View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Queensland eSpace

Recessive mutations in the *INS* gene result in neonatal diabetes through reduced insulin biosynthesis

Intza Garin[a,b,c,1], Emma L. Edghill[d,1], Ildem Akerman[b,e,f,1], Oscar Rubio-Cabezas[d,g,1], Itxaso Rica[a], Jonathan M. Locke[d], Miguel Angel Maestro[b,e,f], Adnan Alshaikh[h], Ruveyde Bundak[i], Gabriel del Castillo[j], Asma Deeb[k], Dorothee Deiss[l], Juan M. Fernandez[m], Koumudi Godbole[n], Khalid Hussain[o], Michele O'Connell[p], Thomasz Klupa[q], Stanislava Kolouskova[r], Fauzia Mohsin[s], Kusiel Perlman[t], Zdenek Sumnik[r], Jose M. Rial[u], Estibaliz Ugarte[v], Thiruvengadam Vasanthi[w], Neonatal Diabetes International Group[2], Karen Johnstone[d], Sarah E. Flanagan[d], Rosa Martínez[a,b,c], Carlos Castaño[b,e,f], Ann-Marie Patch[d], Eduardo Fernández-Rebollo[a,b,c], Klemens Raile[l], Noel Morgan[d], Lorna W. Harries[d], Luis Castaño[a,b,c], Sian Ellard[d], Jorge Ferrer[b,e,f], Guiomar Perez de Nanclares[a,b,c,3], and Andrew T. Hattersley[d,3]

[a]Endocrinology and Diabetes Research Group, Hospital de Cruces, Barakaldo, 48903 Spain; [b]Centro de Investigación Biomédica en Red en Diabetes y Metabolismo; [c]Centro de Investigación Biomédica en Red en Enfermedades Raras, Instituto de Salud Carlos III, 48903 Spain; [d]Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter, United Kingdom; [e]Hospital Clínic de Barcelona, Spain; [f]Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; [g]Department of Endocrinology, Hospital Infantil Universitario Niño Jesús, Madrid, Spain; [h]King Abdulaziz Medical City, National Guard Health Affairs, Jeddah, Kingdom of Saudi Arabia; [i]Department of Pediatric Endocrinology, Istanbul University Hospital, Istanbul, Turkey; [j]Department of Neonatology, Hospital Infantil Los Angeles, San Juan de Pasto, Columbia; [k]Imperial College London Diabetes Centre, Abu Dhabi, United Arab Emirates; [l]Department of Pediatric Diabetology and Endocrinology, Charité Campus, Virchow Children's Hospital, Berlin, Germany; [m]Department of Pediatric Endocrinology, Hospital Universitario San Cecilio, Granada, 18012 Spain; [n]Deenanath Mangehskar Hospital and Research Center, Erandawane, Pune, India; [o]Great Ormond Street Hospital for Children NHS Trust, London, United Kingdom; [p]Royal Children's Hospital, Melbourne, Australia; [q]Department of Metabolic Disease, Jagiellonian University Medical College, Krakow, Poland; [r]Department of Pediatrics, Second Faculty of Medicine, Charles University, Prague, Czech Republic; [s]Pediatric Diabetic Team, Bangladesh Institute of Research and Rehabilitation in Diabetes, Endocrine and Metabolic Disorders, Dhaka, Bangladesh; [t]Division of Endocrinology, Hospital for Sick Children, Toronto, Canada; [u]Department of Pediatric Endocrinology, Hospital Nuestra Señora de la Candelaria, Tenerife, 38010 Spain; [v]Department of Endocrinology, Hospital de Basurto, Bilbao, 48013 Spain; and [w]Kanchi Kamakoti Childs Trust Hospital, Chennai, India

Edited by Donald F. Steiner, University of Chicago, and approved December 22, 2009 (received for review September 15, 2009)

Heterozygous coding mutations in the *INS* gene that encodes preproinsulin were recently shown to be an important cause of permanent neonatal diabetes. These dominantly acting mutations prevent normal folding of proinsulin, which leads to beta-cell death through endoplasmic reticulum stress and apoptosis. We now report 10 different recessive *INS* mutations in 15 probands with neonatal diabetes. Functional studies showed that recessive mutations resulted in diabetes because of decreased insulin biosynthesis through distinct mechanisms, including gene deletion, lack of the translation initiation signal, and altered mRNA stability because of the disruption of a polyadenylation signal. A subset of recessive mutations caused abnormal *INS* transcription, including the deletion of the C1 and E1 *cis* regulatory elements, or three different single base-pair substitutions in a CC dinucleotide sequence located between E1 and A1 elements. In keeping with an earlier and more severe beta-cell defect, patients with recessive *INS* mutations had a lower birth weight (−3.2 SD score vs. −2.0 SD score) and were diagnosed earlier (median 1 week vs. 10 weeks) compared to those with dominant *INS* mutations. Mutations in the insulin gene can therefore result in neonatal diabetes as a result of two contrasting pathogenic mechanisms. Moreover, the recessively inherited mutations provide a genetic demonstration of the essential role of multiple sequence elements that regulate the biosynthesis of insulin in man.

gene regulation | genetic testing | gene expression regulation | RNA instability | promoter regions

Neonatal diabetes is diagnosed within the first 6 months of life (1, 2) and there are two main clinical subtypes: the persistent, permanent neonatal diabetes (PNDM) and the remitting and frequently relapsing, transient neonatal diabetes (TNDM). Recently there have been considerable advances in the understanding of the genetics of neonatal diabetes (3). Most patients with PNDM have activating mutations in *KCNJ11* or *ABCC8*, the genes encoding the potassium ATP-sensitive (K_{ATP}) channel subunits Kir6.2 (4) and SUR1 (5–7), or heterozygous mutations in the preproinsulin (*INS*) gene (8–12). In contrast, abnormalities in chromosome 6q24 are the most common cause of TNDM (13), followed by mutations in the *KCNJ11* and *ABCC8* genes (14). Despite these advances, the etiology of neonatal diabetes is still not known in at least 30% of patients with PNDM, suggesting other genetic causes are still to be found (9).

