A Network of microRNAs Acts to Promote Cell Cycle Exit and Differentiation of Human Pancreatic Endocrine Cells
A Network of microRNAs Acts to Promote Cell Cycle Exit and Differentiation of Human Pancreatic Endocrine Cells

Wen Jin,1 Francesca Mulas,1 Bjoern Gaertner,1 Yinghui Sui,1 Jinzhao Wang,1 Ileana Matta,1 Chun Zeng,1 Nicholas Vinckier,1 Allen Wang,1 Kim-Vy Nguyen-Ngoc,1 Joshua Chiou,1 Klaus H. Kaestner,2 Kelly A. Frazer,3 Andrea C. Carrano,1 Hung-Ping Shih,4 and Maike Sander1,5,*

SUMMARY

Pancreatic endocrine cell differentiation is orchestrated by the action of transcription factors that operate in a gene regulatory network to activate endocrine lineage genes and repress lineage-inappropriate genes. MicroRNAs (miRNAs) are important modulators of gene expression, yet their role in endocrine cell differentiation has not been systematically explored. Here we characterize miRNA-regulatory networks active in human endocrine cell differentiation by combining small RNA sequencing, miRNA over-expression, and network modeling approaches. Our analysis identified Let-7g, Let-7a, miR-200a, miR-127, and miR-375 as endocrine-enriched miRNAs that drive endocrine cell differentiation-associated gene expression changes. These miRNAs are predicted to target different transcription factors, which converge on genes involved in cell cycle regulation. When expressed in human embryonic stem cell-derived pancreatic progenitors, these miRNAs induce cell cycle exit and promote endocrine cell differentiation. Our study delineates the role of miRNAs in human endocrine cell differentiation and identifies miRNAs that could facilitate endocrine cell reprogramming.

INTRODUCTION

The potential to generate pancreatic beta cells from human pluripotent stem cells (hPSCs) or via cell reprogramming from other cell sources holds promise for modeling causes of diabetes and cell replacement therapies (Benthuysen et al., 2016). Knowledge of the molecular underpinnings of pancreas and beta cell development has enabled some success in developing beta cell reprogramming and directed differentiation strategies. In particular, the identification of transcription factors (TFs) governing cell fate decisions has been instrumental for cell reprogramming approaches (Benthuysen et al., 2016). Although TFs play a major role in orchestrating gene expression changes during developmental transitions, recent evidence also shows significant roles for other regulators such as small RNAs.

MicroRNAs (miRNAs) are a group of small non-coding RNAs (~22 nucleotides) with known roles in the regulation of gene expression in development, mature cell function, and disease (Vidigal and Ventura, 2015). Studies in mice and zebrafish have demonstrated important roles for miRNAs in pancreatic endocrine cell development and beta cell function (Kaspi et al., 2014). Pancreatic progenitor cell-specific deletion of Dicer1, an enzyme that is universally required for the functional maturation of miRNAs, results in reduced endocrine cell numbers (Lynn et al., 2007), whereas Dicer1 disruption in beta cells impairs insulin biogenesis (Melkman-Zehavi et al., 2011). At the level of individual miRNAs, miR-375 (Kloosterman et al., 2007; Poy et al., 2009) and miR-7 (Kredo-Russo et al., 2012; Latreille et al., 2014) have been identified as regulators of beta cell differentiation and function.

Generally, miRNAs are thought to repress target mRNAs and act by destabilizing mRNAs through base pairing between the miRNA seed sequence (nucleotides at position 2–8) and a complementary sequence in the target mRNA (Guo et al., 2010; Lim et al., 2005). However, recent evidence suggests that miRNAs can also activate gene expression (Jopling et al., 2008; Valinezhad Orang et al., 2014; Vasudevan et al., 2007). The effects of individual miRNAs on gene expression are generally small, which has led to the concept that miRNAs fine-tune gene expression rather than acting as genetic switches (Vidigal and Ventura, 2015). Consistent with this idea, miRNAs have been shown to promote cell differentiation and to facilitate cell reprogramming when force expressed in conjunction with lineage-determining TFs (Chen et al., 2004, 2006;
Mechanistically, each miRNA has the ability to repress hundreds of mRNA targets, and multiple miRNAs often converge on a single pathway to promote a common developmental outcome (Lim et al., 2005; Vidigal and Ventura, 2015). Therefore, a comprehensive understanding of context-specific contributions of miRNAs to gene regulation requires a systems-level approach where all miRNAs and their targets are considered.

In this study we used genome-wide small RNA sequencing to identify candidate miRNAs with possible roles in human endocrine cell differentiation. By comparing miRNA profiles of hPSC-derived pancreatic progenitors and human cadaveric beta and alpha cells genome-wide, we identified miRNAs that are induced during endocrine cell differentiation. Through gain-of-function experiments during hPSC differentiation, we show that islet cell-enriched miRNAs act to promote cell cycle exit and hence islet cell differentiation. Integrating RNA-seq, CLIP-seq, and chromatin state data, we applied a network modeling approach to identify candidate miRNA-regulated TFs that explain the impact of islet cell-enriched miRNAs on cell cycle regulation during endocrine cell differentiation. Our findings provide a systems-level view of how miRNAs regulate human endocrine cell differentiation, which has implications for programming islet endocrine cells from hPSCs or other cell sources.

RESULTS

Identification of miRNAs Up-Regulated during Endocrine Cell Differentiation

To identify miRNAs that are regulated during pancreatic beta cell differentiation, we conducted genome-wide small RNA sequencing in pancreatic progenitor cells derived from CyT49 human embryonic stem cells (hESCs) (Figure S1) and primary beta cells isolated from cadaveric human islets by fluorescence-activated cell sorting (Kameswaran et al., 2014) (Figure 1A). Although both up- and down-regulated miRNAs could have roles in beta cell differentiation, we here focused on miRNAs that increase in expression during endocrine cell differentiation. By comparing expression levels of individual miRNAs in beta cells and pancreatic endoderm stage (PE) cells, we defined miRNAs induced during beta cell differentiation. This analysis revealed 14 miRNAs that were more highly expressed in beta cells than in PE cells (>5,000 sequence reads in beta cells; > 2.3-fold increase; Figure 1B and Tables S1A and S1B). With the exception of miR-127, miR-204, and miR-99b, the same miRNAs also exhibited higher expression in sorted alpha cells compared with PE cells (Figure 1C and Tables S1A and S1C), suggesting shared roles for most miRNAs in the development of both endocrine cell types. Among the miRNAs induced during endocrine cell differentiation were miR-375, miR-200a/c, and miR-7, which have reported roles in beta cell development, proliferation, function, and survival in mice (Belgardt et al., 2015; Kloosterman et al., 2007; Kredo-Russo et al., 2012; Latreille et al., 2014; Nieto et al., 2012; Poy et al., 2004, 2009; Wang et al., 2013). Most notable was the significantly higher expression of members of the Let-7 miRNA family in both beta and alpha cells compared with PE cells, including Let-7a, Let-7b, Let-7f, Let-7g, and miR-98 (Figures 1B and 1D and Tables S1A and S1B). We confirmed the results from the small RNA sequencing by comparing miRNA levels in PE cells and human cadaveric islets using the Taqman miRNA assay (Figure 1E).

Identifying miRNAs Regulating Human Endocrine Cell Differentiation

To identify miRNAs regulated by these miRNAs, we selected several miRNAs for over-expression in hESC-derived PE cells. We included the top three beta cell-enriched miRNAs (miR-375, miR-127, and Let-7a), as well as Let-7g and miR-200a, as they are highly induced during endocrine cell differentiation. miR-7 was not included because of its inhibitory role in endocrine cell differentiation in mice (Kredo-Russo et al., 2012).

To determine the effects of these miRNAs on gene expression, we next over-expressed Let-7g, Let-7a, miR-200a, miR-375, and miR-127 individually in hESC-derived PE cells (Figure 2A). For these studies, we chose PE cells derived from H1 hESCs because a recently published protocol showed very efficient differentiation of H1 hESCs into beta-like cells in vitro (Rezania et al., 2014). Since our genome-wide small RNA sequencing was performed in PE cells from CyT49 hESCs (Figure 1), we first confirmed that H1 and CyT49 hESC-derived PE cells have similar molecular features. Similar to CyT49 hESC-derived PE cells (Figure S1), 98% of H1 hESC-derived PE cells expressed the pancreatic progenitor marker PDX1 (Figures S2A and S2B). In addition, RNA-seq analysis showed highly concordant transcriptome profiles of H1 and CyT49 hESC-derived PE cells [IR > 0.92; Figure S2C]. Furthermore, we confirmed that Let-7g, Let-7a, miR-200a, and miR-375 were expressed at similarly low levels in H1 and CyT49 hESC-derived PE cells (Figure S2D).
Figure 1. Identification of miRNAs Up-Regulated during Endocrine Cell Differentiation of Human Pancreatic Progenitor Cells

(A) Workflow for genome-wide small RNA profiling of pancreatic progenitors (pancreatic endoderm, PE) and endocrine islet cells. PE cells were differentiated from human embryonic stem cells (hESCs), and human alpha and beta cells were isolated from cadaveric human islets by fluorescence-activated cell sorting (FACS).

(B and C) MA plots comparing miRNA expression levels in PE cells and beta cells (B) or PE cells and alpha cells (C). miRNAs with higher expression in beta and alpha cells than PE are indicated by red circles in B and C, respectively. Blue lines indicate 2-fold change in miRNA expression; yellow line indicates no change.

(D) Heatmap comparing expression levels in PE, alpha cells, and beta cells of the thirteen most highly enriched miRNAs in beta cells compared with PE cells.

