

Positional Cloning of “Lisch-like”, a Candidate Modifier of Susceptibility to Type 2 Diabetes in Mice

Marija Dokmanovic-Chouinard^{1,9}, Wendy K. Chung^{1,9}, Jean-Claude Chevre^{1,9}, Elizabeth Watson^{1,9}, Jason Yonan¹, Beebe Wiegand¹, Yana Bromberg¹, Nao Wakae¹, Chris V. Wright², John Overton¹, Sujoy Ghosh³, Ganesh M. Sathe⁴, Carina E. Ammala⁵, Kathleen K. Brown⁵, Rokuro Ito¹, Charles LeDuc¹, Keely Solomon², Stuart G. Fischer¹, Rudolph L. Leibel^{1*}

1 Naomi Berrie Diabetes Center, Columbia University, New York, New York, United States of America, **2** Vanderbilt University, Nashville, Tennessee, United States of America, **3** Clinical Pharmacology and Discovery Medicine, GlaxoSmithKline, Research Triangle Park, North Carolina, United States of America, **4** Discovery Technology Group, GlaxoSmithKline Pharmaceuticals, Collegeville, Pennsylvania, United States of America, **5** Center of Excellence for Drug Discovery, GlaxoSmithKline, Research Triangle Park, North Carolina, United States of America

Abstract

In 404 *Lep^{ob/ob}* F2 progeny of a C57BL/6J (B6) x DBA/2J (DBA) intercross, we mapped a DBA-related quantitative trait locus (QTL) to distal Chr1 at 169.6 Mb, centered about D1Mit110, for diabetes-related phenotypes that included blood glucose, HbA1c, and pancreatic islet histology. The interval was refined to 1.8 Mb in a series of B6.DBA congenic/subcongenic lines also segregating for *Lep^{ob}*. The phenotypes of B6.DBA congenic mice include reduced β -cell replication rates accompanied by reduced β -cell mass, reduced insulin/glucose ratio in blood, reduced glucose tolerance, and persistent mild hypoinsulinemic hyperglycemia. Nucleotide sequence and expression analysis of 14 genes in this interval identified a predicted gene that we have designated “Lisch-like” (*Ll*) as the most likely candidate. The gene spans 62.7 kb on Chr1qH2.3, encoding a 10-exon, 646-amino acid polypeptide, homologous to *Lsr* on Chr7qB1 and to *lldr1* on Chr16qB3. The largest isoform of *Ll* is predicted to be a transmembrane molecule with an immunoglobulin-like extracellular domain and a serine/threonine-rich intracellular domain that contains a 14-3-3 binding domain. Morpholino knockdown of the zebrafish paralog of *Ll* resulted in a generalized delay in endodermal development in the gut region and dispersion of insulin-positive cells. Mice segregating for an ENU-induced null allele of *Ll* have phenotypes comparable to the B.D congenic lines. The human ortholog, *C1orf32*, is in the middle of a 30-Mb region of Chr1q23-25 that has been repeatedly associated with type 2 diabetes.

Citation: Dokmanovic-Chouinard M, Chung WK, Chevre J-C, Watson E, Yonan J, et al. (2008) Positional Cloning of “Lisch-like”, a Candidate Modifier of Susceptibility to Type 2 Diabetes in Mice. *PLoS Genet* 4(7): e1000137. doi:10.1371/journal.pgen.1000137

Editor: Gregory S. Barsh, Stanford University School of Medicine, United States of America

Received: October 16, 2007; **Accepted:** June 20, 2008; **Published:** July 25, 2008

Copyright: © 2008 Dokmanovic-Chouinard et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work supported by NIH grant RO1-DK66518 (RL), DERC 5P30DK063608, NYORC 5T32DK007599-18. The Berrie Foundation Program in Cellular Therapies of Diabetes, Albert Renold Fellowship of the European Foundation for the Study of Diabetes (EFSO) and grants from the Eli Lilly Company and GlaxoSmithKline.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rl232@columbia.edu

9 These authors contributed equally to this work

Introduction

Type 2 diabetes (T2D) afflicts ~246 million people worldwide, including ~21 million in the United States (7% of the population); another 54 million Americans are pre-diabetic. If the incidence of T2D continues to increase at the present rate, one in three Americans, and one in two minorities born in 2000 will develop diabetes in their lifetimes [1]. Direct medical costs associated with diabetes in the United States exceed \$132 billion a year [2], and consume ~10% of health care costs in industrialized nations.

Peripheral hyporesponsiveness to insulin increases metabolic demands on the insulin-producing β -cells of the pancreatic islets. Many obese individuals are insulin-resistant, but do not become overtly diabetic provided that the increased demand for insulin is effectively met [3,4]. However, if β -cell mass and/or function are insufficient to meet this requirement, overt hyperglycemia and T2D ensue [5]. In autopsy series of subjects with T2D, total β -cell mass is decreased [6,7]. Primary reductions of β -cell mass

predispose to diabetes in rodent models [8,9,10] and in autosomal dominant forms of diabetes (e.g., MODY; maturity onset diabetes of youth) [11]. Such primary reductions might predispose to some instances of T2D.

Susceptibility to T2D is strongly inherited as evidenced by the >80% concordance rates in monozygotic twins [12,13,14,15], familial aggregation, and ethnic predispositions [16]. Heritability of sub-phenotypes related to T2D, e.g. insulin resistance and β -cell hypofunction is even higher [17]. Environmental factors are also important [17,18]. Although several genes for relatively rare monogenic forms of diabetes, including MODY, syndromic (Wolfram syndrome), lipoatrophic, and mitochondrial-inherited diabetes have been identified [2,19], the underlying genetic bases for the genetically complex T2D, accounting for >95% of diabetes patients, have remained elusive. The identification of susceptibility genes is made difficult by the polygenic nature of the phenotype [20], its reflection of convergent, distinct metabolic processes producing identical phenotypes (phenocopies), and the potent

Author Summary

Type 2 diabetes (T2D) accounts for over 90% of instances of diabetes and is a leading cause of medical morbidity and mortality. Twin studies indicate a strong polygenic contribution to susceptibility within the context of obesity. Although approximately ten genes making important contributions to individual risk have been identified, it is clear that others remain to be identified. In this study, we intercrossed obese, diabetes-resistant and diabetes-prone mouse strains to implicate a genetic interval on mouse Chr1 associated with reduced β -cell numbers and elevated blood glucose. We narrowed the region using molecular genetics and computational approaches to identify a novel gene we designated “Lisch-like” (*Ll*). The orthologous human genetic interval has been repeatedly implicated in T2D. Mice with an induced mutation that reduces *Ll* expression are impaired in both β -cell development and glucose metabolism, and reduced expression of the homologous gene in zebrafish disrupts islet development. *Ll* is expressed in organs implicated in the pathophysiology of T2D (hypothalamus, islets, liver, and skeletal muscle) and is predicted to encode a transmembrane protein that could mediate cholesterol transport and/or convey signals related to cell division. Either mechanism could mediate effects on β -cell mass that would predispose to T2D.

gene-gene and gene-environment (e.g. obesity) interactions that characterize the disease. Clear genetic influences on the endophenotypes (intermediate phenotypes) of β -cell mass/function and insulin resistance have been shown, and vary among racial groups. [21,22,23,24]. Some notable earlier successes (e.g. *PPARG*, *CAPN10*), and a recent series of genome-wide association studies of large numbers of well-phenotyped subjects [25,26,27,28,29,30,31] have identified T2D susceptibility loci/genes (e.g. *TCF7L2*) whose functions with regard to the implicated phenotypes are uncertain. As no single implicated gene or allele accounts for more than a small fraction of risk to develop T2D, there are still many genes/molecular mechanisms awaiting identification.

In mice, there is striking strain-dependent susceptibility to T2D in the context of obesity [32]. We exploited the differential diabetes susceptibilities of the B6 and DBA strains segregating for the obesity mutation *Lep^{ob}* [32] to identify a diabetes susceptibility QTL in B6xDBA progeny and then used congenic lines derived from the implicated interval to clone a candidate gene accounting for the QTL. Similar strategies have been used to identify QTLs (and responsible genes) for other complex phenotypes in mice [33] such as type 1 diabetes [34], diet-induced obesity [35], tuberculosis susceptibility [36], atherosclerosis [37], epilepsy [38], schizophrenia [39] and, also, T2D [40,41,42,43].

We identified, “Lisch-like” (*Ll*), a novel gene, encoding multiple, tissue-specific transcripts in brain, liver and islets. The functional consequences of the hypomorphic DBA allele (diabetes-prone) in *Lep^{ob/ob}* mice appear to be late embryonic to early postnatal reductions in β -cell mass due to diminished rates of β -cell replication, some “catch-up” of β -cell mass by 2–3 months, followed by mild glucose intolerance at >6 months of age. These phenotypes are recapitulated in mice with an ENU-induced null allele of *Ll*.

Results

Genetic Map of Diabetes QTL and Related Congenic Lines

We identified a QTL for diabetes-related phenotypes in obese F2 and F3 progeny of an intercross between diabetes-resistant

C57BL/6J (B6) and diabetes-susceptible DBA/2J (DBA) mice segregating for *Lep^{ob}*. Phenotypes including fasting blood glucose, HbA1c and islet histology mapped with LOD >8 around D1Mit110 on distal Chr 1 at 169.6 Mb (details in Methods: Mapping T2D-related Phenotypes). By producing congenic and sub-congenic B6.DBA lines also segregating for *Lep^{ob}*, we refined the interval to 5.0 Mb between rs31968429 at 168.1 Mb and rs31547961 at 173.1 Mb where all four congenic lines overlap for DBA (Figure 1; details in Methods: B6.DBA Congenic Lines: Creation and Fine Mapping).

We further restricted the search (Figure 1) by identifying a haplotype block [44] conserved between B6 and DBA that extends 3.2 Mb from rs30708865 at 169.9 Mb to rs31547961 at 173.1 Mb. Only eleven unvalidated B6 vs. DBA single nucleotide polymorphisms (SNPs) in this interval are listed in the Mouse SNP database (www.ncbi.nlm.nih.gov/SNP/MouseSNP.cgi); however, among fragments we could amplify containing nine of these putative SNPs, we detected no sequence variants. Moreover, we found no coding sequence/expression difference between B6 and DBA among all genes and transcripts in the “conserved” interval by computation, direct sequencing, and quantitative mRNA expression analysis. Thus, it is unlikely that the variant(s) in the genetically-defined interval with peak at 169.6 Mb mediating differential diabetes susceptibility between these two strains is within the “conserved region.” We sequenced the 3 kb interval between rs31968429 and rs33860076 at the centromeric end of subcongenic line ljcdt and detected no variants between the two strains. Therefore, we focused our efforts on the 1.8 Mb B6 vs. DBA “variable” interval, between rs33860076 at 168.1 Mb and rs30708865 at 169.9 Mb.

Metabolic and Anatomic Phenotypes of Congenic Lines

The congenic/sub-congenic lines shown in Figure 1 displayed phenotypes of hypoinsulinemic hyperglycemia in association with relative reductions in β -cell mass due to reduced β -cell proliferation (see Islet Morphology and β -cell Replication and Apoptosis). Phenotypes were generally more salient in male animals. Genotype in the congenic interval (B6 or DBA) *per se* did not affect their body weight or composition. Supporting experiments are described below.

By 4 weeks of age, fasting plasma glucose was elevated in *Lep^{ob/ob}* males who were D/D (DBA/DBA) for the congenic interval ljcd and fed standard (9% fat) chow; glucose concentrations were higher up to 120 days. After 120 days, there were no significant differences in fasting glucose between D/D (DBA/DBA) and B/B (B6/B6) mice (Figure 2A). The decline in pre-prandial blood glucose levels in *Lep^{ob/ob}* males between 90 and 200 days is probably attributable to a slight expansion of β -cell mass in response to transient insulin resistance occurring as a normal consequence of sexual maturation (~60 days of age) [9,45]. To examine diabetes susceptibility in D/D animals that were obese independent of leptin deficiency, we fed lean (*Lep^{+/+}*) ljcd males a high-fat diet (60% kcal from fat) for 13 weeks, starting at 7 weeks of age. These mice became more hyperglycemic than B/B mice (Figure 2B), showing a persistence of this difference – similar to the animals in 2A – up to age ~140 days when the study ended.

To delineate differences in acute glucose handling in D/D vs. B/B animals, we used intraperitoneal glucose tolerance testing (ipGTT). At 60 days (Figure 2C), and even up to 200 days, when the study ended (Figure 2D), *Lep^{ob/ob}* ljcdt males were less glucose tolerant than B/B. The relative reduction in glucose tolerance in D/D vs. B/B animals that are not overtly diabetic is likely related to reduction in the number of islets. The occurrence of the diabetes-related phenotype is independent of *Lep^{ob}*, since 100-day

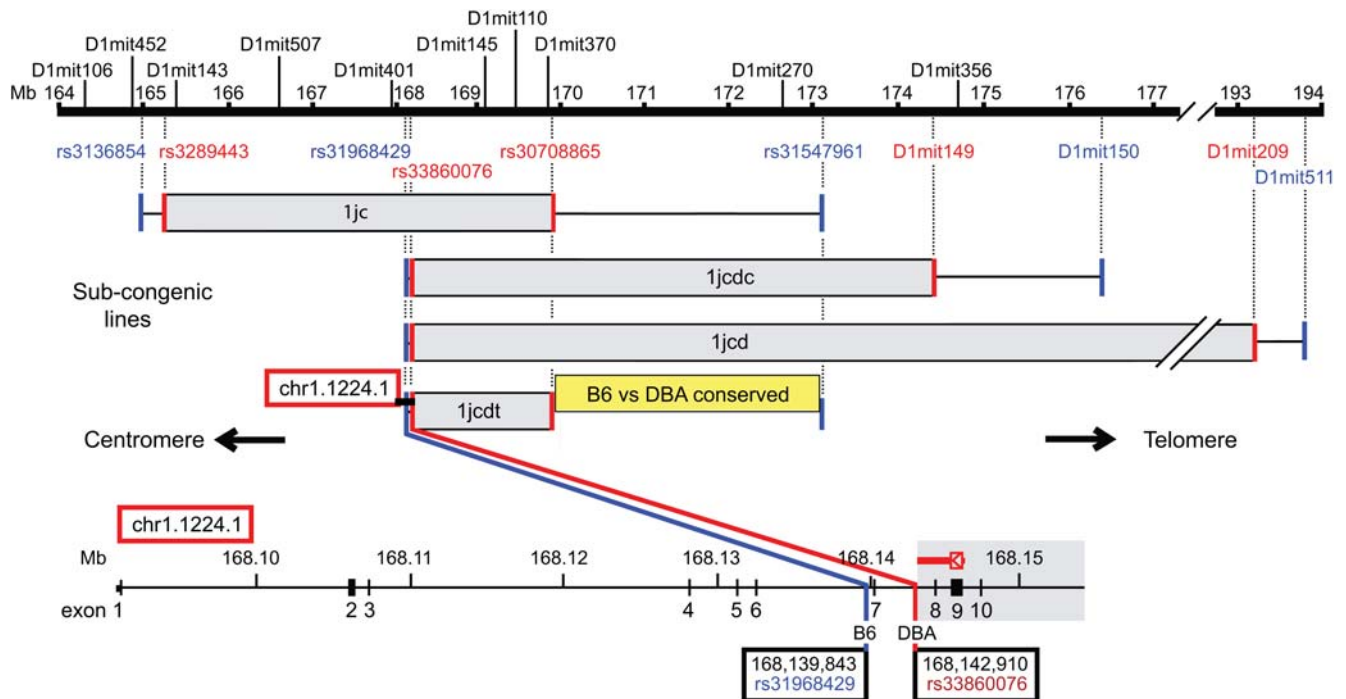


Figure 1. Genetic Map of Sub-Congenic Lines for Diabetes-Related Phenotypes in the Interval Chr 1 164–194 Mb. Genetic map shows sub-congenic lines (1jc, 1jcdc, 1jcd, 1jcdt) in the interval Chr1:164–194 Mb that display hypoinsulinemic hyperglycemia in association with histological evidence of a relative reduction in β -cell mass in the first 21–28 days of life due to reduced β -cell proliferation. An expanded view of the *Ll* gene (chr1.1224.1) is shown at bottom. Above the map scale, in black type, are microsatellite markers that were used to genotype B6 and DBA alleles to establish general boundaries of these congenic intervals. D1mit110 is the peak of the F2/F3 QTL linkage map (see Mapping T2D-related Phenotypes in B6xDBA F2/F3 Progeny). Below map scale, RefSNP (rs) and D-markers in red type identify DBA sequence limits of the respective congenic lines. Markers in blue type identify the closest, confirmed non-DBA (B6) sequence. Sequences in intervals between markers in red and blue type are DBA vs. B6 invariant. Gray bars are DBA-derived sequences. Yellow box corresponds to a 3.2 Mb interval, conserved between DBA and B6. The red box identifies the N-scan predicted gene, chr1.1224.1, subsequently identified as *Lisch-like* (*Ll*), extending centromerically from line 1jcdt. In the expanded view of *Ll*, the B6 boundary (rs31968429) for lines 1jcdc, 1jcd, 1jcdt is 333 bp centromeric of exon 7; the DBA boundary, (rs33860076) is 2,700 bp telomeric of exon 7. 5330438103Rik is an anti-sense transcript described in detail in the text. Marker positions are from the mouse genome annotation (NCBI Build 36, February 2006). doi:10.1371/journal.pgen.1000137.g001

old *Lep*^{+/+} 1jc D/D males fed the Surwit (high fat, high sucrose) diet for 10 weeks were also less glucose tolerant than littermate B/B males (Figure 2E).