Insulin is secreted from islet beta cells of the pancreas. Insufficient secretion of insulin results in hyperglycemia and diabetes, whereas excessive secretion results in hypoglycemia. Insulin biosynthesis and secretion are therefore tightly regulated to maintain blood glucose levels within a narrow physiological range. Extensive studies have dissected an array of *cis* sequence elements in the *INS* promoter region and their cognate DNA binding factors, which together ensure the cellular specificity and rate of *INS* transcription (15–22). In addition, insulin biosynthesis is strongly dependent on posttranscriptional regulatory mechanisms, including the modulation of translation and stability (23–25). The latter is largely mediated through sequences located in the untranslated regions of *INS* transcripts (26–28).

MEDICAL SCIENCES

Author contributions: O.R.-C., K.R., N.M., L.W.H., L.C., S.E., J.F., G.P.d.N., and A.T.H. designed research; I.G., E.L.E., I.A., O.R.-C., I.R., J.M.L., M.A.M., A.A., R.B., G.d.C., A.D., D.D., J.M.F., K.G., K.H., M.O., T.K., S.K., F.M., K.P., Z.S., J.M.R., E.U., T.V., N.D.I.G., K.J., S.E.F., R.M., C.C., A.-M.P., E.F.-R., and K.R. performed research; I.G., I.A., O.R.-C., L.W.H., S.E., J.F., and G.P.d.N. analyzed data; and E.L.E., O.R.-C., J.F., G.P.d.N., and A.T.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]I.G., E.L.E., I.A., and O.R.-C. contributed equally to this work.

[2]Members of the Neonatal Diabetes International Group: Firdevs Bas, Faculty of Medicine, Istanbul University, Istanbul, Turkey; Ondrej Cinek, Motol University Hospital, Czech Republic; Maciek Malecki, Department of Metabolic Disease, Jagiellonian University Medical College, Krakow, Poland; and Marianna Rachmiel, Division of Endocrinology, Hospital for Sick Children, Toronto, Canada.

[3]To whom correspondence may be addressed. E-mail: andrew.hattersley@pms.ac.uk or gnanclares@osakidetza.net.

This article contains supporting information online at www.pnas.org/cgi/content/full/0910533107/DCSupplemental.

Heterozygous missense mutations in the coding region of the *INS* gene have recently been described as a cause of neonatal diabetes (8–12). Most of the reported mutations are predicted to disrupt the folding of the proinsulin molecule. The resulting misfolded protein accumulates in the endoplasmic reticulum (ER), resulting in ER stress and beta-cell apoptosis (29, 30). An alternative potential genetic mechanism would be reduced insulin secretion because of a disruption of the *INS* coding sequence, as seen in the double *Ins1* and *Ins2* knockout mouse (31), or of the sequences that regulate insulin biosynthesis. However, as yet this has not been demonstrated in humans.

We now report recessively acting mutations within the *INS* gene in a series of patients with neonatal diabetes. In contrast to the previously described dominant mutations, these mutations reduce insulin synthesis and thus represent a unique pathogenic mechanism for human diabetes. These mutations also provide genetic evidence for the essential role of distinct nucleotide sequences in the regulation of the human preproinsulin gene.

Results

Recessive *INS* Mutations Cause Neonatal Diabetes. We sequenced 117 unrelated probands with diabetes diagnosed before 6 months (13 offspring of consanguineous parents) in whom the known common genetic causes had been excluded. We identified 10 different *INS* recessive mutations in 15 unrelated families (Figs. 1 and 2). Four homozygous mutations affected the coding region: c.184C > T (p.Q62X), c.3G > T (p.0?), c.3G > A (p.0?), and a large deletion that removes a segment of the promoter, exon 1 and coding exon 2 of *INS* (c.-370-?_186+?del). Five homozygous mutations were found in regulatory regions: c.-331C > A (2 families), c.-331C > G (5 families), c.-218A > C, and a 24-base pair deletion (c.-366_-343del) are located in the promoter region, whereas c.*59A > G is within the 3′ untranslated region. One proband was a compound heterozygote for two regulatory region mutations, c.-331C > G and c.-332C > G.

The mutations were inherited in a recessive manner either homozygous or compound heterozygous, with heterozygous carrier parents being unaffected with neonatal diabetes (Fig. 2). Pathogenicity of mutations was suggested by conservation across species and absence of variants in controls (Table S1).

Recessive *INS* Mutations Uncover Essential Regulatory Sequences in Humans. Further support for the pathogenicity of mutations came from known function of mutated residues and functional studies (Figs. 3 and Figs. 4; see also *SI Results*). Multiple mutation mechanisms were involved in the recessive *INS* mutations, which are described briefly below.

Truncated proteins. The nonsense mutation (p.Q62X) is predicted to give rise to a mutant protein that is truncated within the C-peptide region and will lack the insulin-A chain.

Promoter mutations. The (c.-366_-343del) 24-base pair deletion abolishes the *INS* promoter evolutionary conserved C1 and E1 elements, where MAFA and NEUROD1 bind, respectively (16, 20, 32) (Fig. 3*A*). The c.-218A > C mutation disrupts the CRE3 site that interacts with multiple DNA binding proteins in vitro (22) (Fig. 3*A*). All of these elements have been previously shown to be critically important for the *INS* promoter activity in transient transfection studies (15, 18, 33–35). The c.-331(C > G, C > A) and c.-332C > G mutations were located between the E1 and A1 elements (Fig. 3*A*). This sequence is conserved among a subset of mammalian species (Fig. 3*A*) and mutagenesis of multiple bases neighboring this dinucleotide impairs *INS* promoter activity (36). We constructed insulin promoter fragments carrying the point mutations c.-331(C > G, C > A) and c.-332C > G. The point mutations induced up to 90% reduction in transcriptional activity, while a control mutation, c.-339G > A, did not alter the transcriptional activity in pancreatic beta-cell lines (Fig. 3*B*). Thus, the CC dinucleotide that is mutated in eight unrelated probands with neonatal diabetes forms part of an essential positive *cis* regulatory sequence of the *INS* promoter.