(E) Relative expression of indicated miRNAs determined by Taqman qPCR in PE cells and human islets.

Data are shown as mean ± S.E.M. (n = 3 biological replicates); ns, not significant; **p < 0.01, ***p < 0.001; Student's t test.

See also Figure S1 and Table S1.
Figure 2. Endocrine Cell-Enriched miRNAs Regulate Expression of Cell Cycle Genes in Pancreatic Progenitor Cells

(A) Workflow to identify genes repressed by each indicated miRNA after lentiviral transduction of hESC-derived pancreatic endoderm (PE) cells. Transduced cells were sorted based on mCherry after 48 h, RNA-seq analysis performed (n = 3 biological replicates), and down-regulated genes identified.

(B–F) Venn diagrams showing the overlap between genes down-regulated in islets (n = 3) compared with PE (n = 2) (blue) and genes repressed by Let-7g (purple, B), Let-7a (red, C), miR-200a (yellow, D), miR-375 (green, E), or miR-127 (light blue, F). Top five GO categories enriched among genes repressed by the miRNA and down-regulated in islets compared with PE are shown on the right.

(G) Venn diagram showing overlap between miRNA-repressed genes.

See also Figure S2, Tables S2, and S3.
Inclusion of an mCherry reporter into the miRNA constructs allowed us to monitor transduction efficiencies in PE stage cultures and to isolate transduced cells by FACS. We observed 13%–20% mCherry+ PE cells 2 days after transduction, and this number increased to 34%–49% 6 days after transduction (Figure S2E). The increase is likely explained by the lentiviral expression vector requiring more than 2 days to reach maximum expression. To identify miRNA targets, we analyzed sorted mCherry+ PE cells 2 days after transduction, reasoning that this early time point is best suited for studying the direct effects of miRNA on gene expression. As expected, Let-7g, Let-7a, miR-200a, miR-375, and miR-127 were each expressed at significantly higher levels in cells transduced with the miRNA-expressing vector compared with control vectortransduced cells (Figure S2F). Furthermore, forced expression of Let-7g, Let-7a, miR-200a, miR-375, or miR-127 in hESC-derived PE repressed the expression of genes (p < 0.05, permutation test, Tables S2A–S2E) that were down-regulated between PE and islets (Figure S2G), suggesting that these miRNAs could contribute to gene expression changes during islet cell differentiation.

Islet Cell-Enriched miRNAs Regulate Expression of Cell Cycle Genes in Pancreatic Progenitor Cells

To identify miRNA-regulated transcripts with likely roles in endocrine cell differentiation, we analyzed sets of genes that were down-regulated by forced expression of each miRNA (p < 0.05, permutation test) and also down-regulated in islets as compared with PE cells (p < 0.05, permutation test, Figures 2B–2F). These mRNA subsets comprised 16.5% of Let-7g-, 13.9% of Let-7a-, 18.5% of miR-200a-, 18.7% of miR-375-, and 30.3% of miR-127-repressed mRNAs in PE cells. We then performed Gene Ontology (GO) analysis to define the biological processes regulated by mRNAs that are repressed by individual miRNAs and are also expressed at lower level in islet than PE cells. The top five enriched GO categories for each one of these miRNA-regulated sets of mRNAs comprised processes associated with DNA replication and regulation of the cell cycle (Figures 2B–2F and Tables S3A–S3E). Given that endocrine cell formation is associated with cell cycle exit (Kim et al., 2015; Miyatsuka et al., 2011; Piccand et al., 2014), these findings suggest that miRNAs could control endocrine cell differentiation by regulating mRNAs involved in cell cycle control.

The finding that all five miRNAs regulate cell cycle-associated transcripts raised the question of whether they share similar target genes. Analysis of the extent of overlap between the mRNAs down-regulated by Let-7g, Let-7a, miR-200a, and miR-375 revealed a modest number of shared targets (Figure 2G). Only seven mRNAs (ZNF239, PIF1, CDC45, TMEM114, HIST1H4H, MRPS25, and ESPL1) were repressed by all five miRNAs, indicating distinct regulatory roles for each one of the miRNAs. Together, these results suggest distinct but converging miRNA targets in regulating cell division in pancreatic progenitors.

Endocrine-Enriched miRNAs Regulate Cell Cycle-Associated Transcription Factors

To decipher mechanisms by which Let-7g, Let-7a, miR-200a, and miR-375 regulate cell cycle genes, we sought to distinguish direct and indirect targets of the four miRNAs (Figure 3A). We defined putative direct targets as poly-miR-repressed mRNAs (p < 0.05, permutation test) predicted to be direct targets by TargetScan (based on matching sequence to the miRNA seed region) and/or exhibiting binding to the RNA-binding protein Argonaute, as determined by CLIP-seq in human islets (Kameswaran et al., 2014). From this analysis, 223 putative direct target mRNAs were identified (Figure 3B and Table S4A). Of these, all were predicted by TargetScan and 35 also by HITS-CLIP. These 223 genes represent 9.1% of
Figure 3. Identification of Putative Direct miRNA Target mRNAs in Pancreatic Progenitor Cells

(A) Workflow to identify repressed genes after transduction of hESC-derived pancreatic endoderm (PE) cells with a lentivirus expressing a polycistronic construct for the indicated miRNAs (poly-miR) and mCherry. Transduced cells were sorted after 48 h, RNA-seq analysis performed (n = 3 biological replicates), and down-regulated genes identified. Direct targets of candidate miRNAs were identified based on TargetScan and CLIP-seq analysis.

(B) Pie graph showing percentage of direct (dark gray) and indirect (light gray) targets of candidate miRNAs repressed by poly-miR construct.

(C) GSEA plot showing enrichment of 223 direct target genes of Let-7g, Let-7a, miR-200a, and miR-375 in islets (n = 3) compared with PE (n = 2). False Discovery Rate (FDR) is shown.

(D) mRNA expression levels of transcription factors directly targeted by Let-7g, Let-7a, miR-200a, and miR-375 measured in reads per kilobase per million reads mapped (RPKM).

(E) Predicted network of transcription factors downstream of miRNAs. Transcription factors are indicated by gray triangles, and individual miRNAs are indicated by colored squares.

See also Figure S3, Tables S2, S3, and S4.
all poly-miR-repressed genes. Reinforcing the potential relevance of these predicted direct miRNA targets for endocrine cell development, GSEA analysis showed significantly lower expression of these genes in islets than in PE cells (Figure 3C).

To determine whether miRNAs are direct regulators of cell cycle-associated mRNAs in PE cells, we analyzed enriched GO terms among the 223 predicted direct miRNA targets. We found no enrichment of categories linked to cell cycle-related processes (Table S4B). Moreover, many of the cell cycle regulators that were repressed by the poly-miR, including CCND3, CDC45, MCM7, and CKS1B (Table S2F), were not among the predicted direct miRNA targets (Table S4A). Thus, cell cycle-associated transcripts appear to be indirectly regulated by the miRNAs. We postulated that this indirect effect of miRNAs on the expression of cell cycle genes could be mediated through the regulation of TFs. Consistent with this hypothesis, 21 TFs were among the 223 putative direct miRNA targets (Figure 3D and Tables S4A and S4C). A striking finding was that many of these TFs have documented roles in cell cycle regulation, including E2F2, which is part of the complex controlling cell cycle progression, and numerous TFs are known to regulate cell growth (e.g., ZC3H10, ZNF783, ZBTB46, ZBTB5, ZFYVE26, TPS3, EZH1, HIF3A, DPF2, TEAD3). In addition, TFs predicted to be directly regulated by the miRNAs included TFs involved in the regulation of endocrine cell development and maturation, such as NNX6.1 (Schaffer et al., 2013; Taylor et al., 2013) and the thyroid hormone receptor THRA, consistent with the role of thyroid hormone in beta cell maturation (Matsuda et al., 2017). Reflective of their shared seed sequence, TFs that are predicted to be directly regulated by Let-7g and Let-7a showed complete overlap, whereas miR-200a and miR-375 mostly regulated separate sets of TFs (Figure 3E). This analysis indicates that Let-7g, Let-7a, miR-200a, and miR-375 might jointly change the transcriptional landscape in PE cells by down-regulating expression of different sets of TFs.

miRNAs Regulate a Network of Cell Cycle Genes in Pancreatic Progenitor Cells

Having identified a set of TFs as potential direct miRNA targets, we next sought to determine whether these TFs could act downstream of the miRNAs to regulate cell cycle genes. To test this, we constructed and subsequently probed a miRNA-gene regulatory network, linking the four candidate miRNAs and their putative direct TF targets to poly-miR-regulated genes predicted to be target genes of the TFs (Figures 4A, S4A, and S4B). First, to identify TF-binding events close to poly-miR-regulated genes (down- and up-regulated), we used ATAC-seq data from PE cells and islets and mapped open chromatin regions surrounding transcriptional start sites (TSSs; closest within 10 kb) of these genes (n = 241,922 sites; FDR <0.01, MACS2) (Figures 4A and S4A; see Transparent Methods). Second, to pinpoint identified candidate TF-bound regions with likely impact on gene regulation during the PE to islet transition, we identified ATAC sites exhibiting dynamics in histone modifications between PE cells and islets. We focused on H3K4me3 and H3K27ac, two highly dynamic histone modifications during development (Wang et al., 2015; Xie et al., 2013) that have been associated with active promoters (H3K4me3) and active promoters and enhancers (H3K27ac) (Creyghton et al., 2010; Heintzman et al., 2009). We then tested whether changes in these histone marks are accompanied by expression changes of proximal genes. As predicted, an increase in H3K4me3 and H3K27ac deposition in PE compared with islets was associated with higher mRNA levels (p = 5.3 × 10−34; Mann-Whitney test), whereas a decrease was associated with lower mRNA levels (p = 1.9 × 10−36). Finally, to construct the network, we linked open chromatin regions with dynamic histone marks to miRNA-regulated TFs by identifying those regions with a matching TF-binding motif. Validating our miRNA-gene regulatory network, GO analysis showed that the 1,307 genes comprising the network were enriched for cell cycle regulators (Figure S4C and Table S5).

