Decreased STARD10 Expression Is Associated with Defective Insulin Secretion in Humans and Mice

Gaelle R. Carrat,1 Ming Hu,1 Marie-Sophie Nguyen-Tu,1 Pauline Chabosseau,1 Kyle J. Gaulton,2,3 Martijn van de Bunt,2,4 Afshan Siddiq,5 Mario Falchi,5,6 Matthias Thurner,2,4 Mickaël Canouil,7,8 Francois Pattou,7,8 Isabelle Leclerc,1 Timothy J. Pullen,1 Matthew C. Cane,1 Priyanka Prabhala,1 William Greenwald,2,3 Anke Schulte,9 Piero Marchetti,10 Mark Ibberson,11 Patrick E. MacDonald,12 Jocelyn E. Manning Fox,12 Anna L. Gloyn,2,4,13 Philippe Froguel,5,7,8 Michele Solimena,14,15,16 Mark I. McCarthy,2,4,13 and Guy A. Rutter1,*

Genetic variants near ARAP1 (CENTD2) and STARD10 influence type 2 diabetes (T2D) risk. The risk alleles impair glucose-induced insulin secretion and, paradoxically but characteristically, are associated with decreased proinsulin:insulin ratios, indicating improved proinsulin conversion. Neither the identity of the causal variants nor the gene(s) through which risk is conferred have been firmly established. Whereas ARAP1 encodes a GTPase activating protein, STARD10 is a member of the sterol regulatory element-binding protein (SREBP)-related lipid transfer protein family. By integrating genetic fine-mapping and epigenomic annotation data and performing promoter-reporter and chromatin conformational capture (3C) studies in β cell lines, we localize the causal variant(s) at this locus to a 5 kb region that overlaps a stretch-enhancer active in islets. This region contains several highly correlated T2D-risk variants, including the rs140130268 indel. Expression QTL analysis of islet transcripomes from three independent subject groups demonstrated that T2D-risk allele carriers displayed reduced levels of STARD10 mRNA, with no concomitant change in ARAP1 mRNA levels. Correspondingly, β-cell-selective deletion of StarD10 in mice led to impaired glucose-stimulated Ca2+ dynamics and insulin secretion and recapitulated the pattern of improved proinsulin processing observed at the human GWAS signal. Conversely, overexpression of StarD10 in the adult β cell improved glucose tolerance in high fat-fed animals. In contrast, manipulation of Arap1 in β cells had no impact on insulin secretion or proinsulin conversion in mice. This convergence of evidence and murine data provides compelling evidence that the T2D risk associated with variation at this locus is mediated through reduction in STARD10 expression in the β cell.

Introduction

Normal glucose homeostasis depends on the correct processing of proinsulin and the storage of the mature hormone within secretory granules in the pancreatic β cell.1 Stimulation of insulin secretion by glucose involves the uptake and metabolism of the sugar via glucose transporters (Glut2 and/or Glut1),2 phosphorylation by glucokinase,3 and enhanced mitochondrial ATP synthesis.4,5 Closure of ATP-sensitive K+ channels (KATP),6 plasma membrane depolarization, and Ca2+ influx then prompt the fusion of insulin-containing secretory granules with the plasma membrane. Additional, KATP-Channel-independent mechanisms8 also sensitize the secretory machinery to Ca2+.9 Changes in both the number of β cells10 and in the ability of these cells to respond to glucose11 are involved in the development of type 2 diabetes (T2D [MIM: 125853]) in the face of insulin resistance.12 In addition to lifestyle factors including obesity (MIM: 601665),13 genetic factors also contribute substantially to overall T2D risk.14,15 Genome-wide association studies have identified more than 90 loci associated with type 2 diabetes risk.14,15 In the majority of cases the identified polymorphisms, which usually affect β cell function, lie in intronic or intergenic regions, and neither the identity of the responsible gene(s) nor the mechanism of action is clear.15,16

T2D is typically characterized by a disruption of proinsulin conversion17,18 and carriers of T2D-risk alleles generally display increased proinsulin:insulin ratios compared with those who are homozygous for the protective allele.19 By contrast, the T2D-risk alleles at a locus adjacent to ARAP1 (MIM: 606646) (formerly called CENTD2) and STARD10 on chromosome 11q1320,21 are robustly associated with a
Figure 1. Identification of Likely Causal Variants at the ARAP1/STARD10 Locus

(A and B) Quantitative trait locus association between variants at ARAP1/STARD10 and STARD10 (A) or ARAP1 (B) expression level in human islets. Variants are strongly associated with STARD10 expression level but not ARAP1 level. Colors in each plot represent the extent of linkage disequilibrium between each tested variant and the T2D index variant.

(legend continued on next page)
marked reduction in proinsulin:insulin ratios.\textsuperscript{19,22} This unusual pattern implies preserved or improved proinsulin processing despite increased T2D risk.

\textit{ARAP1} encodes ARF-GAP, Rho-GAP, ankyrin repeat and pleckstrin homology domain-containing protein 1, or centaurein delta 2, an ArfGAP (GTPase activating protein) regulated by phosphatidylinositol 1,4,5-trisphosphate. ARAP1 appears to act on ARF6 (MIM: 600464),\textsuperscript{23} a known regulator of insulin secretion.\textsuperscript{24} \textit{StarD10} (previously termed also apparent in both human islets\textsuperscript{19} and is also highly abundant (second centile of the genes close to the index SNP rs1552224 in \textit{StarD10}).\textsuperscript{24} \textit{StarD10} is strongly expressed in testes, liver, and kidney, but much more weakly in other tissues involved in insulin action and glucose metabolism such as skeletal muscle.\textsuperscript{26} \textit{StarD10} is the most strongly expressed of the genes close to the index SNP rs1552224 in human islets\textsuperscript{19} and is also highly abundant (second centile of mRNAs) in mouse islets, where it is the most highly expressed StarD family member.\textsuperscript{27} \textit{StarD10} expression is also apparent in both human \(\alpha\) and \(\beta\) cells, with similar mRNA levels in each cell type, and \textit{ARAP1} is also expressed in both cell types albeit at levels approximately one-third those of \textit{StarD10}.\textsuperscript{28} Although global inactivation of \textit{StarD10} in mice has previously implicated this protein in bile acid homeostasis,\textsuperscript{29} its role in glucose homeostasis is unknown.

Recent expression quantitative trait loci (eQTL) studies in normoglycemic donors\textsuperscript{30} have suggested that islet \textit{StarD10} expression is correlated with T2D risk variants at this locus, whereas no such islet eQTL association was observed for \textit{ARAP1}. These findings contrast with other results\textsuperscript{31} reporting higher \textit{ARAP1} mRNA synthesis from the T2D risk allele. Expression of other nearby genes at this locus—\textit{FCHSD2} (MIM: 611565), \textit{ATG16L2} (Figure 1C), and \textit{PDE2A} (MIM: 602658)—is relatively low in human islets.\textsuperscript{19,28}

Here, we show first that disease-associated variants in this locus are associated with \textit{StarD10}, but not \textit{ARAP1}, mRNA levels in human pancreatic islets ascertained from both diabetic and non-diabetic individuals. Using genetic and genomic fine mapping and functional analysis in \(\beta\) cells, we identify a region in intron 2 of \textit{StarD10} containing several variants that is likely to mediate T2D risk at this locus. Finally, we generate and characterize a series of mouse strains overexpressing, or inactivated selectively in the adult \(\beta\) cell, for \textit{StarD10} or \textit{Arap1}. These analyses reveal that \textit{StarD10} is required for normal insulin secretion, though its deletion enhances proinsulin processing.

