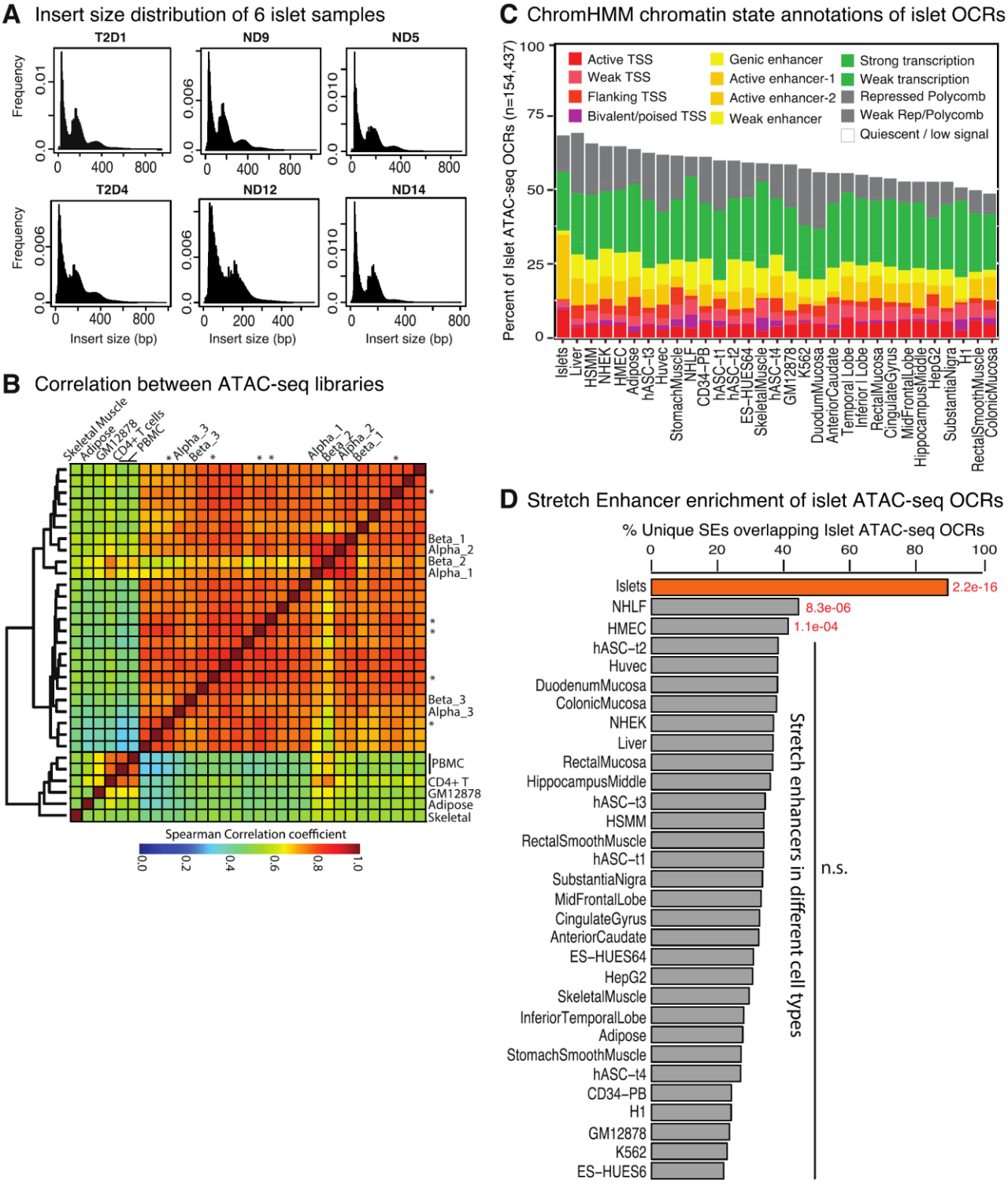
SUPPLEMENTARY DATA

Supplementary Table List

Supplementary Figure S1. Chromatin accessibility profiles of human islet samples. (A) Insert size distributions of six representative islet samples (4 ND, 2 T2D). ATAC-seq libraries capture nucleosome free and mono-, di-nucleosomal regions. (**B**) Pairwise Spearman's correlations between ATAC-seq read distributions of islet samples and other cell types ((1–3); Kursawe and Stitzel, unpublished data). 'Ackerman_Alpha' and 'Ackerman_Beta' represent ATAC-seq profiles of FACS-enriched islet alpha and beta cell populations (4). PBMC = peripheral blood mononuclear cells. (**C**) Stacked bar plot showing percent overlap of islet OCRs (n=154,437) with ChromHMM chromatin states in islets and 30 other cell types. Tissues are ordered based on the overlap of OCRs with the enhancer states. NHLF = normal human lung fibroblasts; HMEC = human mammary epithelial cells; hASC = human adipose derived stem cells; t1-4 denote adipogenesis stages as reported (5); HUVEC = human umbilical vein cells; NHEK = normal human epidermal keratinocytes; HSMM = human skeletal muscle cells; ES-HUES = human embryonic stem cells; CD34-PB = peripheral blood (PB) CD34+ cells; H1 = human embryonic stem cells; K562 = immortalized myelogenous leukemia. TSS = Transcription Start Site; Rep = repressed. (**D**) Fraction of unique tissue-specific stretch enhancers (SEs) overlapping one or more islet OCRs. Fisher's exact test p-values are shown; n.s. = not significant. Cell type abbreviations are as described in S1C above.