Having validated our approach of linking putative TF binding events to changes in gene transcription during the PE to islet transition, we next assembled all data into a structured graph (Figures 4A and S4A) consisting of different types of nodes that represent the individual datasets, namely, the four candidate miRNAs (squares, Let-7g, Let-7a, miR-200a, and miR-375), their predicted target TFs (triangles), predicted TF-binding regions (hexagons), and indirect miRNA target genes (circles). Each connection between nodes (i.e., edge) was given a score representing the strength of their association, as inferred from miRNA-target databases (a), from algorithms matching TF motifs to DNA sequences (b), or from differential regulation of the connected gene (c) (see Figure S4A for details). A combined score was then computed for each possible path in the network from a miRNA to a gene. The score for an individual gene (G1) regulated by a miRNA is the sum of the edge scores (S1 = a1 + b1 + c1). To account for effects of more than one miRNA on an individual gene, we then computed a combined score representing the connectivity of the four miRNAs to G1. According to this scoring system, a higher rank is assigned to genes with strong connectivity of
individual miRNA-mediated paths and characterized by an effect of more than one miRNA (Figure S4B and Table S6). The resulting network of the 40 highest scoring cell cycle genes demonstrates direct connectivity of the four miRNAs with cell cycle regulators through several TFs, including SOX12, RFX1, TP53, E2F2, MBD2, ZNF512B, ZNF783, and ZNF641 (Figure 4B). Of interest is the identification of NEUROD1 as an indirect miRNA target of Let-7 miRNAs and miR200a. Neurod1 is a TF that has been shown to induce cell cycle exit and to regulate endocrine cell differentiation in model organisms (Ahnfelt-Ronne et al., 2007; Mutoh et al., 1998). Our network analysis identifies a core network of miRNAs, TFs as their putative direct targets, and down-stream genes with likely roles in cell cycle regulation and endocrine cell differentiation.
A. Polycistronic miRNA cluster (poly-miR)

2D monolayer → hESCs → Analysis

- Ki67 and insulin staining
- RNA-seq

3D aggregates

B. 3D aggregates

Control

Poly-miR

C. Control vs Poly-miR

D. Control vs Poly-miR

E. Gene expression analysis

- Log2 FC (Poly-miR/control)

F. Gene expression analysis

- Log2 FC (Poly-miR/control)
miRNAs Regulate Endocrine Cell Differentiation by Promoting Cell Cycle Exit

We next determined whether forced expression of Let-7g, Let-7a, miR-200a, and miR-375 represses cell cycle progression in hESC-derived pancreatic progenitor cells, as predicted by our computational analysis. We transduced PE cells with the poly-miR lentiviral construct and differentiated these cells for another 6 days as 3D aggregates to the early pancreatic endocrine (EN) stage, when insulin+ cells are first present (Figure 5A). Sectioned aggregates were then stained for the proliferation marker Ki-67. Consistent with our computational prediction, forced expression of the miRNAs reduced the percentage of Ki-67+ cells in aggregates (Figures 5B and 5D). The miRNAs likely exhibit their anti-proliferative effect in progenitors and not beta cells, as insulin+ cells in EN stage cultures were mostly Ki67− in both control vector- and poly-miR-transduced aggregates (Figure 5A).

Since cell cycle exit and endocrine cell differentiation are tightly coupled (Kim et al., 2015; Miyatsuka et al., 2011; Piccand et al., 2014), we tested whether the reduction in Ki-67+ cells after miRNA over-expression was associated with an increase in the number of insulin+ cells. Indeed, we observed a higher percentage of insulin+ cells in aggregates expressing the poly-miR construct compared with vector-transduced aggregates (Figures 5B and 5D). The bias of our culture conditions for the differentiation of insulin+ cells (Figure 5B) precluded quantification of other endocrine cell types. To further determine how Let-7g, Let-7a, miR-200a, and miR-375 over-expression affects gene expression at the EN stage, we conducted RNA-seq analysis of sorted mCherry+ cells. Consistent with the reduction in Ki67+ cells in poly-miR-transduced cultures (Figure 5C), genes associated with cell cycle regulation, such as CCND1, CDK4, and PCNA, were enriched among genes down-regulated (p < 0.05, permutation test) by forced miRNA expression (Figure 5E and Tables S7A and S7B). Up-regulated genes (p < 0.05, permutation test) comprised endocrine cell-characteristic genes involved in the regulation of insulin secretion and ion transport (e.g., GLPR1, SYT1, CACNA1A, SLC2A1) (Figure 5F and Tables S7A and S7C), further supporting the conclusion that Let-7g, Let-7a, miR-200a, and miR-375 promote endocrine cell differentiation. We note that insulin mRNA levels were slightly decreased rather than increased in poly-miR transduced cells (Table S7A), suggesting that miRNAs or their target genes could affect insulin protein levels at the posttranscriptional level. Taken together, our data support a model whereby endocrine-enriched Let-7g, Let-7a, miR-200a, and miR-375 are part of a gene regulatory network that triggers cell cycle exit to promote endocrine cell differentiation.

DISCUSSION

Here, we identified 14 miRNAs (Let-7g, Let-7a, Let-7f, Let-7b, miR-200a, miR-200c, miR-204, miR-99b, miR-141, miR-127, miR-7, miR-27b, miR-98, and miR-375), that are induced during human beta cell differentiation. We further studied five miRNAs (Let-7g, Let-7a, miR-200a, miR-375, and miR-127) with a high fold-change during endocrine cell differentiation and experimentally show that these miRNAs induce cell cycle exit in pancreatic progenitor cells. By constructing an integrated miRNA-gene regulatory network of endocrine cell differentiation, we show that these miRNAs likely contribute to endocrine cell differentiation by directly regulating different sets of cell cycle-associated TFs.

To analyze how islet cell-enriched miRNAs cooperate to drive endocrine cell differentiation, we developed a computational method to model the relationship of miRNAs, TFs, and miRNA-regulated genes. Our
computational model builds on a previously published approach for constructing miRNA regulatory networks (Gosline et al., 2016) and integrates chromatin state and expression data to build a multi-layer network. Our approach differs in a few key aspects from published methodologies. First, it incorporates predictions from both CLIP-seq data and TargetScan into a combined score that is assigned to network edges. In addition, our scoring system focuses on a set of miRNAs identified experimentally and weighs the number of miRNAs contributing to each path, accounting for potential synergistic effects of miRNAs on downstream gene expression changes. As such, the algorithm presented here can be applied to other cellular contexts with matching miRNA/mRNA/chromatin data and provides a useful framework for the prediction of miRNA effects.

We found that islet cell-enriched miRNAs Let-7g, Let-7a, miR-200a, miR-375, and miR-127 repress different transcripts involved in cell cycle regulation and therefore might synergize in driving cell cycle exit and endocrine cell differentiation. All four miRNAs have been implicated in the regulation of cell proliferation in other contexts. Like the Let-7 family miRNAs studied here, Let-7b inhibits proliferation and induces neural differentiation when over-expressed in neural progenitors (Zhao et al., 2010). Furthermore, Let-7b, miR-200a, and miR-375 have been shown to induce cell cycle arrest in tumor cells (Liu et al., 2012; Uhlmann et al., 2010; Wang et al., 2011). Likewise, acute over-expression of miR-375 in dedifferentiated beta cells reduces their proliferation and promotes their re-differentiation (Nathan et al., 2015). This anti-proliferative effect of miR-375 is opposite to observations in miR-375-deficient mice, which exhibit decreased beta cell proliferation (Poy et al., 2009). Since these mice carry a germline mutation of miR-375, it is possible that the observed decrease in beta cell proliferation is the consequence of a developmental defect rather than a reflection of miR-375 directly regulating inhibitors of cell cycle progression.

Pancreatic endocrine cell differentiation is tightly linked to cell cycle exit. In both mice and humans, endocrine cell differentiation depends on the TF NGN3 (encoded by NEUROG3 (Gradwohl et al., 2000; McGrath et al., 2015), which commits pancreatic progenitors to the endocrine lineage and promotes cell cycle exit by inducing the cell cycle inhibitors Cdkn1a (p21/CIP1) and Pak3 (Miyatsuka et al., 2011; Piccand et al., 2014). We observed no effect of either combined or individual Let-7g, Let-7a, miR-200a, and miR-375 over-expression on NEUROG3 mRNA levels (Tables S2A−S2D, S2F), suggesting that these miRNAs exert their effect on proliferation independent of NGN3. However, Let-7g, Let-7a, miR-200a, and miR-375 expression with the poly-miR construct significantly induced the NGN3 target gene and endocrine differentiation factor NEUROD1 (Ahnfelt-Ronne et al., 2007). Based on our computational model, these miRNAs are predicted to modulate NEUROD1 expression indirectly through down-regulation of NEUROD1 upstream TFs. Given that NEUROD1 can promote cell cycle exit through direct activation of Cdkn1a (Mutoh et al., 1998), miRNA-mediated modulation of NEUROD1 levels likely contributes to the observed effect of islet-enriched miRNAs on cell proliferation and differentiation.