**Material and Methods**

**Materials**

cDNAs encoding human full-length \textit{ARAP1} and mouse \textit{StarD10} were purchased from Genscript and OriGene, respectively. Cell culture medium was from Sigma and fetal bovine serum (FBS) from SeraLab.

**Identification of Causal Variants using MetaboChip and Functional Priors**

We derived causal probabilities for each variant in the following way. We first obtained fine-mapping data of variants at 39 T2D loci (including \textit{ARAP1}) from the Metabochip.\textsuperscript{32} For each of the 49 distinct association signals at the 39 loci, we calculated the Bayesian posterior odds for all variants at each signal using the approach of Wakefield.\textsuperscript{33} Previous studies have demonstrated that sets of T2D risk locus share patterns of functional regulatory annotation in specific cell types and that this information can be used to help prioritize causal variants.\textsuperscript{32,34–37} We thus obtained regulatory chromatin state data previously generated in 12 cell types which included 9 ENCODE cell types (Gm12878, HepG2, HUVEC, Hsmm, hESC, Hmec, NHLF, NHEK, and K562), pancreatic islets (PancIslt), and pre- and mature- adipocytes (hASC1, hASC4).\textsuperscript{32,34,38,39} For each cell type, we used active enhancer (EnhA), weak enhancer (EnhWk), and active promoter (TssA) elements for a total of 36 chromatin annotations. For active enhancer elements, we further defined “stretch” enhancers using a previously described definition of active enhancers greater than 3 kb in size.\textsuperscript{15}

We modeled the effect of these 36 annotations on the posterior odds of variants at the 39 T2D loci using fgwas.\textsuperscript{40} In this procedure we first iteratively added annotations that increased the likelihood of the model. With this joint model, we selected the penalty with the highest penalized cross-validation likelihood. Using the optimal penalty, we maximized the cross-validation likelihood by iteratively removing annotations from the model. We then used the enrichment estimates of each annotation from this final model as functional priors to update the posterior odds for each variant at the \textit{ARAP1/StarD10} locus. We finally calculated the posterior causal probability of each variant from these updated posterior odds.

**Chromatin Accessibility Analysis**

A total of 23 human islet samples were freshly isolated at the Oxford Centre for Islet Transplantation as described previously\textsuperscript{30} and stored for 1–3 days in CMRL or in UW media. The latter were reactivated in CMRL for 1 hr before processing them further. Assay for transposase accessible chromatin (ATAC-seq) was performed on these 23 primary pancreatic islet samples as previously... (C) Top: probability that each variant at \textit{ARAP1/StarD10} is causal for T2D risk using Metabochip fine-mapping and functional priors derived from chromatin state maps of 12 cell types from ENCODE, islets, and adipose tissue.\textsuperscript{15} The indel rs140130268 has the highest causal probability (\(r_e = 42\%\)). Middle: chromatin states for each of the 12 cell types colored by state (abbreviations: Quies, quiescent; TssA, active promoter; EnhWk, weak enhancer; EnhA, active enhancer; Ins, insulator; Txs, transcription; Repr, repressed; TssP, poised promoter). Variants with highest causal probabilities, including rs140130268, fall in a stretch-enhancer region (dark blue) active in islets and largely inactive in other cell types. Variants highlighted in red were characterized individually in functional (promoter-luciferase) assays, those in green were analyzed in combination (see Results). Bottom: islet RNA-seq expression level of each gene transcript in the region.
described. To remove primer dimers, the amplified libraries were additionally purified with Agencourt AMPure beads. Samples were multiplexed using primers Ad2.1-6 and paired-end sequencing using Illumina HiSeq 2500. Raw FASTQ reads were processed with a departmental/in-house pipeline and on the DNase and ChiP pipeline website (Web Resources). Specifically, library/sequencing quality was checked with FASTQC (Web Resources) and reads were mapped to the human genome hg19 via bowtie. For reads that could not be aligned the first time, adapters were removed at the 3’ end with Trim Galore (Web Resources). The resulting trimmed reads were then mapped again with bowtie. Any remaining unmapped and trimmed reads were processed with FLASH and re-mapped a third time with bowtie. For each fine-mapped variant at the ARAP1/STARD10 locus, we then re-mapped reads in the region against allele-specific reference genomes using WASP. We then retained variants with at least five overlapping reads at the variant base from ≥2 different heterozygote samples. For the four resulting variants, we then tested for imbalance in the pooled read counts for each allele from heterozygote samples using a binomial test.

Expression Quantitative Trait Locus Analysis

IMIDIA Samples

Expression data were acquired and normalized from islet organ donors (81 healthy, 19 T2D) or partial pancreatectomy-laser microdissection samples (32 non-diabetic, 35 T2D) from the IMIDIA consortium (Web Resources; M. Solimena, personal communication) with appropriate permissions from donors and/or families. In brief, genotyping analysis was performed on the DNA from the same subjects using HumanOmmi 2.5-8 beadchip from Illumina using a standard Infinium genotyping protocol. Standard quality control assessment was carried out on the genotyping data using PLINK (Web Resources). cis-eQTL analysis was performed with a freely available program called matrix eQTL. Linear model was used as a parameter for the analysis with age and gender as covariates. A “window size” of 100 kb was used as the cisDist; this distance represents the maximum distance at which gene-SNP pair is considered local. Only cis-eQTL results for STAR1D0 and ARAP1 are described in the present manuscript.

Oxford & Edmonton Samples

RNA-seq was performed on 174 human islet preparations collected in Oxford, UK, and Edmonton, Canada (an extension of the data reported in van de Bunt et al.). In brief, samples were genotyped on Illumina HumanOmmi2.5-Exome beadchips followed by imputation from the 1000 Genomes phase 3 panel using SHAPEIT2 and IMPUTE2. Raw RNA-seq reads were aligned to the human genome reference hg19 with STAR and expression quantified at the exon-level. Read count data was normalized to 20M reads, with exons with expression <1 count in >20% of all samples excluded. This was followed by rank normalization per exon, after which 30 hidden factors (accounting for non-genic variability in the samples) were inferred from the count matrix using PEER. cis-eQTL analysis for STAR1D0 and ARAP1 was performed in a window flanking 1 Mb either side of the transcriptional start site using linear regression (with 30 PEER factors as covariates) implemented in FastQTL, with p values adjusted through beta-approximation.

