production. The degree of remodeling of the regulatory network was comparable to that previously shown for macrophages or mouse dendritic cells exposed to similar stimuli[7]. We unveil the activation of approximately 3,600 cytokine-responsive distal *cis*-regulatory elements and reveal a lack of homogeneity in their molecular mechanism of activation. We observe that the induction of a subset of regulatory regions (neo-IREs) require transcription factor binding and chromatin opening, while other chromatin sites are primed to their activation being prebound by islet-specific transcription factors. Our observations suggest a model where binding of tissue-specific transcription factors may facilitate chromatin accessibility at a subset of chromatin sites that can then be promptly activated by the induction of inflammatory response transcription factors. Such a model is supported by very recent findings[33] and it is consistent with observations in murine macrophages[7,34] and dendritic cells[35]; however, thus far, it has not been demonstrated in a highly differentiated and nonimmune-related tissue, such as pancreatic islets. Even though our model suggests that exposure to proinflammatory cytokines causes predominantly induction of gene transcription rather than transcript downregulation, we cannot exclude that a more prolonged stimulus could induce loss of critical β-cell processes resulting from the reduction of β-cell *cis*-regulatory network activity.

Importantly, we show that such regulatory changes are coupled with 3D chromatin remodeling, allowing the newly activated regulatory elements to contact their target genes. Several reports have described the properties of 3D chromatin dynamics in the cell developmental context[36,37], on loss of cell fate[38,39], senescence[40,41] or in response to hormonal exposure[42]. Our observations indicate that the capacity of enhancer loop formation is maintained in a highly differentiated tissue such as the islets and it is coupled with transcriptional regulatory changes in response to an external stimulus.

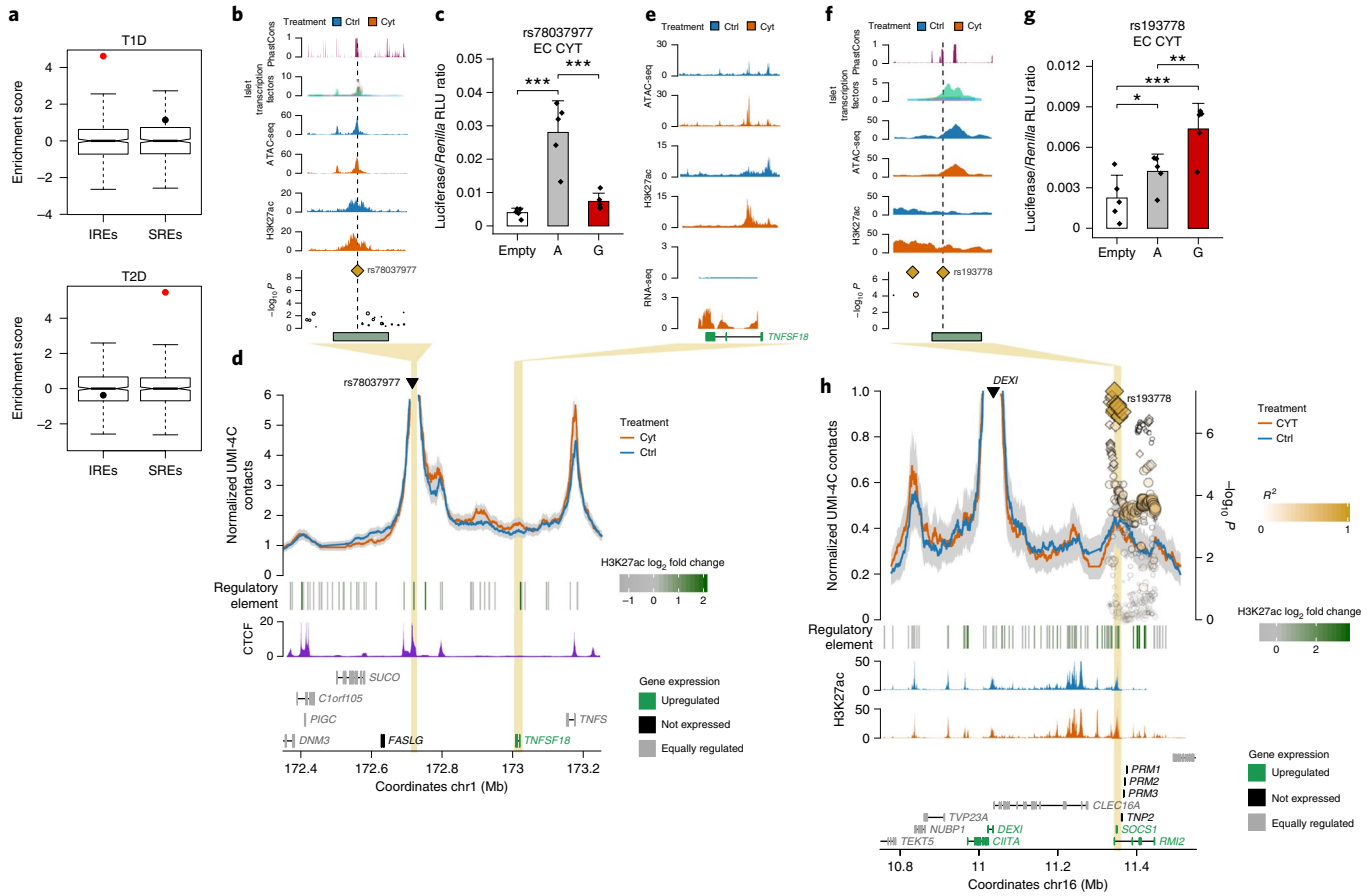


Fig. 4 | Cytokine-induced islet regulatory elements map to T1D-associated regions and guide the identification of functional risk variants. a, Islet IREs are enriched for T1D but not T2D risk variants while the opposite is true for islet SREs. Significant VSE scores are shown in red (Bonferroni-adjusted $P < 0.05$). The boxplot limits show the upper and lower quartiles; the whiskers extend to 1.5× the interquartile range; the notch represents the median confidence interval for distributions of matched null sets (500 permutations). **b**, rs78037977 overlaps an IRE bound by islet-specific transcription factors under basal conditions[9]. PhastCons, evolutionary conservation scores from phylogenetic analysis with space/time models. **c**, Luciferase assays in EC cells exposed to cytokines show that the sequence exerts enhancer activity that is reduced in the T1D-associated allele (G). RLU, relative light units. **d**, UMI-4C in human islets shows that the IRE containing rs78037977 engages multiple distal chromatin contacts. **e**, One induced chromatin contact mapping to the upregulated *TNFSF18* gene. **f**, The rs193778 variant maps to a phylogenetically conserved IRE. **g**, Luciferase assays in EC cells exposed to cytokines show significantly increased enhancer activity of the risk (G) allele compared to the nonrisk (A) allele. **h**, UMI-4C in islets, using the promoter of *DEXI* as the viewpoint, show a chromatin contact with the IRE bearing the T1D risk variant. **d**,**h**, ATAC-seq peaks are represented by rectangles shaded proportionally to the H2K27ac $\log_2$ fold change. **c**,**g**, Statistical significance was determined by one-way ANOVA followed by Bonferroni correction. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. The bars represent the mean ± s.d.

The model used in our study to explore chromatin dynamics is of particular interest because it mimics the inflammatory environment that pancreatic islets may face in the early stages of T1D. While several T1D candidate genes regulating key steps related to danger signal recognition and innate immunity are expressed in human islets[43], T1D-associated variants are enriched for immune cell types but not in stable pancreatic islets regulatory elements[4]. Such apparent contradiction may be reconciled by our findings showing that human islet cytokine-responsive regulatory elements are enriched for T1D risk variants. Our data, supported by recent findings revealing regulatory variants that affect enhancer activation in the immune response[33,44], opens the avenue to identify T1D molecular mechanisms acting at the pancreatic islet cell level.

Although we cannot exclude that functional variants disrupting the β-cell regulatory mechanisms may at the same time affect the regulatory potential of immune-related cell types, the availability of stimulus-responsive *cis*-regulatory maps in pancreatic islets will facilitate hypothesis-driven experiments to uncover how common and lower-frequency genetic variants impact islet cells in T1D. In this study, we researched the human islet responses to a specific proinflammatory stimulus. Future work studying additional immune-mediated stresses potentially affecting β cells at different stages of the disease may uncover other association signals acting at the islet cell level.

More generally, our findings could apply by extension to other diseases where primed enhancers may facilitate cell type-specific responses to ubiquitous signals resulting in tissue-specific genetic susceptibility in autoimmune diseases.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information, details of author contributions and competing interests, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0524-6.

Received: 30 November 2018; Accepted: 27 September 2019;
Published online: 1 November 2019

