Non-coding genome functions in diabetes

Inês Cebola[1] **and Lorenzo Pasquali**[2,3,4]

[1]Department of Medicine, Imperial College London, London W12 0NN, UK
[2]Division of Endocrinology, Germans Trias i Pujol University Hospital and Research Institute, 08916 Badalona, Spain
[3]Josep Carreras Leukaemia Research Institute, 08916 Badalona, Spain
[4]CIBER de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), 08036 Barcelona, Spain

Correspondence should be addressed to L Pasquali
Email
lpasquali@igtp.cat

Abstract

Most of the genetic variation associated with diabetes, through genome-wide association studies, does not reside in protein-coding regions, making the identification of functional variants and their eventual translation to the clinic challenging. In recent years, high-throughput sequencing-based methods have enabled genome-scale high-resolution epigenomic profiling in a variety of human tissues, allowing the exploration of the human genome outside of the well-studied coding regions. These experiments unmasked tens of thousands of regulatory elements across several cell types, including diabetes-relevant tissues, providing new insights into their mechanisms of gene regulation. Regulatory landscapes are highly dynamic and cell-type specific and, being sensitive to DNA sequence variation, can vary with individual genomes. The scientific community is now in place to exploit the regulatory maps of tissues central to diabetes etiology, such as pancreatic progenitors and adult islets. This giant leap forward in the understanding of pancreatic gene regulation is revolutionizing our capacity to discriminate between functional and non-functional non-coding variants, opening opportunities to uncover regulatory links between sequence variation and diabetes susceptibility. In this review, we focus on the non-coding regulatory landscape of the pancreatic endocrine cells and provide an overview of the recent developments in this field.

Key Words

- diabetes (all)
- gene regulation
- pancreatic β cell
- diabetes II
- islet cells

Journal of Molecular Endocrinology
(2016) **56**, R1–R20

Introduction

The prevalence of diabetes mellitus is increasing globally, nowadays assuming the dimensions of a pandemic with more than 500 million predicted to be affected worldwide by 2035 (Guariguata *et al*. 2014). Type 2 diabetes (T2D) is the most prevalent form accounting for >90% of all causes of diabetes. T2D is characterized by decreased insulin sensitivity and defective insulin secretion. The resulting elevated blood glucose levels eventually lead to microvascular damage, making T2D a leading cause of blindness, neuropathy, heart disease, and end-stage renal disease. Even when multiple antidiabetic treatments are applied, blood glucose levels still fluctuate significantly in diabetic patients, making diabetes the sixth leading cause of death in the United States (Jemal *et al*. 2005).

As a prototype of a multifactorial complex disease, T2D arises from an intricate interaction of environmental factors and inherited predisposition. While a sedentary lifestyle and high calorie food intake are well-established risk factors for T2D, family-based and association studies have shown that genetic factors also contribute to disease susceptibility (Köbberling & Tillil 1990, Bell & Polonsky 2001). Accordingly, family history is an important risk

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

factor for this disease. Siblings with T2D confer a four- to sixfold increase in risk (Florez *et al*. 2003). Furthermore, studies with monozygotic twins revealed a concordance rate in the range of 50–92%, whereas similar studies with dizygotic twins showed a much lower concordance (Beck-Nielsen *et al*. 2003). Interestingly, the genetic component of T2D is also supported by studies showing that abnormal glucose homeostasis is also heritable (Poulsen *et al*. 1999). Unfortunately, knowledge about the molecular mechanisms linking genetic variation and environmental factors with diabetes is still limited, which often frustrates attempts to separate individual propensity to develop T2D between the many genetic and environmental components.

In this context, epigenetics may play an important role in interfacing the molecular response of an organism to environmental exposures, orchestrating and modulating tissue- and cell-specific gene expression patterns. Different environmental exposures during development and later on in life can influence disease susceptibility (extensively reviewed in Jiang *et al*. 2013). Hence, understanding the epigenetic processes in the context of T2D will likely shed light on the molecular mechanisms underlying the development and progression of the disease.

Human genetics was shown to be a powerful tool in unmasking disease molecular mechanisms. In several cases of Mendelian diabetes, studies on individuals, families, and closed populations allowed human geneticists to successfully map causal mutations to protein-coding gene regions. For example, mutations in different transcription factors involved in pancreas development, such as *PTF1A* (Sellick *et al*. 2004), *PDX1* (Stoffers *et al.* 1997), *GATA4* (Shaw-Smith *et al*. 2014), *GATA6* (Lango Allen *et al*. 2012), *NKX2.2*, and *MNX1* (Flanagan *et al.* 2014), are now known to cause neonatal diabetes mellitus. Similarly, human genetics uncovered mutations in a handful of genes, including *GCK*, *HNF1B,* or *NEUROD1,* that lead to maturity-onset diabetes of the young (MODY) (reviewed in Siddiqui *et al*. 2015). The identification of these causal mutations allowed the unmasking of key β-cell regulators, opening avenues to the understanding of the molecular mechanisms that control the normal physiology of insulin secreting cells and also allowing clinicians to adequate their therapeutic approaches to the patients (Vaxillaire *et al*. 2012, Siddiqui *et al*. 2015). In these instances, however, the identification of the causality was only possible with the access to affected families segregating highly penetrant rare variants.

In contrast with Mendelian diabetes, T2D is characterized by the contribution of several low penetrant alleles, hence, requiring different methods of study. Genome-wide association studies (GWAS) aim to establish statistical evidence for the association of particular variants with the disease by comparing the genetic traits of large numbers of affected and non-affected individuals. Although GWAS have greatly contributed to the identification of loci associated with T2D, their statistical power is limited by the number of individuals analyzed and by the frequency of variants in the population. Consequently, to date, most studies have only uncovered common variants of small effect size (reviewed in McCarthy *et al*. 2008). These account, even in combination, for at most 5–10% of overall trait variance (Willems *et al*. 2011) and therefore, perhaps, 10–20% of overall heritability. This is still far from a useful platform for disease prediction. We expect that larger cohort studies and GWAS meta-analyses will attribute part of this “missing heritability” to rare variants with intermediate penetrance in the near future (McCarthy *et al*. 2008). Rare variants with stronger effects might account for a fraction of the heritability (Schork *et al*. 2009); however, their identification with GWAS might remain challenging.

Interestingly, most of the variants identified in GWAS, including T2D-associated variants, do not lie in coding regions (Maurano *et al*. 2012, Gusev *et al*. 2014), suggesting that risk variants might affect non-coding elements of the genome, having an impact that is either transcriptional or post-transcriptional rather than altering the sequence of the protein itself.

The assumption that GWAS variants might have a role in transcriptional regulation seems reasonable after large consortia projects, such as the ENCODE and the Epigenome Roadmap, uncovered that a large proportion of the human genome is populated by regulatory elements (ENCODE Project Consortium *et al*. 2012, Roadmap Epigenomics Consortium *et al*. 2015). Applying state-of-the-art techniques coupled with high-throughput sequencing, different consortia and individual laboratories have profiled accessible chromatin, relevant histone modifications, and transcription factor binding sites in an unbiased genome-wide manner for an array of human tissues and cell types, including pancreatic islets of Langerhans. These studies have allowed the tissue-specific mapping of key regulatory elements, such as enhancers, promoters, or insulators. These regulatory elements modulate gene expression in *cis* by binding different sets of transcription factors and chromatin remodelers. Noteworthy, researchers have observed that enhancers tend to

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

cluster in large domains of active chromatin (Gaulton *et al.* 2010, Hnisz *et al.* 2013, Parker *et al.* 2013, Pasquali *et al.* 2014) and that they regulate essential tissue functions, defining genetic programs associated with cellular identity. Although our understanding of genome regulation is currently insufficient to exploit the available genetic findings, annotation of tissue-specific enhancers might help identify causal sequence variants.

In addition to regulatory elements, functional GWAS variants can affect other modulators of gene expression such as non-coding RNAs (ncRNAs). In this view, a variety of ncRNAs, including long non-coding RNAs (lncRNAs) and microRNAs (miRNAs), now starts to be appreciated in pancreatic islets (Morán *et al.* 2012, Nica *et al.* 2013, Kameswaran *et al.* 2014). Furthermore, some GWAS variants might impact gene expression post-transcriptionally. For example, a growing body of evidence shows that GWAS variants can destroy or create miRNA binding sites, hence, altering gene regulation and being associated with disease (Wu *et al.* 2014). An illustrative example is the diabetes-associated gene *HNF4A*, which is regulated by several sequence elements in its 3′-untranslated region (3′-UTR), including miRNA binding sites (Wirsing *et al.* 2011). Other processes affected by GWAS variants include transcript splicing (Karambataki *et al.* 2014) and polyadenylation (Garin *et al.* 2010). Mutations disrupting the polyadenylation signal of the insulin gene may result in neonatal diabetes (Garin *et al.* 2010). Furthermore, variants altering the expression of RNA binding proteins, which are involved in many forms of gene regulation, can also have direct implications in diabetes (Hansen *et al.* 2015). Altogether, these observations call for a better understanding of the non-coding genome functions in tissues that are relevant for diabetes etiology.

For a significant fraction of T2D GWAS loci, genetic variation impacts insulin secretion (Perry & Frayling 2008, Dupuis *et al.* 2010). This observation points to a central role of the pancreatic β cells, whose primary function is to couple glucose levels with insulin secretion, in the pathophysiology of the disease, placing the pancreatic islet as a relevant tissue to study the genetic and molecular mechanisms underlying this disease. Defects in β cell development may also constitute a risk factor for glucose intolerance and β cell failure later on in life, putting pancreatic progenitors in the spotlight as an additional relevant tissue to discern the molecular mechanisms underlying the development of diabetes. In this review, we provide an overview of the recent developments in the analysis of non-coding functions in pancreatic tissues relevant for diabetes research: pancreatic progenitors and adult human islets.

Discovering the non-coding genome functions of pancreatic tissues

In recent years, technological advances in the domain of genome sequencing, together with access to human pancreatic primary tissues, allowed initial annotation of the non-coding genome of tissues such as human islets and pancreatic progenitors. Similarly to observations in other primary tissues and cell lines (Pennacchio & Visel 2010, ENCODE Project Consortium *et al.* 2012, Roadmap Epigenomics Consortium *et al.* 2015), the analysis of pancreatic tissues unmasked relevant regulatory functions of the non-coding genome. Greatly expanding our understanding of pancreatic genomic regulation, some studies focused on the identification of tissue-specific non-coding transcripts (Morán *et al.* 2012, Nica *et al.* 2013, van de Bunt *et al.* 2013, Fadista *et al.* 2014, Kameswaran *et al.* 2014), whereas others mapped genome-wide transcription factor binding sites and chromatin states (Bhandare *et al.* 2010, Gaulton *et al.* 2010, Stitzel *et al.* 2010, Parker *et al.* 2013, Dayeh *et al.* 2014, Pasquali *et al.* 2014, Cebola *et al.* 2015, Wang *et al.* 2015*a*).

Pancreatic islet non-coding RNAs

The advent of next generation sequencing technologies has unveiled that a large proportion of the transcribed genome lacks protein-coding potential, hence, enabling the identification and study of ncRNAs (The FANTOM Consortium 2005, Cabili *et al.* 2011, Iyer *et al.* 2015, Melé *et al.* 2015).

LncRNAs are a subgroup of ncRNAs that have a minimum length of 200 base pairs, and similarly to mRNAs, most of the lncRNAs identified so far are capped, spliced, and polyadenylated, although unspliced and non-polyadenylated variants are also observed. This group of transcripts may represent a novel layer of gene regulation (Guttman *et al.* 2009). In accordance with this hypothesis, a number of lncRNAs have been shown to be involved in the regulation of essential cellular functions, being implicated in many disease scenarios (Esteller 2011).

