MOLECULAR
METABOLISM

Integration of ATAC-seq and RNA-seq identifies human alpha cell and beta cell signature genes

CrossMark

Amanda M. Ackermann [1,2], Zhiping Wang [3], Jonathan Schug [2,4], Ali Naji [5], Klaus H. Kaestner [2,4,*]

ABSTRACT

Objective: Although glucagon-secreting α-cells and insulin-secreting β-cells have opposing functions in regulating plasma glucose levels, the two cell types share a common developmental origin and exhibit overlapping transcriptomes and epigenomes. Notably, destruction of β-cells can stimulate repopulation via transdifferentiation of α-cells, at least in mice, suggesting plasticity between these cell fates. Furthermore, dysfunction of both α- and β-cells contributes to the pathophysiology of type 1 and type 2 diabetes, and β-cell de-differentiation has been proposed to contribute to type 2 diabetes. Our objective was to delineate the molecular properties that maintain islet cell type specification yet allow for cellular plasticity. We hypothesized that correlating cell type-specific transcriptomes with an atlas of open chromatin will identify novel genes and transcriptional regulatory elements such as enhancers involved in α- and β-cell specification and plasticity.
Methods: We sorted human α- and β-cells and performed the "Assay for Transposase-Accessible Chromatin with high throughput sequencing" (ATAC-seq) and mRNA-seq, followed by integrative analysis to identify cell type-selective gene regulatory regions.
Results: We identified numerous transcripts with either α-cell- or β-cell-selective expression and discovered the cell type-selective open chromatin regions that correlate with these gene activation patterns. We confirmed cell type-selective expression on the protein level for two of the top hits from our screen. The "group specific protein" (GC; or vitamin D binding protein) was restricted to α-cells, while CHODL (chondrolectin) immunoreactivity was only present in β-cells. Furthermore, α-cell- and β-cell-selective ATAC-seq peaks were identified to overlap with known binding sites for islet transcription factors, as well as with single nucleotide polymorphisms (SNPs) previously identified as risk loci for type 2 diabetes.
Conclusions: We have determined the genetic landscape of human α- and β-cells based on chromatin accessibility and transcript levels, which allowed for detection of novel α- and β-cell signature genes not previously known to be expressed in islets. Using fine-mapping of open chromatin, we have identified thousands of potential *cis*-regulatory elements that operate in an endocrine cell type-specific fashion.

1. INTRODUCTION

Glucose homeostasis is regulated closely by pancreatic α- and β-cells, which secrete glucagon to raise and insulin to decrease plasma glucose levels, respectively. Despite these distinct functions, the two cell types share a common developmental origin [1] and similar epigenetic regulation of gene expression [2]. Dysfunction of α- and β-cells contributes to the phenotypes of both type 1 and type 2 diabetes

[3,4]; however, the molecular pathophysiological mechanisms by which this occurs are not well understood. Most prior studies investigating the roles of transcriptional regulatory networks under normal and disease conditions have used whole islets, making it difficult to determine which regulatory elements such as promoters or enhancers are specifically active in β-cells versus other islet cell types. This is particularly problematic in studies of human islets, where β-cells make up on average only 54% of all endocrine cells, and can range as low as

[1]Division of Endocrinology and Diabetes, The Children's Hospital of Philadelphia, 3400 Civic Center Boulevard, Philadelphia 19104, PA, USA [2]Institute of Diabetes, Obesity, and Metabolism, The University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia 19104, PA, USA [3]Institute for Biomedical Informatics, Perelman School of Medicine at the University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia 19104, PA, USA [4]Department of Genetics, The University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia 19104, PA, USA [5]Department of Surgery, Perelman School of Medicine at the University of Pennsylvania, 3400 Civic Center Boulevard, Philadelphia 19104, PA, USA

*Corresponding author. University of Pennsylvania, 12-126 Translational Research Center, 3400 Civic Center Blvd, Philadelphia, PA 19104-6145, USA. Tel.: +1 215 898 8759.

E-mails: ackermanna@email.chop.edu (A.M. Ackermann), zhipwang@mail.med.upenn.edu (Z. Wang), jschug@upenn.edu (J. Schug), ali.naji@uphs.upenn.edu (A. Naji), kaestner@mail.med.upenn.edu (K.H. Kaestner).

Abbreviations: ATAC-seq, Assay for Transposase-Accessible Chromatin with high throughput sequencing; FAIRE-seq, Formaldehyde-Assisted Isolation of Regulatory Elements followed by high throughput sequencing; ChIP-seq, Chromatin Immunoprecipitation followed by high throughput sequencing; FACS, fluorescence-activated cell sorting; SNP, single nucleotide polymorphism; DAPI, 4′,6-diamidino-2-phenylindole; GC, group-specific protein; CHODL, chondrolectin; ARX, aristaless related homeobox; GCG, glucagon; DPP4, dipeptidyl-peptidase 4; IRX2, iroquois homeobox 2; MAFA, v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog A; INS, insulin; IGF2, insulin like growth factor 2; NEUROD1, neuronal differentiation 1; SST, somatostatin; PP, pancreatic polypeptide; GHRL, ghrelin

Received December 20, 2015 • Revision received December 30, 2015 • Accepted January 3, 2016 • Available online 11 January 2016
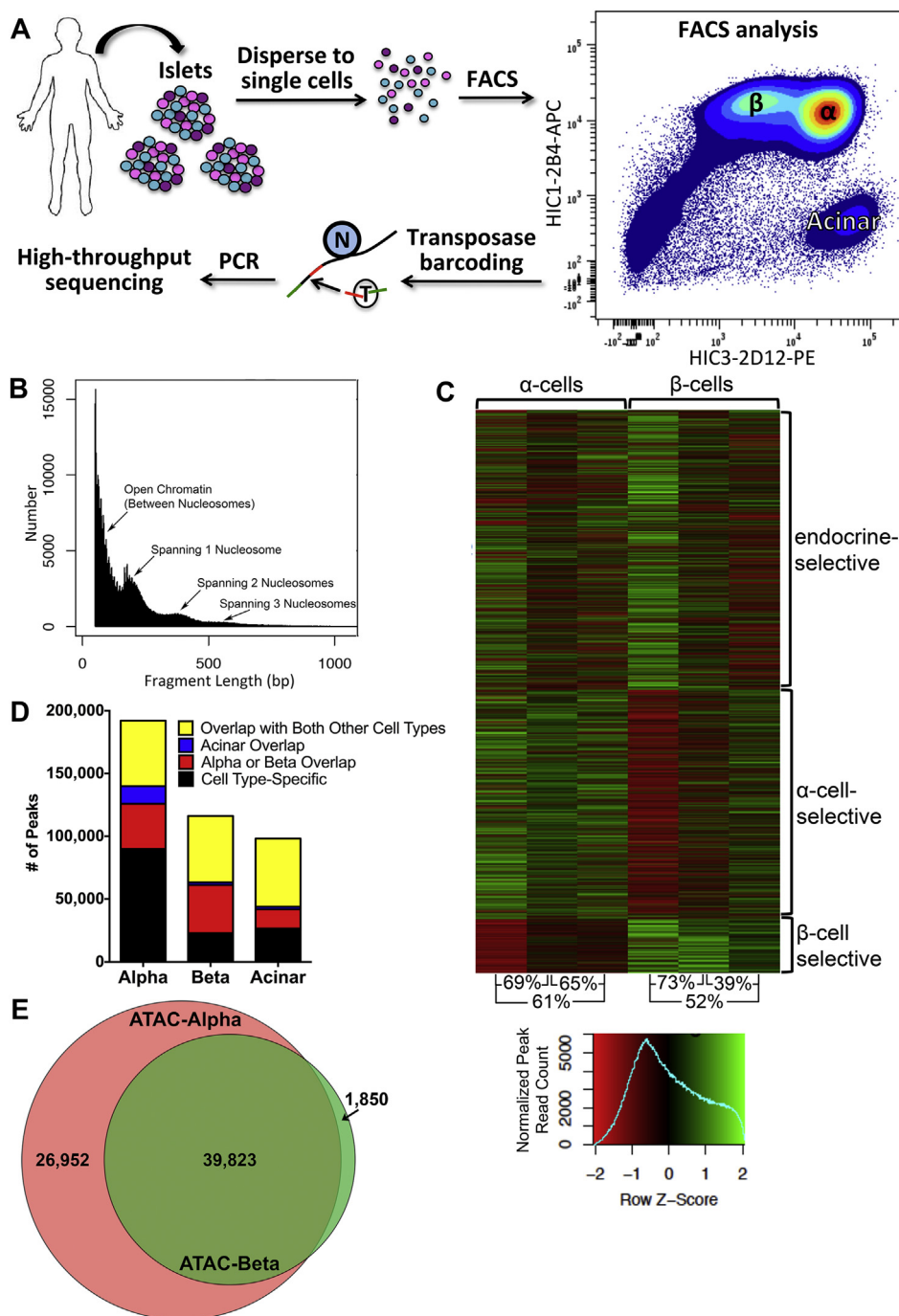
http://dx.doi.org/10.1016/j.molmet.2016.01.002