References

1. Todd, J. A. Etiology of type 1 diabetes. *Immunity* **32**, 457–467 (2010).
2. Ziegler, A.-G. & Nepom, G. T. Prediction and pathogenesis in type 1 diabetes. *Immunity* **32**, 468–478 (2010).
3. Eizirik, D. L., Colli, M. L. & Ortis, F. The role of inflammation in insulitis and β-cell loss in type 1 diabetes. *Nat. Rev. Endocrinol.* **5**, 219–226 (2009).
4. Onengut-Gumuscu, S. et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat. Genet.* **47**, 381–386 (2015).
5. Farh, K. K. et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
6. Ravassard, P. et al. A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *J. Clin. Invest.* **121**, 3589–3597 (2011).
7. Ostuni, R. et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157–171 (2013).
8. Gonzalez-Duque, S. et al. Conventional and neo-antigenic peptides presented by β cells are targeted by circulating naïve CD8+ T cells in type 1 diabetic and healthy donor.*Cell Metab.* **28**, 946–960.e6 (2018).
9. Pasquali, L. et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat. Genet.* **46**, 136–143 (2014).
10. Sung, M.-H., Guertin, M. J., Baek, S. & Hager, G. L. DNase footprint signatures are dictated by factor dynamics and DNA sequence. *Mol. Cell* **56**, 275–285 (2014).
11. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
12. Hemberger, M., Dean, W. & Reik, W. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat. Rev. Mol. Cell Biol.* **10**, 526–537 (2009).
13. Lister, R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
14. Feinberg, A. P. Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440 (2007).
15. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
16. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
17. Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* **14**, 390–403 (2013).
18. Symmons, O. et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
19. Krivega, I., Dale, R. K. & Dean, A. Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **28**, 1278–1290 (2014).
20. Yu, M. & Ren, B. The three-dimensional organization of mammalian genomes. *Annu. Rev. Cell Dev. Biol.* **33**, 265–289 (2017).
21. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800 (2018).
22. Schoenfelder, S. & Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
23. Stadhouders, R., Filion, G. J. & Graf, T. Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* **569**, 345–354 (2019).
24. Miguel-Escalada, I. et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat. Genet.* **51**, 1137–1148 (2019).
25. Schwartzman, O. et al. UMI-4C for quantitative and targeted chromosomal contact profiling. *Nat. Methods* **13**, 685–691 (2016).
26. MacArthur, J. et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).
27. Parker, S. C. J. et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl Acad. Sci. USA* **110**, 17921–17926 (2013).
28. Cooper, N. J. et al. Type 1 diabetes genome-wide association analysis with imputation identifies five new risk regions. Preprint at *bioRxiv* https://doi.org/10.1101/120022 (2017).
29. Fortune, M. D. et al. Statistical colocalization of genetic risk variants for related autoimmune diseases in the context of common controls. *Nat. Genet.* **47**, 839–846 (2015).
30. Ray, A., Basu, S., Williams, C. B., Salzman, N. H. & Dittel, B. N. A novel IL-10-independent regulatory role for B cells in suppressing autoimmunity by maintenance of regulatory T cells via GITR ligand. *J. Immunol.* **188**, 3188–3198 (2012).
31. Davison, L. J. et al. Long-range DNA looping and gene expression analyses identify *DEXI* as an autoimmune disease candidate gene. *Hum. Mol. Genet.* **21**, 322–333 (2012).
32. Dos Santos, R. S. et al. *DEXI*, a candidate gene for type 1 diabetes, modulates rat and human pancreatic beta cell inflammation via regulation of the type I IFN/STAT signalling pathway. *Diabetologia* **62**, 459–472 (2019).
33. Alasoo, K. et al. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat. Genet.* **50**, 424–431 (2018).
34. Heinz, S. et al. Effect of natural genetic variation on enhancer selection and function. *Nature* **503**, 487–492 (2013).
35. Vandenbon, A., Kumagai, Y., Lin, M., Suzuki, Y. & Nakai, K. Waves of chromatin modifications in mouse dendritic cells in response to LPS stimulation. *Genome Biol.* **19**, 138 (2018).
36. Phanstiel, D. H. et al. Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* **67**, 1037–1048.e6 (2017).
37. Mumbach, M. R. et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat. Genet.* **49**, 1602–1612 (2017).
38. Taberlay, P. C. et al. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731 (2016).
39. Barutcu, A. R. et al. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol.* **16**, 214 (2015).
40. Chandra, T. et al. Global reorganization of the nuclear landscape in senescent cells. *Cell Rep.* **10**, 471–483 (2015).
41. Criscione, S. W. et al. Reorganization of chromosome architecture in replicative cellular senescence. *Sci. Adv.* **2**, e1500882 (2016).
42. Le Dily, F. et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014).
43. Op de Beeck, A. & Eizirik, D. L. Viral infections in type 1 diabetes mellitus: why the β cells? *Nat. Rev. Endocrinol.* **12**, 263–273 (2016).
44. Kim-Hellmuth, S. et al. Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat. Commun.* **8**, 266 (2017).

Methods

Human islets and EC cells. Human islets from 14 multiorgan donors without a history of glucose intolerance were isolated in compliance with ethical regulations (Supplementary Note 1) and according to established isolation procedures[45,46] (Supplementary Note 2 and Supplementary Table 5). The human insulin-producing EC cells where kindly provided by R. Scharfmann[6] and cultured in DMEM medium (Supplementary Note 3).

Human islets and EC cells where either exposed or unexposed to a cocktail of proinflammatory cytokines IFN-γ and IL-1β for 48 h. The cytokine concentrations used were those described in previous dose–response experiments[47–49] (Supplementary Note 2). The glucose stimulation index was tested on human islet preparations and EC cell samples to confirm functional competence of the samples (Supplementary Note 4 and Extended Data Fig. 7).

ChIP-seq and ATAC-seq. ATAC-seq library preparations were carried out as described previously[50] with minor modifications[51,52] (Supplementary Note 5). ChIP-seq was carried out using tagmentation (ChIPmentation) as described previously[53] (Supplementary Note 6).

ATAC-seq and ChIP-seq libraries were sequenced on a HiSeq 2500 system (Illumina). Reads were aligned to the hg19 reference genome using Bowtie 2 v.2.3.4.1 (ref. [54]) using default parameters. After alignment, reads mapping to Encyclopedia of DNA Elements blacklist regions[55], noncanonical chromosomes or mitochondrial DNA were discarded. Duplicates were removed using samtools markdup v.1.8 (ref. [56]). See Supplementary Table 6 for the number of mapped reads per experiment and Extended Data Fig. 8 for measures of ATAC-seq quality.

Peaks were called with MACS2 callpeak v.2.1 (ref. [57]) with the parameters '-q 0.05 -nomodel -shift -100 -extsize 200' for ATAC-seq and '-broad -broad-cutoff 0.1 --nomodel' for H3K27ac ChIP-seq. A more detailed description of bioinformatics processing can be found in Supplementary Note 7.

RNA-seq. Total RNA was isolated from EC cells and human islets[8] using the RNeasy Mini Kit (QIAGEN), which retrieves RNA molecules longer than 200 nucleotides, as described in detail previously[58]. RNA integrity number values were evaluated using the Bioanalyzer 2100 (Agilent Technologies). All the samples had RNA integrity number values >8 (Supplementary Note 8).

RNA-seq libraries were sequenced on a HiSeq 2000 platform (Illumina) to produce 100 base pair (bp)-long paired-end reads with an average of 180 million reads per replicate (EC cells, $n=5$). Reads were aligned using TopHat v.2.0.13 (ref. [59]) to the GChr37 genome with default parameters. Then reads were assigned to GENCODE release 18 gene annotation[60] using htseq-count v.0.6.1p1 (ref. [61]) with default parameters. The RNA-seq of five human islet preparations[8] was used for comparison and processed in an identical way.

Differential analysis of ATAC-seq, ChIP-seq and RNA-seq. For both ATAC-seq and ChIP-seq, aligned reads from all replicates were merged into a single BAM file to identify a comprehensive set of peaks. Next, we used the comprehensive peak set to compute read counts separately for each replicate and condition. In the case of the RNA-seq data, the output of htseq-count was used as the input matrix for downstream analysis. The generated matrices were normalized and differential analysis was performed using DESeq2 v.1.24.0 (ref. [62]) using a paired sample design (Supplementary Note 9). Thresholds for significance were set at an FDR-adjusted $P<0.05$ and an absolute $\log_2$ fold change > 1. All regions/genes that did not reach significance or did not pass the $\log_2$ fold change cutoff were classified as stable/equally regulated.

Proteomics. For the proteomics analysis, 1.5 million EC cells treated or not treated with cytokines (IL-1β + IFN-γ) were processed using the metabolite, protein and lipid extraction approach[63] (Supplementary Note 10).

Collected data were processed using Decon2LS_V2 v.2.3.1.4 (ref. [64]) and DtaRefinery v.1.2 (ref. [65]), both using default parameters, to recalibrate the runs and generate peak lists. Peptide identification was done using MS-GF+ v.2017.08.23 (ref. [66]) by searching peak lists against islet protein sequences deduced from a transcriptomics experiment[47] and supplemented with keratin sequences (32,780 total protein sequences) (Supplementary Note 10).

Extracted reporter ion intensities (Supplementary Note 10) were then converted into $\log_2$ and normalized by standard median centering. Proteins were quantified using a Bayesian proteoform discovery methodology (Bayesian proteoform quantification) in combination with standard reference-based median quantification[67] and were considered significant at a cutoff of $P\leq0.05$ based on a paired *t*-test.

Protein–protein interaction network analysis was performed with GeNets[68] using Metanetworks v.1.0, which integrates protein–protein interactions from InWeb v.3 (ref. [69]) and ConsensusPathDB v.32 (ref. [70]). Default parameters were applied and Molecular Signatures Database-enriched v.6.1 (ref. [71]) pathways were overlaid.

Defining classes of IREs. To characterize the dynamics of chromatin accessibility on exposure of human islets and EC cells to proinflammatory cytokines, we processed the results obtained from the DESeq2 differential analysis and computed the overlap between ATAC-seq peaks and H3K27ac-enriched sites, allowing a 200 bp gap. Regions annotated as stable for both ATAC-seq and H3K27ac assays were classified as SREs. Regions classified as either stable or gained in ATAC-seq differential analysis and as gained in H3K27ac were classified as IREs.

IREs were classified in two groups: opening IREs ($n=2,436$), corresponding to regions annotated as gained for both ATAC-seq and H3K27ac; and primed IREs ($n=1,362$) for regions annotated as stable for ATAC-seq and gained for H3K27ac. Since opening IREs include a gradient of cytokine-induced chromatin accessibility changes, we next selected only those opening regions that were completely closed before cytokine exposure. For this purpose, we considered newly open chromatin regions as those opening ATAC-seq peaks that were not called in the control samples using a relaxed threshold ($P\leq0.05$). Such analysis allowed us to identify a subset of 1,716 opening regions that we named neo-IREs. A similar approach was used to identify macrophage latent enhancers[7].

See Supplementary Note 11 for the sequence conservation analysis performed at the different classes of IREs.

Assigning regulatory elements to target genes. To annotate regulatory elements as distal or proximal, we assigned each regulatory element to the nearest TSS of a coding gene (using GENCODE release 18 annotation[60]). Those regions lying within 2 kilobases (kb) from the nearest TSS were annotated as promoters while the rest were considered as distal regulatory elements.

To test the association between different classes of open chromatin and changes in gene expression and protein abundance (Fig. 1d,e and Extended Data Figs. 2e and 3a) in an unbiased manner, we assigned ATAC-seq sites to genes closer than 15 kb from their TSS. To analyze the additive effect of IREs on gene expression changes, we associated to a gene all IREs within 40 kb of their TSS (Extended Data Fig. 2f).