Mutated or absent translational start site. The two point mutations (c.3G > A and c.3G > T) at the first methionine residue (p.Met1) abolish the native translation initiation site for the preproinsulin protein. Quantification of total *INS* mRNA levels by real-time PCR revealed no differences in mRNA abundance for c.3G > A or c.3G > T mutations compared with the wild type. The insulin content of HeLa cells transfected with these mutations was reduced by 86% and 79% for c.3G > A and c.3G > T, respectively, compared to cells transfected with the wild-type sequence (Fig. 4*A* and *SI Results*). The multiexon deletion (exons 1 and 2) removes over half the coding region including the translational start site and is expected to be a null mutation.

Altered mRNA stability through a mutation in the 3′ untranslated region. The c.*59A > G mutation is located in the polyadenylation signal of the 3′ untranslated region and potentially impairs mRNA stability. In a heterozygous lymphoblastoid cell line generated from the proband's mother, the mutant mRNA transcript was present at a very low level compared to the wild-type allele. This is consistent with reduced mRNA stability (Fig. 4*B* and *SI Results*).

Clinical Phenotype of Patients with Recessive *INS* Mutations. The clinical characteristics of patients with recessive *INS* mutations are shown in Table 1 (Table S1 and S2). In keeping with the known actions of insulin before and after birth, the phenotype was limited to markedly reduced fetal growth and diabetes.

The diabetes phenotype within the families is shown in Fig. 2. Nineteen patients had neonatal diabetes (15 probands and 4 family members); 14 had PNDM and were treated with insulin from diagnosis, and 5 patients had TNDM, having gone into remission at a median age of 12 weeks [interquartile range (IQR) 11, 22]. Birth weight was markedly reduced in all patients with neonatal diabetes resulting from recessive mutations [median birth weight 1,680 g (1,420; 2,050), which is −3.2 SD score (−4.1, −2.6)]. In keeping with more severe insulin deficiency, patients with PNDM had a more severe intrauterine growth retardation [median SD score for birth weight −3.9 (−4.4, −2.8) vs. −1.8 (−3.4, −0.9) in TNDM, $P = 0.03$] and diabetes was diagnosed earlier [2 days (1, 9.5) vs. 24 days (5, 62), $P = 0.04$] (Table S3). All patients with mutations that altered the coding region or mRNA stability had PNDM. The noncoding promoter mutations were associated with both PNDM and TNDM. A summary of the remaining 98 patients with neonatal diabetes of unknown origin is given in Table S4.

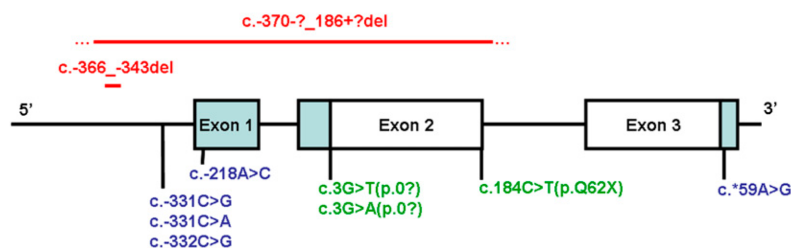

Fig. 1. A schematic of the *INS* gene showing the 10 mutations identified in 15 families. Positions of point mutations are indicated below the exons, while deletions are shown above the gene. The blue shaded regions are noncoding, the red text indicates a deletion, the blue text are noncoding mutations, and the green are coding mutations. The precise breakpoints of the multiexonic deletion are not known; the solid line represents the minimal deleted region. Mutation nomenclature is based on the coding sequence where nucleotide 1 represents translational start site.