Liver Biopsies

Genotyping was performed using Illumina Metabochip DNA arrays. mRNA levels were measured using the Illumina HumanHT-12 Expression Beadchip. cis-eQTL analysis was performed under R (v.3.3.1) using standard linear regression adjusted for age and BMI as implemented in FastQTL, setting a maximum distance from the SNP location of 500 kb.

Animals and Ethics

All in vivo procedures were approved by the UK Home Office according to Animals (Scientific Procedures) Act 1986 (HO License PPL 70/7349) and were performed at the Central Biomedical Service, Imperial College, London, UK. Animals were housed 2 to 5 per individually ventilated cage in a pathogen-free facility with 12 hr light-dark cycle and had free access to standard mouse chow diet unless otherwise stated. For high-fat diet treatment, mice were placed on a high-fat diet at 5 weeks of age (DIO Rodent Purified Diet w/60% energy from fat; Test Diet). Human islet samples were obtained with appropriate local and ethical approval and consent from next of kin as required.

Generation of Transgenic Mice

Human ARAP1 and murine StarD10 coding sequences were amplified from ARAP1-pcDNA3.1+ and StarD10-pCMV-entry6, respectively, with the addition of a single NH2 Flag tag by PCR, and inserted between the Nhel and Xhol sites of the plasmid pB-L Tet (Clontech). The resulting plasmid carried a bidirectional tetracycline-regulated promoter driving expression of both Flags: ARAP1 or Flag::StarD10 and firefly luciferase.

The above expression cassette was excised from the plasmid backbone by AatII and Asel digestion and transferred by pronuclear microinjection into C57BL/6J mouse oocytes. Successful integrants were identified by PCR screening, RIP7-rtTA mice on a C57BL/6 background, expressing the reverse tetracycline transactivator under the control of the rat insulin promoter, were crossed with transgenic mice to permit β-cell-specific, tetracycline-inducible expression of the transgene and luciferase. Heterozygous transgenic mice were crossed to homozygous RIP7-rtTA mice to produce littermates of two genotypes: single transgenic (control, RIP7-rtTA/tet) and double transgenic (ARAP1 or StarD10-tg, RIP7-rtTA/ transgene-Luc).

Generation of StarD10- and Arap1-Null Mice

StarD10 whole body and conditional knockout (KO) mice (C57BL/6NTac background) were generated by the trans-NIH Knock-Out Mouse Project (KOMP) and obtained from the KOMP Repository (Web Resources) via the international mouse phenotyping consortium (IMPC). Mice homozygous for floxed StarD10 (StarD10fl/fl) or Arap1 (Arap1fl/fl) alleles were crossed to mice expressing Cre recombinase from the endogenous Ins1 locus (Ins1-Cre mice). This generated StarD10fl/fl:Ins1Cre−/− (bStarD10 KO) mice, where exon 3 was removed selectively by Cre-mediated excision in pancreatic β cells, or Arap1fl/fl:Ins1Cre−/− (bArap1 KO) mice, where exon 12 of Arap1 was removed. Genotyping was performed by PCR of DNA extracted from ear biopsies by the HotSHOT method (primer sequences in Table S1). Ablation of gene expression from pancreatic islets was assessed by real-time quantitative PCR (qPCR) on islet RNA and western (immuno-) blotting.

In Vivo Physiological Studies

All studies were performed on male mice except when data were comparable between genders, in which case results from males and females were combined to gain statistical power, as indicated.
Intraperitoneal Glucose Tolerance Test
Mice fasted overnight (16 hr) were intraperitoneally injected at -10 am with glucose, 1 g/kg mouse weight. Blood from the tail vein was obtained at 0, 15, 30, 45, 60, 90, and 120 min after injection. Blood glucose levels were measured with the Accu-Chek Aviva glucometer (Roche).

Plasma Insulin Measurement
Mice fasted overnight (16 hr) were intraperitoneally injected at -10 am with 3 g glucose/kg mouse weight. Blood from the tail vein was collected into heparin-coated tubes (microvette) at 0, 15, and 30 min after injection. Plasma was separated by centrifugation at 2,000 x g for 5 min. Plasma insulin levels were measured using an ultrasensitive mouse insulin ELISA kit (Crystal Chem).

Insulin Tolerance Test
Human insulin (Actrapid, Novo Nordisk) was intraperitoneally injected at 3 pm into mice fasted for 5 hr. Blood glucose levels were measured at 0, 15, 30, 45, 60, 90, and 120 min after injection. The quantities of insulin injected were 0.5 U/kg (females, Chow diet), 0.75 U/kg (males, Chow diet and females, high fat diet), or 1 U/kg (males, high fat diet).

Proinsulin Measurement
Blood from the tail vein was collected into heparin-coated tubes from female and male mice fasted 5 hr. Plasma proinsulin levels were measured using a rat/mouse proinsulin ELISA kit (Mercodia).

Islet Isolation and Culture
Mouse islets were isolated after collagenase digestion (Collagenase NB8 Broad Range, 1 mg/mL, Serva Electrophoresis) and subsequently cultured in RPMI1640 medium, containing 11 mM glucose (Sigma) and supplemented with 10% (v/v) fetal bovine serum plus penicillin (100 units/mL) and streptomycin (0.1 mg/mL) at 37°C in an atmosphere of humidified air (95%) and CO2 (5%), for 2 or 3 days prior to experiments.

Insulin Secretion from Mouse Islets
Islets (10/well) were incubated in triplicate for each condition and treatment. Islets were pre-incubated for 1 hr in 3 mM glucose Krebs-Ringer-HEPES-Bicarbonate (KRH) buffer prior to secretion assay (30 min) in 3 mM, 17 mM glucose, or 30 mM KCl. The supernatant were collected and the islets were lysed in 1 mL of acidified ethanol to measure total insulin content. The insulin concentration was measured by radioimmunoassay (LINCO/Millipore) or using an HTRF kit (Cisbio Bioassays).

qRT-PCR
Total cellular RNA from mouse islets or other tissues was obtained using TRIzol reagent (Invitrogen) and reverse transcribed with a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's instructions. Real-time PCR was performed on an ABI-Prism Fast 7500 device (Applied Biosystems) using the Fast SYBR Green Master Mix (Applied Biosystems) according to the manufacturer's instructions. All constructs were subjected to DNA sequencing.

Chromosome Conformation Capture
Fluorescence in situ hybridization (FISH) analysis of the human insulin promoter region was performed using a probe for the human insulin gene (5′-GCTGGTGTGG-3′) and reverse: 5′-ACTGAGTGTCACGCTGCGTCGACGGTGG-3′. The probe was labeled with both FAM and TAMRA. The individual primers covering variant regions were: for NcoI fragment (Q5 site-directed mutagenesis) was carried out by a PCR-based method (QS site-directed mutagenesis kit, New England Biolabs) according to the manufacturer's instructions. All constructs were subjected to DNA sequencing.