Figure 1: ATAC-seq results in sorted human α-, β-, and acinar cells. (A) Experimental design. Islets from deceased organ donors were dispersed and FACS sorted into α-, β-, and acinar cell fractions, and processed for ATAC-seq analysis. N: nucleosome; T: transposase. Red and green bars represent PCR/sequencing barcodes. **(B)** Fragment lengths within a representative ATAC-seq library. The small fragments represent sequence reads in open chromatin, while the peak at ~150 bp results from sequence reads that span one nucleosome, and larger peaks represent progressively more compact chromatin. **(C)** Heatmap of ATAC-seq peak data showing clustering of endocrine-selective peaks (present in α- and β-, but not acinar cells), α-cell-selective peaks, and β-cell-selective peaks. Inter-sample correlation is noted at the bottom. **(D)** Number of peaks identified by ATAC-seq in each cell type that are specific to that cell type versus also found in either of the other two cell types investigated. **(E)** Venn diagram of overlap of α-cell-selective and β-cell-selective ATAC-seq peaks, after removal of peaks also found in acinar cells.

28% [5]. Our laboratory previously reported that although the transcriptomes of sorted human α- and β-cells are fairly distinct, their histone methylation marks are more similar than expected [2]. These findings help to explain how various experimental models result in transdifferentiation of α-cells into β-cells, or vice-versa [6—13].

However, it is unclear how human islet cells maintain cell type specification under normal conditions, yet allow for plasticity under conditions of metabolic stress [10]. We hypothesized that correlating cell type-specific transcriptomes with an atlas of open chromatin could identify novel *cis*-regulatory elements involved in these processes.

In this study, we utilized the "Assay for Transposase-Accessible Chromatin with high throughput sequencing" (ATAC-seq) [14,15] to detect open chromatin regions in highly enriched human α- and β-cells. We then correlated these maps of accessible chromatin with mRNA-seq data from sorted human α- and β-cells to identify the *cis*-regulatory elements that may be responsible for the regulation of cell type-specific signature genes (Figure 1A). Notably, we identified many novel α-cell- and β-cell-selective genes.

2. MATERIAL AND METHODS

2.1. Human islets

Human islets from deceased organ donors (Supplemental Table 1) were provided by the Islet Cell Resource Center of the University of Pennsylvania and the Integrated Islet Distribution Program (iidp.coh.org). After shipment, islets were cultured either in CMRL1066 with 0.5% human albumin, 2 mM L-glutamine, 1% heparin, and 1% penicillin/streptomycin or in Prodo PIM(S) media supplemented with 5% PIM(ABS) and 1% PIM(G) at 37 °C with 5% CO_2 prior to fluorescence-activated cell sorting (FACS). FACS was performed as described in [2] using antibodies against cell surface markers as reported in [16].

2.2. ATAC-seq

ATAC-seq was performed as described [15] on three α-cell, three β-cell, and two acinar FACS sorted cell samples (Supplemental Table 1). Reagent volumes were adjusted according to starting cell number. Libraries were generated using the Ad1_noMX and Ad2.1—2.4 barcoded primers from [14] and were amplified for 7—9 total cycles. Libraries were purified with AMPure beads (Agencourt) to remove contaminating primer dimers. Library quality was assessed using the Agilent Bioanalyzer High-Sensitivity DNA kit. Libraries were quantitated using Qubit.

All libraries were sequenced on the Illumina HiSeq 2500 with 50 bp paired-end reads, then on the Illumina NextSeq 500 with 40 bp paired-end reads. Illumina adapters were trimmed by Trimmomatic [17], and the 50 bp reads were shortened by 10 bp with fastx_-trimmer (http://hannonlab.cshl.edu/fastx_toolkit/). All reads for each sample were combined and aligned to hg18 with bowtie2 [18] using default settings. Reads were aligned to hg18 so that data could be compared to relevant published ChIP-seq (Chromatin Immunoprecipitation followed by high throughput sequencing) and FAIRE-seq (Formaldehyde-Assisted Isolation of Regulatory Elements followed by high throughput sequencing) data [2,19]. Greater than 100 million reads were obtained for each library, and reads mapping to mitochondrial DNA were excluded from the analysis together with low quality reads (MAPQ < 10, duplicates and reads in Encode blacklist regions). Between 30 million and 233 million high-quality reads per sample that mapped to genomic DNA were retained. All mapped reads were offset by +4 bp for the +strand and −5 bp for the −strand [14]. Peaks were called for each sample using MACS2 [20] with parameters "-q 0.05 −nomodel −shift 37 −extsize 73", and differential peaks were identified using the MACS2 *bdgdiff* module comparing different cell types from the same donor. Then peaks were merged for the same cell types using Bedtools [21]. Individual peaks separated by <100 bp were joined together. Peak annotation was performed using HOMER [22]. Motif analysis on peak regions was performed by HOMER function *findMotifsGenome.pl*. All sequencing tracks were viewed using the Integrated Genomic Viewer (IGV 2.3.61). All genomics datasets were deposited in GEO under accession number GSE76268.

2.3. mRNA-seq

mRNA-seq was performed on seven α-cell and eight β-cell sorted samples from ten different islet donors (Supplemental Table 1). Results from three of these samples were previously published [2]. For the results presented here, all mRNA-seq libraries were generated using the Illumina TruSeq Stranded Total RNA LT Sample Prep Kit (Cat. #RS-122-2301). Sequencing and computational analysis were performed as described in [2].

2.4. Protein expression

Whole islets or trypsin-dispersed islet cells were fixed in 4% paraformaldehyde, then washed with phosphate-buffered saline (PBS). Dispersed islet cells were attached to positively-charged slides using a CytoSpin. Islets were incubated in blocking solution (PBS with 0.3% Triton and 10% fetal bovine serum [FBS]), and slides in CAS-Block (ThermoFisher), prior to and concurrent with antibodies. Primary antibodies (rabbit anti-GC [Abcam ab65636], rabbit anti-CHODL [Abcam ab134924], mouse anti-glucagon [Abcam ab82270], guinea pig anti-insulin [Invitrogen 180067], rabbit anti-somatostatin [Santa Cruz sc-13099], goat anti-pancreatic polypeptide [PP; Abcam ab77192], and goat anti-ghrelin [Santa Cruz sc-10368]; all 1:100) were incubated overnight at 4 °C, and secondary antibodies (anti-rabbit Cy2 or Cy3, anti-mouse Cy3 or Cy5, anti-goat Cy3, and anti-guinea pig Cy5 [Jackson ImmunoResearch]; all 1:200) were incubated at room temperature for 3 h. DAPI (4′,6-diamidino-2-phenylindole) was added with the mounting media to counterstain DNA. Whole islets were imaged on a Leica TCS SP8 confocal microscope, and dispersed cells were imaged on a Nikon Eclipse 80i widefield microscope.

3. RESULTS

3.1. Mapping α- and β-cell open chromatin with ATAC-seq

To correlate fine-mapped open chromatin regions with gene expression profiles, we followed the experimental paradigm outlined in Figure 1A. Human islets were dispersed into single cell suspension with trypsin, labeled with antibodies, and subjected to FACS as published previously [2]. Sorted cell fractions were subjected to "transposase barcoding", in which the Tn5 transposase integrates sequencing adapters only into open chromatin regions, but is prevented from doing so in genomic regions tightly packed into nucleosomal arrays [15]. All ATAC-seq libraries yielded the expected distribution of fragment lengths, with the majority of fragments being small, representing internucleosome open chromatin, and progressively fewer fragments of larger size which are spanning nucleosomes (Figure 1B). As predicted by prior expression profiling data, the open chromatin structure of the two endocrine cell types were more similar to each other than either was to that of acinar cells (Figure 1C,D). Approximately 78% of peaks identified by ATAC-seq in β-cells were also present in α-cells, and 47% also occurred in acinar cells (Figure 1D). Similarly, 46% of peaks detected in α-cells were also observed in β-cells, while only 34% were also seen in acinar cells. Endocrine-specific peaks were categorized as those found in α- and β- but not acinar cells, and nearly 40,000 such regions were identified (Figure 1E). Interestingly, nearly 27,000 ATAC-seq peak regions were classified as α-cell-specific, whereas only 1,850 β-cell-specific regions were found.

