Loss-of-function mutations in SLC30A8 protect against type 2 diabetes

A full list of authors and affiliations appears at the end of the article.

Abstract

Loss-of-function mutations protective against human disease provide in vivo validation of therapeutic targets\(^1,2,3\), yet none are described for type 2 diabetes (T2D). Through sequencing or genotyping ~150,000 individuals across five ethnicities, we identified 12 rare protein-truncating variants in SLC30A8, which encodes an islet zinc transporter (ZnT8)\(^4\) and harbors a common variant (p.Trp325Arg) associated with T2D risk, glucose, and proinsulin levels\(^5–^7\). Collectively,
protein-truncating variant carriers had 65% reduced T2D risk ($p=1.7\times10^{-6}$), and non-diabetic Icelandic carriers of a frameshift variant (p.Lys34SerfsX50) demonstrated reduced glucose levels ($-0.17$ s.d., $p=4.6\times10^{-4}$). The two most common protein-truncating variants (p.Arg138X and p.Lys34SerfsX50) individually associate with T2D protection and encode unstable ZnT8 proteins. Previous functional study of SLC30A8 suggested reduced zinc transport increases T2D risk$^{8,9}$, yet phenotypic heterogeneity was observed in rodent Slc30a8 knockouts$^{10–15}$. Contrastingly, loss-of-function mutations in humans provide strong evidence that SLC30A8 haploinsufficiency protects against T2D, proposing ZnT8 inhibition as a therapeutic strategy in T2D prevention.

Genome-wide association studies (GWAS) have identified 65 genomic loci associated with T2D risk$^7$, highlighting previously unidentified pathological pathways. Translation into novel therapeutic targets$^16$ requires identification of causal mutations and genes, as well as the directional relationship between protein activity and disease risk$^17$. Toward this end, loss-of-function (LoF) mutations that protect against disease (without adverse phenotypes) are among the most useful findings from human genetics, suggesting targets that, upon inhibition, may prevent disease in the general population.

To identify T2D-protective LoF variants, in 2009 we sequenced exons of 115 genes near T2D GWAS signals (Supplementary Tables 1–2, Supplementary Fig. 1) in 758 individuals from Finland or Sweden (modeling previous studies$^{18}$). To increase power, we selected individuals at the extremes of T2D risk: 352 young and lean T2D cases and 406 elderly and obese euglycemic controls$^{19}$ (Supplementary Table 3). In total, 1,768 non-synonymous variants were identified (1,683 single nucleotide variants [SNVs] and 85 indels), 1,474 (83%) with minor allele frequency (MAF) <1% and 1,108 (63%) observed in only one individual. We found no evidence of association with T2D when testing individual variants or a burden of rare variants within genes (Supplementary Fig. 2). Genotyping 71 select SNVs (with nominally significant association or predicted to impact protein structure) in 11,288 additional individuals also yielded results consistent with the null distribution (Supplementary Fig. 3).

To increase power to detect association, we used the Illumina Human Exome Array to further genotype a subset of SNVs in 21,096 Finnish or Swedish individuals (10,534 with diabetes and 10,562 without, a superset of the individuals genotyped for the 71 SNVs, Supplementary Table 4). Analysis focused on variants with clear functional interpretation: nonsense, frameshift, or splice site mutations predicted to cause protein truncation (Supplementary Table 5). Six such variants identified via the sequencing were present on the Exome Array.

Of these variants, only a nonsense SNV (c.412C>T, p.Arg138X) in SLC30A8 (transcript accession number NM_173851$^{20}$) showed nominally significant association with T2D (OR=0.46, $p=0.012$, Supplementary Table 6). A second SLC30A8 nonsense variant (c.456G>A, p.Trp152X) was observed in one control (Supplementary Table 5) from sequencing, but was absent from the Exome Array. As a further experiment, we genotyped p.Arg138X in 26,566 additional European individuals (8,210 cases and 18,356 controls; Supplementary Table 7): although only 16 heterozygotes were observed (two cases and 14 controls), the association with T2D risk was directionally consistent ($OR=0.56, p>0.05$).
Based on the combined data, heterozygosity for p.Arg138X was estimated to yield a 53% reduction in T2D risk (p=0.0067, N=48,115).

SLC30A8 encodes an islet zinc transporter ZnT8 (NP_776250), which is necessary for zinc flux into β-cell insulin-secretory granules and subsequent insulin crystallization. Upon co-secretion with insulin, zinc also fulfills auto- and paracrine signaling roles. A previously-identified common SLC30A8 missense variant (rs13266634; c.973T>A, p.Trp325Arg) associates with T2D risk, glucose, and proinsulin, at significance levels beyond genome-wide thresholds.

Cellular characterization has suggested that the risk-increasing allele of p.Trp325Arg reduces ZnT8 zinc transport activity. In Slc30a8 knockout mice, however, the phenotype varies with gender and genetic background: observations range from no effect on insulin secretion or glucose homeostasis, to modest hyperglycemia on a high fat diet. Furthermore, a recent β-cell-specific Slc30a8 knockout proposes a multi-organ effect on the resultant mouse phenotype, with circulating zinc shown to influence hepatic insulin clearance. Thus, the directional relationship between perturbed ZnT8 function and whole organism phenotype is uncertain despite much genetic and biological data.

