



In Vivo Reporter Assays Uncover Changes in Enhancer Activity Caused by Type 2 Diabetes–Associated Single Nucleotide Polymorphisms

Ana Eufrásio, Chiara Perrod, Fábio J. Ferreira, Marta Duque, Mafalda Galhardo, and José Bessa

Diabetes 2020;69:2794–2805 | <https://doi.org/10.2337/db19-1049>

Many single nucleotide polymorphisms (SNPs) associated with type 2 diabetes overlap with putative endocrine pancreatic enhancers, suggesting that these SNPs modulate enhancer activity and, consequently, gene expression. We performed in vivo mosaic transgenesis assays in zebrafish to quantitatively test the enhancer activity of type 2 diabetes–associated loci. Six out of 10 tested sequences are endocrine pancreatic enhancers. The risk variant of two sequences decreased enhancer activity, while in another two incremented it. One of the latter (rs13266634) locates in an *SLC30A8* exon, encoding a tryptophan-to-arginine substitution that decreases *SLC30A8* function, which is the canonical explanation for type 2 diabetes risk association. However, other type 2 diabetes–associated SNPs that truncate *SLC30A8* confer protection from this disease, contradicting this explanation. Here, we clarify this incongruence, showing that rs13266634 boosts the activity of an overlapping enhancer and suggesting an *SLC30A8* gain of function as the cause for the increased risk for the disease. We further dissected the functionality of this enhancer, finding a single nucleotide mutation sufficient to impair its activity. Overall, this work assesses in vivo the importance of disease-associated SNPs in the activity of endocrine pancreatic enhancers, including a poorly explored case where a coding SNP modulates the activity of an enhancer.

Type 2 diabetes affects >300 million people, causing severe complications and premature death (1), yet the underlying molecular mechanisms are largely unknown. This disease is highly complex, multifactorial, and partially characterized by endocrine pancreatic dysfunction, leading to insufficient insulin production (1). Genome-wide association studies

(GWASs) have identified several single nucleotide polymorphisms (SNPs) associated with an increased risk of type 2 diabetes (2–4). Part of these variants are located in non-coding sequences with epigenetic marks associated to enhancer activity known to regulate the expression of their target genes by interacting with their promoters (5) and some overlap with transcription factor (TF) binding sites (TFBSs) required for proper endocrine pancreatic function (6–10). In this way, type 2 diabetes–associated SNPs may ultimately translate into transcriptional changes of the target genes (6–10). Methods to predict enhancers include profiling chromatin accessibility (11) and histone modifications (e.g., H3K27ac, H3K4me1) (12). The majority of the enhancer testing assays are performed in vitro in specific cell lines, missing cellular diversity and physiological contexts. To overcome these limitations, animal models have been used (13,14). The zebrafish has been successfully used for the study of pancreas development and function (15), having an endocrine compartment with the same cell types (α -, β -, δ -, and ϵ -cells) and functions as in mammal pancreas (16,17). Additionally, orthologous TFs operate in the zebrafish pancreas during early development, some of which are also important for adult pancreas maintenance. As in humans, Pdx1 plays a crucial role in zebrafish pancreas development and β -cell maturation and function (18), and Nkx6.1 is required for the identity of endocrine pancreatic progenitors (19,20).

Some coding mutations are associated with the development of type 2 diabetes, while others might confer a protective effect (2,21). Interestingly, *SLC30A8*, a zinc transporter–encoding gene, shows contradicting results. The coding SNP rs13266634 located in the *SLC30A8* gene is associated with an increased risk for type 2 diabetes (3)

i3S–Instituto de Investigação e Inovação em Saúde, Universidade do Porto, and IBMC–Instituto de Biologia Celular e Molecular, Porto, Portugal

Corresponding author: José Bessa, jose.bessa@ibmc.up.pt

Received 17 October 2019 and accepted 2 September 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.12906803>.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

because of a tryptophan-to-arginine switch at protein position 325, causing reduced zinc transport activity (22,23). Zinc is essential for insulin packaging and maturation and secretion in β -cells (24); thus, the decrease of *SLC30A8* activity is the simplest explanation for the increased risk for the disease. Surprisingly, recently identified protein-truncating SNPs in *SLC30A8* have been associated with a protective effect (2,25) by enhanced insulin secretion (26). Further work is needed to clarify the type 2 diabetes association with *SLC30A8* SNPs. An unexplored explanation for these apparently contradicting results could be that the coding SNP rs13266634 exerts a specific impact on adjacent or overlapping *cis*-regulatory sequences.

In this work, we investigated the impact that SNPs located in putative enhancer regions have in enhancer activity, using an *in vivo* approach. Ten sequences that overlap with type 2 diabetes-associated loci and with marks for enhancer activity were tested, one overlapping with an exon of *SLC30A8* (seq132wt). To test sequences for enhancer activity in the endocrine pancreas, we performed *in vivo* mosaic transgenesis assays in zebrafish embryos. We show that this strategy is sensitive, has low noise, and can be quantitative to address enhancer activity. Using this method, we observed that 6 out of 10 tested sequences are endocrine pancreatic enhancers, including the *SLC30A8* exon-containing sequence. In addition, two sequences were found to be pancreatic progenitor enhancers. We also found that the type 2 diabetes-associated SNP (henceforth referred to as risk allele) decreased the enhancer activity of two enhancers, while the risk allele of two sequences resulted in a gain of enhancer activity. Interestingly, the *SLC30A8* coding risk allele (seq132risk) showed an increase in enhancer activity, demonstrating that coding SNPs have the potential to modulate the target gene activity at both the transcriptional and the protein level. To better understand how SNPs can affect enhancer activity, we focused on the seq132 enhancer. We divided seq132wt into different fragments, observing that all are necessary for a robust pancreatic enhancer activity. We also show that common SNPs in seq132 modulate its pancreatic enhancer activity and that a single nucleotide mutation ablates completely the endocrine pancreatic enhancer activity of seq132. Additionally, we observed a chromatin interaction between seq132 and the promoter of *Slc30a8* gene in murine cells and noted that targeting transcriptional modulators to seq132 using CRISPR affects the transcription of *Slc30a8*, strongly suggesting that the seq132 enhancer belongs to the regulatory landscape of *Slc30a8*. Overall, in this work, we use an *in vivo* system to validate enhancers that overlap with type 2 diabetes-associated SNPs, showing several cases where nucleotide variations result in complex changes in enhancer activity, including a classical and poorly understood coding SNP.

RESEARCH DESIGN AND METHODS

Zebrafish Husbandry and Embryo Culture

Zebrafish (*Danio rerio*) were handled according to European animal welfare regulations and standard protocols. Embryos were cultured at 28°C in Petri dishes containing

E3 medium supplemented with 1-phenyl-2-thiourea to delay pigmentation formation (27).

Putative Enhancer Selection

Putative enhancer sequences were selected on the basis of GWAS data that uncovered 163 SNPs associated with type 2 diabetes or glycemic traits ($P < 5e-8$), considering all variants in high linkage disequilibrium (1000 Genomes Project Utah residents with ancestry from northern and western Europe $r^2 > 0.8$), with lead GWAS SNPs being the exception for rs735949 ($P < 3.70e-6$) (6,28–30). Risk alleles associated with seq132, seq117, and seq790 are part of the previously described GWAS credible set SNPs (31). Sequences were analyzed using the Islet Regulome Browser (32), which identifies active enhancers by the presence of H3K4me1, H3K27ac, and H2A.Z epigenetic marks in adult human endocrine pancreatic samples. The analysis was further refined by the presence of PDX1, MAFB, NKX6.1, FOXA2, and NKX2.2 binding obtained by chromatin immunoprecipitation sequencing (ChIP-seq) profiles (6,28), resulting in a list of 10 putative enhancer sequences, each overlapping with one SNP associated with type 2 diabetes (Supplementary Table 1).

In Vivo Mosaic Transgenesis Assays

Zebrafish transgenesis was performed using the *Tol2* transposon system (33). One-cell embryos from the Tg(*sst:mCherry*) zebrafish reporter line were microinjected with 3 nL containing 25 ng/ μ L *Tol2* transposase mRNA and 25 ng/ μ L phenol/chloroform-purified reporter vector. Injections were performed at least two times. *Ins:GFP* reporter was built by isolating the insulin promoter from the *ins*-CFP-NTR vector (SacI and BamHI) (34), cloning it in a pEM-MCS vector (35) and recombining it to a *Tol2*-based transposon containing a Gateway cassette and a GFP reporter gene.