Luciferase Assay
Luciferase constructs containing variant DNA fragments of 500–600 bp in each case were co-transfected, using Lipofectamine 2000 (Life Technologies), with CMV-Renilla construct as internal control, into INs1 (832/13) cells according to manufacturer's instruction. After 48 hr, transfected cells were washed once with PBS and lysed directly in passive cell lysis buffer from Luciferase Assay System (Promega). Cells were incubated on a rotating platform at room temperature for 10 min to ensure complete lysis of cells, and then spun at 10,000 rpm for 1 min to remove cell debris. Luciferase activity was determined with the Dual-luciferase Reporter System on a Lumat LB9507 luminometer (Berthold Technologies).

Identification and Cloning of the ARAP1/STARD10 Variant-Bearing Regions
Genomic DNAs from HEK293 cells were extracted using DNeasy blood and tissue kit according to manufacturer's instruction (Qiagen). PCR reactions using Phusion high-fidelity DNA polymerase (New England Biolabs) were carried out to amplify genomic DNA fragments carrying variants using the primer sets as follow: rs148527516: forward: 5′-GCTGGTGTGGCTGCGTCGACGGTGG-3′ and reverse: 5′-GCAAAGGATGTGGCTGCGTCGACGGTGG-3′; rs140130268: forward: 5′-GCTGGTGTGGCAGCTGCGTCGACGGTGG-3′ and reverse: 5′-GCAAAGGATGTGGCTGCGTCGACGGTGG-3′; rs3862791: forward: 5′-ACTGAGGATCGCGTGGCGAAGAGATCGAAGG-3′ and reverse: 5′-ACTGAGGATCGCGTGGCGAAGAGATCGAAGG-3′; rs7103836: forward: 5′-ACGAGGCCGACGCTGCGTCGACGGTGG-3′; rs7103836: forward: 5′-ACGAGGCCGACGCTGCGTCGACGGTGG-3′; rs7103836: forward: 5′-ACGAGGCCGACGCTGCGTCGACGGTGG-3′; rs7103836: forward: 5′-ACGAGGCCGACGCTGCGTCGACGGTGG-3′; rs7103836: forward: 5′-ACGAGGCCGACGCTGCGTCGACGGTGG-3′.
locate in the promoter 2 region of ARAP1 and the variants, the constant primer was designed to be 5'0 fragment 13, 5'0 fragment 14, 5'0 fragment 15, 5'0 fragment 16. We first explored the possibility that possession of risk alleles of the human CXCL12 gene at this locus may lead to changes in the expression of STARD10. The Taqman reaction was carried out on fast 9700 PCR machine with Taqman Fast advanced reagents. PCR reactions were set as below: 95°C for 10 min, then with 45 cycles at 95°C 30 s and 58°C 45 s. The real crosslinking frequencies were plotted as percentage of that for NcoI fragment 12.

The individual primers covering the variants regions were: fragment 17, 5'0; fragment 18, 5'0; fragment 19, 5'0; fragment 20, 5'0. The probe was labeled with both FAM and TAMRA. The sequence of the qPCR probe was 5'-ACTTCTGTGAGCTCCCTGAGG-3'. The probe was labeled with both FAM and TAMRA.

The individual primers covering the variants regions were: fragment 21, 5'0; fragment 22, 5'0; fragment 23, 5'0; fragment 24, 5'0. The probe was labeled with both FAM and TAMRA.

TGAGCCC-3'; fragment -6: 5'-CCGATGTCTACCCCTTC-3'; fragment -7: 5'-CCGACCTTTCGCCACAGG-3' and fragment -8: 5'-GCACAGTCTAGAGG-3'. The Taqman reaction was carried out on fast 9700 PCR machine with Taqman Fast advanced reagents. PCR reactions were set as below: 95°C for 10 min, then with 45 cycles at 95°C 30 s and 58°C 45 s. The real crosslinking frequencies were plotted as percentage of that for NcoI fragment 12.

To assess the chromatin association between ARAP1 promoter region and the variants, the constant primer was designed to locate in the promoter 2 region of ARAP1 and its DNA sequence was 5'-ACTTCTGTGAGCTCCCTGAGG-3'. The sequence of the qPCR probe was 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'. The probe was labeled with both FAM and TAMRA. The individual primers covering the variants regions were: for NcoI fragment 12, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; fragment 13, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; fragment 14, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; fragment 15, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; fragment 16, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; fragment 17, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'; and fragment 18, 5'-CCGAGCTCTCCCTGTGGCTCTC CGAG-3'.

Statistical Analysis
Data are expressed as means ± SEM. Significance was tested by Student's two-tailed t-test, Mann-Whitney test for non-parametric data, and one- or two-way ANOVA with SIDAK multiple comparison test, as appropriate, using Graphpad Prism software. p < 0.05 was considered significant.

Results

cis-eQTL Analysis Reveals Association between T2D Risk Variants at the ARAP1/STARD10 Locus and STARD10 but Not ARAP1 Expression
We first explored the possibility that possession of risk alleles at this locus may lead to changes in the expression of STARD10 or ARAP1 in human islets from three separate sources. First, we examined two cohorts from a recently described biorepository from the IMIDIA consortium (M. Solimena, personal communication). This consists of samples from non-diabetic and T2D subjects taken either (1) post-mortem (organ donors; OD) or (2) after partial pancreatectomy for pancreatic disease and laser capture microdissection (PP-LCM) (167 samples in total). Associations were determined between signals from the microarray expression probes indicated in Table 1 and the previously defined SNPs at this locus, rs1552224 and rs11603334, which were used as proxies for the likely causal variants (see below) with which they are in linkage disequilibrium (LD). We note that although cis-eQTLs detected with this approach are likely to reflect associations with β cells, we do not exclude a contribution from other islet endocrine cells to the observed signals.

Significant associations were detected between STARD10 mRNA levels and genotype in both OD and PP-LCM groups, irrespective of the SNP analyzed (Table 1). In all analyses, increased expression was associated with possession of the minor (T2D-protective) allele.19–21 By contrast, no such associations were apparent for ARAP1 expression (Table 1). In RNA-seq data from human islet preparations ascertained from 174 normoglycemic ODs in Oxford and Edmonton (this is an extension of a recently reported sample)30 the lead variants at the locus were full LD proxies for the significant (FDR < 1%) cis-eQTL for STARD10 (Figure 1A), but no association with ARAP1 (Figure 1B) mRNA levels was observed.

In the OD and PP-LCM islets from the IMIDIA samples, it was possible to compare STARD10 and ARAP1 expression levels from islets gathered from T2D (n = 54) and non-diabetic (n = 113) subjects. In both OD and PP-LCM, T2D individuals displayed reduced STARD10 expression; whereas a reduction in ARAP1 expression in T2D was observed in OD but not PP-LCM subjects (Table 2).

