Could lncRNAs contribute to β-cell identity and its loss in Type 2 diabetes?

Timothy J. Pullen1 and Guy A. Rutter
Section of Cell Biology, Department of Medicine, Imperial College London, Exhibition Road, London SW7 2AZ, U.K.

Abstract
The progression of Type 2 diabetes is accompanied by diminishing islet β-cell mass and function. It has been proposed that β-cells are lost not only through apoptosis, but also by dedifferentiating into progenitor-like cells. There is therefore much interest in the mechanisms which define and maintain β-cell identity. The advent of genome-wide analyses of chromatin modifications has highlighted the role of epigenetic factors in determining cell identity. There is also evidence from both human populations and animal models for an epigenetic component in susceptibility to Type 2 diabetes. The mechanisms responsible for defining the epigenetic landscape in individual cell types are poorly understood, but there is growing evidence of a role for lncRNAs (long non-coding RNAs) in this process. In the present paper, we discuss some of the mechanisms through which lncRNAs may contribute to β-cell identity and Type 2 diabetes risk.

Introduction
Diabetes is a large and growing problem affecting 347 million people worldwide [1]. T2D (Type 2 diabetes) comprises 90% of this epidemic, and disease risk is determined by a combination of genetic susceptibility and environmental factors. The major mortality and morbidity caused by diabetes is due to the various complications of the disease, including neuropathy, cardiovascular disease, damage to peripheral circulation and blindness. However, these are all consequences of the primary defect which is a failure to correctly regulate blood glucose and lipid levels.

β-Cells, located in the pancreatic islets, are the body’s sole source of insulin. Insulin stimulates the uptake of glucose from the bloodstream by tissues including skeletal muscle and is therefore critical for regulating blood glucose levels. Obesity-induced insulin resistance suppresses hepatic glucose output and thus increases the body’s demand for insulin. Islets compensate for this through increases in β-cell mass and function [2]. However, limitations in the capacity of islets to adequately compensate results in an inability to maintain blood glucose levels within the normal range. A combination of the high demand for insulin coupled with the toxic effects of increased glucose and lipid levels on β-cells (glucolipotoxicity) then contribute to decreases in β-cell mass and β-cell function resulting in T2D [3]. Whereas the relative contributions of the loss of β-cell mass and function to diabetes are debated, a further possibility that has been proposed is that a loss of β-cell identity contributes to the loss of functional β-cell mass. Jonas et al. [4] reported that chronic hyperglycaemia led to a loss of the pattern of gene expression necessary for functional β-cells, including decreases in several transcription factors necessary for maintaining differentiated β-cells. More recently, lineage-tracing studies in transgenic mice lacking FoxO1 (forkhead box O1) specifically in β-cells, showed that the loss of β-cell mass was accompanied by dedifferentiation of β-cells to progenitor-like cells expressing Neurogenin3 [5]. The authors reported a similar effect in animal models of diabetes and proposed this as a model for decreases in β-cell mass in diabetic humans. This highlights the importance of understanding the mechanisms defining β-cell identity. Furthermore, the possibility of modulating these mechanisms to reinforce β-cell identity may offer a new approach for treating T2D.

What defines a β-cell?
β-Cells sense glucose not through a cell-surface receptor, but through its metabolism [6,7]. The consequent increase in cytosolic ATP concentration ultimately triggers the release of insulin stored in granules via a well-defined

Key words: β-cell, chromatin, epigenetics, islet, long non-coding RNA (lncRNA), Type 2 diabetes.
Abbreviations used: GWAS, genome-wide association study; IUGR, intrauterine growth restriction; lncRNA, long non-coding RNA; mRNA, microRNA; NAT, natural antisense transcript; PRC2, polycomb repressive complex 2; SNP, single nucleotide polymorphism; T2D, Type 2 diabetes.

1To whom correspondence should be addressed (email t.pullen@imperial.ac.uk).
pathway involving closure of ATP-sensitive K+ channels [8]. Whereas most cells regulate ATP production in response to energy demand, β-cells have a highly specialized metabolic system which makes ATP levels proportional to circulating glucose concentrations. This is achieved through the specific activation of a number of key genes such as GLUT2, encoding a high-capacity plasma membrane glucose transporter [9], and GCK, encoding glucokinase [10] which regulates the first step of glycolysis. Alternative metabolic pathways which could redirect metabolites from producing ATP or allow other circulating metabolites to increase ATP levels are suppressed through specific repression of the corresponding genes [11–13], which are often widely expressed across other tissues [14,15]. Strikingly, forced up-regulation of one of these, the pyruvate/lactate transporter Slc16a1 (Mct1) mimics the effect of exercise-induced hyperinsulinism wherein muscle-derived pyruvate inappropriately stimulates insulin release during exercise, demonstrating a full physiological role of repressing these genes [16].

