Cell Reports, Volume *26*

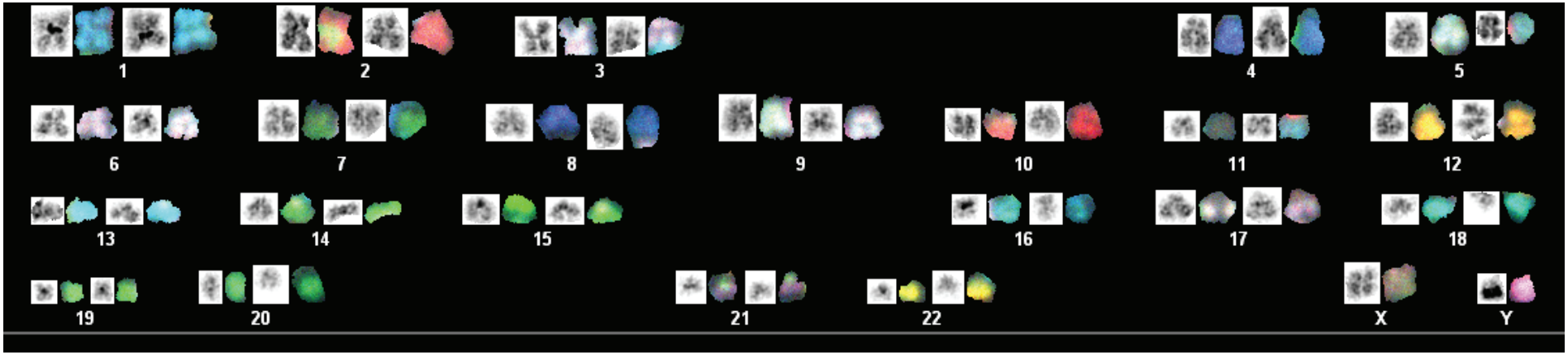
Supplemental Information

Multiomic Profiling Identifies *cis*-Regulatory Networks Underlying Human Pancreatic β Cell Identity and Function

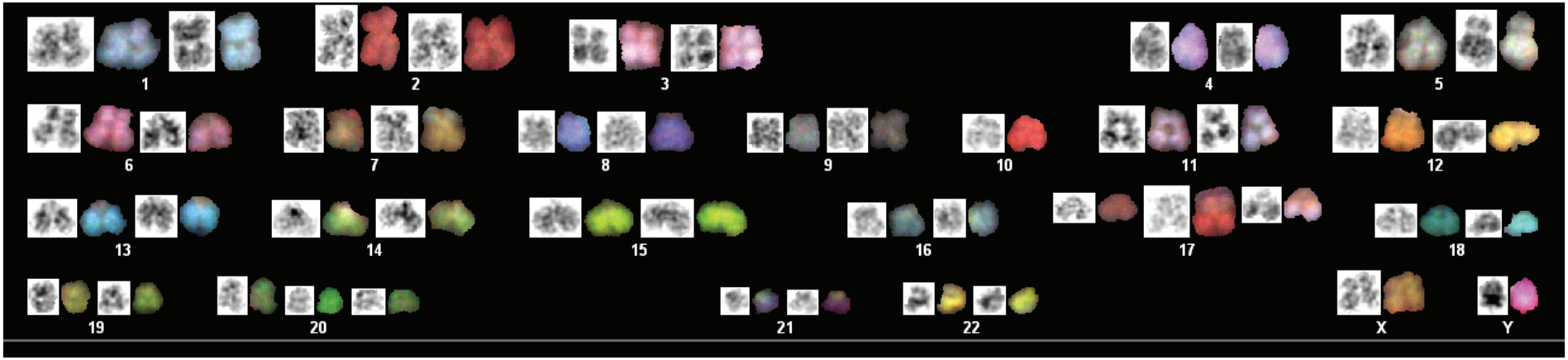
Nathan Lawlor, Eladio J. Márquez, Peter Orchard, Narisu Narisu, Muhammad Saad Shamim, Asa Thibodeau, Arushi Varshney, Romy Kursawe, Michael R. Erdos, Matt Kanke, Huiya Gu, Evgenia Pak, Amalia Dutra, Sheikh Russell, Xingwang Li, Emaly Piecuch, Oscar Luo, Peter S. Chines, Christian Fuchbserger, NIH Intramural Sequencing Center, Praveen Sethupathy, Aviva Presser Aiden, Yijun Ruan, Erez Lieberman Aiden, Francis S. Collins, Duygu Ucar, Stephen C.J. Parker, and Michael L. Stitzel

Common event to multiple metaphases

SKY EndoC-βH1 Metaphase S2.1 46,XY,der(17)t(10;17),der(17)t(3;17),der(17)t(3;17)