Khetan_Fig_S1

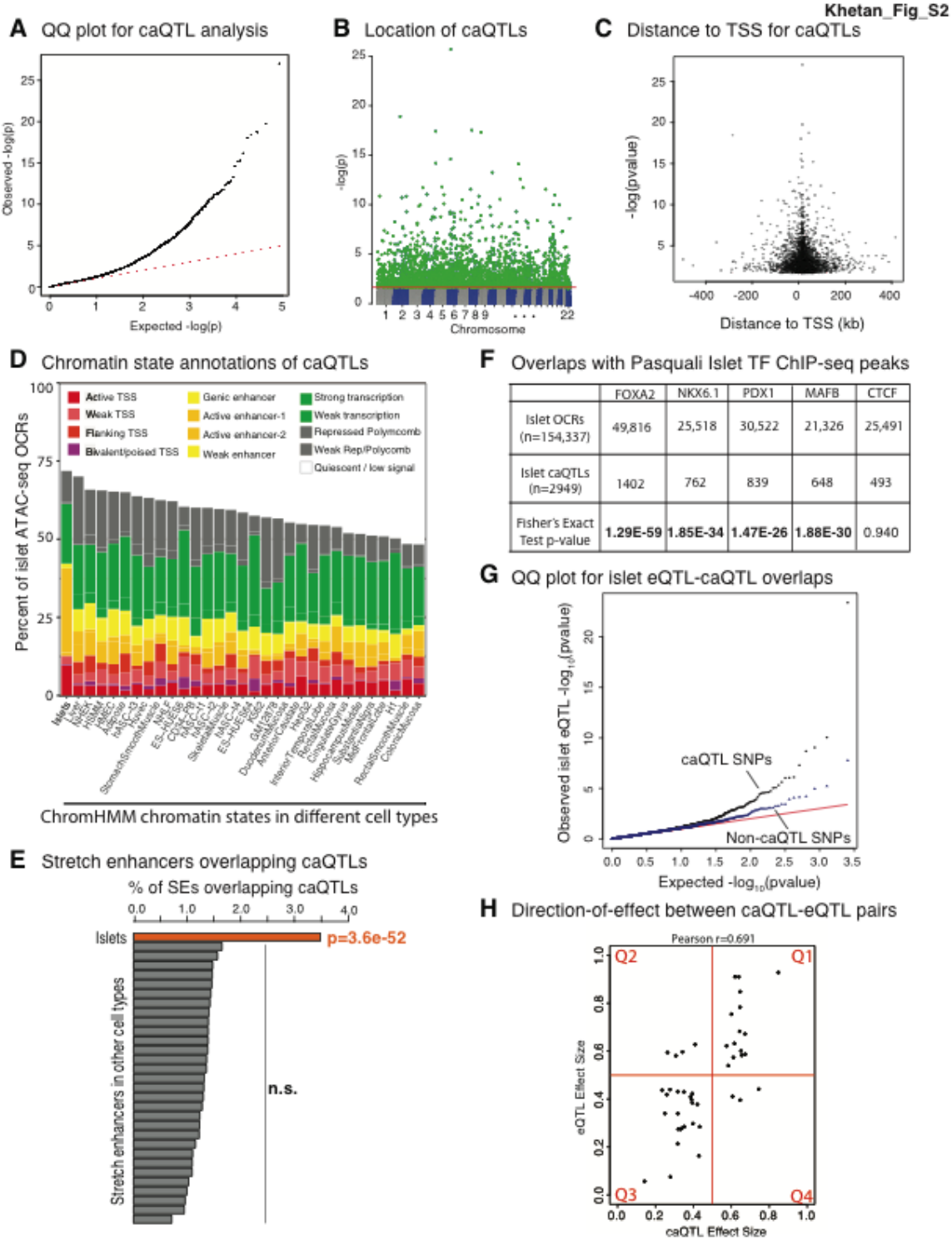


SUPPLEMENTARY DATA

Supplementary Figure S2. Islet chromatin accessibility QTLs. (A) Quantile-quantile (QQ) plot of observed (y-axis) vs. expected (x-axis) association p-values between islet chromatin accessibility and genotype. (**B**) Location of significant caQTLs (marked in green) across the autosomal chromosomes. (**C**) Density scatter plot showing the distance (x-axis; kb=kilobases) between 2949 islet caQTLs (y-axis; FDR 10%) and the transcription start site (TSS) of the nearest islet-expressed gene. The majority (>98.5%) of caQTLs are within the flanking 200 kb of the TSS of the nearest expressed gene. (**D**) Chromatin state annotations of caQTLs in islets vs. other tissues. Tissues are sorted from lowest to highest overlap between OCRs and the 'Quiescent/Low Signal' state. Note the enrichment of islet enhancer states in caQTLs. (**E**) Percent of tissue-specific SEs overlapping one or more islet caQTL. Fisher's exact test p-values are shown; n.s. = not significant. Cell type abbreviations are as described in Figure S1C. (**F**) Table showing the overlaps of islet TF binding sites (i.e., ChIP-seq