Supporting Information

Garin et al. 10.1073/pnas.0910533107

SI Results

Large Deletion Detected by Multiplex Ligation-Dependent Probe Amplification. A large deletion encompassing the promoter region, the noncoding exon 1 and coding exon 2 of *INS* (c.-370-?_186+?del) was identified in a large consanguineous Lebanese family (ISPAD 182). The deletion was identified following failure of PCR amplification of *INS* exons 1 and 2. The multiplex ligation-dependent probe amplification (MLPA) assay showed that the parents had a 50% reduction in dosage for exons 1 and 2 but two copies of exon 3. Testing the proband and affected cousin showed that *INS* exon 1 and 2 MLPA probes did not bind to the target DNA (suggestive of a homozygous deletion).

Effect of the Translation Initiation Codon Mutations on Insulin Content. The homozygous mutations c.3G > T and c.3G > A were identified in two unrelated probands with permanent diabetes diagnosed on the first day of life. Both are nonsynonymous changes affecting the first methionine residue (p.Met1) and abolishing the translation initiation site for the preproinsulin protein. A second methionine residue is located at codon 5 (p.Met5), but the surrounding sequence does not conform to the Kozak consensus for translation because a thymine base is present at position −3 to the ATG codon rather than the required adenine (1). We would therefore predict that translation from the mutated initiator codon would be compromised but without experimental evidence we could not rule out a possible effect on mRNA stability.

To investigate the possibility that these p.0? mutations (c.3G > T and c.3G > A) could affect mRNA stability, we quantified the amount of *INS* mRNA in HeLa cells expressing wild-type or mutation bearing plasmids by real-time PCR. There were no differences in the amount of *INS* mRNA expressed in cells transfected with either mutation compared with cells transfected with wild-type plasmids [*INS* wild-type: *INS* c.3G > T and *INS* wild-type: *INS*: c.3G > A ratios were 1:0.99 (SD = 0.19) and 1:0.95 (SD = 0.15) respectively; $P = 0.87$].

To investigate the possibility of low levels of translation from p.Met5, we used an in vitro radio-immunoassay to measured insulin content in cells expressing the mutations compared to cells expressing wild type (WT) *INS.* The cells expressing *INS* (WT) contained 56 pg of insulin per million cells, while the insulin content was reduced by at least 79% in the cells expressing *INS* c.3G > T and c.3G > A (11.9 and 7.6pg of insulin per million cells respectively) (Fig. 4*A*). These data show that p.Met5 is unlikely to reinitiate translation of preproinsulin in the presence of the c.3G > T or c.3G > A mutations.

Effect of the c.*59A > G Mutation on RNA Stability. To investigate the effect of the c.*59A > G mutation on RNA stability we used real-time PCR to quantify the level of *INS* mRNA transcripts in a heterozygous lymphoblastoid cell line generated from the proband's mother (DM1165). We were unable to discriminate between the transcripts directly via the c.*59A/G alleles, as the mutation was located at the far 3′ end of the transcript. Instead, we used a SNP (rs3842753) to tag the mutation, with the mutation (c.*59G) on the same haplotype as the C allele of the SNP. The mutation-bearing transcript was not detectable in the heterozygous mother (Fig. 4*B*) and in the homozygous state we predict that the affected child will not express the *INS* gene.

SI Materials and Methods

Gene-Dosage Analysis Using MLPA. We designed MLPA assay oligonucleotide probes (sequences available on request) to measure the number of copies of *INS* exons 1 to 3 using probes for *HNF1A* and *HNF4A* as controls (method previously described by ref. 2).