SKY EndoC-βH1 Metaphase 7 47,XY,t(7;18),-10,t(10;17),t(3;17),der(17)t(3;17),+20



SKY EndoC-βH1 Metaphase S2.7 48,XY,t(3;21),-10,+der(17)t(10;17),+20

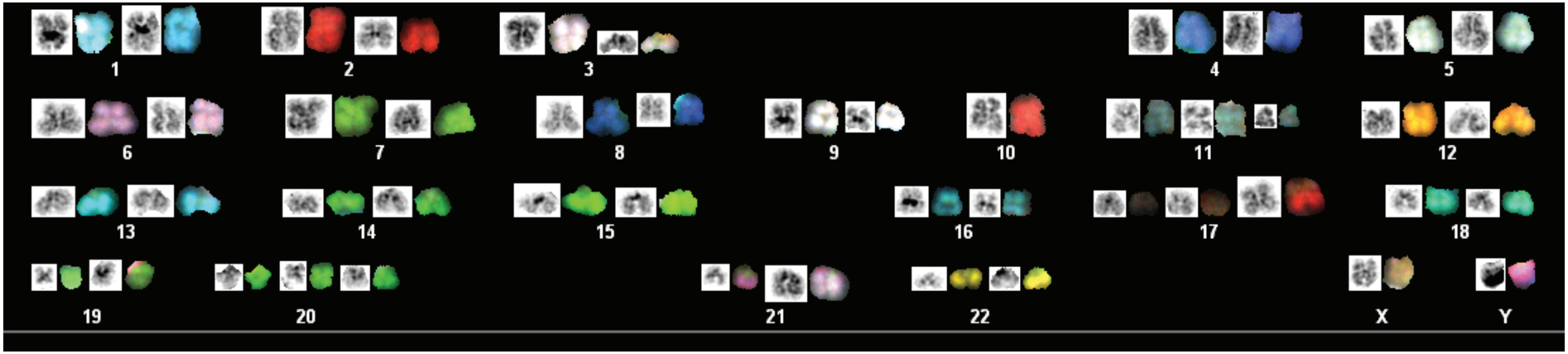


Figure S1, related to Figure 1: Representative spectral karyotypes (SKY) of EndoC-ßH1 cells at metaphase. Specific metaphases shown are S2.1 (top), 7 (middle), and S2.7 (bottom). Common structural or numerical chromosomal aberrations that are evident in multiple metaphases are indicated in green.