LncRNAs can regulate gene expression through a bewildering array of mechanisms, in the nucleus and in the cytoplasm, relying on their secondary and tertiary structures for that (reviewed in Rinn & Chang 2012) (Fig. 1). Nuclear lncRNAs can interact with transcription factors and chromatin remodelers, guiding them to target

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

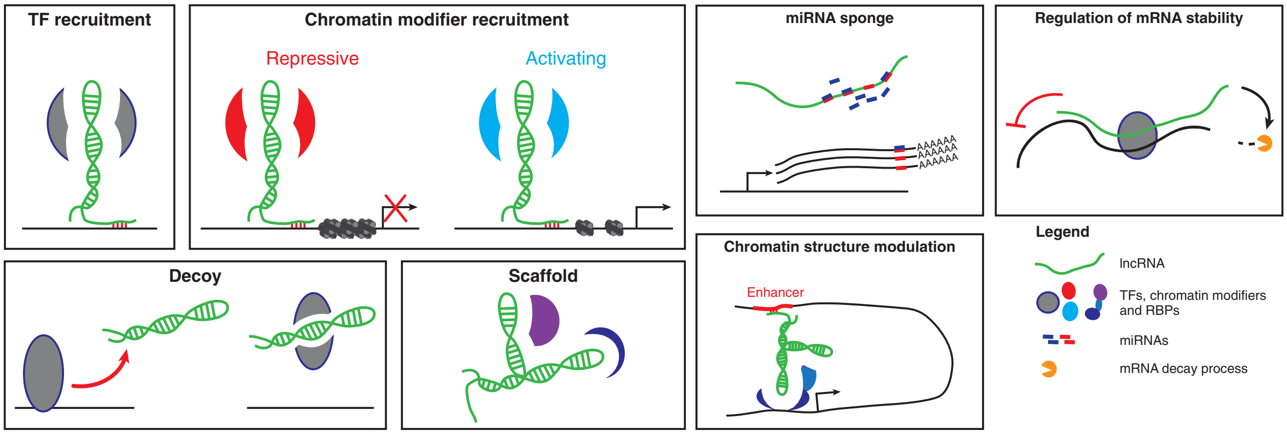


Figure 1
Models of gene expression regulation by lncRNAs. LncRNAs can interact with transcription factors and chromatin remodelers, guiding them to target loci to either activate or repress gene expression. LncRNAs may also function as decoys, competing for DNA-binding proteins, while other nuclear lncRNAs may act by facilitating enhancer-promoter chromosomal looping. In the cytoplasm, lncRNAs can affect gene expression by functioning as miRNA sponges or regulating mRNA stability.

loci to either activate or repress gene expression (Rinn *et al.* 2007, Khalil *et al.* 2009, Yap *et al.* 2010, Aguilo *et al.* 2011). Also in the nucleus, some lncRNAs have been described to function as decoys, competing for DNA-binding proteins, such as transcription factors (TFs), and titrating them away from their binding sites (Prensner *et al.* 2013, Xing *et al.* 2014). Other nuclear lncRNAs have been reported to act as enhancers via chromosomal looping (Ørom *et al.* 2010, Wang *et al.* 2011, Lai *et al.* 2013) or as scaffolds for large protein complexes (Yap *et al.* 2010, Aguilo *et al.* 2011). In the cytoplasm, lncRNAs can regulate mRNA stability (Kretz *et al.* 2013) or affect gene expression acting as miRNA sponges (Salmena *et al.* 2011, de Giorgio *et al.* 2013, Wang *et al.* 2013).

One study provided a comprehensive collection of coding and non-coding transcripts in pancreatic islets, which revealed over a thousand lncRNAs (Morán *et al.* 2012). As previously observed with lncRNAs in other tissues, islet lncRNAs are more tissue specific than their coding counterparts (Morán *et al.* 2012, Nica *et al.* 2013), supporting a possible role for islet lncRNAs in β cell function. Accordingly, while several lncRNAs have been associated with pancreas function and diabetes (recently reviewed by the Kaestner, Rutter, and Sussel teams (Kameswaran & Kaestner 2014, Pullen & Rutter 2014, Arnes & Sussel 2015), Morán *et al.* (2012) showed that a number of islet-specific lncRNAs are upregulated during endocrine-lineage commitment and some are glucose-responsive.

It has been observed that a number of T2D GWAS hits map to islet lncRNAs (Morán *et al.* 2012). However, careful examination is needed to discern this relationship, because the apparent correlation might be due to an overlap between islet lncRNAs and islet regulatory elements, such as enhancers. Even so, some islet lncRNAs were shown to be dysregulated in T2D, suggestive of their possible involvement in the molecular mechanisms of this disease (Morán *et al.* 2012).

MiRNAs, 21–25 base pair long, small non-coding RNAs, are also important epigenetic regulators of β cell function. Several lines of evidence point to their role in imparting robustness to developmental processes and show that miRNAs are interlaced within epigenetic and transcriptional networks for continuous control of lineage-specific gene expression (Kaspi *et al.* 2014). Deletion of *Dicer*, a gene that encodes a miRNA processing enzyme, in adult mouse β cells impairs β cell function (Lynn *et al.* 2007) and leads to diabetes (Melkman-Zehavi *et al.* 2011). Several miRNAs are involved in glucose homeostasis and β cell function, including miR-375 and miR-7a, which are involved in the regulation of glucose-stimulated insulin secretion (Poy *et al.* 2004, Ouaamari *et al.* 2008, Latreille *et al.* 2014). In a first attempt to identify the miRNAs that are enriched in human β cells, van de Bunt *et al.* (2013) profiled the miRNAs expressed in primary human islets and fluorescence-activated cell sorting (FACS) β cells using high-throughput sequencing of small RNAs, identifying 40 islet-enriched miRNAs in comparison to 15 control tissues. Interestingly, the authors observed an enrichment of islet-expressed miRNA targets for T2D association signals, highlighting a possible link between sequence

DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

variation in islet-miRNAs and T2D susceptibility. In addition, comparative studies now start to emerge, pinpointing miRNAs as players in novel molecular mechanisms, dysregulated in diabetic patients. This has been illustrated with the application of small RNA high-throughput sequencing to a small set of islet samples, which allowed the identification of an apoptosis-repressing miRNA cluster that is specifically downregulated in the islets of T2D individuals (Kameswaran *et al.* 2014). In another study, target-specific probe assays in a larger sample set allowed the identification of miR-187, a miRNA consistently overexpressed in islets from T2D individuals and associated with lower glucose-stimulated insulin secretion (Locke *et al.* 2014). Further comparative studies with larger and independent cohorts will further elucidate the role of these and other miRNAs in T2D etiology.

Regulatory element maps of adult human islets

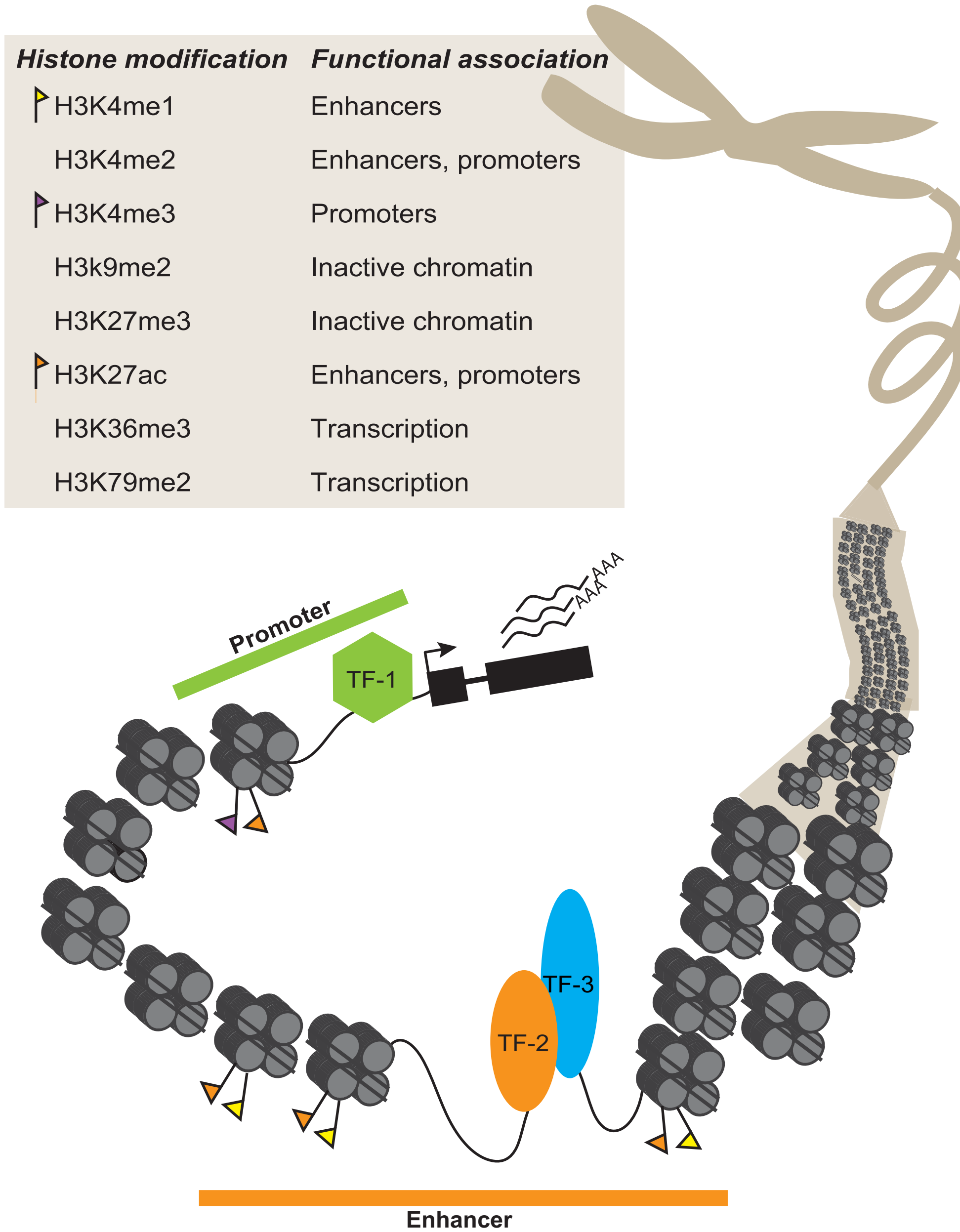
Large consortia such as ENCODE and the Epigenome Roadmap provided extensive epigenetic maps allowing detailed annotation of the non-coding regions of the human genome for a large number of human tissues, including several relevant to diabetes etiology such as adipose tissue and skeletal muscle. However, less accessible primary tissues and organs, such as the endocrine pancreas, were not prioritized in these studies. Due to the central role of human pancreatic islet cells in diabetes pathogenesis, different laboratories embarked in the annotation of non-coding regulatory elements in this tissue, constituting an ongoing effort to dissect the molecular mechanisms of human T2D. While several groups focused on profiling the chromatin landscape of pancreatic islets and on the classification of chromatin states in this tissue (Bhandare *et al.* 2010, Gaulton *et al.* 2010, Stitzel *et al.* 2010, Parker *et al.* 2013, Pasquali *et al.* 2014), others identified the binding sites of transcription factors relevant for β cell function (Khoo *et al.* 2012, Pasquali *et al.* 2014). These initiatives were recently joined by the Epigenome Roadmap project, which released epigenomic profiles of pancreatic islets earlier this year (Roadmap Epigenomics Consortium *et al.* 2015).

Major insights into the epigenetic information encoded within the nucleoprotein structure of chromatin have come from high-throughput genome-wide methods for assaying the accessibility of DNA to the machinery of gene expression, also referred to as chromatin "openness." The application of techniques such as FAIRE (formaldehyde-assisted isolation of regulatory elements) and DNase I hypersensitive site mapping, coupled with high-throughput sequencing, enabled the genome-wide identification of active transcription start sites, enhancers, and insulators in a broad range of cell lines and tissue samples including the pancreatic islets. As a proxy for islet regulatory regions, researchers initially profiled the open chromatin sites of pancreatic islets, providing a first glimpse on the tissue-specific regulatory landscape of human pancreatic islets (Gaulton *et al.* 2010, Stitzel *et al.* 2010).

The chromatin is built of nucleosomes, which are made up of approximately 147 bp of DNA and an octamer of histones. The N-terminal tails of these histones can be chemically modified by a variety of enzymes that are responsible for adding methyl, acetyl, and phosphor groups to histones. These histone modifications affect the chromatin structure and can control chromatin accessibility at certain genomic locations. While some histone modifications such as H3K9me3 contribute to a dense, closed chromatin structure, others are enriched at active genes (e.g., H3K9ac and H3K4me3) or at distal regulatory elements (e.g., H3K27ac and H3K4me1) (Fig. 2). Profiling of specific histone marks has thus enabled the characterization of the regulatory landscape of pancreatic islets and the mapping of distinct chromatin states, including promoters, active enhancers, insulators, and repressed regions (Bhandare *et al.* 2010, Stitzel *et al.* 2010, Parker *et al.* 2013, Pasquali *et al.* 2014); for an overview of the chromatin states and their associated histone marks, see Kellis *et al.* (2014) and Shlyueva *et al.* (2014).

Transcription factors translate cellular signals into regulatory programs. By binding to their target regulatory elements, transcription factors activate specific transcriptional programs that activate tissue- and cell-specific functions. Hence, profiling the binding sites of key islet transcription factors can help decipher the regulatory networks that they govern. As a proof of principle, Khoo *et al.* (2012) profiled the binding sites of PDX1, an essential regulator of pancreas development and β cell function, in mouse and human pancreatic islets, revealing that conserved occupancy sites are near genes with islet-specific activity. Further insights into pancreatic islet gene regulation were obtained by profiling the occupancy sites of NKX6.1, another pancreatic islet-specific transcription factor, in mouse islets, revealing that this transcription factor regulates several genes involved in insulin biosynthesis (Taylor *et al.* 2013).