Finally, to detect all possible IRE gene targets, we assigned to each IRE all upregulated genes whose TSS was closer than 40 kb. When an upregulated gene could not be found in <40 kb, the IRE was assigned to the closest, but <1 Mb far, induced gene (Extended Data Fig. 4a and Supplementary Tables 1 and 2; see Supplementary Note 12).

Sequence composition and transcription factor analysis. De novo motif analysis was performed using HOMER v.4.8.2 (ref. [72]) findMotifGenome.pl tool with the parameter '-size given -bits –mask'. Only enriched sequences present in >1.5% of targets were retained. Selection of best matches was performed as follows: all matches with scores >0.80 were included in the table; for those hits without any match >0.80, the top 3 hits were selected and their score was included in the table (Extended Data Figs. 1g and 3e,f).

To assay motif colocalization, we used all motif instances identified in the de novo analysis in primed enhancers. First we used the findMotifGenome.pl tool from HOMER to map all these motif instances in primed enhancers and SRE enhancers (that is, excluding all sites <2 kb from a TSS). Next, the motif colocalization was calculated by counting motif pairs found in each ATAC-seq peak. Significance was determined by Fisher's exact test comparing the colocalization of motif pairs in distal IREs versus distal SREs. Only significant pairs (Fisher's exact test, FDR-adjusted $P<0.001$) were retained (Extended Data Fig. 3g,h).

To evaluate islet-specific transcription factor occupancy, we used ChIP-seq BAM files for PDX1, NKX2.2, FOXA2, NKX6.1 and MAFB[9]. We computed the read coverage in the regions of interest over 10 bp bins. Reads were quantile-normalized, the mean counts in each bin for each transcription factor were calculated and the mean for all transcription factors was plotted (Fig. 2d).

To identify footprints from the ATAC-seq data, we generated tag directories with all ATAC-seq replicates in each condition using HOMER makeTagDirectory. Neo and primed enhancers were centered on the ISRE motif matrix annotated with annotatePeaks.pl with the option '-center motif1.motif –size given' and tag means for the 5′ and 3′ read ends were obtained using annotatePeaks.pl with the option '-size -100,100 –hist 1 –d tagsDir'. The resulting 5′ ends were plotted using ggplot2 v.3.2.0 (ref. [73]) (Fig. 2e).

To create a nonredundant dataset of motifs for the gene regulatory network analysis (Extended Data Fig. 4a), motifs from primed and opening enhancers were reduced to a nonredundant set with the compareMotifs.pl script from HOMER using a similarity score of 0.7 as the threshold for merging similar motifs. The motifs were then mapped to primed and opening enhancers using annotatePeaks.pl.

Infinium MethylationEPIC array. DNA from EC cells exposed or unexposed to IL-1β and IFN-γ for 48 h, as described earlier (5 replicates per condition), was extracted using QIAamp DNA Mini Kit (QIAGEN); 1 µg DNA aliquots ($n=10$) were processed for 850 K Infinium MethylationEPIC array (Illumina) as described previously[74].

The resulting array signals were processed and analyzed with the RnBeads R package v.3.2.0 (ref. [75]). The method used by RnBeads to assess differences between groups consists of fitting a hierarchal linear model (the empirical Bayesian method from the limma package v.3.40.0 (ref. [76])) using *M* values (log of β values) as metrics to measure methylation levels[77]. All *P* values were corrected for multiple testing using the Benjamini–Hochberg method for controlling the FDR. CpGs

were considered as differentially methylated when FDR-adjusted $P < 0.05$ and the absolute difference in methylation β values between cytokine and control samples was >0.2 (20% change in methylation). Information on the differentially methylated CpGs can be found in Supplementary Table 7.

UMI-4C. UMI-4C was performed as described previously[25] with minor modifications (see Supplementary Note 13). To increase molecular complexity, each library was obtained by pooling 5–10 PCRs per viewpoint. The PCR primers used in UMI-4C are listed in Supplementary Table 8. Each library was sequenced to a depth >1 million 75-bp-long paired-end reads using either the NextSeq 550 or HiSeq 2500 platforms.

Paired-end reads were demultiplexed according to the viewpoint sequence using fastq-multx from ea-utils v.1.30 (ref. [78]) and analyzed with the umi4cPackage v.0.0.0.9000 (ref. [25]). 4C tracks were created by selecting viewpoint-specific reads, aligning them to the genome and extracting the number of UMIs using the p4cCreate4CseqTrack function (see quality control statistics in Supplementary Table 9). Cytokine-treated profiles were then scaled to the control profile using the umi4cPackage function p4cSmoothedTrendComp. Profiles were also smoothened based on the total number of UMIs present in a 2 Mb region centered on the viewpoint and excluding the 3 kb around it. The following formula was used to calculate the minimum UMIs needed for smoothing. If the fragment did not reach this minimum, it was merged with successive fragments until a minimum was reached:

$$\text{Minimum UMIs} = \frac{\sum \text{UMIs}_{\text{region}}}{2{,}000} \times 50$$

To detect differential chromatin contacts we focused on a 2 Mb region centered on the viewpoint, but excluding 1.5 kb on each side of the viewpoint. We then partitioned the region into windows of width proportional to the mean restriction fragment length in the region ($\text{Mean}_{\text{fragment}}$):

$$\text{Width}_{\text{window}} = \text{Mean}_{\text{fragment}} \times 20$$

Differential contact analysis was performed for each of these windows using a chi-squared test, comparing UMIs in such windows with the total number of UMIs in the 2 Mb region. Windows with a chi-squared $P < 0.05$ are highlighted in Fig. 3a,b and Extended Data Fig. 5b–d with small black diamonds. To quantify the chromatin contact changes, we counted the number of cytokine-treated and untreated UMIs for each window and computed their odds ratio (OR) based on the total UMI counts in the region, following the formula:

$$\text{OR}_{\text{window}} = \frac{\text{Ctrl}_{\text{region}} \times \text{Cyt}_{\text{window}}}{\text{Ctrl}_{\text{window}} \times \text{Cyt}_{\text{region}}}$$

where Ctrl and Cyt represent the number of UMIs in unexposed and cytokine-exposed conditions.

Variant set enrichment (VSE) analyses. We used the VSE R package v.0.99 (ref. [79]) to assess the enrichment of T1D and T2D risk variant for IRE and SRE regulatory annotations. We first selected from the NHGRI-EBI GWAS catalog[26] all leading SNPs with disease trait matching either 'type 1 diabetes' or 'type 2 diabetes' (24 April 2019). Next, we extended our collection of associated variants to all those in strong linkage disequilibrium ($R^2 \geq 0.8$, EUR) with the lead SNP (source of linkage disequilibrium information, 1000 Genomes Project phase 3 (ref. [80])). These SNPs and their proxies were used to generate the associated variant set (AVS)[79], resulting in 83 disjointed regions for T1D and 389 for T2D, after removing shared loci between T1D and T2D. A null distribution or matched random variant set, matched in size and structure to the original AVS, was generated from the 1000 Genomes Project phase 3 by permutating the AVS 500 times. The number of independent SNPs from the AVS overlapping the regulatory annotations was computed and compared with the intersections obtained with the matched random variant set. The enrichment score was defined as the number of s.d. that the overlapping tally deviates from the null overlapping tally median. The exact P value was then calculated by fitting a density function to the null distribution derived from the matched random variant set. This P value was finally corrected for multiple testing using the Bonferroni method. Enrichments or depletions with a Bonferroni-adjusted $P < 0.05$ were considered statistically significant (Fig. 4a and Extended Data Fig. 6a).

T1D-associated regions were generated by selecting all SNPs in strong linkage disequilibrium ($R^2 \geq 0.8$, EUR) with the T1D leading SNPs. We defined the risk loci boundaries using the most upstream and downstream SNPs. Next, we merged the overlapping loci to obtain a total of 71 T1D risk regions. All T1D-associated regions containing IREs and T1D risk variants directly overlapping human islet cytokine-induced regulatory elements are shown in Supplementary Table 3. For this analysis, to extract all possible cytokine-induced regulatory elements located at the T1D risk loci, we used a less stringent set of human islet IREs by lowering the H3K27ac $\log_2$ fold change threshold from 1 to 0.8.

For details regarding the GWAS association analysis, see Supplementary Note 14.

Luciferase reporter assays. For episomal reporter assays in the EC cell line, selected human cytokine-induced regulatory elements regions were first amplified from genomic DNA using primers (Supplementary Table 10) containing the XhoI/HindIII restriction sites. The amplicons were then cloned into the pGL4.23[luc2/minP] luciferase reporter vector (Promega Corporation) as described previously[81]. Briefly, the amplicon and the vector were simultaneously digested. Next, the vector was dephosphorylated with FastAP (Thermo Fisher Scientific). The DNA was then purified and ligated with a T4 DNA Ligase (Promega Corporation). Next, the generated reporter vectors were transformed into *Escherichia coli* (DH5α) and purified with the NucleoSpin Plasmid (catalog no. 740588.250; Macherey-Nagel).

Site-directed mutagenesis was used to introduce single-nucleotide variants into the generated construct. The variants were generated by PCR using the primers shown in Supplementary Table 10. The parental supercoiled double-stranded DNA was digested with DpnI (catalog no. R0176S; New England Biolabs) 1 h at 37 °C and the constructs were transformed in competent *E. coli* cells (DH5α) by thermal shock. Finally, the introduced variants were checked using Sanger sequencing.

EC cells were transfected in 24-well plates at a density of 300,000 cells per well, with 200 ng of reporter vectors or empty vectors plus 20 ng of phRL-CMV *Renilla* luciferase to control for transfection efficiency.

Transfections were performed with Lipofectamine 2000 (Thermo Fisher Scientific) for 8 h, according to the manufacturer's instructions. On transfection, the EC medium was supplemented with 2% FCS[82] and exposed or unexposed to the cytokines for 48 h. After 48 h, cells were assayed using the Dual Luciferase Assay (Promega Corporation), according to the manufacturer's instructions. The luciferase units were measured with the VICTOR Multilabel Plate Reader (PerkinElmer). Firefly luciferase activity was normalized to *Renilla* luciferase activity and then divided by the values obtained for the empty pGL4.23. The assays were performed in at least three independent experiments.