This β-cell-specific pattern of gene expression is partly achieved through the particular complement of transcription factors found in mature β-cells including Pdx1, Hnf1α and NeuroD1 [17]. However, there is also evidence that other regulatory systems are involved. For example, we have shown previously that islet-enriched miRNAs (microRNAs) contribute to the β-cell-specific silencing of Slc16a1 [18]. Genome-wide maps of chromatin modifications from purified β-cells also revealed H3K27me3 (me3 indicates trimethylation) marks around this and other genes silenced specifically in β-cells [19], indicating an epigenetic contribution to the β-cell-specific pattern of gene expression. H3K27me3 is a ‘repressive’ chromatin mark and is indicative of polycomb-mediated silencing of gene expression. H3K27me3 is also found at other genes normally silenced specifically in β-cells, including Arx, Hk1 and Rest. In contrast, H3K4me3 is associated with actively transcribed genes, and in β-cells is found at genes whose expression is important for β-cell function such as Pdx1 and Pax6 [19]. This β-cell-specific pattern of chromatin modifications contributes to the β-cell-specific pattern of gene expression and thus to β-cell identity.

An epigenetic component to T2D susceptibility

There is evidence to suggest that epigenetic factors contribute not only to healthy β-cell identity, but also to β-cell dysfunction in T2D. For example, DNA methylation at the imprinted ZAC/HYMAI locus is required for normal insulin release in infants [20]. A further example of an epigenetic component to T2D susceptibility arose from the tragic events of the Dutch Hunger Winter. During a well-defined period between 26 November 1944 and 12 May 1945, part of the Dutch population suffered a famine where daily nutrition for most individuals fell to between 400 and 800 kcal (1 kcal = 4.184 kJ) per day [21]. When expectant mothers were exposed to the famine during the first trimester of pregnancy, there was no clear effect on the birthweight of their children. However, in later life, those exposed to the restricted nutrition in the womb had a greater incidence of T2D [22]. The effects of the intrauterine conditions were stably recorded throughout the lifetime of those individuals only to become apparent decades later. This is an example of mitotically stable changes which do not affect the DNA sequence of the genome and are thus classed as epigenetic changes.

This striking effect has been recapitulated in various animal models of restricted nutrition in utero. For example, in the IUGR (intrauterine growth restriction) model in rats, bilateral uterine artery ligation restricts blood supply to the developing fetus. This produces a ‘diabetic’ phenotype which develops at the age of 15–26 weeks [23]. This presents as a combination of insulin resistance and impaired β-cell function. The animal model also gives an insight into the molecular mechanism underlying this phenomenon: progressive epigenetic silencing of the key β-cell transcription factor Pdx1. The Pdx1 promoter in healthy β-cells is decorated with H3K4me3, typical of active promoters [19]. In IUGR islets, the first effect is a reduction in histone acetylation, which is followed by a reduction in H3K4me3 and an increase in H3K9me2 (me2 indicates dimethylation), a repressive chromatin mark. These combine to decrease Pdx1 transcription, and this is later reinforced through increased DNA methylation at this locus.

Could the down-regulation of this single gene really be responsible for this diabetic phenotype? Pdx1 clearly plays an important role in both islet development and function. Total ablation of Pdx1 expression in mice results in pancreatic agenesis, showing Pdx1 to be essential for pancreatic development. However, even mice heterozygous for the Pdx1-null allele have impaired glucose tolerance with reports of both impaired glucose-stimulated insulin secretion [24] and β-cell mass reduction due to increased apoptosis [25]. This demonstrates the importance of maintaining Pdx1 expression in mature β-cells. Similarly, missense mutations in human PDX1 are responsible for MODY4 (a form of maturity-onset diabetes of the young), in which haploinsufficiency in PDX1 produces β-cell secretory defects [26]. Given the importance of maintaining the level of Pdx1 expression, epigenetic down-regulation is entirely consistent with the β-cell defects observed in the IUGR rat. This highlights further the importance of the epigenetic landscape in regulating gene expression to maintain β-cell identity.