Because the observed protective association between p.Arg138X and T2D risk was statistically modest, we sought additional evidence. Unfortunately, the near absence of p.Arg138X outside Western Finland limited ability to further characterize its effect in other populations (Supplementary Figs 4–5). We thus sought to identify a wider spectrum of protein-truncating variants in SLC30A8, through investigation of the catalog of 35 million variants collected by deCODE genetics through whole-genome sequencing. The p.Arg138X variant was not observed in this dataset. However, an independent protein-truncating variant was observed at 0.17% frequency: a deletion (c.101_107del, p.Lys34SerfsX50; Supplementary Figs 6–7) predicted to cause a frameshift and loss of all six transmembrane domains in the islet specific transcript (NM_173851) of SLC30A8.

Heterozygosity for p.Lys34SerfsX50 was associated with 80% reduced T2D risk, with two observations in 2,953 T2D cases (0.03%) versus 248 observations in 67,919 controls (0.18%; OR=0.18, p=0.004; Supplementary Tables 8–9). Based on the ancestral relationship between Norway and Iceland, we genotyped the variant in 5,714 Norwegians (Supplementary Table 8) and observed zero carriers in 1,645 cases versus three carriers in 4,069 controls. Combining the evidence for p.Lys34SerfsX50 and p.Arg138X strengthened the association between SLC30A8 protein-truncating variants and reduced T2D risk (combined OR=0.32, p=2.4×10^{-4}).

Both rare SLC30A8 variants are bioinformatically predicted to cause ZnT8 truncation and consequently impact activity. To test this prediction, we assessed over-expressed, V5-tagged ZnT8 variants (Trp325, Arg325, X138 [as well as Arg138X], and Ser34fsX50,) in HeLa cells. Despite similar RNA transcript levels for all variants (Supplementary Fig. 8), only Trp325- and Arg325-ZnT8 proteins were easily detectable in cells, with Arg138X, X138-, and Ser34fsX50-ZnT8 present at low to undetectable levels (Fig. 1bc). Similar results were obtained using antibodies against the native protein or the V5-tag, via Western

*Nat Genet.* Author manuscript; available in PMC 2014 June 10.
blot (Fig. 1b) and immunofluorescence (Fig. 1c), and in HeLa as well as Ins1e rat insulinoma cells (Fig. 1d). Co-expression of X138- or Ser34fsX50-ZnT8 with Trp325-ZnT8 did not decrease expression of the full-length allele, nor rescue expression of either truncating variant (Supplementary Fig. 9).

We hypothesized that decreased expression of these two mutants might be due to protein instability and/or enhanced degradation. Following treatment with chloroquine or MG132 (lysosomal and proteasomal inhibitors respectively), higher X138- and Ser34fsX50-ZnT8 expression was detected via immunofluorescence (but remained undetectable via Western blot; Fig. 1e, Supplementary Fig. 10). These results are consistent with (but do not prove) instability and subsequent degradation of these truncated proteins. Through additional experiments (data not shown), we observed zinc transport in cells expressing Arg325- and Trp325-ZnT8 but not X138- or Ser34fsX50-ZnT8 (expected given the low levels of mutant protein). Further experiments are needed to assess the in vivo impact of these variants, including susceptibility to nonsense-mediated decay and potential dominant negative effects on protein oligomerization.

These genetic and functional data suggest SLC30A8 haploinsufficiency reduces T2D risk. However, confidence would be further increased through observation of multiple, additional, putative LoF variants demonstrating protective effects. As part of the T2D-GENES and GoT2D consortia, we sequenced SLC30A8 exons in 12,294 individuals spanning multiple ethnicities (Supplementary Table 10). Nine additional protein-truncating variants were identified – two frameshift indels and two nonsense, four splice site, and one initiator codon SNV – in 23 heterozygous individuals from African American, East Asian, and South Asian ancestries (Supplementary Data Set 1). p.Arg138X was seen in three additional carriers (one case and two controls); p.Lys34SerfsX50 was not observed.

In aggregate, carriers of these additional variants exhibited 60% reduced T2D risk (four case versus 18 control observations, OR=0.38, p=0.0025), with similar effects and statistical significance observed upon analysis of only frameshift or nonsense variant carriers (two case versus 13 control observations, OR=0.37, p=0.0027). Combining all data from sequencing and genotyping in 149,134 subjects, heterozygosity for any of the 12 protein-truncating variants was associated with 65% reduced T2D risk (OR=0.34, p=1.7×10⁻⁶), a statistically significant association even after correction for ~20,000 genes in the human genome (Table 1).