Overall, our ATAC-seq data for sorted α- and β-cells correlated well (60%) with published FAIRE-seq results using whole islets [19] (Figure 2A). However, our ATAC-seq analysis identified many more open chromatin regions and associated genes in sorted α- and β-cells
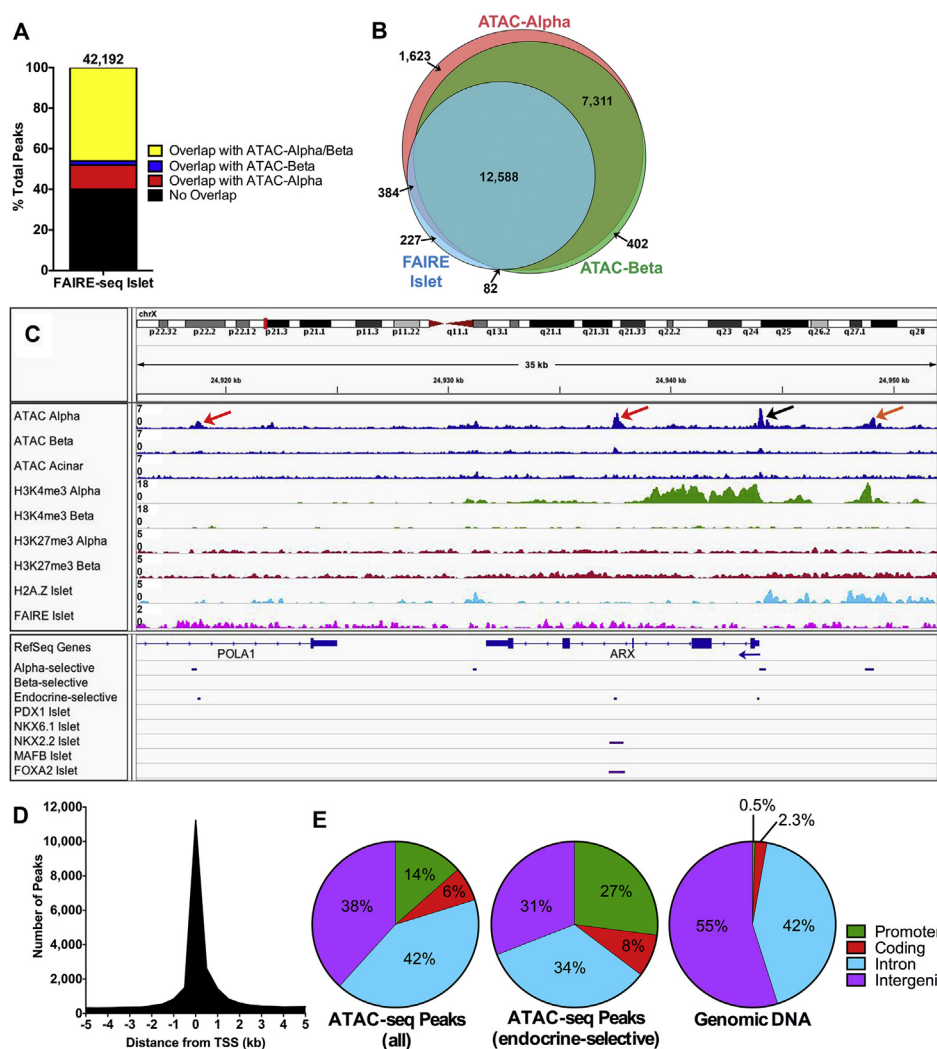
Figure 2: Integration of ATAC-seq data with other genomics datasets. (A) Bar graph of % of overlapping open chromatin regions identified by FAIRE-seq [32] in whole islets versus by ATAC-seq in α- and β-cells (including peaks also found in acinar cells). Total number of FAIRE-seq peaks is noted at top. **(B)** Venn diagram of distinct genes with open chromatin regions in α- and β-cells identified by ATAC-seq (including peaks also found in acinar cells) versus in whole islets identified by FAIRE-seq. **(C)** Sequencing tracks for the *ARX* locus shows distinct α-cell-specific ATAC-seq peaks at the promoter (black arrow), at known intronic and distal enhancers (red arrows), and at a putative 5′ enhancer (orange arrow), none of which were identified by FAIRE-seq. **(D)** Histogram of distance from the nearest transcriptional start site (TSS) for all ATAC-seq peaks within 5 kb of the nearest TSS that were identified in α- and β-cells. Not shown are peaks 5−280 kb from the nearest TSS. **(E)** Proportions of the ATAC-seq peak regions identified in α- and β-cells that represent the various genome annotations, compared to the representation of a given sequence element in the human genome [47].

compared to whole islet FAIRE-seq (Figure 2B), as expected given the much higher sensitivity and resolution of the ATAC-seq methodology. This is easily seen when comparing the ATAC-seq and FAIRE-seq profiles of specific genes, such as the *ARX* locus (Figure 2C). There are strong ATAC-seq peaks in α-cells at the *ARX* promoter and at known enhancers within the third intron and within an intron of a neighboring gene [23] that are not present in β- or acinar cells, while the previously published whole islet FAIRE-seq signals [19] are very broad and do not detect these α-cell-specific open chromatin regions. Furthermore, ATAC-seq identified an α-cell-specific peak approximately 5 kb upstream of the *ARX* promoter that overlapped with α-cell-specific H3K4me3 and whole islet H2A.Z, indicating that this region may function as an enhancer; again, this region was not recognized by whole islet FAIRE-seq [19].

Most ATAC-seq peaks from the α-, β-, and acinar cell samples mapped to within 250 bp of transcriptional start sites (TSS; Figure 2D), marking the accessible chromatin of promoters. In fact, the ATAC-seq dataset was significantly enriched (∼28-fold) for promoter regions compared to the overall abundance of promoters in the genome (Figure 2E). Notably, there was even greater enrichment (∼54-fold) for open promoter regions in the peaks that were specifically identified in α- and β-cells. In addition, many open chromatin regions identified in our analysis were located in introns and intergenic regions, suggestive of enhancers (Figure 2E).