Fine Mapping of Variants at the ARAP1/STARD10 Locus
We next used data from a dense fine-mapping study of 39 T2D loci on the MetaboChip, involving 27.2k cases and

The American Journal of Human Genetics 100, 238–256, February 2, 2017 243

Table 1. Association of cis-eQTLs with rs1552224 and rs11603334

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Probe</th>
<th>OD Control</th>
<th>OD All</th>
<th>LCM Control</th>
<th>LCM All</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1552224</td>
<td>STARD10</td>
<td>223103_at</td>
<td>0.26 (0.10)</td>
<td>0.012*</td>
<td>0.30 (0.09)</td>
<td>0.0023*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>232322_x_at</td>
<td>0.16 (0.09)</td>
<td>0.069</td>
<td>0.19 (0.08)</td>
<td>0.025*</td>
</tr>
<tr>
<td></td>
<td>ARAP1</td>
<td>34206_at</td>
<td>0.12 (0.08)</td>
<td>0.11</td>
<td>0.07 (0.08)</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212516_at</td>
<td>0.11 (0.06)</td>
<td>0.10</td>
<td>0.06 (0.06)</td>
<td>0.33</td>
</tr>
<tr>
<td>rs11603334</td>
<td>STARD10</td>
<td>223103_at</td>
<td>0.26 (0.10)</td>
<td>0.013*</td>
<td>0.30 (0.09)</td>
<td>0.0023*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>232322_x_at</td>
<td>0.16 (0.09)</td>
<td>0.067</td>
<td>0.19 (0.08)</td>
<td>0.023*</td>
</tr>
<tr>
<td></td>
<td>ARAP1</td>
<td>34206_at</td>
<td>0.13 (0.08)</td>
<td>0.095</td>
<td>0.07 (0.08)</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>212516_at</td>
<td>0.11 (0.06)</td>
<td>0.093</td>
<td>0.07 (0.06)</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>225883_at</td>
<td>-</td>
<td>-</td>
<td>0.05 (0.11)</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Probe names are from Human Genome U133 Plus 2.0 Array from Affymetrix. p values calculated by linear model, and beta (b) is measuring the effect size estimate. All subjects were corrected for age and gender as covariates in the analysis. Significant values indicated by asterisk (*).
To further distinguish between these 27 candidate variants, we used the fgwas approach \(^4\) to integrate the T2D fine-mapping data with chromatin state maps from 12 human cell types (including islets; see Material and Methods). We determined the degree of enrichment for each enhancer and promoter annotation with respect to regions of stretch-enhancer elements active in pancreatic islets. Of the variants within this set, the rs140130268 indel accounted for almost half of the re-weighted credible set, were carried forward for functional analysis: two other variants, rs3862791 and rs148527516 (Figure 1C), with low probabilities of contributing to disease risk (rs3862791 \(\pi_c = 0.02\), rs148527516 \(\pi_c = 0.0001\)), served as negative controls. Promoter-reporter studies were performed in the insulin-secreting rat INS1(832/13) cell line. Sites of ~0.5 kb around the human variants were PCR amplified and sub-cloned into plasmids downstream of firefly luciferase cDNA expressed under the control of an SV40 promoter (Figure 2B). Co-transfection with a control vector allowed expression to be normalized to that of Renilla luciferase.

Indel rs140130268 (Figure 1C), which had the strongest posterior causal probability, as well as other variants in the credible set, were carried forward for functional analysis: two other variants, rs3862791 and rs148527516 (Figure 1C), with low probabilities of contributing to disease risk (rs3862791 \(\pi_c = 0.02\), rs148527516 \(\pi_c = 0.0001\)), served as negative controls. Promoter-reporter studies were performed in the insulin-secreting rat INS1(832/13) cell line. Sites of ~0.5 kb around the human variants were PCR amplified and sub-cloned into plasmids downstream of firefly luciferase cDNA expressed under the control of an SV40 promoter (Figure 2B). Co-transfection with a control vector allowed expression to be normalized to that of Renilla luciferase.

To next determine whether variants in the 5 kb fine-mapped region influence local chromatin structure, we obtained chromatin accessibility data in primary pancreatic islets using ATAC-seq. We identified four fine-mapped variants in the 5 kb interval that directly overlapped regions of accessible chromatin and thus could serve as markers of allelic activity. For each variant, we tested for allelic imbalance in ATAC-seq signal in samples heterozygous for all of these variants. We identified evidence for imbalance at three variants (rs7103836 \(p = 3.0 \times 10^{-6}\), rs3862791 \(p = 3.5 \times 10^{-3}\), rs76550717, \(p = 0.02\)). At all three variants, the T2D risk-increasing allele was correlated with lower chromatin accessibility, consistent with the correlation of T2D risk alleles with lower \(\pi_c\). This demonstrates that risk alleles of variants in the 5 kb fine-mapped region are correlated with decreased islet chromatin accessibility in islets and implies that one or several variants in this region directly affects local chromatin structure. Note that indel imbalance could not be assessed accurately given asymmetric read and mapping efficiencies.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Probeset</th>
<th>Log Fold Change</th>
<th>p Value</th>
<th>Adjusted p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STARD10_238911_at</td>
<td>-0.27</td>
<td>0.00618</td>
<td>0.0185*</td>
</tr>
<tr>
<td></td>
<td>STARD10_232322_x_at</td>
<td>-0.15</td>
<td>0.00496</td>
<td>0.0149*</td>
</tr>
<tr>
<td></td>
<td>ARAP1_212516_at</td>
<td>-0.25</td>
<td>0.000479</td>
<td>0.00144*</td>
</tr>
<tr>
<td></td>
<td>ARAP1_34206_at</td>
<td>0.10</td>
<td>0.0658</td>
<td>0.263</td>
</tr>
<tr>
<td></td>
<td>ARAP1_225883_at</td>
<td>0.17</td>
<td>0.265</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Data are from organ (OD, \(n = 81\) and 19 normoglycemic and T2D, respectively) and partial pancreatectomy/laser capture microdissection (LCM, \(n = 32,35\)) donors. (M. Solimena, personal communication). Fold changes indicate the rate of expression in T2D versus non-diabetic islets. Significant values indicated by asterisk (*).
enhancer activity in this assay (not shown). By contrast, the T2D-protective (-GTTT) allele at rs140130268 displayed significantly (~40%) higher activity than the T2D-risk form (Figure 2B).

To determine whether the above region was able to physically associate with the STARD10 promoter, chromatin conformation capture (3C) analysis was performed using human EndoC-βH1 cells. STARD10 may be expressed from one of two promoters (P1 and P2) located at the 5’ end of exon 1 or exon 2, respectively. RNA-seq analysis from human islets (Figure 1C) indicates that transcripts from P2 are the most abundant in islets. The results of 3C analysis using the restriction enzyme NcoI are presented in Figure 2A. Cross-linking frequencies were observed at the genomic DNA fragment (NcoI fragment –4) carrying rs140130268 as well as surrounding regions (NcoI fragments –4 and –7) demonstrating that STARD10 P2 interacts physically with the 5 kb fine-mapped interval and may thus be impacted by variation at rs140130268 as well as other credible set variants.