We investigated potential confounding factors for the observed protective association. We first assessed whether the p.Trp325Arg haplotypic background might influence results. While p.Lys34SerfsX50 and p.Met50Ile variants were isolated to the protective common variant haplotype, the remaining variants (including p.Arg138X) were observed on the risk common variant haplotype. Thus, independent protective protein-truncating variants were observed on opposite p.Trp325Arg haplotypic backgrounds. Second, we tested for a survivor effect, where rare variant carriers with diabetes would die at a younger age. However, (a) carrier ages did not significantly differ from non-carrier ages for either p.Arg138X (69.6±8.4 versus 65.5±11.0 for cases [p>0.1], 46.4±15.7 versus 50.3±15.5 for controls [p>0.1]) or p.Lys34SerfsX50 (70.5±4.5 versus 65.6±13.8 for cases [p>0.1],

*Nat Genet*. Author manuscript; available in PMC 2014 June 10.
48.5±20.1 versus 50.0±23.2 for controls \( [p>0.1] \), and (b) p.Lys34SerfsX50 association attained equivalent significance even when analysis was restricted to age-matched controls. Finally, we acknowledged the noted challenges to control for population stratification in rare variant association studies\(^\text{26}\). We had insufficient data to perform a family-based transmission disequilibrium test (pedigree information was only available for Icelanders, with three carrier parents all transmitting the risk allele to affected children). However, the consistent association of multiple independent protein-truncating variants across multiple cohorts and ancestries argues against population stratification as entirely responsible for the protective association.

These data thus provide compelling evidence that mutations inactivating one copy of \( SLC30A8 \) reduce T2D risk in humans. In addition to T2D risk, the common \( SLC30A8 \) variant (p.Trp325Arg) is associated with proinsulin and fasting plasma glucose levels at genome-wide significance\(^\text{5,6}\), as well as 2-hr glucose levels post-oral glucose tolerance test (OGTT) at nominal significance (Supplementary Table 11)\(^\text{27}\). We asked whether rare protein-truncating \( SLC30A8 \) variants also affected T2D-related phenotypes, particularly glycemic traits that might be indicative of altered islet function.

Among traits analyzed (Supplementary Table 12), the strongest association was observed in Iceland between p.Lys34SerfsX50 and random (non-fasting) glucose: non-diabetic carriers of the protective allele had lower glucose (\( \beta=-0.17\text{s.d.}; N=182 \text{ carriers}; p=4.6\times10^{-4} \)), with a consistent effect seen in three Norwegian carriers (\( \beta=-0.3\text{s.d.}, p>0.1 \)). Glucose was lower at one hour in the small number of p.Lys34SerfsX50 carriers characterized by OGTT (\( \beta=-0.73\text{s.d.}; N=4 \text{ carriers}; p=0.05 \)). We did not observe a significant difference in fasting glucose or insulin, although the directions of effect were consistent with the above: lower for fasting glucose (average \( \beta=-0.10\text{s.d.}; N=146 \text{ carriers}; p>0.1 \)) and higher for fasting insulin (\( \beta=0.24\text{s.d.}; N=52 \text{ carriers}; p=0.09 \)). The co-directionality of glucose levels and T2D risk parallels the pattern observed for p.Trp325Arg, where the T2D-protective allele also associates with lower glucose (Supplementary Table 9), providing further evidence against a survivor effect or population stratification as driving the protective association.

In summary, we identified 12 rare, predicted protein-truncating \( SLC30A8 \) variants (Fig. 2). Carriers of these variants had 65% reduced T2D risk at a level of significance adequate to correct for ~20,000 genes in the human genome (\( p=1.7\times10^{-6} \)). Non-diabetic Icelandic carriers of p.Lys34SerfsX50 also demonstrated lower glucose levels (\( \beta=-0.17\text{s.d.}; p=4.6\times10^{-4} \)). Notably, initial sequencing of 115 genes in 758 extreme individuals produced only two observations of p.Arg138X, without significant evidence of association of low-frequency or rare variants individually or in aggregate for any of the sequenced genes. Rather, establishing the association of \( SLC30A8 \) protein-truncating variants with T2D protection at levels of exome-wide significance (correction for 20,000 genes) required genotyping ~150,000 individuals spanning multiple ethnicities. Detecting similar effects in genes without prior evidence of association may require analysis at a similar or larger scale, for not only T2D but also other complex traits.

Previous modeling of the relationship between ZnT8 activity and T2D risk centered on p.Trp325Arg, where mildly attenuated zinc transport is concomitant with increased T2D...
risk\(^8\), and SLC30A8 knockout mice, where phenotypic heterogeneity is observed\(^{13,15}\). We find a clear and consistent association between putative SLC30A8 LoF variants and T2D risk, across multiple ethnic backgrounds, demonstrating convincingly that a 50% reduction in gene dosage protects against T2D in humans. These data reject the model that SLC30A8 LoF is associated with as little as a 1.2-fold increase in T2D risk (similar to that for the common p.Trp325Arg variant) at significance of \(p \approx 10^{-9}\). Phenotypic interrogation of human mutation carriers is needed to determine the physiological mechanism behind this protective association and establish the effects of SLC30A8 haploinsufficiency in the pancreas and other tissues\(^21\).