peaks) with all islet OCRs and caQTL-containing islet OCRs. Fisher's exact test was used to calculate the significance of overlaps with the caQTL-containing islet OCRs. (**G**) QQ plot of observed (y-axis) vs. expected (x-axis) islet eQTL (eQTL from 112 individuals; (6)) association p-values for islet caQTL SNPs (black) or randomly selected non-caQTL SNPs (blue). Higher enrichment of eQTLs among statistically significant caQTLs links regulation of chromatin accessibility to gene expression. (**H**) Direction-of-effect between caQTL-eQTL pairs in 19 islet cohort. 37/44 islet caQTL-eQTL pairs (84%) show concordant changes (Q1 and Q3) in direction-of-effect on chromatin accessibility (x-axis) and gene expression (y-axis) (Pearson coefficient r=0.691; p=1.992e-07). Effect size < 0.5 signifies that the reference allele is more accessible/expressed than the alternate allele.

SUPPLEMENTARY DATA

Khetan_Fig_S2



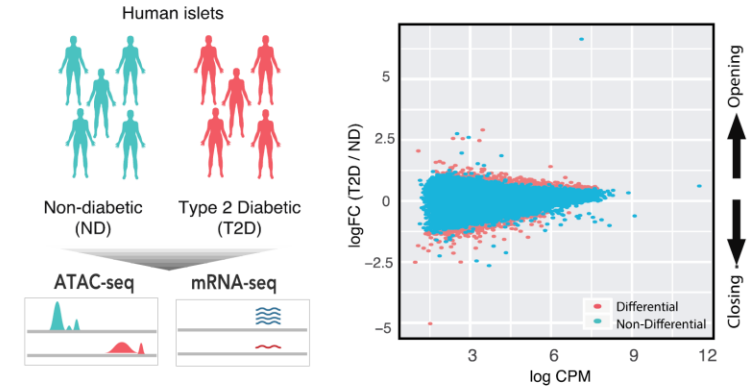
	FOXA2	NKX6.1	PDX1	MAFB	CTCF
Islet OCRs (n=154,337)	49,816	25,518	30,522	21,326	25,491
Islet caQTLs (n=2949)	1402	762	839	648	493
Fisher's Exact Test p-value	**1.29E-59**	**1.85E-34**	**1.47E-26**	**1.88E-30**	0.940