A similar experiment was performed with ARAP1 promoter 2, which is the most active in islets (Figure S2). We also observed an association between the previous index SNP at this locus (rs1552224) which is located in the 5’ UTR of ARAP1 and sites located across the entire locus.

cis-eQTLs for STARD10 and ARAP1 Are Not Detected in the Liver

Given the characteristic effect of this locus on apparent proinsulin processing and the importance of liver, where STARD10 is also highly expressed, for proinsulin clearance, we assessed data from previous reports to determine whether similar T2D-GWAS coincident cis-eQTLs were observed for this tissue. We examined publicly available datasets including samples from 97 (GTEx) and 600 (STARNET) samples, respectively. In contrast to the situation in islets, the previously reported cis-eQTLs for STARD10 in the liver were in very low LD (r² ~0.1) with the identified T2D locus, and no liver cis-eQTL was observed for ARAP1.

In order to further validate the above results, we searched for cis-eQTLs for these genes in a separate cohort of liver biopsies from 186 female subjects in the

---

**Figure 2. A Region Carrying Multiple Risk Variants Including rs140130268 Is Physically Associated with STARD10 Promoter and Influences Enhancer Activity**

(A) 3C-qPCR analysis of long-distance interactions at the STARD10 locus assessed in human-derived EndoC-βH1 cells. The relative level of each ligation product (fragments –4 to –7) is plotted according to its distance (in kb) from the STARD10 promoter P2 (P2). The constant primer and the Taqman probe are indicated in orange. Data were normalized to a CXCL12 loading control. Blue box, STARD10 Promoter2; yellow box, qPCR probe; yellow arrow, qPCR constant primer; red stars, variants. The NcoI restriction fragments are indicated below the graph. NcoI fragments are numbered from fragment –1 to –8. The data represent two or three independent experiments.

(B) The protective allele of variant rs140130268 increases the transcriptional activity of the corresponding region in insulin-secreting cells by luciferase-reporter assay. Diagram of the luciferase reporter constructs carrying either the risk or protective variants. Transcriptional activities measured by dual luciferase assay after transfection of INS1(832/13) cells (see Material and Methods). The risk variant is shown in red and the protective variant in blue. Data are from five independent experiments, represented as the mean ± SEM, and significance was calculated by Mann-Whitney two-tailed test.
Figure 3. Defective Proinsulin Processing and In Vitro Insulin Secretion in βStarD10 KO Mice

(A) Western blot analysis of STARD10 in pancreatic islets of wild-type (Cre−) and littermate βStarD10 KO mice (Cre+).

(B) Growth curves are similar in male WT and βStarD10 KO littermates maintained on a regular chow diet.

(C) Increased fed glycemia in 14-week-old male βStarD10 KO compared to WT littermates (n = 8–9 mice per genotype, *p < 0.05, unpaired two-tailed Student’s t test).

(D) Glucose tolerance at 8, 12, and 16 weeks of age are similar in male WT and βStarD10 KO littermates as determined by intraperitoneal glucose tolerance tests (1 g/kg).

(legend continued on next page)
Defective Glucose Homeostasis and Insulin Secretion in Mice Deleted Selectively for StarD10 in the β Cell
Mice lacking StarD10 selectively in the β cell (βStarD10 KO) were generated by crossing StarD10 floxed mice, in which LoxP sites were present at either side of exon 3, to mice carrying an Ins1Cre allele.57 The latter strain allows efficient (>94%) deletion in β cells, without significant recombination at extra-pancreatic sites including the brain or the expression of human growth hormone.66 Levels of StarD10 mRNA (not shown) and protein (Figure 3A) were markedly reduced in islets from βStarD10 KO mice, demonstrating efficient deletion from β cells. βStarD10 KO mice displayed significantly higher fed glycemia at 14 weeks of age (Figure 3C), although glucose tolerance was not different compared to control Cre’ littermates at 16 weeks of age (Figure 3D). Insulin sensitivity (Figure 3E) and in vivo insulin secretion in response to glucose (Figure 3F) remained unchanged in βStarD10 KO mice versus controls. While vesicle density at the plasma membrane, assessed using total internal reflection of fluorescence (TIRF) microscopy, was not altered by StarD10 deletion (Figures S3A and S3B), increases in the number of exocytotic events in response to glucose or KCl tended to be reduced (Figure S3C). Importantly, proinsulin:insulin ratios were significantly lower in βStarD10 KO mice compared to controls (Figure 3G).

Indicating defects in both glucose sensing and metabolism-independent insulin secretion, cytosolic Ca²⁺ increases in response to high glucose were diminished (Figure 3H), and both glucose- and KCl-induced insulin secretion (Figure 3I) were impaired in islets from βStarD10 KO animals, compared to controls.

Impaired Glucose Homeostasis in Global StarD10-Null Mice
Given the mild glycaemic defects observed in βStarD10 KO animals and the fact that StarD10 is highly expressed in the liver (bioGps, Web Resources), we next explored glucose homeostasis in animals deleted globally (“tm1a allele”)56 (IMPC, Web Resources) for StarD10. This approach generates a null allele through splicing to a lacZ trapping element inserted into the second intron of StarD10. Correspondingly, StarD10 mRNA was eliminated from both the liver and pancreatic islets (data not shown) and StarD10 deletion from islets (Figure 4A) and liver (not shown) was verified by western blotting. Body weights of animals fed a normal chow diet were indistinguishable between genotypes (Figure 4B). However, compared to wild-type controls, male StarD10+/– and StarD10–/– mice exhibited higher fed glycemia from 14 weeks (Figure 4C) and developed glucose intolerance from 16 weeks of age (Figure 4D). StarD10+/– and StarD10–/– mice also showed markedly reduced insulin sensitivity (Figure 4E) and insulin secretion in response to glucose was sharply reduced in vivo (Figure 4F). In common with βStarD10 KO mice, the ratio of circulating proinsulin:insulin was also diminished in StarD10–/– mice versus controls (Figure 4G).

Examined in isolated islets from male and female mice combined, cytosolic Ca²⁺ responses to glucose, though not KCl, were also decreased in StarD10-null animals, compared to wild-type mice (Figure 4H). Correspondingly, insulin secretion (Figure 4I) was markedly impaired in islets from StarD10–/– versus wild-type mice. KCl-stimulated insulin secretion was also strongly diminished in islets from the null mice (Figure 4I, right). By contrast, β cell mass was increased, though β:α cell ratio was not affected, in null animals (Figure S4).

Glucose Homeostasis Is Improved in Mice Overexpressing StarD10 in Pancreatic β Cells
In light of the findings above, we generated mice in which StarD10 was overexpressed selectively in the β cell in adults under the control of an insulin 2 promoter-driven reverse tetracycline trans-activator (Rip7-rtTA).55 Transgene expression was induced by doxycycline (2 g/L) at 5 weeks of age (Figure 5A). Transgenic animals fed a normal chow diet displayed no alterations in body weight (Figure 5B, solid lines) and exhibited similar glucose tolerance compared to wild-type littermates (Figure 5C). However, StarD10-tg male mice fed a high fat diet (HFD, 60% total calories from fat) (Figure 5D, dotted lines) displayed improved intraperitoneal glucose tolerance (16 weeks, AUC: WT: 2,586 ± 153 M.min; StarD10-tg: 1,818 ± 129 M.min; p < 0.05). These animals gained less...
Figure 4. Defective Glucose Homeostasis, Insulin Secretion, and Proinsulin Processing in StarD10 Global KO Mice

(A) Western blot analysis of STARD10 in pancreatic islets of StarD10+/+, StarD10+/−, and StarD10−/− mice.