Hyperglycemia due to relative hypoinsulinemia, was evident in 1jc *Lep*^{ob/ob} D/D animals fed a chow diet as early as 4 weeks (Figure 3A). At mean ages of 30- and 62-days, age-adjusted plasma insulin concentrations per mg blood glucose were lower in D/D than in B/B animals. This difference was due to lower plasma insulin in D/D ($p = 0.0004$) and not higher blood glucose in D/D ($p = 0.916$). Consistent with these ratios, D/D *Lep*^{+/+} males showed a 40% decrease in insulin secretion when clamped at a blood glucose level of 250 mg/dl for an hour (Figure 3B). No difference in insulin sensitivity was detected by euglycemic – hyperinsulinemic clamping (data not shown).

Consistent with their hypoinsulinemic hyperglycemia, 21-day old 1jcd D/D males had smaller islets than their B/B counterparts (Figure 4A). A qualitative cell-autonomous β -cell defect in insulin secretion, however, is unlikely to be the primary functional defect in D/D animals, since islets isolated from 28-day old 1jcd D/D males responded to graded glucose concentrations (2.8 mM–16.8 mM) or 10 mM arginine by secreting amounts of insulin comparable to age- and sex-matched B/B littermates (Figure 4B). Also consistent with insulin/glucose ratios and hyperglycemic clamp results, isolated islets from 60-day old 1jc *Lep*^{ob/ob} males fed normal chow and 100-day old 1jc *Lep*^{+/+} on the Surwit diet

showed reduced insulin secretion at 2.8 mM and 5.6 mM [glucose] in D/D vs. B/B littermates. For reasons indicated below, the early glucose intolerance of D/D mice is probably due, in part, to a deficiency of β -cell mass.

Islet Morphology and β -cell Replication and Apoptosis

The fractional area of the pancreas accounted for by β -cells [46] in *Lep*^{ob/ob} 1jcd males was examined in 20-, 60- and 150-day old mice. By 60 days a trend to reduced β -cell area was apparent in D/D, and by 150 days of age, β -cell mass of the 1jcd D/D sub-congenics was about half that of B/B littermate controls. B/D animals had β -cell masses that were about two-thirds of B/B littermate controls (Figure 5A). These findings are consistent with *in vivo* data showing onset of elevated blood glucose (see Figure 2A) and lower circulating insulin concentrations (relative to glucose) in D/D sub-congenics at ~60 days of age (see Figure 3A), and persistence of decreased glucose tolerance at 200 days of age. The lower relative β -cell mass in D/D animals reflects fewer numbers of β -cells, rather than smaller sized β -cells. There were no differences in pancreatic weight between D/D and B/B male animals.

To assess the basis for the difference in β -cell mass by 60 days, we measured rates of β -cell replication and apoptosis. We co-stained pancreatic sections in 1jcd congenic 1- and 21-day old *Lep*^{ob/ob} male mice with antibodies to insulin and Ki67 antigen, a nuclear marker of proliferation expressed during all stages of the

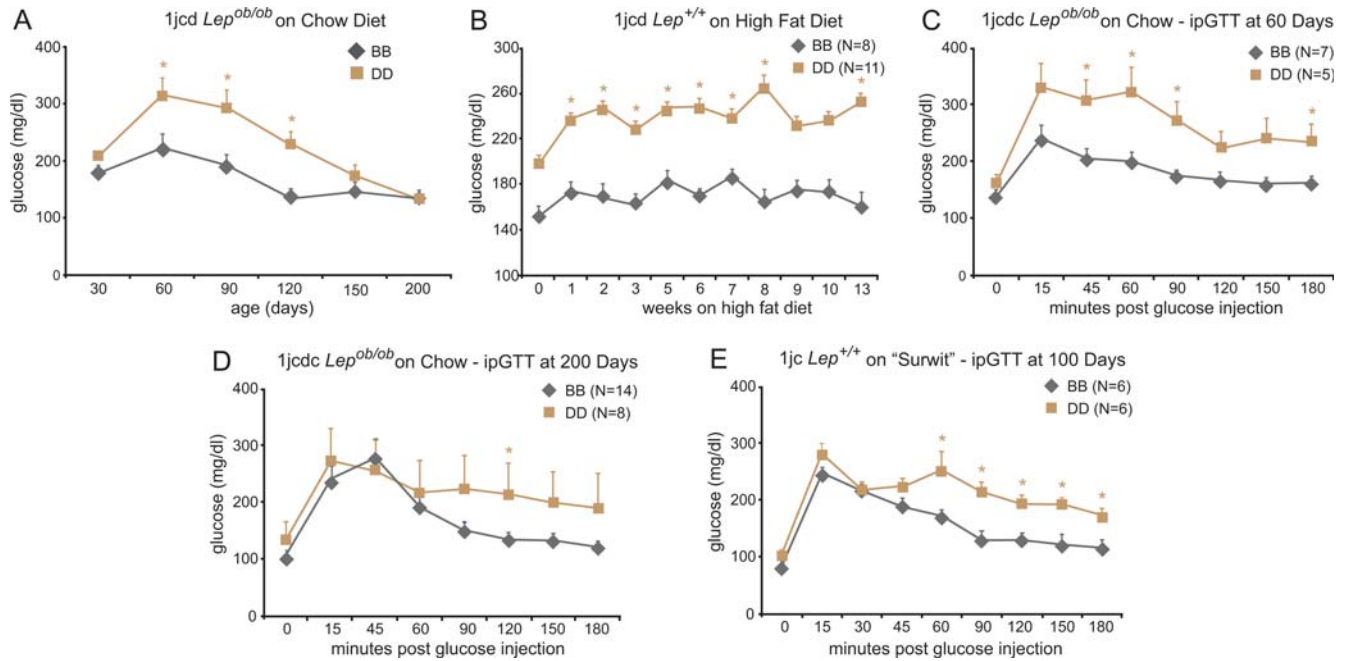


Figure 2. Fasting Blood Glucose and Glucose Tolerance Tests in Congenic Lines. A) Blood glucose in *Lep^{ob/ob}* males congenic for the interval 1jcd (BB; homozygous for B6; DD; homozygous for DBA in the interval) and fed regular mouse chow diet (9% fat) *ad libitum*. Determinations made following a 4 h morning fast. 4–13 animals per genotype group. B) Blood glucose in *Lep^{+/+}* males congenic for the interval 1jcd and fed high fat diet (60% of calories as fat) *ad libitum* for 13 wks, starting at 7 wks of age. Determinations made following a 4 h morning fast. C) ipGTT in 60-day old *Lep^{ob/ob}* males congenic for the interval 1jcd. D) ipGTT in 200-day old *Lep^{ob/ob}* males congenic for the interval 1jcd. E) ipGTT in 14-wk old male *Lep^{+/+}* males congenic for the interval 1j who had been fed the “Surwit” diet for 10 wks. In all panels, * indicates $p < 0.05$ (2 tailed t-test for genotype effect of congenic interval; mean is \pm SEM). doi:10.1371/journal.pgen.1000137.g002

cell cycle except G0 [47]. To estimate the proportion of dividing β -cells, we normalized the number of Ki67 positive β -cells to the total number of insulin positive cells. Groups consisted of 4 B/B and 4 D/D 1-day old mice or 4 B/B, and 8 D/D 21-day old mice. In 1-day old D/D males, the rate of β -cell replication was $\sim 1/3$ that of B/B littermates, whereas there was no difference in 21-day old animals due to normally reduced β -cell replication by the time of weaning (Figure 5B) [48,49,50].

The proportion of small islets ($250\text{--}2000\ \mu\text{m}^2$) in 21-day old *Lep^{ob/ob}* males was greater in D/D (1jcd and 1jcd) mice (73%) than in B/B (60%); whereas the proportion of large islets ($10,000\text{--}50,000\ \mu\text{m}^2$) was lower (9% in D/D and 14% in B/B). This finding is consistent with the β -cell replication studies in P1 mice (Figure 5B), and recently reported evidence that new β -cells are derived from replication of pre-existing β -cells [51].

In 13-day old 1jcd mice, when β -cell apoptosis is active [52], we did not detect significant differences between B/B and D/D islets in β -cell apoptosis using a TUNEL assay [53] and caspase-3 staining [54] (data not shown). Thus, the lower number of β -cells in D/D mice is primarily a result of lower rates of proliferation of β -cells in the perinatal period.

Genes in the Minimal DBA Interval Conveying Diabetes Susceptibility

To identify all genes in the minimal DBA variable interval, (see above for definition) we screened 277 genes and transcripts, computationally predicted by GenScan, TwinScan, FGeneSH, Otto, or SGP2 that map to the interval. We excluded 50 single-exon transcripts (probably pseudogenes [55]) that did not belong to a transcript cluster and were not homologous to transcripts in the syntenic human interval, and 16 ribosomal gene transcripts, unique to this interval, that could not be specifically amplified due to their genomic redundancy, and manually curated the remaining 211 predicted transcripts. We rejected 63 that did not amplify in RNA/cDNA pools from multiple organs/ages of B6 and DBA mice (see Methods: Testing for Predicted Transcripts in cDNA Pools) and, using BLASTn, clustered the remaining 148 transcripts into 14 groups. These, correspond to 11 known genes and 3 predicted genes that we validated by amplification in cDNA pools.

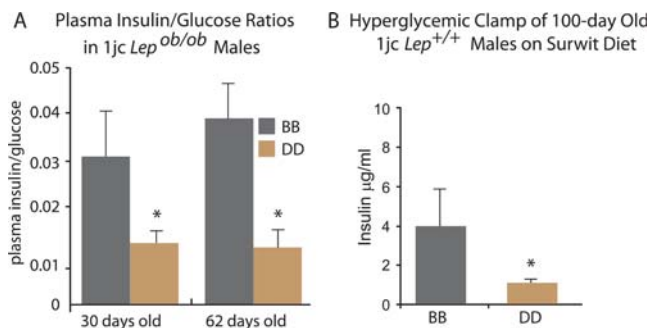


Figure 3. Plasma Insulin/Glucose Ratios and Hyperglycemic Clamps in Age-Grouped 1jcd Congenic Males. A) Fasting plasma insulin/glucose ratios in 30- and 62-day old 1jcd *Lep^{ob/ob}* B/B and D/D male mice, chow-fed since weaning. Raw data are shown in Table S1. Asterisk (*) indicates significant difference between B/B and D/D animals; p -value < 0.05 for 2-tailed t-test. B) Hyperglycemic clamping in 100-day old 1jcd males on Surwit Diet for 18 weeks. 1jcd DD male mice fed a Surwit diet for 18 wks were clamped at a blood glucose concentration of 250 mg/dl for 1 hr and serum insulin concentrations measured at 1 hr. Asterisk (*) indicates p -value < 0.05 for 2-tailed t-test. doi:10.1371/journal.pgen.1000137.g003

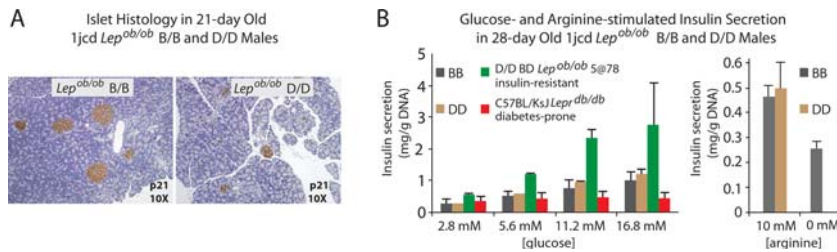


Figure 4. Relationship between Islet Histology and Insulin Secretion. A) Islet histology in 21-day old 1jcd *Lep^{ob/ob}* B/B and D/D male mice. 4 μ m pancreatic sections from 21-day old 1jcd *Lep^{ob/ob}* B/B and D/D male mice were insulin-stained with anti-guinea pig IgG and visualized by light microscopy at 10 \times magnification. In D/D animals, islets were smaller and less numerous. By histomorphometry, the proportion of small islets (250–2000 μ m²) in 21-day old *Lep^{ob/ob}* males was greater in D/D (1jc and 1jcd) mice (73%) than in B/B (60%); whereas the proportion of large islets (10,000–50,000 μ m²) was lower (9% in D/D and 14% in B/B). B) *In vitro* glucose-stimulated insulin secretion in pancreatic islets in 28-day old 1jcd *Lep^{ob/ob}* B/B and D/D males. Each congenic genotype group consisted of 3 male animals. Negative control consisted of 3 4-week old diabetes-prone *Lep^{db/db}* KsJ male animals that are hypo-responsive to glucose stimulation [9]; positive control was 3 4-week old insulin-resistant animals segregating for a diabetes-susceptibility QTL on Chr5 at 78cM, characterized by hyperglycemia and hyperinsulinemia. B/B and D/D show dose response, but no B/B vs. D/D difference at any concentration of glucose. Response to 10 mM arginine in the same animals confirms that the β -cells of the B/B and D/D congenics are comparable with regard to insulin release to a non-glucose stimulus. The 0 mM arginine control in B/B is shown to establish baseline insulin levels.
doi:10.1371/journal.pgen.1000137.g004

A map of the “variable” interval shows 14 genes, flanked by *Mael* and *Pbx1* (Figure 6). We analyzed all transcripts in the entire “variable” region.

Analysis of Genes in the Variable Interval

The genetic variation accounting for differential diabetes-susceptibility in mice segregating B/B vs. D/D in the congenic intervals could be due to: 1) coding sequence variant(s) that alter the amino acid sequence of a protein(s); 2) regulatory variants, including anti-sense transcripts that affect expression and stability, and 3' untranslated region (UTR) variants; or 3) splicing variants. We investigated all hypotheses.

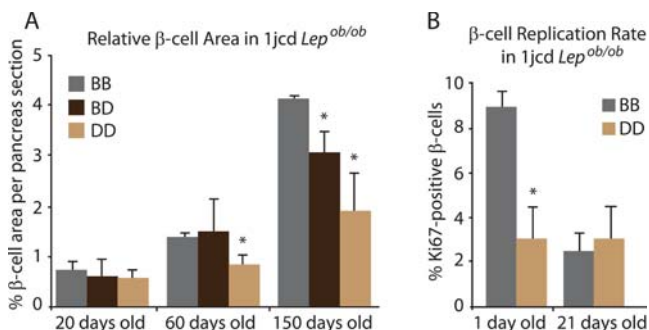


Figure 5. β -Cell Mass and Replication Rates in 1jcd *Lep^{ob/ob}* Males. A) Relative β -cell area in 20-, 60-, and 150-day old *Lep^{ob/ob}* males segregating for B and D 1jcd congenic intervals. In 60 and 150-day old males segregating for the D/D sub-congenic interval, relative β -cell mass was approximately half that of B/B littermates; B/D animals were intermediate at 150 days. N = 10 for each of the 3 groups of animals. Mean \pm SEM. The asterisk (*) indicates that $p < 0.05$ for D/D vs. B/B at 60 days, and D/D and B/D vs. B/B at 150 days. These findings are consistent with *in vivo* data (see Figure 2C) showing onset of elevated blood glucose at rest and during ipGTT by 60 days. B) β -cell replication rates (Ki67) in 1- and 21-day old *Lep^{ob/ob}* B/B and D/D 1jcd males. To estimate the proportion of dividing cells, the number of Ki67-positive β -cells was normalized to the total number of insulin-positive cells. Replication of β -cells in 1-day old D/D males was $\sim 1/3$ that of B/B littermates ($p = 0.017$). This difference, not present in 21-day old animals, was probably due to normally reduced β -cell replication by the time of weaning. Mean \pm SEM. The asterisk (*) indicates that $p < 0.05$ for D/D vs. B/B in 1-day old animals.
doi:10.1371/journal.pgen.1000137.g005

Non-Synonymous Sequence Variants

To identify all non-synonymous B6 vs. DBA sequence variants in the “variable” interval, we collected genomic sequence for B6 and DBA strains from databases at NCBI and Celera [56], filled gaps using bi-directional sequencing to achieve 100% coverage of all coding sequences in both strains, and validated coding sequence variants by bi-directionally re-sequencing gene fragments encompassing each variant in both B6 and DBA strains. Consequently, we identified five non-synonymous single nucleotide variants: one in each of three FMO-like (flavin mono-oxygenase) genes, and two in chr1.1224.1 (Figure 6). The latter gene, we designated “*Lisch-like*” (*L*) because of its sequence similarity to a gene in mouse and rat, formerly known as *Lisch7* (<http://rgd.mcw.edu/>), but now known as *Lsr* (lipolysis stimulated receptor).

Computational analysis of LL and the three FMO-like proteins using SNAP [57], PolyPhen [58], SIFT [59], PAM250 matrix substitution weights [60] and PROFac [61] predicted that all of the amino acid substitutions were benign with respect to function. The SNAP scores obtained for our variant alleles, -1 (FMO13, K282E), -2 (FMO12, V239I), -3 (LL, A647V), and -6 (LL, T587A; FMO9, Q5R), indicate that there is a $\sim 60\%$, $\sim 69\%$, $\sim 79\%$, and $\sim 90\%$ respective chance of the non-synonymous variants being neutral. Similarly, PolyPhen classified all variations as “benign” and SIFT scores were well above 0.05 (neutral). PAM weights of 0 and above suggest interchangeability of the respective amino acids throughout evolution. The % differences were low, suggesting that the DBA and B6 variants are equally likely to occur in related sequences (see Methods: Computational Methods for Evaluating Effects of nsSNPs).

