βlinc1 encodes a long noncoding RNA that regulates islet β-cell formation and function

Luis Arnes,1 Ildefonse Akerman,2,3 Dina A. Balderes,1 Jorge Ferrer,2,3 and Lori Sussel1

1Department of Genetics and Development, Columbia University, New York, New York 10032, USA; 2Genomic Programming of B-Cells Laboratory, Institut d’Investigacions August Pi i Sunyer (IDIBAPS), Barcelona 08036, Spain; 3Department of Medicine, Imperial College London, London W12 0NN, United Kingdom

Pancreatic β cells are responsible for maintaining glucose homeostasis; their absence or malfunction results in diabetes mellitus. Although there is evidence that long noncoding RNAs (lncRNAs) play important roles in development and disease, none have been investigated in vivo in the context of pancreas development. In this study, we demonstrate that βlinc1 (β-cell long intergenic noncoding RNA 1), a conserved lncRNA, is necessary for the specification and function of insulin-producing β cells through the coordinated regulation of a number of islet-specific transcription factors located in the genomic vicinity of βlinc1. Furthermore, deletion of βlinc1 results in defective islet development and disruption of glucose homeostasis in adult mice.

Supplemental material is available for this article.

Received October 22, 2015; revised version accepted February 1, 2016.

Type 2 diabetes (T2D) arises when pancreatic islets are no longer able to compensate for the increasing demand of insulin to maintain glucose homeostatic conditions. Despite the improvement in current therapies, >500 million people live with T2D, and those numbers are expected to almost double in the next 20 years. Accordingly, it is necessary to increase our understanding of the underlying biology of islet dysfunctions associated with T2D and identify novel druggable targets.

Studies in pancreas development have identified numerous transcription factors and lineage-specific regulatory networks that are required for the specification of the different pancreatic cell types, including acini, ducts, and endocrine cells (Arda et al. 2013). This functionally diverse group of cells arises from a common pool of pancreatic progenitors in two major waves of differentiation termed the primary and secondary transitions. Differentiation of the pancreatic β cells occurs during the secondary transition, and the signaling pathways and regulatory factors involved in this important developmental process have been well characterized (for review, see Fan and Wright 2011). These studies have provided the platform to generate functional β cells in vitro for regenerative medicine and suggested novel therapeutic approaches to prevent and treat β-cell dysfunction.

The study of β-cell biology has been recently complemented by genome-wide transcriptome analysis and characterization of epigenetic modifications, suggesting that non-protein-coding regions of the genome are integral to the transcriptional network regulating development and function of β cells (for review, see Arnes and Sussel 2015). Furthermore, a large number of T2D-associated common variants that were identified in genome-wide association studies map to noncoding genomic regions (Pasquali et al. 2014). Although there is increasing evidence that long noncoding RNAs (lncRNAs) play an important role in development and disease (Batista and Chang 2013), none have been investigated in vivo in the context of pancreas development and β-cell function. In this study, we describe the functional characterization of a newly identified conserved lncRNA termed βlinc1 (β-cell long intergenic noncoding RNA 1). We demonstrated that βlinc1 is required for the proper specification and function of endocrine cells through the coordinated regulation of a number of islet-specific transcription factors located in the genomic vicinity of βlinc1. Furthermore, deletion of βlinc1 results in defective islet development and disrupted glucose homeostasis in the adult. These results have important implications for the identification of novel regulatory mechanisms underlying T2D susceptibility and suggest that ncRNAs could represent novel therapeutic targets for the treatment of diabetes.

Results and Discussion

Characterization of βLINC1

lncRNAs expressed in the pancreatic islet are often highly tissue-specific, associated with clusters of open chromatin, and located in the genomic vicinity of transcription factors involved in β-cell development and/or function (Ku et al. 2012; Moran et al. 2012). We analyzed βLINC1 (formally HI-LNC15), a 6.8-kb post-transcriptionally processed human islet-specific transcript without coding potential (Moran et al. 2012) that is located in a region of open chromatin ~20 kb upstream of NKK2.2, an essential islet homeobox transcription factor gene (Supplemental Fig. 1; Sussel et al. 1998). βLINC1 resides in a large syntenic block located on chromosome 20 in humans and chromosome 2 in mice (Fig. 1A). It is conserved in mammals, with the largest stretches of homology located at the putative promoter region, similar to many characterized lncRNAs (Carninci et al. 2005). This high level of conservation enabled us to identify an orthologous βlinc1 transcript in mice (Fig. 1B). Mouse βlinc1 is also predicted to have no coding potential (CPC score ~0.261; CPAT 0.055 |Supplemental Fig. 2a|, and comparative sequence analysis between the mouse and human βlinc1 transcripts.