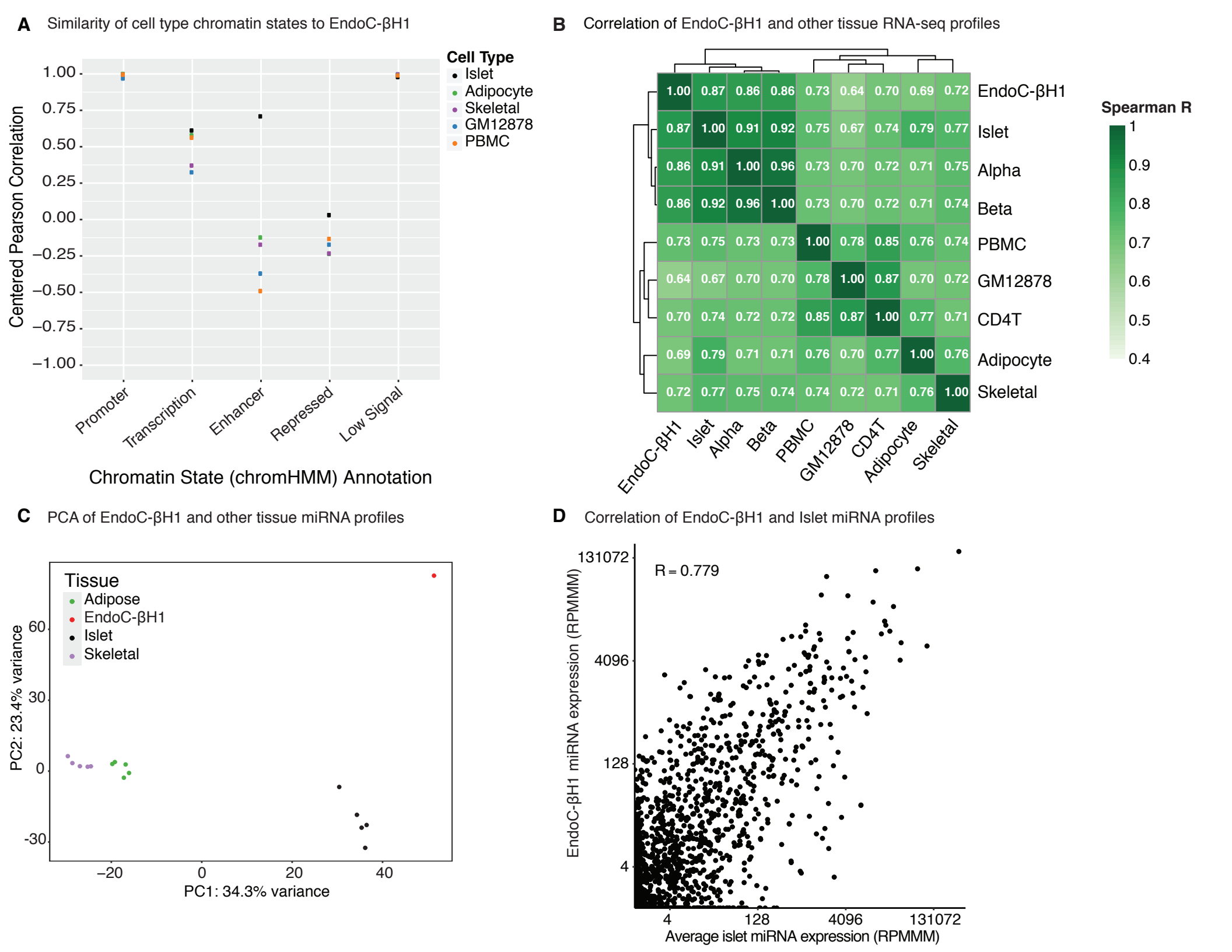

A
Similarity of cell type chromatin states to EndoC-βH1
Cell Type
Islet
Adipocyte
Skeletal
GM12878
PBMC
Centered Pearson Correlation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Promoter
Transcription
Enhancer
Repressed
Low Signal
Chromatin State (chromHMM) Annotation
B
Correlation of EndoC-βH1 and other tissue RNA-seq profiles
Spearman R
1
0.9
0.8
0.7
0.6
0.5
0.4
1.00 0.87 0.86 0.86 0.73 0.64 0.70 0.69 0.72 EndoC-βH1
0.87 1.00 0.91 0.92 0.75 0.67 0.74 0.79 0.77 Islet
0.86 0.91 1.00 0.96 0.73 0.70 0.72 0.71 0.75 Alpha
0.86 0.92 0.96 1.00 0.73 0.70 0.72 0.71 0.74 Beta
0.73 0.75 0.73 0.73 1.00 0.78 0.85 0.76 0.74 PBMC
0.64 0.67 0.70 0.70 0.78 1.00 0.87 0.70 0.72 GM12878
0.70 0.74 0.72 0.72 0.85 0.87 1.00 0.77 0.71 CD4T
0.69 0.79 0.71 0.71 0.76 0.70 0.77 1.00 0.76 Adipocyte
0.72 0.77 0.75 0.74 0.74 0.72 0.71 0.76 1.00 Skeletal
EndoC-βH1
Islet
Alpha
Beta
PBMC
GM12878
CD4T
Adipocyte
Skeletal
C
PCA of EndoC-βH1 and other tissue miRNA profiles
Tissue
Adipose
EndoC-βH1
Islet
Skeletal
PC2: 23.4% variance
60
30
0
-30
−20
0
20
40
PC1: 34.3% variance
D
Correlation of EndoC-βH1 and Islet miRNA profiles
R = 0.779
EndoC-βH1 miRNA expression (RPMMM)
131072
4096
128
4
4
128
4096
131072
Average islet miRNA expression (RPMMM)

Figure S2, related to Figure 2: Comparison of EndoC-ßH1 and islet epigenomic and transcriptomic profiles.
(A) Similarity (centered Pearson correlation) of human islet, adipocyte, skeletal muscle, GM12878, and PBMC chromatin state annotations (ChromHMM) to those of EndoC-ßH1 at EndoC-ßH1 OCRs. The similarity matrix was calculated using the "simil" function within the proxy version 0.4 R package (Meyer and Buchta, 2018).
(B) Spearman correlation between EndoC-ßH1 RNA-seq and the corresponding RNA-seq from islets, sorted alpha or beta cells (Ackermann et al., 2016), and other cell types and tissues. PBMC=peripheral blood mononuclear cells, GM12878 = lymphoblastoid cell line, CD4T = CD4+ T immune cell, skeletal = skeletal muscle, Alpha = primary islet alpha cells, Beta = primary islet beta cells.
(C) Principal component analysis (PCA) of miRNA-seq profiles from EndoC-ßH1 and 5 representative human islet, skeletal muscle, and adipose tissue samples. (D) Scatter plot illustrating the resemblance of miRNA expression levels between EndoC-ßH1 and human islets. RPMMM = reads per million mapped miRNA, R = Pearson R.