Human sequences were PCR amplified from human genomic DNA using specific primers (Supplementary Table 2), cloned in TOPO vector (pCR8/GW/TOPO TA Cloning KIT; Invitrogen) and recombined *in vitro* to the Z48 transgenesis vector (36) by the Gateway system. Risk SNPs from seq58, seq68, seq73, seq219, and seq460 were inserted by site-directed mutagenesis using specific primers containing the risk variant (Supplementary Table 2). Injected embryos showing expression of GFP in the midbrain were selected for immunohistochemistry at 48 hours post fertilization (hpf).

sst:mCherry Reporter Line

To identify the zebrafish endocrine pancreatic domain, we developed an *in vivo* reporter line that drives expression of mCherry in δ -cells. Primers for the somatostatin (*sst*) promoter amplification were designed (Supplementary Table 2) and the amplified fragment cloned in a *Tol2* transposon containing mCherry as reporter gene. The vector was microinjected in one-cell embryos using the *Tol2* transposon system. Embryos were selected for *sst:mCherry*-positive cells and raised until adulthood and a stable transgenic line was isolated.

Immunohistochemistry

The 48-hpf microinjected embryos were dechorionated and fixed in 4% formaldehyde in PBS (PBS1×) overnight at 4°C and then washed in PBS-T (0.5% Triton X-100 in PBS1×) with 1% Triton X-100 in PBS1× (2 h) and 5% BSA in PBS-T (0.1%). Embryos were incubated with anti-Nkx6.1 (1:75) (F55A12; Developmental Studies Hybridoma Bank) and anti-insulin (1:50) (ab210560; Abcam) diluted in 5% BSA-PBS-T followed by washing. Embryos were then incubated with anti-mouse Alexa Fluor 647 (1:800), anti-rabbit Alexa Fluor 647 (1:800) (Thermo Fisher Scientific), and DAPI (1:1,000) (Invitrogen) diluted in 5% BSA-PBS-T. Embryos were washed and stored in 50% glycerol in PBS1×. Microscopy slides were prepared using 50% glycerol in PBS1×. Confocal imaging was performed using a Leica SP5II confocal microscope.

Assessment of Enhancer Activity

Embryos were analyzed, using confocal microscopy, for the presence of GFP-positive cells in the endocrine pancreatic domain (*sst:mCherry* reporter domain) or in the endocrine progenitor domain (anti-Nkx6.1). One embryo was considered positive if at least one GFP-positive cell was detected within the endocrine pancreatic domain or progenitor domain. Quantifications are presented as percentages of positive embryos to ensure the quantification of different transposon integrations.

Prediction of TFBSs Affected by Type 2 Diabetes Risk Variants

The wild-type (wt) and risk variant sequences were analyzed using 719 specific position weight matrices for vertebrate TFs using JASPAR software (37). TFs were ranked by position-specific score matrix. The relative score is a threshold score between 0 and 1 and is calculated by $(\text{score} - \text{minimum score}) / (\text{maximum score} - \text{minimum score})$, meaning 1 is the highest affinity and 0 is no affinity of binding (37). TFs that showed differential binding affinity in wt and risk variants were selected and filtered by presence of H3K4me3 at their promoters (6,32).

ChIP on Plasmid

Seq119wt and risk were cloned into a pLVX lentiviral backbone (#125839; Addgene) as KpnI-ApaI (Anza) fragments. Lentiviral particles were produced in HEK-293 cells (packaging plasmids psPAX2, #12260, and pCMV-VSV-G, #8454; Addgene) and used to infect MIN6 cells (a gift from Lorenzo Pasquali) according to standard procedures. Infected cells were selected with puromycin (1 μg/mL; Sigma-Aldrich) for 12 days, starting 48 h after infection. Three to 10 million cells were used for ChIP (12) with 4 μg of Nkx6.1 antibody (F5510; Developmental Studies Hybridoma Bank) and magnetic Dynabeads (Thermo Fisher Scientific). Eluted chromatin was purified with a MiniElute Kit (QIAGEN). Immunoprecipitated DNA was dissolved in water and further analyzed by real-time PCR (iQaq Universal SYBR Green Supermix, CFX 384; Bio-Rad).

Real-time Quantitative Expression Analysis

MIN6 cells were harvested for RNA extraction with TRIzol (Ambion) and treated with DNase (Thermo Fisher Scientific). Five hundred nanograms to 1 μg of DNA-free RNA was retrotranscribed with iScript cDNA Synthesis Kit (Bio-Rad). cDNA was used for quantitative PCR (iQaq Universal SYBR Green Supermix, CFX 384). *Slc30a8* expression was calculated by the ΔCt method to *actb* housekeeping mRNA.

4C Sequencing

The 4C sequencing (4C-seq) was performed on 10 million MIN6 cells using sequential *DpnII* and *Csp6I* as previously described (38), with minor modifications. The 4C template was purified using an Amicon Ultra-15 Centrifugal Filter Unit (Millipore). Two libraries were independently prepared with the Expand Long Template PCR System (Roche) using specific primers (Supplementary Table 2). Libraries were purified with QIAquick PCR Purification Kit (QIAGEN) followed by the Agencourt AMPure XP reagent (Beckman Coulter). Libraries were sequenced on an Ion S5 XL System (Ion 540 Chip, Ion Torrent; Thermo Fisher Scientific). Previously described processing (39,40) was used with a custom Perl script. More than 3.5 million reads were aligned to the mouse genome (mm10) using Bowtie2 (default parameters, global alignment mode) (41). Reads within fragments flanked by restriction sites of the same enzyme (checked with bedtools) or fragments <40 base pairs (bp) were filtered out. Mapped reads were then converted to reads-per-first-enzyme-fragment-end units and smoothed using a 30-fragment mean running window algorithm.

CRISPR Inactivation and CRISPR Activation Targeting

Twenty-nucleotide single guide RNAs (sgRNAs) targeting the murine *Slc30a8* enhancer with high predicted cleavage were selected from the UCSC Genome Browser CRISPR target track and cloned into the lentiviral backbone for enhancer inactivation (CRISPRi) [in Lenti-(BB)-hPGK-KRAB-dCas9-2A-BlastR, #118155; Addgene] or activation (CRISPRa) (in lentiSAMv2, #75112; Addgene) as previously described (42). Lentiviral particles were produced in HEK-293 cells (packaging plasmids: pRSV-rev, #12253; pMDLg/pRRE, #12251; and pMD2G, #12259; Addgene) and used to infect MIN6 cells according to standard procedures. Infected cells were selected by blasticidin (8 μg/mL) (Sigma-Aldrich) for 12 days, starting 48 h after infection.

Statistical Analysis

Statistical analyses were performed by using the χ^2 test with Fisher correction and unpaired *t* test, applying a significance level of $P \leq 0.05$. For real-time expression experiments, statistical analysis was performed with the Mann-Whitney test.

Data and Resource Availability

All data generated or analyzed during this study are included in the published article (and its Supplementary Material), with the exception of 4C-seq sequencing data. The 4C-seq data sets have been deposited in the European Nucleotide Archive at EMBL-EBI under accession number

PRJEB39688 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB39688>). The *sst:mCherry* reporter line generated during the current study is available from the corresponding author upon reasonable request.

RESULTS

Identification of Endocrine Pancreatic Enhancers In Vivo by Mosaic Transgenesis in Zebrafish

In vivo enhancer reporter assays can be performed in zebrafish either by generating stable transgenic lines, a time-consuming approach, or by mosaic transgenesis (43) on the basis of the analysis of many independent integration events. To test sequences for endocrine pancreatic enhancer activity, we used a *Tol2* transposon (44) containing a minimal promoter, a GFP reporter gene, and a midbrain-specific enhancer (Z48) acting as an internal control of transgenesis (36) (Fig. 1A). As an endocrine marker, we developed an in vivo reporter line that drives expression of mCherry in δ -cells (*sst:mCherry*). To validate *sst:mCherry* as an endocrine pancreatic reporter, we generated double-positive embryos for the *sst:mCherry* and insulin (*ins:GFP*) reporter transgenes (Fig. 1B). At 48 hpf, all GFP-positive cells (*ins:GFP*) were located within the *sst:mCherry* expression pattern (Fig. 1C), indicating that the *sst:mCherry* reporter line can be used to define the endocrine pancreatic domain. Next, to understand whether the mosaic strategy to identify endocrine pancreatic regulatory elements was sensitive enough, we mobilized a *Tol2* transposon containing the insulin promoter upstream of GFP. The mosaic analysis of 48-hpf-injected embryos, using confocal microscopy, revealed that 69% ($n = 23$) showed GFP expression in the endocrine pancreatic domain (Fig. 1D and Supplementary Fig. 1). Random integrations of the Z48 transgenesis vector can generate noise as a result of the influence of regulatory elements located in the genomic landscapes of each integration, termed position effect (43). To test whether the Z48 transgenesis vector was prone to position effect, we mobilized this vector without a sequence to test (negative control [NC]). These injected embryos did not show expression of GFP in the endocrine pancreas (0%, $n = 43$) at 48 hpf. In total, these results show that the use of mosaic transgenic embryos is sensitive enough to identify endocrine pancreatic regulatory sequences and that the associated noise as a result of the position effect is very low (Fig. 1D and Supplementary Fig. 1). Because endocrine pancreatic enhancers might also be active in pancreatic progenitor cells, we asked whether mosaic transgenesis assays could also be applied for these cell types. To test this hypothesis, we used an anti-Nkx6.1 antibody to define the endocrine pancreatic progenitor domain (Fig. 1E), and we mobilized the Z48 vector containing a known progenitor enhancer from the human *SOX9* locus (36). Twenty-seven percent ($n = 11$) of embryos showed expression of GFP within the pancreatic progenitor domain labeled by Nkx6.1, contrasting with the NC, for which GFP was not detected (0%, $n = 23$) (Fig. 1F and Supplementary Fig. 1).