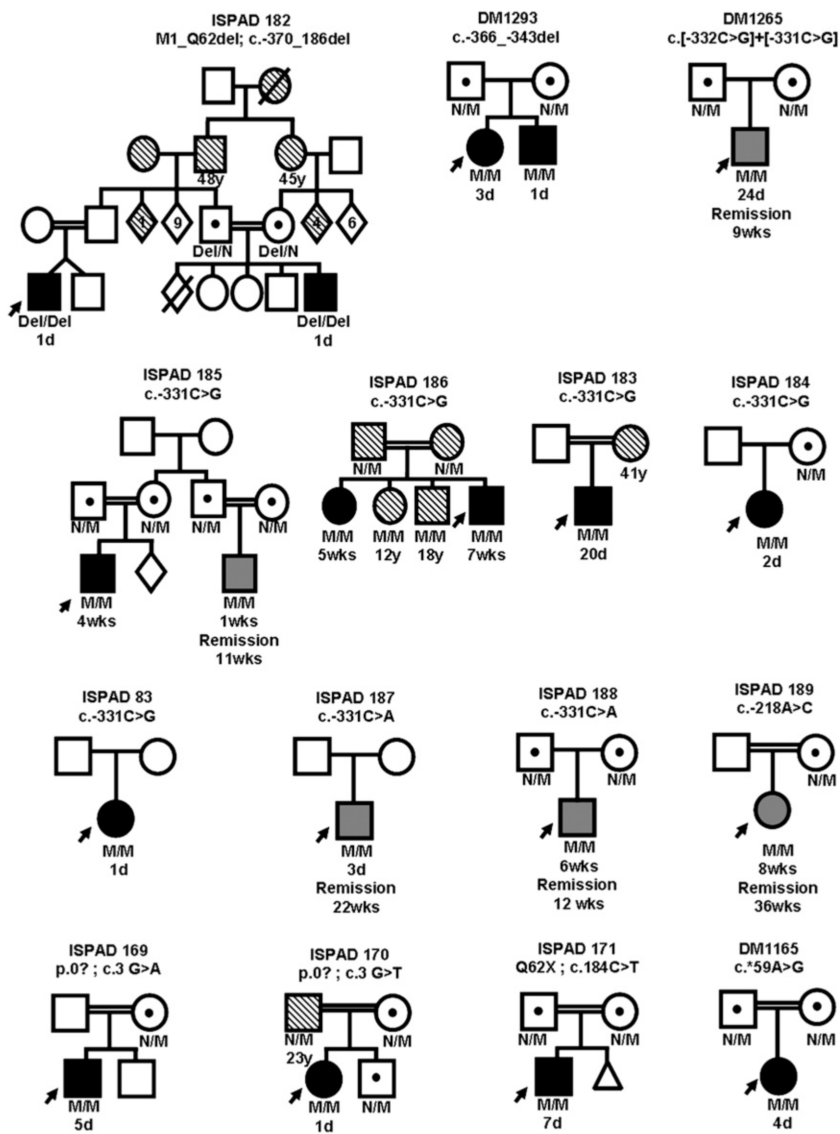

Fig. 2. Partial pedigrees of the 15 families with recessive *INS* mutations. (Del, deletion; *n*, Normal allele; M, mutation). Solid black-filled shapes represent patients with permanent neonatal diabetes, gray filled shapes represent patients with transient neonatal diabetes, and shapes filled with diagonal lines represent those patients diagnosed with diabetes after 6 months of age. Age at diagnosis and remission (where applicable) are shown below the symbols.

Differences in the Clinical Phenotype with Recessive and Dominant *INS* Mutations. To identify whether the different mutation mechanisms in the same gene resulted in phenotypic differences, we compared the clinical characteristics of patients with neonatal diabetes as a result of recessive *INS* mutations with patients with the previously identified dominant mutations in *INS* (Table 1). Patients with neonatal diabetes resulting from recessive *INS* mutations had a markedly different phenotype, with lower birth weight [median SD score −3.2 (IQR −4.1, −2.6) vs. −2.0 (−2.5, −1.0), $P < 0.001$] and an earlier age of diagnosis [median age in weeks 1 (0, 3) vs. 10 (5, 22), $P < 0.001$]. TNDM is only seen in patients with recessive mutations (26 vs. 0%, $P = 0.001$). Overall, recessive *INS* mutations accounted for 3.7% PNDM ($n = 300$) and 2.2% TNDM ($n = 134$) in a consecutive series of patients with isolated neonatal diabetes referred to the two laboratories for genetic testing.

Discussion

We have shown that recessively acting mutations in the preproinsulin gene (*INS*) are a cause of neonatal diabetes. They act by reducing synthesis of the preproinsulin peptide because of a truncated protein, abnormal transcription, reduced mRNA stability, or disrupted translation. These mutations usually cause PNDM but may manifest as TNDM or diabetes outside the neonatal period. In keeping with the recessive inheritance, many probands (60%) were the offspring of consanguineous parents.

The clinical manifestations of recessive *INS* mutations reflect the consequences of insulin deficiency in humans during pre- and postnatal life. The birth weight was markedly reduced [median SD score −3.2 (−4.1, −2.6)], consistent with the major role of insulin in fetal growth. The early onset of neonatal diabetes (median 1 week) reflects severe insulin deficiency postnatally. In contrast to many other subtypes of neonatal diabetes, there are no extrapancreatic features.

Differences in the underlying pathophysiology explain why patients with recessive *INS* mutations are diagnosed earlier and have a lower birth weight than patients with heterozygous *INS* mutations (8–12). The disrupted insulin synthesis seen with recessive mutations occurs as soon as the fetal beta cell starts to secrete insulin. In contrast, insulin secretion is required before beta-cell dysfunction develops in patients with heterozygous mutations, which result in misfolding of the preproinsulin peptide, accumulation of the misfolded protein in the ER, and hence the destruction of the beta cell through ER stress. These two distinct disease mechanisms are supported by phenotypic studies in mouse models, where reduced insulin secretion at birth or progressive ER