Statistical differences were calculated using a one-way ANOVA. P values were then Bonferroni-corrected.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Datasets for the IREs are available for download and visualization at the Islet Regulome Browser[83] (www.isletregulome.com).

Raw sequencing reads for the different high-throughput assays can be accessed at the Gene Expression Omnibus with the following identifiers: GSE123404 (ATAC-seq); GSE133135 (H3K27ac data); GSE137136 (RNA-seq); and GSE136865 (UMI-4C). Raw proteomics data can be accessed at the ProteomeXchange with the identifier PXD011902.

Code availability

The code and scripts used in this study are available from the corresponding author upon reasonable request.

References

45. Bucher, P. et al. Assessment of a novel two-component enzyme preparation for human islet isolation and transplantation. *Transplantation* **79**, 91–97 (2005).
46. Melzi, R. et al. Role of CCL2/MCP-1 in islet transplantation. *Cell Transplant.* **19**, 1031–1046 (2010).
47. Eizirik, D. L. et al. The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. *PLoS Genet.* **8**, e1002552 (2012).
48. Colli, M. L., Moore, F., Gurzov, E. N., Ortis, F. & Eizirik, D. L. MDA5 and PTPN2, two candidate genes for type 1 diabetes, modify pancreatic β-cell responses to the viral by-product double-stranded RNA. *Hum. Mol. Genet.* **19**, 135–146 (2010).
49. Ortis, F. et al. Cytokine-induced proapoptotic gene expression in insulin-producing cells is related to rapid, sustained, and nonoscillatory nuclear factor-κB activation. *Mol. Endocrinol.* **20**, 1867–1879 (2006).
50. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
51. Lavin, Y. et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* **159**, 1312–1326 (2014).
52. Raurell-Vila H., Ramos-Rodríguez M., & Pasquali L. Assay for transposase accessible chromatin (ATAC-Seq) to chart the open chromatin landscape of human pancreatic islets. *Methods Mol. Biol.* **1766**, 197–208 (2018).
53. Schmidl, C., Rendeiro, A. F., Sheffield, N. C. & Bock, C. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat. Methods* **12**, 963–965 (2015).
54. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

55. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
56. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
58. Juan-Mateu, J. et al. SRp55 regulates a splicing network that controls human pancreatic β-cell function and survival. *Diabetes* **67**, 423–436 (2018).
59. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
60. Harrow, J. et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).
61. Anders, S., Pyl, P. T. & Huber, W. HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
62. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
63. Nakayasu, E. S. et al. MPLEx: a robust and universal protocol for single-sample integrative proteomic, metabolomic, and lipidomic analyses. *mSystems* **1**, e00043-16 (2016).
64. Mayampurath, A. M. et al. DeconMSn: a software tool for accurate parent ion monoisotopic mass determination for tandem mass spectra. *Bioinformatics* **24**, 1021–1023 (2008).
65. Petyuk, V. A. et al. DtaRefinery, a software tool for elimination of systematic errors from parent ion mass measurements in tandem mass spectra data sets. *Mol. Cell. Proteomics* **9**, 486–496 (2010).
66. Kim, S. & Pevzner, P. A. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat. Commun.* **5**, 5277 (2014).
67. Webb-Robertson, B.-J. M. et al. Bayesian proteoform modeling improves protein quantification of global proteomic measurements. *Mol. Cell. Proteomics* **13**, 3639–3646 (2014).
68. Li, T. et al. GeNets: a unified web platform for network-based genomic analyses. *Nat. Methods* **15**, 543–546 (2018).
69. Li, T. et al. A scored human protein–protein interaction network to catalyze genomic interpretation. *Nat. Methods* **14**, 61–64 (2017).
70. Herwig, R., Hardt, C., Lienhard, M. & Kamburov, A. Analyzing and interpreting genome data at the network level with ConsensusPathDB. *Nat. Protoc.* **11**, 1889–1907 (2016).
71. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
72. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
73. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).
74. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–399 (2016).
75. Assenov, Y. et al. Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* **11**, 1138–1140 (2014).
76. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
77. Du, P. et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* **11**, 587 (2010).
78. Aronesty, E. Comparison of sequencing utility programs. *Open Bioinforma. J.* **7**, 1–8 (2013).
79. Ahmed, M. et al. Variant Set Enrichment: an R package to identify disease-associated functional genomic regions. *BioData Min.* **10**, 9 (2017).
80. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
81. Jordà, M. et al. Upregulation of MMP-9 in MDCK epithelial cell line in response to expression of the Snail transcription factor. *J. Cell Sci.* **118**, 3371–3385 (2005).
82. Brozzi, F. et al. Cytokines induce endoplasmic reticulum stress in human, rat and mouse beta cells via different mechanisms. *Diabetologia* **58**, 2307–2316 (2015).
83. Mularoni, L., Ramos-Rodríguez, M. & Pasquali, L. The pancreatic Islet Regulome Browser. *Front. Genet.* **8**, 13 (2017).

Acknowledgements

L.Pasquali was supported by grants from the Spanish Ministry of Economy and Competitiveness (nos. BFU2014-58150-R and SAF2017-86242-R), Marató TV3 (no. 201624.10) and a young investigator award from the Spanish Society of Diabetes (Ayuda SED a Proyectos de Investigación, no. 2017-SED). L.Pasquali is a recipient of a Ramon y Cajal contract from the Spanish Ministry of Economy and Competitiveness (no. RYC-2013-12864). The Pasquali lab is further supported by Instituto de Salud Carlos III (no. PIE16/00011). M.R. is supported by an FI Agència de Gestió d'Ajuts Universitaris i de Recerca PhD fellowship (no. 2019FI_B100203). J.J. was supported by a Marie Skłodowska-Curie Actions fellowship grant from the Horizons 2020 European Union (EU) Programme (project no. 660449). M.I.A. was supported by a FRIA fellowship from the Fonds de la Recherche Scientifique (FNRS) (no. 26410496). Human islets were provided through the European Consortium for Islet Transplantation distribution program for basic research supported by JDRF (no. 31-2008-416). D.L.E. was supported by the Walloon Region through the FRFS-WELBIO Fund for Strategic Fundamental research (no. CR-2015A-06s and CR-2019C-04) and by grants from the Fonds National de la Recherche Scientifique (no. T003613F), the Horizon 2020 Programme (project T2Dsystems, no. GA667191), Brussels Capital Region Innoviris (project DiaType, no. 2017-PFS-24), Dutch Diabetes Research Fundation (project Innovate2CureType1, DDRF; no. 2018.10.002) and the Innovative Medicines Initiative 2 Joint Undertaking (project INNODIA, no. 115797). This Innovative Medicines Initiative 2 Joint Undertaking receives support from the EU's Horizon 2020 Research and Innovation Programme and European Federation of Pharmaceutical Industries and Associations, JDRF and the Leona M. and Harry B. Helmsley Charitable Trust (project INNODIA, no. 115797). T.O.M. and D.L.E. were supported by a grant from the National Institutes of Health-National Institute of Diabetes and Digestive and Kidney Diseases-Human Islet Research Network Consortium (no. 1UC4DK104166-01). Part of the work was performed at the Environmental Molecular Sciences Laboratory, a US Department of Energy national scientific user facility located at Pacific Northwest National Laboratory. Battelle operates the Pacific Northwest National Laboratory for the Department of Energy (contract no. DE-AC05-76RLO01830). We thank J. Mercader (Broad Institute) and M. Guindo Martínez (Barcelona Supercomputing Center) for helpful discussions regarding GWAS enrichment analyses and A. Castela (Université Libre de Bruxelles Center for Diabetes Research) for helping with the glucose-stimulated insulin secretion experiments.

Author contributions

L.Pasquali, M.R. and D.L.E. designed the experiments. H.R., M.L.C., M.A., J.J., R.N. and E.N. performed and analyzed the experiments. M.R. performed the bioinformatic analyses with contribution from M.S., J.-V.T., B.W. and J.I. L.Piemonti, P.M., M.E. and T.O.M. provided materials and resources. L.Pasquali, D.L.E., T.O.M. and J.A.T. supervised the study. L.Pasquali and M.R. coordinated and conceived the project, and wrote the manuscript with contribution from D.L.E. All authors reviewed the final manuscript.

Competing interests

The authors declare no competing interests.