Identification of Endocrine Pancreatic Enhancers Overlapping With Type 2 Diabetes-Associated SNPs

We selected 10 sequences (Supplementary Table 1) that overlap with SNPs previously associated with type 2 diabetes that are enriched for enhancer marks (H3K4me1, H3K27ac, and H2A.Z) and for TFBSs of endocrine pancreas TFs (FOXA2, NKX2.2, NKX6.1, MAFB, and PDX1) (6) (Fig. 2A and Supplementary Fig. 2A–H). Nine of these sequences are noncoding while one, seq132, partially overlaps with a coding exon of the *SLC30A8* gene. The respective sequences that do not contain the type 2 diabetes-associated variant (wt alleles) were cloned in Z48 transgenesis vector, and in vivo enhancer assays were performed by mosaic transgenesis in zebrafish embryos (36). Out of the 10 tested sequences, 6 showed a consistent expression of GFP in the endocrine pancreatic domain, therefore being endocrine pancreatic enhancers (seq58wt, seq68wt, seq73wt, seq132wt, seq219wt, and seq460wt) (Fig. 2B and D and Supplementary Figs. 3–5). Stable transgenic lines were generated for three of these sequences (seq132wt, seq460wt, and seq58wt) to confirm their endocrine pancreatic enhancer activity (Supplementary Fig. 6). Interestingly, for at least three of the tested sequences (seq58, seq68, and seq73), the GFP signal was also detected adjacent to the endocrine pancreatic domain (Supplementary Fig. 7), suggesting that these sequences may be pancreatic progenitor enhancers. To address this hypothesis, we labeled embryos injected with reporters of seq58, seq68, and seq73 with anti-Nkx6.1 antibody, showing that seq68 and seq73 drive GFP expression in the Nkx6.1-positive progenitor domain (45% [$n = 13$] and 25% [$n = 12$], respectively) (Fig. 2C and E), being therefore identified as pancreatic progenitor enhancers. To further characterize the identified enhancers, we determined in which endocrine pancreatic cell types they drive expression. For that, we injected the Z48 transgenesis vector containing the respective enhancers in a *gcga:mCherry* reporter line (α -cells), counterstaining these embryos with anti-insulin to label β -cells. We found that the majority of tested enhancers were able to drive expression in β -cells and that most of them were able to drive expression in more than one cell type (Supplementary Fig. 8).

SNPs Associated With Increased Risk of Type 2 Diabetes Modulate Enhancer Activity

To address the possible impact that type 2 diabetes-associated SNPs have in overlapping enhancers, we tested the corresponding variants (risk alleles), performing enhancer assays for endocrine pancreas. Of the six previously identified endocrine pancreatic enhancers, two (seq58risk and seq219risk) showed a decreased enhancer activity for the respective risk allele and two an increase (seq68 and seq132) compared with the wt allele (Fig. 3A and B and Supplementary Figs. 9 and 10). Strikingly, for seq132, the risk allele is in a coding exon of *SLC30A8* (seq132risk 56%, $n = 36$; seq132wt 23%, $n = 34$) (Fig. 3B). For one sequence, the risk allele was able to drive GFP expression in the endocrine pancreas above the established threshold,

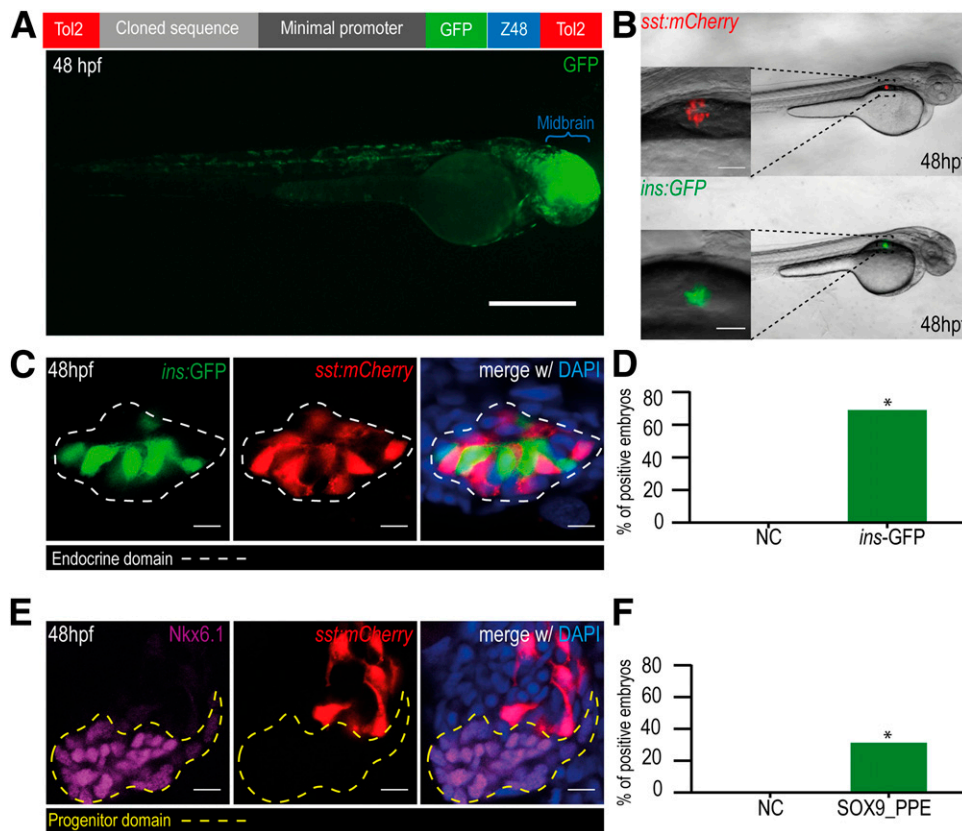


Figure 1—A: Schematic representation of the Z48 vector (top) and a representative image of a Z48-injected embryo (bottom) showing GFP expression in the midbrain, mediated by the Z48 enhancer (blue), at 48 hpf. The expression of GFP in the midbrain functions as an internal control of transgenesis. Scale bar = 200 μ m. B: Representative images of *sst:mCherry* (top) and *ins:GFP* (bottom) reporter lines at 48 hpf, showing mCherry and GFP expression in δ - and β -cells, respectively. Scale bars = 50 μ m. C: Confocal images showing the endocrine pancreatic domain (dashed line), defined by the cross of the *sst:mCherry* and *ins:GFP* reporter lines. The 48-hpf embryos were counterstained with the nuclear marker DAPI. Scale bars = 10 μ m. D: Percentage of embryos showing GFP-positive cells within the endocrine domain when injected with a vector containing GFP as reporter gene under the control of the insulin promoter (*ins:GFP*) (69%, $n = 23$) or with the Z48 vector without an endocrine enhancer (NC). E: Confocal images from 48-hpf embryos stained with anti-Nkx6.1 antibody (purple) to define the progenitor pancreatic domain (dashed line) in the *sst:mCherry* reporter line and the nuclear marker DAPI. Scale bars = 10 μ m. F: Graph representing the percentage of embryos with GFP expression in progenitor cells when injected with a pancreatic progenitor enhancer (SOX9_PPE) (27%, $n = 11$) or the NC (0%, $n = 43$). * $P < 0.05$, by χ^2 test.

while the wt allele did not (seq119risk 14%, $n = 28$; seq119wt 4%, $n = 27$) (Fig. 3B). Overall, these results demonstrate that type 2 diabetes-associated SNPs have the potential to modulate the activity of enhancers in a sequence-specific manner.

Differential binding of TFs to wt and risk alleles could explain the observed differential enhancer activity. To test this hypothesis, we generated a stable mouse MIN6 β -cell line containing human wt and risk sequences of the seq119 enhancer. Seq119 risk showed both an increased predicted affinity to Nkx6.1 binding (Supplementary Fig. 11A) and increased enhancer activity (Fig. 3B). Performing ChIP-PCR, we demonstrated that Nkx6.1 binds with higher affinity to the seq119 risk variant (Supplementary Fig. 11B). Additionally, we predicted bioinformatically TFBSs for wt and risk alleles of each sequence (Supplementary Table 3). Sequences were then clustered in two groups: sequences that had shown differential enhancer activity between wt and risk alleles (seq58, seq68, seq119, seq132,

and seq219) and sequences that did not (seq72, seq73, seq117, seq460, and seq790). Although both groups showed a similar number of predicted TFBSs in the wt allele, the differential activity enhancers group showed a higher number of predicted differential binding between wt and risk alleles (Supplementary Fig. 11C). These results suggest that differential binding of TFs might control the regulatory output of wt and risk variants.