Gain- and loss-of-function studies in model organisms have shown that the repressive effects of miRNAs on their targets is mostly modest, which has led to the view that miRNAs act to fine-tune gene expression. Consistent with this view, we observed relatively small effects of miRNA over-expression on gene expression, cell proliferation, and endocrine cell differentiation. However, these results do not mean that the miRNAs are not important for endocrine cell differentiation. We over-expressed islet cell-enriched miRNAs in an in vitro system where growth factor conditions have been optimized for efficient beta cell differentiation. Therefore, the miRNAs might not be limiting in the context of these optimized conditions. Studies in model organisms underscore the idea that miRNAs confer robustness to developmental processes and become limiting only under conditions of stress. For example, loss of miR-7 has little effect on Drosophila sensory organ development under normal conditions, but when environmental stresses are added to the developing organism, miR-7 becomes necessary (Li et al., 2009). Similar examples exist in worms and mice, where miRNA deletions lead to significant developmental perturbations only on sensitized backgrounds or under stress (Brenner et al., 2010; Chivukula et al., 2014). Further illustrating that miRNAs can have significant biological effects in specific contexts, miRNAs have been shown to drastically augment reprogramming efficiencies (Anokye-Danso et al., 2011; Yao et al., 2011). Therefore, the here-identified islet cell-enriched miRNAs could help develop still missing protocols for robust direct reprogramming of human endocrine cells.

Limitations of the Study
One limitation of our approach for identifying endocrine cell differentiation-relevant miRNAs is the focus on miRNAs that repress mRNAs. There is evidence that miRNAs can activate gene expression or directly
reduce protein levels (El Ouaamari et al., 2008; Jopling et al., 2008; Vasudevan et al., 2007). It is possible that some of the identified miRNAs regulate endocrine cell differentiation through these mechanisms. Another limitation is that we compared miRNA profiles in pancreatic progenitors and mature human endocrine cells. Therefore, we do not know how these miRNAs are regulated during postnatal endocrine cell maturation. Finally, our network modeling approach predicts synergist effects of Let-7g, Let-7a, miR-200a, and miR-375 on cell cycle regulation. This predicted synergy will have to be experimentally validated.

METHODS
All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY
All RNA-seq and ATAC-seq data generated in this study can be found at GEO with accession number GSE115327.

Accession numbers for additional data used in this study are as follows: GSE52314 (small RNA-seq, sorted alpha and beta cells); GSE51924 (CLIP-seq, human islets); E-MTAB-1086 (RNA-seq, PE cells); GSE54471 (H3K27ac, H3K4me3 ChIP-seq, PE cells); GSE51311, E-MTAB-1919, E-MTAB-189, and E-MTAB-191 (H3K27ac, H3K4me3 ChIP-seq, human islets).

SUPPLEMENTAL INFORMATION
Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.10.063.

ACKNOWLEDGMENTS
We thank Medhavi Mallick and Gaowei Wang for help with figure preparation. Human pancreatic islets were provided by the Integrated Islet Distribution Program (UC4 DK098085). We acknowledge the UCSD IGM Genomics Center for next-generation sequencing (P30 DK063491). This work was supported by the National Institutes of Health (R01 DK068471 and R01 DK078803 to M.S.), the California Institute for Regenerative Medicine (RB4-06144 to M.S.), and postdoctoral fellowships from the Juvenile Diabetes Research Foundation Ltd (3-2015-83 to W.J., 3-2012-177 to A.W., and 3-2017-386 to K.-V.N.-N.) and the Larry L. Hillblom Foundation (2015-D-021-FEL to B.G.).

AUTHOR CONTRIBUTIONS
W.J., F.M., and M.S. conceived the project and designed the experiments. W.J., B.G., J.W., C.Z., K.-V.N.-N., I.M., and H.-P.S. performed experiments. F.M., Y.S., A.W., and J.C. conducted bioinformatics analysis. N.V. contributed ATAC-seq analysis, K.H.K. provided small RNA-seq data, and K.A.F. provided islet RNA-seq and islet ATAC-seq data. W.J., F.M., B.G., and M.S. interpreted data. W.J., F.M., A.C.C., and M.S. wrote the manuscript.

DECLARATION OF INTERESTS
The authors declare no competing interests.

Received: October 3, 2018
Revised: September 30, 2019
Accepted: October 28, 2019
Published: November 22, 2019

REFERENCES


Supplemental Information

A Network of microRNAs Acts to Promote Cell Cycle Exit and Differentiation of Human Pancreatic Endocrine Cells

Wen Jin, Francesca Mulas, Bjoern Gaertner, Yinghui Sui, Jinzhao Wang, Ileana Matta, Chun Zeng, Nicholas Vinckier, Allen Wang, Kim-Vy Nguyen-Ngoc, Joshua Chiou, Klaus H. Kaestner, Kelly A. Frazer, Andrea C. Carrano, Hung-Ping Shih, and Maike Sander
Figure S1: Pancreatic endoderm differentiated from CyT49 hESCs. Related to Figure 1.  
(A) Schematic of the hESC-based differentiation strategy. (B) Immunofluorescence staining of PE cell aggregate sections for PE-specific markers PDX1 and SOX9. Scale bar, 50 µm. (C) Representative flow cytometry analysis at PE stage for PDX1. hESC, human embryonic stem cells; DE, definitive endoderm; GT, gut tube; FG posterior foregut; PE, pancreatic endoderm; EN, endocrine cell stage.
Figure S2

A

PDX1/SOX9/DAPI (H1 PE)

B

H1 PE

% of Max

-10^0 10^0 10^1 10^2 10^3 10^4 10^5 10^6
PDX1

C

CyT49 PE[Log2(RPKM+1)]

R=0.92

D

Relative miRNA levels

(fold change to CyT49 PE)

CyT49 PE

H1 PE

Islet

Let-7a

Let-7g

miR-200a

miR-375

E

miRNA transduction (48 hours)

miRNA transduction (6 days)

Let-7a

Let-7g

miR-200a

miR-375

mCherry

% of Max

-10^0 10^0 10^1 10^2 10^3 10^4 10^5 10^6

F

Relative miRNA expression

(fold change to control)

***

C

miRNA

C

miRNA

C

miRNA

C

miRNA

C

miRNA

G

Enrichment score

Let-7g

Let-7a

miR-127

FDR<0.001

Islet

PE

Islet

PE

Islet

PE

FDR<0.001

FDR<0.001

FDR<0.001
Figure S2. Forced expression of individual miRNAs in hESC-derived pancreatic progenitor cells. Related to Figure 2.

(A) Immunofluorescence staining of pancreatic endoderm (PE) differentiated from H1 hESCs for PDX1 and SOX9. Scale bar, 50 µm. (B) Representative flow cytometry analysis at PE stage for PDX1. (C) Scatter plot showing correlation in mRNA expression between PE cells derived from H1 hESCs and CyT49 hESCs. (D) Expression of indicated miRNAs in H1-derived PE cells and islets relative to CyT49-derived PE cells determined by Taqman qPCR. Data are shown as mean ± S.E.M. (n = 3 technical replicates). (E) Representative flow cytometry analysis for mCherry 48 h (top row) and 6 days (bottom row) after lentiviral transduction with miRNA-mCherry constructs. Gating for cell sorting is shown. (F) Relative expression of indicated miRNAs determined by Taqman qPCR in H1 PE cells 48 h after lentiviral transduction with miRNAs or vector control (C). Data are shown as mean ± S.E.M. (n = 3 biological replicates). **P < 0.01, ***P < 0.001; Student’s t-test. (G) GSEA plots showing enrichment of genes repressed by Let-7g, Let-7a, miR-200a, miR-375, and miR-127 in islets (n = 3) compared to PE (n = 2). False Discovery Rate (FDR) is shown.
Figure S3. Forced expression of a polycistronic construct for four miRNAs in hESC-derived pancreatic progenitor cells. Related to Figure 3. 

(A) Representative flow cytometry analysis for mCherry 48 h (left) and 6 days (right) after lentiviral transduction with poly-miR-mCherry construct. Gating for cell sorting is shown. (B) Relative expression of indicated miRNAs determined by Taqman qPCR in H1 PE cells 48 h after lentiviral transduction with a vector control or a polycistronic construct containing Let-7g, Let-7a, miR-200a, and miR-375 (poly-miR). Data are shown as mean ± S.E.M. (n = 3 biological replicates). **P < 0.01, ***P < 0.001; Student’s t-test. (C) GSEA plot showing enrichment of genes repressed by the poly-miR construct in islets compared to PE. False Discovery Rate (FDR) is shown. (D) Venn diagram showing the overlap between genes down-regulated in islets compared to PE (blue) and genes repressed by the poly-miR construct (grey). Top five GO categories enriched among genes repressed by the poly-miR construct and down-regulated in islets compared to PE are shown on the bottom.
Figure S4. Approach to identify core network of miRNA-regulated genes in pancreatic progenitor cells. Related to Figure 4.

(A,B) Schematic of approach to identify core network of miRNA-regulated transcription factors and cell cycle genes. Building of network (A) and probing of network (B) is shown. The nodes of the graph represent miRNAs [squares; Let-7g (purple), Let-7a (red), miR-200a (yellow), and miR-375 (green)], transcription factors (triangles), transcription factor binding regions (hexagons) and genes (circles). In (A), the source for each node is indicated on the left, while evidence (indicated by a, b, and c) used to calculate scores is indicated on the right. In (B), path-based scoring of an individual gene [e.g. gene 1 (G1)] is shown. See Materials and Methods for details. (C) Top five GO categories enriched in genes comprising the miRNA-regulated core network in (A). TF, transcription factor; G, gene; S, score.
Figure S5. Endocrine cell differentiation from H1 hESCs. Related to Figure 5. **(A)** Representative images showing immunofluorescence staining for insulin (INS), Ki-67, mCherry, and DAPI at the endocrine cell (EN) stage for control vector (top) or poly-miR (bottom) transduced aggregates. **(B)** Immunofluorescence staining of sections from EN stage aggregates for insulin (INS), glucagon (GCG), and somatostatin (SST). Scale bar, 50 µm.
Transparent Methods

Contact for reagent and resource sharing

Further information and requests for reagents may be directed and will be fulfilled by the corresponding author Maike Sander (masander@ucsd.edu).