The observed human genetics data present several implications for SLC30A8 function in T2D pathophysiology. The identification of multiple disease-associated protein-altering variants in SLC30A8 unambiguously (albeit unsurprisingly) documents SLC30A8 as the causal gene behind GWAS association signals. The observation that protein-truncating variants protect against T2D defines the directional relationship between SLC30A8 activity and T2D risk in humans. The expanded SLC30A8 allelic series offers a more functionally-informative catalog of variation versus p.Trp325Arg alone, enabling future experiments investigating potential mechanisms. Although significant work is required to understand how reduced SLC30A8 activity lowers T2D risk, the current observations motivate experiments to test ZnT8 inhibition in T2D treatment in human populations.

**Methods**

**Sequencing and genotyping**

Individuals were selected for initial sequencing from several population-based cohorts from Finland and Sweden. A custom hybrid selection array was used to target genes, which were sequenced on an Illumina HiSeq 2000. Additional individuals from these same cohorts, as well as other cohorts drawn from different European populations, were genotyped for the SLC30A8 nonsense SNV p.Arg138X through the Illumina HumanExome v1.1 array. All sequenced individuals were also genotyped, with data showing 100% concordance.

Icelandic individuals were genotyped for the frameshift p.Lys34SerfsX50 variant using a combination of whole-genome sequencing and imputation (either direct imputation based on chip-genotyping, or through familial-based imputation). Sanger sequencing was used to confirm carriers. Norwegian individuals were genotyped with a fragment-length-based method using differentially-labeled fluorescent primers, with Sanger sequencing again used to confirm carriers. Further SLC30A8 sequencing (aimed at identifying further rare variant carriers) was performed as part of a whole-exome sequencing experiment, with the Agilent SureSelect Human All Exon platform used to capture exons and an Illumina HiSeq 2000 used for sequencing.

These studies were performed using protocols approved by the ethics committees of Helsinki University Hospital, Finland and Lund University, the Data Protection Commission of Iceland and the National Bioethics Committee of Iceland, the Regional Committee for Research Ethics and the Norwegian Data Inspectorate, and the Massachusetts Institute of
Technology Institutional Review Board, as well as with informed consent from all participants.

**Association analysis**

Association analysis was performed separately for three groups of variants: p.Arg138X, p.Lys34SerfsX50, and the remaining variants. For p.Arg138X, association analysis was separate for each analyzed cohort, and used a linear mixed model so as to account for sample structure including population stratification and genetic relatedness. Results were combined via a fixed-effects meta-analysis. For p.Lys34SerfsX50, association analysis was performed in Iceland using logistic regression, with controls matched to cases based on how informative the imputed genotypes were, and in Norway using a simple logistic regression with significance calculated via the score statistic. For the remaining variants, all individuals were analyzed jointly via a collapsing method, treating carriers of any variant indistinguishably, and regressing phenotype on the presence of any variant, with a linear mixed model used to account for sample structure. The resulting three association statistics were combined via a random-effects meta-analysis to obtain combined estimates of effect size and statistical significance.

For further details, see Supplementary Information.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Authors**

Jason Flannick¹,²,³, Gudmar Thorleifsson⁴, Nicola L. Beer¹,⁵, Suzanne B. R. Jacobs¹, Niels Grarup⁶, Noël P. Burtt¹, Anubha Mahajan⁷, Christian Fuchsberger⁸, Gil Atzmon⁹,¹⁰, Rafn Benediktsson¹¹, John Blangero¹², Don W. Bowden¹³,¹⁴,¹⁵,¹⁶, Ivan Brandslund¹⁷,¹⁸, Julia Brosnan¹⁹, Frank Burslem²⁰, John Chambers²¹,²²,²³, Yoon Shin Cho²⁴, Cramer Christensen²⁵, Desiree A. Douglas²⁶, Ravindranath Duggirala¹², Zachary Dynek¹, Yossi Farjoun¹, Timothy Fennell¹, Pierre Fontanillas¹, Tom Forsén²⁷,²⁸, Stacey Gabriel¹, Benjamin Glaser²⁹,³⁰, Daniel F. Gudbjartsson⁴, Craig Hanis³¹, Torben Hansen³², Astradur B. Hreidarsson¹¹, Kristian Hveem³³, Erik Ingelsson⁷,³⁴, Bo Isomaa³⁵,³⁶, Stefan Johannsson³⁷,³⁸,³⁹, Torben Jørgensen⁴⁰,⁴¹,⁴², Marit Eika Jørgensen⁴³, Sekar Kathiresan¹,⁴⁴,⁴⁵,⁴⁶, Augustine Kong⁴, Jaspal Kooner²²,²³,⁴⁷, Jasmina Kravik⁴⁸, Markku Laakso⁴⁹, Jong-Young Lee⁵⁰, Lars Lind⁵¹, Cecilia M Lindgren¹,⁷, Allan Linneberg⁴⁰,⁴¹,⁵², Gisli Masson⁴, Thomas Meitinger⁵³, Karen L Mohlke⁵⁴, Anders Molven³⁷,³⁸,³⁹, Andrew P. Morris⁷,⁵⁷, Shobha Potluri⁵⁸, Rainer Rauramaa⁵⁹,⁶⁰, Rasmus Ribel-Madsen⁶, Ann-Marie Richard¹⁹, Tim Rolph¹⁹, Veikko Salomaa⁶¹, Ayllet V. Segré¹,², Hanna Skástrønd⁶², Valgerdur Steinthorsdottir⁴, Heather M. Stringham⁸, Patrick Sulem⁴, E Shyong Tai⁶²,⁶³,⁶⁴, Yuk Ying Teo⁶²,⁶⁵,⁶⁶,⁶⁷,⁶⁸, Tanya Teslovich⁸, Unnur Thorsteinsdottir⁶⁹, Jeff K. Trimmer¹⁹, Tiinamaija Tuomi³⁵,³⁶, Jaakko Tuomilehto⁷¹,⁷²,⁷³, Fariba Vaziri-Sani²⁶, Benjamin F. Voight¹,⁷⁴,⁷⁵, James G. Wilson⁷⁶, Michael Boehnke⁸, Mark I. McCarthy⁷,⁷⁷,⁷⁸, Pål R. Njølstad¹,³⁷,⁷⁹, Oluf
Pedersen⁶, the Go-T2D consortium⁸⁰, the T2D-GENES consortium⁸¹, Leif Groop⁴⁸,⁸², David R. Cox⁵⁸, Kari Stefansson⁴,⁶⁹,ʻ, and David Altshuler¹,²,³,⁴⁴,⁴⁵,⁸³,⁸⁴,ʻ