SUPPLEMENTARY DATA

Supplementary Figure S3. T2D disease state-associated chromatin accessibility changes. (A) Schematic of comparative analyses of five T2D and five ND islet samples. (**B**) The weighted average proportion variance explained by meta-variables available from the 10 islet samples before and after batch correction using Surrogate Variable Analysis (SVA) (7). Note that SVA reduces the variance attributed to all meta-variables, except for the disease state (ND or T2D). (**C**) MA plot for differential OCR analyses (at FDR 10%). Every dot represents an OCR considered in the analyses (n=52,387). Positive log fold change (logFC) indicates that the chromatin is opening in T2D islet samples, and negative logFC indicates that the chromatin is closing in T2D islet samples (CPM=counts per million). Differential OCRs are red; non-differential ones are shown in blue. (**D**) Transcription factor (TF) motif enrichment p-values and q-values for opening and closing OCRs. Mutually exclusive lists of TFs are enriched in closing and opening OCRs. (**E**) Gene expression changes (measured in log fold change of read counts from RNA-seq profiles) at promoters with OCRs (n=7499). Genes with opening OCRs exhibit positive fold change, i.e., increased expression in T2D islet samples; genes with closing OCRs show negative fold change, i.e., decreased expression in T2D islet samples (p=3.186e-06, Welch's Two Sample two-sided t-test). (**F**) Venn diagram showing the number of differentially accessible OCRs overlapping islet caQTLs. 90/1515 (5.9%) of differential OCRs overlap islet caQTLs.