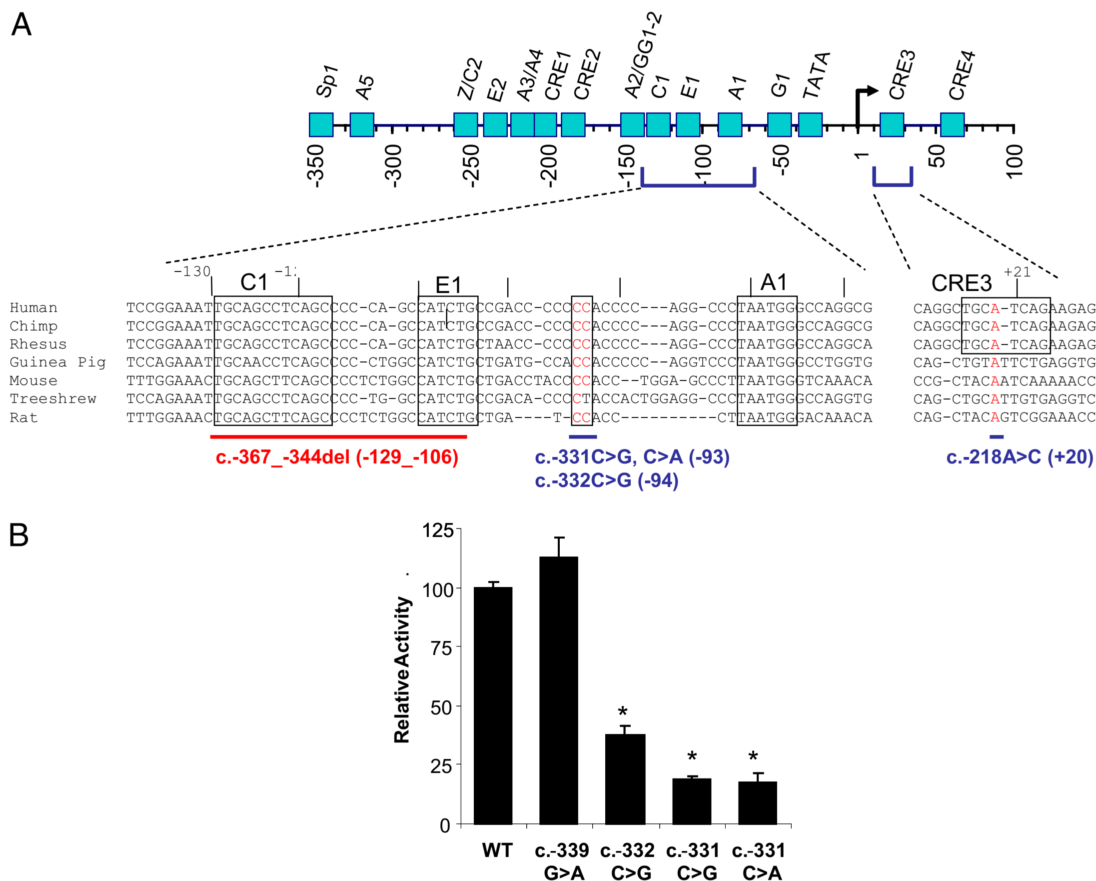

Fig. 3. Functional evidence for the pathogenicity of recessive promoter *INS* mutations. (*A*) Schematic of the genomic sequence of the *INS* promoter structure with major *cis* regulatory elements, and the sequence context of mutated elements in several mammalian species that do not exhibit major divergence in these regions. Mutated bases are highlighted in red. The numbering of promoter landmarks is relative to the transcription start site (genomic numbering, where g.1 is equivalent to c.-238) consistent with the convention used in previous studies. Mutations are described according to Human Genome Variation Society guidelines (http://www.hgvs.org/mutnomen/) (cDNA numbering according to the translational start site where c.1 is equivalent to g.238), and distance to the conventional *INS* transcriptional start site is shown in parenthesis. (*B*) Evidence for loss-of-function of the c.-331(C > G, C > A) and c.-332C > G mutations. Firefly luciferase expression is compared in constructs containing the wild-type (WT) *INS* promoter sequence (*INS* WT), or c.-331 C > G, c.-331 C > A, c.-332 C > G, c.-339G > A mutations, after transfection in MIN6 β-cells. Data shown are means (+/−SE) from three independent constructs for each mutation ($n = 3$ replicates). c.-339G > A is a control mutation that does not impair *INS* transcription. Results are corrected for transfection efficiency using a vector that constitutively expresses Renilla luciferase, and are expressed relative to the *INS* WT results. The asterisks denote $P < 0.0001$ in ANOVA for the difference between *INS* WT and mutant constructs.

stress and cell death have been described in mice carrying analogous recessive or dominant mutations, respectively (30, 31, 37).

The majority of the patients with neonatal diabetes have PNDM, but 26% (5 of 19) have TNDM. TNDM is only found in patients with noncoding mutations and they have a higher birth weight and are diagnosed later. This finding is consistent with TNDM resulting from a less severe insulin deficiency, and is comparable to the situation with mutations in the Kir6.2 subunit of the K_{ATP} mutations (38), where TNDM mutations have less severe functional consequences. The mechanism of remission in recessive *INS* mutation carriers is not understood but is likely to reflect a variation in demand or the ability of the beta cell to meet this demand, as a similar timing of remission is seen in some patients with less severe mutations, resulting in channelopathies (14, 38) and pancreatic developmental defects (39, 40).

The mutations identified in this study illustrate multiple mechanisms by which insulin biosynthesis can be disrupted. These include absent or altered translation because of coding-sequence deletions or mutations, reduced transcription because of mutations of the promoter, or abnormal mRNA stability. Our functional studies established that the 3′ UTR mutation that abolishes the polyadenylation signal results in severe RNA instability and that the initiation codon mutations result in reduced transcription of the preproinsulin gene. The promoter mutations are highly informative because they provide human genetic evidence that discrete *INS cis* regulatory elements are essential. Numerous studies have demonstrated that multiple *cis* elements are required for the activity of episomal *INS* reporter constructs in cultured cells (15, 17–19, 21, 22). However, it is not known if each of those *cis* elements is truly necessary in vivo, because such studies can only partially predict their function in the integrated chromatin environment of bona fide differentiated cells. Studies in other selected genes have addressed this by targeted deletion of transcriptional regulatory elements in mice (41). Our findings now reveal that the C1/E1, CC, and CRE3 elements are essential for *INS* gene transcription in humans. The discovery of three separate mutations that target the CC dinucleotide sequence is particularly significant. This element forms part of a canonical CCACC binding-site motif for Kruppel-like zinc-finger proteins. Two recent studies have shown that artificial mutations of this CCACC element lead to decreased *INS* promoter activity (36, 42). Although previous studies failed to identify protein complexes interacting with this region in beta-cell nuclear extracts (17, 36), a recent study showed that in vitro translated GLIS3, a zinc-finger transcription factor that is mutated in patients with neonatal diabetes and congenital hypothyroidism (43), exhibits sequence-specific binding to this region (42). Multiple DNA binding factors may