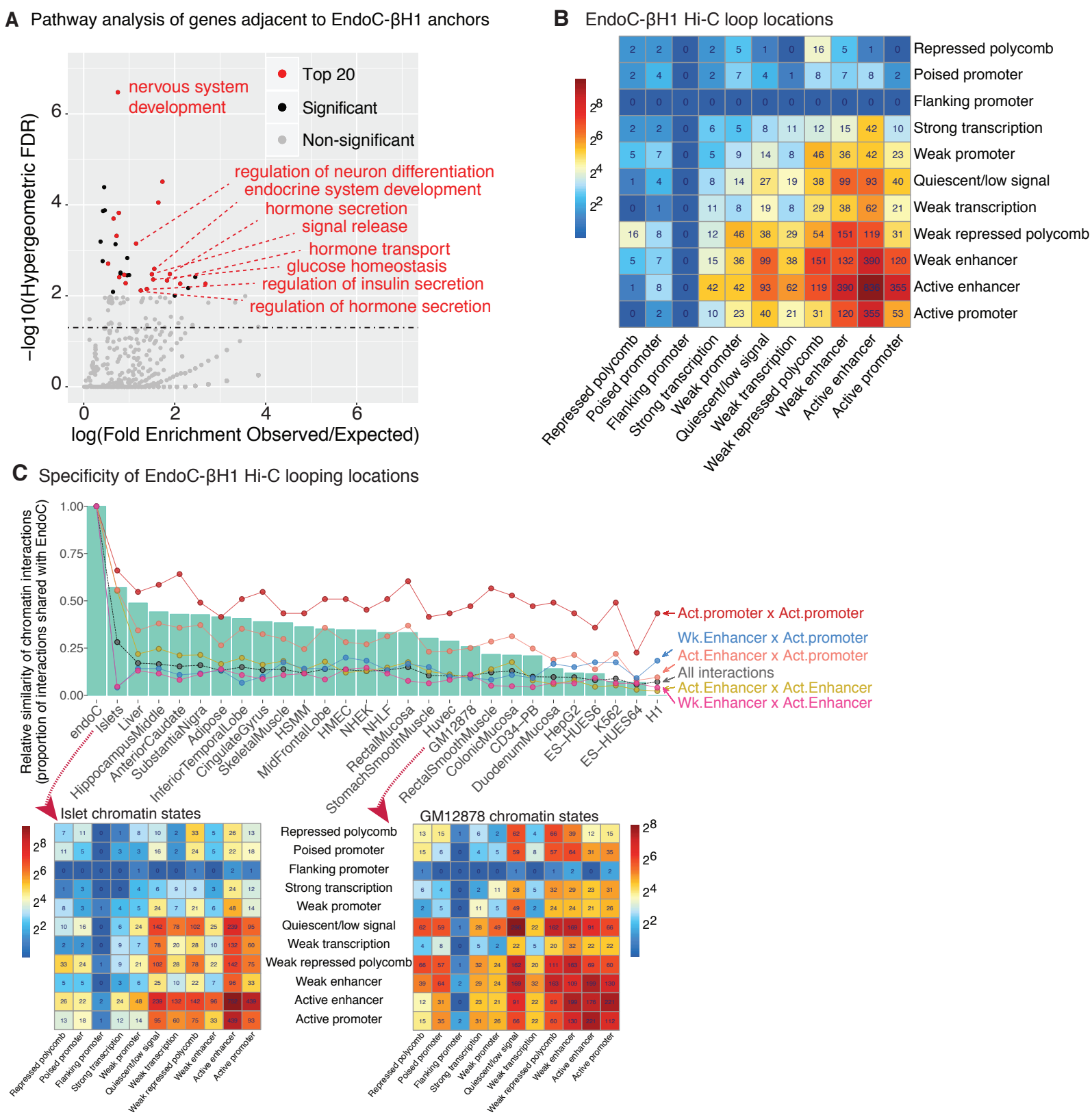
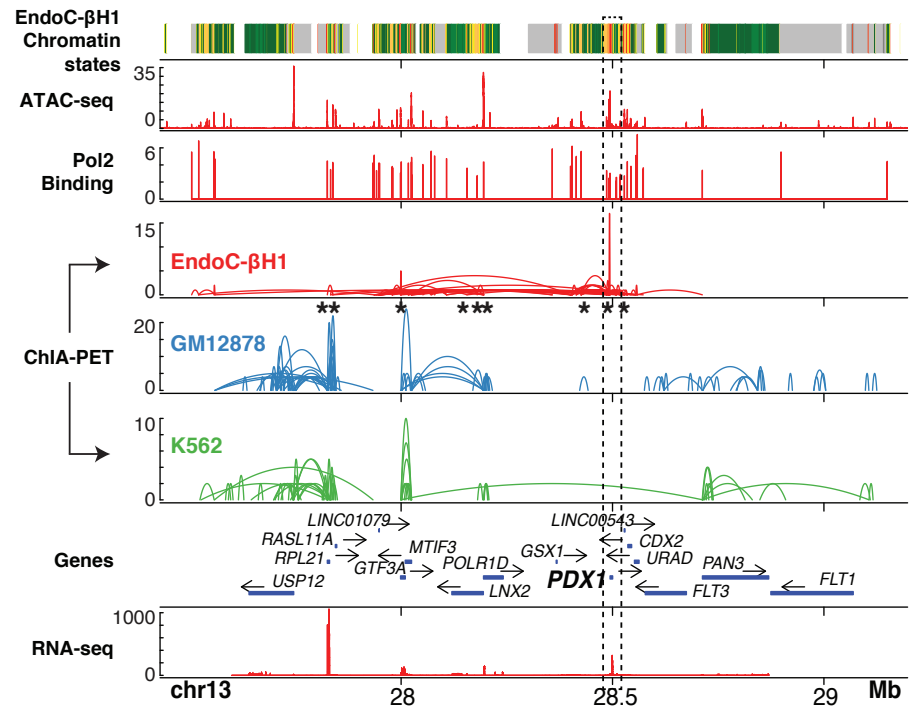
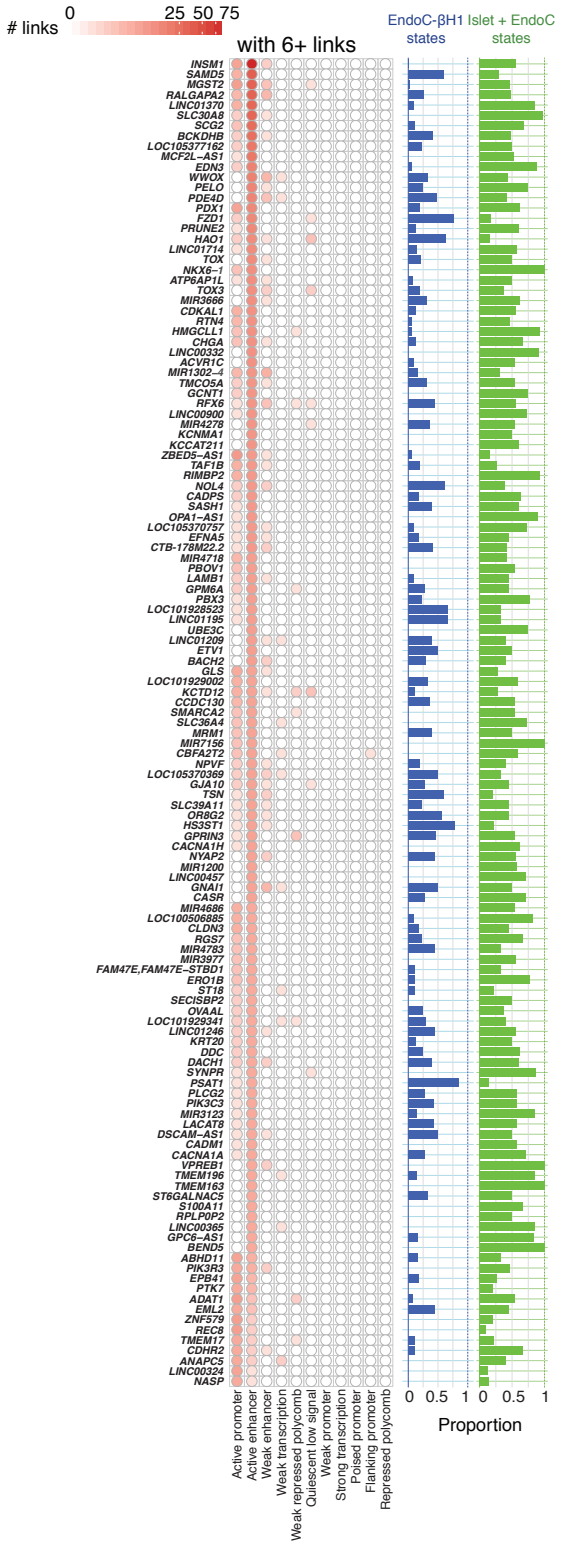