The Enhancer Seq132mm Belongs to the *Slc30a8* Regulatory Landscape

Among the detected enhancers, we found that seq132, a sequence that partially overlaps with an exon of *SLC30A8*, is an endocrine pancreatic enhancer. Additionally, we showed that the type 2 diabetes-associated risk allele (seq132risk), which encodes a tryptophan-to-arginine substitution causing a decrease in the function of *SLC30A8* (24), has increased enhancer activity compared with the wt allele (seq132wt). To determine whether seq132 belongs to

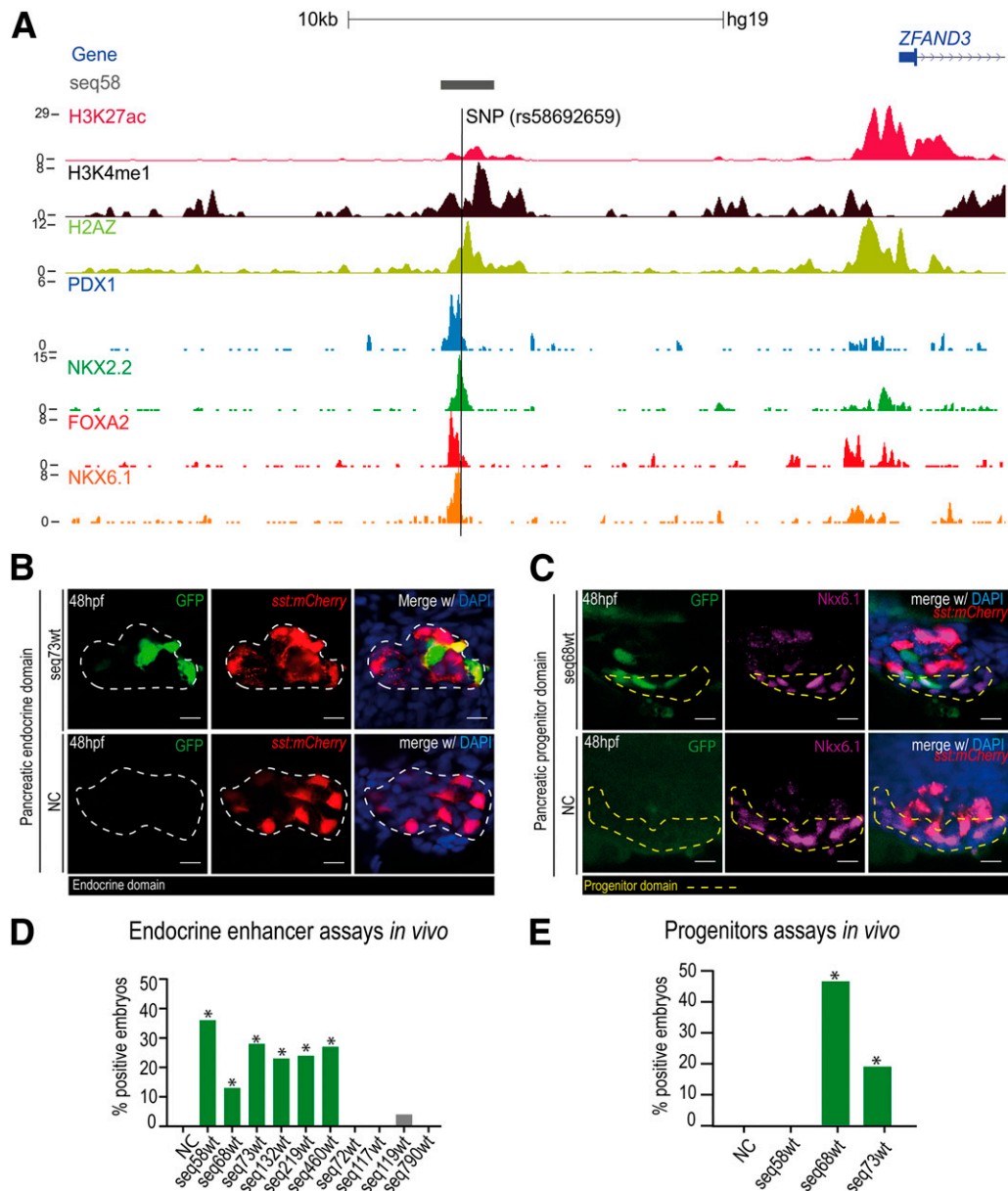


Figure 2—A: Genomic landscapes of the putative enhancer seq58. Tracks represent H3K27ac, H3K4me1, histone variant H2A.Z, and TF binding (PDX1, NKX2.2, FOXA2, and NKX6.1) from ChIP-seq data of human endocrine pancreatic samples. Human *ZFAND3* is the nearest gene to the putative enhancer (blue). The location of the type 2 diabetes-associated SNP (rs58692659) is represented as a vertical black line. B: In vivo reporter assay for endocrine pancreatic enhancers. Top panels show a representative confocal image of a *sst:mCherry* zebrafish embryo injected with the Z48 enhancer reporter vector containing the seq73wt sequence, showing GFP-positive cells within the endocrine pancreatic domain (dashed line), contrasting that absent in embryos injected with NC (bottom). Scale bars = 10 μ m. C: In vivo reporter assay for pancreatic progenitor enhancers. Confocal analysis of seq68 reporter assay shows colocalization of GFP-positive cells with *Nkx6.1* progenitor marker, contrasting with NC for the pancreatic progenitor domain (dashed line) that did not show GFP-positive cells. All vectors were injected in the *sst:mCherry* reporter line and embryos analyzed at 48 hpf and stained with DAPI. Scale bars = 10 μ m. D: Graph representing the percentage of embryos with GFP expression in endocrine pancreatic domain at 48 hpf for each sequence analyzed: seq58wt (36%, $n = 56$), seq68wt (13%, $n = 47$), seq73wt (28%, $n = 47$), seq132wt (23%, $n = 34$), seq219wt (24%, $n = 38$), seq460wt (27%, $n = 36$), seq72wt (0%, $n = 27$), seq117wt (0%, $n = 21$), seq119wt (4%, $n = 27$), seq190wt (0%, $n = 20$), and NC (0%, $n = 43$). E: Graph representing the number of embryos with GFP expression in pancreatic progenitor domain, defined by *Nkx6.1* staining at 48 hpf, for each sequence analyzed: seq58wt (0%, $n = 12$), seq68wt (46%, $n = 13$), seq73wt (25%, $n = 12$), and NC (0%, $n = 43$). * $P < 0.05$, by χ^2 test.

the regulatory landscape of *SLC30A8*, we used the MIN6 cell line to detect chromatin interaction points, since enhancers contact the promoter of the genes that they control. First, we performed enhancer assays in zebrafish, demonstrating that the mouse orthologous sequence

(seq132mm) of the seq132 human enhancer is also an endocrine pancreas enhancer (Fig. 4A). Then, using 4C-seq (38), we observed the existence of an interaction between the *Slc30a8* promoter and seq132mm (Fig. 4B and Supplementary Fig. 12). To further validate that seq132mm