Expression Differences

We used Affymetrix microarrays to quantify those transcripts in the minimum congenic interval that we had validated by PCR-amplification (see Methods: Testing for Predicted Transcripts in cDNA Pools). We examined hypothalamus, islets, liver, soleus and EDL (extensor digitorum longus) skeletal muscle from DD and BB *Lep^{ob/ob}* congenic animals (see Methods: Microarray Gene Expression Analysis). These arrays did not contain elements for all of the 14 genes we confirmed in the interval: missing from the array were the 3 FMO genes. Therefore, we also used real-time qPCR, to quantify expression of each gene and confirmed transcript in tissues and organs central to diabetes (pancreatic islets, liver, skeletal muscle, adipose tissue and hypothalamus) in 90-day old male *Lep^{ob/ob}* 1jc D/D and B/B animals (see Methods:

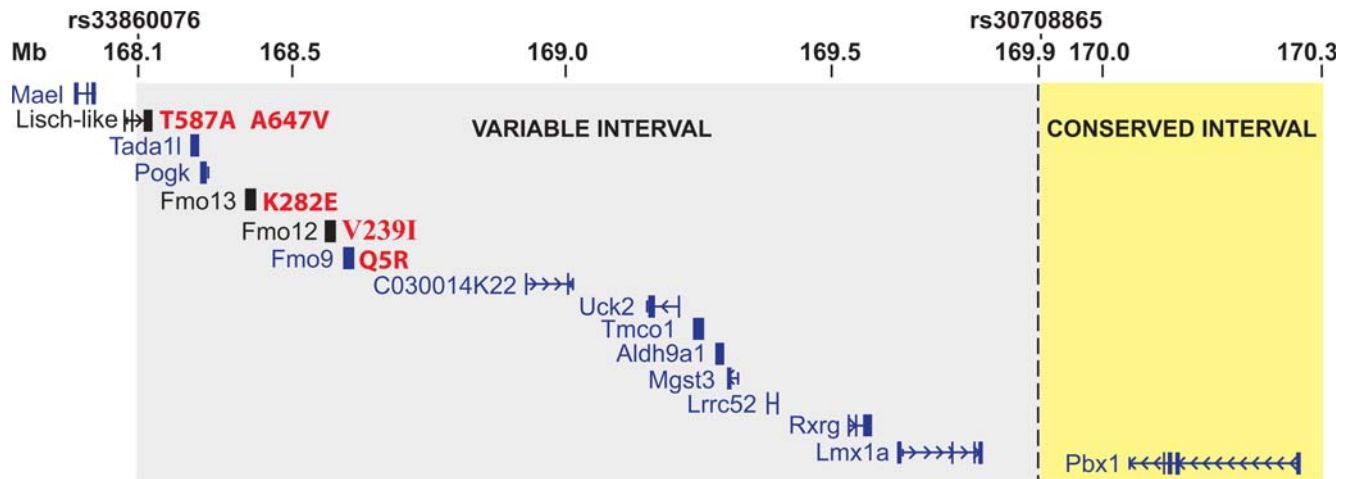


Figure 6. Genes in the Minimal Congenic Interval on Chr1:168.1–170.3 Mb. Gray background corresponds to the minimal DBA “variable” interval from 168.1 Mb–169.9 Mb, between markers rs33860076 and rs30708865. Yellow background corresponds to the centromeric end of the DBA vs. B6 “conserved” interval (i.e. nominally invariant). Genes in blue are from RefSeq; genes in black are predicted and locally confirmed as described in the text. The N-scan predicted gene chr1.1224.1 is designated here as “Lisch-like”. Amino acid variants are shown in red to the right of the corresponding gene. Nucleotide substitutions were confirmed by bidirectional sequencing in both C57BL/6J and DBA/2J DNA. doi:10.1371/journal.pgen.1000137.g006

real time qPCR). Results of the microarray and qPCR experiments are shown in Table 1 and summarized in Figure 7A.

Among genes in the region, including *Lmx1a* [62], and *Rxrg* [63], that constitute candidates for susceptibility to T2D, we identified no non-synonymous SNPs (nsSNPs) and no multi-organ differences in expression levels between B/B and D/D animals. The most prominent and consistent differences in expression we did observe were for chr1.1224.1 (*Ll*), which was two to four-fold lower in 21-day old *Lep^{ob/ob}* D/D mice than in B/B mice in the diabetes-relevant tissues/organs by microarray analysis and up to twenty-fold lower by qPCR (Figure 7A). (We later show that *Ll* protein in hypothalamus is strikingly reduced in *ljc* D/D vs. B/B; see Figure 11A). The difference in *Ll* gene expression in liver persists with age (Figure 7B) as does the difference in glucose tolerance in response to overt glucose challenge (see Figure 2D). Whether the differences in hepatic *Ll* expression are mechanistically related to differences in glucose homeostasis are unknown at this point; LL may influence hepatic gluconeogenesis, or the hepatic differences could simply mirror parallel and more physiological relevant changes in β -cells.

We also detected (by PCR) *Ll* transcripts in e7, e11, e15, and e17 whole mouse embryos, and in testis, kidney, heart, lung, uterus, eye, thymus and spleen. For the anti-sense interval between intron 9 and intron 7 (see below and Figures 1 and 8), we found higher expression levels in liver and hypothalamus of D/D v. B/B animals. This difference is consistent with a possible suppressive role for the D/D anti-sense transcript (see below). The *Aldh9a1* gene, known to be highly expressed in human embryonic brain and involved in glycolysis and fatty acid metabolism, showed qualitative changes comparable to those seen in *Ll*. The mapping statistical signals related to T2D phenotypes would be expected to enrich for regions in which several genes might contribute to the phenotypes. Although *Aldh9a1* may be such a gene, we chose to focus initially on *Ll*, since it showed the most striking quantitative differences in expression between D/D and B/B animals.

“Lisch-like” (*Ll*) Gene Structure and Splice Variants

Complete Gene Sequence. To identify 3′ and 5′ UTRs flanking the isolated transcripts of Lisch-like (*Ll*), we mapped each

transcript onto the UCSC Mouse Genome Browser (<http://genome.ucsc.edu/>) and included contiguous 5′ and 3′ ESTs. The *Ll* gene spans 62,714 bp on mouse Chr. 1, from 168,090,795–168,153,508 (Figure 8). The full-length, 10-exon transcript, isoform 1 (iso1), is 8,279 nucleotides. It comprises a 301 nt 5′ non-coding sequence, a 1941 nt coding sequence (including stop codon), encoding a 646 amino acid polypeptide, and a 6,037 nt 3′ UTR. The predicted protein includes a cleavable, signal peptide (SP; exon 1), an extra-cellular domain (ECD; exons 2–4), a transmembrane domain (TMD; the amino-half of exon 5), and a large intra-cellular domain (ICD; from the cysteine-rich, carboxy-half of exon 5–exon 10). Exons 2 and 3 of the ECD are immunoglobulin-like (Ig-like) V-type domains. Exon 6 is proline-rich and the ICD is overall serine/threonine-rich.

Isoforms. We isolated complete transcripts for 7 isoforms of *Ll* by PCR amplification of cDNAs using primer-pairs flanking the first and last predicted exons (see Methods: Cloning and Sequencing of Lisch-like Isoforms). We identified 4 major isoforms shown in Figure 8 and 3 minor isoforms. Exons 5 and 6 are absent in iso5; exon 9 is absent in iso6; and exons 5–9 are absent in iso7.

5′ Upstream Interval. The 5′ upstream interval shown (Figure 8A) includes 569 nt upstream of the predicted first transcribed base of the 5′ UTR. A CpG island is predicted to overlap the 5′ UTR. By sequencing this interval in DBA BAC 95f9 (MM_DBA library, Clemson University Genomics Institute; www.genome.clemson.edu/), we discovered 8 DBA vs. B6 nucleotide variants not in the public database. Of these, only one variant, (a C to T substitution within a CpG island) is outside a repeat element.

Anti-sense Interval. An unspliced 2,845 nt anti-sense transcript (Figure 8B) of *Ll*, from adult male mouse B6 pituitary gland (5330438I03Rik; red bar in Figure 1), starts 42 bp telomeric of exon 9, crosses exons 9 and 8, and terminates in the intron between exons 7 and 8. This transcript (see Figure 7A) is expressed 2–3 fold higher in DBA vs. B6 in hypothalamus and liver. The centromeric end of the anti-sense transcript is just 506 bp from rs33860076 at the centromeric end of the region of DBA overlap among congenic lines *ljcd*, *ljcdt* and *ljcdc*. An open reading frame (ORF) encodes a predicted polypeptide of 271 amino acids,

Table 1. BB/DD Transcript Ratios of Genes in the Variable Interval 168.1–169.9 Mb.

Confirmed Genes ^a	BB/DD Transcript ratio									
	Liver		Brain		Islets		EDL	Soleus	Muscle	Adipose
	μ -array ^{b,c,d,e}	qPCR ^{f,g}	μ -array	qPCR	μ -array	qPCR	μ -array	μ -array	qPCR	qPCR
<i>chr1.1224.1</i>	2.5	20 ^h	1.7	2.0	2.9	3.0	2.6	4.1	2.9	2.7
<i>(Lisch-like)</i>	9×10^{-4}		5×10^{-7}		2×10^{-3}		4×10^{-4}	7×10^{-4}		
<i>Lisch-like antisense</i>	0.5		0.3		NE ⁱ		NE	NE		
	6×10^{-3}		3×10^{-9}							
<i>Tada11</i>	1	1.4	1.4	0.9	0.9	1.05	1.1	0.9	0.5	1.1
	NS ^j		1×10^{-8}		NS		NS	NS		
<i>Pogk</i>	0.9	1.3	0.8	0.73	1.0	1.2	1.0	0.9	0.7	1.05
	NS		1×10^{-3}		NS		NS	NS		
<i>FMO13</i>	Not on array - inclusive-only ^{k,l}									
<i>FMO12</i>	Not on array - inclusive-only ^{k,l}									
<i>FMO9</i>	Not on array - inclusive-only ^{k,l}									
<i>C030014K22</i>	NE	1.1	NE	1.9	NE	NE	NE	NE	1.4	1.1
<i>Uck2</i>	NE	1.4	0.6	0.8	1.4	1.2	1.0	1.3	0.7	0.8
			NS		2×10^{-2}		NS	NS		
<i>Tmco1</i>	0.8	1.5	0.9	0.8	0.7	1.0	0.8	0.9	0.7	1.5
	2×10^{-5}		3×10^{-4}		2×10^{-4}		2×10^{-2}	3×10^{-2}		
<i>Aldh9a1</i>	1.4	1.3	1.5	1.4	1.9	1.5	1.6	1.7	1.3	1.6
	1×10^{-6}		3×10^{-7}		4×10^{-5}		3×10^{-4}	3×10^{-2}		
<i>Mgst3</i>	0.4	1.0	0.6	0.6	1.3	0.8	0.7	0.8	0.6	0.7
	1×10^{-9}		2×10^{-8}		NS		7×10^{-6}	NS		
<i>Lrrc52</i>	NE		NE		NE		NE	NE		
<i>Rxrg</i>	NE	0.7	NE	1.6	NE	1.0	1.2	NE	0.9	1.0
							NS			
<i>Lmx1a</i>	NE	nd	NE	2.0	NE	0.9	NE	NE	1.0	1.7

^aBoldface indicates RefSeq gene; regular type is locally-confirmed, predicted transcript.

^bExpression profiling was carried out using Affymetrix #430A mouse gene chips.

^cTen 21-day old DD and BB *Lep^{ob/ob}* 1jc congenic males were analyzed.

^dRatio is average B/B signal divided by average D/D signal in the organ.

^eLower line in μ -array box is p-value, 2-sided t-test, comparing each set of 10 BB and 10 DD mice in the same organ.

^fSamples for qPCR prepared from 5 BB and 5 DD 90-day old *Lep^{ob/ob}* 1jc males on 2 occasions.

^gRatios represent relative numbers of transcripts (BB/DD).

^hPrimer-pairs amplify transcripts of *Ll* isoforms 1, 2, 4, and 5; these comprise >90% of *Ll* transcripts.

ⁱExpression not detected.

^jNot significant ($p > 0.05$).

^kInclusive-only transcripts were detected in a cDNA pool that included whole embryos, 1-day old pups, and other tissues, but not in the cDNA pool prepared from diabetes-relevant organs.

^lProbes for these genes were neither on the Affymetrix #430A nor analyzed by qPCR.

doi:10.1371/journal.pgen.1000137.t001

but with no identifiable domain, and homologous only to ORF segments in anti-sense strands of *Ll* in other species. The interval contains 45 DBA vs. B6 variants, five of which, underlying exon 9, are listed in dbSNP. One newly discovered variant in the intron preceding exon 8, is an insertion in DBA of a 37 nt unique sequence that is homologous to a sequence in an intron of the mouse otoancorin gene on chromosome 7 and to an intronic sequence of an N-scan predicted gene on chromosome 11.

3' UTR. Of 52 B/D sequence variants in the long (6 kb) 3' UTR of the *Ll* transcript (Figure 8C), 20 were newly discovered by our “in-house” sequencing.

Cross-Species Comparisons of *Ll* Sequence

From the Ensembl database, we identified zebra fish orthologs of *Ll* and *Lsr*. The clustalW pair-wise similarity scores for the

predicted protein coded for by the zebra fish gene *zgc:114089* (*Lsr* ortholog) is 42 vs. the mouse LSR protein, and 29 vs. the mouse LL protein. The similarity scores for the predicted protein coded for by the zebra fish gene *zgc:110016* (*Lisch-like* ortholog) are 36 vs. LL and 28 vs. LSR. We performed clustalW analysis (Figure 9) between the mouse LL-iso1 protein and three related proteins: 1) the human Clorf32 protein at 1q24.1 (chr.1 165,154,620–165,211,185; NCBI Build 36.1), which is the product of a gene highly expressed in the developing human retina and brain [64]; 2) the predicted protein sequence for the zebra fish *Lisch-like* ortholog, *zgc:110016* located on zebra fish chromosome 9 at 31.6 Mb; and 3) the mouse *LSR* protein, transcribed from a gene on chromosome 7 at 30.7 Mb. Pair-wise similarity scores for the intact proteins and major domains are shown in the legend. The human homolog is similar throughout, but diverges slightly in the

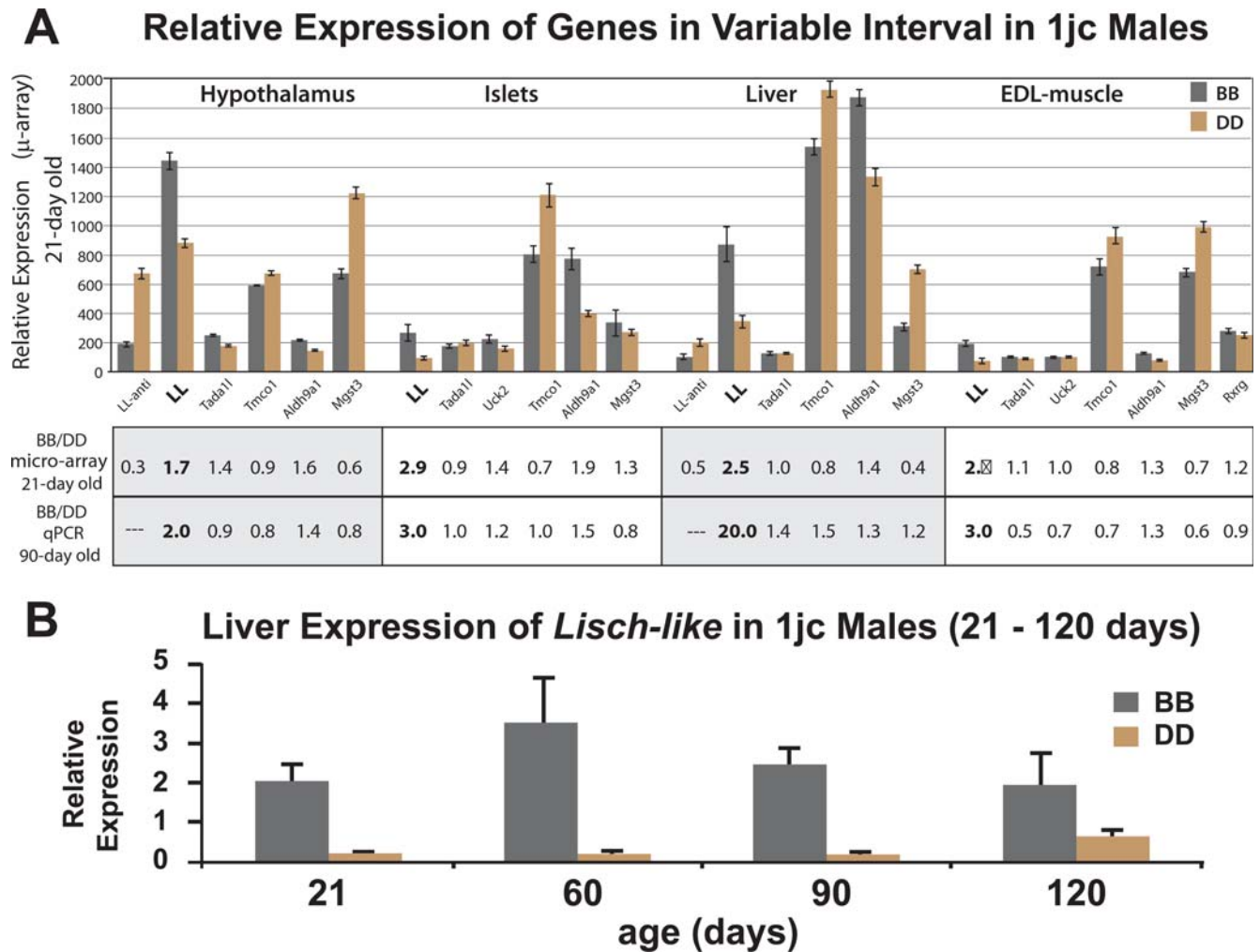


Figure 7. Expression Analysis of Candidate Genes and Liver Expression of *Lisch-like*. A) Tissue-specific expression analysis of genes in the “variable” portion of the minimum congenic interval. Data for relative expression (B/B to D/D) from Table 1 for hypothalamus, islets, liver and EDL-muscle are displayed graphically and numerically below the graph. 21-day old DD and BB *Lep^{ob/ob}* 1jc congenic males were analyzed using Affymetrix #430A microarrays. B) Liver expression of *Lisch-like* in 1jc B/B and D/D males from 21–120 days. Samples from *Lep^{ob/ob}* 1jc males were analyzed by qPCR.

doi:10.1371/journal.pgen.1000137.g007

putative ICD. The zebra fish *Lisch-like* ortholog and mouse *LSR* proteins are most alike in the TMD, less so in the Ig-like domain, and most dissimilar in the ICD. The *Lsr* protein has a short extension to exon 6, and no exon 8 equivalent. *Ll* and *Lsr* also have splicing patterns similar to the mouse *Ildr1* (Ig-like domain receptor 1) gene [65], and the proteins they encode all belong to the *Lisch7* family (IPR008664; www.ebi.ac.uk/interpro).