Methodology for the Investigation of the Effect of Promoter Mutations on Insulin Gene Promoter Activity. A plasmid containing nucleotides c.-489 to c.-228 of the human insulin gene promoter (pSOUAPRL-251hINS-Luc) linked to the firefly luciferase gene (Roland Stein, Vanderbilt University) was used to perform site-directed mutagenesis with the QuikChange kit (Stratagene). Three independent clones were created for each base substitution. The plasmids were verified by direct sequencing, and assayed in transient transfection assays in triplicate on three separate occasions, as described (3). MIN6 β cells were trypsinized 24 h before transfection, plated into 24-well plates (1×10^5 cells per well) and maintained in DMEM supplemented with 15% FBS, 2 mM Glutamine, and 50 nM 2-Mercaptoethanol. Cells were transfected with promoter-luciferase constructs (500 ng) using Metafectene PRO (Biontex), following manufacturers instructions. pGL4.75[hRluc/CMV] (10 ng) was included in the transfections to normalize for transfection efficiency. Twenty-four hours posttransfection, Firefly and Renilla luciferase activity was assayed using Dual-Luciferase Reporter Assay System (Promega), and the ratios were expressed relative to the wild-type insulin promoter values.

Methodology for Investigating the Effect of the Translation Initiation Codon Mutations (c.3G > T and c.3G > A). cDNA encoding wild-type *INS* and both mutant forms of *INS*-p.0?,c.3G > T and c.3G > A, were cloned into pcDNA 3.1/*myc*-His A, which uses a CMV and T7 promoter to drive high level expression of the recombinant protein. Correct insertion was verified by sequencing of the plasmids. Next, 4×10^6 HeLa cells were cotransfected with 8-μg pcDNA 3.1 and 8-μg pMAX GFP using Nucleofector technology (Amaxa/Lonza) according to the manufacturer's instructions. GFP expression was examined by fluorescence microscopy to monitor transfection efficiency. Control cells were transfected with empty pcDNA3.1 vector.

Noninsulin expressing HeLa cells were cultured in DMEM Glutamax medium (Gibco Life Technologies) supplemented with 5% FBS, 100 units/mL penicillin, and 100 μg/mL streptomycin at 37 °C in a fully humidified atmosphere of 5% CO_2.

To rule out an effect of these mutations on mRNA stability, we quantified the amount of *INS* mRNA relative to the Beta-2-microglobulin (*B2M*) gene in transfected cells. Total RNA was extracted from $\approx 1 \times 10^6$ HeLa cells using the Perfect RNA Mini RNA kit (Eppendorf). Next, 4.5 μg mRNA was treated with 2 units of RNase-free DNase (TURBO DNase kit, Ambion) to remove Genomic DNA by following the manufacturer's instructions. cDNA was synthesized from mRNA using the Thermoscript RT-PCR system (Life Technologies) with an incubation temperature of 50 °C and a random hexamer primer. PCR products were detected by the use of a probe situated across the boundary of exons 2 and 3 to ensure amplification from cDNA rather than genomic DNA (forward primer; CGGGAGGCAGAGGACCT, reverse primer; AGGCTGCCTGCACCAG, probe; 6FAM- ACCTGCCCCACCTGC-MGB). Reactions contained 5-μL TaqMan Fast Universal PCR Master Mix, no AmpErase, 0.9 μM each primer and 0.25 μM probe in a total volume of 10 μL on the ABI prism 7900HT platform (Applied Biosystems). Amplification conditions were a single cycle of 95 °C for 20 s followed by 60 cycles of 95 °C for 1 s, and 60 °C for 20 s. The results given are an average of quantifications from three replicate amplifications.

Expression levels of the mutant and wild-type transcripts were measured by comparing the number of cycles at which the *INS* and *B2M* PCR products cross a specific threshold. The relative abundance is defined by the difference in number of cycles to achieve the same quantity of product and calculated using the equation $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct$ is the difference between the crossing points (ΔCt^{test}) in a test sample normalized to a reference sample (ΔCt^{ref}) (Applied Biosystems). Differences in the mRNA abundance for wild-type and mutant transcripts were examined for statistical significance by Kruskal–Wallis test.