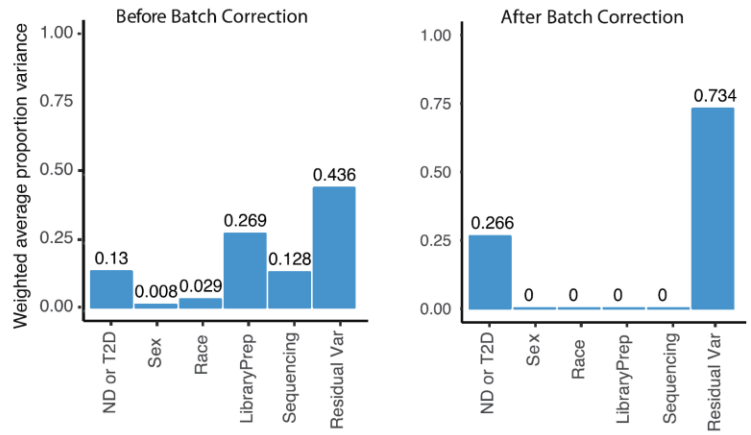
SUPPLEMENTARY DATA

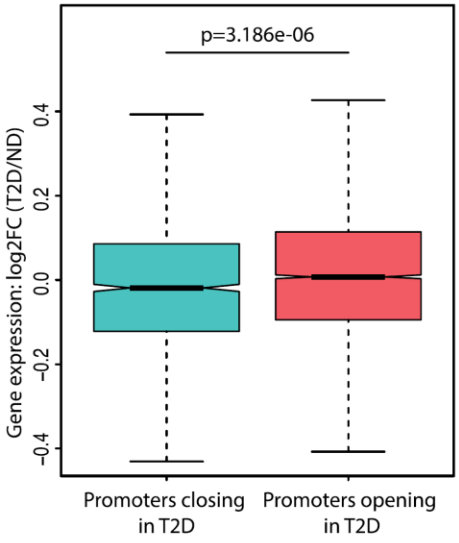
Khetan_Fig_S3

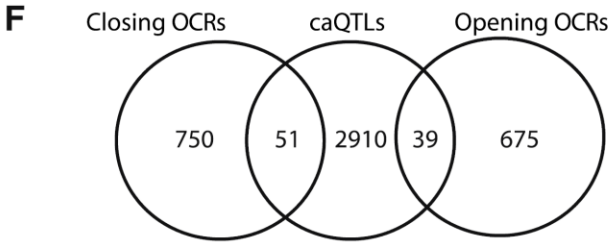


D TF motif enrichment at differential OCRs

	TF	Closing q-value	Closing p-value	Opening q-value	Opening p-value
Motifs enriched in opening	NeuroG2	1	1	3.64E-07	1e-09
	AR-Half	1	1	1.21E-06	1e-08
	SCL	1	1	1.21E-06	1e-08
	Atoh1	1	1	7.28E-06	1e-07
	Ascl1	1	1	7.28E-06	1e-07
	NeuroD1	1	1	4.04E-05	1e-06
	ZNF519	1	1	4.04E-05	1e-06
	NF1-Half	1	1	4.04E-05	1e-06
	Olig2	1	1	4.04E-05	1e-06
	Tcf21	1	1	1.00E-04	1e-05
Motifs enriched in closing	Atf3	3.64E-47	1e-49	1	1
	AP-1	1.82E-46	1e-48	1	1
	BATF	9.10E-45	1e-46	1	1
	JunB	9.10E-45	1e-46	1	1
	Fra1	7.28E-44	1e-45	1	1
	Fra2	6.07E-38	1e-39	1	1
	Fosl2	5.20E-24	1e-25	1	1
	Jun-AP1	4.55E-18	1e-19	1	1
	Hoxb4	4.04E-08	1e-09	1	0.6285
	Pdx1	3.64E-07	1e-08	1	0.6921
	Bach1	2.00E-04	1e-05	1	1