Knockdown of *Ll* and *Lsr* Orthologs in Zebra Fish

To assess the function of *Ll* in islet/ β -cell ontogenesis, we examined expression patterns and the effects of morpholino-mediated knockdown in zebra fish embryos. Morpholinos are modified anti-sense oligonucleotides that produce a strong hypomorphic “knockdown” phenotype [66] either by inhibiting proper splicing of the pre-RNA transcript [66] or by ATG-blocking of translation [67]. Morpholino knockdown has been used to demonstrate a role for the endocrine hormones GnRH, GHRH and PACAP during development [68,69,70,71]. Many of the molecular mechanisms regulating pancreas development appear to be conserved among zebra fish and other vertebrates

[72], and the single zebra fish islet provides an excellent model of vertebrate development.

Using whole mount *in situ* hybridization (Figure 10A), we observed that the *Lisch-like* ortholog zgc:110016 was expressed in the brain and otocyst by 48 hours post fertilization (hpf), and by 72 hpf expression was evident in the intestine. The *Lsr* ortholog zgc:114089, located on Chr 15 at 39.0 Mb, was expressed in pancreas at 48 and 72 hpf, (similar to our postnatal observations in mouse with *Ll*), intestine, liver, pharynx, pronephros and otocyst for 48 hpf (72 hpf not shown), and, at 34 hpf, in both pancreatic buds. Since the anterior bud gives rise to exocrine tissue, pancreatic duct, and a small number of endocrine cells, while the posterior bud gives rise only to endocrine tissue [69], expression of the *Lsr-like* paralog throughout this stage is consistent with a role in the ontogeny of pancreatic endocrine tissue.

The close structural similarities among *Lisch*-related genes (see Figure 9) suggested that functional data on both zebra fish genes could be physiologically relevant and, therefore, we studied the involvement in islet development of both orthologs. We injected (in separate experiments) morpholinos for both genes into embryos

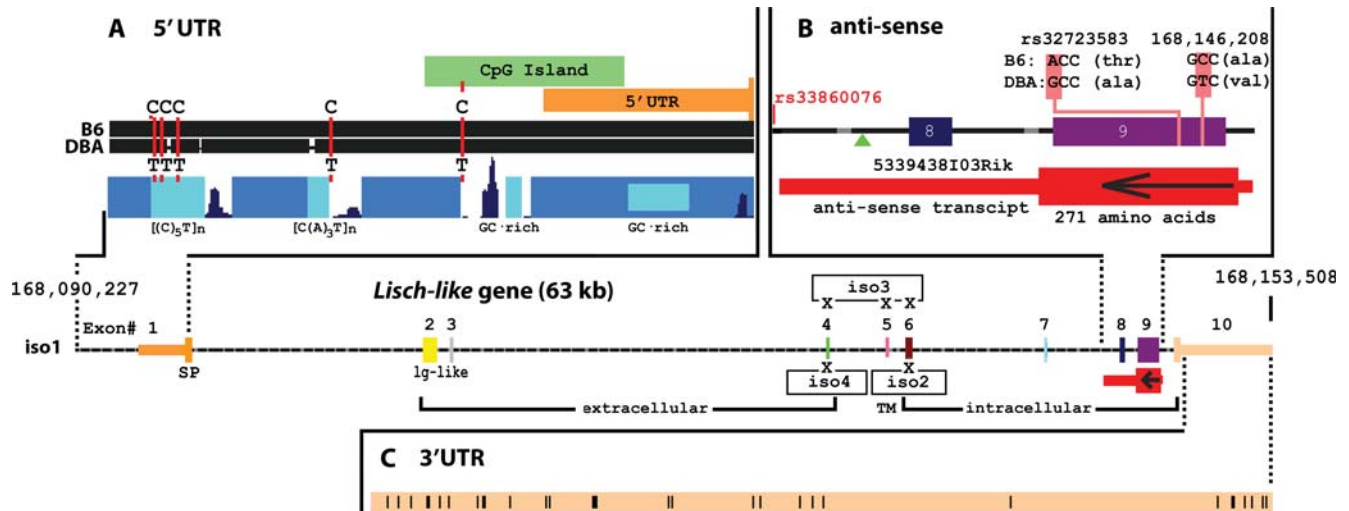


Figure 8. Predicted Structure of *L/* Gene with Expanded Views of Critical Regions. *Lisch-like* gene (middle of figure) is the full-length, 10-exon, splice variant (iso1) and includes 872 bp upstream of the transcriptional start site. Predicted domains are below exons. Exon 1 includes the 5' UTR (narrow orange bar) and cleavable signal peptide (SP). Exons 2–4 are extra-cellular, within which exons 2–3 code for an lg-like domain. Exon 5 includes the TMD with a very cysteine-rich cluster in the carboxyl half; exons 6–10 code for a serine- and proline-rich intracellular domain; exon 10 also includes a long 3' UTR. The “xs” identify exons deleted in isoforms 2–4. A. 5' upstream interval (expanded view); Black bars correspond to BLAT displays vs. the reference B6 genome. DBA variants are below the DBA bar. Annotations are composites of displays from the UCSC Genome Browser on Mouse February 2006 Assembly. Royal blue peaks correspond to sequences with predicted ESPERR regulatory potential (<http://www.bx.psu.edu/projects/esperr>). The darker blue blocks correspond to evolutionarily conserved sequences, from the UCSC “Conservation” track. Lighter blue blocks show positions of simple sequence motifs, with the consensus motif shown below. The CpG island track, provided by the UCSC Genome Browser, was generated using the unpublished *cpgh* program from Washington University (St. Louis) Genome Sequencing Center. B. Anti-sense interval corresponds to the sequences overlapping the Riken transcript 5339438I03Rik. The green triangle identifies a 37 nt unique sequence insertion in DBA. The two non-synonymous sequence variants in exon 9 sense transcript are shown. The SNP rs33860076 is the DBA marker at the centromeric end of congenic lines 1jcdt, 1jcd, and 1jcdt. C. 3' UTR interval; vertical black bars represent positions of 52 B6 vs. DBA nucleotide sequence variants. doi:10.1371/journal.pgen.1000137.g008

	EXON 1	EXON 2	EXON 3	EXON 4	EXON 5	EXON 6	EXON 7	EXON 8	EXON 9	EXON 10	
Mm_L1	-----MDRVVLGWTAVFWLTAM-VEGLQVTVPDFKKKRVAMLFQPTVLRCHFSTSSHQPA-VVQWKFPSYCDRMGESLQ---MSSPRAQALSQRNLEWDFYLDCLDSRRTV	100									
Hs_C1orf32	-----MDRVLLRWISLFWLTAM-VEGLQVTVPDFKKKRVAMLFQPTVLRCHFSTSSHQPA-VVQWKFPSYCDRMGESLQ---MSSPRAQALSQRNLEWDFYLDCLDSRRTV	100									
Dr_L1	-----MFLLFHAFWILETFLFSLQSCDGVQVVKWDEKRFAMLFSSIVLPCHYTHTSTQTA-VVQWYKSYCTDRTRDSFT---DLGASSHDCSDNSRRTV	99									
Mm_LSR	MAPAASACAGAPGSHFATTIFVCLFLIICYPRASAIQVTVPDFYHVVLFPQVTLHCTYQMSNTLTAPIVIVKYSFCRDRVADAFSPASVDNQLNAQAAAGNPGYNYVECCQDSVRTV	120									
Mm_L1			RVVASKQGSIVTLGDFYRGREITIVHDADLQIGKLMWGDGSLYYCIITPDDLEGKNEDSVELLVLRGTGLLADLLPSFAVEIMPEWVFLVILGIFLFFVLVVICWCQCCPHSCCCYV	220							
Hs_C1orf32			RVVASKQGSIVTLGDFYRGREITIVHDADLQIGKLMWGDGSLYYCIITPDDLEGKNEDSVELLVLRGTGLLADLLPSFAVEIMPEWVFLVILGIFLFFVLVVICWCQCCPHSCCCYV	220							
Dr_L1			RIVASGGQASMTLAEHYKGRDISIINKADLHIGQLWQDGSVYFCKVIIISDDLEGKNEQVELLVLRGTGLVLDLILPEFDLEIMPEWAFVGVVVVGSILFLFLVVICWCQCCPHSCCCYV	219							
Mm_LSR			RVVATKQGNNAVTLGDYYQGRRTITGNADLTFEQTAGWDSGVYYCSVVSAQDLGNNEAYAEILVLRGTSEAPELPPGFRAQPLEDWLFVVVYCLASLLFFLLGLICWCQCCPHSCCCYV	240							
Mm_L1			RCPCCPDSCCCPQALYEAGKAAKAGYPPSVSGVFGPYSIPSVPLG-GAPSSGMLMDKPHPPPLAPSDSTGGSHSVRKYRIQADKERDSMKVLVYVEKELAQDFPARRMRGRYNNITISEL	339							
Hs_C1orf32			RCPCCPDSCCCPQALYEAGKAAKAGYPPSVSGVFGPYSIPSVPLG-GAPSSGMLMDKPHPPPLAPSDSTGGSHSVRKYRIQADKERDSMKVLVYVEKELAQDFPARRMRGRYNNITISEL	339							
Dr_L1			RCPCCPDSCCCPQALYEAGKAAKAGYPPSVSGVFGPYSIPSVPLG-GAPSSGMLMDKPHPPPLAPSDSTGGSHSVRKYRIQADKERDSMKVLVYVEKELAQDFPARRMRGRYNNITISEL	299							
Mm_LSR			RCPCCPDSCCCPQALYEAGKAAKAGYPPSVSGVFGPYSIPSVPLG-GAPSSGMLMDKPHPPPLAPSDSTGGSHSVRKYRIQADKERDSMKVLVYVEKELAQDFPARRMRGRYNNITISEL	306							
									VGGHSSQVPLLREVDSVSE	366	
Mm_L1			SSLHDDSNFRQSYHQMPNKQFPMGDLSESNPDYWSGVMGSGTNGRP-ALEYNKEDRESFRHSQQRSKSEMLSRKNFATG-VPAVSMDELAAPADSYGQRARRANGNSHEARAGSR-F	456							
Hs_C1orf32			SSLHEEDSNFRQSYHQMPNKQFPMGDLSESNPDYWSGVMGSGTNGRP-ALEYNKEDRESFRHSQQRSKSEMLSRKNFATG-VPAVSMDELAAPADSYGQRARRANGNSHEARAGSR-F	457							
Dr_L1			SSLHGDVDFRQTYRQVQKALPPIIDHLDPEPLRTASIG-----HGLRPSHQSDHSLDDEHDNRWNCRSEHLPRKAFDSR-GRTVSLELEEFAVSYGPHGRRG-----DIRGPQDF	408							
Mm_LSR			-----AMSEVTSLHEDDWRSRPSRAPALTPIRDEEWRHSRPSRPTWEQEPLEQPRGG-	426							
Mm_L1			ERS-ESRAHGAFFYQDGSLEDEYYG-RGRSREPPGDGERGWTYSPAPARRRPPEDAPLRLVSRTPGTAPKYDHSYLSVSLERQARPESSSRGGSLETSPKLGALGPRASASYAWSPPPTY	574							
Hs_C1orf32			ERS-ESRAHSGFYQDGSLEDEYYGQRRSREPLTDADRGNWAFS--PARRRPAEDAHLRPLVSRTPGTAPKYDHSYLSGASRERQARPEGASRGGSLTPSKRQAQLGPRASASYAWSPPPTY	574							
Dr_L1			EMAPRTRDHPTSYRN-----GFRYLREDDSDWHRGSPSPFKRR-----DTADSERVSRQR-----SYDDTYLNSLLEKARGHGE-RGGVDDSDPTPSKGSKSS-----DCY	506							
Mm_LSR			WGSGRPRARSVDALDDINRPGSTESGRSSPPSSGRGRAYAPPRSRRDDLYLDDPDDPRDLPHSR-----DPHYDDLRSRDPRADPRSRQRSHDPRDAGFRSRDPOYDG-----RL	532							
Mm_L1										KAGASEGEDEDDAAEDALPPYSELELRGELSR-----GPSYGRDLFSHNS-----EKRRKKEPAKPGDFPFRMSLVV	646
Hs_C1orf32										KAGSSQDDQED--ASDDALPPYSELELRG-----PSYGRDLFSHNS-----EKRRKKEPAKPGDFPFRMSLVV	639
Dr_L1										QRSRPSNRPPE-----EDLPPYSEREGEERFTEPTGRERYTADPAMRPFYSYRPPHGLSQTLERRERDRDKPRKLTTHLSRDLIV	590
Mm_LSR										LEELAKKKKAG-----ERRRVYREEEEEEE-----EGHYPPAPPYSET-----DSQASRERRRKKNLALSRESLVV	594

Figure 9. ClustalW Analysis of *Lisch-like* Homologs and the *LSR* Protein. ClustalW analysis was performed on the EMBL-EBI server (www.ebi.ac.uk/clustalw/) using their default settings. We modified the display to emphasize exonic alignments. Positions of the two non-synonymous variants in exon 9 of *L/* are identified by blue shading. The non-homologous extension of mouse *Lsr* exon 6 (pink background) is shown beneath exon 7. Mm_L1; *Mus musculus Lisch-like*; Hs_C1orf32, *Homo sapiens C1orf32*; Dr_L1, *Danio rerio* (zebra fish) *Lisch-like* ortholog; Mm_LSR, *Mus musculus LSR-like* ortholog. Pair-wise similarity scores by isoform and domain are shown in Table S2. doi:10.1371/journal.pgen.1000137.g009

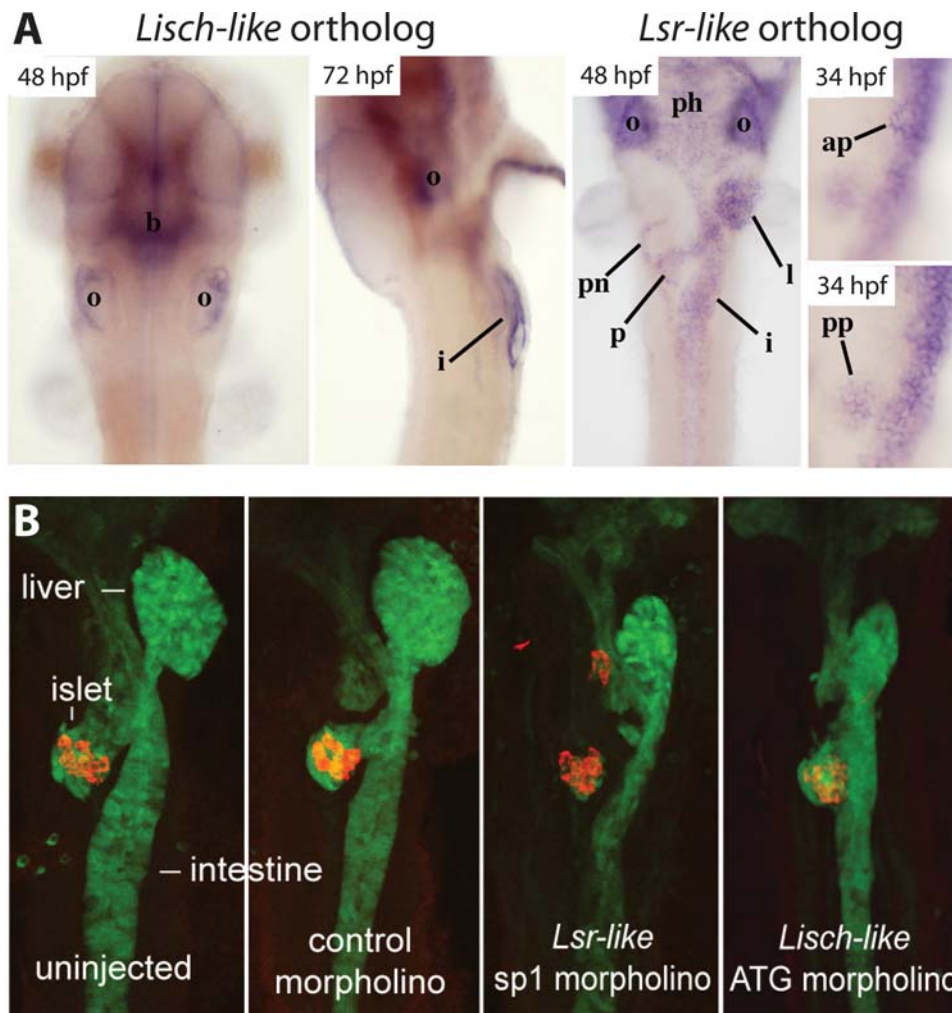


Figure 10. Expression Patterns and Morpholino Knockdown in Zebra Fish Embryos. A) Developmental expression of zebra fish *Lisch-like* and *Lsr-like* orthologs. *Lisch-like* RNA was hybridized *in situ* to whole-mount zebra fish embryos at 48 hours post-fertilization (hpf), dorsal view with anterior towards the top; and 72 hpf, lateral view with anterior towards the top, ventral towards the right and yolk removed. *Lsr-like* RNA was hybridized at 48 hpf and 34 hpf. *Ll* panels show ventral views of embryos with yolks removed and anterior towards the top. *Lsr-like* panels show the same image captured in the focal plane of the anterior (ap) and posterior (pp) pancreatic buds, respectively. i, intestine; ph, pharynx; pn, pronephric ducts; l, liver; ap, anterior pancreatic bud; pp, posterior pancreatic bud; p, pancreas (after anterior and posterior bud fusion); b, brain; o, otic vesicle. B) Morpholino knockdown of *Lisch-like* and *Lsr-like* orthologs at 48 hpf. Two dimensional ventral views (anterior towards top) of confocal stacks of 48 hpf embryos, uninjected or injected with 15 ng morpholino: control, *Lsr-like* sp1, and *Lisch-like* ATG. Gut-GFP transgene expression (green); insulin immunolabelling (red).
doi:10.1371/journal.pgen.1000137.g010

homozygous for the gut-GFP (green fluorescent protein) transgene to visualize developing endodermal organs (Figure 10B) [73]. We assessed β -cell development with an anti-insulin antibody at 48 hpf or by insulin *in situ* hybridization at 24 hpf (not shown). To assess morpholino specificity, we analyzed the effects of two separate, non-overlapping morpholinos for each gene. Both morpholinos for each ortholog independently produced similar phenotypes, providing evidence that the effects (described below) were the result of specific gene knockdown and not due to nonspecific morpholino-related effects.