Experimental model and subject details

Human islets

Human cadaveric pancreatic islets for the Taqman miRNA analysis (Donor ID: 1-3) and for the RNA-seq and ATAC-seq analysis (Donor ID: 4-7) were obtained through the Integrated Islet Distribution Program (IIDP). The islets had ≥ 90% purity and ≥ 90% viability. Upon receipt, islets were handpicked and immediately processed for RNA extraction or isolation of nuclei.

<table>
<thead>
<tr>
<th>Donor ID</th>
<th>Donor Age</th>
<th>Donor Sex</th>
<th>Diabetes*</th>
<th>BMI</th>
<th>Race</th>
<th>Cause of Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53</td>
<td>Male</td>
<td>No</td>
<td>27.2</td>
<td>Caucasian</td>
<td>CVA/STROKE</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>Female</td>
<td>No</td>
<td>31.0</td>
<td>American Indian or Alaska Native</td>
<td>CVA/STROKE</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>Male</td>
<td>No</td>
<td>27.7</td>
<td>Caucasian</td>
<td>CVA/STROKE</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>Male</td>
<td>No</td>
<td>29.8</td>
<td>Black or African American</td>
<td>CVA/STROKE</td>
</tr>
<tr>
<td>5</td>
<td>59</td>
<td>Female</td>
<td>No</td>
<td>24.7</td>
<td>Caucasian</td>
<td>CNS tumor</td>
</tr>
<tr>
<td>6</td>
<td>55</td>
<td>Male</td>
<td>No</td>
<td>23.2</td>
<td>Caucasian</td>
<td>Head trauma</td>
</tr>
<tr>
<td>7</td>
<td>56</td>
<td>Female</td>
<td>No</td>
<td>33.4</td>
<td>Black or African American</td>
<td>CVA/STROKE</td>
</tr>
</tbody>
</table>
*Diabetes status was defined by the patient’s medical record and, when available, hemoglobin A1c levels.

**Maintenance and differentiation of H1 hESCs**

hESC research was approved by the University of California, San Diego, Institutional Review Board and Embryonic Stem Cell Research Oversight Committee. All hESC experiments were performed in H1 hESCs with the exception of miRNA expression profiling, for which CyT49 hESCs were used.

H1 hESCs were maintained and differentiated as described with some modifications (Rezania et al., 2014). In brief, hESCs were cultured in mTeSR1 media (Stem Cell Technologies) and propagated by passaging cells every 3 to 4 days using Accutase (eBioscience) for enzymatic cell dissociation. For differentiation of H1 cells, we employed a 2D monolayer culture format up to day 11 of differentiation. Cells were then dissociated using accutase for 10 min, reaggregated by plating the cells in a low attachment 6-well plate on an orbital shaker (100 rpm) in a 37 °C incubator. Cells were subsequently cultured in suspension from Days 11-14.

On Day 0, dissociated hESCs were resuspended in mTeSR1 media (see media compositions below) and seeded onto Matrigel-coated 12-well plates by adding 1 ml of cell suspension (~8 x 10^5 cells/well) to each well. The following day, undifferentiated cells were washed in stage 1 medium and then differentiated using a multi-step protocol with stage-specific media (see below) and daily media changes.
All stage-specific base media were comprised of MCDB 131 medium (Thermo Fisher Scientific) supplemented with NaHCO$_3$, GlutaMAX, D-Glucose, and BSA using the following concentrations:

Stage 1/2 medium: MCDB 131 medium, 1.5 g/L NaHCO$_3$, 1X GlutaMAX, 10 mM D-Glucose, 0.5% BSA

Stage 3/4 medium: MCDB 131 medium, 2.5 g/L NaHCO$_3$, 1X GlutaMAX, 10 mM D-glucose, 2% BSA

Stage 5 medium: MCDB 131 medium, 1.5 g/L NaHCO$_3$, 1X GlutaMAX, 20 mM D-glucose, 2% BSA

Media compositions for each stage were as follows:

Stage 1 (Days 0-2): base medium, 100 ng/ml Activin A, 25 ng/ml Wnt3a (Day 0). Day 1-2: base medium, 100 ng/ml Activin A

Stage 2 (Days 3-5): base medium, 0.25 mM L-Ascorbic Acid (Vitamin C), 50 ng/mL FGF7

Stage 3 (Days 6-7): base medium, 0.25 mM L-Ascorbic Acid, 50 ng/mL FGF7, 0.25 µM SANT-1, 1 µM Retinoic Acid, 100 nM LDN193189, 1:200 ITS-X, 200 nM TPB

Stage 4 (Days 8-10): base medium, 0.25 mM L-Ascorbic Acid, 2 ng/mL FGF7, 0.25 µM SANT-1, 0.1 µM Retinoic Acid, 200 nM LDN193189, 1:200 ITS-X, 100nM TPB
Stage 5 (Days 11-14): base medium, 0.25 µM SANT-1, 0.05 µM Retinoic Acid, 100 nM LDN193189, 1:200 ITS-X, 1 µM T3, 10 µM ALK5 inhibitor II, 10 µM ZnSO4, and 10 µg/mL Heparin, 10 µM ROCK inhibitor

End of stage 1 = definitive endoderm
End of stage 2 = gut tube
End of stage 3 = posterior foregut
End of stage 4 = pancreatic endoderm
End of stage 5 = endocrine cells

**Maintenance and differentiation of CyT49 hESCs**

CyT49 hESCs were maintained and differentiated as described (Xie et al., 2013). Propagation of CyT49 hESCs was carried out by passing cells every 3 to 4 days using Accutase™ (eBioscience) for enzymatic cell dissociation, and with 10% (v/v) human AB serum (Valley Biomedical) included in the hESC medium the day of passage. hESCs were seeded into tissue culture flasks at a density of 50,000 cells/cm².

CyT49 hESC medium was comprised of DMEM/F12 (Corning; 45000-346) supplemented with 10% (v/v) KnockOut™ Serum Replacement (Thermo Fisher Scientific), 1X MEM non-essential amino acids (Thermo Fisher Scientific), 1X GlutaMAX™ (Thermo Fisher Scientific), 1% (v/v) penicillin-streptomycin (Thermo Fisher Scientific), 0.1mM 2-mercaptoethanol (Thermo Fisher Scientific), 10ng/mL Activin A (R&D Systems), and 10ng/mL Heregulin-β1 (PeproTech).
Pancreatic differentiation of CyT49 hESCs was performed as previously described (Schulz et al., 2012). Briefly, we employed a suspension-based format using rotational culture. Undifferentiated hESCs were aggregated by preparing a single-cell suspension in hESC media at $1 \times 10^6$ cells/mL and overnight culture in six-well ultra-low attachment plates (Costar) with 5.5ml per well on an orbital rotator (Innova2000, New Brunswick Scientific) at 95 rpm. The following day, undifferentiated aggregates were washed in RPMI media (Corning) and then differentiated using a multistep protocol with daily media changes and continued orbital rotation at either 95 rpm or at 105 rpm on Days 4 to 8.

Stage 1/2 medium: RPMI medium (Corning), 0.2 % (vol/vol) FBS, 1X GlutaMAX

Stage 3/4 medium: DMEM High Glucose medium (HyClone), 0.5X B-27 Supplement, 1X GlutaMAX

Media compositions for each stage were as follows:

Stage 1 (Days 0-1): Day 0: RPMI/FBS, 100ng/mL Activin A, 50ng/mL mouse Wnt3a, 1:5000 ITS. Day 1: RPMI/FBS, 100 ng/mL Activin A, 1:5000 ITS

Stage 2 (Days 2-4): Day 2: RPMI/FBS, 2.5µM TGFβ R1 kinase inhibitor IV, 25ng/mL KGF, 1:1000 ITS. Days 3-4: RPMI/FBS, 25ng/mL KGF, 1:1000 ITS

Stage 3 (Days 5 -7): DMEM/B27, 3nM TTNPB, 0.25mM KAAD-Cyclopamine, 50ng/mL Noggin

Stage 4 (Days 7-10): DMEM/B27, 50ng/mL KGF, 50ng/mL EGF
End of stage 1 = definitive endoderm
End of stage 2 = gut tube
End of stage 3 = posterior foregut
End of stage 4 = pancreatic endoderm

Cell line

HEK293T cells were maintained in DMEM containing 100 units/mL penicillin and 100 mg/mL streptomycin sulfate supplemented with 10% fetal bovine serum (FBS).

Method details

Immunocytochemistry

Cells were washed twice before fixation with 4% paraformaldehyde in PBS for either 30 min at room temperature, or overnight at 4°C. Cells were then washed three times with PBS and incubated in 30% sucrose at 4°C overnight before mounting in Optimal Cutting Temperature Compound (Tissue-Tek) and sectioning at 10 µm. Immunocytochemistry was performed as described (Xie et al., 2013). The following primary antibodies and dilutions were used: guinea pig anti-PDX1 (gift from Dr. Christopher Wright, Vanderbilt University) 1:1000; rabbit anti-SOX9 (Millipore, AB5535) 1:1000; rabbit anti-Ki-67 (ThermoFisher, RM-9106-S1) 1:200; guinea pig anti-insulin (LINCO, 4011-01) 1:1000. Secondary antibodies were Cy5-, Cy3-, or Alex488-conjugated donkey antibodies against guinea pig or rabbit (Jackson Immuno Research Laboratories). Images were
acquired on a Zeiss Axio-Observer-Z1 microscope with a Zeiss AxioCam digital camera and figures prepared with Adobe Photoshop CS6/Illustrator CS5.