© 2016 Arnes et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genesdev.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.
βlinc1 regulates islet β-cell development

βlinc1 RNA is retained in the nuclear fraction of β cells, further suggesting a role for βlinc1 in transcriptional regulation [Fig. 1E]. Although βlinc1 is expressed at relatively low levels, the half-life of the βlinc1 transcript is equivalent to that of Tbp [Supplemental Fig. 4b,c]. This suggests that low transcript levels are not due to the degradation of aberrant transcripts, similar to what has been documented for other low-expressing lncRNAs [Clark et al. 2012].

Although there are limited tools available to predict the function of lncRNAs based on nucleotide sequence or genomic location, there is growing evidence that a subset of nuclear lncRNAs functions locally to regulate neighboring genes [Sauvageau et al. 2013; Vance et al. 2014]. Consistently, siRNA-mediated knockdown of βlinc1 RNA in MIN6 cells resulted in the down-regulation of the adjacent coding gene Nkx2.2 [Fig. 1F], suggesting that the βlinc1 transcript positively regulates the expression of Nkx2.2.

βlinc1 knockout mice are glucose-intolerant

To determine the in vivo function of βlinc1, we generated βlinc1-null mice [Supplemental Fig. 5]. βlinc1+/− and βlinc1−/− mice were viable and fertile. Consistent with the expression of βlinc1 being restricted to the developing endocrine pancreas and adult islet, βlinc1−/− mice were mildly glucose-intolerant compared with βlinc1+/− or βlinc1+/+ littermate controls [Fig. 2A]. Furthermore, βlinc1−/− mice displayed abnormal fasting plasma insulin levels and increased insulin secretion under low-glucose conditions and failed to elevate circulating insulin in response to glucose stimulation [Fig. 2B, Supplemental Fig. 6a]. Interestingly, βlinc1−/− mice were also mildly insulin-resistant despite the absence of βlinc1 expression in muscle and subcutaneous white adipose tissue [Supplemental Fig. 6b,c]. We did not detect additional phenotypes in βlinc1−/− mice, particularly in other tissues where Nkx2.2 is expressed, such as the brain and intestine, which is also consistent with the restricted expression and function of the βlinc1 transcript to the endocrine pancreas.

Endocrine development is affected in βlinc1−/− mice

Given the expression of βlinc1 in the embryonic pancreas, it was possible that aberrant glucose homeostasis was caused by defects in pancreas development. Perinatal βlinc1−/− mice did not exhibit overt defects in pancreas morphology but displayed a 50% reduction in β cells and an increase in the number of somatostatin-expressing cells [Supplemental Fig. 7a,b; data not shown]. We did not detect alterations in islet cell ratios in βlinc1−/− neonates [data not shown]. The changes in islet cell ratios in the βlinc1−/− animals appeared to be initiated during development at the secondary transition. At embryonic day 15.5 [E15.5], the number of insulin-producing β cells was already significantly reduced, and there was an approximately threefold increase in the number of somatostatin-producing cells [Fig. 2C,D; Supplemental Fig. 7c]. There was also a smaller but significant reduction in the number of glucagon- and ghrelin-producing cells [Supplemental Fig. 7d–g]. Changes in the expression of the corresponding hormone genes and β-cell-specific transcription
Consistent with the changes in endocrine cell ratios, these data indicated a general dysregulation of endocrine system development in the absence of $\beta linc1$ [Supplemental Fig. 9a,b]. Notably, we did not detect changes in genes coding for transcription factors involved in the specification and maintenance of pancreas progenitors; however, in the absence of $\beta linc1$, a large number of genes involved in the specification of endocrine progenitors and in the maturation and function of $\beta$ cells were dysregulated [Supplemental Fig. 9c]. This supports a role for $\beta linc1$ in the proper specification of endocrine progenitors during the secondary transition of pancreas development, when the majority of $\beta$-cell differentiation occurs.