The insulin content of transfected HeLa cells was determined 24 h posttransfection by radio-immunoassay. Briefly, cells were lysed in 200 μL of acidified ethanol (0.8 M HCl:98% ethanol mixed in a ratio of 1:3) for 2 h at −20 °C and samples then neutralized by addition of NaOH. In triplicate, 50 μL of each experimental sample or crystalline recombinant human insulin standard was mixed with 50 μL of ^{125}I insulin (Linco; diluted to yield ≈2,000 cpm/50 μL) and 50 μL Guinea Pig anti-bovine insulin antibody (ICN) diluted 1:20,000 in insulin assay buffer (IAB; 150 mM NaCl, 30 mM Na_2HPO4, 10 mM KH_2PO_4, 10 mM EDTA, 5 mg/mL BSA, pH 7.4). Tubes were incubated at 4 °C overnight before addition of 50 μL of donkey anti-Guinea Pig coated cellulose beads (IDS Ltd.) (diluted 1:1 with IAB). After incubation for 1 h, 2-mL dH_2O was added and tubes centrifuged for 5 min at 1,000 × *g*. The supernatant was aspirated under vacuum and the radioactivity retained in each pellet measured on a WALLAC gamma counter. The insulin content of each sample was then calculated by reference to a standard curve constructed with recombinant human insulin.

Methodology for Investigating the Effect of the c.*59A > G Mutation on mRNA Stability. Cell lines were established from peripheral blood lymphocytes derived from the proband's heterozygous mother (DM1165) and an unaffected control by EBV transformation. Cell lines were maintained in 1× RPMI-1640 (Gibco Life Technologies), supplemented with 10% FCS (Gibco Life Technologies).

Total RNA was extracted from ≈1 × 10^6 EBV-transformed lymphoblastoid cells using the Perfect RNA Mini RNA kit (Eppendorf). Next, 4.5 μg mRNA was treated with 2 units of RNase-free DNase (TURBO DNase kit, Ambion) 30 min at 37 °C followed by 85 °C for 5 min for nuclease inactivation. CDNA (cDNA) was synthesized from mRNA using the Thermoscript RT-PCR system (Life Technologies) with an incubation temperature of 50 °C and a random hexamer primer.

Ectopic mRNA transcripts were amplified from lymphoblastoid cells using a single tube TaqMan approach. It was not possible to design probes for the c.*59A > G mutation because of its location at the extreme 3′ end of the transcript. Instead we used a heterozygous single nucleotide polymorphism (rs3842753) located 37 nucleotides upstream of the poly-A tail to differentiate the mutation-bearing and normal transcripts. Probe and primer sequences were designed to this variant and validated by standard curve analysis [Forward – 5′ ggagaactactgcaactagac 3′, Reverse – 5′ catctctctcggtgcaggag 3′, Probe (WT) – VIC-cagccccAcacccg-MGB, Probe (MT) – 6-FAM-agccccCcacccg-MGB].

Two microliters cDNA from the proband's mother (heterozygous for c.*59A > G and rs3842753) was used for real-time PCR quantification and cDNA from a normal cell line (heterozygous for rs3842753 only). PCR products were detected by the use of mutation-specific probes which were identical except for the site of the mutation. Reactions contained 5-μL TaqMan Fast Universal PCR Master Mix, no AmpErase, 0.36 μM each primer, and 0.08 μM each probe in a total volume of 10 μL on the ABI prism 7900HT platform (Applied Biosystems). Amplification conditions were a single cycle of 95 °C for 20 s, followed by 60 cycles of 95 °C for 1 s, and 60 °C for 20 s. The results given are an average of quantifications from three triplicate amplifications derived from two separate RNA extractions.

To validate the real-time assay, we conducted standard curve analysis, which indicated that the assay was accurate and quantitative over seven serial 1:2 dilutions. The efficiency of amplification as assessed by the gradient of the standard curves for mutant and wild-type probes were −3.1 and −3.0, respectively. The correlation between crossing point and input template (r^2) was 0.95 and 0.92 for the mutant and wild-type probes, respectively. The A- and C-allele transcripts of rs3842753 are present in approximately equal amounts in the normal heterozygous cell line, indicating that the presence of this SNP does not adversely affect the stability of either allele.