To determine the percentage of Ki-67+ cells in the mCherry+ cell population, at least ten sections from different aggregates were analyzed per hPSC differentiation. For each condition, three independent hPSC differentiations were performed. Ki-67+ and mCherry+ cells were quantified using HALO software (PerkinElmer Inc).

Fluorescence-activated cell sorting and intracellular flow cytometry

hESC-derived PE cells were dissociated to a single-cell suspension with Accutase (Stemcell Technologies) at 37°C for 10 min. Accutase was neutralized with FACS sorting buffer [1% (wt/vol) FBS, 1 mM EDTA, 25mM Hepes, PBS]. FACS was performed on a FACS Fortessa equipped with FACS DiVa software (BD Biosciences). Cells were sorted into Trizol for RNA analysis. For intracellular flow cytometry, dissociated cells were fixed, permeabilized with BD Cytoperm/Cytofix (BD Bioscience), and stained with anti-PDX1-PE conjugated antibody (BD Biosciences, 562161; 1:20) at room temperature for 30 min, washed, and resuspended in FACS buffer. Flow cytometry analysis was performed on FACSCanto II (BD Biosciences) and analyzed with FlowJo software (FlowJo LLC).
**TaqMan microRNA assay**

qRT-PCR for miRNAs was performed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems, Cat. No. 4366596). Briefly, 10 ng of total RNA was reverse transcribed using RT primers from the TaqMan MicroRNA Assay kit [Applied Biosystems; probe catalogue numbers: Let-7a (000377), Let-7g (002282), miR-127 (000452), miR-200a (000502), miR-375 (000564), miR-7 (000268), miR-99b (000436), and RUN44 (001094)]. qRT-PCR was performed on a Bio-rad CFX96 real-time system using the TaqMan Universal PCR Master Mix (Applied Biosystems, Cat. No. 4324018) and TaqMan probes from the TaqMan MicroRNA Assay kit. miRNA levels were determined on three independent samples and values were normalized to endogenous snoRNA RNU44.

**miRNA expression vector construction**

To generate miRNA expression vectors, 270 nt of the miRNA gene primary transcript, including the 22 nt mature miRNA and 125 nt of genomic sequence flanking each side of the miRNA (Chen et al., 2004), were amplified. Let-7a and Let-7g were expressed with mutations in their loop sequence to block LIN28 binding and ensure proper miRNA processing (Piskounova et al., 2008). For transduction of PE cells, a modified version of pLKO.3G was used, in which GFP was exchanged for mCherry (pLKO.mcherry). For the polycistronic miRNA expression vector, a gBlock gene fragment encompassing miR-375, Let-7a, Let-7g, and miR-200a was cloned into pLKO.mcherry.
Lentivirus production and transduction of PE cells

High-titer lentiviral supernatants were generated by co-transfection of the miRNA expression vector and the lentiviral packaging construct into HEK293T cells as described (Xie et al., 2013). Briefly, miRNA expression vectors were cotransfected with the pCMV-R8.74 (Addgene, #22036) and pMD2.G (Addgene, #12259) expression plasmids into HEK293T cells using a 1mg/ml PEI solution (Polysciences). Lentiviral supernatants were collected at 48 hr and 72 hr after transfection. Lentiviruses were concentrated by ultracentrifugation for 120 min at 19,500 rpm using a Beckman SW28 ultracentrifuge rotor at 4°C. The titer routinely achieved was $5 \times 10^8$–$10^9$ TU/ml. For PE cell transductions, H1 hESCs were differentiated to the PE stage (Day 8 of differentiation) in monolayer cultures and transduced with lentivirus at a MOI of 2. For RNA analysis, cells were collected 48 hr after transduction.

Small RNA sequencing and data analysis

Small RNA-seq data from sorted human alpha and beta cells have been described (Kameswaran et al., 2014). RNA from PE stage CyT49 hESC cultures was isolated using the miRVana miRNA Isolation kit (Thermo Fisher Scientific). 3 μg of RNA was used for library preparation using the TruSeq Small RNA sample preparation kit (Illumina) and a Pippin Prep (Sage Science) for size selection with a 3% cassette (CSD3010). RNA was prepared for sequencing using the Illumina protocol (Illumina FC-102-1009) and amplified libraries were sequenced on an Illumina Genome Analyzer II (Illumina FC-104-1003). Sequenced libraries were
processed as described (Kameswaran et al., 2014). miRNAs with sample values below 1 RPM were excluded from the analysis. There was one replicate each for hESC-derived PE, alpha, and beta cells. Each miRNA expression value was log₂-transformed and displayed in a heatmap.

**RNA sequencing, mapping and data analysis**

RNA quality was assessed using TapeStation (Agilent Technologies). Libraries were prepared according to the instructions of Illumina's TruSeq RNA library prep kit. Libraries were quantified using High Sensitivity DNA screen tape (Agilent Technologies) and Qubit dsDNA High Sensitivity (Life Technologies) assays. Finally, libraries were multiplexed and sequenced on a HiSeq 2500 (Illumina) sequencer using single-end sequencing.

RNA-seq samples were mapped to the UCSC human transcriptome (hg19/GRCh37) by the Spliced Transcripts Alignment to a Reference (STAR) aligner (STAR-STAR_2.4.0f1), allowing for up to 10 mismatches (Dobin et al., 2013). Only reads aligned uniquely to one genomic location were retained for subsequent analysis. Expression levels of all genes were quantified by Cufflink (https://github.com/cole-trapnell-lab/cufflinks) using only reads with exact matches. Genes with average RPKM above 1 were retained for further analyses.

Differentially expressed genes were identified using a permutation test, with the number of permutations set to 1000. Briefly, all the samples were shuffled, fold changes were computed to obtain a null distribution, and a $P$-value was estimated for each gene's fold change as a cumulative probability from the
null distribution. For comparison of PE and islet data and poly-miR versus control data at the EN stage, batch effects were removed using ComBat (Johnson et al., 2007).

**Gene Set Enrichment Analysis (GSEA) and Gene Ontology (GO) analysis**

We applied GSEA ([http://www.broad.mit.edu/gsea](http://www.broad.mit.edu/gsea)), which scores a-priori defined gene sets in two different conditions (Subramanian et al., 2005). GSEA ([http://www.broad.mit.edu/gsea](http://www.broad.mit.edu/gsea)) was run with the number of permutations for *P*-value computation set to 1000. We used genes significantly repressed by miRNAs (*P* < 0.05, permutation test) as gene sets to determine coordinated regulation in islets compared to PE samples. Gene sets with a false discovery rate of < 0.05 were considered significantly enriched. Enrichment of gene sets for Gene Ontology (GO) terms was tested using Metascape (Tripathi et al., 2015).

**ATAC-seq sample preparation**

Roughly 50,000 PE or primary human islet cells were used for each ATAC-seq assay as described (Buenrostro et al., 2013). Briefly, cell nuclei were isolated using cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 and 0.1% IGEPAL CA-630). The nuclei pellet was resuspended in the transposase reaction mix; 25 μL 2x TD buffer, 2.5 μL transposase (Illumina) and 22.5 μL nuclease-free water at 37°C for 30 min. Then transposed DNA fragments were purified using the Qiagen MinElute kit and amplified 10-12 cycles using the
Nextera (Illumina) PCR primers. Libraries were sequenced on HiSeq4000 platform.

**ChIP-seq and ATAC-seq data analysis**

ChIP-seq and ATAC-seq reads were mapped to the human genome (hg19/GRCh37) using Bowtie (Langmead et al., 2009) and BWA (Li and Durbin, 2009), respectively, and visualized using the UCSC Genome Browser (Kent et al., 2002). Unmapped reads were discarded. After mapping, SAMtools (Li et al., 2009) was used to remove duplicate sequences and merge samples. Here, "SAMtools view -Sbq 30" was used to filter out reads with mapping quality less than 30, "SAMtools rmdup" was used to remove duplicated reads, and "SAMtools merge" was used to merge files of the same histone marker or input condition. ChIP-seq and ATAC-seq analysis was performed in two biological replicates for PE and 4-5 donors for islet. The Pearson correlation among biological replicates ranged from 64% to 96% for human islets and 91% to 96% for PE.

HOMER (Heinz et al., 2010) as used to call ChIP-seq peaks using “findPeaks function” with “–style histone” to call peaks. Stage- and condition-matched input DNA controls were used as background when calling peaks. MACS2 (Zhang et al., 2008) was used to call peaks from ATAC-seq data, with parameters “shift set to 100 bps, smoothing window of 200 bps” and with “nolambda” and “nomodel” flags on.

To link changes in chromatin to gene expression changes, we first defined differential H3K27ac and H3K4me3 peaks in PE and islet (adjusted $P<0.05$,
“getDifferentialPeaksReplicates” function in HOMER) and then used BEDtools (Quinlan and Hall, 2010) to identify overlapping ATAC peaks in PE or islet using a ± 1.5 kb window from the summit of the ATAC peak. Next, we identified the nearest TSS within a 10kb window of the H3K27ac or H3K4me3 peak. We then assessed the concordance of the directionality of changes in gene expression and histone marks by evaluating whether genes near regions showing gain or loss of H3K27ac or H3K4me3 in PE versus islet exhibit significant concordant expression changes (Mann-Whitney test).