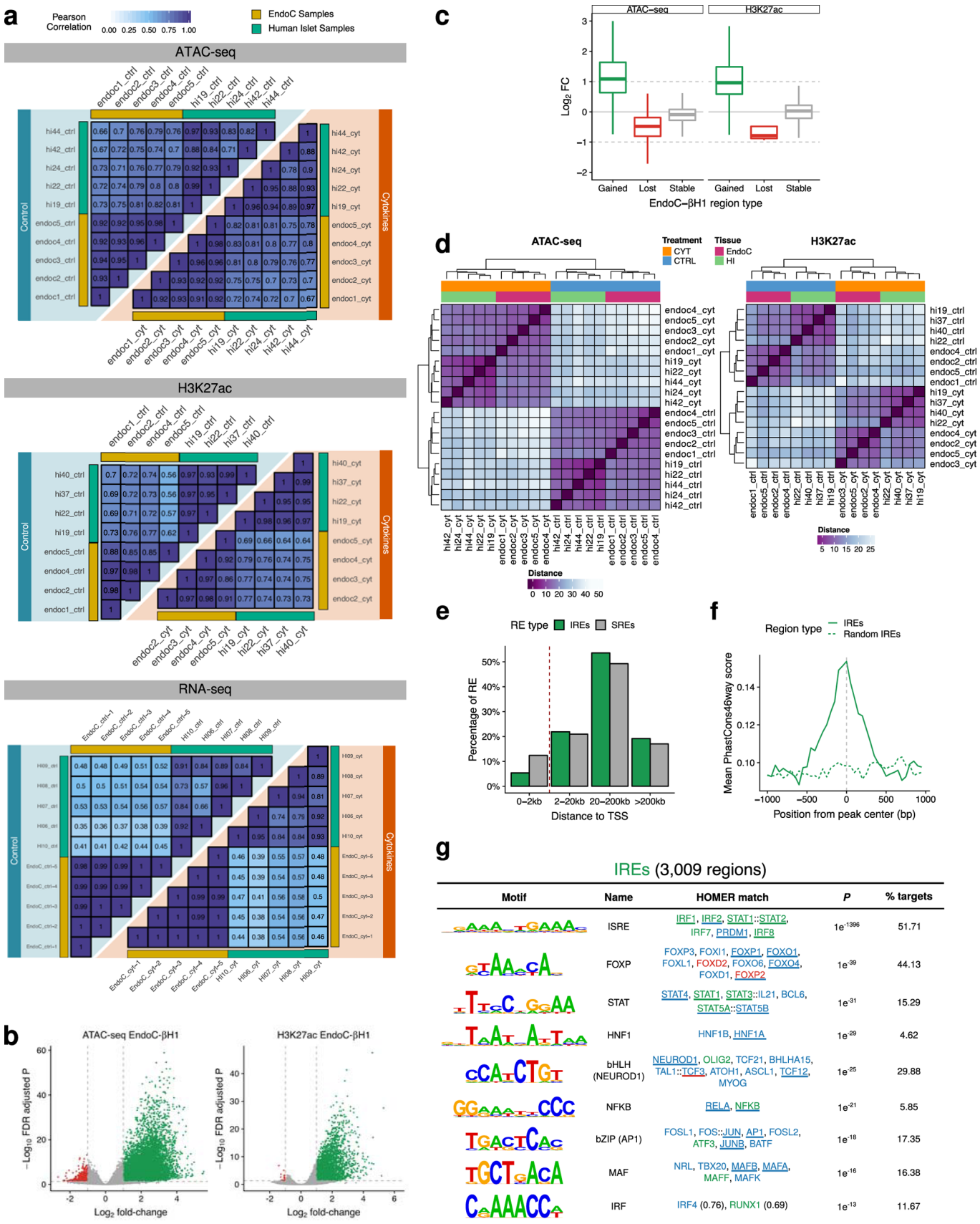
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-019-0524-6.

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0524-6.

Correspondence and requests for materials should be addressed to L.P.

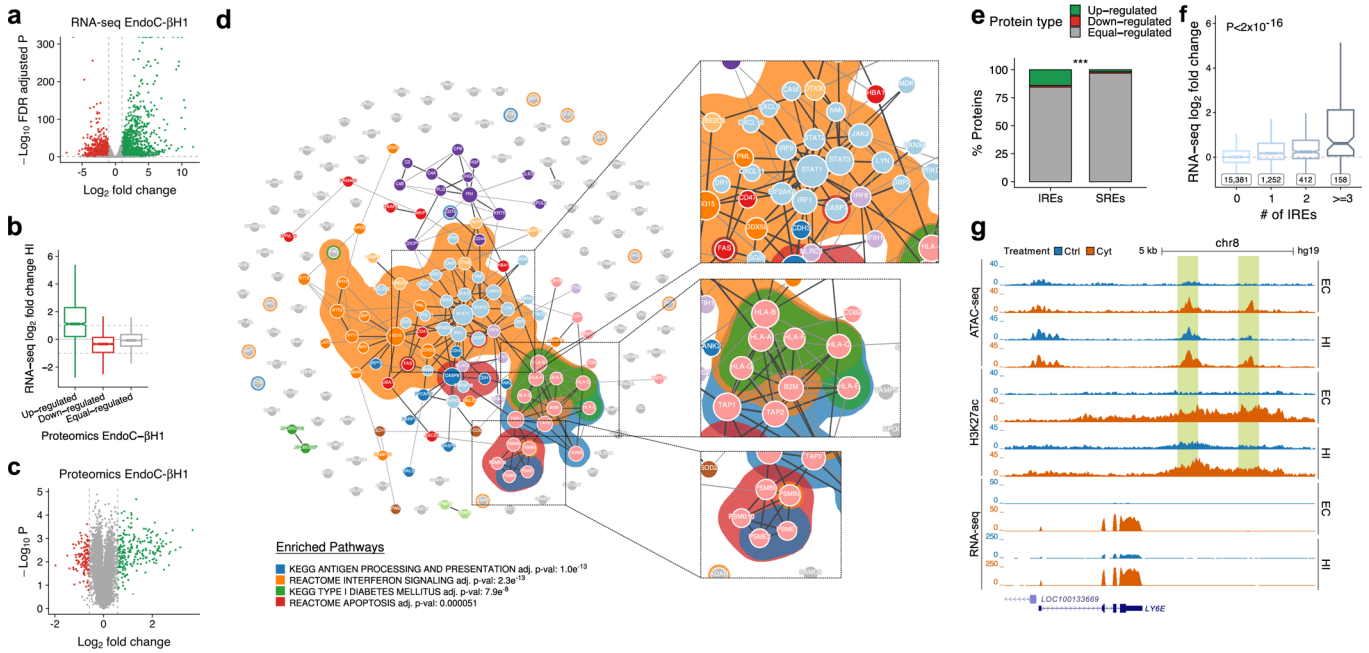
Reprints and permissions information is available at www.nature.com/reprints.



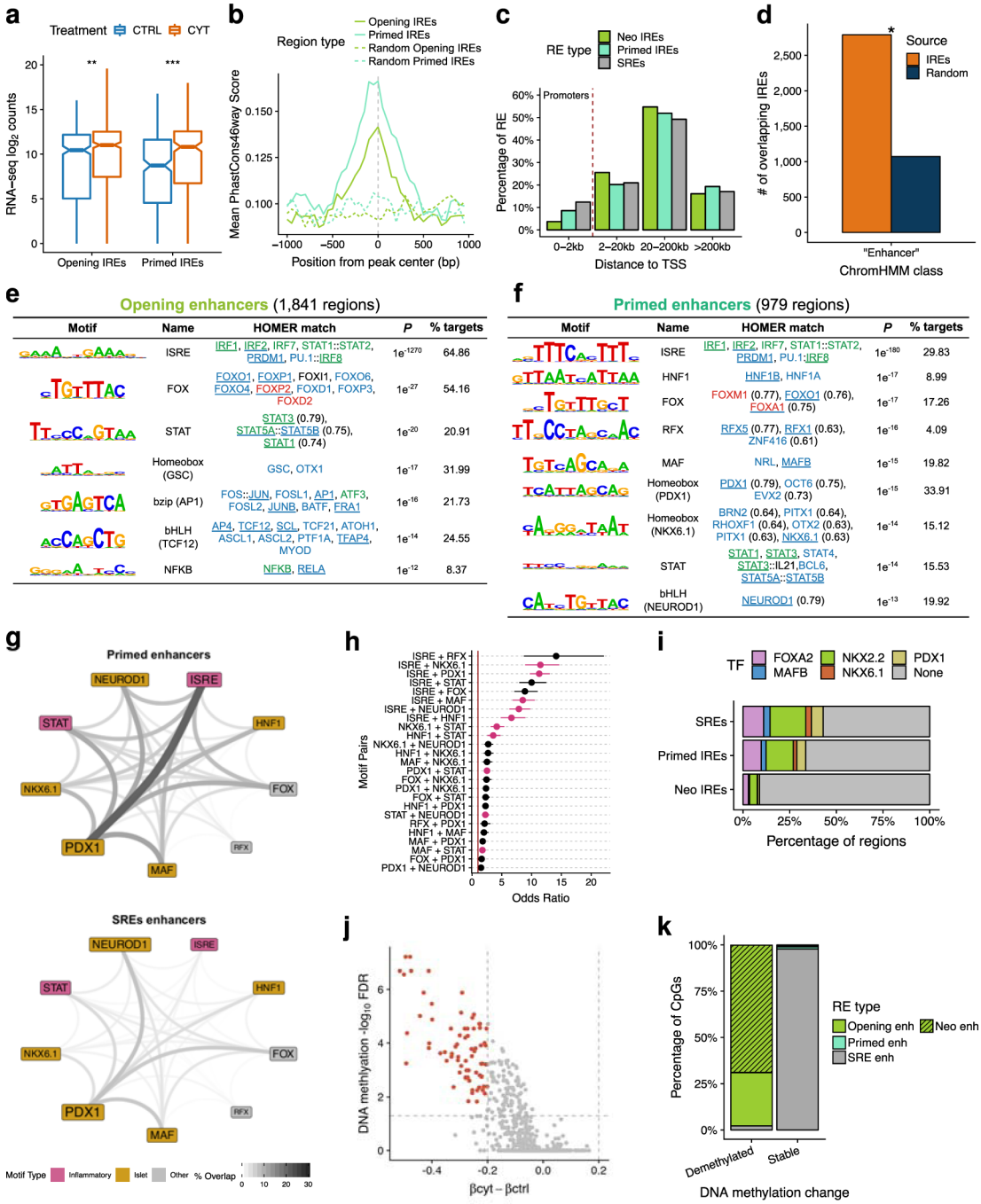
Motif	Name	HOMER match	P	% targets
	ISRE	IRF1, IRF2, STAT1::STAT2, IRF7, PRDM1, IRF8	$1e^{-1396}$	51.71
	FOXP	FOXP3, FOXI1, FOXP1, FOXO1, FOXL1, FOXD2, FOXO6, FOXO4, FOXD1, FOXP2	$1e^{-39}$	44.13
	STAT	STAT4, STAT1, STAT3::IL21, BCL6, STAT5A::STAT5B	$1e^{-31}$	15.29
	HNF1	HNF1B, HNF1A	$1e^{-29}$	4.62
	bHLH (NEUROD1)	NEUROD1, OLIG2, TCF21, BHLHA15, TAL1::TCF3, ATOH1, ASCL1, TCF12, MYOG	$1e^{-25}$	29.88
	NFKB	RELA, NFKB	$1e^{-21}$	5.85
	bZIP (AP1)	FOSL1, FOS::JUN, AP1, FOSL2, ATF3, JUNB, BATF	$1e^{-18}$	17.35
	MAF	NRL, TBX20, MAFB, MAFA, MAFF, MAFK	$1e^{-16}$	16.38
	IRF	IRF4 (0.76), RUNX1 (0.69)	$1e^{-13}$	11.67

Extended Data Fig. 1 | See next page for caption.