Expression levels of the mutant and wild-type transcripts were measured using the allele-specific assays by comparing the number of cycles at which the two different PCR products cross a specific threshold. The relative abundance is defined by the difference in number of cycles to achieve the same quantity of product and calculated using the equation $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct$ is the difference between the crossing points [(ΔCt^{test}) in a test sample normalized to the difference between mutant and normal crossing points in a 50:50 mixture ($\Delta Ct^{50\%}$) (Applied Biosystems (4)].

1. Kozak M (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol Cell Biol* 9:5134–5142.
2. Ellard S, et al. (2007) Partial and whole gene deletion mutations of the GCK and HNF1A genes in maturity-onset diabetes of the young. *Diabetologia* 50:2313–2317.
3. Hansen SK, et al. (2002) Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function. *J Clin Invest* 110:827–833.
4. Applied Biosystems (2001) Relative quantitation of gene expression. *User Bulletin* 2:11–15.

Table S1. Assessment of pathogenicity of noncoding mutations giving data on the conservation of the nucleotides and the number of control chromosome sequences tested

Variant	*n*	Conservation					Control chromosomes sequenced for variant	Experimental evidence for loss of function
		Chimp	Rhesus	Tree shrew	Mouse	Rat		
c.-332C > G	1	Yes C	Yes C	Yes C	Yes C	Yes C	496	Yes
c.-331C > G	7	Yes C	Yes C	No T	Yes C	Yes C	496	Yes
c.-331C > A								
c.-218A > C	1	Yes A	Yes A	Yes A	Yes A	Yes A	496	ND
c.*59A > G	1	Yes A	Yes A	Yes A	Yes A	Yes A	496	Yes

The variants were not detected in any of the control chromosomes. ND, not done.

Table S2. Clinical characteristics of the 21 patients with recessive *INS* mutations

Family identifier	Sex	Mutation	Relationship to proband	Ethnicity	Known Consang.	Age at diag. (days)	Glucose at diag. (mmol/L)	Disease status	Age at remission/ relapse (wk)	Birth weight (kg)	Birth length (cm)	GA (wk)	Current Age (yr)	Current treatment (U/kg/d)	Current HbA1c
DM 1293	F	c.-366_-343del	Proband	Caucasian	No	3	13	PNDM	NA	1.45	44	41	23	Insulin	6.2
DM 1293	M	c.-366_-343del	Brother	Caucasian	No	1	14	PNDM	NA	1.5	43	36	14.2	Insulin	7.6
DM 1265	M	c.-332C > G + c. c.331C > G	Proband	Caucasian	No	24	34	TNDM	9	2.5	48	37	1.9	None	4.8
ISPAD185	M	c.-331C > G	Proband	Arabic	Yes	28	20	PNDM	NA	1.8	NA	40	4	Insulin (0.5)	NA
ISPAD185	M	c.-331C > G	Cousin	Arabic	Yes	7	18	TNDM	11	1.8	NA	40	2	None	9.0
ISPAD186	M	c.-331C > G	Proband	Arabic	Yes	50	27	PNDM	NA	2.4	NA	40	15	Insulin (1.0)	11.5
ISPAD186	F	c.-331C > G	Sister	Arabic	Yes	35	NA	PNDM	NA	NA	NA	NA	25	Insulin	9.0
ISPAD186	F	c.-331C > G	Sister	Arabic	Yes	4380	NA	Diabetes	NA	NA	NA	NA	22	Insulin	11.6
ISPAD186	M	c.-331C > G	Brother	Arabic	Yes	6570	10	Diabetes	NA	NA	NA	NA	19	Insulin	14.0
ISPAD183	M	c.-331C > G	Proband	Turkish	Yes	20	28	PNDM	NA	1.5	NA	40	21	Insulin (1.0)	8.5
ISPAD184	F	c.-331C > G	Proband	Caucasian	No	2	40	PNDM	NA	1.8	45	37	10	Insulin (0.6)	7.3
ISPAD 83	F	c.-331C > G	Proband	Caucasian	No	1	16	PNDM	NA	1.0	NA	36	10	NA	7.1
ISPAD 187	M	c.-331C > A	Proband	Caucasian	No	3	29	TNDM	22	1.9	NA	38	3.2	None	NA
ISPAD 188	M	c.-331C > A	Proband	Bangladeshi	No	42	15	TNDM	12	2.7	NA	40	2	None	6.1
ISPAD 189	F	c.-218A > C	Proband	Indian	Yes	56	27	TNDM	36/48	2.6	NA	40	1.1	Insulin (1.0)	NA
ISPAD 169	M	c.3G > A (p.0?)	Proband	Hispanic	Yes	5	28	PNDM	NA	1.8	45	37	1.7	Insulin (0.6)	5.6
ISPAD 170	F	c.3G > T (p.0?)	Proband	Pakistani	Yes	1	16	PNDM	NA	1.2	39	36	5.7	Insulin (0.9)	8.5
ISPAD 171	M	c.184C > T (p.Q62X)	Proband	Indian	Yes	7	59	PNDM	NA	1.3	39	35	0.8	Insulin (0.6)	8.3
DM1165	F	c.*59A > G	Proband	Caucasian	Yes	4	NA	PNDM	NA	1.5	42	39	27.5	Insulin	6.2
ISPAD 182	M	c.-370-?_186+?del	Proband	Lebanese	Yes	1	NA	PNDM	NA	1.3	NA	35	3.9	Insulin	7.5
ISPAD 182	M	c.-370-?_186+?del	Cousin	Lebanese	Yes	1	1	PNDM	NA	1.6	43.5	36	8.9	NA	8.7