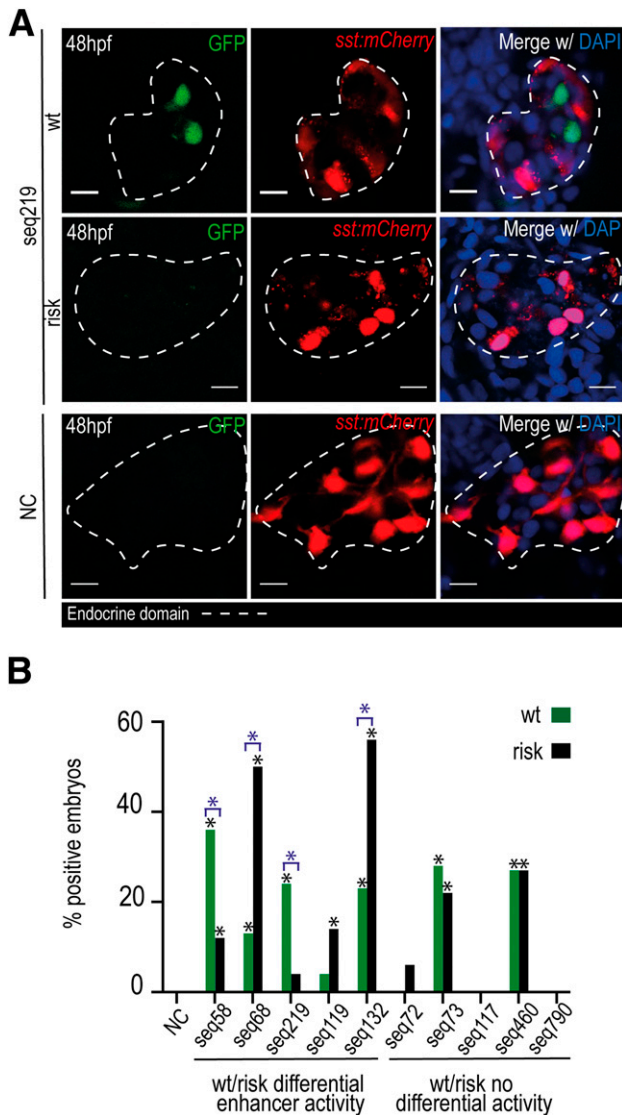


Figure 3—A: Representative confocal images for seq219wt, seq219risk, and NC. Seq219wt showed GFP expression in endocrine pancreatic domain (dashed line), defined by the *sst:mCherry* reporter line. The 48-hpf embryos were stained with DAPI. Scale bars = 10 μ m. B: Graph showing the total percentage of positive embryos in wt and risk alleles for each of the 10 sequences analyzed: seq58wt/risk (36%, $n = 56$; 12%, $n = 43$, respectively), seq68wt/risk (13%, $n = 47$; 50%, $n = 32$), seq73wt/risk (28%, $n = 47$; 22%, $n = 36$), seq132wt/risk (23%, $n = 34$; 56%, $n = 36$), seq219wt/risk (24%, $n = 38$; 6%, $n = 35$), seq119wt/risk (4%, $n = 27$; 14%, $n = 28$), seq72wt/risk (0%, $n = 27$; 6%, $n = 30$), seq117wt/risk (0%, $n = 21$; 0%, $n = 8$), seq460wt/risk (27%, $n = 36$; 27%, $n = 30$), and seq790wt/risk (0%, $n = 20$; 0%, $n = 14$). Six sequences showed differential enhancer activity between wt and risk allele. * $P < 0.05$, by χ^2 test.

belongs to the regulatory landscape of *Slc30a8*, we targeted seq132mm using the CRISPR/Cas9 system with a dCas9 fused to a transcriptional activation domain (CRISPRa) and another to a repressor domain (CRISPRi), observing a significant increase and decrease of *Slc30a8* expression levels, respectively (Fig. 4C). These results strongly suggest that seq132mm belongs to the regulatory landscape of

Slc30a8, and because of the remarkable conservation in the activity and sequence of this enhancer, we propose that this regulatory mechanism is conserved in humans.

The *SLC30A8* Seq132 Enhancer Is Divided Into Different Functional Domains

Next, we wanted to understand whether seq132 is divided into different functional domains. For that, we divided seq132wt into four fragments (Fig. 5A) and performed enhancer assays for each (Fig. 5B). Fragments seq132wt1 (872 bp) and seq132wt2 (967 bp) showed a milder endocrine pancreatic enhancer activity than the seq132wt total fragment (seq132wt 23%, $n = 34$; seq132wt1 9.8%, $n = 41$; seq132wt2 6.5%, $n = 31$). We also tested another fragment, seq132wt3 (899 bp), that contains seq132wt1 and extends to the end of the coding sequence of *SLC30A8*. Seq132wt3 was able to drive GFP expression in endocrine pancreatic cells (10%, $n = 30$), as was the remaining fragment seq132wt4 (788 bp), although with a decreased efficiency (4%, $n = 23$) (Fig. 5C). From these results, we conclude that seq132 has several functional domains spread through this sequence, and the sum of these parts is necessary for this enhancer to be fully functional. These results also suggest that other SNPs could potentially affect the output of this enhancer. To test this, we have performed enhancer assays with seq132risk, which contains three other common SNPs with no known association to type 2 diabetes (seq132risk#: rs2466296, rs2466295, and rs2466294) (Fig. 6A). Interestingly, seq132risk# was a less active enhancer than seq132risk, showing an activity similar to seq132wt (Fig. 6B and C) and demonstrating that the impact of disease risk alleles in the activity of enhancers might be modulated by other adjacent polymorphisms.

A Single Nucleotide Mutation Impairs Seq132 Enhancer

Focusing on seq132, we wanted to further determine whether a single nucleotide mutation could lead to the complete ablation of the activity of this enhancer. Previous results have shown that Pdx1, an important TF required for proper pancreatic function, controls the activity of one endocrine pancreatic enhancer located in the second intron of the mouse *Slc30a8* gene (45). On the basis of this, we hypothesized that seq132 could also be controlled by PDX1 binding. After performing TFBS analysis (Supplementary Table 3), we found that within seq132, there is a high score-predicted binding site for PDX1 (JASPAR score 0.9654) (Fig. 7A). To test whether this binding site is required for enhancer activity, we performed transgenesis assays using seq132wt containing a mutation in the predicted binding site of PDX1 (seq132wtPDX1) (Fig. 7B and C). This is an adenine-to-guanine substitution in the PDX1 consensus binding site, resulting in a predicted ablation of the binding of PDX1 (PDX1 predicted binding score: seq132wt 0.9654; seq132wtPDX1 0). Comparing endocrine pancreatic enhancer activity between seq132wt and seq132wtPDX1, we found that the seq132wtPDX1 sequence is unable to drive GFP expression in the endocrine pancreas (0%, $n = 20$) (Fig. 7D). Because we previously

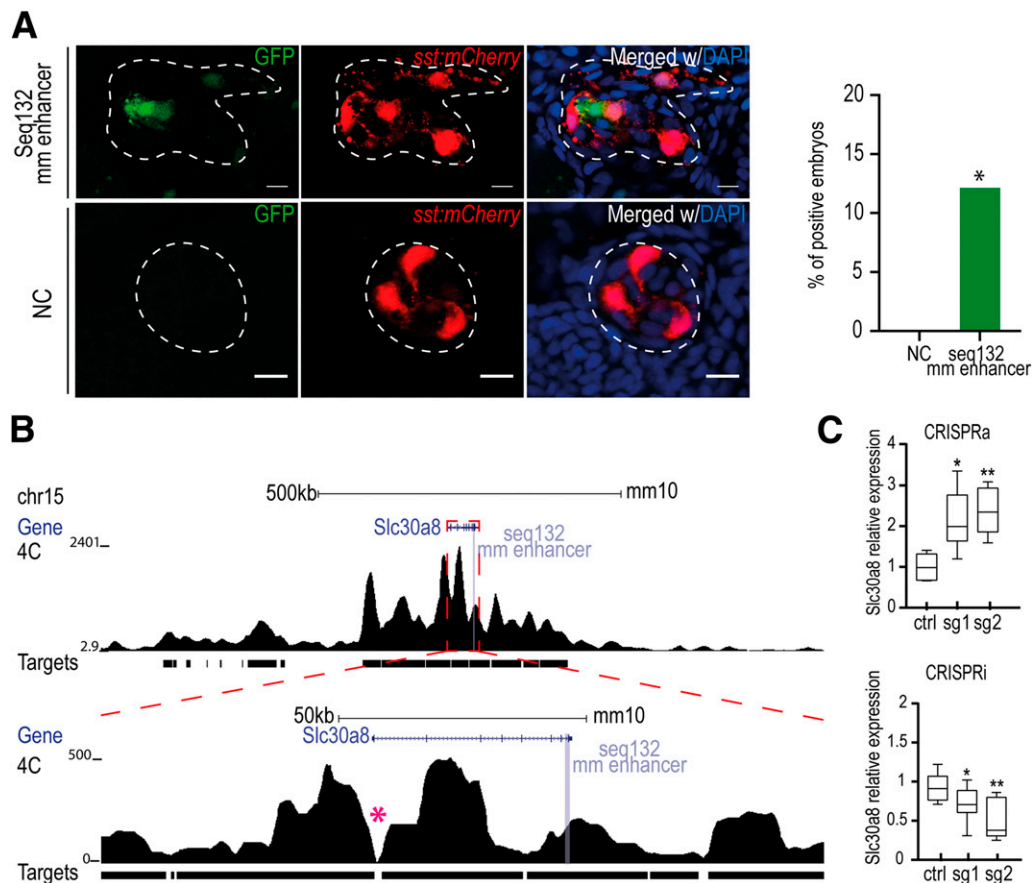


Figure 4—A: Representative confocal images for the mouse seq132 (seq132mm enhancer chr15:52334298 + 52335281). Seq132mm enhancer showed GFP-positive cells within the endocrine pancreatic domain (dashed line) defined by the *sst:mCherry* reporter line. The 48-hpf embryos were stained with DAPI. Scale bars = 10 μ m. Representative graph showing the total percentage of positive embryos for the sequence analyzed (13%, $n = 35$). * $P < 0.05$, by χ^2 test. B: Genomic landscape of the mouse *Slc30a8* gene (blue), showing 4C-seq profiles (black), with view point in *Slc30a8* promoter (pink asterisk) in the Min6 cell line; zoom-out (top) and zoom-in (bottom) of the *Slc30a8* gene and seq132mm enhancer. The targets line represents the regions where the interaction is significant. C: sgRNAs targeting murine *Slc30a8* enhancer in CRISPRa or CRISPRi assay in Min6 cells. *Slc30a8* expression was calculated relative to the β -actin housekeeping gene by quantitative PCR. Dot blots represent the 10–90% quantile of six biological replicates. * $P < 0.05$, ** $P < 0.01$. ctrl, control.

observed that seq132risk had increased endocrine pancreatic enhancer activity, we explored the possibility of the risk SNP rescuing the loss of function observed for the seq132wtPDX1 sequence. We observed that seq132riskPDX1 also showed no endocrine pancreatic enhancer activity (0%, $n = 23$); thus, the risk SNP is not sufficient to rescue the loss of the PDX1 binding site (Fig. 7D).