3.2. Integration of ATAC-seq and mRNA-seq results
To determine whether cell type-selective open chromatin regions from the ATAC-seq analysis correlated with cell type-selective gene
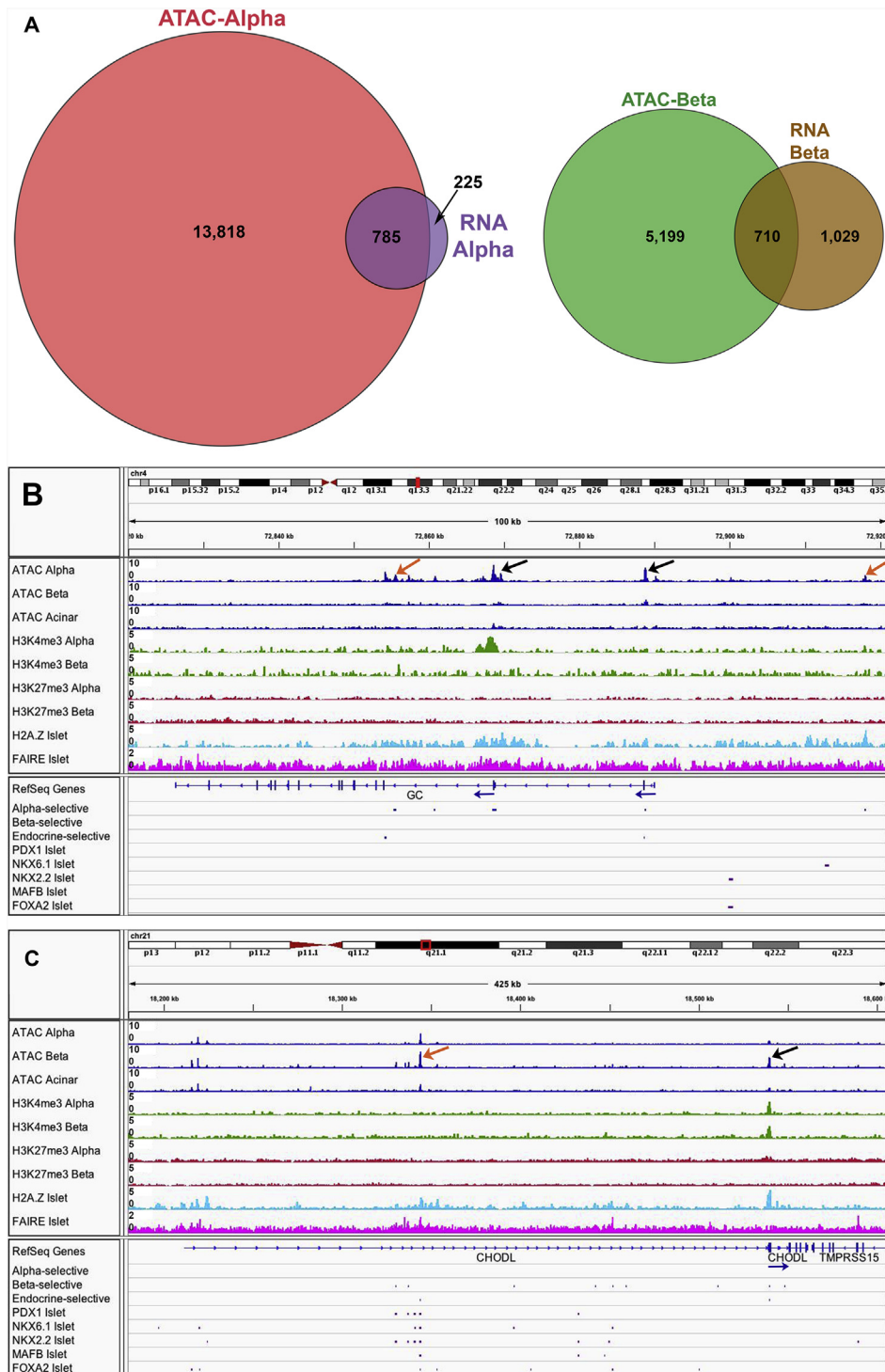
Figure 3: Integration of ATAC-seq data with mRNA-seq, histone marks, FAIRE-seq, and transcription factor binding data for novel α- and β-cell genes. (A) Venn diagrams indicating the proportion of genes that are differentially expressed in α- or β-cells (based on mRNA-seq) that have cell type-selective ATAC-seq peaks (identified only in α- or β-cells, including peaks also identified in acinar cells). **(B)** Pooled sequencing tracks for the *GC* locus from α-cell, β-cell, and acinar cell ATAC-seq data (blue); α-cell and β-cell H3K4me3 (green) and H3K4me27 (red) ChIP-seq data; whole islet H2A.Z ChIP-seq data (light blue); and whole islet FAIRE-seq data (pink). Cell type-selective peaks are identified in blue below the row labeled "RefSeq Genes", followed by islet transcription factor ChIP-seq peaks in purple. Promoter regions are indicated with black arrows, and putative enhancer regions are labeled with orange arrows. **(C)** Pooled sequencing tracks for the *CHODL* gene. Annotation as in (B).

Original article

expression, we integrated our α- and β-cell ATAC-seq data with α- and β-cell mRNA-seq data. Overall, 785 genes that were expressed at significantly higher levels in α- versus β-cells (defined as ≥2-fold difference, with a false discovery rate [FDR] <0.1) had at least one associated α-cell-specific open chromatin region that was not identified in β- or acinar cells (Figure 3A), which accounted for 78% of differentially-expressed α-cell genes. In contrast, only 41% of differentially expressed β-cell genes were similarly identified as having β-cell-specific open chromatin regions. These results suggest that open chromatin may be a better predictor of gene activation in α-cells than in β-cells, perhaps due to inherent differences in gene regulation in these two different cell populations, or possibly due to a higher degree of cellular heterogeneity within the β-cell versus α-cell population. Furthermore, there were many more genes in both α- and β-cells with cell type-selective ATAC-seq peaks that were not differentially expressed in a cell type-selective manner. Only 5% of the α-cell-specific and 12% of the β-cell-specific ATAC-seq peaks mapped to differentially expressed genes (Figure 3A). These results concur with the growing understanding that gene activation depends on multiple regulatory regions, many of which are located far from the gene locus itself. In fact, further peak annotation analysis revealed no enrichment for open promoter regions compared to intronic, intergenic, or coding regions in differentially expressed genes in either cell population (Supplemental Figure 1A—H).

Narrowing the integrated ATAC-seq and mRNA-seq results to the genes that were the most highly and significantly expressed in each cell type (defined as ≥10-fold expression difference between α- and β-cells, with FDR <0.05), we identified 33 α-cell-selective and 35 β-cell-selective transcripts (Table 1). Several of the α-cell-enriched transcripts were expected from the literature (including ARX, GCG, DPP4, and IRX2), and we found that all of these loci had α-cell-selective open chromatin regions. Furthermore, α-cell-selective open promoter regions were specifically noted for ARX (Figure 2C), GCG, and DPP4 (Supplemental Figure 2A). Interestingly, 28 of the α-cell-enriched transcripts had not previously been described to be expressed in islets (Table 1). Of these novel α-cell genes, 24 were marked by α-cell-selective ATAC-seq peaks, with 12 of these having open chromatin identified in promoter regions.

Similarly, the β-cell-selective transcripts included those of MAFA, INS, and IGF2 as expected, but also 22 transcripts that had not previously been identified in islets. Of these novel β-cell genes, 5 had β-cell-specific ATAC-seq peaks, 2 of which had peaks in promoter regions. In contrast to what was observed in α-cells, most of the genes selectively expressed in β-cells have open chromatin regions that were identified in both α- and β- but not acinar cells (i.e. endocrine-specific, Supplemental Figure 2B). Together, these findings suggest that many genes with preferential expression in β-cells nevertheless are in an at least partially open chromatin state in α-cells, which would favor α- to β-cell transdifferentiation over the reverse process. These results thus extend and support the notion of "epigenomic plasticity" in human α-cells made previously based on the mapping of active and repressive histone marks [2].

Table 1 — α- and β-cell-selective transcripts based on mRNA-seq.