(B) Growth curves are similar in StarD10-WT (black), heterozygous (light blue), and null (dark blue) male littermates maintained on a regular chow diet.

(legend continued on next page)
Unaltered Glucose Homeostasis and Proinsulin Processing after Deletion of *Arap1* Selectively in the Pancreatic β Cell

We next examined the impact of ablating *Arap1* expression selectively in the β cell. Mice bearing floxed alleles of *Arap1*, in which *LoxP* sites were present at either side of exon 12, were bred to mice carrying an Ins1Cre allele as above. Islet ARAP1 immunoreactivity was reduced by >70%–80% (Figure 5A), reflecting selective deletion from β cells. Mice null for *Arap1* displayed normal changes in body weight (Figure 5B) and fed glycemia (Figure 5C), unaltered glucose tolerance (Figure 5D), insulin sensitivity (Figure 5E), and glucose-stimulated insulin secretion in vivo (Figure 5F). In marked contrast to βStarD10 KO mice, fasting proinsulin:insulin ratios (Figure 5G), intracellular Ca\(^2+\) dynamics (Figure 5H), and insulin secretion from isolated islets (Figure 5I) were indistinguishable between β*Arap1* KO mice and littermate controls.

Overexpression of *ARAP1* in Pancreatic β Cells Does Not Affect Glucose Homeostasis in Mice

Given that previous studies\(^1\) reported an association between the possession of risk alleles and increased expression of *ARAP1*, we also explored the impact of overexpressing *ARAP1* in β cells (Figure S6A). There were no differences in body weight between control (black) and transgenic (red) animals (Figure S6B) fed either a regular chow (solid lines) or a high fat diet (HFD; dotted lines). Transgenic animals fed a regular chow diet (Figure S6C, solid lines) showed little evidence of abnormal glucose tolerance until 16 weeks of age, other than a small increase in peak blood glucose observed at 8 weeks of age. Similarly, transgenic animals fed a HFD (Figure S6C, dotted lines) presented with no abnormalities in intraperitoneal glucose tolerance. Intracellular free Ca\(^2+\) increases (Figure S6D) and insulin secretion in response to high glucose (Figure S6E) were also identical between the two genotypes.

These data further argue against a role for increased *ARAPI* expression in the β cell as responsible for the diabeticogenic effects of T2D risk alleles at this locus.

Given the existence of insulin resistance in global *StarD10* mice, likely to be the result of changes in liver function, we reassessed the impact of variants near *ARAPI* and *STAR10* on this parameter in humans by consulting previously published data.\(^5\) HOMA-IR revealed no changes in insulin sensitivity in risk allele carriers (results not shown).

Discussion

Our primary aim in the present study was to establish which gene(s) at the previously identified T2D association signal at the *ARAPI/STAR10* locus contribute to altered T2D risk in humans, and the likely tissue through which these effects were observed. Using a combination of human pancreatic islet transcriptome data, in vitro studies, and molecular genetics in mice, we provide evidence, discussed in detail below, that the pathogenic action is mediated, at least in large part, via decreases in *STAR10*, but not *ARAPI*, expression in the β cell.

First, we identify a 5 kb region in intron 2 of *STAR10* that: (1) captures the T2D GWAS association signal; (2) is associated with *STAR10*, but not *ARAPI*, mRNA expression in human islets and physically interacts with the promoter for the *STAR10* isoform dominant in islets; (3) overlays a stretch-enhancer in islets; (4) is correlated with local chromatin structure in islets; and (5) contains allelic variants shown experimentally to alter enhancer function. Though this region contains several highly correlated candidate variants, which may individually or jointly contribute to T2D risk, we demonstrate that the variant most likely to underlie the associations has a direct effect.
Figure 5. Normal Glucose Homeostasis, Insulin Secretion, and Proinsulin Processing in βArap1 KO Mice

(A) Western blot analysis of ARAP1 in pancreatic islets of WT (Cre−) and Arap1 β cell KO mice (Cre+). The arrows depict the short (130 kDa) and long (160 kDa) variants of ARAP1.

(B) Growth curves are similar in male WT and βArap1 KO littermates maintained on a regular chow diet.

(C) Fed glycemia are identical in 14-week-old male WT and βArap1 KO littermates.

(legend continued on next page)
on local enhancer function. Of note, the risk allele of indel rs140130268, a gain of a GTTT repeat, is associated with lowered enhancer activity (Figure 2B). Such repeats are associated with Crohn disease and neurodegenerative disorders, in each case decreasing the transcription of nearby genes.

In contrast to recent findings, but in line with other studies, we obtained no evidence of an association between the risk allele at rs1552224 and increased ARAPI expression levels in islets. On the other hand, when we extended a previous study of human pancreatic islet expression that had associated increased STARD10 expression with the minor (T2D-protective) allele at the GWAS proxy variant rs1552224 and added human islet data from two new independent subject groups provided by the IMIDIA consortium, derived from both organ donors and partial pancreatectomy patients, we could show reproducible association with STARD10 expression. The latter samples provide data for the first time from T2D subjects: interestingly, both STARD10 and ARAPI expression were decreased in T2D islets versus non-diabetic controls, the former in common with earlier findings, and further suggesting that STARD10 depletion may impair insulin production in the diseased state. Given the absence of a cis-eQTL for ARAPI, we interpret the reduced expression of this gene in T2D versus non-diabetic islets as being reactive rather than causal.

