Original article

MOLECULAR METABOLISM

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

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.

Keywords Islet; Alpha cell; Beta cell; Diabetes; Epigenetics; Open chromatin

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

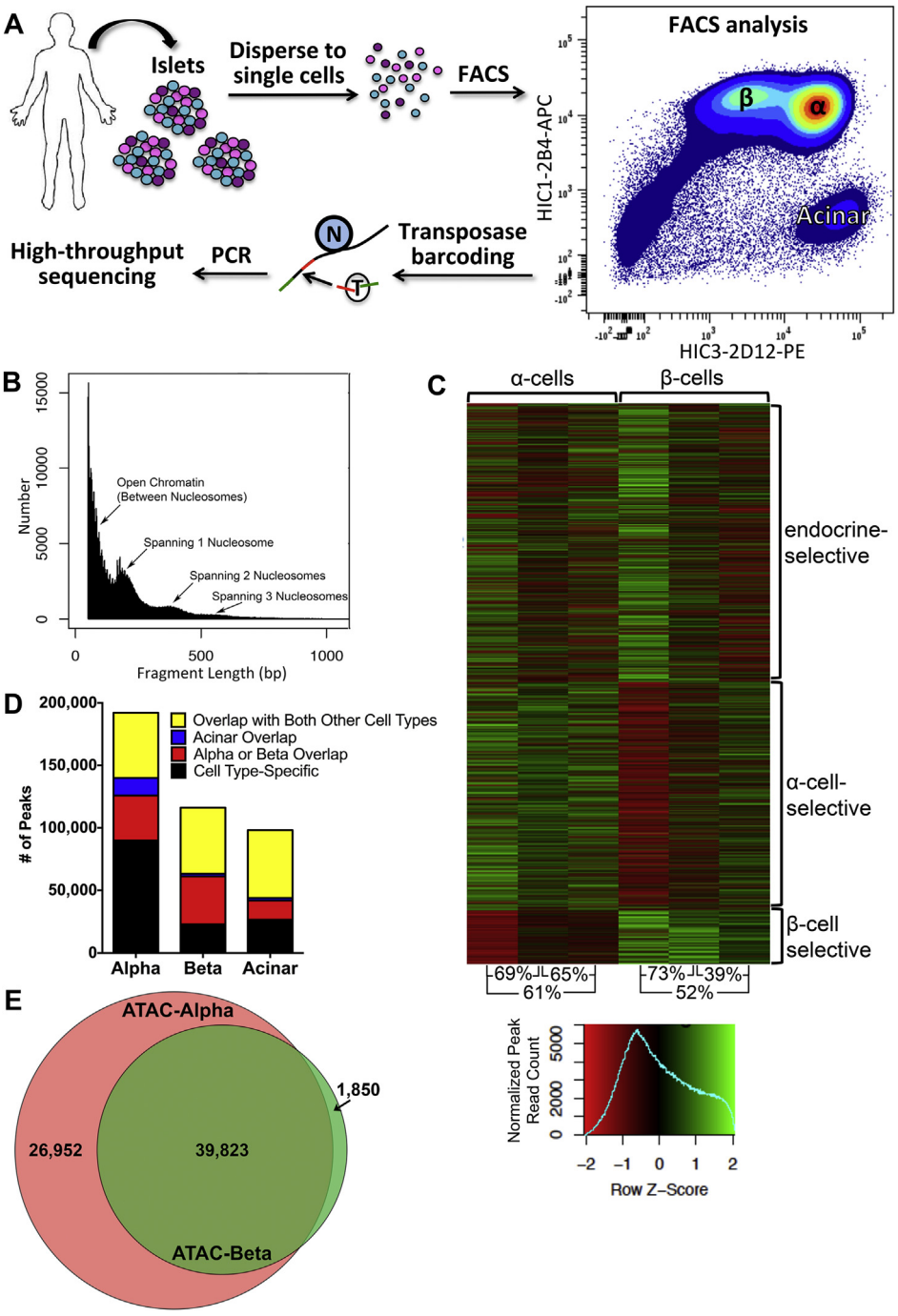

Figure 1: ATAC-seq results in sorted human α-, β-, and acinar cells. (A) Experimental design. Islets from deceased organ donors were dispersed and FACS sorted into α-, β-, and acinar cell fractions, and processed for ATAC-seq analysis. N: nucleosome; T: transposase. Red and green bars represent PCR/sequencing 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.