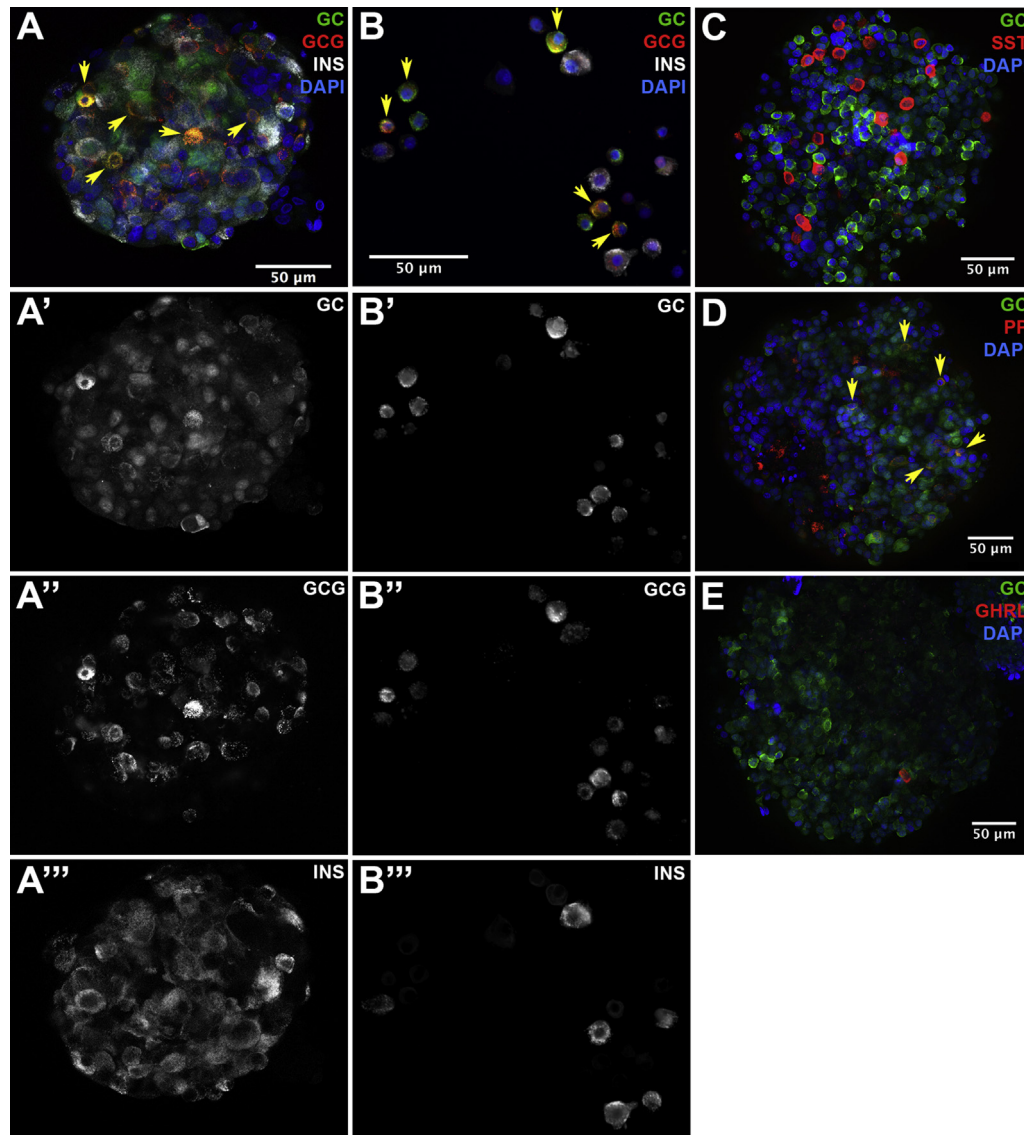| α-Cell signature genes				β-Cell signature genes			
Gene symbol	Any ATAC-seq peak	Peak at promoter	α-cell-specific peak	Gene symbol	Any ATAC-seq peak	Peak at promoter	β-cell-specific peak
APOH				ACPP			
ARX	X	X	X	ADCYAP1	X	X	
BVES	X	X	X	ANXA8			
CARTPT	X	X		ASB9			
CRH	X	X		ASCL1	X		
CRYBA2				BMP5			
DPP4	X	X	X	CAPN13			
F10	X	X	X	CHODL	X	X	X
FAM83B	X	X	X	CST2			
FAP	X	X		CXCL5			
GC	X	X	X	DLK1			
GCG	X	X	X	ESR1			
GJA3	X	X	X	GPM6A	X		
IGFBP2	X	X		GRIN2A			
IRX2	X	X		IGF2			
LOXL4	X	X		IGSF11			
LPAR1	X	X		INS			
MBOAT4	X	X	X	KRTAP4-7			
MUC13	X	X		LAPTM5			
MYO10	X	X		LRFN2			
NPNT				LRRTM3	X		
POPDC3	X	X	X	MAFA			
PTPRT	X	X		P2RY1	X		
SERPINA1	X	X	X	PCDH7	X		
SERTM1	X	X	X	PLCH2			
SPOCK3	X	X		PPAPDC1A			
STK32B	X	X	X	PTGS2			
SYNDIG	X	X		RGS16			
SYTL5				SIX3			
TM4SF4	X	X	X	SLC27A6	X		X
TMEM236				STEAP3	X	X	
TMEM45B	X	X		SULF2			
TTR	X	X	X	TFCP2L1	X	X	
				TGFBR3	X	X	
				UNC5D	X		

Figure 4: Immunofluorescent labeling for the novel α-cell-selective protein GC. Whole mount immunofluorescent labeling of human islets **(A, C-E)** and dispersed human islet cells **(B)** for GC and the islet hormones as indicated in the figure labels. Yellow arrowheads indicate cells expressing GC and islet hormones. GC: group-specific component; GCG: glucagon; INS: insulin; SST: somatostatin; PP: pancreatic polypeptide; GHRL: ghrelin; DAPI: 4′,6-diamidino-2-phenylindole.

3.3. Integration of ATAC-seq results with additional epigenetic marks

To distinguish which open chromatin regions in the α- and β-cell-selective genes may represent gene regulatory regions, we integrated our ATAC-seq data with previously published α- and β-cell ChIP-seq data for H3K4me3 (an activating histone mark) and H3K27me3 (a repressive histone mark) [2], as well as with whole islet H2A.Z ChIP-seq and FAIRE-seq data [19]. We first verified that known promoter and enhancer regions for well-established cell type-specific genes were identified within our dataset including those at the *ARX* [23], *DPP4*, and *MAFA* loci (Figure 2C, Supplemental Figure 2A,B), as well as for well-known pan-endocrine expressed genes such as *NEUROD1* (Supplemental Figure 2C). Importantly, the genes expressed exclusively in α-cells exhibited α-cell-selective open promoter regions associated with α-cell-specific H3K4me3 marks. In contrast, as mentioned above, the genes expressed exclusively in β-cells exhibited

open promoter regions in both α- and β-cells, and were associated with bivalent H3K4me3 (a mark of active promoters) and H3K27me3 (a mark of a repressed chromatin state) in α-cells. Thus, as has been suggested previously [2], our data indicate that many β-cell signature genes are bivalently marked and thus poised for activation in α-cells, which correlates with the propensity of α-cells to transdifferentiate into β-like cells under conditions of extreme metabolic stress, such as complete induced β-cell ablation [6].

Next, we identified cell type-selective open chromatin regions at the promoters of the novel α-cell signature gene *GC* (Figure 3B) and of the novel β-cell signature gene *CHODL* (Figure 3C), as examples. In addition, α-cell-selective ATAC-seq peaks were also present within an intron of *GC*, as well as ∼28 kb upstream of the *GC* promoter, associated with H2A.Z, suggesting that this region may function as an α-cell-specific enhancer (Figure 3B). Similarly, multiple β-cell-selective ATAC-seq peaks are present 5′ upstream of and within an intron of
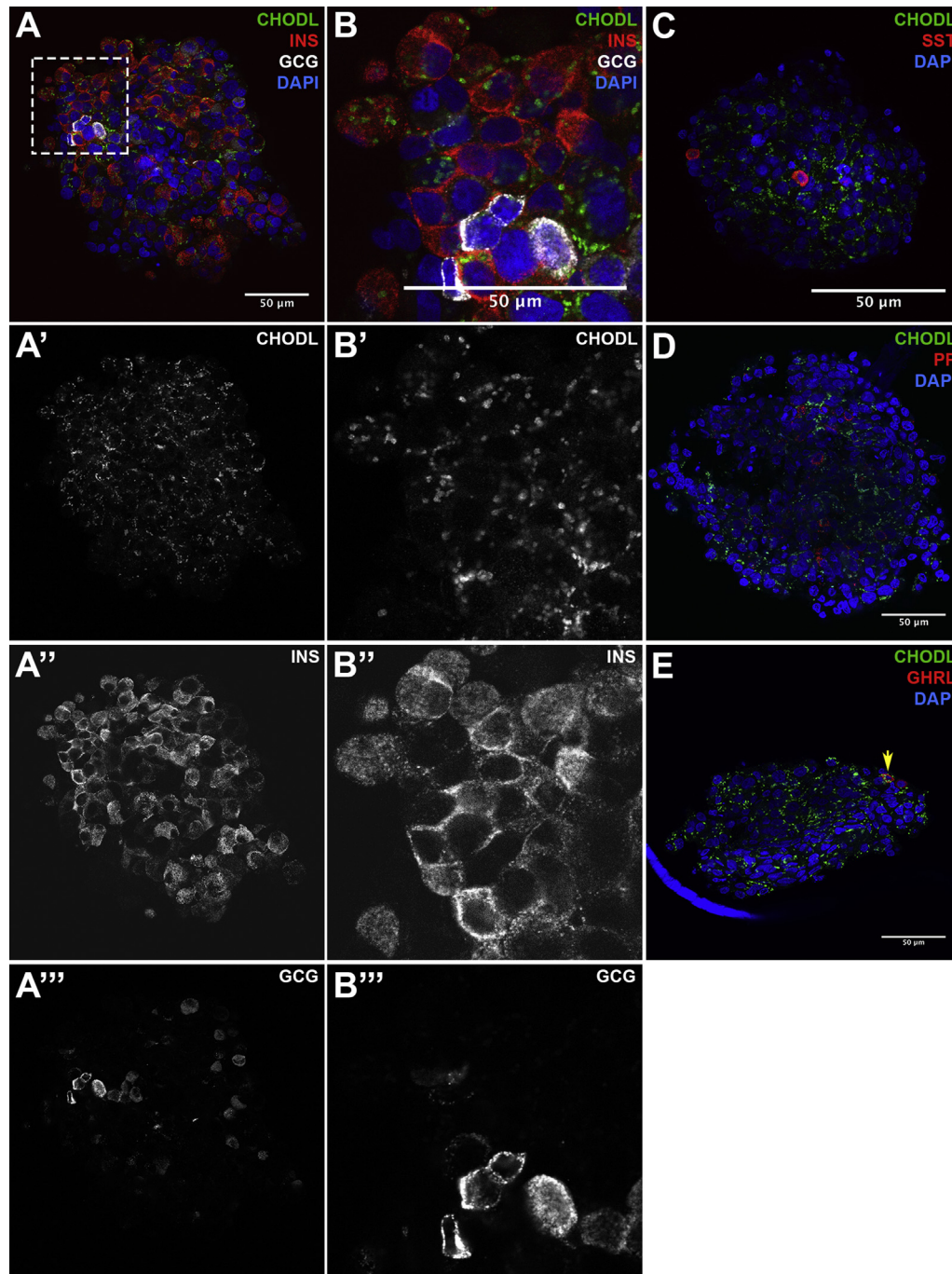
Figure 5: Immunofluorescent labeling for the novel β-cell-selective protein CHODL. (A—E) Whole mount immunofluorescent labeling of human islets for CHODL and the islet hormones as indicated in the figure labels. The boxed region in **(A)** is magnified in **(B)**. CHODL: chondrolectin; GCG: glucagon; INS: insulin; SST: somatostatin; PP: pancreatic polypeptide; GHRL: ghrelin; DAPI: 4′,6-diamidino-2-phenylindole.