DISCUSSION

In this work, we demonstrate the feasibility to perform enhancer assays using mosaic transgenesis in zebrafish. The zebrafish pancreas, as its mammal counterpart, is composed mainly by α -, β -, and δ -cells that secrete the hormones glucagon, insulin, and somatostatin, respectively. The malfunction of these cell types can contribute to type 2 diabetes development. Therefore, enhancer assays evaluating type 2 diabetes-associated alleles should consider all these cell types. In the current work, we used an *sst:mCherry* reporter construct to determine the zebrafish endocrine pancreatic domain in vivo, making available the inherent

cellular complexity of a fully functional pancreas to define the activity of enhancers. In contrast, most of in vitro assays are limited to only one endocrine pancreatic cell type, in many cases not fully functional (46). In vivo assays have, however, their own limitations. This is particularly important when studying SNPs that might modulate the activity of enhancers and, therefore, transcriptional levels rather than binary activation or inactivation states. In this work, we demonstrate that a mosaic transgenesis method is sensitive enough and has low levels of noise, making it possible to assay quantitatively the activity of enhancers. In the current study, variation of enhancer activity comprises at least three parameters predicted to have an impact on the number of embryos with GFP expression in the endocrine pancreas: 1) transcriptional output, affecting the amount of GFP expression per cell and its detection potential defined by the GFP detection threshold; 2) expression domain, defined by the potential of GFP expression in different endocrine pancreatic cell types; and 3) robustness, defined by the stability of GFP expression. Although

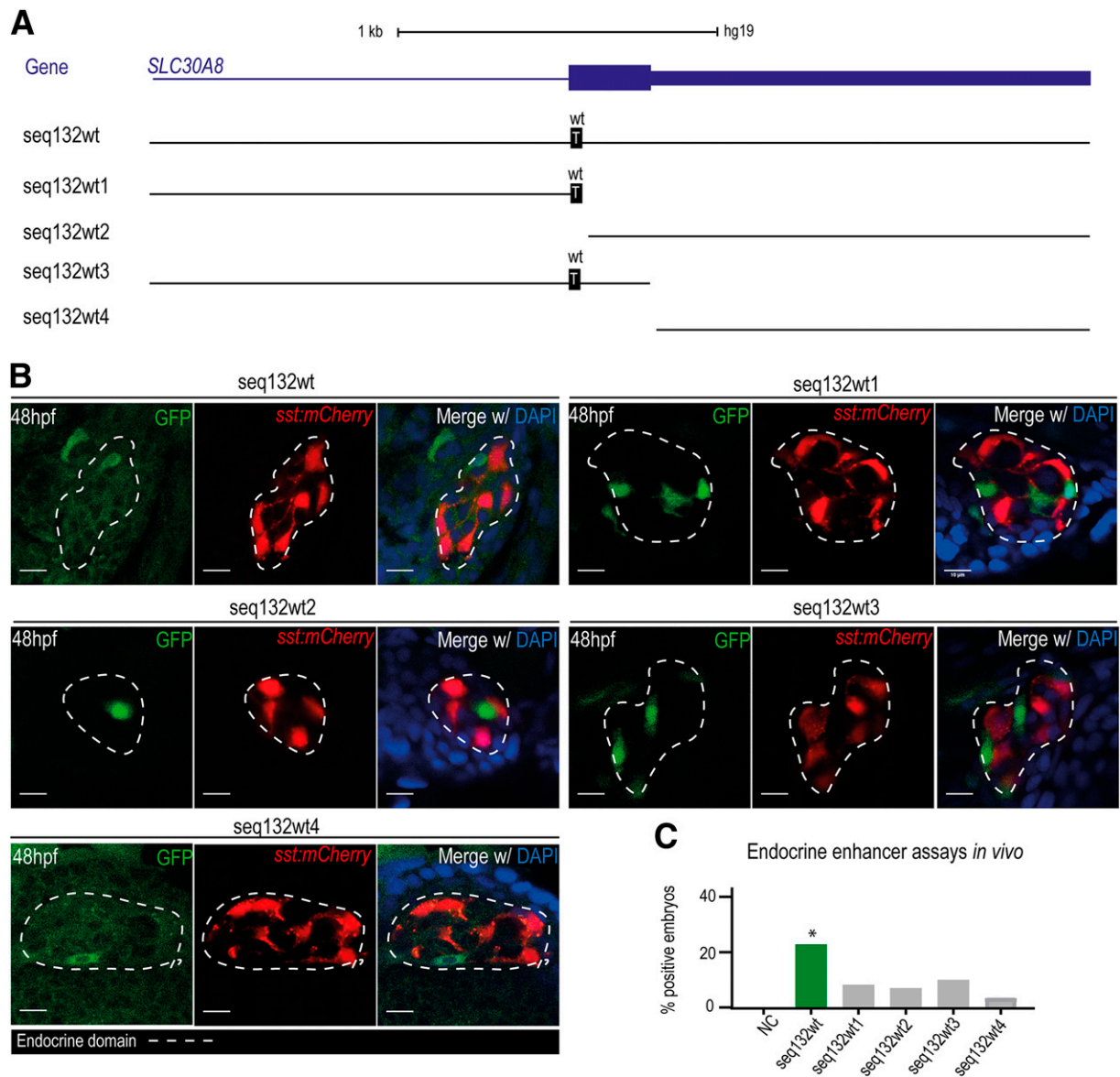


Figure 5—*A*: Schematic representation of the four analyzed fragments derived from seq132wt: seq132wt1, seq132wt2, seq132wt3, and seq132wt4. The gene *SLC30A8* is represented in blue. The wt allele is discriminated in black boxes. *B*: Representative confocal images for the total sequence, seq132wt, and the four analyzed fragments showing GFP-positive cells in the endocrine pancreatic domain (dashed line) defined by the *sst:mCherry* reporter line. The 48-hpf embryos were stained with DAPI. Scale bars = 10 μ m. *C*: Graph representing the total percentages of positive embryos for seq132wt (23%, $n = 34$), seq132wt1 (9.8%, $n = 41$), seq132wt2 (6.5%, $n = 31$), seq132wt3 (10%, $n = 30$), and seq132wt4 (4%, $n = 23$). * $P < 0.05$, by χ^2 test.

the present assay does not discriminate these three sources of differential enhancer activity, the introduction of further improvements can offer this possibility. Expression domain can be discriminated if specific endocrine pancreatic cell markers are used, as shown for wt alleles (Supplementary Fig. 7), and differences in transcriptional output can be observed if an internal control is used and GFP expression per cell is quantified.

Using mosaic transgenesis in zebrafish, we validated 6 out of 10 tested sequences as human endocrine pancreatic enhancers, two of which are also pancreatic progenitor enhancers, underlining the accuracy of enhancers prediction.

Although species-specific regulatory outputs cannot be completely excluded, genetic networks and TFs that operate in enhancers are usually highly conserved between distantly related vertebrates as zebrafish and human, making interspecies enhancers assays reliable.