Recently, one study integrated the profiling of pancreatic islet-specific transcription factors binding sites with mapping and annotation of chromatin states in human pancreatic islets (Pasquali *et al.* 2014). As observed



Histone modification	Functional association
H3K4me1	Enhancers
H3K4me2	Enhancers, promoters
H3K4me3	Promoters
H3k9me2	Inactive chromatin
H3K27me3	Inactive chromatin
H3K27ac	Enhancers, promoters
H3K36me3	Transcription
H3K79me2	Transcription

Figure 2

Histone post-translational modifications and their functional associations. A landmark of regulatory regions, such as enhancers, is their chromatin accessibility to transcription factors (TF), whereas densely positioned nucleosomes are associated with chromatin inactivity. Different combinations of post-translational histone modifications are associated with global and local chromatin states that eventually correlate with gene expression. Histones that flank active enhancers are often marked by histone H3 lysine 27 acetylation (H3K27ac) and histone H3 lysine 4 monomethylation (H3K4me1). Active promoters may be flanked by nucleosomes with H3K27ac and H3K4me3 modifications. Highlighted in the table are the major post-translational histone modifications and their functional associations.

in other tissues, the co-occupancy of transcription factors tends to coincide with active enhancers more frequently than for other similarly accessible chromatin states. Interestingly, further analysis revealed that transcription factor binding on open chromatin of different classes is associated with considerably different regulatory functions. Islet-selective transcription factors were unexpectedly found to bind to thousands of ubiquitously expressed promoters, as well as to the CTCF-bound sites and H3K4me1-enriched transcriptionally silent regions. Tissue-specific gene regulation was instead linked to large domains of active chromatin characterized by a high

DOI: 10.1530/JME-15-0197

Printed in Great Britain

Published by Bioscientifica Ltd.

density of enhancers bound by multiple transcription factors. These observations suggest that transcription factor networks establish functionally distinct epigenomic contacts and control tissue-specific functions by means of *cis*-regulatory networks. Together with evidence showing overrepresentation of T2D-associated variants in enhancers, these data suggest a possible mechanism by which non-coding disease variants have an impact in islet function.

Regulatory element maps of human pancreatic progenitors

Pancreatic islets are central in diabetes etiology, but variants in loci involved in early pancreas development such as *FOXA2* and *PDX1* can also be associated with T2D (Manning *et al.* 2012, Scott *et al.* 2012), suggesting that defects during pancreas development might also contribute to the onset of the disease in the adult.

Due to the limited access to cadaveric human fetal tissue, to date, mouse knockout models have been the most powerful tools to study embryogenesis and to uncover the role of many transcription factors in pancreas development (Offield *et al.* 1996, Jacquemin *et al.* 2000, Haumaitre *et al.* 2005, Seymour *et al.* 2007, Gao *et al.* 2008, Carrasco *et al.* 2012, Xuan *et al.* 2012). However, marked differences between mouse and human pancreas development limit the application of such models (reviewed in Nair & Hebrok 2015). This is especially true when trying to apply mouse genetics to understand the impact of human genetic variation in pancreas development and its contribution to the various forms of diabetes.

The differentiation of human embryonic stem cells (hESC) and human-induced pluripotent stem cells (hiPSC) now provides an unlimited source of pancreatic progenitors that are amenable to be studied and manipulated in different settings (Cho *et al.* 2012, Pagliuca *et al.* 2014, Russ *et al.* 2015). Two recent works have employed hESC-derived pancreatic progenitors to map active enhancers (Cebola *et al.* 2015, Wang *et al.* 2015*a*). Importantly, in one of the studies, the identified regions were validated with human fetal tissue, supporting the applicability of the *in vitro* model in the study of human pancreas development and disease (Cebola *et al.* 2015).

Similarly to observations in adult pancreatic islets (Pasquali *et al.* 2014) and other tissues, the tissue-specific regulatory program of pancreatic progenitors is orchestrated by a combinatorial code of transcription factors that lie in enhancer regions (Cebola *et al.* 2015). In fact, the introduction of mutations in such *cis*-regulatory modules at specific transcription factor binding sites completely abrogates their regulatory activity. Current and future works will help characterize disease-associated genetic variation in the context of pancreatic progenitor transcriptional regulation.

Other non-coding genome functions in adult human islets

In addition to the analysis of non-coding RNAs and chromatin structure and packaging, a number of studies have investigated other non-coding functions of the genome. In this section we provide a summarized overview of these processes and their implications in pancreas biology and disease.

Similarly to histone mark enrichment, DNA methylation is an epigenetic mechanism able to modulate genomic regulation by altering chromatin accessibility and is frequently dysregulated in pathological settings (Robertson 2005, Jones 2012). In mammalian cells, DNA methylation has been more extensively studied in the context of CpG dinucleotides but can also occur outside CpG sequences (Lister *et al.* 2009, Schultz *et al.* 2015). Importantly, DNA methylation function is context dependent (Jones 2012, Schübeler 2015). CpG-rich regions, known as CpG islands, tend to be located near transcription start sites, being their methylation associated with transcription initiation blockade and consequent gene silencing. On the other hand, gene body methylation tends to be associated with transcription elongation. DNA methylation is also involved in transposable elements suppression, promoting genome stability. Even though transcriptional enhancers tend to be CpG-poor, there is mounting evidence for a close relationship between their methylation status, TFs occupancy, and transcriptional activity (Wiench *et al.* 2011, Hon *et al.* 2014, Lu *et al.* 2014).

DNA methylation profiles are highly tissue specific, being involved in the regulation of key tissue-specific functions, including β cell maturation (Dhawan *et al.* 2015). When comparing islets from T2D patients with islets from non-diabetic donors, investigators observed alterations in the methylation levels of several genes (Volkmar *et al.* 2012, Dayeh *et al.* 2014), including well-known T2D-risk loci such as *KCNQ1* and *IRS* (Dayeh *et al.* 2014).

DNA methylation can be a dynamic process, and a recent study in rodents suggests that β cell DNA methylation is modulated during aging, a major risk factor for diabetes (Avrahami *et al.* 2015). In this study, increased DNA methylation was observed at the promoters

Published by Bioscientifica Ltd.

of cell cycle genes, which was associated with reduced proliferative capacity. This observation relates well to the reduced ability of old β cells to regenerate. Surprisingly, however, the authors observed demethylation of enhancers near genes that regulate metabolic functions in aged mice, which is associated with improved β cell function. These results suggest that, at least in rodents, β cell function may improve with age to counteract their decreased proliferative potential. Similar comparisons in humans could help elucidate the role of aging in T2D risk.

Although few studies have addressed this issue, mounting evidence now links DNA sequence variation and DNA methylation, and it is possible that methylation changes act together with particular genetic traits to confer higher disease susceptibility (Bell *et al*. 2010, Dayeh *et al*. 2013, Petersen *et al*. 2014, Orozco *et al*. 2015).

Several epigenetic processes affecting transcripts rather than chromatin structure, such as RNA editing and RNA methylation, are now starting to be investigated in depth and will likely be the subject of future efforts to understand the molecular mechanisms underlying human diseases, including T2D.

Adenosine-to-inosine (A-to-I) RNA editing is a common post-transcriptional modification of RNA molecules that has been implicated in several human diseases (Maas *et al*. 2006). In pancreatic islets, Fadista *et al*. (2014) reported potential RNA editing at several loci, highlighting the potential of this mechanism to modulate key β cell genes and, hence, confer susceptibility to T2D.

The most prevalent type of RNA methylation, N^6-methyladenosine (m^6A), is broadly distributed in both coding and non-coding RNAs (Dominissini *et al*. 2012, Fu *et al*. 2014). Recent reports show that RNA molecules carrying the m^6A modification are less stable and more efficiently translated, being associated with dynamic and fast-response cellular processes (Fustin *et al*. 2013, Wang *et al*. 2014, Wang *et al*. 2015*b*). Preliminary studies showed that T2D patients tend to show higher m^6A demethylase (FTO) expression, which correlates with lower m^6A in peripheral blood RNA (Shen *et al*. 2014). Further studies should elucidate the functional implications of differential m^6A levels in T2D individuals.

Additional mechanisms by which transcripts can be differentially processed, stabilized, localized, or translated involve the action of RNA-binding proteins (RBPs) (reviewed in Keene 2007). RBPs interact with target transcripts via specific sequence motifs, such as pyrimidine-rich, CG-rich, or AU-rich sequences. In the pancreas, RBPs regulate key features of β cell function, including the regulation of insulin mRNA stability and translation (Magro & Solimena 2013). Consequently, GWAS variants affecting RBPs may have a direct impact on β cell function (Hansen *et al*. 2015). RBPs have an essential role in insulin regulation, mediated by specific regulatory sequences present in the preproinsulin mRNA UTRs. These pyrimidine-rich sequence recruits RBPs that in turn stabilize the transcript (Tillmar *et al*. 2002). A conserved element at the 5′-UTR is instead required for glucose-regulated proinsulin translation (Wicksteed *et al*. 2007). These examples highlight the regulatory potential of RBPs and sequence motifs in transcripts. Given the broad spectrum of RBP functions and regulatory sequences in mRNAs and other types of transcript, it is likely that many more RBP-related functions will be discovered in the context of pancreas function and disease.

Interconnections of non-coding genomic functions

The features described above – non-coding RNAs, histone modifications, DNA and RNA methylation, RNA editing, and regulation by RBPs – are only part of the vast array of non-coding functions of the human genome. Furthermore, these regulatory mechanisms are not isolated from each other but are, in reality, interconnected. In this section we provide a few examples showing how different epigenetic processes and regulatory elements can be interlinked to modulate gene expression at loci that may be linked to diabetes (Fig. 3).

An interesting example of interconnections of different epigenetic mechanisms that can contribute to T2D can be observed in the *DLK1-MEG3* locus, which contains an islet-specific miRNA cluster. *DLK1-MEG3* is an imprinted locus in which, under normal conditions, a cluster of islet-specific miRNAs is expressed from the maternal allele together with the lncRNA *MEG3*. In T2D, researchers have observed significant DNA hypermethylation of *MEG3*, which is associated with a downregulation of the miRNA cluster (Kameswaran *et al*. 2014). Among the targets of these miRNAs, the authors of the study identified genes essential for islet function such as *IAPP* and *TP53INP1* (p53), which are involved in β cell apoptosis in T2D (Fig. 3A).

An example of the interplay of lncRNAs, histone modifications, and DNA methylation resides in the *CDKN2A* locus, a hot spot in GWAS for a variety of diseases, including T2D (Pasmant *et al*. 2011). In brief, studies in cancer cells have found that the lncRNA *ANRIL* regulates the expression of *CDKN2A* by directly recruiting the polycomb repressive complexes-1 and -2 (PRC1 and PRC2) (Yap *et al*. 2010, Aguilo *et al*. 2011). This recruitment

Published by Bioscientifica Ltd.

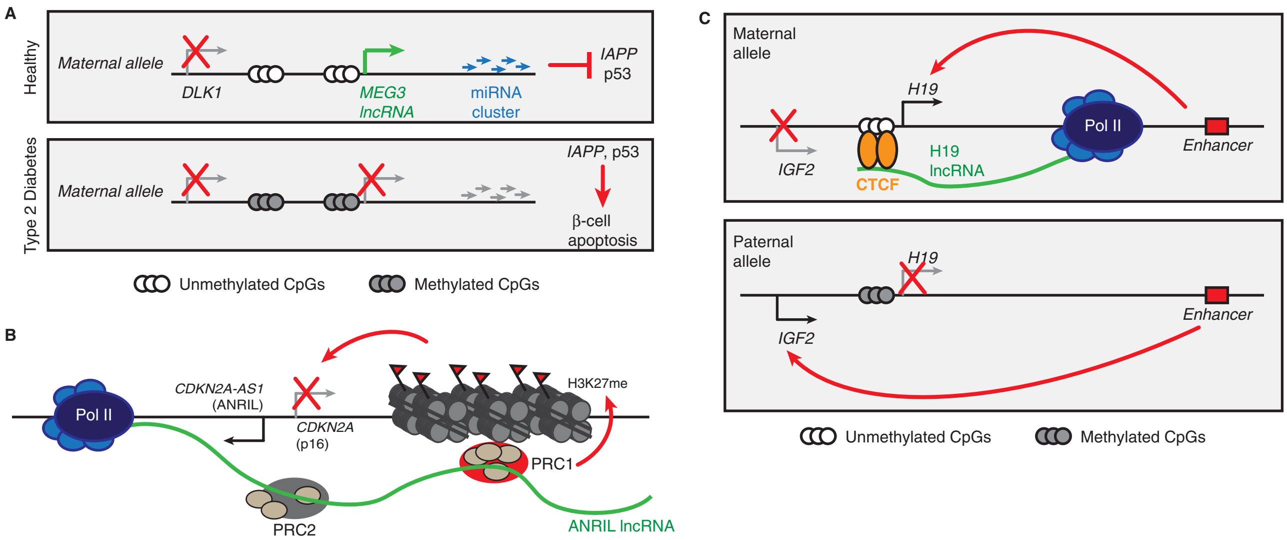


Figure 3

The non-coding functions of the genome are interconnected. (A) *DLK1-MEG3* is an imprinted locus in which, under normal conditions, a cluster of islet-specific miRNAs is expressed from the maternal allele together with the lncRNA *MEG3*. In T2D, *MEG3* is hypermethylated, which is associated with a downregulation of the miRNA cluster. These miRNAs target genes essential for islet function such as *IAPP* and *TP53INP1* (p53), which are involved in β cell apoptosis in T2D. (B) In prostate cancer tissues, the lncRNA *ANRIL* interacts with polycomb complexes to induce gene repression in *CDKN2A* locus. (C) Control of the imprinted locus *H19-IGF2*. Methylation of an imprinting controlled region in the paternal allele maintains the lncRNA *H19* silence and allows the interaction of downstream enhancers with the promoter of *IGF2*, contributing to its expression. In the maternal allele, this region is unmethylated, allowing expression of *H19* and binding of CTCF, a factor involved in the establishment of insulators, blocking the interaction of downstream enhancers to *IGF2*. This results in an enhanced interaction of the same set of enhancers with *H19*.

results in the enrichment of repressive histone modifications (H3K27me3) in the region and subsequent gene silencing (Fig. 3B). While in mouse β cells, *Cdkn2a* expression increases during aging and is associated with a decline in islet regenerative potential (Krishnamurthy *et al.* 2006), similar mechanisms as those described in cancer cells could be implicated the regulation of *CDKN2A* in β cells or other tissues relevant to diabetes.