IncRNAs (long non-coding RNAs) and cell identity

The β-cell-specific pattern of chromatin modifications revealed by genome-wide studies, underlies β-cell identity. Yet the mechanisms by which they are laid down and maintained are far from clear. The protein complexes responsible for many of the histone modifications have been elucidated, such as the PRC2 (polycomb repressive complex 2) which mediates trimethylation of H3K27. However, these complexes lack intrinsic DNA-binding capability and the mechanisms that target them to particular genomic loci...
in a cell-type-specific manner are not known. Evidence is mounting for a role for lncRNA in this process of defining the cell-type-specific epigenetic landscape.

RNA has traditionally been viewed as an intermediate step between the DNA of the genome and functional proteins. Yet advances in sequencing technologies have revealed an increasing number of RNA molecules which do not appear to encode proteins. High-throughput sequencing of RNA (RNA-Seq) presents an unbiased view of the transcriptome compared with microarrays which are generally restricted to analysing previously annotated features [27]. Genome-wide maps of histone modifications have also allowed genes to be identified by the pattern of modifications typically found at the start, and along the length of actively transcribed genes [28]. These two approaches have identified thousands of novel transcripts which lack any protein-coding capability. In their comprehensive investigation into the functional human genome, investigators of the ENCODE project have catalogued over 9640 lncRNA loci and have systematically analysed the epigenetic landscape around, and expression from these loci across a range of cell lines [29].

Many of these transcripts have a similar structure to protein-coding mRNAs: they are transcribed by RNA Pol (polymerase) II, and are capped, spliced and polyadenylated. lncRNAs can be classified according to the structure of the loci from which they are expressed. Those expressed from regions away from other genes have been termed lincRNAs (long intergenic RNAs) [28]. However, many lncRNAs overlap neighbouring genes, often in an antisense orientation. These antisense transcripts can be divided further into those with overlapping exons and those in which the exons of one transcript only overlap intronic sequences of the other.

Although lncRNAs and mRNAs share many features, their patterns of expression are strikingly different, with lncRNAs being expressed in a far more cell-type-specific manner than protein-coding mRNAs. This observation, initially made after analysis of a smaller dataset [27], has since been confirmed by the more comprehensive ENCODE data. Whereas 55.8% of protein-coding genes were expressed across all the cell lines investigated, only 10.7% of lncRNA genes were. In contrast, expression of nearly a quarter of lncRNA genes (23.5%) was restricted to just one cell line, whereas this was the case for approximately one in 20 protein-coding genes (5.7%) [29]. Thus the pattern of lncRNAs expressed is highly cell-type-specific with many being expressed in just a single cell type.

**Mechanisms of lncRNA action**

The restricted distribution of many lncRNAs makes them well placed to perform highly cell-type-specific tasks such as regulating the epigenetic landscape and pattern of gene expression which determine cell identity. One piece of evidence supporting such a role is the observation that many lncRNAs interact with chromatin-modifying complexes [30]. Furthermore, lncRNAs have been proposed to recruit chromatin-modifying complexes to particular genomic loci providing a model for highly specific changes to the chromatin landscape. For example, the lncRNA HOTAIR has been shown to recruit PRC2 to particular loci throughout the genome [30]. In addition to a role in targeting protein complexes, lncRNAs may also act as scaffolds for the assembly of different combinations of chromatin-modifying complexes, transcription factors and other proteins. Indeed, these two functions may be overlapping, as in the case of HOTAIR, which recruits both the LSD1 (lysine-specific demethylase 1)–CoREST (co-repressor for element-1-silencing transcription factor) complex and the PRC2 complex [31]. The combined effect of this is to both demethylate H3K4 and methylate H3K27 which both contribute to the down-regulation of gene expression. A growing number of other lncRNAs have also been shown to target chromatin-modifying complexes including Air, Kcnq1ot1 and Xist [32–34]. This is clearly one of the major mechanisms through which lncRNAs act; however, we await detailed functional data on many more lncRNAs before we can appreciate what contribution this makes to defining the epigenetic landscape and determining cell identity.