Exploring the impact of the risk-associated SNPs in enhancer activity, we found that for seq58 and seq219, the risk allele dramatically decreased the enhancer activity while for seq68, seq119, and seq132 the opposite result was observed, suggesting that disease-associated SNPs have the potential to be translated into a loss or gain of function of target genes (47). For one case, seq119, we

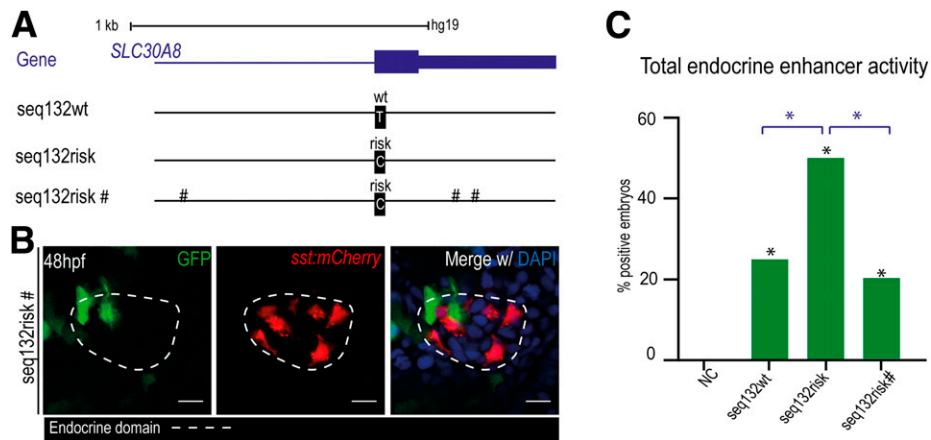


Figure 6—A: Schematic representation of the three different analyzed versions of seq132: seq132wt, seq132risk, and seq132risk#. The gene *SLC30A8* is represented in blue. The symbol # represents common variants with no association with type 2 diabetes. The wt and risk alleles are represented in black boxes. B: Representative confocal images for the seq132risk#, showing GFP expression in the endocrine pancreatic domain (dashed line) defined by the *sst:mCherry* reporter line. The 48-hpf embryos were stained with DAPI. Scale bars = 10 μ m. C: Graph representing the total percentages of positive embryos for the seq132wt (23%, $n = 34$), seq132risk (56%, $n = 36$), and seq132risk# (20%, $n = 20$). * $P < 0.05$, by χ^2 test.

further demonstrated the increased binding affinity of Nkx6.1 in the risk allele, suggesting that the enhancer activity outputs might be explained by the differential binding of TFs to wt and risk alleles, consistent with previous works (48,49).

Seq132 overlaps with exon 8 of *SLC30A8*. Strikingly, the risk allele rs13266634 (4), which results in an amino acid substitution impairing the function of *SLC30A8* (22), showed a significantly higher level of enhancer activity than the wt

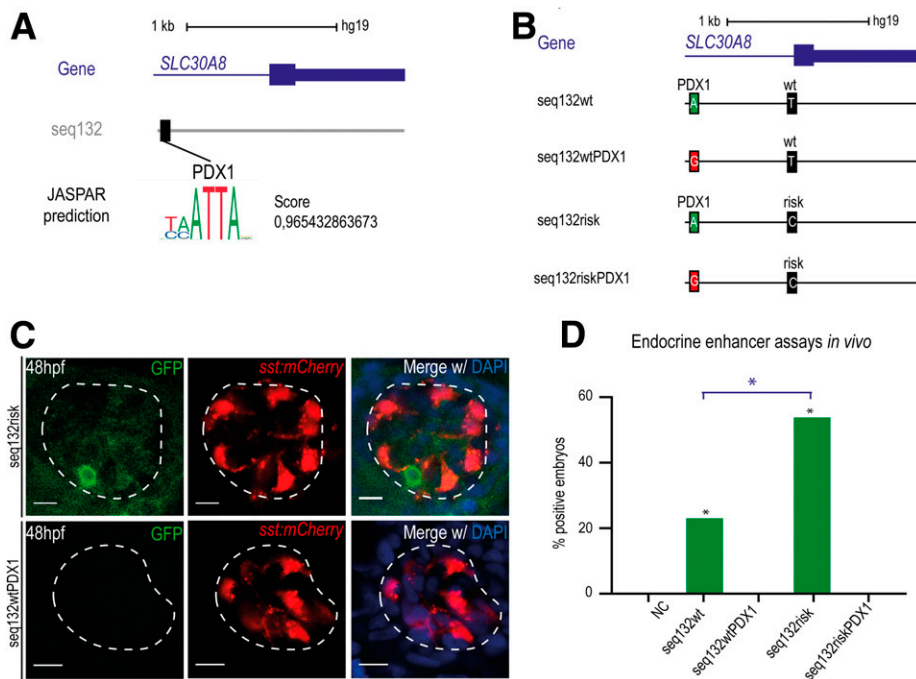


Figure 7—A: PDX1 binding site prediction by JASPAP software. The human *SLC30A8* is represented in blue. The vertical black line represents the single nucleotide mutation that overlaps with a PDX1 putative binding site. B: Schematic representation of the analyzed sequences: seq132wt, seq132wtPDX1, seq132risk, and seq132riskPDX1. The wt and risk alleles are discriminated in black boxes and the mutation in red boxes. The green boxes represent the binding site for PDX1. C: Representative confocal images for the analyzed sequence containing the risk variant in the absence (top) of the mutation in the putative binding site for PDX1, showing GFP expression in endocrine pancreatic domain (dashed line) defined by the *sst:mCherry* reporter line. The sequence containing the wt allele and the mutation (bottom) was not able to drive GFP expression. The 48-hpf embryos were stained with DAPI. Scale bars = 10 μ m. D: Graph representing the total percentages of positive embryos for the different sequences represented in panel B: seq132wt (23%, $n = 34$), seq132wtPDX1 (0%, $n = 21$), seq132risk (56%, $n = 36$), and seq132riskPDX1 (0%, $n = 20$). * $P < 0.05$, by χ^2 test.

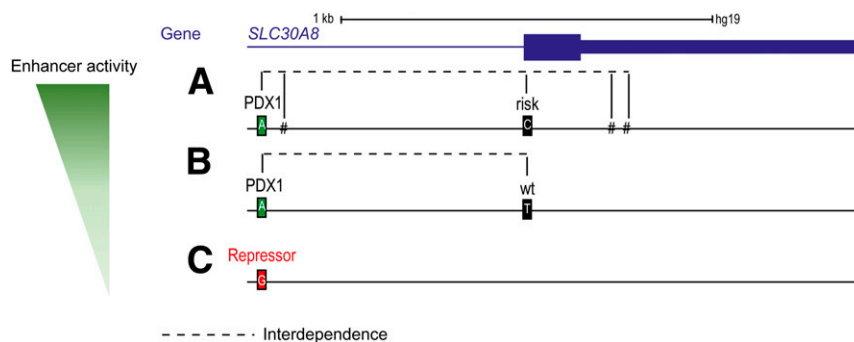


Figure 8—Theoretical model of the functional domains of seq132. Schematic representation from the three possible sequence versions and the respective associated enhancer activity level: high (A), low (B), and no enhancer activity (C). The gene *SLC30A8* is represented in blue. The symbol # represents common variants with no association with type 2 diabetes. The wt and risk alleles are represented in black boxes, the putative repressor in red, and the binding site for PDX1 in green. The dashed line shows interdependence.

sequence. These results demonstrate that coding SNPs have the potential to modulate the activity of overlapping enhancers, a mechanism poorly explored but compatible with described overlapping exon/enhancer functions (50). *SLC30A8* encodes a zinc transporter, an ion essential for insulin maturation and secretion in β -cells (22,23). Therefore, the rs13266634-associated increased risk for type 2 diabetes is commonly attributed to the decrease of the zinc transporter activity. In opposition to this hypothesis, the loss of function of *SLC30A8* was shown to enhance insulin secretion (26), and 12 identified truncating SNPs in the *SLC30A8* gene have been associated with a protective effect against type 2 diabetes (2,25). This incongruence could be explained if the rs13266634 association with type 2 diabetes results, not from the reduced *SLC30A8* zinc transporter activity, but from an increase of its transcription caused by a more active enhancer.

Further exploring the seq132 enhancer, we observed that the increase of its activity when containing the risk allele can be reverted by the presence of three other common SNPs, highlighting how combinatorial variations in single nucleotides can alter enhancer activity. Furthermore, a mutation that disrupts the predicted binding site of PDX1 in seq132 results in the complete ablation of the enhancer. This suggests that the loss of the binding of PDX1 might coincide with the gain of a transcriptional repressor, since this mutation ablates the activity of nonoverlapping seq132 subfragments that have mild autonomous enhancer activity (Fig. 8). Overall, in this work, we open new avenues on the understanding of the code embedded in noncoding regulatory sequences and show the complexity of effects that one single nucleotide variation might have in the activity of enhancers and its possible impact on human disease.

Acknowledgments. The authors thank Lorenzo Pasquali (Josep Carreras Leukaemia Research Institute, Barcelona, Spain) for helpful suggestions and critical reading of the manuscript. The authors acknowledge the contribution of Joana Teixeira (i3S-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, and IBMC-Instituto de Biologia Celular e Molecular, Porto, Portugal) for

the ChIP-seq data plotting, Sílvia Naranjo (CABD - Centro Andaluz de Biología del Desarrollo, Universidad Pablo de Olavide, Seville, Spain) for the *sst:mCherry* vector, Ana Maia (i3S-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, and IBMC-Instituto de Biologia Celular e Molecular, Porto, Portugal) for the *ins:GFP* construct, the support of i3S Scientific Platform Advanced Light Microscopy, members of the national infrastructure Portuguese Platform of Bio-Imaging (supported by POCI010145FEDER022122), and the assistance of the Genomics i3S Scientific Platform (supported by POCI-01-0145-FEDER-022184).