Figure 10B shows that both *Lsr-like* and *Ll* morpholinos injected at 15 ng/embryo produced general developmental delay in the endodermal organs, evidenced by a smaller liver, a smaller, straighter intestine, and a smaller pancreas that does not extend as much as in wild-type. The *Lsr-like* morpholinos disrupt β -cells more severely (note ectopic insulin-positive cells in the cephalad

region of the pancreas) than do the *Ll* morpholinos (note the milder local dispersion of insulin-positive cells); 48/72 and 25/144 embryos injected with morpholinos targeting *Lsr-like* and *Ll*, respectively, displayed a scattered β -cell phenotype. These effects were rarely observed in uninjected sibling embryos (0/25) or embryos injected with a control morpholino (1/35). Lower doses of *Lsr-like* and *Ll* morpholinos (~7–10 ng) resulted in a lower frequency of β -cell scattering and higher doses (~20–25 ng) resulted in embryonic toxicity, which is common with high doses of morpholinos. The efficacy of the splice-blocking *Lsr-like* and *Ll* morpholinos was assessed via RT-PCR and all were found to strongly and specifically inhibit proper splicing of their respective target transcripts at the 15 ng dose (not shown). In combination, the expression analyses and morpholino knockdown studies provide support for a role of *Lisch* gene family members in endodermal development, and suggest specific effects on the

embryonic β -cell. The relevance of such zebra fish studies to mammalian pancreas development has been shown earlier for *Pf1a* [74,75] and for *Pdx1* [76].

W87* Stop Mutation of *Ll* in C3HeB/FeJ Mice

To examine phenotypes of mice segregating for a null allele for *Ll*, we screened a repository of ENU-generated (N-ethyl-N-nitrosourea) mutant sperm DNAs from 18,000 C3HeB/FeJ G1 males (Ingenium; <http://www.ingenium-pharmaceuticals.com/>) for mutations in *Lisch-like* [77]. We detected a G/A substitution that encodes an amber stop mutation at threonine-87 [W87*] and also creates an EcoN1 cleavage site, which we used to genotype for the mutation. By *in vitro* fertilization, we generated W87* heterozygotes on the C3HeB/FeJ background, and bred these animals to generate progeny that were homozygous wild-type (+/+), homozygous mutant (-/-) or heterozygous (+/-) for the W87* mutation. Progeny were born at the anticipated Mendelian ratios, and the -/- animals did not appear grossly compromised.

To verify that the W87* homozygous mutant was hypomorphic for LL protein, we compared a Western blot of hypothalamic extracts prepared from C3HeB/FeJ wild-type (+/+) and mutant (-/-) mice, with a second blot of hypothalamic extracts prepared from B/B and ljc-D/D congenic mice. We probed both sets of filters with a polyclonal rabbit antibody generated to a conjugated polypeptide, corresponding to exons 7 and 8 of isoform 1, in the predicted ICD of LL. As anticipated, LL protein was greatly reduced in the brains of D/D vs B/B congenics and in the ENU-treated W87* homozygotes vs. the wild-type animals (Figure 11A).

In mice at 14 days of age we can detect reductions in β -cell replication rates that are similar to those seen in the DD congenic lines (Figure 5B). There is a >2-fold difference in the proportion of Ki67-positive β -cells in 14-day old wild-type (3.75%) vs. homozygous W87* mice (1.75%), with heterozygotes intermediate (2.5%) (Figure 11B). Plasma insulin concentrations in *Ll* W87* homozygotes are reduced by the time of sexual maturation (Figure 11C) and, consistent with this difference, at 50 days of age, homozygous W87* males show an increased glucose AUC during iPGTT (Figure 11D). A significant decrease in β -cell mass is also detected in W87* homozygotes ($1.05 \pm .117$, $n = 3$, $p = .0113$) v. +/- littermates ($2.74 \pm .364$; $n = 3$) at 150 days of age.

It is important to note that these phenotypes were detected despite the segregation of the mutation on a different background strain (C3HeB/FeJ) than our congenics (C57BL/6J), and in the absence of co-segregation of the *Lept^{ob}*. These preliminary data strongly support the candidacy of *Ll* as the gene accounting for the diabetes-related phenotypes of the DD congenic lines.

Discussion

Based upon a QTL analysis of modifiers of T2D in B6xDBA F2 *Lept^{ob/ob}* mice, we identified a novel gene, Lisch-like (*Ll*), whose apparent effect on β -cell development, and possibly other aspects of β -cell/islet biology, qualify it as a strong candidate mediator of susceptibility to T2D. On the C57BL/6J strain background, the presence of the DBA/2J congenic interval(s) produced mild hypoinsulinemic hyperglycemia (in association with reduced β -cell replication and mass). Our preliminary data in ENU-mutagenized mice with a null *Ll* allele are consistent with a role for LL in β -cell development.

Three of the *Ll* subcongenic lines (ljcd, ljcdt and ljcdc) contain only DBA DNA 3' of exon 7, while line ljc is DBA for the entire gene and extends DBA for another 3 Mb 5' of *Ll*. We infer, therefore, that coding and/or non-coding DBA vs. B6 variant(s) in the region of DBA overlap accounts for the phenotypic differences

between the DBA congenic lines and animals segregating for B6 alleles in this region. In the region of overlap that includes the DBA vs. B6 "variable region" (Figure 6), *Ll* is the only gene showing anticipated differences in coding sequence and gene expression. These findings strongly support, but do not prove, the putative role of *Ll* alleles in conveying the phenotypic differences seen between the various DD and BB congenic lines. The phenotypes of the *Ll* W87* C3H mice also support our inferences regarding the candidacy of *Ll* based upon the B.D congenics.

There are two non-synonymous SNPs in *Ll* within the region of overlap among the congenic lines, in exon 9. However, their effects on protein function are predicted to be minor and it is unlikely that they determine the differences in either transcript abundance or protein level seen in the congenics. Variants in other regions of the gene are likely more relevant.

In the 5' UTR, all but one of the eight variants are in simple repeats, where they are likely less significant. The interval underlying the anti-sense transcript contains 45 D/B variants, including a long, unique insertion. A regulatory role for the *Ll* anti-sense transcript is suggested by the similar location of anti-sense transcripts at the 3' ends of the human *C10rf32* (human ortholog of *Ll*) gene (e.g., DA322725 from hippocampus), the human *LSR* gene (DA320945, also from hippocampus), the human *ILDR1* gene (AW851103), and the mouse *Lsr* gene (BY747866). Moreover, comparative interspecies transcriptomic analysis has identified the 3' regions of transcripts as important in anti-sense regulation, and conserved overlap between species may be evidence of function [78]. For a recent review of anti-sense regulatory mechanisms, see [79].

We identified 52 B/D variants in the 3' UTR, and it is estimated that the stability of 35% of yeast transcripts are regulated by motifs in the 3' UTR [80]. Regulatory motifs, at a similar density, have been identified in the 3' UTRs of several mammals, including mice [81]. A 3' UTR polymorphism between two putative mRNA destabilizing motifs in *PPP1R3* (muscle-specific glycogen-targeting regulatory PP1 subunit) has been genetically [82] and functionally [83] related to T2D. Variants in the 3' UTR may also affect regulation by microRNAs (miRNAs). The 3' UTR is the target of mammalian microRNAs (miRNAs) [84] and their relevance to diabetes is underscored by the finding that mouse islet-specific miR-375 affects insulin secretion [85].

The physiological role of *Ll* is unknown. Based upon the effects of D alleles of *Ll* on β -cell proliferation rates, β -cell mass, *in vivo* insulin release and glucose tolerance, (Figure 5) it is likely that *Ll* influences early β -cell differentiation/turnover in a manner that predisposes obese animals to later failure of β -cells by effects on mass and possibly function [86,87]. The fact that these phenotypes are substantially recapitulated in W87* *Ll* C3H mice supports this inference.

In the neonatal rodent, extensive remodeling of β -cells occurs as a result of simultaneous activation of both apoptosis and β -cell replication [49]. Between 4 and 24 weeks, postnatally, β -cell mass is estimated to increase 10 fold, related in part to increased body mass [49]. Compensation for β -cell stress/loss in adult rodents is primarily by β -cell hypertrophy and β -cell proliferation [51]. In rats, β -cell proliferation rates decline from ~20% per day in pups, to ~10% per day at 6–8 weeks, and to ~2% shortly thereafter [88]. However, even this low rate of turnover apparently does not persist in adulthood. Using continuous long term BrdU labeling in B6x129Sv and BALB/C one year-old mice, replacement rates as low as ~1/1400 mature β -cells/day have been reported [89]. Consistent with this finding, pancreas mass in the mouse was recently shown to be irreversibly constrained by the size of a progenitor pool in the embryonic pancreatic bud [87]. These data suggest that β -cell mass established in the first 6–8 weeks of life may be critical to the ability to meet

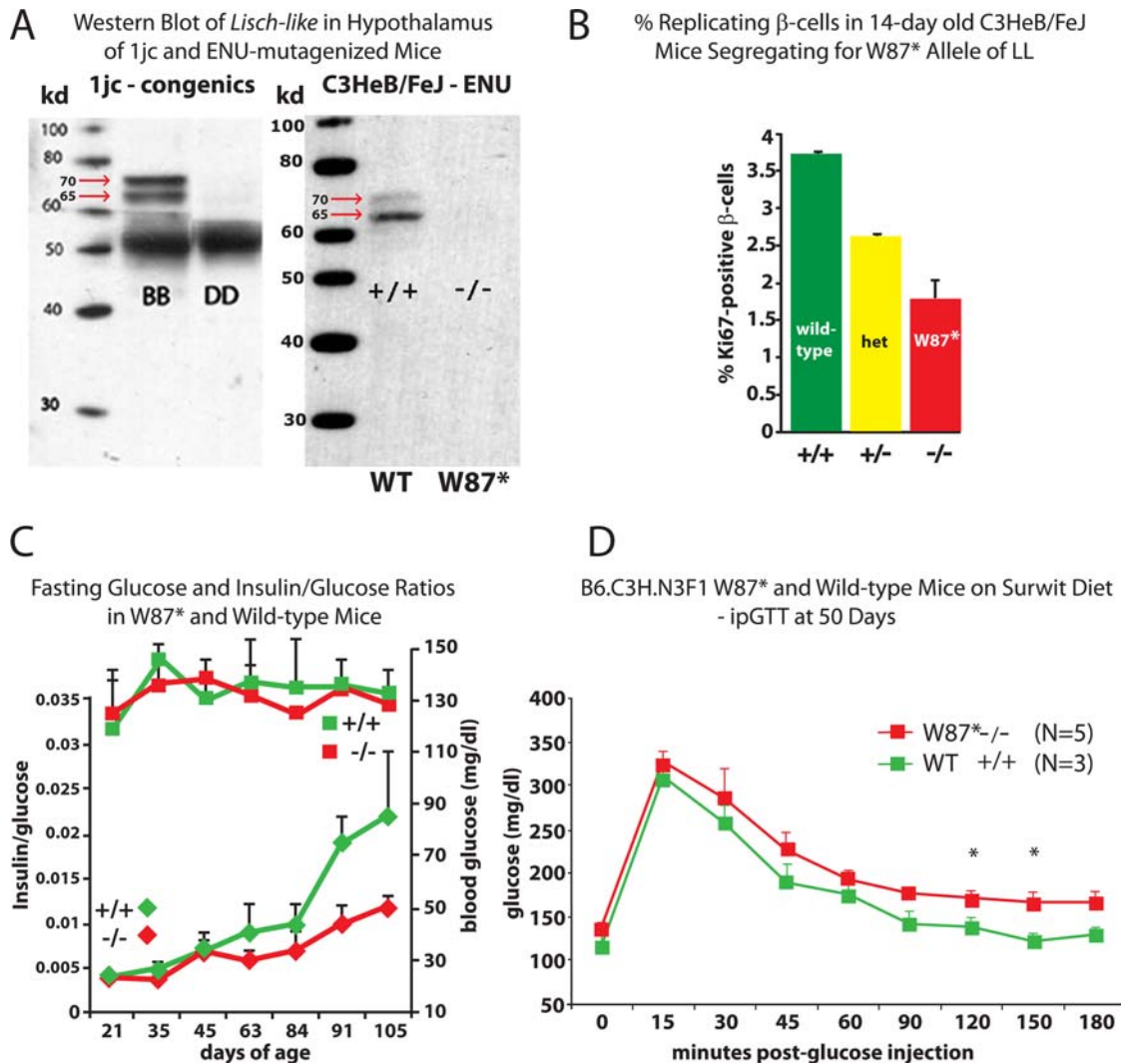


Figure 11. Phenotypes of Mice Segregating for the W87* Allele of *Lisch-like*. A) Western analysis of *Lisch-like* in hypothalamus of 1jc and homozygous W87* mice. The Western immunoblot shows differences in *Ll* expression in hypothalami of 1jc *Lep^{ob/ob}* B/B vs. D/D congenic males (left panel), and between wild-type C3HeB/FeJ and W87* C3HeB/FeJ males (right panel). The right panel immunoblot was incubated with rabbit anti-LL antiserum, prepared against a polypeptide corresponding to exons 7 and 8 of the ICD. The LL transcript isomers are visible as a 65 and 70 kD doublet in the B/B and C3HeB/FeJ wild-type lanes, but absent in the lanes of the 1jc-D/D congenic and C3HeB/FeJ W87* homozygous ENU mutants. B) Percent Replicating β -cells in 14-day old ENU-mutagenized mice. The percentage of Ki67⁺ β -cells was estimated in 14-day old C3HeB/FeJ ENU-mutagenized mice, who were either homozygous wild-type (+/+), heterozygous (+/-), or homozygous for the W87* LL amber mutation (-/-). At 14 days there was a 2-fold difference in the % of Ki67⁺ β -cells in +/+ (3.75%) vs. -/- (1.75%) ENU W87* mice; +/- were intermediate (2.5%). Non-overlapping images of longitudinal pancreatic sections (200 μ m apart) were acquired and analyzed using ImageJ software version 1.37 (NIH) to count insulin-positive and Ki67⁺ cells. Pancreatic weights of +/+ and -/- were not different. C) Fasting blood glucose (squares) and insulin/glucose ratios (diamonds) in W87* (-/-) and wild-type (+/+) littermates. P-value <0.05 for 2-tailed t-test at 63 days of age. Data points at other ages show trends. D) ipGTT on 50-day old Surwit-fed B6.CH3. N3F1 W87* males. Glucose intolerance is seen in W87* mice. Mice were fasted overnight prior to dextrose injection (50% dextrose solution, 0.5 g/kg, ip). Capillary tail bleeds were performed at the specified time points to determine circulating glucose levels by glucometer (FreeStyle Flash, Abbott). Blood glucose concentrations that are marked with an asterisk are significantly different (t-test; p<0.05; mean \pm SEM). Area under curve +/+ vs. -/- (p=0.02). doi:10.1371/journal.pgen.1000137.g011

subsequent stresses on β -cell function imposed by e.g. obesity, hyperglycemia, and dyslipidemia. The molecular regulation of these processes is incompletely understood, but even transient interruptions may, based upon this formulation, result in permanent effects on cell mass, or function, or both [90]. Hypoactivity of the candidate T2D modifier gene (*Ll*) reported here could mediate such effects on establishment of initial β -cell mass, and/or later responses of cell hypertrophy/replication by β -cell-autonomous effects or in response to an exogenous ligand for this putative receptor.

Observations that expression levels of *Ll* are most strikingly affected in liver, the effects of the zebra fish knockdowns on general endodermal development, and structure/function considerations raised by the homologous LSR molecule [91], are consistent with the possibility that the mechanism(s) by which *Ll* conveys effects on cell mass/function might relate, in part, to consequences of putative effects on hepatic development/function. IGF1 [92] and hepatic growth factor [93] are examples of such β -cell “hepatokines” affecting β -cell function.