**Affiliations**

¹Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA ²Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA ³Diabetes Unit, Massachusetts General Hospital, Boston, MA, USA ⁴deCODE genetics/Amgen, Inc., Reykjavik, Iceland ⁵Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Oxford, U.K ⁶The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark ⁷Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK ⁸Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA ⁹Department of Medicine, Albert Einstein College of Medicine, Bronx, NY, USA ¹⁰Department of Genetics, Albert Einstein College of Medicine, Bronx, NY, USA ¹¹Department of Endocrinology and Metabolism, Landspitali-University Hospital, Reykjavik, Iceland ¹²Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA ¹³Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA ¹⁴Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA ¹⁵Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA ¹⁶Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA ¹⁷Department of Clinical Biochemistry, Vejle Hospital, Vejle, Denmark ¹⁸Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark ¹⁹Cardiovascular & Metabolic Diseases Research Unit, Pfizer Inc., Cambridge, MA, USA ²⁰Cardiovascular and Metabolic Diseases Practice, Prescient Life Sciences, London, UK ²¹Department of Epidemiology and Biostatistics, Imperial College London, London, UK ²²Imperial College Healthcare NHS Trust, London, UK ²³Ealing Hospital National Health Service (NHS) Trust, Middlesex, UK ²⁴Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, Korea ²⁵Department of Internal Medicine and Endocrinology, Vejle Hospital, Vejle, Denmark ²⁶Department of Clinical Sciences, Unit of Diabetes and Celiac Diseases, Lund University, Malmö, Sweden ²⁷University of Helsinki, Department of General Practice and Primary Health Care, Finland ²⁸Vaasa Health Care Centre, Diabetes Care Unit, Vaasa, Finland ²⁹Endocrinology and Metabolism Service, Hadassah-Hebrew University Medical Center, Jerusalem, Israel ³⁰Israel Diabetes Research Group (IDRG), Israel ³¹Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA ³²Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark ³³Department of Public Health, Faculty of Medicine, Norwegian University of Science and Technology, Levanger, Norway ³⁴Molecular Epidemiology and Science for Life Laboratory, Department of Medical Sciences, Uppsala University, Uppsala, Sweden ³⁵Folkhalsan Research Centre, Helsinki, Finland ³⁶The Department of Social

*Nat Genet. Author manuscript; available in PMC 2014 June 10.*
Services and Health Care, Jakobstad, Finland 37KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway 38Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway 39Department of Biomedicine, University of Bergen, Bergen, Norway 40Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark 41Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark 42Faculty of Medicine, University of Aalborg, Aalborg, Denmark 43Steno Diabetes Center, Gentofte, Denmark 44Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA 45Cardiovascular Research Center, Cardiology Division, Massachusetts General Hospital, Boston, MA, USA 46Department of Medicine, Harvard Medical School, Boston, MA, USA 47National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK 48Department of Clinical Sciences, Diabetes and Endocrinology, Lund University and Lund University Diabetes Centre, Malmö, Sweden 49Department of Medicine, University of Eastern Finland, Kuopio Campus and Kuopio University Hospital, Kuopio, Finland 50Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Cheongwon-gun, Gangoе-myеon, Yeонje-ри, Korea 51Department of Medical Sciences, Uppsala University, Uppsala, Sweden 52Department of Clinical Experimental Research, Glostrup University Hospital, Glostrup, Denmark 53Institute of Human Genetics, Technical University Munich, Munich, Germany 54Department of Genetics, University of North Carolina, Chapel Hill, NC, USA 55The Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway 56Department of Pathology, Haukeland University Hospital, Bergen, Norway 57Department of Biostatistics, University of Liverpool, Liverpool, UK 58Applied Quantitative Genotherapeutics, Pfizer Inc., South San Francisco, CA, USA 59Kuopio Research Institute of Exercise Medicine, Kuopio, Finland 60Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland 61THL-National Institute for Health and Welfare, Helsinki, Finland 62Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore, Singapore 63Department of Medicine, National University of Singapore, National University Health System, Singapore, Singapore 64Duke-National University of Singapore Graduate Medical School, Singapore, Singapore 65Centre for Molecular Epidemiology, National University of Singapore, Singapore, Singapore 66Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore 67Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore 68Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore 69Faculty of Medicine, University of Iceland, Reykjavik, Iceland 70Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland 71Centre for Vascular Prevention, Danube-University Krems, Krems, Austria 72Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland 73King Abdulaziz University, Jeddah, Saudi Arabia 74Department of