Although we did not detect altered islet cell ratios or glucose metabolism defects in the heterozygous $\beta linc1^{+/−}$ pancreata, [Fig. 2a; Supplemental Fig. 7], $\beta linc1^{+/−}$ embryos had reduced $\beta linc1$ transcript and displayed defects in the expression of several islet transcription factors, including Nkx2.2, Pdx6, and Mafb [Fig. 3B]. To identify additional gene expression changes that were not confounded by a change in islet cell type ratios, we performed transcriptome analysis on E15.5 $\beta linc1^{−/−}$ pancreata. Strikingly, there were many genes dysregulated in pancreata lacking one copy of $\beta linc1$ [Fig. 4A; Supplemental Table 1]. This analysis also revealed that, in addition to Nkx2.2, Pdx6, and Mafb, five of the top 10 most significantly dysregulated genes in the $\beta linc1^{−/−}$ mice were also located within an ∼55-Mb region on chromosome 2 [Fig. 4B,C]. qPCR analysis verified that these five down-regulated genes on chromosome 2 were significantly decreased in both $\beta linc1^{+/−}$ and $\beta linc1^{−/−}$ pancreata [Fig. 4D].

Dysregulated genes in $\beta linc1^{+/−}$ localize to the genomic vicinity of $\beta linc1$

To determine whether the genes regulated by $\beta linc1$ were biased based on their genomic location, we positionally mapped all genes that were significantly dysregulated ($P < 0.05$) in $\beta linc1^{+/−}$ pancreata versus the wild-type controls. Remarkably, a disproportionate number of dysregulated genes mapped to a discrete region of chromosome 2, within 40 Mb of the $\beta linc1$ locus. Importantly, this is not due to an increased number of genes that are expressed on chromosome 2 during pancreas development [Supplemental Fig. 10a,b]. Moreover, gene ontology analysis

**Figure 2.** $\beta linc1$ knockout mice are glucose-intolerant and have defects in endocrine specification. [A] $\beta linc1^{−/−}$ 16-wk-old mice are mildly glucose-intolerant. [AUC] Area under the curve. [B] Plasma insulins levels in 16-wk-old $\beta linc1^{−/−}$ mice (green) ($n = 10$) and littermate $\beta linc1^{+/−}$ control mice (blue) ($n = 7$) at 0 and 30 min after glucose injection. [C] Immunofluorescence analysis of pancreas development at E15.5 showing a reduction of $\beta$ cells and an increase of somatostatin-expressing $\delta$ cells in $\beta linc1^{−/−}$ mice compared with $\beta linc1^{+/−}$ and $\beta linc1^{+/+}$ mice. There are no apparent morphological defects in the exocrine compartment, which was visualized by immunostaining against CPA1. Images are representative of $n > 3$ mice. Bar, 50 µm. [D] Quantification of hormone-producing cells in E15.5 pancreata. $n = 4$. Error bars represent ±SEM. (∗) P < 0.05 $\beta linc1^{−/−}$ versus $\beta linc1^{+/+}$; (#) P < 0.05 $\beta linc1^{−/−}$ versus $\beta linc1^{+/−}$; Student’s t-test.

$\beta linc1$ deficiency affects the expression of several genes related to endocrine cell differentiation and $\beta$-cell function

To identify the pathways regulated by $\beta linc1$, we performed global transcriptome analysis on E15.5 $\beta linc1^{−/−}$ versus $\beta linc1^{+/−}$ pancreata [Gene Expression Omnibus [GEO] accession no. GSE73711] [Supplemental Table 1].

**Figure 3.** $\beta linc1$ regulates the expression of endocrine-specific genes. [A] qPCR analysis of hormones in E15.5 pancreata. General reduction of hormone expression with the up-regulation of somatostatin. $n ≥ 5$. [B] qPCR analysis of $\beta linc1$ and several transcription factors involved in pancreas development in E15.5 pancreata. $n ≥ 5$. Error bars represent ±SEM. (∗) P < 0.05 $\beta linc1^{−/−}$ versus $\beta linc1^{+/−}$; (#) P < 0.05 $\beta linc1^{−/−}$ versus $\beta linc1^{+/+}$; Student’s t-test.
βlinc1 regulates islet β-cell development

showed that the dysregulated genes on chromosome 2 are associated with endocrine development and islet morphology. Ingenuity Pathway Analysis also indicated that the predominant upstream regulators affected by loss of one allele of βlinc1 are the known islet-specific transcription factors Pdx1 and NeuroD1. These pancreas-related categories are lost if the same analysis is performed with all of the genes that are located in this region of chromosome 2 (Supplemental Fig. 10a,b).