CHODL, many of which overlap with β-cell-specific (PDX1, NKX6.1) and islet (NKX2.2, MAFB, FOXA2) transcription factor binding sites (Figure 3C).

3.4. Confirmation of cell type-specific protein expression of novel signature genes

Next, we wanted to determine the cell type-selective expression of our newly identified α- and β-cell signature genes extended to the protein level. For most of these genes no antibodies are commercially available

at present; however, we were able to test protein expression of two genes by immunolabeling. We stained both whole human islets and dispersed islet cells with an antibody against the GC protein, and detected strong GC immunoreactivity in α-cells (Figure 4A,B), validating our RNA-seq analysis. GC protein was also present in a small subset of pancreatic polypeptide (PP)-expressing cells, but not in insulin, somatostatin, or ghrelin expressing cells (Figure 4C—E). GC, or "group-specific component", is more commonly known as vitamin D binding protein (DBP). Its primary role is to bind and transport vitamin D

to its receptor in the nucleus, which then transcriptionally activates target genes. There is currently no known role for the vitamin D receptor in α-cells, but vitamin D deficiency and metabolism have been associated with type 1 and type 2 diabetes [24]. Furthermore, common *GC* non-coding variants have been associated with increased risk of gestational diabetes mellitus [25], as well as altered fasting insulin levels in normoglycemic individuals, although this was attributed primarily to changes in insulin sensitivity [26]. Thus, GC activity in α-cells may influence β-cell function via indirect intercellular or paracrine interactions.

β-cell-specific expression of the CHODL (chondrolectin) protein was also confirmed in whole human islets (Figure 5). CHODL currently has no recognized role in β-cells, but it is a member of the C-type lectin superfamily of proteins, which contain a single transmembrane domain and a calcium-dependent carbohydrate binding domain [27]. Thus, it can bind to carbohydrate moieties on glycosylated proteins. CHODL and other C-type lectin family members have been implicated in intercellular adhesion, extracellular matrix interactions, and intracellular protein transport.

3.5. Transcription factor binding motifs in α- and β-cell-selective open chromatin regions

Because many of the genes with annotated α- or β-cell-selective ATAC-seq peaks were not differentially expressed (Figure 3A), we sought to determine whether the cell type-selective peak regions enriched for particular transcription factor binding sites. Not surprisingly, α-cell open chromatin regions were enriched for binding sites of transcription factors known to play important roles in α-cells, including the FOX factors, ISL1, and MAFB (Table 2). Other enriched binding sites included those for FRA1, TFAP4, RFX5, CTCF, ATF2, STAT1, and GATA3. Expression of all of these transcription factors was confirmed in α-cells by mRNA-seq (data not shown).

Binding site motifs for FRA1, TFAP4, and the FOX factors were also enriched in β-cell open chromatin regions, as was the motif for SMAD2, which is known to mediate TGF-β signaling and affect pancreatic endocrine cell development as well as mature β-cell function [28,29] (Table 2). Binding sites for other transcription factors that do not have known roles in β-cells were also enriched. Again, expression of these transcription factors in human β-cells was confirmed by mRNA-seq (data not shown).

3.6. Diabetes risk loci within open chromatin regions identified by ATAC-seq

We also sought to determine whether α- or β-cell-selective open chromatin regions in our ATAC-seq dataset included single nucleotide polymorphism (SNP) loci previously found in genome-wide association studies (GWAS) to be associated with increased risk of type 1 or type 2 diabetes mellitus [30,31]. Notably, two risk loci for type 2 diabetes were identified in endocrine-specific open chromatin by ATAC-seq: rs7732130 and rs7903146, which are located within introns of *TCF7L2* and *ZBED3-AS1*, respectively (Table 3). Eleven additional type 2 diabetes risk loci were located near cell type-specific open chromatin regions, of which 6 were specific to α-cells. These findings suggest that multiple type 2 diabetes risk loci might be associated with α-cell dysfunction.

4. DISCUSSION

In this study, we utilized the sensitive ATAC-seq technology to determine regions of open chromatin in highly enriched populations of human α-, β-, and acinar cells. This is the first evaluation of open

Table 2 — Transcription factor binding motifs enriched in α- or β-cell-specific open chromatin regions.

Factor	Binding Motif	% Regions	P-value
α-Cells			
FRA1/FOSL1		21.4%	10^{-1145}
FOX		19.7%	10^{-432}
RFX5		5%	10^{-386}
CTCF		3.9%	10^{-349}
AP4/TFAP4		41.2%	10^{-235}
RFX4		42.5%	10^{-181}
ATF2		5.4%	10^{-139}
ISL1		26%	10^{-128}
STAT1		40.1%	10^{-121}
GATA3		12.1%	10^{-111}
E2F6		32.6%	10^{-97}
IRF4		25.8%	10^{-86}
MAFK		6.9%	10^{-84}
MAFB		76.9%	10^{-62}
HAND1		13.6%	10^{-44}
β-Cells			
FRA1/FOSL1		47.2%	10^{-264}
LHX2		51.4%	10^{-92}
AP4/TFAP4		23%	10^{-32}
CUX1		22.4%	10^{-30}
FOX		18.6%	10^{-29}
PIT1/POU1F1		6.3%	10^{-19}
JUN-FOS		4.2%	10^{-18}
SMAD2		7%	10^{-14}
MEF2C		3.2%	10^{-12}

chromatin in distinct islet cell subtypes. Previous studies have performed FAIRE-seq on whole human islets [19,32], but because of the multiple cell types included in such analyses and the variability of islet composition (28–70% β-cells, 10–60% α-cells, 1–20% δ-cells, as well as up to 25% each of contaminating duct and acinar cells [5,33]),

Table 3 — Diabetes risk-associated SNPs near α- or β-cell-specific ATAC-seq peaks.