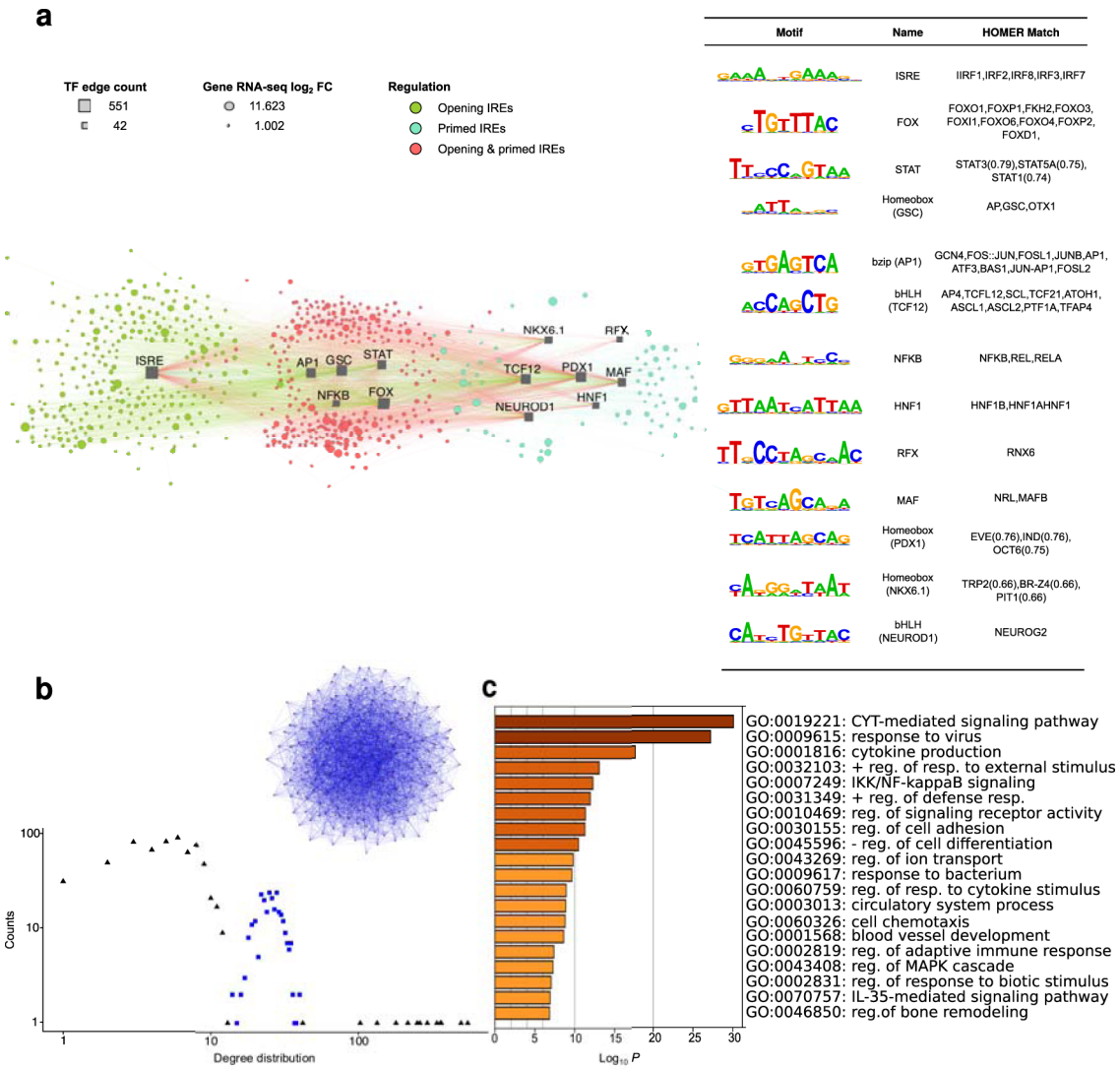
Extended Data Fig. 1 | Chromatin characterization of human pancreatic β cells exposed to pro-inflammatory cytokines. a, Pearson correlation values between replicates in different assays and conditions (see Supplementary Note 2). **b**, Volcano plots of ATAC-seq (left) and H3K27ac ChIP-seq (right) changes obtained after exposure of EndoC-βH1 to IFN-γ and IL-1β; green and red dots correspond to sites with absolute $\log_2$ fold change > 1 and FDR adjusted $P < 0.05$ as calculated by fitting a negative binomial model in DESeq2. Chromatin changes are classified as 'gained' and 'lost' chromatin sites whereas non-significant changes are defined as 'stable'. **c**, Chromatin accessibility and H3K27ac enrichment changes observed in EndoC-βH1 are largely replicated in human pancreatic islets as illustrated by the distribution of $\log_2$ fold change at regions as classified in **b** in EndoC-βH1. Dotted lines indicate $\log_2$ fold change thresholds (absolute $\log_2$ fold change > 1). Box plot limits show upper and lower quartiles, whiskers extend to 1.5 times the interquartile range and the notch represents the confidence interval around the median. **d**, Hierarchical clustering using normalized ATAC-seq and H3K27ac read counts at EndoC-βH1 IREs shows that samples cluster according to treatment, suggesting that the differences caused by the proinflammatory cytokines are greater than those derived by the sample heterogeneity. HI = Human pancreatic islets, EndoC = EndoC-βH1 **e**, Distribution of distances to nearest TSS for the different types of regulatory elements, showing that IREs, compared with stable regulatory elements (SREs), are preferentially located distally to TSS. **f**, Mean sequence conservation score of IREs and a randomized set of IREs in placental mammals. Peaks were extended from the center 1 kb to each direction and mean score was calculated in 50 bp windows. **g**, Sequence composition analysis of IREs (n = 3,009) illustrating the top identified *de novo* motifs. Colors for matched genes correspond to RNA-seq (name) or protein (underlined) status (red = down-regulated, blue = equal-regulated, green = up-regulated, black/no line = not expressed/detected).



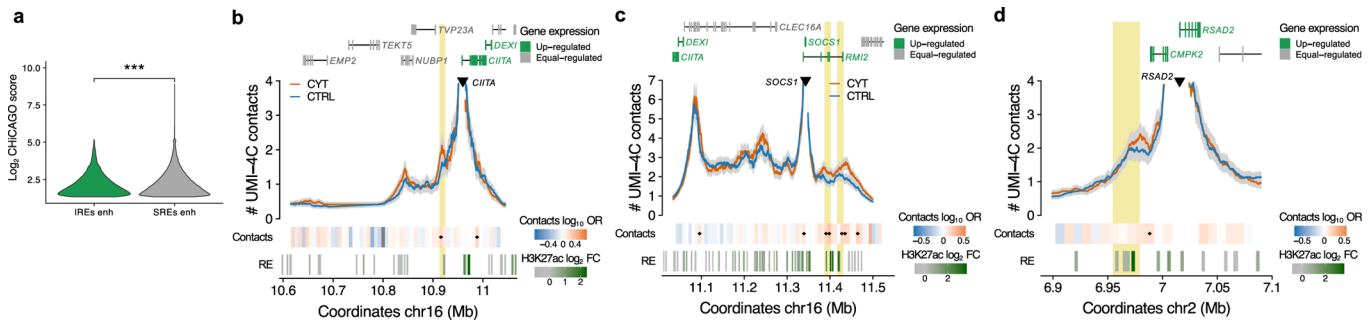
Extended Data Fig. 2 | Exposure to pro-inflammatory cytokines drives changes in the transcriptome and proteome of pancreatic β cells. a, Volcano plot of RNA-seq genes, showing up-regulated genes (green) and down-regulated genes (red) upon exposure of EndoC-βH1 to cytokines. Vertical lines indicate the $\log_2$ fold change threshold (absolute $\log_2$ fold change > 1) and horizontal line indicates the FDR adjusted *P* cutoff for significance (FDR adjusted $P < 0.05$) calculated by fitting a negative binomial model in DESeq2. **b**, Distribution of RNA-seq counts in human islet samples in the genes previously classified as up, down or equal-regulated in EndoC-βH1 cells. Boxplot limits show upper and lower quartiles, whiskers extend to 1.5 times the interquartile range and the notch represents the confidence interval around the median. **c**, Volcano plot for multiplex proteomics, showing in green the up-regulated proteins and in red the down-regulated, which have a Q-value < 0.1 and absolute $\log_2$ fold change > 0.58. Vertical lines indicate the $\log_2$ fold change thresholds. **d**, Protein-protein Interaction (PPI) network generated from up-regulated proteins after cytokine exposure. Node color indicates belonging to same interacting community and background corresponds to specific pathway enrichment. **e**, Proportion of up, equal or down-regulated proteins encoded by genes located <15 kb from IREs or SREs. *** Chi-squared test $P < 0.001$. **f**, An additive effect on gene up-regulation was observed for multiple IREs located at <40 kb of a gene. Box plot limits show upper and lower quartiles, whiskers extend to 1.5 times the interquartile range and the notch represents the confidence interval around the median. ANOVA $P < 2.2 \times 10^{-16}$. **g**, View of the *LY6E* locus, whose expression is induced after cytokine exposure and is coupled with chromatin changes in the vicinity.



