Cell Reports, Volume *35*

Supplemental information

Neonatal diabetes mutations disrupt a chromatin pioneering function that activates the human insulin gene

Ildem Akerman, Miguel Angel Maestro, Elisa De Franco, Vanessa Grau, Sarah Flanagan, Javier García-Hurtado, Gerhard Mittler, Philippe Ravassard, Lorenzo Piemonti, Sian Ellard, Andrew T. Hattersley, and Jorge Ferrer

Figure S1

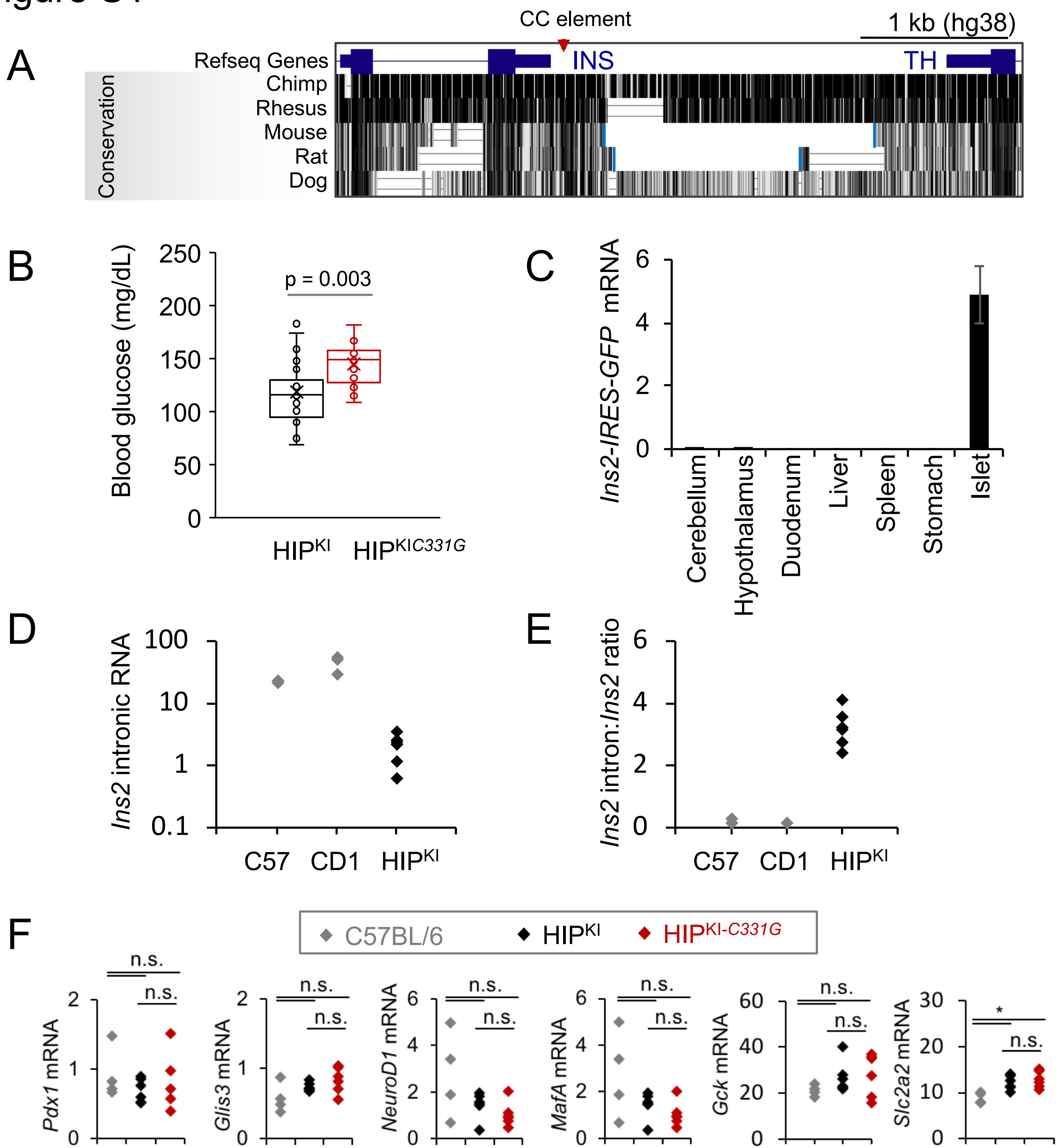


Figure S1. Related to Main Figure 1
(A) UCSC genome browser snapshot of the region between the INS gene and the TH gene. Multiz Alignment representing the level of conservation between the displayed species is shown.
(B) Glycemia from *ad libitum*-fed HIP^{KI} and $HIP^{KI\text{-}C331G}$ mice. P values were calculated with Student's t test.
(C) Quantification of *Ins2-IRES-GFP* mRNA in islets and a panel of tissues in HIP^{KI} mice. Bars represent average values normalised to Actb mRNA from n=6 mice. Cerebellum and hypothalamus display ~50,000 and ~850,000 fold less Ins2-IRES-GFP mRNA respectively, while no transcription is detectable in other tissues.
(D) Measurements of intronic Ins2 transcript levels. Ins2 intronic transcripts, representing the amount of nascent transcription, was quantified using quantitative PCR with taqman probes in islets isolated from C57BL/6 mice (n=4) or HIPKI (n=6).