A Pathway analysis of genes adjacent to EndoC-βH1 anchors
Top 20
Significant
Non-significant
nervous system development
regulation of neuron differentiation
endocrine system development
hormone secretion
signal release
hormone transport
glucose homeostasis
regulation of insulin secretion
regulation of hormone secretion
−log10(Hypergeometric FDR)
log(Fold Enrichment Observed/Expected)
B EndoC-βH1 Hi-C loop locations
Repressed polycomb
Poised promoter
Flanking promoter
Strong transcription
Weak promoter
Quiescent/low signal
Weak transcription
Weak repressed polycomb
Weak enhancer
Active enhancer
Active promoter
C Specificity of EndoC-βH1 Hi-C looping locations
Relative similarity of chromatin interactions (proportion of interactions shared with EndoC)
Act.promoter x Act.promoter
Wk.Enhancer x Act.promoter
Act.Enhancer x Act.promoter
All interactions
Act.Enhancer x Act.Enhancer
Wk.Enhancer x Act.Enhancer
endoC
Islets
Liver
HippocampusMiddle
AnteriorCaudate
SubstantiaNigra
Adipose
InferiorTemporalLobe
CingulateGyrus
SkeletalMuscle
HSMM
MidFrontalLobe
HMEC
NHEK
NHLF
RectalMucosa
StomachSmoothMuscle
Huvec
GM12878
RectalSmoothMuscle
ColonicMucosa
CD34−PB
DuodenumMucosa
HepG2
ES−HUES6
K562
ES−HUES64
H1
Islet chromatin states
GM12878 chromatin states