All mutations are homozygous except for DM 1265, which is compound heterozygous for the two mutations (inherited in *trans*). GA, gestational age; NA, not applicable/available; PNDM, permanent neonatal diabetes; TNDM, transient neonatal diabetes.

Table S3. Comparison of clinical characteristics in patients with permanent or transient neonatal diabetes caused by recessive *INS* mutation

Characteristic	Transient	Permanent	*P*-value
n	5	14	NA
Type of mutation (% regulatory)	100	64	0.3
Sex (% male)	80	57	0.6
Birth weight (g)	2,540 (1,860; 2,650)	1,500 (1,300; 1,800)	0.007
Gestational age (wk)	38 (37.5, 40)	37 (36, 40)	0.3
Birth weight (SD score)	−1.8 (−3.4, −0.9)	−3.9 (−4.4, −2.8)	0.03
Age at diagnosis (days)	24 (5, 62)	2 (1, 9.5)	0.04
Glucose at diagnosis (mmol/L)	27 (18, 31)	22 (15, 31)	0.7
Bicarbonate at diagnosis (mmol/L)	9 (7, 12)	23 (20, 24)	0.06
Remission (%)	100	0	NA
Age at remission (wk)	12 (11, 22)	NA	NA
Age at relapse (yr)	1 (1 case)	NA	NA
Current age (yr)	2 (1.5, 3)	10 (4, 18)	0.02
Current HbA1c (%)	5.5 (4.8, 6.1)	8.3 (6.7, 8.9)	0.04

Data are median (interquartile range). NA, not applicable.

Table S4. Clinical characteristics in patients with isolated neonatal diabetes without mutations in any of the reported genes

Characteristic	Transient	Permanent	*P*-value
n	13	85	NA
Sex (% male)	64	56	0.8
Birth weight (g)	2250 (2077, 2933)	2725 (1900, 3213)	0.5
Gestational age (wk)	40 (37, 40)	39 (37, 40)	0.8
Birth weight (SD score)	−1.6 (-2.9, -0.8)	−1.1 (-2.5, -0.3)	0.4
Age at diagnosis (wk)	3 (1, 8)	5 (1, 22)	0.4
Remission (%)	100	0	NA
Age at remission (wk)	17 (6, 48)	NA	NA
Relapse (%)	38.5	0	NA
Age at relapse (yr)	4 (2, 5.5)	NA	NA

Data are median (interquartile range). NA, not applicable.