DNA methylation is a key regulatory mechanism controlling gene expression of coding genes and lncRNAs. An example of this is observed in the imprinted locus *H19-IGF2*, which consists of the paternally expressed *IGF2* gene (coding for insulin-like growth factor 2, an important fetal growth factor) and the maternally expressed lncRNA *H19* (involved in cell proliferation) (reviewed in Kameswaran & Kaestner 2014). In the maternal allele, an imprinting controlled region (ICR) is unmethylated, which allows binding by CTCF, a factor involved in the establishment of insulator elements, blocking the interaction of downstream enhancers with the promoter of *IGF2* and promoting their interaction with *H19* (Bell & Felsenfeld 2000, Hark *et al.* 2000). Conversely, in the paternal allele, the ICR is methylated, silencing *H19* and inhibiting CTCF binding, which in turn and allows the interaction of distal enhancers with *IGF2* (Fig. 3C). This example highlights how lncRNAs and coding genes can compete for the same set of enhancers to modulate their transcriptional levels and how regulatory elements such as differentially methylated regions can be involved in this process.

The molecular mechanisms described above exemplify the high level of complexity required for maintaining and fine-tuning gene regulation at loci potentially involved in human diabetes. Thus, to improve our current understanding of the molecular basis of diabetes mellitus, we need to carefully characterize the interconnection of non-coding genome functions in human pancreatic islets as well as in other tissues or developmental stages involved in the onset and progression of the disease.

Enhancer clusters and pancreatic islet-cell identity

Initial studies in human pancreatic islets and other tissues revealed that tissue-specific regulatory elements are not evenly distributed along the genome but instead are contained in large clusters of open regulatory elements (COREs) (Gaulton *et al.* 2010, Song *et al.* 2011). More detailed regulatory maps including the profiling of histone

Published by Bioscientifica Ltd.

modifications and transcription factor binding maps have further unmasked a pervasive link between enhancer clusters, also referred to as super- or stretch-enhancers, and tissue-specific gene activity (Hnisz *et al.* 2013, Parker *et al.* 2013, Pasquali *et al.* 2014). This link is illustrated, in islets and other tissues, by the fact that genes near regions with a high density of enhancers in a given cell type tend to be involved in functions that define the identity of that specific cell type. For example, islet enhancer clusters tend to be near genes involved in insulin biosynthesis and secretion (Fig. 4).

Further supporting the hypothesis that enhancer clusters are key to defining the genetic program associated with islet-cell identity, gain and loss of function experiments have demonstrated that the subset of transcription factors binding enhancer clusters is functionally linked to islet-specific gene activity (Pasquali *et al.* 2014). Surprisingly, genes bound by the same transcription factors only at promoter or other open chromatin sites are not modulated on perturbation. These experiments suggest a complex regulatory architecture, involving clusters of enhancers, that controls the transcriptional programs required to establish islet-cell-specific functions.

Studies on the chromatin architecture demonstrated that the genome is functionally organized in chromosomal territories (Gilbert *et al.* 2004, Dillon 2006, Guelen *et al.* 2008). Such higher-order conformation of the chromatin pointed to the possibility that gene regulation relies on functional domains. Recently, chromosome conformation capture (3C)-based techniques have confirmed the compartmentalization of the genome and its further organization into smaller topologically associated domains (TADs). Importantly, while TAD borders are predominantly conserved among different cell-types, TADs often harbor active chromatin domains that undergo dynamic and cell-type-specific interactions. In pancreatic islets, high-resolution conformation capture experiments (4C) showed that islet-specific promoters frequently interact with tissue-specific enhancer clusters (Pasquali *et al.* 2014). In fact, subsequent analyses revealed that these interactions are always confined within TAD borders (Fig. 4). These observations show that clusters of

Journal of Molecular Endocrinology

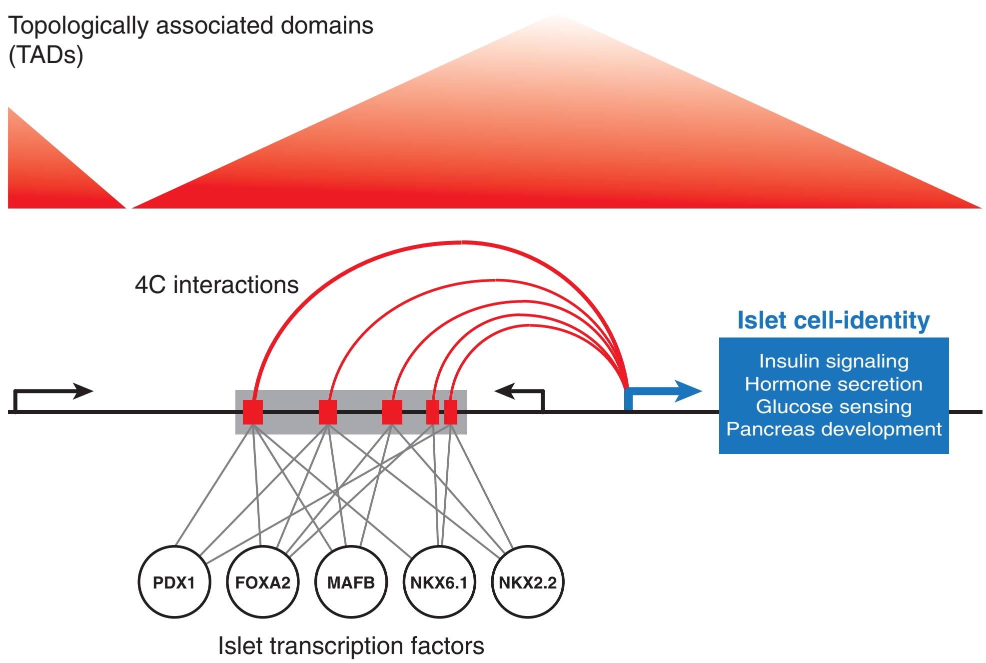


Figure 4
Enhancer clusters and islet-cell identity. Transcription factors essential for β cell differentiation and function, such as PDX1, FOXA2, MAFB, NKX6.1, and NKX2.2, regulate the transcriptional program of this cell type by binding tissue-specific enhancer elements (red boxes). This intricate regulatory network is further organized into genomic regions with high density of enhancers, called enhancer clusters (highlighted as a grey box), which regulate defining functions of islet-cell identity, including insulin biosynthesis and secretion. Enhancer clusters form complex 3D chromatin structures at tissue-specific expressed loci. Enhancer-promoter contacts may be captured by 3C-based techniques such as 4C-seq (4C-seq contacts are schematically depicted by the red arches). Enhancer clusters and their target genes can be megabases away and may undergo dynamic cell type-specific interactions within their topologically associated domains (TAD) boundaries.

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

enhancers participate in complex 3D structures at loci that are specifically expressed in islets, further disclosing their tight association with islet-cell identity.

A better understanding of the tissue-specific regulatory architecture is likely to provide insights to identify novel pancreatic islet regulators linked to clusters of enhancers, decipher the pancreatic islet sequence regulatory code, and pinpoint disease-relevant non-coding genetic variation affecting tissue-specific genome regulation.

On the other hand, more compelling questions on the architectural regulatory functions are now starting to be addressed by the scientific community. What is the level of redundancy among the enhancers comprised in these large active chromatin domains? Which are the dynamic proprieties of such chromatin structures in development and disease? Recently, Lupiáñez *et al.* (2015) showed how high-order chromatin architecture disruption could have pathological implications. In this study, the authors showed that human limb malformations could arise from chromosomal rearrangements spanning TAD boundaries. These structural rearrangements cause abnormal interactions between promoters and regulatory elements, resulting in erroneous gene expression regulation. Even though these observations were made in other cell types, it is possible to envision that a similar mechanism could be responsible for diseases affecting other human tissues and organs, including the pancreas.

Genetic variation in enhancers and diabetes

Despite the numerous T2D-associated genetic variants revealed by GWAS, the identification of causal variants remains a challenge. Disease-associated variants lie in non-coding regions and their functional role cannot be explained by protein sequence changes, thus suggesting a regulatory impact. Some of the first studies to associate GWAS variants with regulatory elements integrated genetic risk variants with regulome maps generated through epigenomic profiling (Ernst *et al.* 2012, Maurano *et al.* 2012). Two major observations emerged from these studies. First, sites with an enhancer signature are highly enriched for genetic risk variants relative to other chromatin-defined elements such as promoters and insulators. Second, risk variants preferentially map to enhancers specific to disease-relevant cell types. Furthermore, as observed for other disease traits, single nucleotide polymorphisms (SNPs) associated with T2D and fasting glycaemia levels are enriched in pancreatic islet-enhancer clusters and stretch enhancers (Parker *et al.* 2013, Pasquali *et al.* 2014) rather than non-clustered enhancers. This indicates that the regulatory variation that affects islet-specific gene regulation is relevant to T2D pathophysiology. These conclusions are supported by the observation that at least a few T2D variants seem to be linked with allele-specific gene expression changes (Locke *et al.* 2015).

Some groups have taken advantage of the publicly available maps of open chromatin and enhancer maps from pancreatic islets to identify functional T2D variants (Table 1). Exploiting this data enabled the functional characterization of risk variants that disrupt transcription factor-binding motifs and that have an impact on the activity of islet enhancers. Although still scarce, these studies proved the regulatory potential of a few selected T2D-associated variants and encourage researchers to apply epigenomic maps to better understand the genetic basis of this disease. Future studies, modeling such variants in pancreatic islet cells, may shed light on the molecular mechanisms underlying T2D.

Typically the functional validation of putative causal variants involves transcription factor binding analysis (*in silico* and *in vitro*), as well as enhancer activity assays and allele-specific expression quantification. Nevertheless, such experiments characterize the functional potential of the associated variant without providing information on the regulated gene target. A landmark study recently showed that obesity-associated variants at an intronic *FTO* region are located in enhancers that unexpectedly regulate *IRX3*, a gene that maps 0.5 Mb downstream of *FTO* (Smemo *et al.* 2014). Such results highlight the need of integrating functional characterization of variants with computational analysis and other molecular biology techniques, enabling the systematic identification of genes influenced by T2D-susceptibility variants. 3C-based techniques have thus the potential to reconstruct the 3D chromatin structure of T2D or fasting glycemia-associated loci, unmasking genes in physical contact with enhancers carrying disease variants. The analysis of natural variation in expression quantitative trait loci (eQTL) studies, as well as the use of targeted mutations in experimental models, will also provide a deeper understanding of the mechanistic and functional relationships between enhancers and target genes. This knowledge will be the basis for understanding how enhancer variants influence human disease and glucose-related traits in particular.

Similar issues arise from the ever-growing number of non-coding variants uncovered by whole genome sequencing of Mendelian diabetes patients. Thus, as for GWAS, to make the translation of Mendelian genome sequencing findings meaningful, it is critical to build platforms to prioritize variants according to their functional likelihood.

Published by Bioscientifica Ltd.

Table 1 Examples of diabetes-associated non-coding functional variants with impact in pancreatic tissues[a]

Locus[b]	Functional variant	Experimental evidence	Phenotype	Reference
INS	NM_000207.2:c.-331C>G NM_000207.2:c.-331C>A NM_000207.2:c.-332C>G NM_000207.2:c.-218A>C	Episomal reporter assay	Neonatal diabetes	Garin *et al.* (2010)
PTF1A[c]	hg19 chr10:g.[23508437A>G] hg19 chr10:g.[23508363A>G] hg19 chr10:g.[23508305A>G] hg19 chr10:g.[23508365A>G] hg19 chr10:g.[23508446A>C] hg19 chr.10:g.[23502416-23510031del]	Episomal reporter assay, EMSA, 3C	Neonatal diabetes	Weedon *et al.* (2014)
BLK	hg18 chr8:g.[11369157G>A] hg18 chr8:g.[11459364T>G] hg18 chr8:g.[11459531G>T] hg18 chr8:g.[11468050C>T]	Episomal reporter assay	MODY	Borowiec *et al.* (2009)
ZFAND3[c]	rs58692659	Episomal reporter assay, EMSA	T2D	Pasquali *et al.* (2014)
JAZF1[c]	rs1635852	Episomal reporter assay, EMSA	T2D	Fogarty *et al.* (2013)
CDC123[c]	rs11257655	EMSA, Allele-specific ChIP	T2D	Fogarty *et al.* (2014)
ARAP1	rs11603334 rs1552224	Episomal reporter assay, EMSA	T2D	Kulzer *et al.* (2014)
TCF7L2[c]	rs7903146	Episomal reporter assay, Allele specific FAIRE	T2D	Gaulton *et al.* (2010)

EMSA, electrophoretic mobility shift assay; ChIP, chromatin immunoprecipitation.
[a]This table shows a non-exhaustive set of variants with an impact on pancreatic tissues, aiming to highlight regulatory variants in enhancers.
[b]Nearest gene to the variant.
[c]Regulatory variants in enhancers.