Figure S3, related to Figure 3: Enrichment analysis and annotation of EndoC-ßH1 Hi-C loops.
(A) Biological process gene ontology terms enriched in genes adjacent to EndoC-ßH1-specific Hi-C anchors. Enrichment analysis was performed with GREAT (McLean et al., 2010) using the single gene whose transcription start site (TSS) was nearest to each anchor. Results with an adjusted (Benjamini & Hochberg) hypergeometric p-value < 0.05 were regarded as statistically significant. (B) Frequency of EndoC-ßH1 Hi-C loop anchors and their corresponding chromatin state (ChromHMM) annotations. (C) (Top) Bar plot depicting the average similarity of chromatin state (ChromHMM) annotations for each cell type at EndoC-ßH1 Hi-C loop anchor positions. Overlaid line plots highlight the relative similarity for various chromatin state interactions (e.g. active promoter vs. active promoter). (Bottom) Select heat maps showing the frequency of human islet (left) and GM12878 (right) chromatin states at EndoC-ßH1 loop anchors.


A
Connectivity at PDX1 locus
EndoC-βH1 Chromatin states
ATAC-seq
Pol2 Binding
ChIA-PET
EndoC-βH1
GM12878
K562
Genes
LINC01079
LINC00543
RASL11A
MTIF3
CDX2
RPL21
GTF3A
POLR1D
GSX1
URAD
PAN3
USP12
LNX2
PDX1
FLT3
FLT1
RNA-seq
chr13
28
28.5
29
Mb


B
Network connectivity of EndoC-βH1 anchor genes
links
with 6+ links
EndoC-βH1 states
Islet + EndoC states
Active promoter
Active enhancer
Weak enhancer
Weak transcription
Weak repressed polycomb
Quiescent low signal
Weak promoter
Strong transcription
Poised promoter
Flanking promoter
Repressed polycomb
Proportion