Similarities to Trans-Membrane Receptors LSR AND ILDR1

Insight into the function(s) of the mouse *Lisch-like* protein may be gained from similarities in structure, expression, and cellular location with the human paralog, *C1orf32*, and with genes encoding related trans-membrane receptors, *Ildr1* [65] and *Lsr* [91]. Splicing patterns of these genes generate isoforms, similar to those of *Ll*. Each gene's largest isoform includes an extra-cellular Ig-like domain, a single TMD, and a similar set of ICDs in related order. In one isoform of each protein, the TMD and cysteine-rich domains are absent. An evolutionary, regulatory relationship is suggested by the observation that the *Ll*-paralog and *Ildr1* are adjacent in the zebra fish genome (Zv6 assembly, UCSC Genome Browser). All three genes are abundantly expressed in the brain, liver and pancreas (and islets, where studied), and all are predicted to have 14-3-3 interacting domains (thus far experimentally verified for the human *LSR*) [94]. Although 14-3-3 interacting domains may be present on as many as 0.6% of human proteins, their occurrence on all of these *Lisch*-related proteins is notable, since among known 14-3-3-interacting proteins is phosphodiesterase-3B, which is implicated in diabetes and pancreatic β -cell physiology [95,96,97], and others, such as the *Cdc25* family members, important in regulating cell proliferation and survival [98,99].

T2D Genetics for Region of *C1orf32*: Chr1q23

The human ortholog of *Ll*, *C1orf32*, which is 90% identical to *Ll* at the amino acid level, maps to a region of Chr1q23 that has been implicated in T2D in seven ethnically diverse populations including Caucasians (Northern Europeans in Utah) [100], Amish Family Study [101,102], United Kingdom Warren 2 study [103], French families [104], and Framingham Offspring study [105], Pima Indians [106], and Chinese [96] with LOD scores as high as 4.3. The mouse congenic interval examined here is in the middle of, and physically $\sim 10\times$ smaller than, the 30 Mb human interval. Recent analysis of the broad interval ascertained in Utah identified two peaks, one of which, at D1S2762 (at 163.6 Mb), is just 12 kb telomeric to the 5' end of *C1orf32* [107]. The genes, and gene order, are generally conserved between mouse and human in the region syntenic to the congenic interval. The metabolic phenotypes documented in human subjects with T2D linked to 1q23 resemble diabetic phenotypes observed in congenic mice segregating for the DBA interval in B6.DBA congenics examined here [108], suggesting that the diabetes-susceptibility gene in congenic mice and human subjects may be the same gene, or among the genes, acting in the same genetic pathway(s). The syntenic interval in the Goto-Kakizaki (GK) rat also correlates with diabetes-susceptibility [109].

Summary

We report the molecular cloning and preliminary characterization of a candidate gene for a mouse QTL modifying T2D phenotypes in mice. The gene, *Lisch-like*, is novel in structure among diabetes susceptibility genes, and appears to modify β -cell development. Amino acid sequence analysis is consistent with the possibility that hypomorphism for this gene could affect β -cell development by a number of possible molecular mechanisms. Proof of the role of this gene in the imputed phenotypes and molecular processes awaits its further analysis in transgenic animals and cell-based systems.

Methods

Animal Husbandry

Mice were housed in a barrier facility in ventilated Plexiglas cages under pathogen-free conditions at room temperature ($22\pm 1^\circ\text{C}$) with a 12 h light/dark cycle. Mice were weaned at 21 d and given *ad libitum* access to water and 9% Kcal fat Picolab Rodent Chow 20

(Purina Mills; www.purinamills.com/). The high fat diet protocol used in some animals is described below. Columbia University's Institutional Animal Care and Use Committee (IACUC) approved all protocols. After a 4 h morning fast, mice were sacrificed by carbon dioxide asphyxiation and phenotyped for weight, naso-anal length, and glycosuria. Blood was collected by cardiac puncture and aliquoted into microfuge tubes containing an anticoagulant cocktail of 10 μl of 1 mM EDTA and 1.5 mg/ml aprotinin (Sigma A-6279). Plasma and red blood cell pellets were used to measure glucose, insulin, and glycosylated hemoglobin as previously described [110]. Tissues (skeletal muscle, pancreas/pancreatic islets, liver, brain, hypothalamus, kidney, spleen, heart, visceral fat, retroperitoneal fat) were collected and immediately frozen in liquid N_2 , and stored at -80°C for further studies. Pancreata were dissected under stereoscope, weighed, and fixed in Z-fix zinc-formalin fixative (Anatech; www.anatechtdusa.com/).

Genotyping

Liver tissue or tail tips were used for genomic DNA isolation according to standard procedures [111]. A mutation-specific assay was used to confirm that all phenotypically obese animals were *Lep^{ob}/Lep^{ob}* and all lean animals either *+/+* or heterozygous at the *Lep* locus [112]. Animals were genotyped using MapPairs Microsaellite Markers (Invitrogen; www.invitrogen.com/) as previously described [113].

Mapping T2D-related Phenotypes in B6xDBA F2 Progeny

Maps were created using MapMarkerQTL (www.broad.mit.edu/genome_software/other/qtl.html) on a dataset representing 404 obese F2 progeny of a B6xDBA cross segregating for *Lep^{ob}* at 120–150 days of age. The QTL for T2D was most significantly associated with fasting blood glucose, glycosylated hemoglobin, and islet histology in male mice to a region of Chr1, with peak statistical significance at D1Mit110 at 169.6 Mb from the centromere ($p < 10^{-8}$) (Figure 12). Other QTLs were identified on other chromosomes (for example Chr5 at 78cM), but none had as great an effect on the phenotype or demonstrated consistent effects on all aspects of the phenotype. We tested for interactions for QTLs and identified a modest interaction between the locus on Chr1 and a second locus at D4Mit286 ($p = 0.008$).

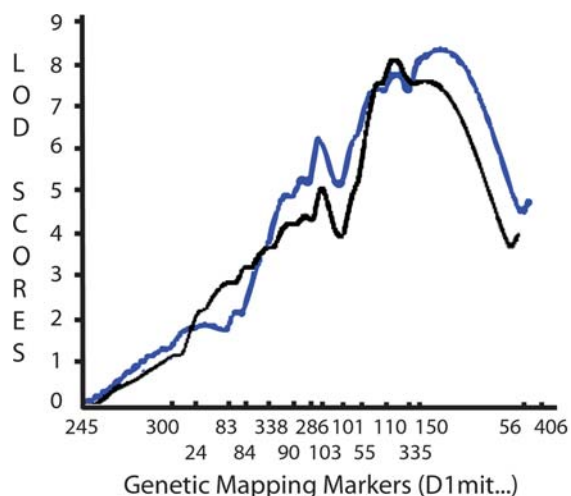


Figure 12. LOD Scores for Markers along Chromosome 1. LOD scores are shown for fasting blood glucose (black) and pancreatic grade (blue). Terminal phenotypes by genotype at D1Mit110 at 169.6 Mb are summarized in Table S5.

doi:10.1371/journal.pgen.1000137.g012

B6.DBA Congenic Lines: Creation and Fine Mapping

B6.DBA congenic mice were generated by intercrossing *Lep^{ob}/Lep⁺* B6 X DBA mice from Jackson Laboratory (www.jax.org/) to generate F1 progeny, followed by backcrossing to the recurrent B6 strain using a “speed congenic” approach in subsequent generations [114]. At the eighth backcross, a genome scan was performed in all breeders using polymorphic markers at 20 cM intervals. In the mouse line that was continued, all non-contiguous markers outside the DBA interval were homozygous B6. Over the next two generations, there were two recombination events, one that eliminated a telomeric portion of the DBA interval (line 1jc) and one that preserved approximately half of the originally defined DBA interval (line 1jcd). The 1jcd mouse was bred repeatedly to B6 mice, giving rise, by meiotic recombination, to two additional subcongenic lines (1jcdt and 1jcdc) (see Figure 1). Preservation of the phenotypes present in the original B6xDBA and DBAxB6 F2/F3 progeny was assessed by longitudinal and end-point measurements of fasting glucose, insulin, glycosylated hemoglobin and islet morphology. At N12, *Lep^{ob/+}* mice B6/DBA (B/D) for the respective congenic intervals were intercrossed to produce N12F1 progeny. Obese progeny were used for fine mapping and phenotyping experiments. *Lep^{ob/+}* animals D/D for the congenic interval were recurrently intercrossed or crossed to B6 *Lep^{ob/+}* animals to generate *ob/ob Lep^{ob}/Lep^{ob}* animals with D/D and B/D genotypes for the Chr1 interval, respectively.

Studies of Glucose Homeostasis

For longitudinal phenotyping studies, mice were fasted for 4 h and restrained for blood collection by a trained individual. Blood was collected from unanesthetized animals by capillary tail bleed into heparinized tubes and stored at -80°C . Glucose was measured with a FreeStyle Flash Blood Glucose Monitor (Abbott; www.abbottdiabetescare.com/). Insulin was measured by ultra-sensitive rat insulin ELISA (ALPCO; www.alpco.com/). HbA1c was measured by affinity chromatography (Mega Diagnostics; www.mega-dx.com/). Urine ketones were measured using Chemstrip Test Strips (Roche Diagnostics; <http://us.labsystems.roche.com/index.shtml>). For ipGTT, mice were fasted overnight and 0.5 g/kg body weight of 50% dextrose was administered intra-peritoneally at time 0. Plasma glucose was measured at 15–30 min intervals for 3 h, as above. Terminal phenotypic characterization consisted of measurements of fasting glucose, insulin, glycosuria, and glycosylated hemoglobin as previously described [110]. To control for stress-induced hyperglycemia at the time of sacrifice, tail blood glucose was also measured by glucometer one day prior to sacrifice.

High Fat and “Surwit” Diet Studies

High fat chow pellets (#D12492i; 60% kcal from fat, 20% kcal from protein, 20% kcal from carbohydrate) and “Surwit” [115] (#D12331i; 58% kcal from fat, 16.4% kcal from protein, 25.5% kcal from carbohydrate) (Research Diets; www.researchdiets.com/) were used as described in the text.

Morphometric and β -cell replication analysis of Pancreatic Islets

Pancreatic tissues were dissected under stereoscope to avoid contamination with adipose tissue, and weighed.

Islet Morphometry

Non-overlapping images of longitudinal pancreatic sections were acquired and analyzed using ImageProPlus software version 5.0 (Media Cybernetics; www.mediacy.com/) to calculate insulin-positive area, insulin-positive area as % total area, and number of

islets (defined by an area containing a minimum of 8 contiguous insulin-positive cells). For β -cell replication studies, we recorded the number of Ki67-positive or negative, insulin-positive cells. Replication of β -cells was expressed as % of cells (Ki67-positive and insulin-positive)/ total insulin-positive. For replication studies, ~ 100 islets were examined per animal from several different non-overlapping sections through the pancreas. ImageProPlus or Image J (1.37 V; NIH) were used to determine the relative area of each section occupied by β -cells or the actual number of β -cells for each representative longitudinal pancreatic section (50 μm apart) that had been immunochemically stained for insulin as previously described [116]. We analyzed 5–7 sections from different regions of the pancreas. Apoptosis rates were assessed using the DeadEnd Fluorometric TUNEL System G3250 (Promega; www.promega.com/) TUNEL assay and cleaved Caspase-3 (Asp175) Antibody 9661S (Cell Signaling Technology; www.cellsignal.com/).

Pancreatic Islet Isolation

Pancreatic perfusion and islet collection were performed as previously described [117]. Each pancreas was perfused via the bile duct with 1.5 mg/ml collagenase P (Roche Applied Science; www.roche-applied-science.com/) and incubated at 37°C for 17 min. Following disaggregation of pancreatic tissue, pancreata were rinsed with M199 medium containing 10% NCS. Islets were collected by density- gradient centrifugation in Histopaque (Sigma-Aldrich; www.sigmaaldrich.com/) [117], and washed several times with M199 medium. For glucose-stimulated insulin release studies [118,119], islets were incubated overnight in RPMI medium 1640 (Invitrogen).

Glucose-Stimulated Insulin Secretion (GSIS)

The GSIS procedure has been described previously [120]. Islets were hand-picked into tissue culture dishes containing cold Krebs's buffer (118.5 mM NaCl, 2.54 mM CaCl_2 , 1.19 mM KH_2PO_4 , 1.19 mM MgSO_4 , 10 mM HEPES, pH 7.4), and 2% BSA (Sigma-Aldrich), 5.5 mM glucose, and incubated overnight at 37°C . Islets were hand-picked and incubated another 15 min. in Krebs's buffer+BSA, containing 11.2 mM glucose. Hand-picked islets are then resuspended in Krebs's buffer plus BSA, supplemented with 2.8 mM glucose, and shaken at 37°C for 15 min. The pellet was spun down gently and resuspended in triplicate (5–10 islets each) in 500 μl Krebs's buffer, supplemented with glucose at final concentrations of 2.8 mM, 5.6 mM, 11.2 mM or 16.8 mM, or supplemented with 10 mM arginine and incubated for 1 h in a water bath at 37°C with constant shaking (300 rpm). After 1 h incubation, islets were gently pelleted and the supernatant collected and assayed for insulin by ELISA. Islet pellets were dissolved in high salt buffer (2.15 M NaCl, 0.01 M NaH_2PO_4 , 0.04 M Na_2HPO_4 , EDTA 0.672 g/L, pH 7.4) and sonicated at 4–5 W for 30 s and DNA concentration was measured using a TKO100 fluorometer (Hoefer; www.hoeferinc.com/) with Hoechst #33258 dye (Polysciences; www.polysciences.com). Results were expressed as concentration of secreted insulin/[DNA]/h.

Testing for Predicted Transcripts in cDNA Pools

Putative transcripts, identified from public annotation and local sequencing, were validated by PCR-amplification from tissue-specific cDNA pools prepared from male and female B6 mice. Two cDNA pools were used: 1. An inclusive cDNA pool was prepared from E7 and E20 fetuses and P1 pups, and included the

following tissues of 60-day old mice: eyes, large intestine, skin, tongue, spinal cord, kidney, testes/ovaries, pancreatic islets, whole brain, hypothalamus, skeletal muscle, and liver. This pool was used for transcript validation. 2. **A diabetes-relevant cDNA pool**, from 90-day old mice, was comprised of only the following tissues and organs: pancreatic islets, whole brain, hypothalamus, skeletal muscle, liver, and adipose tissue. This pool was used to quantify transcripts identified by computational approaches and the microarrays. Nominal intron-spanning primers were generated using the Primer3 program (www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Amplification was first performed on the diabetes-relevant pool at an annealing temperature of 60°C. If we detected no PCR-product, we performed gradient temperature PCR on the same pool using eight different annealing temperatures from 58–68°C. Gradient temperature PCR was then used to amplify the inclusive cDNA pool. If no product was detected in this pool, a 2nd set of intron-spanning primers was used before we interpreted negative amplification as failure to substantiate a predicted transcript. Positive amplification products of predicted sizes, and those that did not match the expected sizes, were gel-purified and sequenced for confirmation. The final set of primer-pairs is listed in Real-time qPCR.

Microarray Gene Expression Analysis

RNA extraction, purification, labeling, hybridization and analysis were performed as described [121]. 10 BB and 10 DD 21-day old *Lep^{ob/ob}* 1jc males were dissected and RNA was extracted from hypothalamus, liver, isolated islets, EDL muscle, and soleus muscle. Individually labeled RNA (by mouse and organ) was interrogated with Affymetrix MOE-430A expression arrays. For further details, see legends to Table 1 and Figure 7. For all transcripts in the region of interest, where possible, only probes that spanned multiple exons and clearly represented each of the 14 genes in the interval were used. If >1 probe met these conditions, we used only, the probe that gave the strongest signal. Organs were grouped into two groups by genotype and were compared using a two tailed T-test. The Affymetrix probe IDs selected for this analysis are shown in Table S3.

Real-Time qPCR

Effects of the DBA congenic interval on the levels of confirmed transcripts expressed in diabetes-relevant organs were assessed on an organ-specific basis. We made separate pools from 90-day old *Lep^{ob/ob}* 1jc D/D and B/B mice for each of the diabetes-relevant organs (see above). Each individual organ pool was generated on 2 occasions from 5 mice. RNA was extracted from organs with TRIzol acid-phenol reagent (Invitrogen). 2 µg of RNA were reverse-transcribed using SuperScript III reverse transcriptase (cDNA First Synthesis Kit, Invitrogen) with random hexamer priming. The cDNA was diluted 4-fold using nuclease-free water (QIAGEN; www.qiagen.com). 2 µl of diluted cDNA were amplified by PCR in Roche LightCycler. A standard curve for each transcript was generated using cDNA diluted 1:1, 1:10, and 1:100. We assessed the number of mRNA molecules in each sample using the slope and intercepts of PCR product appearance during the exponential phase of the PCR reactions optimized for transcript-specific product using specific primers. Each sample was run in triplicate in the same LightCycler run. Using LightCycler Software, we calculated the crossing point (CP) for each sample. The CP is the first maximum of the second derivative of the fluorescence curve, and is equivalent to the number of cycles at which the fluorescence first exceeds background. In the exponential phase, the relationship between CP and initial transcript

concentration is linear. We calculated relative concentration ratios, normalized to actin, as follows:

$$R = \eta_{\text{gene}}^{(\Delta\text{CP}_{\text{gene}}(\text{sample-ref}))} / \eta_{\text{actin}}^{(\Delta\text{CP}_{\text{hg}}(\text{sample-ref}))}$$

In this expression, $\Delta\text{CP}_{\text{gene}}$ is the CP of the gene in the sample minus the CP of the gene in the relevant reference; $\Delta\text{CP}_{\text{hg}}$ is the CP of the housekeeping gene in the sample minus the CP of the housekeeping gene in the reference (“ref”) sample; and η is the efficiency (where 2 is perfectly efficient) as determined by the negative slope of the plot generated when CP is plotted as a function of the log of initial concentration determined in the standard curve. Each CP listed is the mean of CP values of the triplicates for each sample. Results are summarized in Table 1. Primers used are listed in Table S4 (A).