(E) Nascent and spliced Ins2 RNA ratios in wild type and HIP^{KI} mice. Intronic Ins2 transcript levels and Ins2 transcript levels were quantified using taqman probes in islets isolated from C57BL/6 mice (n=4) or HIP^{KI} (n=6).
(F) Reverse transcription quantitative PCR for *Pdx1, Glis3, NeuroD1, and MafA* mRNAs from pancreatic islets from 3-5 month control C57BL/6 (n=4), HIP^{KI} (n=6) and $HIP^{KI\text{-}C331G}$ (n=6) mice. Values were normalized to *Actb* or *Hprt* mRNAs. (*) Asterisk indicate significance <0.05 by ANOVA unpaired t=test.

Figure S2

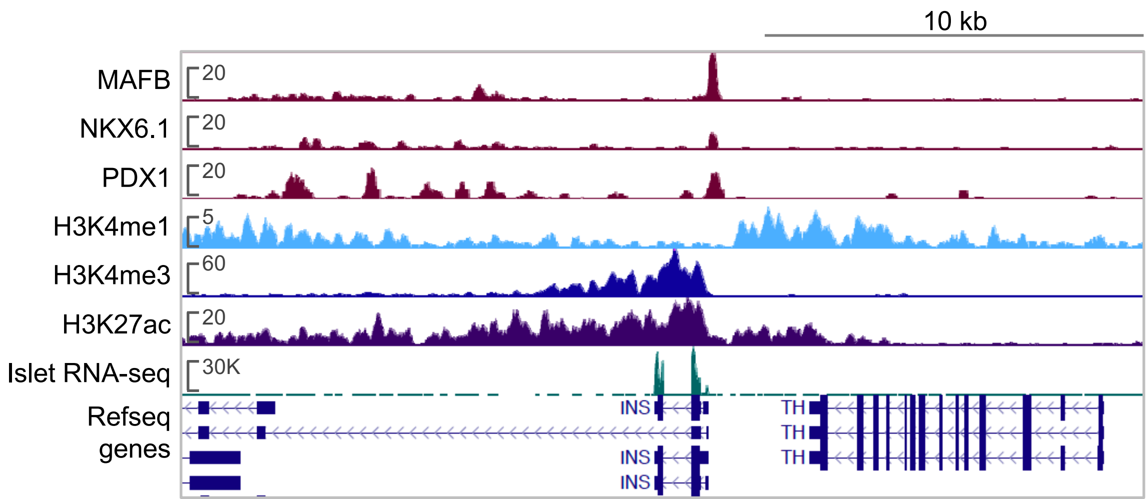


Figure S2. Related to Main Figure 4.
Chromatin landscape of the human *INS* locus in human pancreatic islets. ChIP-seq profiles of activating histone marks H3K4 trimethylation and H3K27 acetylation[12]. All scales represent RPKMs.

Figure S3

A

Transcription factor	Identification method	Activation of wild type *INS* in episomal assay	Activation of endogenous *INS*
Klf13	SILAC	-	-
Klf16	SILAC	-	-
Maz (PUR1)	SILAC	-	-
Zfp37	SILAC	-	-
SP1	in silico motif search	-	na
Glis2	Literature (PMID:21127075)	-	-
Glis3	Literature (PMID:23927931, 19264802)	+	+
Klf10	In silico motif search	-	-
Klf11	Literature (PMID:21592955, 21592955)	-	-