The remarkable observation that many βlinc1-regulated genes were located on the same chromosome as βlinc1 prompted us to investigate whether the observed gene expression changes were due to loss of βlinc1 transcript rather than the deletion of a DNA enhancer element, similar to siRNA- and shRNA-mediated knockdown experiments of the lincRNA in mouse MIN6 cells [Figs. 1F, 4E] and human EndoC-βH1 cells recapitulated some of the gene changes observed in vivo. Furthermore, the βlinc1 genomic locus does not appear to have enhancer activities in MIN6 cells (Supplemental Fig. 3); however, limitations of these in vitro studies prevented us from ascribing all of the βlinc1 phenotypes to loss of the βlinc1 transcript. Interestingly, although we demonstrated that the βlinc1 transcript regulates several important islet transcription factors, only a subset of the βlinc1−/− phenotypes could be attributed to reduction of Nkx2.2, Pax6, and/or MafB expression (Supplemental Table 2; Pan and Wright 2011), suggesting that βlinc1 is involved in novel lineage regulatory pathways.

Transcriptome analysis of pancreata lacking one copy of βlinc1 also revealed the dysregulation of a set of functionally related β-cell genes located in discrete genomic regions. It is particularly intriguing that βlinc1 specifically regulates three essential islet transcription factors, ChGB, and additional β-cell genes on chromosome 2. Long-range interchromosomal and intrachromosomal interactions through chromosomal looping and transcription factories are known to be involved in the transcriptional coregulation of functionally related genes in a cell type-specific manner (Schoenfelder et al. 2010; Papantoni et al. 2012). For example, insulin expression and calcium signaling have been shown to be coupled through a long-range chromosomal interaction that spans >65 Mb [Xu et al. 2014]. IncRNAs have also been implicated in higher-order gene regulation through chromatin organization, and there is evidence of IncRNAs being highly enriched at transcription “factories” (Caoudr-Herger et al. 2015). Therefore, it is possible that βlinc1 coordinately regulates β-cell gene expression through the structural organization of the chromatin in β cells. Analysis of the known chromatin features and topologically associated domains (TADs) in and around βlinc1 failed to reveal any obvious domains or features that could explain the regulation of the genes on chromosome 2. HiC analyses suggested that βlinc1 may be located at the edge of a TAD (Supplemental Fig. 11; Shen et al. 2012); however, the presence of this domain does not correlate with the observed in vivo gene expression changes. For example, the TAD data include all of the genes located between Insm1 and Foxa2, but only two of the 13 genes within this domain [Nkx2.2 and Insml] are down-regulated in βlinc1−/− mice.

Although we are just beginning to understand a role for IncRNAs in the cell, many findings have suggested that IncRNAs provide a layer of cell-specific gene regulation that contributes to cellular diversity [Mattick 2001]. It is therefore intriguing that βlinc1 has such a limited expression domain, and the phenotype of the global βlinc1−/− mice is primarily restricted to β-cell formation and function. Furthermore, our discovery of βlinc1 as a novel islet-specific transcriptional regulator has important implications for understanding β-cell biology and suggests that ncRNAs could represent novel therapeutic targets for the treatment of diabetes.
Materials and methods

Generation of the \textit{linc1} knockout allele

The \textit{linc1} knockout allele was generated using the Recombinease Mediate Cassette Exchange protocol as previously described (Arnes et al. 2012) with some modifications. Short arms (≤500 base pairs) homologous to the flanking region of \textit{linc1} and to a downstream region (2.9 kb) were generated by PCR and cloned into pLC.A.71/2272 and pMCS-DTA. The BAC clone (RP23-236P19) was modified in two consecutive recombineering steps in SW106 cells: The \textit{linc1} sequence was replaced with the \textit{paztk-EM7-kan} cassette, and the DTA-\textit{Amp} cassette was inserted 2.9 kb downstream from \textit{linc1}, resulting in the replacement of 1 kb of genomic DNA (Supplemental Fig. 5b). Positive clones were validated by PCR and DNA sequencing, and a correctly modified BAC was electroporated into mouse embryonic stem cells at Columbia University (Herbert Irving Transgenic Center) were passaged and maintained following standard techniques provided by Dr. Jeffrey Zigman, University of Texas Southwestern Medical Center (provided by Dr. Jeffrey Zigman, University of Texas Southwestern Medical Center) were passaged and maintained following standard techniques in 5% CO\textsubscript{2} and 95% air. MIN6 cells were transfected with 10 nM siRNA targeting \textit{linc1} and a scrambled control [Silencer Select, Ambion] using Lipoconnectamine 2000 following the manufacturer’s instructions (Life Technologies). siRNAs targeting \textit{linc1} were designed using the algorithm provided by the manufacturer, and sequences are listed in Supplemental Table 3.