Cloning and Sequencing of *Lisch-like* Isoforms

We amplified full-length *Ll* cDNAs from either B6 islets (isolated by us) or from Clontech MTC Panels 1 #636745 and 3 #636757, containing pooled tissue cDNAs from 8–12 week old BALB/c mice and from Swiss Webster embryos. In a final volume of 50 µl, we added 0.5 µl LA Taq (TaKaRa; www.takara-bio.com/) to a cocktail containing TaKaRa GC Buffer II, 400 µm each dNTP, 1 µl cDNA and 1 µl each primer (300 ng/µl). Primers are listed in Table S4 (B). Samples were cycled in an MJ Tetrad Thermalcycler (BioRad; www.bio-rad.com) using a Touchdown protocol of a 2 min. extension and decreasing annealing temperature from 60°C to 55°C for 10 cycles, followed by 25 cycles with an annealing temperature of 55°C. Each sample was TOPO TA cloned (Invitrogen) and plated. From all three libraries, a total of 140 colonies were picked and grown overnight in LB buffer. Inserts were amplified by colony PCR and sized by gel-fractionation. Inserts representing each unique size were then sequenced. The isoforms and the exons deleted (Δ): iso1 (intact 10 exons); iso2, Δ 6; iso3, Δ 4,5,6; iso4, Δ 4; iso5, Δ 5,6; iso6, Δ 9; iso7, Δ 5,6,7,8,9.

Zebra Fish Analyses

A. Zebra Fish Strains and Embryo Culture. Zebra fish and embryos were raised, maintained and staged according to standard procedures [122]. The AB* (Eugene, OR) line and Tg(gut GFP)_{s854} transgenic line (gutGFP; [73]) were used in natural matings to obtain embryos. The gutGFP line was provided by Didier Stainier. Embryos examined at stages later than 24 hpf were maintained in embryo medium containing 0.003% phenylthiourea to inhibit pigmentation.

B. Morpholino Injections. Morpholino antisense and control oligonucleotides, listed in Table S4 (C), were purchased from Gene Tools (www.gene-tools.com/) and injected into 1–2 cell stage embryos at concentrations from 7–20 ng/embryo as previously described [67].

C. RT-PCR. Total RNA was extracted from morpholino-injected and uninjected sibling embryos at 29 hpf with TRIzol; cDNA was synthesized with SuperScript II Reverse Transcriptase (Invitrogen) using primer-pairs shown in Table S4 (D).

D. Immunofluorescence and RNA *in situ* Hybridization. Zebra fish gene sequences were amplified using the primer-pairs shown in Table S4 (E) and cloned into the PSTBlue-1 vector (Novagen) and used for antisense probe synthesis with T7 RNA polymerase after *Xho*I linearization (*Lsr-like*) and SP6 polymerase following *Bam*HI linearization (*Lisch-like*). Whole-mount *in situ* hybridization was performed as described [123]. For

immunofluorescence, embryos were fixed at room temperature (rt) in 4% paraformaldehyde for 2 h. After fixation, yolks were manually removed and embryos were permeabilized in acetone at -20°C for 7 min. Embryos were washed briefly in PBS +0.1% Triton $\times 100$ (PBSTx) and incubated for 1 h in antibody hybridization buffer (PBSTx with 2% DMSO, 2% BSA and 2% sheep serum). Guinea pig anti-insulin antibody (Biomedica V2024) was diluted 1:1000 in antibody hybridization buffer and incubated with embryos for 2 h at rt. Following antibody hybridization, embryos were washed extensively with PBSTx and incubated with Cy3-labelled donkey anti-guinea pig secondary antibody diluted 1:500 in antibody hybridization buffer for 2 h at rt. Embryos were washed extensively with PBSTx and cleared in 80% glycerol/20% PBS. Images of optical sections were captured using a confocal microscope and 2-D projections were generated from optical sections using MetaMorph software.

Computational Methods for Evaluating Effect of nsSNPs

We used five methods to compute the likelihood of a functional change due to single amino acid substitutions (see Figure 9). SNAP, PolyPhen, and SIFT predict changes in protein function due to a single amino acid substitution. SNAP [57] is a neural-network based method that considers protein features predicted from sequence (*e.g.*, residue solvent accessibility and chain flexibility). Scores from -9 to $+9$ are estimates of accuracy of prediction, computed using a testing set of $\sim 80,000$ mutants. A low negative score indicates confidence in prediction of neutrality (functional change absent), whereas a high positive score indicates confidence in prediction of non-neutrality (functional change present). Accuracy was computed for neutrals using the equation below:

$$Accuracy_{neutral} = \frac{\text{number of correct neutral predictions}}{\text{total number of neutral predictions}}$$

PolyPhen considers structural and functional information and alignments. Predictions are sorted into 4 classes: benign, possibly damaging, probably damaging, and unknown.

SIFT predictions. SIFT [59] is a statistical method that only considers alignments. Scores range from 0 to 1. Scores >0.05 indicate neutrality of a substitution.

PAM250 matrix substitutions. PAM matrix [124] (Percent Accepted Mutations) reflects frequency of amino acid interchange throughout evolution (by evaluating alignments of proteins in a family). Scores range from a low of -8 for rare substitutions (*e.g.* W to C) to a high of 17 (same residue found in almost all proteins in alignment).

Percentage in alignment (PROFac). The score is reported as the difference in observed percentages of wild-type and mutated residues in alignments against a non-redundant UniProt [125] and PDB [126] database (at 80% sequence identity). Scores range from -100 (if the mutant is observed in all instances) to $+100$ (if the wild type is observed in all instances); 0 if the mutant is observed as often as the wild type. Scores near 0 favor the likelihood of a mutation being neutral.

DBA BAC Shotgun Sequencing

BAC 95f9 DNA (5 μg) was fragmented to 1–5 kb using a nebulizer supplied with the TOPO Shotgun Subcloning kit (Invitrogen) and checked for size and quantity on an agarose gel. The shotgun library was constructed with 2 μg of sheared DNA. Blunt-end repair, dephosphorylation, ligation into PCR 4Blunt-

TOPO vector, and transformation into TOP10 Electrocompetent *E. coli* were performed with the TOPO Shotgun Subcloning kit, following the manufacturer's protocol. Phenol:chloroform extraction of the dephosphorylated DNA was replaced with Qiagen QIAquick PCR Purification spin columns (QIAGEN). Recombinant colonies were selected by blue/white screening and incubated in LB medium supplemented with 50 $\mu\text{g}/\text{ml}$ ampicillin for 20 h at 37°C in 96-well deepwell plates. Plasmid miniprep was conducted in 96-well plates using QIAGEN Turbo Miniprep kits on a QIAGEN BioRobot 9600. DNA sequencing was performed on a 3730xl Genetic Analyzer (Applied Biosystems; www.appliedbiosystems.com/) using BigDye[®] Terminator v3.1 Cycle Sequencing Kits with M13 forward and reverse sequencing primers.

Statistical Analyses

ANOVA and ANCOVA were used to assess effects of genotype in congenic interval. Comparisons at individual time points, or pairs of means were performed using Student's t-test. P values are 2-tailed. The Statistica package (StatSoft; www.statsoft.com/) was used for ANOVAE; Excel (Microsoft, <http://office.microsoft.com/en-us/default.aspx>) for t-testing.

Western Blot

Hypothalamic extracts were prepared using M-PER Mammalian Protein Extraction Reagent (Pierce Biotechnology, www.piercenet.com/). Hypothalamic extracts (85 mg for B/B and D/D congenics and 175 mg for wild-type and mutant ENU mice) were resolved by 8% SDS-PAGE, transferred to nitrocellulose membrane (Invitrogen). We generated a set of polyclonal rabbit antibodies (Covance Research Products; www.covance.com) against the predicted ICD, spanning residues 298–401 (exons 7,8) and verified that the α -ICD rabbit antibodies detected the appropriate fusion proteins, with only minor cross-reactivity in cultured cells. We hybridized the blot with anti-LL anti-sera at a dilution of 1:5,000 in TBS/0.05% Tween/5% milk (TBSTM) or with blocked anti-LL anti-sera diluted 1:10,000 in TBSTM. To prepare blocked anti-sera, liver sections from C3HeB/FeJ knock-out mice were fixed overnight in phosphate-buffered paraformaldehyde at 4°C and rinsed in PBS. Sections equivalent to one-third of a liver were fragmented and mixed with 1 ml anti-sera diluted 1/1000 in PBS/0.1% Triton. Liver fragments were spun out and the supernatant was used to probe filters from ENU mice. We detected bound antibody with horseradish peroxidase-coupled antibody against rabbit IgG (Amersham Biosciences; www.amershambiosciences.com) at a dilution of 1:5,000 using the SuperSignal West Pico Chemiluminescent Substrate kit (Pierce Biotechnology).

Immunohistochemical and Immunofluorescence Analysis of Pancreatic Islets

For β -Cell Replication Studies. Pancreata were fixed overnight in 10% formalin, embedded the specimens in paraffin, and consecutive 5 μm -thick sections were mounted on slides. For immunofluorescence and diaminobenzidine (DAB) staining of Ki67 and for insulin immunoreactivity, tissue sections were de-waxed in xylene, hydrated through a descending ethanol series and subjected to an antigen retrieval step using a heated citrate buffer solution. Several longitudinal sections $>100 \mu\text{m}$ apart were used to assess β -cell replication and double staining for the nuclear proliferation marker Ki67 and insulin. Sections were incubated with Novocastra rabbit polyclonal anti-Ki67 antibody (Leica Microsystems; www.leica-microsystems.com) diluted 1:200 and an insulin polyclonal guinea pig anti-swine antibody (Vector Lab; www.vectorlabs.com/) diluted 1:2000 overnight at 4°C .

For Immunofluorescence Detection. Sections were washed in PBS and incubated with secondary anti-guinea pig IgG (1:200) and fluorescein isothiocyanate-conjugated rabbit secondary antibody (1:200) (Vector Labs) for 1 hr and counterstained with DAPI before the addition of mounting medium. Non-overlapping images of longitudinal pancreatic sections were acquired using a Nikon Eclipse microscope and images imported into ImageJ (1.37 V, NIH) to count insulin-positive and Ki67-insulin-positive cells. β -cell replication is expressed as % Ki67-positive+insulin-positive/total insulin-positive cells. For diaminobenzidine staining, sections were incubated with secondary biotinylated rabbit and guinea pig IgG for 1 hr and then subjected to an avidin:biotinylated enzyme complex (ABC Kit; Vector Labs) with DAB as substrate. Sections were counterstained with hematoxylin. Images of pancreatic sections were acquired using SpotAdvanced version 5 software (Diagnostic Instruments; www.diaginc.com/) and analyzed using Image Pro Plus software to calculate the % of β -cell area occupied by Ki67-positive cells. We examined 30–50 islets per animal from several non-overlapping sections through the pancreas.

Accession Numbers

Genbank (www.ncbi.nlm.nih.gov/(Genbank) accession numbers for the *M. musculus* genes: *Lisch-like*, lipolysis-stimulated remnant receptor-related (XM_001473525); *Lsr* (NM_017405); *Ildr1* (NM_134109); *Tada11* SPT3-associated factor 42 (NM_030245); *Pogk* pogo transposable element with KRAB domain (NM_175170); *FMO13*, flavin-containing monooxygenase family; FMO-like (XM_136366); *FMO9*, flavin-containing monooxygenase family; FMO-like (NM_172844) *FMO12*, flavin-containing monooxygenase family; FMO-like (XM_136368); *C030014K22Rik*, unknown (NM_175461); *Uck2*, uridine monophosphate kinase (NM_030724); *Tmco1*, membrane protein of unknown function (NM_001039483); *Aldh9a1*, aldehyde dehydrogenase 9, subfamily A1 (NM_019993); *Mgst3*, microsomal glutathione-S-transferase 3 (NM_025569); *Lrrc52*, leucine-rich repeat (LRR) protein of unknown function (NM_001033382); *Rxyg*, retinoid X receptor, gamma (NM_009107); *Lmx1a*, LIM homeobox transcription factor 1, α (NM_033652); *Pbx1* (NM_008783); *H.sapiens C1orf32* (NM_199351); *LSR* (NM_015925); *ILDR1* (NM_175924); *D. rerio Ll* ortholog, zgc:110016 (NM_001030192.1); *D. rerio Lsr* ortholog, zgc:114089 (NM_001025472.1); *R. rattus Lsr* (NM_032616).

The Genbank accession numbers for protein sequences: *M. musculus* *Lisch-like* (amino acid residues 150–795, XP_001473575); (*Lsr*) (NP_059101); *Ildr1* (NP_598870); *H. sapiens* *C1orf32*

(NP_955383); *LSR* (NP_057009); *ILDR1* (NP_787120); *D. rerio* *Lisch-like* (NP_001025363); *D. rerio* *Lsr* (NP_001020643); *R. rattus* *Lsr* (NP_116005)

Supporting Information

Table S1 Data for Figure 3A: Plasma Insulin/Glucose Ratios in Age-Grouped 1jc *Lep^{ob/ob}* Males.

Found at: doi:10.1371/journal.pgen.1000137.s001 (0.03 MB DOC)

Table S2 Pair-Wise Similarity Scores by Isoform and Domain for Figure 9: ClustalW Analysis of *Lisch-like* Homologs and the *LSR* Protein.

Found at: doi:10.1371/journal.pgen.1000137.s002 (0.03 MB DOC)

Table S3 ID numbers for Affymetrix MOE-430A Probes used in Methods: Microarray Gene Expression Analysis.

Found at: doi:10.1371/journal.pgen.1000137.s003 (0.03 MB DOC)

Table S4 Oligonucleotide Sequences.

Found at: doi:10.1371/journal.pgen.1000137.s004 (0.05 MB DOC)

Table S5 Data for Figure 12: Terminal Phenotypes in 404 Obese F2 Progeny by Genotype at D1mit110.

Found at: doi:10.1371/journal.pgen.1000137.s005 (0.04 MB DOC)

Acknowledgments

This manuscript is dedicated to our friend and associate, Dr. Marija Dokmanovic-Chouinard, who died on 25 July 2006. Her efforts were critical to the project reported here. We thank Rick Rausch, Brandi Davis, Rosa Rosario and Qiong Li for technical assistance, Michael Rosenbaum for help with statistical analyses, Streamson C. Chua for technical assistance and helpful discussions, Larry L. Shapiro for technical help with antibodies, and Domenico Accili for helpful advice. Derek J. Nunez (GlaxoSmithKline) provided helpful consultations.

Author Contributions

Conceived and designed the experiments: MDC WKC JCC EW NW CVW JO CEA KKB RI CL KS RLL. Performed the experiments: MDC WKC JCC EW NW CVW JO SG GMS CEA KKB RI CL KS RLL. Analyzed the data: JY BW YB SGF RLL. Wrote the paper: MDC WKC JCC KS SGF RLL.

References

- Cowie C (2003) Prevalence of diabetes and impaired fasting glucose in adults - United States, 1999–2000. *MMWR* 52: 833–837.
- Saltiel AR (2001) New perspectives into the molecular pathogenesis and treatment of type 2 diabetes. *Cell* 104: 517–529.
- Haffner SM (2006) Relationship of metabolic risk factors and development of cardiovascular disease and diabetes. *Obesity* (Silver Spring) 14 Suppl 3: 121S–127S.
- Hossain P, Kavar B, El Nahas M (2007) Obesity and diabetes in the developing world—a growing challenge. *N Engl J Med* 356: 213–215.
- DeFronzo RA, Bonadonna RC, Ferrannini E (1992) Pathogenesis of NIDDM. A balanced overview. *Diabetes Care* 15: 318–368.
- Kloppel G, Lohr M, Habich K, Oberholzer M, Heitz PU (1985) Islet pathology and the pathogenesis of type 1 and type 2 diabetes mellitus revisited. *Surv Synth Pathol Res* 4: 110–125.
- Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, et al. (2003) Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes* 52: 102–110.
- Miralles F, Portha B (2001) Early development of beta-cells is impaired in the GK rat model of type 2 diabetes. *Diabetes* 50 Suppl 1: S84–88.
- Leiter EH (1989) The genetics of diabetes susceptibility in mice. *Faseb J* 3: 2231–2241.
- Zucker LM, Antoniades HN (1972) Insulin and obesity in the Zucker genetically obese rat “fatty”. *Endocrinology* 90: 1320–1330.
- Frayling TM, Evans JC, Bulman MP, Pearson E, Allen L, et al. (2001) beta-cell genes and diabetes: molecular and clinical characterization of mutations in transcription factors. *Diabetes* 50 Suppl 1: S94–100.
- Barnett AH, Eff C, Leslie RD, Pyke DA (1981) Diabetes in identical twins. A study of 200 pairs. *Diabetologia* 20: 87–93.
- Lo SS, Tun RY, Hawa M, Leslie RD (1991) Studies of diabetic twins. *Diabetes Metab Rev* 7: 223–238.
- Kahn CR, Vicent D, Doria A (1996) Genetics of non-insulin-dependent (type-II) diabetes mellitus. *Annu Rev Med* 47: 509–531.
- Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD (1999) Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. *Diabetologia* 42: 146–150.
- Jun H, Bae HY, Lee BR, Koh KS, Kim YS, et al. (1999) Pathogenesis of non-insulin-dependent (type II) diabetes mellitus (NIDDM) - genetic predisposition and metabolic abnormalities. *Adv Drug Deliv Rev* 35: 157–177.