B

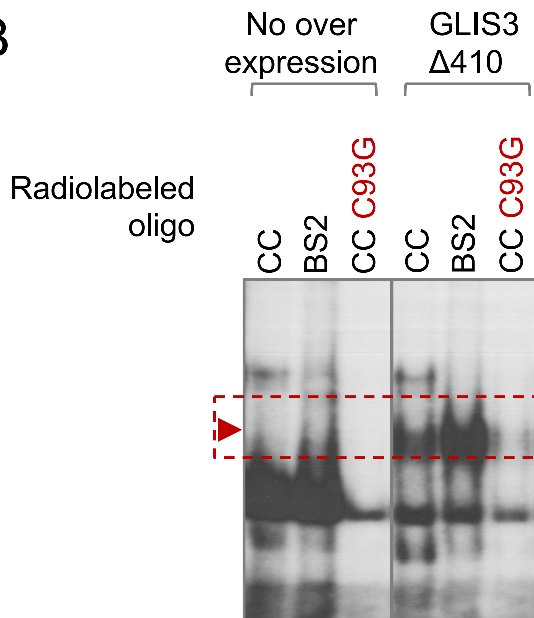


Figure S3. Related to Main Figure 6
(A) Candidate DNA-binding transcription factor regulators of the *INS* gene that underlie the deleterious effects of the c.-331C>G mutation. Four transcription factors were selected based on differential binding to c.-331C vs. c.-331C>G double stranded oligonucleotides in SILAC experiments in MIN6 b cells, while others were selected based on *in silico* predicted differential binding to c.-331C vs. c.-331C>G *INS* sequences or published studies from indicated references. The summary table emphasizes that amongst these candidates only GLIS3 led to activation of the unmodified episomal insulin promoter plasmid, and activated *INS* mRNA in non-pancreatic cell lines in the presence of islet transcription factors as shown in Figure S5.
(B) Electromobility shift assays (EMSA) from HEK 293T cells transfected with human GLIS3Δ410 cDNA show binding to CC element oligonucleotides and BS2, another previously reported GLIS3 recognition sequence in the INS 5' flanking regions[29], but not to the CC element carrying the -331C>G mutation.

Figure S4

A

Top PWMs for known TFs

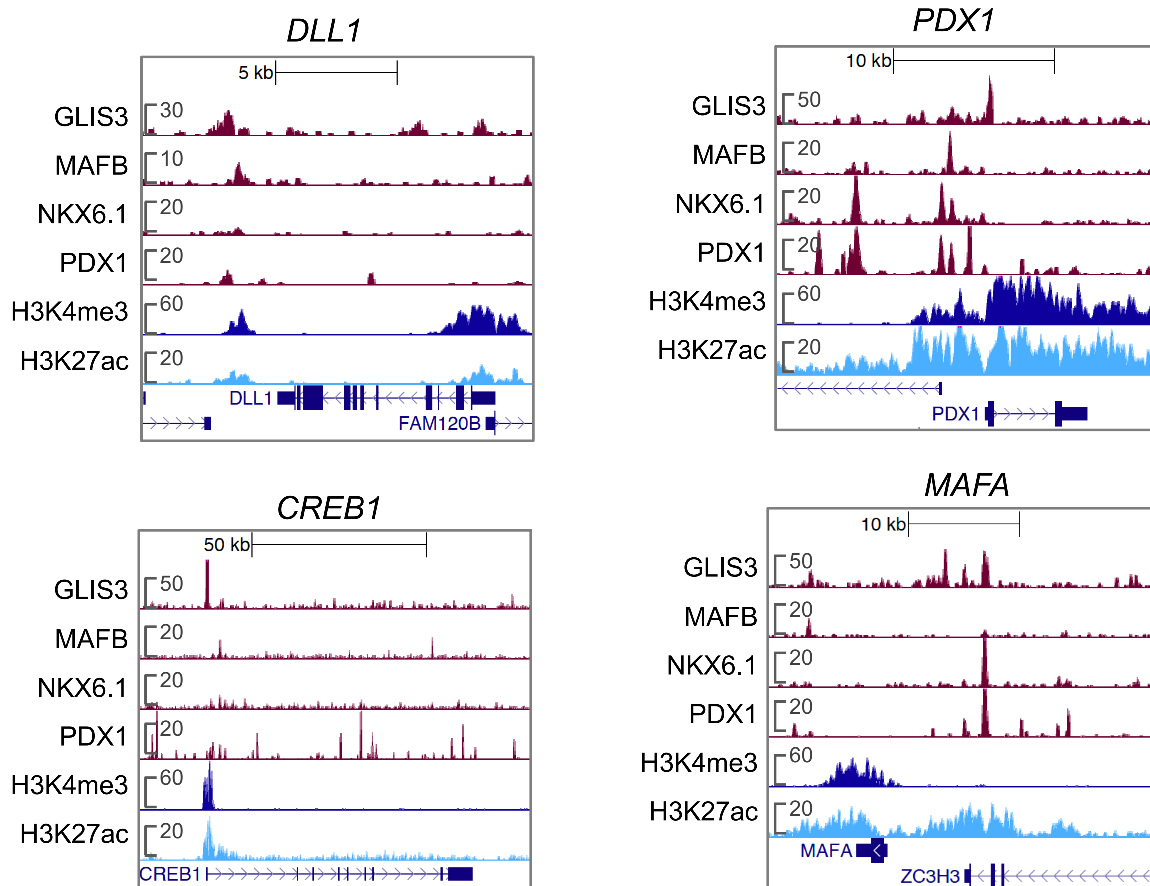
Sequence motif	Name	p-value	% of Targets Sequences with Motif	% of Background Sequences with Motif
CTCCCTGGGAGGCCT	**GLIS3**(Zf)/Thyroid-Glis3.GFP-ChIP-Seq(GSE103297)	1e-52	38.18%	15.21%
GGCCCCCTGCTGGGA	**Zic3**(Zf)/mES-Zic3-ChIP-Seq(GSE37889)	1e-27	15.62%	4.86%
CACAGCAGGGGG	Unknown-ESC-element/mES-Nanog-ChIP-Seq(GSE11724)	1e-18	10.95%	3.53%