Recently, the systematic analysis of variants enabled the discovery of regulatory mutations associated with isolated pancreas agenesis (no extra-pancreatic features) (Weedon *et al*. 2014). In this study, whole genome sequencing of two unrelated patients was combined with epigenomic maps to restrict the search of causal mutations. The analysis revealed several recessive mutations in an enhancer of *PTF1A* (pancreas transcription factor 1), a gene essential for pancreas development. In addition to a large deletion encompassing the enhancer element, point mutations disrupting the binding sites of key pancreas developmental transcription factors were also identified. As the affected *PTF1A* enhancer is only active in pancreatic progenitors (Cebola *et al*. 2015, Wang *et al*. 2015*a*), without access to human pancreatic progenitors, or in their place, *in vitro* differentiated pancreatic progenitors, the identification of the causal mutations would not have been possible. Strikingly, to date, this is the most frequent cause of isolated pancreas agenesis.

Given its phenotypical heterogeneity, an appropriate diagnosis of MODY is essential to provide adequate treatment to patients. For example, while individuals with *HNF1A* and *HNF4A* mutations are sensitive to low-dose sulphonylureas, individuals with *GCK* mutations do not require pharmacological treatment. Traditional genetic diagnosis methods only cover small coding regions and lack the power to detect all causal mutations. Until recently, molecular diagnosis was not possible in over 70% of the cases referred for genetic testing, for which coding mutations at known culprit genes could not be identified (Shields *et al*. 2010). Efforts to fill this gap arise from high-throughput sequencing methods, such as targeted sequencing of large panels of genes, which constitutes a cost-effective method of genetic diagnosis (Ellard *et al*. 2013). In this study, Ellard *et al*. applied protein-coding sequence targeted assays to a cohort of patients with an unknown cause of MODY (classified as MODY-X) or neonatal diabetes, finding novel MODY coding mutations in 15% of the cases. Noteworthily, the successful identification of causal mutations allowed the redirection of patients from insulin to sulphonylurea treatment. However, even when applying high-throughput methods, 85% of clinically diagnosed MODY patients lacked causal coding mutations, suggesting that regulatory mutations might also play a role in monogenic forms of diabetes. Similar targeting approaches could be applied to pancreatic progenitor and islet-enhancer maps to address this question and help phenotypically characterize MODY patients with unknown cause.

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

While we herein described examples of Mendelian diabetes in which a single enhancer variant causes congenital disease, GWAS loci often contain several variants in linkage disequilibrium, raising the question of whether the causal variant acts alone. Alternatively, complex risk haplotypes, containing more than one causal variant, might be in place. In such a scenario, multiple variants in linkage disequilibrium might affect a cluster of enhancers and cooperatively affect target-gene expression. In fact, work on autoimmune disorders has showed that this "multiple enhancer variant'' hypothesis may underlie part of the missing heritability of complex diseases (Corradin *et al*. 2014).

Interestingly, the molecular mechanisms of MODY and T2D cannot be fully dissociated. According to current knowledge, MODY is predominantly caused by coding mutations in genes that are involved in the transcriptional control of glucose homeostasis and β cell development and function. Likewise, these molecular pathways also seem to be affected in T2D, as the associated non-coding variants tend to disrupt binding sites for the same transcription factors (Maurano *et al*. 2012). Thus, better knowledge of the genes affected in MODY might provide clues to discover novel genes implicated in T2D etiology and vice versa.

Systematic identification of functional variants

As enhancers and other regulatory elements tend to be more conserved than random genomic sequences, sequence conservation scores, such as Genomic Evolutionary Rate Profiling (GERP) and phyloP, can be applied to prioritize variants with functional potential. However, many regulatory elements are not highly conserved at the sequence level (Gulko *et al*. 2015). Recently, other computational methods have been applied to uncover general features of *cis*-regulatory variants by integrating experimental data, *in silico* transcription factor binding site predictions, selective pressure, tissue-specific regulatory maps, and regional patterns of polymorphism.

The web tool RegulomeDB (http://regulomedb.org) comprises one of the first attempts to systematically prioritize and identify functional non-coding variants, integrating experimental data from ENCODE, the Epigenome Roadmap, and other sources, as well as *in silico* predictions and manual annotations (Boyle *et al*. 2012). Posterior studies have also focused on the tissue specificity of regulatory variants, providing more customizable tools, such as GWAVA, CADD and fitCons, which can incorporate different kinds of tissue-specific epigenomic annotations to prioritize putative functional variants (Kircher *et al*. 2014, Ritchie *et al*. 2014, Gulko *et al*. 2015).

Transcription factor co-occupancy at enhancers is a recurrent event in many human tissues, including pancreatic progenitors and islets (Pasquali *et al*. 2014, Cebola *et al*. 2015). Claussnitzer *et al*. (2014) exploited this feature together with selective pressure to explore the regulatory code at T2D GWAS loci. This integrative computational analysis revealed a striking accumulation of homeobox transcription factor binding sites, including PDX1 and other transcription factors known to be important in pancreas biology, and resulted in a framework to guide the identification of *cis*-regulatory functional variants.

While the starting point for Claussnitzer *et al*. was GWAS loci, in a recent study, Lee *et al*. (2015) detected tissue-specific regulatory codes by comparing putative tissue-specific regulatory regions derived from open chromatin assays with a matched negative control. Such tissue-specific regulatory codes were then applied to predict the functional impact of sequence variation at base-pair resolution. Noteworthy, the authors were only able to accurately identify causal variants when the computational tool was trained with regulatory regions from an appropriate tissue. These results further demonstrate the requirement for regulatory maps of disease-relevant tissues to find causal variants.

The tools described above provide unbiased methods to prioritize putative causal variants; however, they do not integrate enhancer-gene interaction information, which is key when translating regulatory sequence variation to its biological impact. Non-coding variants are often found in gene deserts and megabases away from their target genes, which are thus difficult to pinpoint (Maurano *et al*. 2012). In fact, even though earlier studies attributed enhancers to their nearest gene, the application of 3C-based techniques demonstrated that this is not always the case (Smemo *et al*. 2014). To address this issue, Corradin *et al*. (2014) have developed PreSTIGE, a publicly available tool that integrates enhancer histone marks and gene expression analysis from a panel of cell and tissue types to identify tissue-specific interactions.

Data visualization and easy access to regulatory information is also vital to correctly design hypothesis-driven functional experiments. In this sense, RegulomeDB allows the visualization of variants of interest in their genomic context, providing functional annotations for an array of tissues and transcription factor binding motif analysis (Boyle *et al*. 2012). Specifically focused on pancreatic gene regulation, the Islet Regulome Browser

DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

(http://www.isletregulome.com) provides interactive access to a wealth of information, allowing the visualization of different classes of regulatory elements, together with enhancer clusters, transcription factor binding sites, and binding motifs, which are integrated with publicly available T2D and fasting glycemia GWAS datasets.

The computational analysis and visualization tools mentioned here provide frameworks to systematically prioritize regulatory variants for further *in vitro* and *in vivo* functional validation. These experiments will hopefully accelerate the discovery of disease-relevant variants and, in the future, the eventual translation of GWAS findings to the clinic.

Next challenges for T2D variant discovery

Future progress in understanding the impact of genetic variants on tissue-specific epigenomes in the context of T2D will necessarily need to go through whole genome sequencing of T2D patients with identification of low-frequency variants associated with the disease and epigenetics map charting in T2D-relevant tissues including early and late stages of development, as well as pertinent metabolic states. These advances will enable researchers to dissect the contribution of genetic variation to disease development while further functional studies including allelic expression, 3C assays, and genome editing will unmask mechanistic links within tissue-specific gene regulation processes.

As studies have shown, there is an excess of recent rare variants associated with T2D in the human population (Coventry *et al*. 2010, Bonnefond *et al*. 2012). Thus, the expansion of association studies to rare or personal variants will certainly improve the estimates of variance explained. However, rare variants are unlikely to completely explain the predisposition. An open avenue in the attempt of unmasking the unexplained fraction of disease variance may pass through the epigenetic characterization of humans at risk of T2D. So far, the few studies that addressed this issue were predominantly focused on the DNA methylation of selected CpG sites, identifying aberrantly regulated genes in T2D pancreatic islets (Ling *et al*. 2008, Volkmar *et al*. 2012, Yang *et al*. 2012). However, these observations need to be considered carefully, as epigenetic variation can either contribute or be a consequence of the disease. Aging, which is associated with T2D onset, promotes the accumulation of DNA methylation errors. Conversely, altered metabolic regulation in T2D could induce sustained epigenetic changes.

The first T2D epigenome-wide association studies (EWASs) are now starting to be performed (Dick *et al*. 2014, Hidalgo *et al*. 2014, Petersen *et al*. 2014, Yuan *et al*. 2014, Chambers *et al*. 2015, Kulkarni *et al*. 2015). However, so far, T2D EWAS have only been performed with whole-blood cells instead of pancreatic islets or other disease-relevant tissues, only being able to grasp early developmental epigenetic changes, which can be present in multiple tissues, and alterations derived from inflammatory processes, which are often detectable in circulating blood. Furthermore, similarly to the initial GWAS, these first studies were limited by low statistical power and rare follow-up replication. In the near future, EWAS will almost certainly rely on centralized community efforts due to the high experimental costs and the difficulty of accessing large numbers of samples from disease-relevant tissue and/or cell types. These studies will improve our understanding of several aspects of T2D participating factors: the contribution of the epigenome rather than the sequence composition to the disease development; the interaction between sequence variation and personal epigenome; and the manner in which epigenome translates environmental risk factors into disease susceptibility.

Altogether, integration of genetics and epigenetics data will allow a clearer picture of the molecular mechanisms behind the development of T2D.

Concluding remarks and perspectives

GWAS have provided large collections of T2D-associated variants in the recent years. Nevertheless, despite better methodologies such as meta-analysis of large cohorts, trans-ethnic GWAS, or fine mapping with dense genotyping (Farh *et al*. 2015), identifying the functional variants remains challenging in most cases.

The identification of regulatory elements in relevant cell and tissues types – pancreatic progenitors and islets – will allow us to refine the search for disease-relevant variants. In the upcoming years, a large collection of T2D-associated variants overlapping promoters, enhancers, miRNAs, lncRNAs, and other non-coding elements of pancreatic progenitors and islets will be unmasked. Increased resolution and types of regulatory maps will help prioritize truly functional variants but will not suffice to reveal the mechanism of how disease-susceptibility is conferred.

Affordable genome-editing tools, such as the clustered regularly interspaced short palindromic repeats (CRISPR) system (Cong *et al*. 2013), will allow us to directly study

DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

the impact of a given variant in its *cis*-regulatory context. The introduction of T2D-associated variants or other forms of diabetes-associated variants in relevant cell lines or animal models will be crucial to isolate the impact of each variant on β cell function and/or on pancreas development. Ultimately, genome editing of associated variants will also enable the study of more complex and realistic scenarios, including genotypes containing several interacting functional variants. Furthermore, CRISPR-enabled epigenome editing tools have been recently developed (Hilton *et al.* 2015). By coupling CRISPRs with either repressor or activating protein domains, researchers will now be able to target specific genomic regions and alter the regulatory landscape, which will result in controlled gene expression manipulation. This line of research has the potential to identify key molecular mechanisms underlying diabetes and other human diseases, possibly uncovering etiological therapeutic targets.

To date, the functional study of genetic variants-associated diabetes development has been greatly frustrated by the limited access to human pancreatic islets, as well as by the lack of appropriate *in vitro* cellular models to study pancreatic β cells. The groundbreaking discovery of induced pluripotent stem cells (iPSCs) by Yamanaka (2007) has opened new doors in the field of personalized medicine. Similarly to many other human diseases, it is now possible to generate iPSC from diabetic patients (Maehr *et al.* 2009, Kudva *et al.* 2012, Hua *et al.* 2013, Teo *et al.* 2013, Thatava *et al.* 2013). These *in vitro* cellular models could also be exploited to better characterize patient-specific features and to perform drug discovery studies.

Even though promising results have already been shown, the differentiation of iPSC into pancreatic progenitors and, more importantly, into glucose-responsive β cells is still undergoing improvement (Hrvatin *et al.* 2014, Pagliuca *et al.* 2014, Rezania *et al.* 2014, Russ *et al.* 2015). In the past decades, different rodent β cell lines were established, allowing a detailed study of rodent β cells, but the generation of functional human β cell lines proved more challenging. Recently, Ravassard *et al.* (2011) applied targeted oncogenesis in human fetal pancreatic buds, which, coupled with grafting into SCID mice, allowed cell maturation and the establishment of the first functional human pancreatic β cell, EndoC-βH1. EndoC-βH1 cells express a number of pancreatic β cell markers but do not show a significant expression of markers of other pancreatic cell types. Furthermore, these cells secrete insulin in a glucose-responsive manner, and their transplantation reverses chemically induced diabetes in mice (Ravassard *et al.* 2011). Subsequent work from the same team allowed fine-tuning of this methodology, resulting in the generation of the EndoC-βH2 line, which allows excision of the transgenes that confer cell immortalization and, hence, a better approximation to the physiological features of true β cells (Scharfmann *et al.* 2014).