In addition to a role at the transcriptional level, lncRNAs can also regulate gene expression post-transcriptionally. lncRNAs and mRNAs with overlapping exons transcribed in opposite orientations will produce mature transcripts with a stretch of perfectly complementary sequence. These have been termed NATs (natural antisense transcripts) and they have the potential for direct interactions between the complementary regions of the transcripts. In the case of the gene encoding human β-secretase 1 (BASE1), an overlapping antisense transcript (BASE1-AS) has been shown to increase mRNA stability [35]. However, perfect complementarity between lncRNAs and mRNAs is not required for direct interactions, as demonstrated by the discovery that imperfect base-pairing between lncRNAs and mRNAs could recruit Staufen 1 (STAU1) leading to mRNA degradation [36]. One further mechanism of post-transcriptional regulation is through acting as ‘sponges’ for miRNAs, thus preventing them from interacting with their target mRNAs (e.g. linc-MD1 [37]). lncRNAs can therefore exert both positive and negative post-transcriptional regulation on mRNA expression. lncRNAs have also been proposed to perform a number of other functions including assembling nuclear structures (NEAT1 [38,39]), and regulating alternative mRNA splicing (Malat1 [40,41]).

A further distinction to be made in the action of lncRNAs is whether they act locally to affect expression of nearby genes, or act globally throughout the genome. NATs are most likely to act locally on the transcript with which they overlap. However, lncRNAs interacting with mRNA with imperfect complementarity are likely to act on mRNA transcribed from distant loci. Regulation at the transcriptional level has also been found to occur both locally (e.g. ANRIL, Air and Kcnq1ot1) and globally (e.g. HOTAIR). lncRNAs expressed from loci overlapping protein-coding genes often appear to regulate that gene, whereas a large-scale study of intergenic
IncRNAs found that they primarily affected gene expression in trans [42]. IncRNAs are clearly a mixture of different classes which act through different mechanisms and with different scopes. So far, the functional characterization of individual IncRNA functions has lagged behind the discovery of novel IncRNA genes. However, as the depth of experimental evidence for IncRNA functions mounts, it is hoped that this mass of IncRNAs may be organized into groups with more tightly defined mechanisms of action.

IncRNAs and T2D

As discussed above, there is a clear role for epigenetic factors regulating β-cell identity and diabetes risk. There is also clear evidence that IncRNAs affect the gene expression in part through altering the chromatin landscape. Can these be combined to show a role for IncRNAs in determining β-cell identity and to participate in the misregulation of gene expression during T2D? One source of evidence for such a role is the GWASs (genome-wide association studies) searching for T2D susceptibility loci [43]. Several SNPs (single nucleotide polymorphisms) associated with increased risk localize to an IncRNA expressed from the INK4 locus, called ANRIL [44]. The INK4 locus encodes three tumour-suppressor genes, p15INK4b, p14ARF and p16INK4a, which inhibit cell-cycle progression by inhibiting cyclin-dependent kinases 4/6. ANRIL has been shown to recruit PRC2 to the p15INK4b locus, down-regulating its expression [45]. SNPs which alter the expression or function of ANRIL would likely have effects on the proliferative capacity of β-cells, potentially increasing diabetes susceptibility through limiting the capacity for compensatory increases in β-cell mass. GWASs identified a further SNP mapping to a known IncRNA, Kcnq1ot1 [46].

In addition to these previously identified IncRNAs, a genome-wide search for human islet-expressed IncRNAs has identified >1100 IncRNAs, with 55% of the intergenic IncRNAs being islet-specific [47]. Interestingly, of 55 T2D susceptibility loci, nine contained an islet IncRNA within 150 kb of the main SNP. This indicates that misregulation of IncRNAs may explain part of the heritable susceptibility to T2D.

Although these data give a tantalizing glimpse into the potential for IncRNAs to regulate β-cell identity and dysfunction, their relative importance will only become apparent when more of the >1000 islet IncRNAs have been functionally characterized. In the meantime, it is tempting to speculate that their highly cell-type-specific expression makes them attractive potential drug targets with limited off-target effects in other cell types.

References

15. Thorez, L., Laudadio, I., Van Deun, K., Quintens, R., Hendrickx, N., Servitja, J.M. and Ferrer, J. (2004) Transcriptional networks controlling β-cell identity and diabetes risk. There is also evidence for lncRNA functions mounts, it is hoped that this mass of IncRNAs may be organized into groups with more tightly defined mechanisms of action.

Funding

T.J.P. is funded by a Non-Clinical Research Fellowship from the Diabetes Research and Wellness Foundation. G.A.R. is supported by a Senior Investigator Award from the Wellcome Trust.


34 Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Shandier, O., Chinappi, M., Tramontano, A. and Bozzoni, I. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 147, 358-369


Received 18 December 2012 doi:10.1042/BST20120355