*Nat Genet.* Author manuscript; available in PMC 2014 June 10.
Pharmacology, The University of Pennsylvania - Perelman School of Medicine, Philadelphia, PA, USA 76Department of Genetics, The University of Pennsylvania - Perelman School of Medicine, Philadelphia, PA, USA 77Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA 78Oxford Centre for Diabetes, Endocrinology, and Metabolism, Churchill Hospital, University of Oxford, Oxford, UK 79Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK 80Department of Pediatrics, Haukeland University Hospital, Bergen, Norway 81The Go-T2D Consortium; Full lists of members and affiliations are provided in the Supplementary Note 82The T2D-GENES Consortium; Full lists of members and affiliations are provided in the Supplementary Note 83Finnish Institute for Molecular Medicine (FIMM), Helsinki University, Helsinki, Finland 84Department of Genetics, Harvard Medical School, Boston, MA, USA
85Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Acknowledgments

This manuscript is dedicated to the memory of David R. Cox, our dear friend and colleague, who was relentlessly supportive of this work – and more generally the use of human genetics to improve human health. He is missed, but his legacy goes on.

We gratefully acknowledge the contribution of all ~150,000 participants from the various population studies that contributed to this work.

Funding information

JF was supported in part by NIH Training Grant 5-T32-GM007748-33. DA was supported by funding from the Doris Duke Charitable Foundation (2006087). NLB was supported by a Fulbright Diabetes UK Fellowship (BDA 11/0004348).

This work was supported in part by funding to the Broad Institute (principal investigator, DA) from Pfizer Inc. Funding for the GoT2D and T2D-GENES studies was provided by grants 5U01DK085526 (Multiethnic Study of Type 2 Diabetes Genes), DK088389 (Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes), and U54HG003067 (Large Scale Sequencing and Analysis of Genomes), as well as NIH U01’s DK085501, DK085524, DK085545, and DK085584. The Malmö Preventive Project and the Scania Diabetes Registry were supported by grants from the Swedish Research Council (Dnr 521-2010-3490 to LG and Dnr 349-2006-237 to the Lund University Diabetes Centre) as well as by an ERC grant (GENETARGET T2D, GA269045) and two EU grants (ENGAGE [2007-201413] and CEED3 [2008-223211]) to LG. The Botnia study was supported by funding from Sigrid Juselius Foundation and Folkhälsan Research Foundation. PRN was funded by the European Research Council (AdG 293574), Research Council of Norway (197064/V50), KG Jebsen Foundation, University of Bergen, Western Norway Health Authority, EASD Sabbatical Leave Programme and Innovest. The Danish studies were supported by the Lundbeck Foundation (The Lundbeck Foundation Centre for Applied Medical Genomics in Personalised Disease Prediction, Prevention and Cure [LuCamp], www.lucamp.org) and The Danish Council for Independent Research. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk). The PIVUS/ULSAM cohort was supported by Wellcome Trust Grants WT098017, WT064890, WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council and the Swedish Heart-Lung Foundation. The METSIM study was supported by the Academy of Finland (contract 124243), the Finnish Heart Foundation, the Finnish Diabetes Foundation, Tekes (contract 1510/3106), the Commission of the European Community (HEALTH-F2-2007-201681) and R01DK062370, R01DK072193 and Z01HG000024. The FUSION study was supported by R01DK062370, R01DK072193 and Z01HG000024. The DR’s EXTRA Study was supported by the Ministry of Education and Culture of Finland (627;2004–2011), Academy of Finland (102318; 123885), Kuopio University Hospital, Finnish Diabetes Association, Finnish Heart Association, Päivikki and Sakari Sohlberg Foundation and by grants from European Commission FP6 Integrated Project (EXGENESIS); LSHM-CT-2004-005272, City of Kuopio and Social Insurance Institution of Finland (4/26/2010). VSalomaa is funded by the Academy of Finland, grant number 139635, and the Finnish Foundation for Cardiovascular Disease. Sequencing and genotyping of British individuals

Nat Genet. Author manuscript; available in PMC 2014 June 10.
was supported by Wellcome Trust funding WT090367, WT090532, WT098381, and NIDDK U01-DK085545. Funding for the Jackson Heart Study (JHS) was provided by the NHLBI and the National Institute on Minority Health and Health Disparities (N01-HC-95170, N01-HC-95171 and N01-HC-95172). APM acknowledges support from Wellcome Trust grants WT098017, WT090532, and WT064890. FVS and HS were supported by the EU 7th Framework Programme: DIAPREPP (Diabetes type 1 Prediction, Early Pathogenesis and Prevention, grant agreement 202013) and The Swedish Child Diabetes Foundation (Barndiabetesfonden)

References


Nat Genet. Author manuscript; available in PMC 2014 June 10.