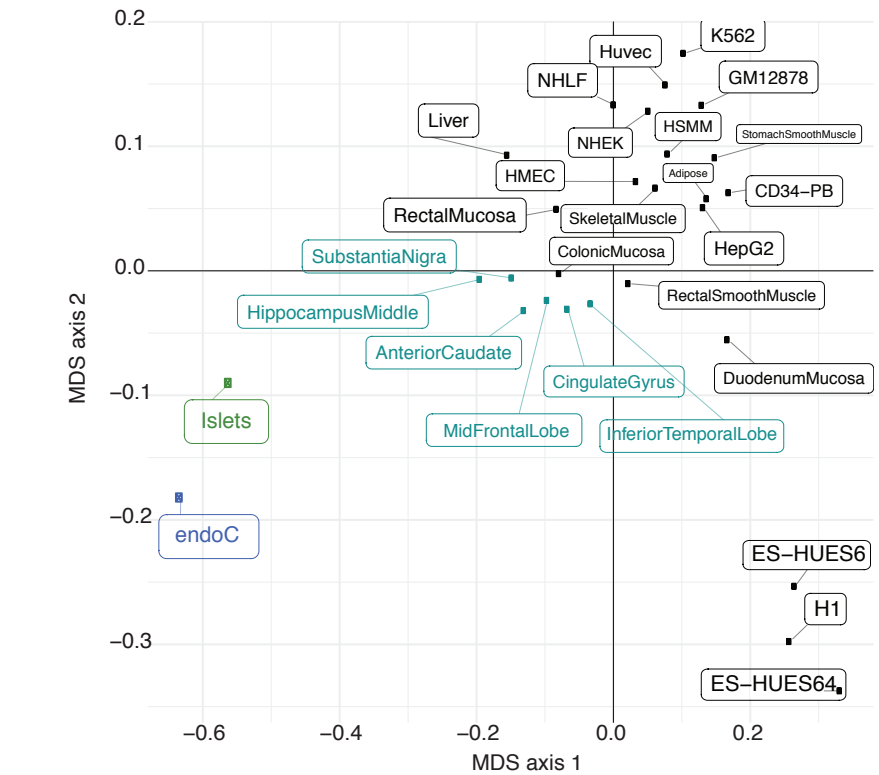
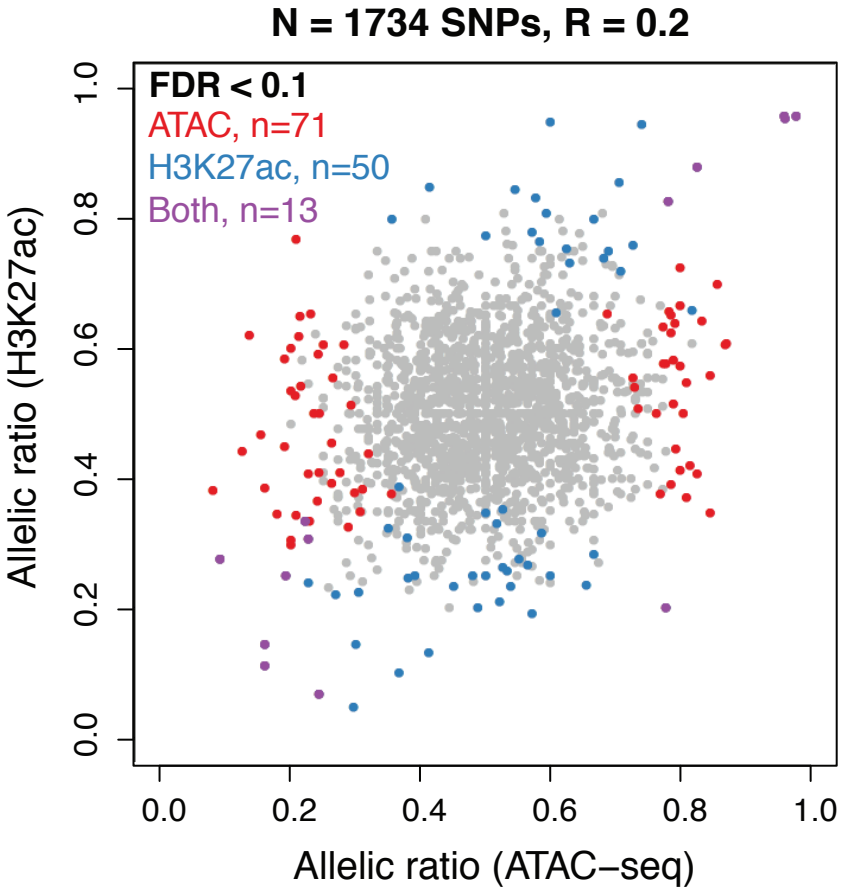

C
MDS based on dissimilarity with EndoC-βH1 chromHMM states
MDS axis 1
MDS axis 2
K562
Huvec
NHLF
GM12878
Liver
HSMM
NHEK
StomachSmoothMuscle
HMEC
Adipose
CD34−PB
RectalMucosa
SkeletalMuscle
HepG2
SubstantiaNigra
ColonicMucosa
RectalSmoothMuscle
HippocampusMiddle
AnteriorCaudate
CingulateGyrus
DuodenumMucosa
Islets
MidFrontalLobe
InferiorTemporalLobe
endoC
ES−HUES6
H1
ES−HUES64

Figure S4, related to Figure 4: Annotation and evaluation of EndoC-ßH1 RNA Pol2 ChIA-PET interactions.
(A) Genome-wide view of RNA polymerase 2-mediated (Pol2 ChIA-PET) chromatin interactions around the PDX1 locus on chromosome 13 in EndoC-ßH1. Asterisks under EndoC-ßH1 (red) ChIA-PET interactions indicate interacting sites/anchors identified by targeted 4C-seq analyses in this locus in human islets (Pasquali et al., 2014). (B) Connectivity of EndoC-ßH1 ChIA-PET interactions when considering nodes with 6 or more links to other regulatory elements. (Left) Circular plots depict the number of links that occur with corresponding regulatory elements (e.g., active promoter, active enhancer). (Right) bar plots illustrate the proportion of interacting nodes that exhibit the active regulatory element chromatin states exclusively in EndoC-ßH1 (blue) or identical chromatin states in both EndoC-ßH1 and islet (green). (C) Multidimensional scaling (MDS) plot based on pairwise Chi-square distances of vectors of proportions of chromatin states in EndoC-ßH1, islets, and additional Epigenomics Roadmap cell and tissue types at EndoC-ßH1 defined ChIA-PET interacting nodes.