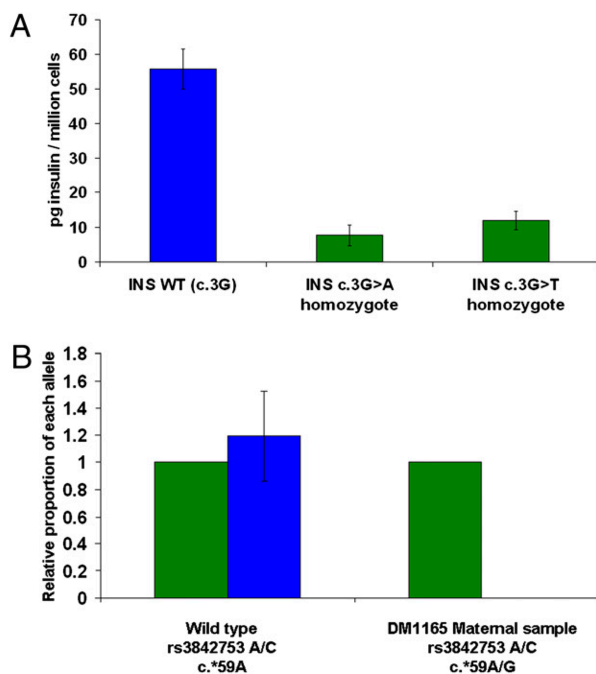

Fig. 4. Functional evidence for the pathogenicity of recessive *INS* mutations affecting translation and mRNA stability (*A*) Homozygous mutations in the translation initiation codon of the *INS* gene result in reduced insulin content of transfected HeLa cells. The insulin content of HeLa cells was measured by RIA after transfection with wild-type insulin (*INS* WT) or either of two *INS* mutant constructs, as shown. Both nucleotide changes were identified in patients with permanent neonatal diabetes. Nonspecific values obtained with HeLa cells transfected with empty vector were subtracted from all samples and those data are presented as mean +/– SE (n = 3 replicates). (*B*) Allele-specific quantitative real-time PCR of c.*59A > G and normal transcripts. The graph shows the relative abundance of the wild-type and mutant RNA transcripts in mutant and normal cell lines. The rs3842753 A allele tags the c.*59A (wild type, shown in *green*); the c.*59G (mutant) was tagged by rs3842753 C allele (*blue*). The graph shows the level of transcripts in the control sample heterozygous only for rs3842753 and in the maternal sample (family DM1165), which is heterozygous for both rs3842753 and c.*59A/G. The level of the mutant transcript is reduced to less than 3×10^{-4} percent compared with the normal transcript in the heterozygous c.*59A > G cell line. Experimental error as calculated from the standard deviation of the replicate experiments is indicated. The standard deviation for the quantification of the c.*59G allele in the maternal sample is 3×10^{-6}, and thus the experimental error is not visible in the figure.

bind to CCACC elements in vitro, and thus future studies are warranted to determine if additional factors act through this element in vivo in beta cells. Our findings, therefore, demonstrate that the natural CC element mutations that cause diabetes disrupt *INS* gene activity and establish the importance of this *cis* regulatory element.

In conclusion, we have shown that homozygous *INS* mutations are a unique cause of neonatal diabetes. The mutations result in reduced synthesis of the insulin polypeptide through a variety of mechanisms and may yield further insights into the regulation of insulin biosynthesis.

Materials and Methods

Cohort Characteristics. We studied an international cohort of 117 unrelated patients (67 males) with diabetes diagnosed before 6 months (median age 4 weeks) and without a known genetic etiology, which were referred to the Exeter (n = 105) or Bilbao laboratories (n = 12). Thirteen patients were offspring of consanguineous parents (second-degree relatives or closer). In the 100 probands with PNDM, we excluded mutations in *KCNJ11*, *ABCC8*, *GCK*, and previously described heterozygous coding mutations in *INS* (9). In the 17 patients with TNDM, we excluded 6q24 anomalies, *KCNJ11*, and *ABBC8* mutations. Remission was defined as the disappearance of clinical symptoms with normalization of blood glucose or HbA1c for a period longer than 15 days after withdrawal of insulin therapy. Babies born before 33 weeks of gestation were excluded to avoid hyperglycemia of prematurity. Studies were approved by Cruces Hospital committee and North and East Devon Research Ethics Committee. Informed consent was obtained from all patients or their parents and the studies were conducted in line with the Declaration of Helsinki. Clinical data were obtained from the patients' clinical records. We calculated standard deviation scores for birth weight (44).

Table 1. Comparison of clinical characteristics in patients with isolated neonatal diabetes with recessive and dominant *INS* mutations

Characteristic	INS recessive	INS dominant	*P*-value
n	19	46	NA
Sex, % male	63.2	47.8	0.3
Birth weight, g	1,680 (1,410; 2,050)	2,530 (2,350; 2,900)	<0.001
Gestational age, wk	37.5 (36, 40)	40 (38.5, 40)	0.008
Birth weight, SD score	−3.2 (−4.1, −2.6)	−2.0 (−2.5, −1.0)	<0.001
Age at diagnosis, wk	1 (0, 3)	10 (5, 22)	<0.001
Remission, %	26	0	0.001
Age at remission, wk	12 (11, 22)	NA	NA
Age at relapse, yr	1 (only 1 case)	NA	NA
Current age, yr	5 (2, 14)	11 (4, 23)	0.2

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

Molecular Genetic Analysis. Genomic DNA was extracted from peripheral leukocytes using standard procedures. Regulatory elements up to 450 bp upstream of the transcriptional start site and exons 1 to 3 of the *INS* gene (Fig. 1) were amplified by the PCR in three amplicons (primers and conditions available on request). Unidirectional sequencing was carried out on an ABI3730 (Applied Biosystems) and analyzed using Mutation Surveyor v3.20. Sequences were compared with the published sequence (Ensembl sequence ENSG00000129965) and published polymorphisms. The genomic reference sequence nucleotide 1 is the transcriptional start site (g.1A or c.-238A), whereas the translational start site is located at g.238 (c.1). Mutation nomenclature is shown in compliance with the Human Genome Variation Society, where nucleotide 1 represents the A of the translational start-site codon ATG (c.1). Suspected mutations were tested for conservation across species and cosegregation within families. Putative gene deletions were investigated using multiplex ligation-dependent probe amplification assay oligonucleotide probes specific for the three exons of *INS* (see *SI Materials and Methods*).