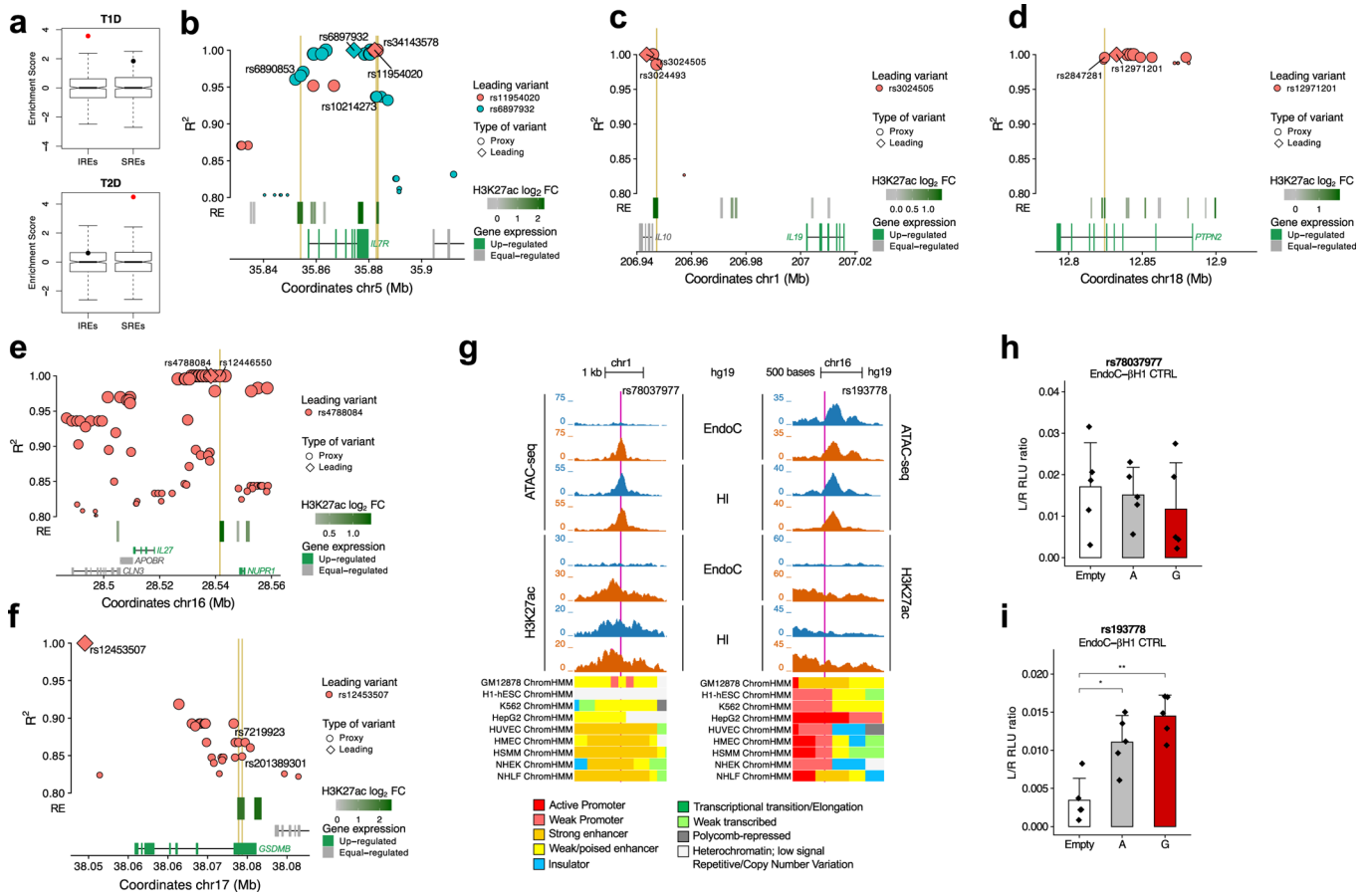
Extended Data Fig. 3 | Characterization of β-cell IREs. a, Genes associated to different classes of IREs (classified as in Fig. 2a) show cytokine-induced expression in EndoC-βH1. CYT = cytokine exposed, CTRL = control. Boxplot limits show upper and lower quartiles, whiskers extend to 1.5 times the interquartile range and notch represents the median confidence interval. ***Wilcoxon test $P < 0.001$. **b**, Sequence conservation score of IREs and a corresponding randomized set used as control. **c**, Distribution of distances to nearest TSS of the different classes of open chromatin sites. Line indicates the threshold used to classify them as 'promoters'. **d**, Number of IREs overlapping regions annotated as 'Strong' or 'Weak' enhancers by ENCODE ChromHMM. *Chi-squared $P < 2 \times 10^{-16}$. **e**, **f**, Top hits for *de novo* motif analysis in opening (**e**) and primed enhancers (**f**). Colors for matched genes correspond to RNA-seq (name) or protein (underlined) status (red = down-regulated, blue = equal-regulated, green = up-regulated, black/no-line = not-expressed/detected). **g**, Diagram showing the percentage of colocalization between the TF binding sites identified by *de novo* motif analysis in SRE and primed enhancers (that is excluding sites < 2Kb from a TSS). Label size indicates number of regions containing the TF binding sites and line width/intensity percentage of regions in which two motifs colocalize. **h**, Odds-ratio for finding a motif pair in the same enhancer in primed vs. SRE. Only significant pairs (FDR-adjusted Fisher's Exact test $P < 0.001$) are shown. Immune and islet-specific TF motifs colocalize more often in primed compared to SRE chromatin sites. **i**, Percentage of overlap between EndoC-βH1 different classes of open chromatin and islet-specific TFs obtained by ChIP-seq in untreated human islets. **j**, Volcano plot showing differentially methylated sites (depicted in red) in EndoC-βH1 exposed or not to cytokines. Dotted lines indicate the threshold for methylation differences or significance using limma moderated t-test. **k**, Distribution of demethylated and stable CpGs according to different classes of open chromatin.



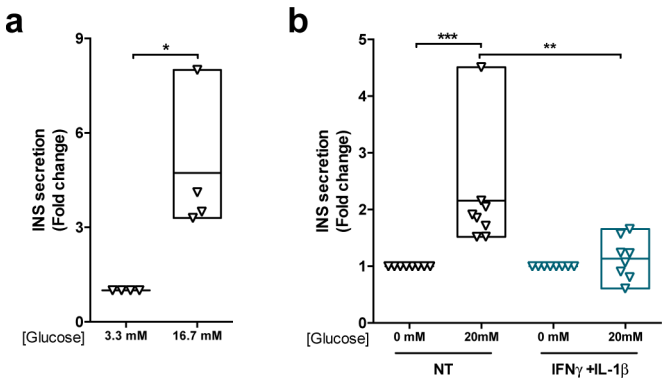
Extended Data Fig. 4 | Deconstructing cytokine induced cis-regulatory networks in β cells. a, Gene Regulatory Network (GRN) derived from IREs and their putative target genes. Squares represent the IREs inferred TF binding sites (motifs logos and TF matches are shown on the right side) and the ellipses represent their putative target genes (see Methods). The size of the squares reflects the number of connections (edge count) while the gene node size reflects the $\log_2$ fold change of RNA expression after cytokine exposure. The resulting GRN is an interconnected scale-free network composed of 648 nodes and 3,589 edges. Genes regulated exclusively by primed IREs are represented in blue while green depicts opening IREs regulated genes. Red denotes genes regulated by both types of IREs. In each of these three groups the representation of the hierarchy is based on the principle of network centrality where authoritative nodes are located more proximal to the core. **b**, Comparison between the degree distribution of the observed GRN (black triangles) and a random generated network (blue squares) having the same number of nodes and edges. The bell-shaped degree distribution of random graph denotes a statistically homogeneity in the degree of small and large nodes. In contrast, the observed network showed a high frequency of small degree nodes and a low frequency of highly connected nodes as is typical of a scale-free network. **c**, Bar plot of gene ontology biological process enrichment analysis. Gene-ontology analysis was performed using all target genes in the GRN. Functional enrichment analysis was performed by Metascape (http://metascape.org). Only terms with $P < 0.001$ and with at least 3 enriched genes were considered as significant. Color is proportional to their P values.



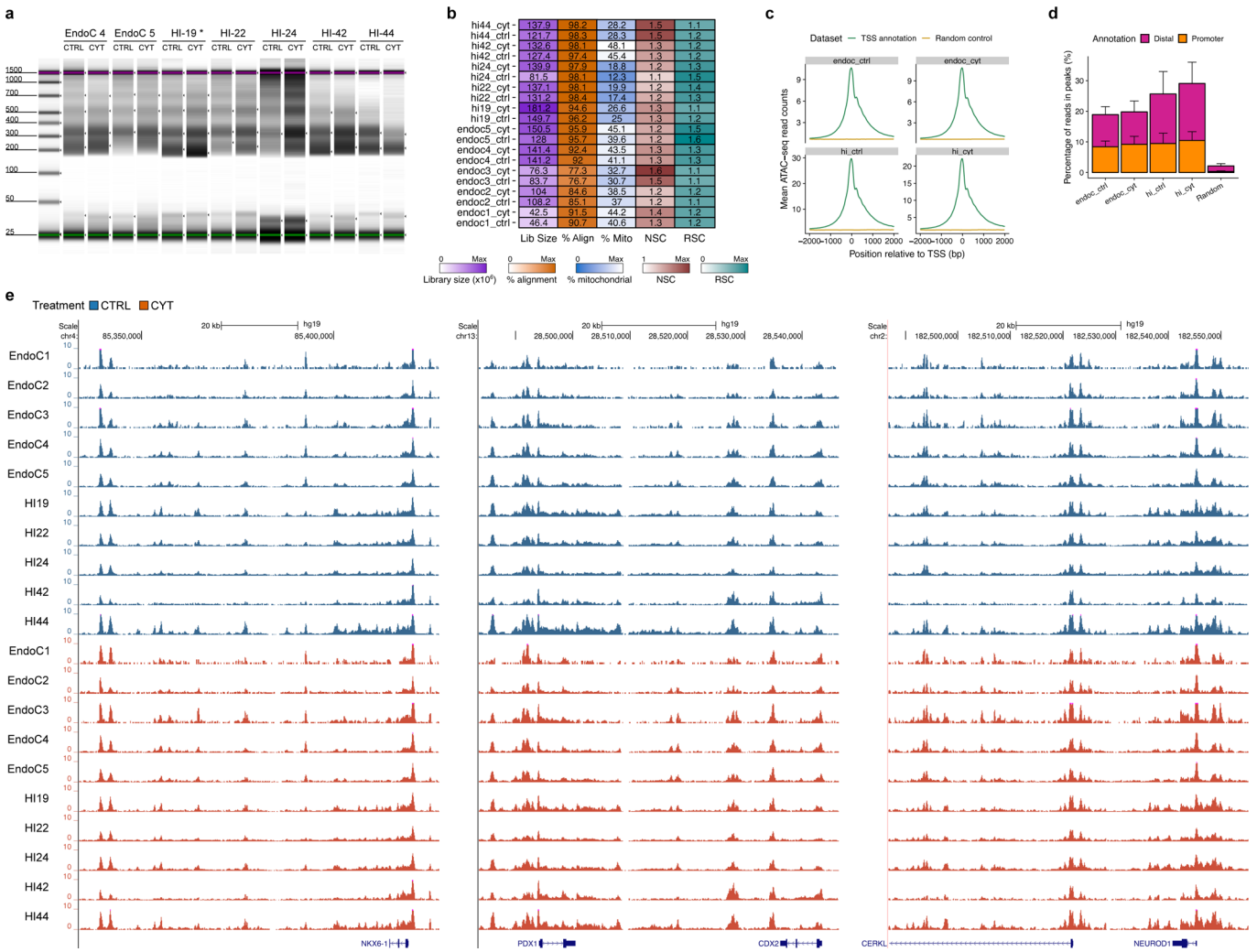
Extended Data Fig. 5 | 3D chromatin changes induced by exposure of human islets to pro-inflammatory cytokines. a, Violin plots showing the distribution of CHiCAGO scores of contacts, detected by promoter capture HiC experiments in untreated human islets[24], between stable and induced enhancers and their target genes. SREs engage chromatin contacts with higher interaction scores compared to those detected for IREs. *** Wilcoxon test $P < 0.001$. **b**, **c**, **d**, Views of the 3D chromatin contacts of *CIITA* (**b**), *SOCS1* (**c**) and *RSAD2* (**d**) promoters obtained by UMI-4C performed in islets exposed or not to pro-inflammatory cytokines. In yellow we highlight those IREs that gain contacts with the up-regulated gene promoter. A heatmap under the 4C track represents the $\log_{10}$ odds ratio (OR) of the UMI-4C contacts difference in cytokine vs. control and a small black diamond on top of the contact heatmap indicates a significant difference in contacts between cytokine-treated and control samples (Chi-squared $P < 0.05$). ATAC-seq peaks are represented by rectangles, shaded from gray to green proportionally to the cytokine-induced H2K27ac $\log_2$ fold change observed at that site.