De novo enriched motif

ACCTCCC

p-value:	1.00E-15
% of Target Sequences with motif	18.29%
% of Background Sequences with motif	8.68%

B



C

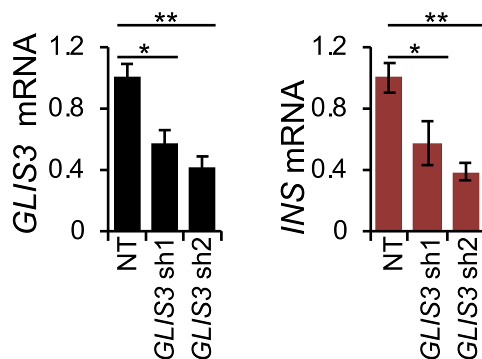


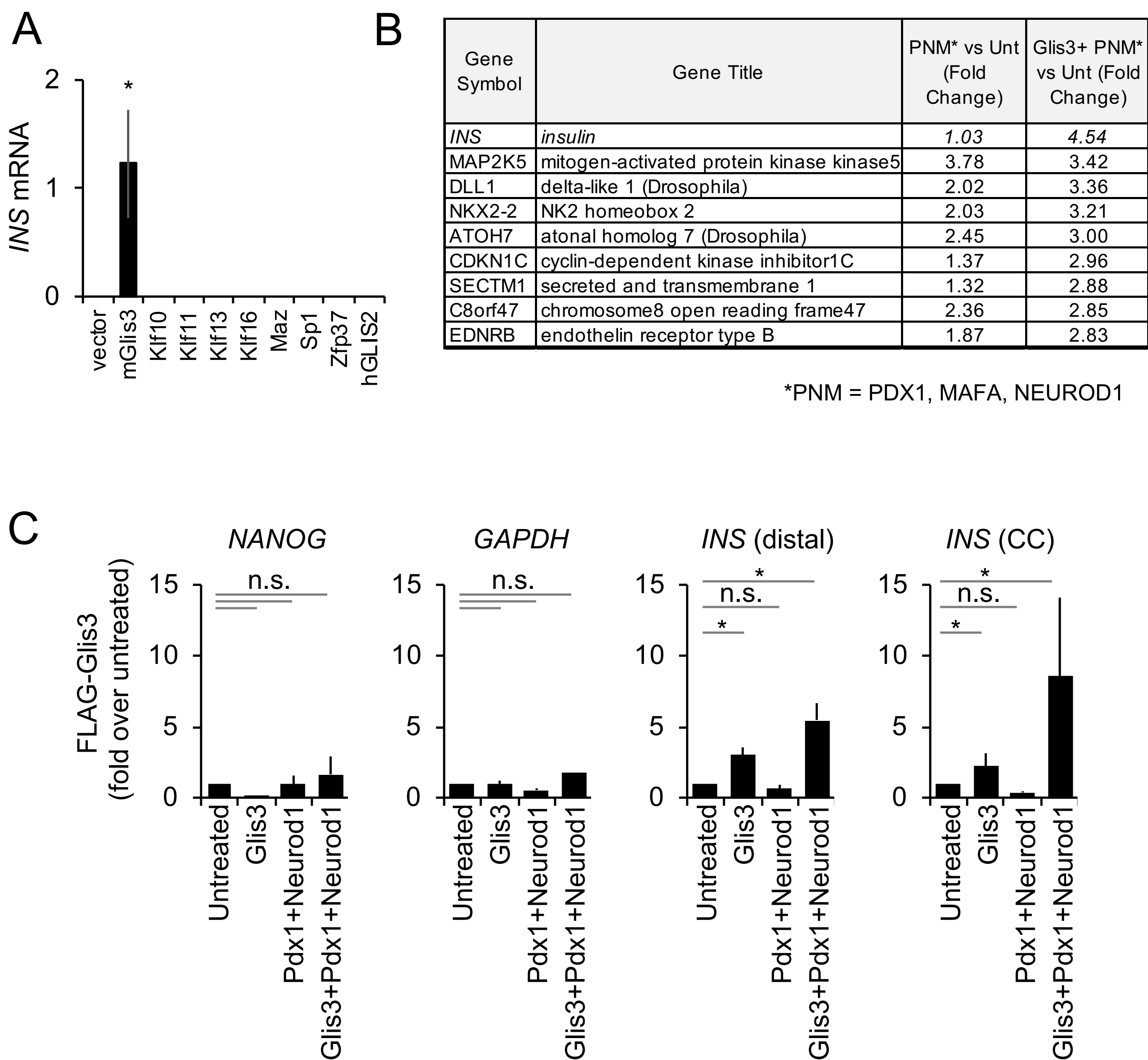
Figure S4. Related to Main Figure 6

(A) The three most enriched position weight matrixes (PWMs) at GLIS3-bound regions in human islets included a previously identified GLIS3-bound PWM (left panel). Discovery of *de novo* enriched PWMs identified sequence PWMs matching known GLIS3 recognition sequences, including one matching the *INS* CC element (right panel).

(B) Examples of GLIS3-bound regions in human pancreatic islets at selected loci, along with binding profiles of other human pancreatic islet transcription factors.

(C) Quantitative reverse-transcription PCR of INS and GLIS3 mRNAs after lentiviral-mediated knockdown of GLIS3, using two independent shRNA sequences, in EndoCb-H1 human b-cells. Student's t-test * $p < 0.05$ or ** $p < 0.001$.

Figure S5



Gene Symbol	Gene Title	PNM* vs Unt (Fold Change)	Glis3+ PNM* vs Unt (Fold Change)
INS	*insulin*	*1.03*	*4.54*
MAP2K5	mitogen-activated protein kinase kinase5	3.78	3.42
DLL1	delta-like 1 (Drosophila)	2.02	3.36
NKX2-2	NK2 homeobox 2	2.03	3.21
ATOH7	atonal homolog 7 (Drosophila)	2.45	3.00
CDKN1C	cyclin-dependent kinase inhibitor1C	1.37	2.96
SECTM1	secreted and transmembrane 1	1.32	2.88
C8orf47	chromosome8 open reading frame47	2.36	2.85
EDNRB	endothelin receptor type B	1.87	2.83

Figure S5. Related to Main Figure 6