A Correlation of cis-regulatory element allelic imbalance



B Cis-regulatory element imbalance at human islet eQTL SNP-gene pairs

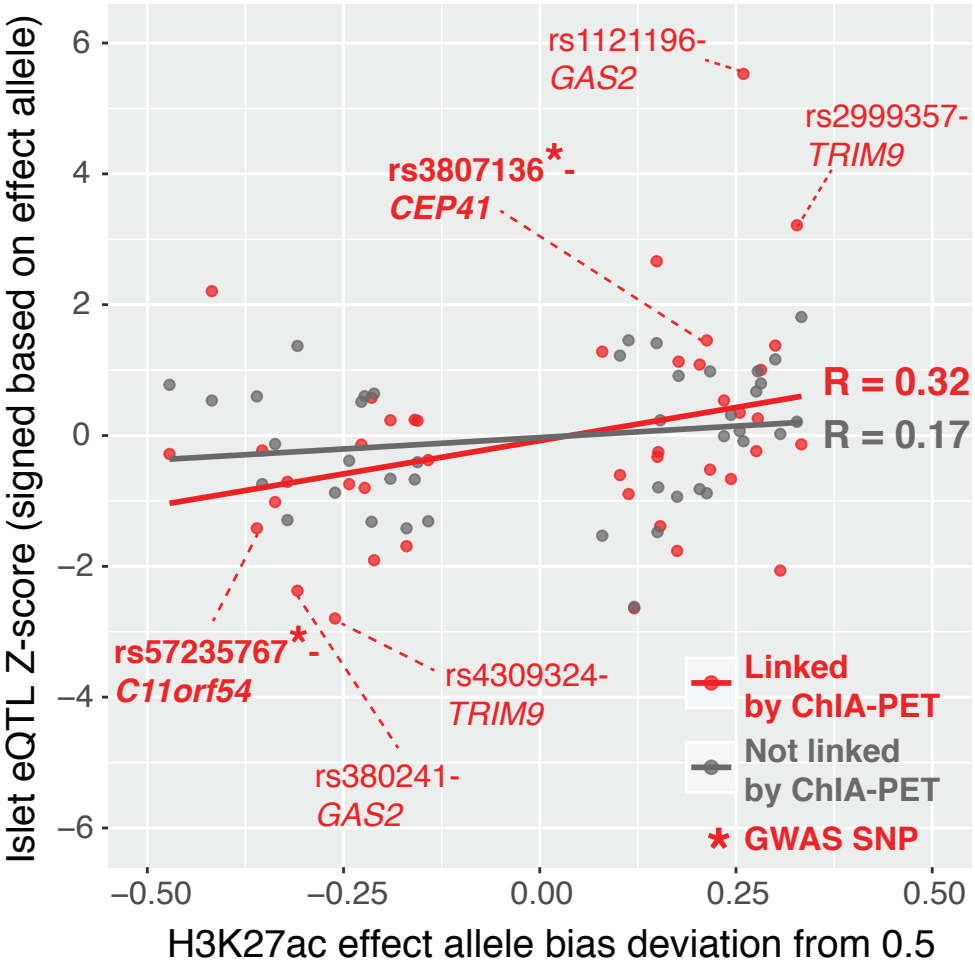


Figure S5, related to Figure 5: Assessment of cis-regulatory element allelic imbalance in EndoC-ßH1 and islets.

(A) Correlation of cis-regulatory element allelic imbalance for SNPs with heterozygous genotypes in EndoC-ßH1. For a given heterozygous SNP (n = 1,734 total with 20X coverage in EndoC-ßH1 ATAC-seq and H327ac data), allelic imbalance ratios were calculated and plotted. Colored points signify SNPs with significant allelic imbalance (FDR < 10%) in ATAC-seq (red), H3K27ac (blue), or both (purple) datasets. (B) Correlation of human islet eQTL SNP-gene pair direction of effect (z-score) from (Varshney et al., 2017) and their corresponding H3K27ac effect allele bias deviation from 0.5 in EndoC-ßH1. Red points indicate eQTL SNP-gene pairs that are also linked by an EndoC-ßH1 ChIA-PET interaction. Asterisks indicate eQTL SNPs that are also diabetes-associated GWAS SNPs.