Although the more marked glycemic phenotype of global versus β cell-selective StarD10 mice might appear to suggest an action of T2D-associated variants at the ARAPI/STAR10 locus via the liver or other insulin-sensitive tissues, several lines of evidence point away from this possibility. First, changes in fed glycemia were similar in both global and β cell-selective StarD10-null mice and the lowering of plasma proinsulin:insulin ratio, which characterizes carriers of the human risk alleles at this locus, was similarly recapitulated in both models. The latter observation strongly suggests that alterations in proinsulin processing in the β cell, rather than preferential clearance of proinsulin by the liver in the absence of StarD10 in the latter tissue, are responsible for the altered circulating levels of the two forms of insulin (Figures 3 and 4). Deletion of StarD10 selectively in liver cells will be needed in the future to confirm or refute this point. Second, previous GWASs and meta-analyses in man demonstrated that risk variants at this locus are associated with a negative log HOMA-B, implying decreased β cell function. By contrast, no associations were found with HOMA-IR, thus indicating unaltered insulin sensitivity. In line with this, decreased insulin secretion during OGTt was observed in non-diabetic carriers of risk alleles, with no alteration of the insulin sensitivity index. Third, a change in STARD10 (but not ARAPI) expression associated with genotype was clearly observed in islets as well as from pancreatic tissue obtained by LCM and thus partially enriched in β cells (Table 1). Fourth, analysis of active histone marks and DNase hypersensitivity reveals that the implicated variants at this locus reside in large stretch-enhancer specific to islets but largely absent from liver-derived cells (Figures 1C and S1). Fifth, the identified variants were located in a region shown by 3C analysis to be physically associated with the STARD10 promoter in a human β cell-derived line (Figure 2A). Interestingly, a physical association was also detected between the ARAPI promoter and the same region (Figure S2), though the functional significance of this is obscure. Thus, sixth, eQTLs in significant LD with the T2D locus are not detected in human liver samples from the GTEx or STARNET consortia as well as samples from the ABOS consortium. Based on the power calculation provided by the GTEx Consortium, and given that the minor allele frequency for variants at the STARD10/ARAPI locus is 15%, then with a sample size of n = 186 (ABOS) or n = 600 (STARNET), we would expect a power of 90% or 100%, respectively, to detect a β of 0.15 at α 2.5 × 10^-2. In neither of these cases, nor in the GTEx samples (n = 97), are liver cis-eQTLs detected for either STARD10 or ARAPI. The above power calculations are somewhat imprecise given differing expression levels, the latter two studies are thus adequately powered to detect cis-eQTLs for STARD10 or ARAPI in the liver of the same size or smaller than that seen in islets. Instead, our analysis of more than 800 liver samples from three separate groups fails to identify any liver cis-eQTLs for either ARAPI or STARD10 of comparable size to that detected for STARD10 in islet samples from 341 subjects (Figure 1, Table 1). Nevertheless, we do not exclude the possibility that future, even larger eQTL studies, might identify more subtle effects on the expression of additional genes at this locus.

Studies in mice also provided further functional evidence for STARD10 as the gene most likely to confer effects...
on T2D risk at this locus. Thus, highly targeted disruption of Arap1 in β cells had no effect on insulin secretion in vivo or in vitro nor on circulating proinsulin:insulin ratios. Furthermore, inducible overexpression of ARAPI in β cells in adult mice failed to exert substantial effects on insulin secretion or glucose tolerance. Again, global or liver-specific deletion of Arapi will be useful in the future to exclude any possible contribution of this gene to the action of risk variants through the latter tissue.

Together, this body of evidence points toward an action of the T2D-associated variants via the β cell and through changes in STARD10 expression. Further work will need to be undertaken to explore the molecular mechanisms through which the observed variants affect STARD10 expression, including the identification of transcription factors which bind in this region. While interrogation of the Islet Regulome browser (Web Resources) reveals that none of five key β cell transcription factors examined by ChIP-seq in human islets (FOXA2, MAFB, NKX2.2, NKX6.1, and PDX1) bind at this site, the binding of High Mobility Group Box 1 (HMGB1) is predicted to be affected by GTTT deletion. However, our preliminary experiments (not shown) have failed to provide any evidence to support this possibility.

We would note that the relevance of changes in insulin sensitivity in StarD10-null mice as regards impact of T2D variants at the ARAPI/STARD10 locus is questionable, given the absence of any evidence for an impact of this locus on insulin signaling (see above). Nevertheless, the more striking glycemic phenotype of the global versus the β cell-selective StarD10-null mouse (Figure 3 versus Figure 4) means that the extra-pancreatic actions of therapeutic agents that seek to target STARD10 will need to be given careful consideration.

Possible Mechanisms of STARD10 Action on Insulin Processing and Secretion

By what means may STARD10 depletion impair insulin secretion while preserving proinsulin processing? First, since STARD10 is concentrated in sperm flagella, a site of vigorous energy consumption, the enzyme might conceivably be required for normal glucose metabolism and signaling in the β cells. Supporting this view, glucose-induced Ca\textsuperscript{2+} dynamics, likely reflecting glucose-dependent ATP generation, were impaired in StarD10-null β cells. On the other hand, preliminary lipidomic analysis in liver (not shown) indicates changes that, if they also affect the β cell granule or plasma membranes, may impair exocytosis or favor intracellular retention of unprocessed insulin. Finally, impaired exocytosis may of itself improve processing by increasing dwell time in the maturing granule. Interestingly, we saw no significant change in the number of morphologically docked granules at the plasma membrane by TIRF imaging (Figure S3), despite a trend (p = 0.06) toward a reduction in secretion in response to depolarization with KCl, implying impaired exocytotic competence of granules. This may be consistent with altered granule membrane lipid composition and defective incorporation of SNAP/SNARE proteins.

Conclusions

The present study has used multiple complementary approaches to assess the identity of the gene(s) and the site(s) of action of variants at a locus on chromosome 11q associated with T2D risk in several earlier studies. This multi-faceted approach has been adopted since, in our view, no one single piece of evidence can be considered definitive in connecting true GWAS signals to their downstream effectors as each has intrinsic limitations. Rather, it is the alignment of multiple types of data, each supporting the same hypothesis which, as here, provides compelling evidence to establish a particular gene as the effector transcript at a given locus.

Although it is anticipated that future studies will bring important additional insights, our findings challenge the existing view that alterations in ARAPI expression in the β cell represent the sole or most important mechanism but instead imply a role for STARD10 in this tissue. The molecular mechanisms through which the encoded lipid transfer protein affects β cell physiology, and in particular proinsulin processing, should provide exciting avenues for future research and possibly therapeutic exploitation.

Accession Numbers

Raw sequence files for the ATAC-seq data described in this manuscript are available from the European Nucleotide Archive (ENA) under accession number PRJEB18684.

Supplemental Data

Supplemental Data include six figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2017.01.011.

Acknowledgments

G.A.R. was supported by a Wellcome Trust Senior Investigator Award (WT098424/AIA), an MRC Programme (MR/J000304/1), Biotechnology and Biological Sciences Research Council (BB/J015873/1), Diabetes UK (BDA/11/0004210, BDA/15/0005275), Imperial Confidence in Concept (ICiC) grants, and a Royal Society Wolfson Research Merit Award. G.R.C. received a Fellowship from the Société Francophone du Diabète. This work was also supported by an MRC Experimental Challenge grant (MR/L020149/1) to the DIVA consortium and by the Innovative Medicines Initiative Initiative Joint Undertaking under grant agreement 155005 (IMIDIA), resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP7/2007-2013) and European Federation of Pharmaceutical Industries and Associations (EFPIA) companies. We thank Raphael Scharfmann (Inserm U1016, Université Paris Descartes, Paris, France) for advice on the use of EndoC-Beta1 cells and Professor Timothy Frayling...
Web Resources

ClinicalTrials.gov, http://clinicaltrials.gov
DNase and ChIP pipeline, http://userweb.molbiol.ox.ac.uk/public/telenius/PipeSite.html
European Nucleotide Archive, http://www.ebi.ac.uk/ena
IMIDIA, http://www.imidia.org
International Mouse Phenotyping Consortium, http://www.mousephenotype.org/
Islet regulome, http://gattaca.imppc.org/isletregulome/home
Knockout Mouse Project (KOMP) Repository, https://www.komp.org/
PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/

References

diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. Diabetes 63, 2158–2171.