Functional Studies. ***Investigating the effect of INS promoter mutations on transcriptional activity.*** To determine the functional impact of the c.-331(C > G, C > A) and c.-332C > G mutations we performed site-directed mutagenesis of an *INS* promoter firefly luciferase reporter construct (pSOUAPRL-251hINS-Luc), and compared the activity of control and mutated promoters in MIN6 β-cells, using a Renilla luciferase promoter (pGL4.75) to correct for differences in transfection efficiency (see *SI Materials and Methods*).

Investigating the effect of the translation initiation mutations (c.3G > T and c.3G > A). To determine the effect of these mutations on insulin production, we transfected HeLa cells, which do not express insulin, with wild-type or mutant *INS* and analyzed intracellular insulin content using radio-immunoassay (*SI Materials and Methods*).

Investigating the effect of the c.*59A > G mutation on mRNA stability. We determined the effect of the c.*59A > G mutation on insulin mRNA stability using real time PCR to measure the relative levels of the *INS* mRNA transcripts in a heterozygous lymphoblastoid cell line derived from the proband's mother. We used a heterozygous SNP, rs3842753, to identify the mutation bearing allele (*SI Materials and Methods*).

Statistical Analysis. Clinical numeric data are given as median (IRQ range). Functional data are given as mean (SE). The clinical features of patients were compared using Kruskal-Wallis, χ^2 (Fisher's exact) tests or Mann-Whitney U in the statistical package SPSS version 13. Student's t-test or analysis of variance was used for expression studies.

MEDICAL SCIENCES

ACKNOWLEDGMENTS. We thank Chris Boustred, Piers Fulton, Amna Khamis, Andrew Parrish, Gustavo Pérez-Nanclares, Xavi Garcia, and Eleftheria Eiakogiannaki for technical assistance; Ann Kelly, Anthony Dixon, Srikanth Bellary, and Abigail Britten for supplying the Pakistani control DNAs; Dr Roland Stein (Vanderbilt University) for providing a human insulin promoter plasmid; and Dr Jun-ichi Miyazaki for the MIN6 cells. This work was funded by the European Union (Integrated Project EURODIA LSHM-CT-2006-518153 in the Framework Programme 6 and CEED3 RTD REG/F.2(2009)D/518731 in the Framework Programme) To A.T.H. and J.F, Wellcome Trust Grants 067463/Z/2/Z (to A.T.H.) and 081278/Z/06/Z (to L.W.H.); the Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas and PI06/0690 from the Instituto de Salud Carlos III of the Spanish Ministry of Health (to G.P.d.N.); Grant BFI06.266 from the Basque Department of Education (to I.G.); an "Ayuda para contratos post-Formación Sanitaria Especializada" from the "Instituto de Salud Carlos III" (FIS CM06/00013) (to O.R.-C.); FIS researcher Grant CP03/0064 from the Instituto de Salud Carlos III of the Spanish Ministry of Health (to G.P.d.N.); and Grants 021620819 from the Czech Ministry of Education and 64203 from the Czech Ministry of Health (to Z.S.). Work in Poland is supported by the Polish Diabetes Association. A.T.H. is a Wellcome Trust Research Leave Fellow, L.W.H. is a Research Councils United Kingdom Academic Fellow, and J.M.L. is funded by a Peninsula Medical School PhD studentship.