Network building
CLIP-seq signal, mRNA expression, ATAC-seq and ChIP-seq binding data were encoded in a graphical model depicted in Figures 4 and Figure S4 by adapting a previously published algorithm (Gosline et al., 2016). Nodes arranged in four different layers, corresponding to miRNAs, transcription factors (TFs), DNA regions, and genes, were identified and connected as follows: For each of the four selected miRNAs, predicted target genes were retrieved through the TargetScan repository (http://www.targetscan.org/vert_71/). For the pairs of miRNA-target gene identified, the corresponding CLIP-seq signal was collected from previously published data (Kameswaran et al., 2014). Among the targets, TFs in the second layer of the network were selected based on down-regulation by poly-miR transfection compared to control with $P < 0.05$ and with an annotation in the TF Animal Database (Zhang et al., 2015). In the third layer of the network, we selected DNA regions showing significant changes in PE versus
islet for H3K27ac or H3K4me3 (see previous paragraph) with the nearest gene showing down-regulation by poly-miR transfection, hereafter referred to as Rsel. The Rsel DNA regions were filtered for links to the selected TFs by scoring the match of their binding motifs with the DNA regions in the network. Briefly, motifs of selected TFs were extracted from a collection of databases, including JASPAR (http://jaspar.genereg.net/cgi-bin/jaspar_db.pl), Hocomoco (http://hocomoco11.autosome.ru/), and ENCODE-related data sets (Aylward et al., 2018) and scored for matches with narrow regions spanning 300bp around the peak summits of each Rsel. Log-odds scores and corresponding $P$-values were obtained using the MEME Suite tool FIMO (http://meme-suite.org/tools/fimo) with default parameters. The last layer was defined by considering genes proximal to DNA regions, as described above, and filtering for those differentially regulated in poly-miR versus control ($P < 0.05$).

A score representing the strength of the association was computed for each pair of connected nodes in the different network layers, as follows: Given the TargetScan context++ score $S_i$ (Agarwal et al., 2015) and the CLIP-seq signal $S_j$ of each miRNA-TF association, their values were normalized in a 0-1 range and combined as $a = (1 - S_i)(1 - S_j)$ (Szkłarczyk et al., 2015). Scores of edges connecting TFs to Rsel regions were defined as the $b = 1 - q$, $q$ being the q-value returned by FIMO, representing the probability of obtaining the log-odds ratio scores of the matches by chance. Scores from DNA regions to genes were defined as the absolute value of the Log2 Fold Change of each gene $x$ in poly-miR versus control data: $c = |Log2FC(x)|$. A combined score was computed for
each possible path in the network starting from a miRNA to a gene, by adding the contribution of the different layers, as: \( S = a + b + c \).

**Network-based gene ranking**

Given an individual gene, G1, all network paths connecting a miRNA to G1 were considered with their corresponding scores and compared for gene ranking as follows: Among the paths connecting the same miRNA to G1, only the one with the highest score was retained, obtaining a score \( S_i \) for each of the \( k \) miRNAs showing an indirect link to G1, \( k \leq N \), with \( N \) equal to the number of miRNAs in the network (\( N=4 \)). A first score summarizing the strength of the association of these retrieved network paths was computed as \( S_{\text{net}}(G1) = \text{mean}(S_i) \), \( i \) ranging from one to \( k \). A second score, accounting for the synergistic effect of several miRNAs on the same gene, was computed as proportion of miRNAs in at least one network path linking to the selected gene G1: \( S_{\text{mir}}(G1) = k/N \). Finally, a combined score for G1 was obtained as a weighted sum of \( S_{\text{net}} \) and \( S_{\text{mir}} \), i.e. \( S_{\text{comb}}(G1) = w \times S_{\text{net}}(G1) + w \times S_{\text{mir}}(G1) \), with \( w \) set to 0.5. This procedure was applied to score individual genes annotated to cell cycle regulation in the Gene Ontology (GO:0051726) and genes were ranked based on \( S_{\text{comb}} \) values.

**Quantification and statistical analysis**

Statistical parameters including FDR, R, and \( P \)-values are reported in the Figures and the Figure Legends.
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabbit anti-PDX1</td>
<td>Abcam</td>
<td>Cat# ab47267, RRID:AB_777179</td>
</tr>
<tr>
<td>Rabbit anti-SOX9</td>
<td>Millipore</td>
<td>Cat# AB5535, RRID:AB_2239761</td>
</tr>
<tr>
<td>Guinea pig anti-INS</td>
<td>Dako</td>
<td>Cat# A0564, RRID:AB_10013624</td>
</tr>
<tr>
<td>Mouse anti-GCG</td>
<td>Sigma-Aldrich</td>
<td>Cat# G2654, RRID:AB_259852</td>
</tr>
<tr>
<td>Rabbit anti-SST</td>
<td>Agilent</td>
<td>Cat# A056601-2</td>
</tr>
<tr>
<td>Mouse anti-Ki67</td>
<td>Lab Vision</td>
<td>Cat# RM-9106-S1, RRID:AB_149792</td>
</tr>
<tr>
<td>Mouse anti-PDX1, PE conjugate</td>
<td>BD Biosciences</td>
<td>Cat# 562161, RRID:AB_10893589</td>
</tr>
<tr>
<td>Mouse IgG1, kappa isotype control (PE conjugate)</td>
<td>BD Biosciences</td>
<td>Cat# 556650, RRID:AB_396514</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPBS</td>
<td>Corning</td>
<td>Cat# 45000-434</td>
</tr>
<tr>
<td>Fatty Acid-Free BSA</td>
<td>Proliant Biologicals</td>
<td>Cat# 68700</td>
</tr>
<tr>
<td>D-(-)-Glucose Solution, 45%</td>
<td>Sigma-Aldrich</td>
<td>Cat# G8769</td>
</tr>
<tr>
<td>Accutase®</td>
<td>eBioscience</td>
<td>Cat# 00-4555-56</td>
</tr>
<tr>
<td>Penicillin-Streptomycin</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 15140122</td>
</tr>
<tr>
<td>GlutaMAX™</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 35050061</td>
</tr>
<tr>
<td>MEM Non-Essential Amino Acids Solution (100X)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 11140050</td>
</tr>
<tr>
<td>Sodium Bicarbonate</td>
<td>Sigma-Aldrich</td>
<td>Cat# NC0564699</td>
</tr>
<tr>
<td>Matrigel®</td>
<td>Corning</td>
<td>Cat# 356231</td>
</tr>
<tr>
<td>ROCK Inhibitor Y-27632</td>
<td>STEMCELL Technologies</td>
<td>Cat# 72305</td>
</tr>
<tr>
<td>mTeSR1 Complete Kit - GMP</td>
<td>STEMCELL Technologies</td>
<td>Cat# 85850</td>
</tr>
<tr>
<td>RPMI 1640 1X, w/o L-Glutamine</td>
<td>Corning</td>
<td>Cat# 45000-404</td>
</tr>
<tr>
<td>DMEM/F12 with L-Glutamine, HEPES</td>
<td>Corning</td>
<td>Cat# 45000-350</td>
</tr>
<tr>
<td>DMEM/F12 w/o L-Glutamine</td>
<td>Corning</td>
<td>Cat# 45000-346</td>
</tr>
<tr>
<td>HyClone Dulbecco’s Modified Eagles Medium</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# SH30081.FS</td>
</tr>
<tr>
<td>MCDB 131</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 10372-019</td>
</tr>
<tr>
<td>CTS™ KnockOut™ SR XenoFree Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# A1099202</td>
</tr>
<tr>
<td>Insulin-Transferrin-Selenium (ITS-G) (100X)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 41400045</td>
</tr>
<tr>
<td>Insulin-Transferrin-Selenium-Ethanolamine (ITS-X) (100X)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 51500-056</td>
</tr>
<tr>
<td>B-27™ Supplement (50X)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 17504044</td>
</tr>
<tr>
<td>Product Name</td>
<td>Manufacturer</td>
<td>Catalog Number</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Bovine Albumin Fraction V (7.5%)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 15260037</td>
</tr>
<tr>
<td>2-Mercaptoethanol</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 21985-023</td>
</tr>
<tr>
<td>Human AB Serum</td>
<td>Valley Biomedical</td>
<td>Cat# HP1022</td>
</tr>
<tr>
<td>Activin A</td>
<td>R&amp;D Systems</td>
<td>Cat# 338-AC/CF</td>
</tr>
<tr>
<td>Heregulinβ-1</td>
<td>Peprotech</td>
<td>Cat# 100-03</td>
</tr>
<tr>
<td>ALK5 Inhibitor II</td>
<td>Enzo Life Sciences</td>
<td>Cat# ALX-270-445</td>
</tr>
<tr>
<td>KGF/FGF7</td>
<td>R&amp;D Systems</td>
<td>Cat# 251-KG</td>
</tr>
<tr>
<td>EGF</td>
<td>R&amp;D Systems</td>
<td>Cat# 236-EG</td>
</tr>
<tr>
<td>Retinoic Acid</td>
<td>Sigma-Aldrich</td>
<td>Cat# R2625</td>
</tr>
<tr>
<td>LDN-193189</td>
<td>Stemgent</td>
<td>Cat# 04-0074</td>
</tr>
<tr>
<td>SANT-1</td>
<td>Sigma-Aldrich</td>
<td>Cat# S4572</td>
</tr>
<tr>
<td>TPB</td>
<td>Calbiochem</td>
<td>Cat# 565740</td>
</tr>
<tr>
<td>Noggin</td>
<td>R&amp;D Systems</td>
<td>Cat# 3344-NG-050</td>
</tr>
<tr>
<td>Wnt3a</td>
<td>R&amp;D Systems</td>
<td>Cat# 1324-WN-010</td>
</tr>
<tr>
<td>3,3',5-Triiodo-L-thyronine sodium salt (T3)</td>
<td>Sigma-Aldrich</td>
<td>Cat# T6397</td>
</tr>
<tr>
<td>TGFβ R1 kinase inhibitor IV</td>
<td>EMD Biosciences</td>
<td>Cat# 616454</td>
</tr>
<tr>
<td>KAAD-Cyclopamine</td>
<td>Toronto Research Chemicals</td>
<td>Cat# K171000</td>
</tr>
<tr>
<td>Heparin</td>
<td>Sigma-Aldrich</td>
<td>Cat# H3149</td>
</tr>
<tr>
<td>TTNPB</td>
<td>Sigma-Aldrich</td>
<td>Cat# T3757</td>
</tr>
<tr>
<td>Zinc Sulfate</td>
<td>Sigma-Aldrich</td>
<td>Cat# Z0251</td>
</tr>
<tr>
<td>TRizol®</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 15596018</td>
</tr>
<tr>
<td>Polyethylenimine (PEI)</td>
<td>Polysciences</td>
<td>Cat# 23966-1</td>
</tr>
<tr>
<td>O.C.T. Compound</td>
<td>Sakura Finetek USA</td>
<td>Cat# 25608-930</td>
</tr>
<tr>
<td>Vectashield Antifade Mounting Medium</td>
<td>Vector Laboratories</td>
<td>Cat# H-1000</td>
</tr>
</tbody>
</table>