**Enhancer activity**

A 4.2-kb fragment of the \textit{linc1} genomic locus [chr2: 147,030,443–147,034,638, mm9] was cloned into the PGL4.27 luciferase vector. One microgram of the experimental vector PGL4.27-\textit{linc1} and the positive controls (C-DRK2BAS enhancer, Ins2 promoter, and NeuroD promoter/enhancer) were individually cotransfected with 0.1 µg of pRL into MIN6 cells in triplicate. Luciferase activity was measured after 48 h. PGL4.23-CDKN2BAS was a gift from Jorge Ferrer [Addgene plasmid no. 60296]. PGL3-Ins2 and PGL3-NeuroD have been previously described (Raum et al. 2006; Anderson et al. 2009). Luciferase values were normalized to Renilla activity to account for transfection efficiencies and were expressed as fold increase over the empty vector.

**RNA in situ hybridization**

The \textit{linc1} probe was generated by PCR from E15.5 pancreas cDNA and cloned into pCRII-TOPO (Life Technologies). Sense and antisense probes were labeled with the DIG RNA-labeling mix (Roche Applied Science). RNA in situ hybridization was performed as previously described [Mastracci et al. 2013], including a short treatment of proteinase K digestion.

**Immunofluorescence**

Tissue processing and immunofluorescence analysis were performed as previously described (Arnes et al. 2012). Primary antibodies are listed in Supplemental Table 3. DAPI (1:1000; Invitrogen) was applied for 30 min following secondary antibody incubation: DyLight-488, DyLight-549, DyLight-649, Alexa-488, and Alexa-647 [Jackson Immunoresearch]. Images were acquired with either an epifluorescence (Leica DM5500) or a confocal (Zeiss LSM 710) microscope.

**Morphometric analysis and cell counting**

Morphometric analysis was performed in E15.5 and postnatal day 2 pancreata. The entire organ was sectioned, and at least six evenly distributed sections were analyzed. The number of endocrine cells was determined relative to the total pancreatic area. Pancreas area was quantified using Image Pro Plus 5.0.1 software [Media Cybernetics].

**Cellular fractionation**

MIN6 cells were grown to confluency, detached by trypsinization, and pelleted. Half of the pellet was used for total RNA isolation, and the other half was used for nuclear and cytoplasmic isolation using the PARIS kit (Ambion) following the manufacturer’s instructions.

**RNA extraction and quantitative RT–PCR analysis**

Total RNA was isolated and analyzed as previously described [Arnes et al. 2012]. Primers and probes are listed in Supplemental Table 3.

**DNA sequencing (RNA-seq) analysis**

Total RNA from \textit{linc1}−/−, \textit{linc1}+/−, and \textit{linc1}+/+ mice was converted into cDNA libraries (TruSeq RNA sample preparation kit version 2, Illumina) using poly-A pull-down for mRNA enrichment. Sequencing was performed to a depth of 30 million pairs in three biological replicates per condition. Differential expression between replicates was assessed using DESeq [R package]. All samples had RNA integrity (RIN) values >9.0 as determined with Agilent Bioanalyzer 2100. Complete RNA-seq data are available through GEO accession number GSE73711.

**TADs**

TADs in the genomic vicinity of \textit{linc1} were identified from the mouse encode project at the Ren laboratory (http://chromosome.sdsc.edu/mouse).

**Statistics**

All data were expressed as mean ± SEM. Mean and SEM values were calculated from at least three biological replicates. The statistical significance was measured by two-tailed Student’s \textit{t}-test.

**Acknowledgments**

We thank the Sussel laboratory for critical discussions throughout the project and reading of the manuscript. We are especially grateful to Chyuan-Sheng [Victor] Lin for assistance with the generation of the \textit{linc1} knockout mouse, Jiguang Wang for computational support, and Jeffrey Zigman for providing the PG1 cells. The project was initiated with the National Institutes of Health B-Cell Biology Consortium collaborative project grant [U19 DK072473] to L.S. and J.F. Additional support was provided by Juvenile Diabetes Research Foundation Grant [CU13-252-L.A.], the Russell Berrie Foundation [L.A.], the Foundation for Diabetes Research [L.S.], and Juvenile Diabetes Research Foundation Grant [117-2012054-5]. L.S. and J.F. conceived the study. L.A. and J.F. collected and analyzed the data on human \textit{linc1}. L.A. generated and processed the data on mouse \textit{linc1} with assistance from D.A.B. L.S. and L.A. designed the study, analyzed the data, and wrote the manuscript. All authors discussed the results and edited the manuscript.