1. Iafusco D, et al.; Early Onset Diabetes Study Group of the Italian Society of Paediatric Endocrinology and Diabetology (2002) Permanent diabetes mellitus in the first year of life. *Diabetologia* 45:798–804.
2. Edghill EL, et al. (2006) HLA genotyping supports a nonautoimmune etiology in patients diagnosed with diabetes under the age of 6 months. *Diabetes* 55:1895–1898.
3. Aguilar-Bryan L, Bryan J (2008) Neonatal diabetes mellitus. *Endocr Rev* 29:265–291.
4. Gloyn AL, et al. (2004) Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *N Engl J Med* 350:1838–1849.
5. Proks P, et al. (2006) A heterozygous activating mutation in the sulphonylurea receptor SUR1 (ABCC8) causes neonatal diabetes. *Hum Mol Genet* 15:1793–1800.
6. Ellard S, et al. (2007) Permanent neonatal diabetes caused by dominant, recessive, or compound heterozygous SUR1 mutations with opposite functional effects. *Am J Hum Genet* 81:375–382.
7. Babenko AP, et al. (2006) Activating mutations in the *ABCC8* gene in neonatal diabetes mellitus. *N Engl J Med* 355:456–466.
8. Støy J, et al.; Neonatal Diabetes International Collaborative Group (2007) Insulin gene mutations as a cause of permanent neonatal diabetes. *Proc Natl Acad Sci USA* 104: 15040–15044.
9. Edghill EL, et al.; Neonatal Diabetes International Collaborative Group (2008) Insulin mutation screening in 1,044 patients with diabetes: mutations in the INS gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in childhood or adulthood. *Diabetes* 57:1034–1042.
10. Colombo C, et al.; Early Onset Diabetes Study Group of the Italian Society of Pediatric Endocrinology and Diabetes (SIEDP) (2008) Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *J Clin Invest* 118: 2148–2156.
11. Molven A, et al.; Norwegian Childhood Diabetes Study Group (2008) Mutations in the insulin gene can cause MODY and autoantibody-negative type 1 diabetes. *Diabetes* 57:1131–1135.
12. Polak M, et al.; French ND (Neonatal Diabetes) Study Group (2008) Heterozygous missense mutations in the insulin gene are linked to permanent diabetes appearing in the neonatal period or in early infancy: a report from the French ND (Neonatal Diabetes) Study Group. *Diabetes* 57:1115–1119.
13. Temple IK, et al. (2000) Transient neonatal diabetes: widening the understanding of the etiopathogenesis of diabetes. *Diabetes* 49:1359–1366.
14. Flanagan SE, et al. (2007) Mutations in ATP-sensitive K+ channel genes cause transient neonatal diabetes and permanent diabetes in childhood or adulthood. *Diabetes* 56: 1930–1937.
15. Karlsson O, Edlund T, Moss JB, Rutter WJ, Walker MD (1987) A mutational analysis of the insulin gene transcription control region: expression in beta cells is dependent on two related sequences within the enhancer. *Proc Natl Acad Sci USA* 84:8819–8823.
16. Naya FJ, Stellrecht CM, Tsai MJ (1995) Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Genes Dev* 9:1009–1019.
17. Boam DS, Clark AR, Docherty K (1990) Positive and negative regulation of the human insulin gene by multiple trans-acting factors. *J Biol Chem* 265:8285–8296.
18. Sharma A, Stein R (1994) Glucose-induced transcription of the insulin gene is mediated by factors required for beta-cell-type-specific expression. *Mol Cell Biol* 14: 871–879.
19. German M, et al. (1995) The insulin gene promoter. A simplified nomenclature. *Diabetes* 44:1002–1004.
20. Olbrot M, Rud J, Moss LG, Sharma A (2002) Identification of beta-cell-specific insulin gene transcription factor RIPE3b1 as mammalian MafA. *Proc Natl Acad Sci USA* 99: 6737–6742.
21. Melloul D, Marshak S, Cerasi E (2002) Regulation of insulin gene transcription. *Diabetologia* 45:309–326.
22. Hay CW, Docherty K (2006) Comparative analysis of insulin gene promoters: implications for diabetes research. *Diabetes* 55:3201–3213.
23. Permutt MA, Kipnis DM (1972) Insulin biosynthesis. I. On the mechanism of glucose stimulation. *J Biol Chem* 247:1194–1199.
24. Wicksteed B, et al. (2007) A *cis*-element in the 5′ untranslated region of the preproinsulin mRNA (ppIGE) is required for glucose regulation of proinsulin translation. *Cell Metab* 5:221–227.
25. Goodge KA, Hutton JC (2000) Translational regulation of proinsulin biosynthesis and proinsulin conversion in the pancreatic beta-cell. *Semin Cell Dev Biol* 11:235–242.
26. Sander M, Griffen SC, Huang J, German MS (1998) A novel glucose-responsive element in the human insulin gene functions uniquely in primary cultured islets. *Proc Natl Acad Sci USA* 95:11572–11577.
27. Wicksteed B, et al. (2001) Cooperativity between the preproinsulin mRNA untranslated regions is necessary for glucose-stimulated translation. *J Biol Chem* 276: 22553–22558.
28. 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. *J Biol Chem* 277:1099–1106.
29. Liu M, Hodish I, Rhodes CJ, Arvan P (2007) Proinsulin maturation, misfolding, and proteotoxicity. *Proc Natl Acad Sci USA* 104:15841–15846.
30. Herbach N, et al. (2007) Dominant-negative effects of a novel mutated Ins2 allele causes early-onset diabetes and severe beta-cell loss in Munich Ins2C95S mutant mice. *Diabetes* 56:1268–1276.
31. Duvillié B, et al. (1997) Phenotypic alterations in insulin-deficient mutant mice. *Proc Natl Acad Sci USA* 94:5137–5140.
32. Matsuoka TA, et al. (2003) Members of the large Maf transcription family regulate insulin gene transcription in islet beta cells. *Mol Cell Biol* 23:6049–6062.
33. Docherty HM, et al. (2005) Relative contribution of PDX-1, MafA and E47/beta2 to the regulation of the human insulin promoter. *Biochem J* 389:813–820.
34. Inagaki N, et al. (1992) c-Jun represses the human insulin promoter activity that depends on multiple cAMP response elements. *Proc Natl Acad Sci USA* 89:1045–1049.
35. Hay CW, et al. (2005) Glucagon-like peptide-1 stimulates human insulin promoter activity in part through cAMP-responsive elements that lie upstream and downstream of the transcription start site. *J Endocrinol* 186:353–365.
36. Niu X, et al. (2007) Human Krüppel-like factor 11 inhibits human proinsulin promoter activity in pancreatic beta cells. *Diabetologia* 50:1433–1441.
37. Wang J, et al. (1999) A mutation in the insulin 2 gene induces diabetes with severe pancreatic beta-cell dysfunction in the Mody mouse. *J Clin Invest* 103:27–37.
38. Gloyn AL, et al. (2005) Relapsing diabetes can result from moderately activating mutations in KCNJ11. *Hum Mol Genet* 14:925–934.
39. Yorifuji T, et al. (2004) Neonatal diabetes mellitus and neonatal polycystic, dysplastic kidneys: Phenotypically discordant recurrence of a mutation in the hepatocyte nuclear factor-1beta gene due to germline mosaicism. *J Clin Endocrinol Metab* 89: 2905–2908.
40. Edghill EL, et al. (2006) Hepatocyte nuclear factor-1 beta mutations cause neonatal diabetes and intrauterine growth retardation: support for a critical role of HNF-1beta in human pancreatic development. *Diabet Med* 23:1301–1306.
41. Fujitani Y, et al. (2006) Targeted deletion of a *cis*-regulatory region reveals differential gene dosage requirements for Pdx1 in foregut organ differentiation and pancreas formation. *Genes Dev* 20:253–266.
42. Kang HS, et al. (2009) Transcription factor Glis3, a novel critical player in the regulation of pancreatic beta-cell development and insulin gene expression. *Mol Cell Biol* 29:6366–6379.
43. Senée V, et al. (2006) Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism. *Nat Genet* 38:682–687.
44. Freeman JV, et al. (1995) Cross sectional stature and weight reference curves for the UK, 1990. *Arch Dis Child* 73:17–24.

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-cagcccc**A**cacccg-MGB, Probe (MT) – 6-FAM-agcccc**C**cacccg-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.-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.