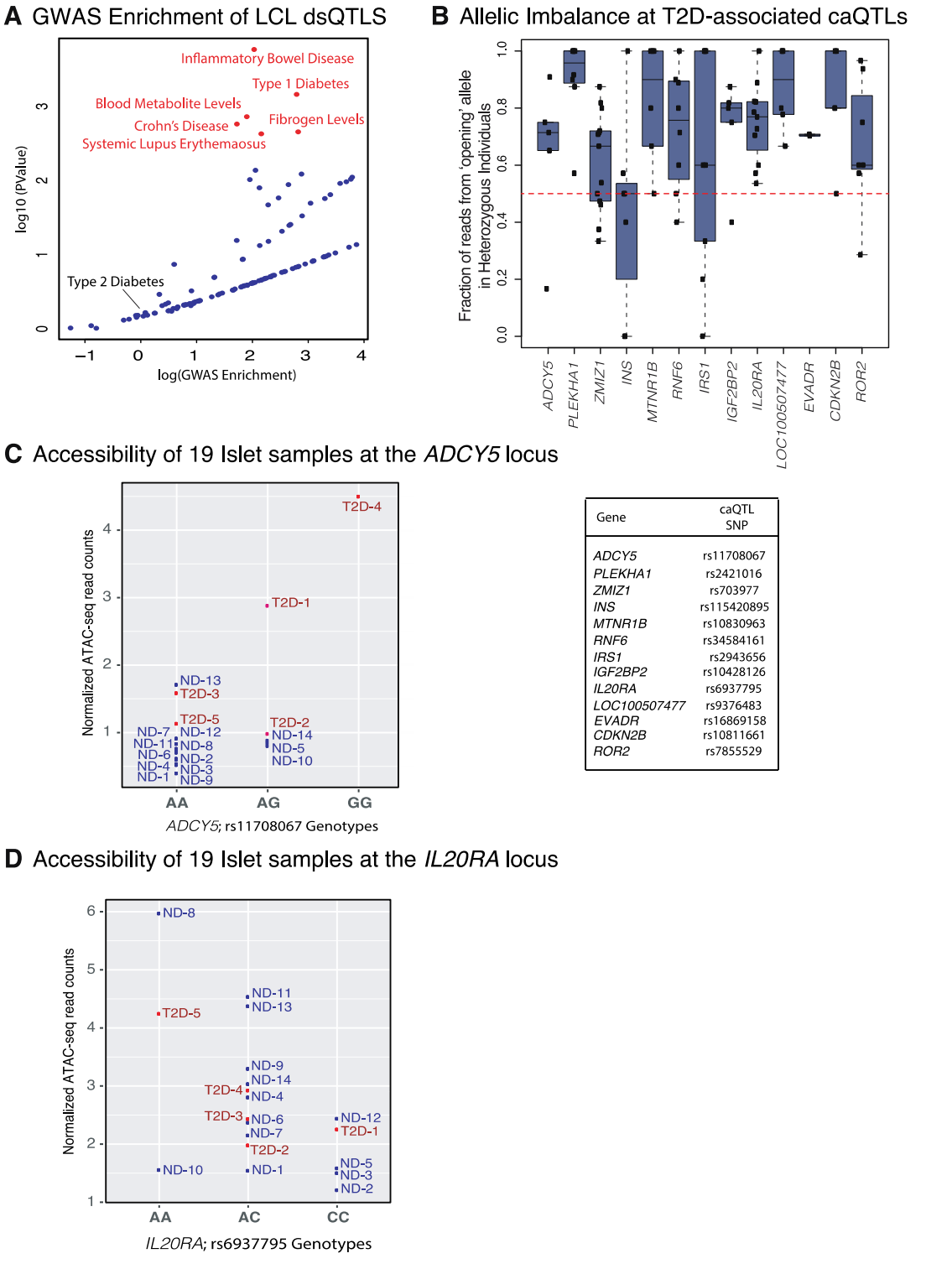




SUPPLEMENTARY DATA

Supplementary Figure S4. Chromatin accessibility QTLs in islets. (A) Lymphoblastoid cell line (LCL) DNase I sensitivity QTLs (dsQTLs; (8)) are enriched for GWAS SNPs associated with immune-related diseases (e.g., Inflammatory Bowel Disease, Crohn's Disease, lupus, Type 1 Diabetes) and not for those associated with T2D or fasting glucose as observed with islet caQTLs. Diseases/traits with GWAS SNPs significantly overlapping LCL dsQTLs at FDR 5% are highlighted red. (**B**) Allelic imbalance of ATAC-seq reads for T2D-associated islet caQTL SNPs in heterozygous islet donors. Dot-and-box plots show the fraction of total ATAC-seq reads containing the 'opening' allele for each heterozygous individual. caQTL SNP rsID for each corresponding locus is indicated in the table below. The dashed red line at y=0.5 denotes equal representation of each allele. (**C-D**) Chromatin accessibility of 19 islet samples stratified by genotypes at caQTL SNP rs11708067 in the *ADCY5* locus (**C**) and at rs6937795 in the *IL20RA* locus (**D**). Note that T2D islet samples (highlighted in red) are distributed across all three genotypes at each locus.

Khetan_Fig_S4



Gene	caQTL SNP
ADCY5	rs11708067
PLEKHA1	rs2421016
ZMIZ1	rs703977
INS	rs115420895
MTNR1B	rs10830963
RNF6	rs34584161
IRS1	rs2943656
IGF2BP2	rs10428126
IL20RA	rs6937795
LOC100507477	rs9376483
EVADR	rs16869158
CDKN2B	rs10811661
ROR2	rs7855529

SUPPLEMENTARY DATA

Supplemental References Cited:

1. Scott LJ, Erdos MR, Huyghe JR, Welch RP, Beck AT, Wolford BN, et al. The genetic regulatory signature of type 2 diabetes in human skeletal muscle. Nat Commun. 2016;7:11764.

2. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013 Dec;10(12):1213–8.

3. Ucar D, Márquez EJ, Chung C-H, Marches R, Rossi RJ, Uyar A, et al. The chromatin accessibility signature of human immune aging stems from CD8(+) T cells. J Exp Med. 2017 Oct 2;214(10):3123–44.

4. Ackermann AM, Wang Z, Schug J, Naji A, Kaestner KH. Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes. Mol Metab. 2016 Mar;5(3):233–44.

5. Mikkelsen TS, Xu Z, Zhang X, Wang L, Gimble JM, Lander ES, et al. Comparative epigenomic analysis of murine and human adipogenesis. Cell. 2010 Oct 1;143(1):156–69.

6. Varshney A, Scott LJ, Welch RP, Erdos MR, Chines PS, Narisu N, et al. Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. Proc Natl Acad Sci U S A. 2017 Feb 13;

7. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinforma Oxf Engl. 2012 Mar 15;28(6):882–3.

8. Degner JF, Pai AA, Pique-Regi R, Veyrieras J-B, Gaffney DJ, Pickrell JK, et al. DNase I sensitivity QTLs are a major determinant of human expression variation. Nature. 2012 Feb 16;482(7385):390–4.