Taken together, we expect that these cellular models will allow a deeper understanding of the non-coding regulatory functions of the genome and how *cis*-regulatory networks can be affected by specific sequence variants in the context of the development of common and rare forms of diabetes.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this review.

Funding

L P is a recipient of a Ramon y Cajal contract from the Spanish Ministry of Economy and Competitiveness (RYC 2014-0069) and a Rising Star Award from the European Foundation for the Study of Diabetes (EFSD). CIBER of Diabetes and Associated Metabolic Diseases (CIBERDEM) is an initiative from Instituto de Salud Carlos III.

References

Aguilo F, Zhou M-M & Walsh MJ 2011 Long noncoding RNA, polycomb, and the ghosts haunting INK4b-ARF-INK4a expression. *Cancer Research* **71** 5365–5369. (doi:10.1158/0008-5472.CAN-10-4379)

Arnes L & Sussel L 2015 Epigenetic modifications and long noncoding RNAs influence pancreas development and function. *Trends in Genetics* **31** 290–299. (doi:10.1016/j.tig.2015.02.008)

Avrahami D, Li C, Zhang J, Schug J, Avrahami R, Rao S, Stadler MB, Burger L, Schübeler D, Glaser B & Kaestner KH 2015 Aging-dependent demethylation of regulatory elements correlates with chromatin state and improved β cell function. *Cell Metabolism* **22** 619–632. (doi:10.1016/j.cmet.2015.07.025)

Beck-Nielsen H, Vaag A, Poulsen P & Gaster M 2003 Metabolic and genetic influence on glucose metabolism in type 2 diabetic subjects–experiences from relatives and twin studies. *Best Practice & Research. Clinical Endocrinology & Metabolism* **17** 445–467. (doi:10.1016/S1521-690X(03)00041-1)

Bell AC & Felsenfeld G 2000 Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405** 482–485. (doi:10.1038/35013100)

Bell GI & Polonsky KS 2001 Diabetes mellitus and genetically programmed defects in β-cell function. *Nature* **414** 788–791. (doi:10.1038/414788a)

Bell CG, Finer S, Lindgren CM, Wilson GA, Rakyan VK, Teschendorff AE, Akan P, Stupka E, Down TA *et al.* 2010 Prokopenko I. Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS ONE* **5** e14040. (doi:10.1371/journal.pone.0014040)

Bhandare R, Schug J, Le Lay J, Fox A, Smirnova O, Liu C, Naji A & Kaestner KH 2010 Genome-wide analysis of histone modifications in

Published by Bioscientifica Ltd.

human pancreatic islets. *Genome Research* **20** 428–433. (doi:10.1101/gr.102038.109)

Bonnefond A, Bonnefond A, Clément N, Clément N, Fawcett K, Fawcett K, Yengo L, Vaillant E, Vaillant E, Guillaume J-L *et al.* 2012 Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nature Genetics* **44** 297–301. (doi:10.1038/ng.1053)

Borowiec M, Liew CW, Thompson R, Boonyasrisawat W, Hu J, Mlynarski WM, Khattabi El I, Kim S-H, Marselli L, Rich SS *et al.* 2009 Mutations at the BLK locus linked to maturity onset diabetes of the young and β-cell dysfunction. *PNAS* **106** 14460–14465. (doi:10.1073/pnas.0906474106)

Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S *et al.* 2012 Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research* **22** 1790–1797. (doi:10.1101/gr.137323.112)

van de Bunt M, Gaulton KJ, Parts L, Morán I, Johnson PR, Lindgren CM, Ferrer J, Gloyn AL & McCarthy MI 2013 The miRNA profile of human pancreatic islets and β-cells and relationship to type 2 diabetes pathogenesis. *PLoS ONE* **8** e55272. (doi:10.1371/journal.pone.0055272)

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A & Rinn JL 2011 Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes and Development* **25** 1915–1927. (doi:10.1101/gad.17446611)

Carrasco M, Delgado I, Soria B, Martín F & Rojas A 2012 GATA4 and GATA6 control mouse pancreas organogenesis. *Journal of Clinical Invesigation* **122** 3504–3515. (doi:10.1172/JCI63240)

Cebola I, Rodríguez-Seguí SA, Cho CH-H, Bessa J, Rovira M, Luengo M, Chhatriwala M, Berry A, Ponsa-Cobas J, Maestro MA *et al.* 2015 TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nature Cell Biology* **17** 615–626. (doi:10.1038/ncb3160)

Chambers JC, Chambers JC, Loh M, Loh M, Lehne B, Lehne B, Drong A, Drong A, Kriebel J, Motta V *et al.* 2015 Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes & Endocrinology* **3** 526–534. (doi:10.1016/S2213-8587(15)00127-8)

Cho CHH, Hannan NRF, Docherty FM, Docherty HM, Joåo Lima M, Trotter MWB, Docherty K & Vallier L 2012 Inhibition of activin/nodal signalling is necessary for pancreatic differentiation of human pluripotent stem cells. *Diabetologia* **55** 3284–3295. (doi:10.1007/s00125-012-2687-x)

Claussnitzer M, Dankel SN, Klocke B, Grallert H, Glunk V, Berulava T, Lee H, Oskolkov N, Fadista J, Ehlers K *et al.* 2014 Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156** 343–358. (doi:10.1016/j.cell.2013.10.058)

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.* 2013 Multiplex genome engineering using CRISPR/Cas systems. *Science* **339** 819–823. (doi:10.1126/science.1231143)

Corradin O, Saiakhova A, Akhtar-Zaidi B, Myeroff L, Willis J, Lari RC-S, Lupien M, Markowitz S & Scacheri PC 2014 Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Research* **24** 1–13. (doi:10.1101/gr.164079.113)

Coventry A, Bull-Otterson LM, Liu X, Clark AG, Maxwell TJ, Crosby J, Hixson JE, Rea TJ, Muzny DM, Lewis LR *et al.* 2010 Deep resequencing reveals excess rare recent variants consistent with explosive population growth. *Nature Communications* **1** 131. (doi:10.1038/ncomms1130)

Dayeh TA, Olsson AH, Volkov P, Almgren P, Rönn T & Ling C 2013 Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia* **56** 1036–1046. (doi:10.1007/s00125-012-2815-7)

Dayeh T, Volkov P, Salö S, Hall E, Nilsson E, Olsson AH, Kirkpatrick CL, Wollheim CB, Eliasson L, Rönn T *et al.* 2014 Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genetics* **10** e1004160. (doi:10.1371/journal.pgen.1004160)

Dhawan S, Tschen S-I, Zeng C, Guo T, Hebrok M, Matveyenko A & Bhushan A 2015 DNA methylation directs functional maturation of pancreatic β cells. *Journal of Clinical Invesigation* **125** 2851–2860. (doi:10.1172/JCI79956)

Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, Meduri E, Morange P-E, Gagnon F, Grallert H *et al.* 2014 DNA methylation and body-mass index: a genome-wide analysis. *Lancet* **383** 1990–1998. (doi:10.1016/S0140-6736(13)62674-4)

Dillon N 2006 Gene regulation and large-scale chromatin organization in the nucleus. *Chromosome Research* **14** 117–126. (doi:10.1007/s10577-006-1027-8)

Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M *et al.* 2012 Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485** 201–206. (doi:10.1038/nature11112)

Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL *et al.* 2010 New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics* **42** 105–116. (doi:10.1038/ng.520)

Ellard S, Lango Allen H, De Franco E, Flanagan SE, Hysenaj G, Colclough K, Houghton JAL, Shepherd M, Hattersley AT, Weedon MN *et al.* 2013 Improved genetic testing for monogenic diabetes using targeted next-generation sequencing. *Diabetologia* **56** 1958–1963. (doi:10.1007/s00125-013-2962-5)

ENCODE Project Consortium , Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M *et al.* 2012 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489** 57–74. (doi:10.1038/nature11247)

Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M *et al.* 2012 Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473** 43–49. (doi:10.1038/nature09906)

Esteller M 2011 Non-coding RNAs in human disease. *Nature Review Genetics* **12** 861–874. (doi:10.1038/nrg3074)

Fadista J, Vikman P, Laakso EO, Mollet IG, Esguerra JL, Taneera J, Storm P, Osmark P, Ladenvall C, Prasad RB *et al.* 2014 Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *PNAS* **111** 13924–13929. (doi:10.1073/pnas.1402665111)

Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJH, Shishkin AA *et al.* 2015 Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518** 337–343. (doi:10.1038/nature13835)

Flanagan SE, De Franco E, Lango Allen H, Zerah M, Abdul-Rasoul MM, Edge JA, Stewart H, Alamiri E, Hussain K, Wallis S *et al.* 2014 Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. *Cell Metabolism* **19** 146–154. (doi:10.1016/j.cmet.2013.11.021)

Florez JC, Hirschhorn J & Altshuler D 2003 The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annual Review of Genomics and Human Genetics* **4** 257–291. (doi:10.1146/annurev.genom.4.070802.110436)

Fogarty MP, Panhuis TM, Vadlamudi S, Buchkovich ML & Mohlke KL 2013 Allele-specific transcriptional activity at type 2 diabetes-associated single nucleotide polymorphisms in regions of pancreatic islet open chromatin at the JAZF1 locus. *Diabetes* **62** 1756–1762. (doi:10.2337/db12-0972)

Fogarty MP, Cannon ME, Vadlamudi S, Gaulton KJ & Mohlke KL 2014 Identification of a regulatory variant that binds FOXA1 and FOXA2 at the CDC123/CAMK1D type 2 diabetes GWAS locus. *PLoS Genetics* **10** e1004633. (doi:10.1371/journal.pgen.1004633)

Fu Y, Dominissini D, Rechavi G & He C 2014 Gene expression regulation mediated through reversible m6A RNA methylation. *Nature Review Genetics* **15** 293–306. (doi:10.1038/nrg3724)

Fustin J-M, Doi M, Yamaguchi Y, Hida H, Nishimura S, Yoshida M, Isagawa T, Morioka MS, Kakeya H, Manabe I *et al.* 2013 RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **155** 793–806. (doi:10.1016/j.cell.2013.10.026)

Gao N, LeLay J, Vatamaniuk MZ, Rieck S, Friedman JR & Kaestner KH 2008 Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes and Development* **22** 3435–3448. (doi:10.1101/gad.1752608)

Garin I, Edghill EL, Akerman I, Rubio-Cabezas O, Rica I, Locke JM, Maestro MA, Alshaikh A, Bundak R, del Castillo G *et al.* 2010 Recessive mutations in the INS gene result in neonatal diabetes through reduced insulin biosynthesis. *PNAS* **107** 3105–3110. (doi:10.1073/pnas.0910533107)

Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG, Fogarty MP, Panhuis TM, Mieczkowski P, Secchi A, Bosco D *et al.* 2010 A map of open chromatin in human pancreatic islets. *Nature Genetics* **42** 255–259. (doi:10.1038/ng.530)

Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP & Bickmore WA 2004 Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* **118** 555–566. (doi:10.1016/j.cell.2004.08.011)

de Giorgio A, Krell J, Harding V, Stebbing J & Castellano L 2013 Emerging roles of competing endogenous RNAs in cancer: insights from the regulation of PTEN. *Molecular and Cell Biology* **33** 3976–3982. (doi:10.1128/MCB.00683-13)

Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U & Shaw JE 2014 Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Research and Clinical Practice* **103** 137–149. (doi:10.1016/j.diabres.2013.11.002)

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, deKlein A, Wessels L, de Laat W *et al.* 2008 Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453** 948–951. (doi:10.1038/nature06947)

Gulko B, Hubisz MJ, Gronau I & Siepel A 2015 A method for calculating probabilities of fitness consequences for point mutations across the human genome. *Nature Genetics* **47** 276–283. (doi:10.1038/ng.3196)

Gusev A, Lee SH, Trynka G, Finucane H, Vilhjálmsson BJ, Xu H, Zang C, Ripke S, Bulik-Sullivan B, Stahl E *et al.* 2014 Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *American Journal of Human Genetics* **95** 535–552. (doi:10.1016/j.ajhg.2014.10.004)

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP *et al.* 2009 Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458** 223–227. (doi:10.1038/nature07672)

Hansen TH, Vestergaard H, Jørgensen T, Jørgensen ME, Lauritzen T, Brandslund I, Christensen C, Pedersen O, Hansen T & Gjesing AP 2015 Impact of PTBP1 rs11085226 on glucose-stimulated insulin release in adult Danes. *BMC Medical Genetics* **16** 4. (doi:10.1186/s12863-015-0162-7)

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM & Tilghman SM 2000 CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405** 486–489. (doi:10.1038/35013106)

Haumaitre C, Barbacci E, Jenny M, Ott MO, Gradwohl G & Cereghini S 2005 Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *PNAS* **102** 1490–1495. (doi:10.1073/pnas.0405776102)

Hidalgo B, Irvin MR, Sha J, Zhi D, Aslibekyan S, Absher D, Tiwari HK, Kabagambe EK, Ordovas JM & Arnett DK 2014 Epigenome-wide association study of fasting measures of glucose, insulin, and HOMA-IR in the Genetics of Lipid Lowering Drugs and Diet Network study. *Diabetes* **63** 801–807. (doi:10.2337/db13-1100)

Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE & Gersbach CA 2015 Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology* **33** 510–517. (doi:10.1038/nbt.3199)

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA & Young RA 2013 Super-enhancers in the control of cell identity and disease. *Cell* **155** 934–947. (doi:10.1016/j.cell.2013.09.053)

Hon GC, Song C-X, Du T, Jin F, Selvaraj S, Lee AY, Yen C-A, Ye Z, Mao S-Q, Wang B-A *et al.* 2014 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Molecular Cell* **56** 286–297. (doi:10.1016/j.molcel.2014.08.026)

Hrvatin S, O'Donnell CW, Deng F, Millman JR, Pagliuca FW, DiIorio P, Rezania A, Gifford DK & Melton DA 2014 Differentiated human stem cells resemble fetal, not adult, β cells. *PNAS* **111** 3038–3043. (doi:10.1073/pnas.1400709111)

Hua H, Hua H, Shang L, Shang L, Martinez H, Martinez H, Freeby M, Freeby M, Gallagher MP, Gallagher MP *et al.* 2013 iPSC-derived β cells model diabetes due to glucokinase deficiency. *Journal of Clinical Invesigation* **123** 3146–3153. (doi:10.1172/JCI67638)

Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S *et al.* 2015 The landscape of long noncoding RNAs in the human transcriptome. *Nature Genetics* **47** 199–208. (doi:10.1038/ng.3192)

Jacquemin P, Durviaux SM, Jensen J, Godfraind C, Gradwohl G, Guillemot F, Madsen OD, Carmeliet P, Dewerchin M, Collen D *et al.* 2000 Transcription factor hepatocyte nuclear factor 6 regulates pancreatic endocrine cell differentiation and controls expression of the proendocrine gene ngn3. *Molecular and Cell Biology* **20** 4445–4454. (doi:10.1128/MCB.20.12.4445-4454.2000)

Jemal A, Ward E, Hao Y & Thun M 2005 Trends in the leading causes of death in the United States, 1970–2002. *Journal of the American Medical Association* **294** 1255–1259. (doi:10.1001/jama.294.10.1255)

Jiang X, Ma H, Wang Y & Liu Y 2013 Early life factors and type 2 diabetes mellitus. *Journal of Diabetes Research* **2013** article 485082. (doi:10.1155/2013/485082)

Jones PA 2012 Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Review Genetics* **13** 484–492. (doi:10.1038/nrg3230)

Kameswaran V & Kaestner KH 2014 The missing lnc(RNA) between the pancreatic β-cell and diabetes. *Frontiers in Genetics* **5** 200. (doi:10.3389/fgene.2014.00200)

Kameswaran V, Bramswig NC, McKenna LB, Penn M, Schug J, Hand NJ, Chen Y, Choi I, Vourekas A, Won K-J *et al.* 2014 Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. *Cell Metabolism* **19** 135–145. (doi:10.1016/j.cmet.2013.11.016)

Karambataki M, Malousi A & Kouidou S 2014 Risk-associated coding synonymous SNPs in type 2 diabetes and neurodegenerative diseases: genetic silence and the underrated association with splicing regulation and epigenetics. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **770** 85–93. (doi:10.1016/j.mrfmmm.2014.09.005)

Kaspi H, Pasvolsky R & Hornstein E 2014 Could microRNAs contribute to the maintenance of β cell identity? *Trends in Endocrinology and Metabolism* **25** 285–292. (doi:10.1016/j.tem.2014.01.003)

Keene JD 2007 RNA regulons: coordination of post-transcriptional events. *Nature Review Genetics* **8** 533–543. (doi:10.1038/nrg2111)

Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J *et al.* 2014 Defining functional DNA elements in the human genome. *PNAS* **111** 6131–6138. (doi:10.1073/pnas.1318948111)

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, iveaMorales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A *et al.* 2009 Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *PNAS* **106** 11667–11672. (doi:10.1073/pnas.0904715106)

Khoo C, Yang J, Weinrott SA, Kaestner KH, Naji A, Schug J & Stoffers DA 2012 Research resource: the pdx1 cistrome of pancreatic islets. *Molecular Endocrinology* **26** 521–533. (doi:10.1210/me.2011-1231)

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM & Shendure J 2014 A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics* **46** 310–315. (doi:10.1038/ng.2892)

Köbberling J & Tillil H 1990 Genetic and nutritional factors in the etiology and pathogenesis of diabetes mellitus. *World Review of Nutrition and Dietetics* **63** 102–115.

Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J *et al.* 2013 Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493** 231–235. (doi:10.1038/nature11661)

Krishnamurthy J, Ramsey MR, Ligon KL, Torrice C, Koh A, Bonner-Weir S & Sharpless NE 2006 p16INK4a induces an age-dependent decline in islet regenerative potential. *Nature* **443** 453–457. (doi:10.1038/nature05092)

Kudva YC, Ohmine S, Greder LV, Dutton JR, Armstrong A, DeLamo JG, Khan YK, Thatava T, Hasegawa M, Fusaki N *et al.* 2012 Transgene-free disease-specific induced pluripotent stem cells from patients with type 1 and type 2 diabetes. *Stem Cells Translational Medicine* **1** 451–461. (doi:10.5966/sctm.2011-0044)

Kulkarni H, Kos MZ, Neary J, Dyer TD, Göring HHH, Cole SA, Comuzzie AG, Almasy L, Mahaney MC, Curran JE *et al.* 2015 Novel epigenetic determinants of type 2 diabetes in mexican american families. *Human Molecular Genetics* **24** 5330–5344. (doi:10.1093/hmg/ddv232)

Kulzer JR, Stitzel ML, Morken MA, Huyghe JR, Fuchsberger C, Kuusisto J, Laakso M, Boehnke M, Collins FS & Mohlke KL 2014 A common functional regulatory variant at a type 2 diabetes locus upregulates ARAP1 expression in the pancreatic β cell. *American Journal of Human Genetics* **94** 186–197. (doi:10.1016/j.ajhg.2013.12.011)

Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA & Shiekhattar R 2013 Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* **494** 497–501. (doi:10.1038/nature11884)

Lango Allen H, Flanagan SE, Shaw-Smith C, DeFranco E, Akerman I, Caswell R, International Pancreatic Agenesis Consortium, Ferrer J, Hattersley AT & Ellard S 2012 GATA6 haploinsufficiency causes pancreatic agenesis in humans. *Nature Genetics* **44** 20–22. (doi:10.1038/ng.1035)

Latreille M, Hausser J, Stützer I, Zhang Q, Hastoy B, Gargani S, Kerr-Conte J, Pattou F, Zavolan M, Esguerra JLS *et al.* 2014 MicroRNA-7a regulates pancreatic β cell function. *Journal of Clinical Invesigation* **124** 2722–2735. (doi:10.1172/JCI73066)

Lee D, Gorkin DU, Baker M, Strober BJ, Asoni AL, McCallion AS & Beer MA 2015 A method to predict the impact of regulatory variants from DNA sequence. *Nature Genetics* **47** 955–961. (doi:10.1038/ng.3331)

Ling C, Del Guerra S, Lupi R, Rönn T, Granhall C, Luthman H, Masiello P, Marchetti P, Groop L & Del Prato S 2008 Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia* **51** 615–622. (doi:10.1007/s00125-007-0916-5)

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, Human DNA *et al.* 2009 methylomes at base resolution show widespread epigenomic differences. *Nature* **462** 315–322. (doi:10.1038/nature08514)

Locke JM, da Silva Xavier G, Dawe HR, Rutter GA & Harries LW 2014 Increased expression of miR-187 in human islets from individuals with type 2 diabetes is associated with reduced glucose-stimulated insulin secretion. *Diabetologia* **57** 122–128. (doi:10.1007/s00125-013-3089-4)

Locke JM, Hysenaj G, Wood AR, Weedon MN & Harries LW 2015 Targeted allelic expression profiling in human islets identifies cis-regulatory effects for multiple variants identified by type 2 diabetes genome-wide association studies. *Diabetes* **64** 1484–1491. (doi:10.2337/db14-0957)

Lu F, Liu Y, Jiang L, Yamaguchi S & Zhang Y 2014 Role of Tet proteins in enhancer activity and telomere elongation. *Genes and Development* **28** 2103–2119. (doi:10.1101/gad.248005.114)

Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R *et al.* 2015 Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161** 1012–1025. (doi:10.1016/j.cell.2015.04.004)

Lynn FC, Skewes-Cox P, Kosaka Y, McManus MT, Harfe BD & German MS 2007 MicroRNA expression is required for pancreatic islet cell genesis in the mouse. *Diabetes* **56** 2938–2945. (doi:10.2337/db07-0175)

Maas S, Kawahara Y, Tamburro KM & Nishikura K 2006 A-to-I RNA editing and human disease. *RNA Biology* **3** 1–9. (doi:10.4161/rna.3.1.2495)

Maehr R, Chen S, Snitow M, Snitow M, Ludwig T, Ludwig T, Yagasaki L, Yagasaki L, Goland R, Goland R *et al.* 2009 Generation of pluripotent stem cells from patients with type 1 diabetes. *PNAS* **106** 15768–15773. (doi:10.1073/pnas.0906894106)

Magro MG & Solimena M 2013 Regulation of β-cell function by RNA-binding proteins. *Molecular Metabolism* **2** 348–355. (doi:10.1016/j.molmet.2013.09.003)

Manning AK, Hivert M-F, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, Rybin D, Liu C-T, Bielak LF, Prokopenko I *et al.* 2012 A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature Genetics* **44** 659–669. (doi:10.1038/ng.2274)

Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J *et al.* 2012 Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337** 1190–1195. (doi:10.1126/science.1222794)

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA & Hirschhorn JN 2008 Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Review Genetics* **9** 356–369. (doi:10.1038/nrg2344)

Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ *et al.* 2015 Human genomics. The human transcriptome across tissues and individuals. *Science* **348** 660–665. (doi:10.1126/science.aaa0355)

Melkman-Zehavi T, Oren R, Kredo-Russo S, Shapira T, Mandelbaum AD, Rivkin N, Nir T, Lennox KA, Behlke MA, Dor Y *et al.* 2011 miRNAs control insulin content in pancreatic β-cells via downregulation of transcriptional repressors. *EMBO Journal* **30** 835–845. (doi:10.1038/emboj.2010.361)

Morán I, Akerman I, van de Bunt M, Xie R, Benazra M, Nammo T, Arnes L, Nakić N, García-Hurtado J, Ríodrguez-Seguí S *et al.* 2012 Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metabolism* **16** 435–448. (doi:10.1016/j.cmet.2012.08.010)

Nair G & Hebrok M 2015 Islet formation in mice and men: lessons for the generation of functional insulin-producing β-cells from human pluripotent stem cells. *Current Opinion in Genetics & Development* **32** 171–180. (doi:10.1016/j.gde.2015.03.004)

Nica AC, Ongen H, Irminger J-C, Bosco D, Berney T, Antonarakis SE, Halban PA & Dermitzakis ET 2013 Cell-type, allelic, and genetic signatures in the human pancreatic β cell transcriptome. *Genome Research* **23** 1554–1562. (doi:10.1101/gr.150706.112)

Offield MF, Jetton TL, Labosky PA, Ray M, Stein RW, Magnuson MA, Hogan BL & Wright CV 1996 PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* **122** 983–995.

Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C & Huang Q 2010 Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143** 46–58. (doi:10.1016/j.cell.2010.09.001)

Orozco LD, Morselli M, Rubbi L, Guo W, Go J, Shi H, Lopez D, Furlotte NA, Bennett BJ, Farber CR *et al.* 2015 Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell Metabolism* **21** 905–917. (doi:10.1016/j.cmet.2015.04.025)

Ouaamari EA, Baroukh N, Martens GA, Lebrun P, Pipeleers D & vanObberghen E 2008 miR-375 targets 3′-phosphoinositide-dependent

Published by Bioscientifica Ltd.

protein kinase-1 and regulates glucose-induced biological responses in pancreatic β-cells. *Diabetes* **57** 2708–2717. (doi:10.2337/db07-1614)

Pagliuca FW, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, Peterson QP, Greiner D & Melton DA 2014 Generation of functional human pancreatic β cells in vitro. *Cell* **159** 428–439. (doi:10.1016/j.cell.2014.09.040)

Parker SCJ, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, vanBueren KL, Chines PS, Narisu N, NISC Comparative Sequencing Program *et al.* 2013 Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *PNAS* **110** 17921–17926. (doi:10.1073/pnas.1317023110)

Pasmant E, Sabbagh A, Vidaud M & Bièche I 2011 ANRIL, a long, noncoding RNA, is an unexpected major hotspot in GWAS. *FASEB Journal* **25** 444–448. (doi:10.1096/fj.10-172452)

Pasquali L, Gaulton KJ, Roídrguez-Seguí SA, Mularoni L, Miguel-Escalada I, Akerman I, Tena JJ, Morán I, Gómez-Marín C, van deunt M *et al.* 2014 Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature Genetics* **46** 136–143. (doi:10.1038/ng.2870)

Pennacchio LA & Visel A 2010 Limits of sequence and functional conservation. *Nature Genetics* **42** 557–558. (doi:10.1038/ng0710-557)

Perry JRB & Frayling TM 2008 New gene variants alter type 2 diabetes risk predominantly through reduced β-cell function. *Current Opinion in Clinical Nutrition and Metabolic Care* **11** 371–377. (doi:10.1097/MCO.0b013e32830349a1)