Figure 1. Over-expression of p.Arg138X- and p. Ser34fsX50-ZnT8 in HeLa cells
We sought to experimentally evaluate whether the p.Arg138X or p. Ser34fsX50 ZnT8 variants resulted in decreased ZnT8 expression and/or activity. (a) Depiction of SLC30A8 open reading frames in C-terminal V5-tagged constructs (tag highlighted in green). (b) Western blot of HeLa lysates following transient over-expression of V5-tagged ZnT8 variants (anti-V5-tag). Antibody against tubulin was used as a loading control for each sample, and untransfected cell lysate was used to demonstrate specificity of anti-V5 antibody. (c, d) Immunofluorescent staining of ZnT8 variant expression in (c) HeLa and (d) Ins1e cells. ZnT8 was detected using antibodies against the C-terminal V5-tag (anti-V5) or the N-terminus of the endogenous protein (anti-ZnT8), as indicated. BFP-V5 and
untransfected HeLa cells serve as controls. Cells were co-stained with Hoechst-33342 to mark nuclei. Within each row of images for the indicated antibody and objective, identical exposure times were used across all proteins. (e) ZnT8 variant expression, as detected by anti-V5 immunostaining, following 4hr treatment with inhibitors of the lysosome (chloroquine, 100 μM) or the proteasome (MG132, 10 μM). Images were acquired using a 10x objective and identical exposure times. Scale bars, 100 μM.
Figure 2. Protein-truncating variants identified in SLC30A8

Through sequencing and genotyping of nearly 150,000 individuals across 5 ethnicities, we identified 12 SLC30A8 variants – each rare and predicted to cause premature protein truncation. (a) Shown is the position of each variant on the islet-specific SLC30A8 transcript (NM_173851). p.Met50Ile is predicted to alter the initiator codon in other transcripts of SLC30A8. Lines are drawn from each variant to ethnicities for which carriers were observed, with greater widths corresponding to ethnicities with more observations. Lines are further drawn from each ethnicity to the populations (cohorts) from which carriers were identified. From left, cohorts are: JHS, WFS, Botnia, Danish, deCODE, Finnish, HUNT2, KORA, Malmo, PIVUS/ULSAM, WTCCC, LOLIPOP, Singapore Indians, and KARE (cohort information in Supplementary Information). Ethnicities or cohorts with no observations are not shown. (b) Graphical representation of the case and control frequencies for each observed variant; case frequencies in red (above) and control frequencies in blue (below). Wider bars correspond to variants with more observations. A quantitative and complete representation of these data is given in Table 1.
### Table 1