Funding. This study was supported by the H2020 European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. ERC-2015-StG-680156-ZPR), the Fundação para a Ciência e a Tecnologia (FCT) (IF/00654/2013), and the European Regional Development Fund (Norte-01-0145-FEDER-000029). A.E., M.D., and F.J.F. are PhD fellows from FCT (grants SFRH/BD/147762/2019 to A.E., SFRH/BD/135957/2018 to M.D., and PD/BD/105745/2014 to F.J.F.). J.B. acknowledges FCT for a scientific stimulus grant (CEECIND/03482/2018).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. A.E. carried out the experiments. A.E., M.D., and J.B. wrote the manuscript. A.E. and J.B. conceived, designed, and analyzed the data. C.P. performed the ChIP-seq experiment and CRISPRa and CRISPRi assays. F.J.F. performed the 4C-seq assay. M.D. contributed to the development of the *sst:mCherry* reporter line. M.G. performed the bioinformatic analysis of the 4C-seq experiment. J.B. designed and supervised the study. All authors revised the manuscript. J.B. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Chatterjee S, Khunti K, Davies MJ. Type 2 diabetes. *Lancet* 2017;389:2239–2251
- Flannick J, Mercader JM, Fuchsberger C, et al.; Broad Genomics Platform; DiscovEHR Collaboration; CHARGE; LuCamp; ProDiGY; GoT2D; ESP; SIGMA-T2D; T2D-GENES; AMP-T2D-GENES. Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* 2019;570:71–76
- Saxena R, Voight BF, Lyssenko V, et al.; Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007;316:1331–1336
- Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007;445:881–885
- Ong C-T, Corces VG. Enhancers: emerging roles in cell fate specification. *EMBO Rep* 2012;13:423–430
- Pasquali L, Gaulton KJ, Rodríguez-Seguí SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet* 2014;46:136–143

7. Stitzel ML, Sethupathy P, Pearson DS, et al.; NISC Comparative Sequencing Program. Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metab* 2010;12:443–455
8. Bhandare R, Schug J, Le Lay J, et al. Genome-wide analysis of histone modifications in human pancreatic islets. *Genome Res* 2010;20:428–433
9. Parker SCJ, Stitzel ML, Taylor DL, et al.; NISC Comparative Sequencing Program; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc Natl Acad Sci U S A* 2013;110:17921–17926
10. Gaulton KJ, Nammo T, Pasquali L, et al. A map of open chromatin in human pancreatic islets. *Nat Genet* 2010;42:255–259
11. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 2013;10:1213–1218
12. Rada-Iglesias A, Bajpai R, Swigut T, Bruggmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 2011;470:279–283
13. Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res* 2007;35:D88–D92
14. Bessa J, Tena JJ, de la Calle-Mustienes E, et al. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev Dyn* 2009;238:2409–2417
15. Prince VE, Anderson RM, Dalgin G. Zebrafish pancreas development and regeneration: fishing for diabetes therapies. *Curr Top Dev Biol* 2017;124:235–276
16. Youson JH, Al-Mahrouki AA, Amemiya Y, Graham LC, Montpetit CJ, Irwin DM. The fish endocrine pancreas: review, new data, and future research directions in ontogeny and phylogeny. *Gen Comp Endocrinol* 2006;148:105–115
17. Youson JH, Al-Mahrouki AA. Ontogenetic and phylogenetic development of the endocrine pancreas (islet organ) in fish. *Gen Comp Endocrinol* 1999;116:303–335
18. Kimmel RA, Dobler S, Schmitner N, Walsen T, Freudenblum J, Meyer D. Diabetic pdx1-mutant zebrafish show conserved responses to nutrient overload and anti-glycemic treatment. *Sci Rep* 2015;5:14241
19. Nelson SB, Schaffer AE, Sander M. The transcription factors Nkx6.1 and Nkx6.2 possess equivalent activities in promoting beta-cell fate specification in Pdx1+ pancreatic progenitor cells. *Development* 2007;134:2491–2500
20. Binot A-C, Manfroid I, Flasse L, et al. Nkx6.1 and nkx6.2 regulate alpha- and beta-cell formation in zebrafish by acting on pancreatic endocrine progenitor cells. *Dev Biol* 2010;340:397–407
21. Mahajan A, Wessel J, Willems SM, et al.; ExomeBP Consortium; MAGIC Consortium; GIANT Consortium. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet* 2018;50:559–571
22. Nicolson TJ, Bellomo EA, Wijesekara N, et al. Insulin storage and glucose homeostasis in mice null for the granule zinc transporter ZnT8 and studies of the type 2 diabetes-associated variants. *Diabetes* 2009;58:2070–2083
23. Kim I, Kang ES, Yim YS, et al. A low-risk ZnT-8 allele (W325) for post-transplantation diabetes mellitus is protective against cyclosporin A-induced impairment of insulin secretion. *Pharmacogenomics J* 2011;11:191–198
24. Rutter GA, Chimienti F. SLC30A8 mutations in type 2 diabetes. *Diabetologia* 2015;58:31–36
25. Flannick J, Thorleifsson G, Beer NL, et al.; Go-T2D Consortium; T2D-GENES Consortium. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet* 2014;46:357–363
26. Dwivedi OP, Lehtvirta M, Hastoy B, et al. Loss of ZnT8 function protects against diabetes by enhanced insulin secretion. *Nat Genet* 2019;51:1596–1606
27. Merriman C, Huang Q, Rutter GA, et al. Lipid-tuned zinc transport activity of human ZnT8 protein correlates with risk for type-2 diabetes. *J Biol Chem* 2016;291:26950–26957
28. Mohlke KL, Boehnke M. Recent advances in understanding the genetic architecture of type 2 diabetes. *Hum Mol Genet* 2015;24:R85–R92
29. Morris AP, Voight BF, Teslovich TM, et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of Anthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981–990
30. Scott RA, Lagou V, Welch RP, et al.; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* 2012;44:991–1005
31. Thurner M, van de Bunt M, Torres JM, et al. Integration of human pancreatic islet genomic data refines regulatory mechanisms at type 2 diabetes susceptibility loci. *eLife* 2018;7:e31977
32. Mularoni L, Ramos-Rodríguez M, Pasquali L. The pancreatic islet regulome browser. *Front Genet* 2017;8:13
33. Kawakami K, Shima A, Kawakami N. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proc Natl Acad Sci U S A* 2000;97:11403–11408
34. Curado S, Anderson RM, Jungblut B, Mumm J, Schroeter E, Stainier DYR. Conditional targeted cell ablation in zebrafish: a new tool for regeneration studies. *Dev Dyn* 2007;236:1025–1035
35. PME-MCS - Tol2kit. Accessed 27 August 2019. Available from <http://tol2kit.genetics.utah.edu/index.php/PME-MCS>
36. Cebola I, Rodríguez-Seguí SA, Cho CH-H, et al. TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat Cell Biol* 2015;17:615–626
37. Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* 2004;32:D91–D94
38. van de Werken HJG, de Vree PJP, Splinter E, et al. 4C technology: protocols and data analysis. *Methods Enzymol* 2012;513:89–112
39. Splinter E, de Wit E, van de Werken HJG, Klous P, de Laat W. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* 2012;58:221–230
40. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The dynamic architecture of Hox gene clusters. *Science* 2011;334:222–225
41. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359
42. Miguel-Escalada I, Bonàs-Guarch S, Cebola I, et al. Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat Genet* 2019;51:1137–1148
43. Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 1993;74:505–514
44. Kwan KM, Fujimoto E, Grabher C, et al. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* 2007;236:3088–3099
45. Pound LD, Hang Y, Sarkar SA, et al. The pancreatic islet β -cell-enriched transcription factor Pdx-1 regulates Slc30a8 gene transcription through an intronic enhancer. *Biochem J* 2011;433:95–105
46. Skelin M, Rupnik M, Cencic A. Pancreatic beta cell lines and their applications in diabetes mellitus research. *ALTEX* 2010;27:105–113
47. Vezzoli G, Terranegra A, Arcidiacono T, et al. R990G polymorphism of calcium-sensing receptor does produce a gain-of-function and predispose to primary hypercalciuria. *Kidney Int* 2007;71:1155–1162
48. Tjian R, Maniatis T. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 1994;77:5–8
49. Ong C-T, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* 2011;12:283–293
50. Birnbaum RY, Clowney EJ, Agamy O, et al. Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res* 2012;22:1059–1068