Petersen A-K, Zeilinger S, Kastenmüller G, Römisch-Margl W, Brugger M, Peters A, Meisinger C, Strauch K, Hengstenberg C, Pagel P *et al.* 2014 Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits. *Human Molecular Genetics* **23** 534–545. (doi:10.1093/hmg/ddt430)

Poulsen P, Kyvik KO, Vaag A & Beck-Nielsen H 1999 Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance – a population-based twin study. *Diabetologia* **42** 139–145. (doi:10.1007/s001250051131)

Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, MacDonald PE, Pfeffer S, Tuschl T, Rajewsky N, Rorsman P *et al.* 2004 A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432** 226–230. (doi:10.1038/nature03076)

Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, Vergara IA, Davicioni E, Erho N, Ghadessi M *et al.* 2013 The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nature Genetics* **45** 1392–1398. (doi:10.1038/ng.2771)

Pullen TJ & Rutter GA 2014 Roles of lncRNAs in pancreatic β cell identity and diabetes susceptibility. *Frontiers in Genetics* **5** 193. (doi:10.3389/fgene.2014.00193)

Ravassard P, Hazhouz Y, Pechberty S, Bricout-Neveu E, Armanet M, Czernichow P & Scharfmann R 2011 A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *Journal of Clinical Investigation* **121** 3589–3597. (doi:10.1172/JCI58447)

Rezania A, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, O'Dwyer S, Quiskamp N, Mojibian M, Albrecht T *et al.* 2014 Reversal of diabetes with insulin-producing cells derived *in vitro* from human pluripotent stem cells. *Nature Biotechnology* **32** 1121–1133. (doi:10.1038/nbt.3033)

Rinn JL & Chang HY 2012 Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry* **81** 145–166. (doi:10.1146/annurev-biochem-051410-092902)

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E *et al.* 2007 Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129** 1311–1323. (doi:10.1016/j.cell.2007.05.022)

Ritchie GRS, Dunham I, Zeggini E & Flicek P 2014 Functional annotation of noncoding sequence variants. *Nature Methods* **11** 294–296. (doi:10.1038/nmeth.2832)

Roadmap Epigenomics Consortium , Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z & Wang J 2015 Integrative analysis of 111 reference human epigenomes. *Nature* **518** 317–330. (doi:10.1038/nature14248)

Robertson KD 2005 DNA methylation and human disease. *Nature Review Genetics* **6** 597–610. (doi:10.1038/nrg1655)

Russ HA, Russ HA, Parent AV, Parent AV, Ringler JJ, Ringler JJ, Hennings TG, Hennings TG, Nair GG, Nair GG *et al.* 2015 Controlled induction of human pancreatic progenitors produces functional β-like cells *in vitro*. *EMBO Journal* **34** 1759–1772. (doi:10.15252/embj.201591058)

Salmena L, Poliseno L, Tay Y, Kats L & Pandolfi PP 2011 A ceRNA hypothesis: the rosetta stone of a hidden RNA language? *Cell* **146** 353–358. (doi:10.1016/j.cell.2011.07.014)

Scharfmann R, Pechberty S, Hazhouz Y, Bülow von M, Bricout-Neveu E, Grenier-Godard M, Guez F, Rachdi L, Lohmann M, Czernichow P *et al.* 2014 Development of a conditionally immortalized human pancreatic β cell line. *Journal of Clinical Invesigation* **124** 2087–2098. (doi:10.1172/JCI72674)

Schork NJ, Murray SS, Frazer KA & Topol EJ 2009 Common vs. rare allele hypotheses for complex diseases. *Current Opinion in Genetics & Development* **19** 212–219. (doi:10.1016/j.gde.2009.04.010)

Schübeler D 2015 Function and information content of DNA methylation. *Nature* **517** 321–326. (doi:10.1038/nature14192)

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H *et al.* 2015 Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523** 212–216. (doi:10.1038/nature14465)

Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, Mägi R, Strawbridge RJ, Rehnberg E, Gustafsson S *et al.* 2012 Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature Genetics* **44** 991–1005. (doi:10.1038/ng.2385)

Sellick GS, Barker KT, Stolte-Dijkstra I, Fleischmann C, Coleman RJ, Garrett C, Gloyn AL, Edghill EL, Hattersley AT, Wellauer PK *et al.* 2004 Mutations in PTF1A cause pancreatic and cerebellar agenesis. *Nature Genetics* **36** 1301–1305. (doi:10.1038/ng1475)

Seymour PA, Freude KK, Tran MN, Mayes EE, Jensen J, Kist R, Scherer G & Sander M 2007 SOX9 is required for maintenance of the pancreatic progenitor cell pool. *PNAS* **104** 1865–1870. (doi:10.1073/pnas.0609217104)

Shaw-Smith C, De Franco E, Allen HL, Batlle M, Flanagan SE, Borowiec M, Taplin CE, van Alfen-van der Velden J, Cruz-Rojo J, de Nanclares GP *et al.* 2014 GATA4 mutations are a cause of neonatal and childhood-onset diabetes. *Diabetes* **63** DB_140061–DB_142894. (doi:10.2337/db14-0061)

Shen F, Huang W, Huang J-T, Xiong J, Yang Y, Wu K, Jia G-F, Chen J, Feng Y-Q, Yuan B-F *et al.* 2014 Decreased N6-methyladenosine in peripheral blood RNA from diabetic patients Is associated with FTO expression rather than ALKBH5. *Journal of Clinical Endocrinology and Metabolism* **100** E148–E154. (doi:10.1210/jc.2014-1893)

Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT & Ellard S 2010 Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* **53** 2504–2508. (doi:10.1007/s00125-010-1799-4)

Shlyueva D, Stampfel G & Stark A 2014 Transcriptional enhancers: from properties to genome-wide predictions. *Nature Review Genetics* **15** 272–286. (doi:10.1038/nrg3682)

Siddiqui K, Musambil M & Nazir N 2015 Maturity onset diabetes of the young (MODY)–history, first case reports and recent advances. *Gene* **555** 66–71. (doi:10.1016/j.gene.2014.09.062)

Smemo S, Tena JJ, Kim K-H, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF *et al.* 2014 Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507** 371–375. (doi:10.1038/nature13138)

Song L, Zhang Z, Grasfeder LL, Boyle AP, Giresi PG, Lee B-K, Sheffield NC, Gräf S, Huss M, Keefe D *et al.* 2011 Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Research* **21** 1757–1767. (doi:10.1101/gr.121541.111)

http://jme.endocrinology-journals.org
DOI: 10.1530/JME-15-0197

Published by Bioscientifica Ltd.

Stitzel ML, Sethupathy P, Pearson DS, Chines PS, Song L, Erdos MR, Welch R, Parker SCJ, Boyle AP, Scott LJ *et al.* 2010 Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metabolism* **12** 443–455. (doi:10.1016/j.cmet.2010.09.012)

Stoffers DA, Zinkin NT, Stanojevic V, Clarke WL & Habener JF 1997 Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nature Genetics* **15** 106–110. (doi:10.1038/ng0197-106)

Taylor BL, Liu F-F & Sander M 2013 Nkx6.1 is essential for maintaining the functional state of pancreatic β cells. *Cell Reports* **4** 1262–1275. (doi:10.1016/j.celrep.2013.08.010)

Teo AKK, Teo AKK, Windmueller R, Windmueller R, Johansson BB, Johansson BB, Dirice E, Njolstad PR, Njolstad PR, Tjora E *et al.* 2013 Derivation of human induced pluripotent stem cells from patients with maturity onset diabetes of the young. *Journal of Biological Chemistry* **288** 5353–5356. (doi:10.1074/jbc.C112.428979)

Thatava T, Kudva YC, Edukulla R, Edukulla R, Squillace K, Squillace K, DeLamo JG, Khan YK, Sakuma T, Sakuma T *et al.* 2013 Intrapatient variations in type 1 diabetes-specific iPS cell differentiation into insulin-producing cells. *Molecular Therapy* **21** 228–239. (doi:10.1038/mt.2012.245)

The FANTOM Consortium 2005 The transcriptional landscape of the mammalian genome. *Science* **309** 1559–1563. (doi:10.1126/science.1112014)

Tillmar L, Carlsson C & Welsh N 2002 Control of insulin mRNA stability in rat pancreatic islets. Regulatory role of a 3′-untranslated region pyrimidine-rich sequence. *Journal of Biological Chemistry* **277** 1099–1106. (doi:10.1074/jbc.M108340200)

Vaxillaire M, Bonnefond A & Froguel P 2012 The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. *Best Practice & Research. Clinical Endocrinology & Metabolism* **26** 171–187. (doi:10.1016/j.beem.2011.12.001)

Volkmar M, Dedeurwaerder S, Cunha DA, Ndlovu MN, Defrance M, Deplus R, Calonne E, Volkmar U, Igoillo-Esteve M, Naamane N *et al.* 2012 DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. *EMBO Journal* **31** 1405–1426. (doi:10.1038/emboj.2011.503)

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA *et al.* 2011 A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472** 120–124. (doi:10.1038/nature09819)

Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, Wu M, Xiong J, Guo X & Liu H 2013 Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog and Sox2 in human embryonic stem cell self-renewal. *Developmental Cell* **25** 69–80. (doi:10.1016/j.devcel.2013.03.002)

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G *et al.* 2014 N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505** 117–120. (doi:10.1038/nature12730)

Wang A, Yue F, Li Y, Xie R, Harper T, Patel NA, Muth K, Palmer J, Qiu Y, Wang J *et al.* 2015*a* Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* **16** 386–399. (doi:10.1016/j.stem.2015.02.013)

Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H & He C 2015*b* N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161** 1388–1399. (doi:10.1016/j.cell.2015.05.014)

Weedon MN, Cebola I, Flanagan SE, DeFranco E, Caswell R, Roídrguez-Seguí SA, Shaw-Smith C, Lango Allen H, Vallier L, International Pancreatic Agenesis Consortium *et al.* 2014 Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. *Nature Genetics* **46** 61–64. (doi:10.1038/ng.2826)

Wicksteed B, Uchizono Y, Alarcon C, McCuaig JF, Shalev A & Rhodes CJ 2007 A cis-element in the 5′ untranslated region of the preproinsulin mRNA (ppIGE) is required for glucose regulation of proinsulin translation. *Cell Metabolism* **5** 221–227. (doi:10.1016/j.cmet.2007.02.007)

Wiench M, John S, Baek S, Johnson TA, Sung M-H, Escobar T, Simmons CA, Pearce KH, Biddie SC, Sabo PJ *et al.* 2011 DNA methylation status predicts cell type-specific enhancer activity. *EMBO Journal* **30** 3028–3039. (doi:10.1038/emboj.2011.210)

Willems SM, Mihaescu R, Sijbrands EJG, van Duijn CM & Janssens ACJW 2011 A methodological perspective on genetic risk prediction studies in type 2 diabetes: recommendations for future research. *Current Diabetes Reports* **11** 511–518. (doi:10.1007/s11892-011-0235-6)

Wirsing A, Senkel S, Klein-Hitpass L & Ryffel GU 2011 A systematic analysis of the 3′UTR of HNF4A mRNA reveals an interplay of regulatory elements including miRNA target sites. *PLoS ONE* **6** e27438. (doi:10.1371/journal.pone.0027438)

Wu D, Yang G, Zhang L, Xue J, Wen Z & Li M 2014 Genome-wide association study combined with biological context can reveal more disease-related SNPs altering microRNA target seed sites. *BMC Genomics* **15** 669. (doi:10.1186/1471-2164-15-669)

Xing Z, Lin A, Li C, Liang K, Wang S, Liu Y, Park PK, Qin L, Wei Y, Hawke DH *et al.* 2014 lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. *Cell* **159** 1110–1125. (doi:10.1016/j.cell.2014.10.013)

Xuan S, Borok MJ, Decker KJ, Battle MA, Duncan SA, Hale MA, Macdonald RJ & Sussel L 2012 Pancreas-specific deletion of mouse Gata4 and Gata6 causes pancreatic agenesis. *Journal of Clinical Invesigation* **122** 3516–3528. (doi:10.1172/JCI63352)

Yamanaka S 2007 Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell* **1** 39–49. (doi:10.1016/j.stem.2007.05.012)

Yang BT, Dayeh TA, Volkov PA, Kirkpatrick CL, Malmgren S, Jing X, Renström E, Wollheim CB, Nitert MD & Ling C 2012 Increased DNA methylation and decreased expression of PDX-1 in pancreatic islets from patients with type 2 diabetes. *Molecular Endocinology* **26** 1203–1212. (doi:10.1210/me.2012-1004)

Yap KL, Li S, Muñoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ & Zhou M-M 2010 Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular Cell* **38** 662–674. (doi:10.1016/j.molcel.2010.03.021)

Yuan W, Xia Y, Bell CG, Yet I, Ferreira T, Ward KJ, Gao F, Loomis AK, Hyde CL, Wu H *et al.* 2014 An integrated epigenomic analysis for type 2 diabetes susceptibility loci in monozygotic twins. *Nature Communications* **5** 5719. (doi:10.1038/ncomms6719)

Received in final form 30 September 2015
Accepted 2 October 2015
Accepted Preprint published online 2 October 2015

Published by Bioscientifica Ltd.