(A) GLIS3, but not other candidate transcription factors that bind CC, activate *INS* in a heterologous cell type. HEK 293T cells were either left untreated (Unt) or transfected with plasmids overexpressing the indicated transcription factors along with NEUROD1 and PDX1. *INS* mRNA are *INS* to *GAPDH* mRNA ratios X 1000. Statistical significance was calculated relative to untreated samples (n = 3 independent experiments). * Student's t-test, $p < 0.05$. Analogous experiments were performed in the presence of MAFA, or NEUROG3 instead of NEUROD1, yielding similar results.
(B) Table showing the list of genes most impacted by the overexpression of GLIS3+ PNM (PDX1, NEUROD1, MAFA) vs. PNM alone, in HEK 293T cells. Cells were either left untreated or transfected with the indicated transcription factors. At 3 days post transfection, RNA was quantified using microarrays. Values indicate fold-difference in expression relative to control cells. This experiment was performed with a single replicate and provides an unbiased confirmation that *INS* was the single most induced gene when GLIS3 was added to the transcription factor cocktail.
(C) ChIP of FLAG-tagged GLIS3 in HEK 293T cells transfected with the indicated transcription factors. ChIP DNA was quantified by RT-PCR and expressed as percentage of input DNA and as fold enrichment over untreated HEK 293T cells. Statistical significance was calculated relative to untreated (Unt) samples using Student's t-test (n=3 experiments). As expected, *NANOG* promoter and *GAPDH* enhancer showed no changes between treatments.