SNP	Associated genes	Location in gene	ATAC-seq peak cell type-specificity	Preferential mRNA expression	TF binding sites [19]
Type 2 diabetes mellitus					
rs7732130[a]	TCF7L2	Intron	Endocrine cells	Same	NKX6.1, NKX2.2, FOXA2, MAFB
rs7732130[a]	ZBED3-AS1	Intron	Endocrine cells	β-cells	NKX2.2
rs1169288	HNF1A	5′ UTR	Endocrine cells	Same	NKX2.2, FOXA2
rs1800574	HNF1A	5′ UTR	Endocrine cells	Same	NKX2.2, FOXA2
rs5215	KCNJ11	5′ UTR	Endocrine cells	β-cells	FOXA2
rs17066842	MC4R	n/a (5′)	Endocrine cells	Same	NKX2.2
rs13266634	SLC30A8	Intron	Endocrine cells	Same	NKX2.2, FOXA2
rs11708067[a]	ADCY5	Intron	α-cells	Same	none
rs10811660[a]	CDKN2A-CDKN2B	n/a (5′)	α-cells	Same	MAFB, FOXA2
rs10757283	CDKN2A-CDKN2B	n/a (5′)	α-cells	Same	MAFB, FOXA2
rs2237895	KCNQ1	Intron	α-cells	Same	NKX2.2, FOXA2
rs74046911	KCNQ1	Intron	α-cells	Same	NKX2.2, FOXA2
rs458069	KCNQ1	Intron	α-cells	Same	NKX2.2, FOXA2
Type 1 diabetes mellitus					
rs11755527	BACH2	Intron	Endocrine cells	Same	None
rs10517086	none	n/a	Endocrine cells	n/a	None

TF: transcription factor; UTR: untranslated region; n/a: not applicable.
[a] Located within ATAC-seq peak region. Other SNPs are located within 1,000 bp of ATAC-seq peak region and TF binding sites.

these prior studies were not designed to detect cell type-specific chromatin states. ATAC-seq also provides several advantages over the older techniques such as FAIRE-seq or DNase-seq for the identification of open chromatin regions. First, ATAC-seq requires many fewer cells than the two older techniques, making subtype analysis possible, and second, ATAC-seq results have a much higher signal-to-noise ratio than FAIRE-seq, as documented above, and map open chromatin regions much more precisely [34].

Using ATAC-seq for purified islet cell subtypes, we identified cell type-specific open chromatin regions that were not apparent from the whole islet FAIRE-seq data. By integrating ATAC-seq data with mRNA-seq data from sorted human α- and β-cells, we were also able to locate open chromatin regions in genes that previously were not known to be expressed in these cell types. Importantly, we defined novel sets of both α- and β-cell 'signature genes', i.e. genes with high expression in one cell type or the other, but not both. These data will provide useful guideposts for the efforts in the field to derive β-cells from human embryonic stem cells or induced pluripotent stem cells, as they allow for an easy distinction between intermediate endocrine cells and mature β-cells.

A limitation of this study is reliance on obtaining cells for analysis from human deceased organ donors, among which there is unknown but likely significant inherent genetic, epigenetic, and physiological variation. Furthermore, there is significant variability among the samples due to the fact that each islet sample must be sorted and undergo the transposase reaction individually. In an attempt to limit the effects of such variation, we performed differential peak calling of α- versus β-cells separately for each donor, then pooled the results and discarded any peaks that were differentially represented in the other cell type from another donor, in order to arrive at cell type-selective peaks. However, it is possible that such criteria resulted in inclusion of peaks that were only significant in one sample and thus perhaps not truly reflective of canonical α- or β-cell-specific open chromatin regions. This could contribute to the fact that a large number of cell type-selective ATAC-seq peaks were identified in genes that were not differentially expressed in these cell types. Furthermore, because different donor samples were used for ATAC-seq and mRNA-seq, it is not surprising that there is incomplete overlap of these two datasets. However, as previously mentioned, that result could also reflect the

fact that small regions of open chromatin alone are not sufficient for gene activation.

Similar to our previous findings in murine liver [35], this study indicates that nucleosome structure in mammalian cell types is remarkably consistent. In that study, nucleosome positions were mapped genome-wide through micrococcal nuclease digestion followed by sequencing of the remaining, nucleosome-protected DNA. If nucleosome position were highly variable from cell to cell, then we would have expected to find few if any nucleosome peaks within a specific cell population, but instead we found strong, well-spaced peaks in the nucleosome maps of hepatocytes [35], indicating that these positions are similar between at least the majority of hepatocytes. Likewise, in the current study, the fact that ATAC-seq mapping of open chromatin regions in human α- and β-cells produced clearly defined peaks indicates that the flanking nucleosomes are fixed in place, or nearly so.

The current and previous [2] results indicate that many genes in α-cells are "poised" to be activated, and that many of these "poised" genes are β-cell signature genes. This is reflected in the ATAC-seq dataset by the finding of many more open chromatin regions in α-cells versus β-cells. We also performed de novo transcription factor binding site motif analysis using α- and β-cell-selective open chromatin regions. In addition to the expected enrichment for binding sites of transcription factors already known to play important roles in α- and β-cells, we also identified several additional transcription factors whose binding site motifs were enriched in our dataset. This exploration uncovered novel putative transcriptional regulators within human α- and β-cells.

For example, we found the CTCF consensus motif significantly enriched in open chromatin regions of α-cells (Supplemental Figure 2D). CTCF has been shown to repress Pax6 expression in murine α-cells, and global constitutive over-expression of CTCF in a transgenic mouse model significantly impaired α-cell differentiation and glucagon expression [36], similar to what was observed in Pax6 deletion mouse models [37—40]. However, only α-cells were affected in the CTCF overexpression model, while in the Pax6 deletion models, both α- and β-cell populations were affected, although β-cells to a lesser degree. These results suggest that CTCF regulates PAX6 expression only in α-cells but not β-cells, despite PAX6 being expressed to similar levels in both cell types, and despite α- and β-

cells having similar open chromatin maps at the *PAX6* locus (Supplemental Figure 2D). Because CTCF functions as a general transcriptional insulator [41], it may be responsible for binding to open chromatin regions in α-cells and preventing or limiting gene activation, which may be one of the mechanisms by which α-cells are able to maintain their cell-specific gene expression pattern despite displaying an open chromatin at many β-cell signature loci.

STAT1, another transcription factor whose binding site motif was enriched in α-cell-selective open chromatin regions, has been implicated in antiviral intracellular responses in α-cells [42], but its specific targets are still unknown. Furthermore, FRA1 and TFAP4 binding sites were highly enriched in both α- and β-cells, suggesting that these factors may function in both cell types. FRA1 has been implicated in mediating TGF-β signaling with SMAD2 in other cell types [43]. TFAP4 has been shown to regulate proliferation in other cell types by directly repressing transcription of the cell cycle inhibitors p16 and p21 [44], both of which are important regulators of β-cell replication [45,46]. Thus, further study of these transcription factors in pancreatic islets is warranted.

In summary, we provide here a novel resource for identifying open chromatin regions in human α-, β-, and acinar cells. This dataset builds upon the growing body of genome-wide epigenetic studies performed in purified human α- and β-cells. We have shown that integrating ATAC-seq data with other epigenetic information enhances interpretation of computational results, and we anticipate that this ATAC-seq dataset will be useful for integration with future genomic analyses of human islets. Such studies are important for understanding the tightly regulated gene expression networks that make and sustain functioning α- and β-cells, determining how these networks become dysregulated in disease states such as diabetes, and identifying ways in which these networks can be manipulated for diabetes therapy.

5. CONCLUSIONS

We present the first analysis of open chromatin in purified human α- and β-cells using highly sensitive ATAC-seq. This technique allowed for precise mapping of potential α- and β-cell-specific gene regulatory regions. Integration with human α- and β-cell transcriptomes led to the identification of novel signature genes for these two cell types. Further mining of the open chromatin regions defined by ATAC-seq revealed overlap with several diabetes risk SNPs, as well as enrichment for novel transcription factors that may play important roles in α- and β-cells. This study provides an important dataset and multiple new avenues of investigation for the future.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST

None declared.

REFERENCES

[1] Gradwohl, G., Dierich, A., LeMeur, M., Guillemot, F., 2000. Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. Proceedings of National Academy of Sciences U S A 97(4):1607—1611.

[2] Bramswig, N.C., Everett, L.J., Schug, J., Dorrell, C., Liu, C., Luo, Y., et al., 2013. Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. Journal of Clinical Investigation 123(3):1275—1284.

[3] Rickels, M.R., Fuller, C., Dalton-Bakes, C., Markmann, E., Palanjian, M., Cullison, K., et al., 2015. Restoration of glucose counterregulation by islet transplantation in long-standing type 1 diabetes. Diabetes 64(5):1713—1718.

[4] Unger, R.H., Cherrington, A.D., 2012. Glucagonocentric restructuring of diabetes: a pathophysiologic and therapeutic makeover. Journal of Clinical Investigation 122(1):4—12.