**Critical Commercial Assays**

<table>
<thead>
<tr>
<th>Product Name</th>
<th>Manufacturer</th>
<th>Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>MinElute Reaction Cleanup Kit</td>
<td>QIAGEN</td>
<td>Cat#28204</td>
</tr>
<tr>
<td>miRNeasy Mini Kit</td>
<td>QIAGEN</td>
<td>Cat# 74104</td>
</tr>
<tr>
<td>mirVana™ miRNA Isolation Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# AM1560</td>
</tr>
<tr>
<td>Nextera DNA Library Preparation Kit</td>
<td>Illumina</td>
<td>Cat#FC-121-1030</td>
</tr>
<tr>
<td>TruSeq Stranded mRNA Library Preparation Kit</td>
<td>Illumina</td>
<td>Cat# 20020594</td>
</tr>
<tr>
<td>TruSeq Small RNA Library Preparation Kit</td>
<td>Illumina</td>
<td>Cat# RS-200-0012</td>
</tr>
<tr>
<td>RNA 6000 Nano Kit</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-1511</td>
</tr>
<tr>
<td>High Sensitivity D1000 ScreenTape</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5584</td>
</tr>
<tr>
<td>RNA ScreenTape</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5576</td>
</tr>
<tr>
<td>RNA ScreenTape Sample Buffer</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5577</td>
</tr>
<tr>
<td>RNA ScreenTape Ladder</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5578</td>
</tr>
<tr>
<td>Qubit ssDNA assay kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# Q10212</td>
</tr>
<tr>
<td>iScript™ cDNA Synthesis Kit</td>
<td>Bio-Rad</td>
<td>Cat# 1708890</td>
</tr>
<tr>
<td>iQ™ SYBR® Green Supermix</td>
<td>Bio-Rad</td>
<td>Cat# 1708880</td>
</tr>
<tr>
<td>Cytofix/Cytoperm W/Golgi Stop Kit</td>
<td>BD Biosciences</td>
<td>Cat# 554715</td>
</tr>
</tbody>
</table>
### Deposited Data

| RNA-seq and ATAC-seq data | GEO     | GSE115327 |

### Experimental Models: Cell Lines

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Source</th>
<th>RRID</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>WiCell Research Institute</td>
<td>NIHhESC-10-0043, RRID:CVCL_9771</td>
</tr>
<tr>
<td>CyT49</td>
<td>ViaCyte, Inc.</td>
<td>NIHhESC-10-0041, RRID:CVCL_B850</td>
</tr>
<tr>
<td>HEK293T</td>
<td>ATCC</td>
<td>Cat# CRL-3216, RRID:CVCL_0063</td>
</tr>
</tbody>
</table>

### Oligonucleotides

<table>
<thead>
<tr>
<th>Assay ID</th>
<th>Description</th>
<th>Source</th>
<th>Cat#</th>
</tr>
</thead>
<tbody>
<tr>
<td>000377</td>
<td>Taqman miRNA assay; hsa-let-7a</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>002282</td>
<td>Taqman miRNA assay; hsa-let-7g</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>00452</td>
<td>Taqman miRNA assay; hsa-miR127-3p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>00502</td>
<td>Taqman miRNA assay; hsa-miR200a-3p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>002300</td>
<td>Taqman miRNA assay; hsa-miR200c-3p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>000377</td>
<td>Taqman miRNA assay; hsa-let-7a</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>00564</td>
<td>Taqman miRNA assay; hsa-miR-375</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>00268</td>
<td>Taqman miRNA assay; hsa-miR-7-5p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>00436</td>
<td>Taqman miRNA assay; hsa-miR-99b-5p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>001182</td>
<td>Taqman miRNA assay; hsa-miR-124-3p</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>001094</td>
<td>Taqman miRNA assay; RNU control</td>
<td>Thermo Fisher</td>
<td>4427975</td>
</tr>
<tr>
<td>ATTTgaattcagtgcagttctcaggtgagggcggaggctagcggggcgctgtgcagcactgagctcgcggaagaccaggaccaggagatcccgagggcgaccgccaggccccgggccctccgctcccgccccgcgacgagcccctgcacaaaccggacctgagcgttttgttcgttcggctcgcgtgaggcaggggcggcctctcagcaccagcccgggggccggcctgatcgccacgcaggcacctgcggccgcacccgccacccgccatcacaaccgtacgggtggagag</td>
<td>This study</td>
<td>IDT gBlock</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
</tbody>
</table>

**Software and Algorithms**

<table>
<thead>
<tr>
<th>Tool</th>
<th>Version</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAR</td>
<td>2.4.0f1</td>
<td>(Dobin et al., 2013)</td>
</tr>
<tr>
<td>Cufflinks</td>
<td>2.2.1</td>
<td>(Trapnell et al., 2010)</td>
</tr>
<tr>
<td>MACS2</td>
<td></td>
<td>(Zhang et al., 2008)</td>
</tr>
<tr>
<td>Software</td>
<td>Reference</td>
<td>URL</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>BEDTools 2.17.0</td>
<td>(Quinlan and Hall, 2010)</td>
<td><a href="https://bedtools.readthedocs.io/en/latest/content/installation.html">https://bedtools.readthedocs.io/en/latest/content/installation.html</a></td>
</tr>
<tr>
<td>ComBat (part of the sva Bioconductor package)</td>
<td>(Johnson et al., 2007) (Leek, 2014)</td>
<td><a href="http://bioconductor.org/packages/release/bioc/html/sva.html">http://bioconductor.org/packages/release/bioc/html/sva.html</a></td>
</tr>
<tr>
<td>GSEA</td>
<td>(Mootha et al., 2003; Subramanian et al., 2005)</td>
<td><a href="http://www.broad.mit.edu/gsea">http://www.broad.mit.edu/gsea</a></td>
</tr>
<tr>
<td>Metascape</td>
<td>(Tripathi et al., 2015)</td>
<td><a href="http://metascape.org/gp/index.html#main/step1">http://metascape.org/gp/index.html#main/step1</a></td>
</tr>
<tr>
<td>TargetScan</td>
<td>(Agarwal et al., 2015)</td>
<td><a href="http://www.targetscan.org/vert_72/">http://www.targetscan.org/vert_72/</a></td>
</tr>
<tr>
<td>JASPAR</td>
<td>(Khan et al., 2018)</td>
<td><a href="http://jaspar.genereg.net/cgi-bin/jaspar_db.pl">http://jaspar.genereg.net/cgi-bin/jaspar_db.pl</a></td>
</tr>
<tr>
<td>Hocomoco</td>
<td>(Kulakovskiy et al., 2018)</td>
<td><a href="http://hocomoco11.autosome.ru/">http://hocomoco11.autosome.ru/</a></td>
</tr>
<tr>
<td>FIMO</td>
<td>(Grant et al., 2011)</td>
<td><a href="http://meme-suite.org/tools/fimo">http://meme-suite.org/tools/fimo</a></td>
</tr>
<tr>
<td>R 3.5.0</td>
<td></td>
<td><a href="https://cran.r-project.org/">https://cran.r-project.org/</a></td>
</tr>
<tr>
<td>SAMtools 1.3</td>
<td>(Li et al., 2009)</td>
<td><a href="https://github.com/samtools/samtools">https://github.com/samtools/samtools</a></td>
</tr>
<tr>
<td>HOMER 4.10</td>
<td>(Heinz et al., 2010)</td>
<td><a href="http://homer.ucsd.edu/homer/download.html">http://homer.ucsd.edu/homer/download.html</a></td>
</tr>
<tr>
<td>HALO™ Image Analysis Software</td>
<td>PerkinElmer</td>
<td><a href="http://www.perkinelmer.com/product/halo-plus-3-ws-license-cls141255">http://www.perkinelmer.com/product/halo-plus-3-ws-license-cls141255</a></td>
</tr>
<tr>
<td>Adobe Illustrator CS5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adobe Photoshop CS5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplemental References