Extended Data Fig. 6 | Cytokine-induced islet regulatory elements are enriched in T1D associated variants. a, EndoC-βH1 cytokine-induced regulatory elements (IREs) overlap more often than expected T1D associated variants while the opposite is true for T2D. EndoC-βH1 cytokine-invariant regulatory elements (SREs) are instead enriched for T2D, but not T1D associated variants. Each dot denotes the Varian Set Enrichment (VSE) score in IREs or SREs regions. Boxplot shows the enrichment distribution of the matched null permutated data sets. Red dots indicate that the difference is statistically significant as determined by VSE (Bonferroni adjusted $P < 0.05$). Box plot limits show upper and lower quartiles, whiskers extend to 1.5 times the interquartile range and the notch represents the confidence interval around the median. **b-f**, Representative regional plots of different T1D risk loci containing T1D variants overlapping IREs and up-regulated genes. R^2 values are based on 1KG phase 3 EUR and the leading SNPs in the locus is represented by a diamond. If different leading variants are present in the same locus, their proxies are depicted in different colors. Yellow squares highlight those variants that overlap a human islet IRE. IREs are depicted as boxes, with the filling color corresponding to the H3K27ac $\log_2$ fold change. **g**, The IRE bearing the T1D associated variant rs78037977 is marked by the ENCODE ChromHMM classification as a 'strong enhancer' (orange) in other non β-cell lines (left). ENCODE ChromHMM classification in non β-cell lines for the IRE bearing the T1D associated variant rs193778. **h, i**, Allele-specific luciferase experiments for rs78037977 (h) and rs193778 (i) in untreated EndoC-βH1. ANOVA followed by Bonferroni correction * $P < 0.05$; ** $P < 0.01$. Bars represent mean ± sd.



Extended Data Fig. 7 | Human islets and EndoC-βH1 Glucose-Stimulated Insulin Secretion (GSIS). GSIS was assessed, in pancreatic human islets (**a**) and EndoC-βH1 cells (**b**). In the case of EndoC-βH1 cells, the experiments were performed upon exposure or not to IFNγ (1000 U/ml) +IL1β (50U/ml) for 48 h. Data are mean plus range of four to eight independent experiments, and are expressed as the ratio between glucose stimulated and basal insulin secretion. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, for the indicated comparisons (paired t test (a) or ANOVA followed by Bonferroni correction (b)). NT = Non treated.



Extended Data Fig. 8 | ATAC-seq quality control. a, Agilent TapeStation profiles obtained by chromatin tagmentation of human islets and EndoC-βH1 samples showing the laddering pattern of ATAC-seq libraries. The band sizes correspond to the expected nucleosomal pattern. *Notice that samples HI-19 CTRL and CYT were used as examples to illustrate the expected fragment distribution pattern in ATAC-seq experiments in Raurell-Vila et al.[52]. **b**, Summary of per-replicate sequencing metrics, showing total library sizes, percentage of aligned reads, percentage of mitochondrial aligned reads, normalized strand cross-correlation coefficient (NSC, values significantly lower than 1.1 (<1.05) tend to have low signal to noise or few peaks) and relative strand cross-correlation coefficient (RSC, values significantly lower than 1 (<0.8) tend to have low signal to noise). **c**, TSS enrichment over a 4 kb window centered on genes TSS compared to a set of genes randomized along the genome, showing the expected pattern of open chromatin centered on the TSS. **d**, Percentage of total reads found at called open chromatin peaks classified as distal (>2 kb from TSS) or promoters (≤2 kb from TSS) compared to a randomized set of peaks. **e**, UCSC views at islet-specific loci (*NKX6.1*, *PDX1* and *NEUROD1*) showing the high reproducibility of ATAC-seq profiles among independent replicates.

nature research

Corresponding author(s): Lorenzo Pasquali

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	All commercial or open source code used in this analysis is referenced in the Methods section. All custom scripts are described in detail as well.
Data analysis	All commercial or open source code used in this analysis is referenced in the Methods section. All custom scripts are described in detail as well.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Datasets for induced regulatory elements (IREs) are available for download and visualization at the Islet Regulome Browser (www.isletregulome.com).

Raw sequencing reads for the different high-throughput assays can be accessed at GEO with the following identifiers: GSE123404 (ATAC-seq), GSE133135 (ChIP-seq H3K27ac), GSE137136 (RNA-seq) and GSE136865 (UMI-4C). Raw proteomics data can be accessed at ProteomeXchange with the identifier PXD011902.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size was calculated in this study.
Data exclusions	No data was excluded from this study.
Replication	Methods to determine reproducibility between replicates are described in the Methods section.
Randomization	Not applicable. All human islets used in this study came from de-identified cadaveric donors without previous history of glucose intolerance.
Blinding	Not applicable. All human islets used in this study came from de-identified cadaveric donors without previous history of glucose intolerance.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Rabbit polyclonal to Histone H3 (acetyl K27) Abcam Cat#ab4729.
Validation	The use of antibody for H3K27ac in human islets for ChiP-Seq was validated in Pasquali et al 2014 (PMID: 24413736).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The human pancreatic beta cell lines EndoC-βH1 were generated and provided by Dr. R. Scharfmann, University of Paris, France, and described in J Clin Invest. 2014 May;124(5):2087-98. doi: 10.1172/JCI72674.
Authentication	EndoC-βH1 was authenticated by comparing its epigenome and transcriptome to that of human islets (which contain about 70% of β cells) and by testing their function (insulin release, expression of human beta cell markers).
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	- GSE123404 ATAC-seq - GSE133135 H3K27ac ChIP-seq - GSE137136 RNA-seq - GSE136865 UMI-4C - PXD011902 Proteomics
Files in database submission	-- Supplementary Tables: Supplementary Table 1. List of induced regulatory elements in EndoC- βH1 cells and annotation to their closest up-regulated gene. Supplementary Table 2. List of induced regulatory elements in human pancreatic islets and annotation to their closest up-regulated gene -- Supplementary Data Sets: Supplementary Data Set 1. ATAC-seq peaks in EndoC-βH1. Supplementary Data Set 2. ATAC-seq peaks in human islets. Supplementary Data Set 3. H3K27ac ChIP-seq peaks in EndoC-βH1. Supplementary Data Set 4. H3K27ac ChIP-seq peaks in human islets.
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	- 5 ATAC-seq assays per condition in EndoC-βH1. - 5 ATAC-seq assays per condition in human islets. - 4 H3K27ac ChIP-seq assays per condition in EndoC-βH1. - 4 H3K27ac ChIP-seq assays per condidtion in human islets. - 5 RNA-seq assays per condition in EndoC-βH1. - 5 RNA-seq assays per condition in human islets. - 5 850K Infinium MethylationEPIC Arrays per condition in EndoC-βH1. - 3 multiplex proteomic assays per condition in EndoC-βH1. - 5 human islets used in UMI-4C assays. - >=3 luciferase assays per condition and vector in EndoC-βH1.
Sequencing depth	Sequencing depth for ATAC-seq, ChIP-seq and RNA-seq is provided in Supplementary Table 8. Sequencing depth for UMI-4C samples is provided in Supplementary Table 10.
Antibodies	Rabbit polyclonal to Histone H3 (acetyl K27) Abcam Cat#ab4729.
Peak calling parameters	ATAC-seq peaks were called with MACS2 (version 2.1) callpeak using the following parameters "-q 0.05 --nomodel --shift -100 --extsize 200". H3K27ac ChIP-seq peaks were identified with the same software using the following parameters "--broad --broad-cutoff 0.1 --nomodel".
Data quality	For both assays (ATAC-seq and ChIP-seq), peaks were called separately for each replicate as a measure of quality control. Merged BAM files were then used to call peaks with Q-value > 0.05. Data quality of ATAC-seq samples is shown in detail in Supplementary Figure 7.
Software	Reads were aligned using Bowtie 2 (version 2.3.4.1). Unaligned reads, reads mapping to ENCODE blacklist regions, to mitocondrial DNA or to non-autosomal chromosomes were discarded using samtools (version 1.8). Duplicates were removed with Picard MarkDuplicates (version 2.5.0). Peaks were called with MACS2 (version 2.1). Differential analysis was performed using DESeq2 (version 1.20.0).