[5] Brissova, M., Fowler, M.J., Nicholson, W.E., Chu, A., Hirshberg, B., Harlan, D.M., et al., 2005. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. Journal of Histochemistry & Cytochemistry 53(9):1087—1097.

[6] Thorel, F., Nepote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., et al., 2010. Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature 464(7292):1149—1154.

[7] Chung, C.H., Hao, E., Piran, R., Keinan, E., Levine, F., 2010. Pancreatic beta-cell neogenesis by direct conversion from mature alpha-cells. Stem Cells 28(9):1630—1638.

[8] Collombat, P., Xu, X., Ravassard, P., Sosa-Pineda, B., Dussaud, S., Billestrup, N., et al., 2009. The ectopic expression of Pax4 in the mouse pancreas converts progenitor cells into alpha and subsequently beta cells. Cell 138(3):449—462.

[9] Spijker, H.S., Ravelli, R.B., Mommaas-Kienhuis, A.M., van Apeldoorn, A.A., Engelse, M.A., Zaldumbide, A., et al., 2013. Conversion of mature human beta-cells into glucagon-producing alpha-cells. Diabetes 62(7):2471—2480.

[10] Talchai, C., Xuan, S., Lin, H.V., Sussel, L., Accili, D., 2012. Pancreatic beta cell dedifferentiation as a mechanism of diabetic beta cell failure. Cell 150(6): 1223—1234.

[11] Dhawan, S., Georgia, S., Tschen, S.I., Fan, G., Bhushan, A., 2011. Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. Developmental Cell 20(4):419—429.

[12] Courtney, M., Gjernes, E., Druelle, N., Ravaud, C., Vieira, A., Ben-Othman, N., et al., 2013. The inactivation of Arx in pancreatic alpha-cells triggers their neogenesis and conversion into functional beta-like cells. PLoS Genetics 9(10): e1003934.

[13] Gao, T., McKenna, B., Li, C., Reichert, M., Nguyen, J., Singh, T., et al., 2014. Pdx1 maintains beta cell identity and function by repressing an alpha cell program. Cell Metabolism 19(2):259—271.

[14] Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Methods 10(12):1213—1218.

[15] Buenrostro, J.D., Wu, B., Chang, H.Y., Greenleaf, W.J., 2015. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Current Protocols in Molecular Biology 109, 21.29.1-21.29.9.

[16] Dorrell, C., Abraham, S.L., Lanxon-Cookson, K.M., Canaday, P.S., Streeter, P.R., Grompe, M., 2008. Isolation of major pancreatic cell types and long-term culture-initiating cells using novel human surface markers. Stem Cell Research 1(3):183—194.

[17] Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30(15):2114—2120.

[18] Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26(5):589—595.

[19] Pasquali, L., Gaulton, K.J., Rodriguez-Segui, S.A., Mularoni, L., Miguel-Escalada, I., Akerman, I., et al., 2014. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. Nature Genetics 46(2): 136—143.

[20] Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., et al., 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biology 9(9): R137.

[21] Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841—842.

[22] Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., et al., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular Cell 38(4):576—589.

[23] Liu, J., Hunter, C.S., Du, A., Ediger, B., Walp, E., Murray, J., et al., 2011. Islet-1 regulates Arx transcription during pancreatic islet alpha-cell development. Journal of Biological Chemistry 286(17):15352—15360.

[24] Mathieu, C., 2015. Vitamin D and diabetes: where do we stand? Diabetes Research and Clinical Practice 108(2):201—209.

[25] Wang, Y., Wang, O., Li, W., Ma, L., Ping, F., Chen, L., et al., 2015. Variants in vitamin D binding protein gene are associated with gestational diabetes mellitus. Medicine 94(40):e1693.

[26] Hirai, M., Suzuki, S., Hinokio, Y., Hirai, A., Chiba, M., Akai, H., et al., 2000. Variations in vitamin D-binding protein (group-specific component protein) are associated with fasting plasma insulin levels in Japanese with normal glucose tolerance. Journal of Clinical Endocrinology & Metabolism 85(5):1951—1953.

[27] Zelensky, A.N., Gready, J.E., 2005. The C-type lectin-like domain superfamily. FEBS Journal 272(24):6179—6217.

[28] Nomura, M., Zhu, H.L., Wang, L., Morinaga, H., Takayanagi, R., Teramoto, N., 2014. SMAD2 disruption in mouse pancreatic beta cells leads to islet hyperplasia and impaired insulin secretion due to the attenuation of ATP-sensitive K+ channel activity. Diabetologia 57(1):157—166.

[29] El-Gohary, Y., Talachan, S., Guo, P., Welsh, C., Wiersch, J., Prasadan, K., et al., 2013. Smad signaling pathways regulate pancreatic endocrine development. Developmental Biology 378(2):83—93.

[30] Gaulton, K.J., Ferreira, T., Lee, Y., Raimondo, A., Magi, R., Reschen, M.E., et al., 2015. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. Nature Genetics 47(12):1415—1425.

[31] Barrett, J.C., Clayton, D.G., Concannon, P., Akolkar, B., Cooper, J.D., Erlich, H.A., et al., 2009. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nature Genetics 41(6):703—707.

[32] Gaulton, K.J., Nammo, T., Pasquali, L., Simon, J.M., Giresi, P.G., Fogarty, M.P., et al., 2010. A map of open chromatin in human pancreatic islets. Nature Genetics 42(3):255—259.

[33] Pisania, A., Weir, G.C., O'Neil, J.J., Omer, A., Tchipashvili, V., Lei, J., et al., 2010. Quantitative analysis of cell composition and purity of human pancreatic islet preparations. Lab Investigation 90(11):1661—1675.

[34] Tsompana, M., Buck, M.J., 2014. Chromatin accessibility: a window into the genome. Epigenetics & Chromatin 7(1):33.

[35] Li, Z., Schug, J., Tuteja, G., White, P., Kaestner, K.H., 2011. The nucleosome map of the mammalian liver. Nature Structural & Molecular Biology 18(6): 742—746.

[36] Tsui, S., Gao, J., Wang, C., Lu, L., 2012. CTCF mediates effect of insulin on glucagon expression. Experimental Cell Research 318(8):887—895.

[37] Ahmad, Z., Rafeeq, M., Collombat, P., Mansouri, A., 2015. Pax6 inactivation in the adult pancreas reveals ghrelin as endocrine cell maturation marker. PLoS One 10(12):e0144597.

[38] Hart, A.W., Mella, S., Mendrychowski, J., van Heyningen, V., Kleinjan, D.A., 2013. The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. PLoS One 8(1):e54173.

[39] Sander, M., Neubuser, A., Kalamaras, J., Ee, H.C., Martin, G.R., German, M.S., 1997. Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes & Development 11(13):1662—1673.

[40] St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., Gruss, P., 1997. Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. Nature 387(6631):406—409.

[41] Ong, C.T., Corces, V.G., 2014. CTCF: an architectural protein bridging genome topology and function. Nature Reviews Genetics 15(4):234—246.

[42] Marroqui, L., Lopes, M., dos Santos, R.S., Grieco, F.A., Roivainen, M., Richardson, S.J., et al., 2015. Differential cell autonomous responses determine the outcome of coxsackievirus infections in murine pancreatic alpha and beta cells 4:e06990.

[43] Sundqvist, A., Zieba, A., Vasilaki, E., Herrera Hidalgo, C., Soderberg, O., Koinuma, D., et al., 2013. Specific interactions between Smad proteins and AP-1 components determine TGFbeta-induced breast cancer cell invasion. Oncogene 32(31):3606—3615.

[44] Jackstadt, R., Jung, P., Hermeking, H., 2013. AP4 directly downregulates p16 and p21 to suppress senescence and mediate transformation. Cell Death & Disease 4:e775.

[45] Salas, E., Rabhi, N., Froguel, P., Annicotte, J.S., 2014. Role of Ink4a/Arf locus in beta cell mass expansion under physiological and pathological conditions. Journal of Diabetes Research 2014:873679.

[46] Blandino-Rosano, M., Alejandro, E.U., Sathyamurthy, A., Scheys, J.O., Gregg, B., Chen, A.Y., et al., 2012. Enhanced beta cell proliferation in mice overexpressing a constitutively active form of Akt and one allele of p21Cip. Diabetologia 55(5):1380—1389.

[47] Alexander, R.P., Fang, G., Rozowsky, J., Snyder, M., Gerstein, M.B., 2010. Annotating non-coding regions of the genome. Nature Reviews Genetics 11(8): 559—571.