**Association of SLC30A8 variants with type 2 diabetes**

Through sequencing and genotyping of ~150,000 individuals across five ethnicities, a spectrum of 12 rare, predicted protein-truncating variants were identified in *SLC30A8*. Shown for each variant are: ethnicity, cohort, number of genotyped cases and controls (N), number of cases and controls observed to carry a variant (Carriers), and observed allele frequencies in cases and controls (Allele Frequency). Odds ratios (OR) and p-values were computed separately for three groups of variants: p.Arg138X, p.Lys34SerfsX50, and the remaining variants. For p.Arg138X and p.Lys34SerfsX50, for which more than ten carriers were observed, statistics were computed separately for each cohort (see Methods, Supplementary Information) and then combined via a fixed-effects meta-analysis. For the remaining variants, an score of association was computed by comparing the aggregate frequency of variant carriers between cases and controls. These three statistics were combined via a random-effects meta-analysis to produce combined estimates of risk and statistical significance (bottom row). Variant counts and frequencies were computed based on all studied individuals, while OR and p-values were computed with correction for sample structure (population stratification and genetic relatedness; see Supplementary Information); thus, displayed ORs differ from those computed solely from frequency estimates.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Ethnicity</th>
<th>Country</th>
<th>Cohort</th>
<th>N</th>
<th>Carriers</th>
<th>Allele Frequency</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Case</td>
<td>Ctrl</td>
<td>Case</td>
<td>Ctrl</td>
<td>Case</td>
</tr>
<tr>
<td>p.Arg138X</td>
<td>European</td>
<td>Finland</td>
<td>Botnia</td>
<td>3,727</td>
<td>5,440</td>
<td>9</td>
<td>39</td>
<td>0.12%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Sweden</td>
<td>Malmo</td>
<td>6,960</td>
<td>5,480</td>
<td>2</td>
<td>3</td>
<td>0.014%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Sweden</td>
<td>PIVUS/ULSAM</td>
<td>270</td>
<td>1,734</td>
<td>1</td>
<td>3</td>
<td>0.19%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Denmark</td>
<td>Danish</td>
<td>3,889</td>
<td>7,869</td>
<td>0</td>
<td>9</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Finland</td>
<td>Finnish</td>
<td>4,050</td>
<td>8,696</td>
<td>1</td>
<td>2</td>
<td>0.012%</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>Singapore</td>
<td>Singapore Indians</td>
<td>562</td>
<td>585</td>
<td>1</td>
<td>1</td>
<td>0.089%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>UK</td>
<td>UKT2D</td>
<td>321</td>
<td>319</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p.Lys34SerfsX50</td>
<td>European</td>
<td>Iceland</td>
<td>deCODE</td>
<td>2,953</td>
<td>67,919</td>
<td>2</td>
<td>248</td>
<td>0.034%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Norway</td>
<td>HUNT2</td>
<td>1,645</td>
<td>4,069</td>
<td>0</td>
<td>3</td>
<td>0.0%</td>
</tr>
<tr>
<td>c.71+2T&gt;A</td>
<td>African-American</td>
<td>USA</td>
<td>WFS</td>
<td>501</td>
<td>527</td>
<td>1</td>
<td>0</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>African-American</td>
<td>USA</td>
<td>JHS</td>
<td>530</td>
<td>533</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p.Met50lle</td>
<td>European</td>
<td>Germany</td>
<td>KORA</td>
<td>97</td>
<td>91</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>c.271+G&gt;A</td>
<td>East Asian</td>
<td>Korea</td>
<td>KARE</td>
<td>520</td>
<td>551</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>Singapore</td>
<td>Singapore Indians</td>
<td>562</td>
<td>585</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>c.419-1G&gt;C</td>
<td>South Asian</td>
<td>UK</td>
<td>LOIPOP</td>
<td>530</td>
<td>537</td>
<td>1</td>
<td>0</td>
<td>0.094%</td>
</tr>
<tr>
<td>p.Trp152X</td>
<td>European</td>
<td>Finland</td>
<td>Botnia</td>
<td>134</td>
<td>180</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>Variant</td>
<td>Ethnicity</td>
<td>Country</td>
<td>Cohort</td>
<td>N</td>
<td>Carriers</td>
<td>Allele Frequency</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
<td>--------------</td>
<td>-----</td>
<td>----------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Case</td>
<td>Ctrl</td>
<td>Case</td>
<td>Ctrl</td>
</tr>
<tr>
<td>p.Arg138X</td>
<td>European</td>
<td>Finland</td>
<td>Botnia</td>
<td>3,727</td>
<td>540</td>
<td>9</td>
<td>39</td>
<td>0.12%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Sweden</td>
<td>Malmö</td>
<td>6,960</td>
<td>5,480</td>
<td>2</td>
<td>3</td>
<td>0.014%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Sweden</td>
<td>PIVUS/ULSAM</td>
<td>270</td>
<td>1,734</td>
<td>1</td>
<td>3</td>
<td>0.19%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Denmark</td>
<td>Danish</td>
<td>3,889</td>
<td>7,869</td>
<td>0</td>
<td>9</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>Finland</td>
<td>Finnish</td>
<td>4,050</td>
<td>8,696</td>
<td>1</td>
<td>2</td>
<td>0.012%</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>Singapore</td>
<td>Singapore Indians</td>
<td>562</td>
<td>585</td>
<td>1</td>
<td>1</td>
<td>0.089%</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>UK</td>
<td>UKT2D</td>
<td>321</td>
<td>319</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p.Gln174X</td>
<td>South Asian</td>
<td>UK</td>
<td>LOLIPOP</td>
<td>530</td>
<td>537</td>
<td>1</td>
<td>5</td>
<td>0.094%</td>
</tr>
<tr>
<td>c.572+1G&gt;A</td>
<td>African-American</td>
<td>USA</td>
<td>JHS</td>
<td>530</td>
<td>533</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p.Tyr284X</td>
<td>South Asian</td>
<td>UK</td>
<td>LOLIPOP</td>
<td>530</td>
<td>537</td>
<td>0</td>
<td>2</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>South Asian</td>
<td>Singapore</td>
<td>Singapore Indians</td>
<td>562</td>
<td>585</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p.Ile291PhefsX2</td>
<td>African-American</td>
<td>USA</td>
<td>JHS</td>
<td>530</td>
<td>533</td>
<td>0</td>
<td>1</td>
<td>0.0%</td>
</tr>
<tr>
<td>p Ser327ThrfsX5</td>
<td>African-American</td>
<td>USA</td>
<td>WFS</td>
<td>501</td>
<td>527</td>
<td>0</td>
<td>2</td>
<td>0.0%</td>
</tr>
<tr>
<td>Combined</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>30,433</td>
<td>118,701</td>
<td>19</td>
<td>326</td>
<td>-</td>
</tr>
</tbody>
</table>

CI = confidence interval; Ctrl = Control.