17. Permutt MA, Wasson J, Cox N (2005) Genetic epidemiology of diabetes. *J Clin Invest* 115: 1431–1439.
18. Florez JC, Hirschhorn J, Altshuler D (2003) The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet* 4: 257–291.
19. Khanim F, Kirk J, Latif F, Barrett TG (2001) WFS1/wolframin mutations, Wolfram syndrome, and associated diseases. *Hum Mutat* 17: 357–367.
20. Cox NJ, Xiang KS, Fajans SS, Bell GI (1992) Mapping diabetes-susceptibility genes. Lessons learned from search for DNA marker for maturity-onset diabetes of the young. *Diabetes* 41: 401–407.
21. Pimenta W, Korytkowski M, Mitrakou A, Jensen T, Yki-Jarvinen H, et al. (1995) Pancreatic beta-cell dysfunction as the primary genetic lesion in NIDDM. Evidence from studies in normal glucose-tolerant individuals with a first-degree NIDDM relative. *Jama* 273: 1855–1861.
22. Gelding SV, Andres C, Nithyananthan R, Gray IP, Mather H, et al. (1995) Increased secretion of 32,33 split proinsulin after intravenous glucose in glucose-tolerant first-degree relatives of patients with non-insulin dependent diabetes of European, but not Asian, origin. *Clin Endocrinol (Oxf)* 42: 255–264.
23. Knowler WC, Saad MF, Pettitt DJ, Nelson RG, Bennett PH (1993) Determinants of diabetes mellitus in the Pima Indians. *Diabetes Care* 16: 216–227.
24. Hanley AJ, Williams K, Gonzalez C, D'Agostino RB Jr, Wagenknecht LE, et al. (2003) Prediction of type 2 diabetes using simple measures of insulin resistance: combined results from the San Antonio Heart Study, the Mexico City Diabetes Study, and the Insulin Resistance Atherosclerosis Study. *Diabetes* 52: 463–469.
25. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881–885.
26. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, et al. (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 39: 770–775.
27. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38: 320–323.
28. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, et al. (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316: 889–894.
29. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336–1341.
30. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, et al. (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316: 1341–1345.
31. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316: 1331–1336.
32. Clee SM, Attie AD (2007) The genetic landscape of type 2 diabetes in mice. *Endocr Rev* 28: 48–83.
33. Flint J, Valdar W, Shifman S, Mott R (2005) Strategies for mapping and cloning quantitative trait genes in rodents. *Nat Rev Genet* 6: 271–286.
34. Todd JA (1999) From genome to aetiology in a multifactorial disease, type 1 diabetes. *Bioessays* 21: 164–174.
35. York B, Lei K, West DB (1996) Sensitivity to dietary obesity linked to a locus on chromosome 15 in a CAST/Ei x C57BL/6J F2 intercross. *Mamm Genome* 7: 677–681.
36. Mitsos LM, Cardon LR, Fortin A, Ryan L, LaCourse R, et al. (2000) Genetic control of susceptibility to infection with *Mycobacterium tuberculosis* in mice. *Genes Immun* 1: 467–477.
37. Welch CL, Bretschger S, Latib N, Bezouevski M, Guo Y, et al. (2001) Localization of atherosclerosis susceptibility loci to chromosomes 4 and 6 using the Ldlr knockout mouse model. *Proc Natl Acad Sci U S A* 98: 7946–7951.
38. Legare ME, Bartlett FS, 2nd, Frankel WN (2000) A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res* 10: 42–48.
39. Joobar R, Zarate JM, Rouleau GA, Skamene E, Boksa P (2002) Provisional mapping of quantitative trait loci modulating the acoustic startle response and prepulse inhibition of acoustic startle. *Neuropsychopharmacology* 27: 765–781.
40. Clee SM, Yandell BS, Schueller KM, Rabaglia ME, Richards OC, et al. (2006) Positional cloning of Sorcs1, a type 2 diabetes quantitative trait locus. *Nat Genet* 38: 688–693.
41. Goodarzi MO, Lehman DM, Taylor KD, Guo X, Cui J, et al. (2007) SORCS1: a novel human type 2 diabetes susceptibility gene suggested by the mouse. *Diabetes* 56: 1922–1929.
42. Freeman H, Shimomura K, Horner E, Cox RD, Ashcroft FM (2006) Nicotinamide nucleotide transhydrogenase: a key role in insulin secretion. *Cell Metab* 3: 35–45.
43. Freeman HC, Hugill A, Dear NT, Ashcroft FM, Cox RD (2006) Deletion of nicotinamide nucleotide transhydrogenase: a new quantitative trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes* 55: 2153–2156.
44. Wade CM, Kulbokas EJ, 3rd, Kirby AW, Zody MC, Mullikin JC, et al. (2002) The mosaic structure of variation in the laboratory mouse genome. *Nature* 420: 574–578.
45. Leiter EH, Chapman HD, Coleman DL (1989) The influence of genetic background on the expression of mutations at the diabetes locus in the mouse. V. Interaction between the db gene and hepatic sex steroid sulfotransferases correlates with gender-dependent susceptibility to hyperglycemia. *Endocrinology* 124: 912–922.
46. Kido Y, Burks DJ, Withers D, Bruning JC, Kahn CR, et al. (2000) Tissue-specific insulin resistance in mice with mutations in the insulin receptor, IRS-1, and IRS-2. *J Clin Invest* 105: 199–205.
47. Stanton KJ, Sidner RA, Miller GA, Cummings OW, Schmidt CM, et al. (2003) Analysis of Ki-67 antigen expression, DNA proliferative fraction, and survival in resected cancer of the pancreas. *Am J Surg* 186: 486–492.
48. Bonner-Weir S (2000) Life and death of the pancreatic beta cells. *Trends Endocrinol Metab* 11: 375–378.
49. Bonner-Weir S (2000) Perspective: Postnatal pancreatic beta cell growth. *Endocrinology* 141: 1926–1929.
50. Bonner-Weir S (2001) beta-cell turnover: its assessment and implications. *Diabetes* 50 Suppl 1: S20–24.
51. Dor Y, Brown J, Martinez OI, Melton DA (2004) Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature* 429: 41–46.
52. Scaglia L, Cahill CJ, Finegood DT, Bonner-Weir S (1997) Apoptosis participates in the remodeling of the endocrine pancreas in the neonatal rat. *Endocrinology* 138: 1736–1741.
53. Gavrieli Y, Sherman Y, Ben-Sasson SA (1992) Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J Cell Biol* 119: 493–501.
54. Fernandes-Alnemri T, Litwack G, Alnemri ES (1994) CPP32, a novel human apoptotic protein with homology to *Caenorhabditis elegans* cell death protein Ced-3 and mammalian interleukin-1 beta-converting enzyme. *J Biol Chem* 269: 30761–30764.
55. Wang J, Li S, Zhang Y, Zheng H, Xu Z, et al. (2003) Vertebrate gene predictions and the problem of large genes. *Nat Rev Genet* 4: 741–749.
56. Lindblad-Toh K, Lander ES, McPherson JD, Waterston RH, Rodgers J, et al. (2001) Progress in sequencing the mouse genome. *Genesis* 31: 137–141.
57. Bromberg Y, Rost B (2007) SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res* 35: 3823–3835.
58. Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 30: 3894–3900.
59. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31: 3812–3814.
60. Dayhoff M (1978) Atlas of Protein Sequence and Structure. In: Dayhoff M, ed. Washington, DC: National Biochemical Research Foundation. pp 353–358.
61. Rost B, Sander C (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* 19: 55–72.
62. German MS, Wang J, Fernald AA, Espinosa R, 3rd, Le Beau MM, et al. (1994) Localization of the genes encoding two transcription factors, LMX1 and CDX3, regulating insulin gene expression to human chromosomes 1 and 13. *Genomics* 24: 403–404.
63. Hsieh CH, Liang KH, Hung YJ, Huang LC, Pei D, et al. (2006) Analysis of epistasis for diabetic nephropathy among type 2 diabetic patients. *Hum Mol Genet* 15: 2701–2708.
64. Schulz H (2003) Towards a Comprehensive Description of the Human Retinal Transcriptome: Identification and Characterization of Differentially Expressed Genes [PhD dissertation]: University of Wurzburg. pp 5.
65. Hauge H, Patzke S, Delabie J, Aasheim HC (2004) Characterization of a novel immunoglobulin-like domain containing receptor. *Biochem Biophys Res Commun* 323: 970–978.
66. Draper BW, Morcos PA, Kimmel CB (2001) Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis* 30: 154–156.
67. Nasevicius A, Ekker SC (2000) Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* 26: 216–220.
68. Kim HJ, Sumanas S, Palencia-Desai S, Dong Y, Chen JN, et al. (2006) Genetic analysis of early endocrine pancreas formation in zebrafish. *Mol Endocrinol* 20: 194–203.
69. Field HA, et al. (2003a) Formation of the digestive system in zebrafish. II. Pancreas morphogenesis. *Dev Biol* 261: 197–208.
70. Sherwood NM, Wu S (2005) Developmental role of GnRH and PACAP in a zebrafish model. *Gen Comp Endocrinol* 142: 74–80.
71. McGonnell IM, Fowkes RC (2006) Fishing for gene function—endocrine modelling in the zebrafish. *J Endocrinol* 189: 425–439.
72. Gnugge L, Meyer D, Driever W (2004) Pancreas development in zebrafish. *Methods Cell Biol* 76: 531–551.
73. Field HA, Ober EA, Roeser T, Stainier DY (2003) Formation of the digestive system in zebrafish. I. Liver morphogenesis. *Dev Biol* 253: 279–290.
74. Zecchin E, Mavropoulos A, Devos N, Filippi A, Tiso N, et al. (2004) Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates. *Dev Biol* 268: 174–184.
75. Lin JW, Biankin AV, Horb ME, Ghosh B, Prasad NB, et al. (2004) Differential requirement for ptf1a in endocrine and exocrine lineages of developing zebrafish pancreas. *Dev Biol* 274: 491–503.
76. Yee NS, Yusuff S, Pack M (2001) Zebrafish pdx1 morphant displays defects in pancreas development and digestive organ chirality, and potentially identifies a multipotent pancreas progenitor cell. *Genesis* 30: 137–140.

77. Augustin M, Sedlmeier R, Peters T, Huffstadt U, Kochmann E, et al. (2005) Efficient and fast targeted production of murine models based on ENU mutagenesis. *Mamm Genome* 16: 405–413.
78. Numata K, Okada Y, Saito R, Kiyosawa H, Kanai A, et al. (2007) Comparative analysis of cis-encoded antisense RNAs in eukaryotes. *Gene* 392: 134–141.
79. Lapidot M, Pilpel Y (2006) Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO Rep* 7: 1216–1222.
80. Shalgi R, Lapidot M, Shamir R, Pilpel Y (2005) A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biol* 6: R86.
81. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, et al. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 434: 338–345.
82. Xia J, Scherer SW, Cohen PT, Majer M, Xi T, et al. (1998) A common variant in PPP1R3 associated with insulin resistance and type 2 diabetes. *Diabetes* 47: 1519–1524.
83. Xia J, Bogardus C, Prochazka M (1999) A type 2 diabetes-associated polymorphic ARE motif affecting expression of PPP1R3 is involved in RNA-protein interactions. *Mol Genet Metab* 68: 48–55.
84. Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP, et al. (2007) MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol Cell* 27: 91–105.
85. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, et al. (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* 432: 226–230.
86. Prentki M, Nolan CJ (2006) Islet beta cell failure in type 2 diabetes. *J Clin Invest* 116: 1802–1812.
87. Stanger BZ, Tanaka AJ, Melton DA (2007) Organ size is limited by the number of embryonic progenitor cells in the pancreas but not the liver. *Nature* 445: 886–891.
88. Finegood DT, Scaglia L, Bonner-Weir S (1995) Dynamics of beta-cell mass in the growing rat pancreas. Estimation with a simple mathematical model. *Diabetes* 44: 249–256.
89. Teta M, Long SY, Wartschow LM, Rankin MM, Kushner JA (2005) Very slow turnover of beta-cells in aged adult mice. *Diabetes* 54: 2557–2567.
90. Hales CN, Barker DJ (2001) The thrifty phenotype hypothesis. *Br Med Bull* 60: 5–20.
91. Yen FT, Masson M, Clossais-Besnard N, Andre P, Grosset JM, et al. (1999) Molecular cloning of a lipolysis-stimulated remnant receptor expressed in the liver. *J Biol Chem* 274: 13390–13398.
92. Leahy JL, Vandekerkhove KM (1990) Insulin-like growth factor-I at physiological concentrations is a potent inhibitor of insulin secretion. *Endocrinology* 126: 1593–1598.
93. Garcia-Ocana A, Takane KK, Syed MA, Philbrick WM, Vasavada RC, et al. (2000) Hepatocyte growth factor overexpression in the islet of transgenic mice increases beta cell proliferation, enhances islet mass, and induces mild hypoglycemia. *J Biol Chem* 275: 1226–1232.
94. Jin J, Smith FD, Stark C, Wells CD, Fawcett JP, et al. (2004) Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol* 14: 1436–1450.
95. Onuma H, Osawa H, Yamada K, Ogura T, Tanabe F, et al. (2002) Identification of the insulin-regulated interaction of phosphodiesterase 3B with 14-3-3 beta protein. *Diabetes* 51: 3362–3367.
96. Xiang K, Wang Y, Zheng T, Jia W, Li J, et al. (2004) Genome-wide search for type 2 diabetes/impaird glucose homeostasis susceptibility genes in the Chinese: significant linkage to chromosome 6q21-q23 and chromosome 1q21-q24. *Diabetes* 53: 228–234.
97. Pozuelo Rubio M, Geraghty KM, Wong BH, Wood NT, Campbell DG, et al. (2004) 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. *Biochem J* 379: 395–408.
98. Meek SE, Lane WS, Pivnicka-Worms H (2004) Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins. *J Biol Chem* 279: 32046–32054.
99. Hermeking H, Benzinger A (2006) 14-3-3 proteins in cell cycle regulation. *Semin Cancer Biol* 16: 183–192.
100. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ (1999) A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes* 48: 1175–1182.
101. Hsueh WC, St Jean PL, Mitchell BD, Pollin TI, Knowler WC, et al. (2003) Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. *Diabetes* 52: 550–557.
102. St. Jean P, Hsueh W-C, Mitchell B, Ehm M, Wagner M, Burns D, Shuldiner AR (2000) Association between diabetes, obesity, glucose and insulin levels in the Old Amish and SNP's on 1q21-23 (Abstract). *American Journal of Human Genetics* 67: 1848.
103. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, et al. (2001) A genome-wide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *Am J Hum Genet* 69: 553–569.
104. Vionnet N, Hani El H, Dupont S, Gallina S, Francke S, et al. (2000) Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. *Am J Hum Genet* 67: 1470–1480.
105. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA (2002) A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study. *Diabetes* 51: 833–840.
106. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, et al. (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. *Am J Hum Genet* 63: 1130–1138.
107. Das SK, Elbein SC (2007) The search for type 2 diabetes susceptibility Loci: the chromosome 1q story. *Curr Diab Rep* 7: 154–164.
108. McCarthy M, Shuldiner AR, Bogardus C, Hanson RL, Elbein S (2004) Positional Cloning of a Type 2 Diabetes Susceptibility Gene on Chromosome 1q: A collaborative effort by the Chromosome 1q Diabetes Positional Cloning Consortium. pp 1–39.
109. Chung WK, Zheng M, Chua M, Kershaw E, Power-Keohoe L, et al. (1997) Genetic modifiers of *Leprfa* associated with variability in insulin production and susceptibility to NIDDM. *Genomics* 41: 332–344.
110. Chung WK, Power-Keohoe L, Chua M, Chu F, Aronne L, et al. (1997) Exonic and intronic sequence variation in the human leptin receptor gene (*LEPR*). *Diabetes* 46: 1509–1511.
111. Amar LC, Arnaud D, Cambrou J, Guenet JL, Avner PR (1985) Mapping of the mouse X chromosome using random genomic probes and an interspecific mouse cross. *Embo J* 4: 3695–3700.
112. Chung WK, Chua SC, Lee GH, Leibel RL (1997) Polymerase chain reaction-restriction fragment length polymorphisms (PCR-RFLP) and electrophoretic assays for the mouse obese (*Lepob*) mutation. *Obes Res* 5: 183–185.
113. Phan LK, Lin F, LeDuc CA, Chung WC, Leibel RL (2002) The mouse mahoganoid coat color mutation disrupts a novel C3HC4 RING domain protein. *J Clin Invest* 110: 1449–1459.
114. Visscher PM (1999) Speed congenics: accelerated genome recovery using genetic markers. *Genet Res* 74: 81–85.
115. Surwit RS, Kuhn CM, Cochrane C, McCubbin JA, Feinglos MN (1988) Diet-induced type II diabetes in C57BL/6J mice. *Diabetes* 37: 1163–1167.
116. Finegood DT, McArthur MD, Kojwang D, Thomas MJ, Topp BG, et al. (2001) Beta-cell mass dynamics in Zucker diabetic fatty rats. Rosiglitazone prevents the rise in net cell death. *Diabetes* 50: 1021–1029.
117. Guillam MT, Dupraz P, Thorens B (2000) Glucose uptake, utilization, and signaling in GLUT2-null islets. *Diabetes* 49: 1485–1491.
118. Lacy PE, Kostianovsky M (1967) Method for the isolation of intact islets of Langerhans from the rat pancreas. *Diabetes* 16: 35–39.
119. Gotoh M, Maki T, Kiyozumi T, Satomi S, Monaco AP (1985) An improved method for isolation of mouse pancreatic islets. *Transplantation* 40: 437–438.
120. Eizirik DL, Sandler S, Hallberg A, Bendtzen K, Sener A, et al. (1989) Differential sensitivity to beta-cell secretagogues in cultured rat pancreatic islets exposed to human interleukin-1 beta. *Endocrinology* 125: 752–759.
121. Weisberg SP, McCann D, Desai M, Rosenbaum M, Leibel RL, et al. (2003) Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest* 112: 1796–1808.
122. Westerfield M (2000) *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio rerio)*. Eugene, OR: University of Oregon Press.
123. Thisse C, Thisse B (1998) High resolution whole-mount in situ hybridization. *Zebrafish Sci Monit* 5: 8–9.
124. Schwartz RM, Dayhoff MO (1978) Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* 199: 395–403.
125. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, et al. (2005) The Universal Protein Resource (UniProt). *Nucleic Acids Res* 33: D154–159.
126. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.