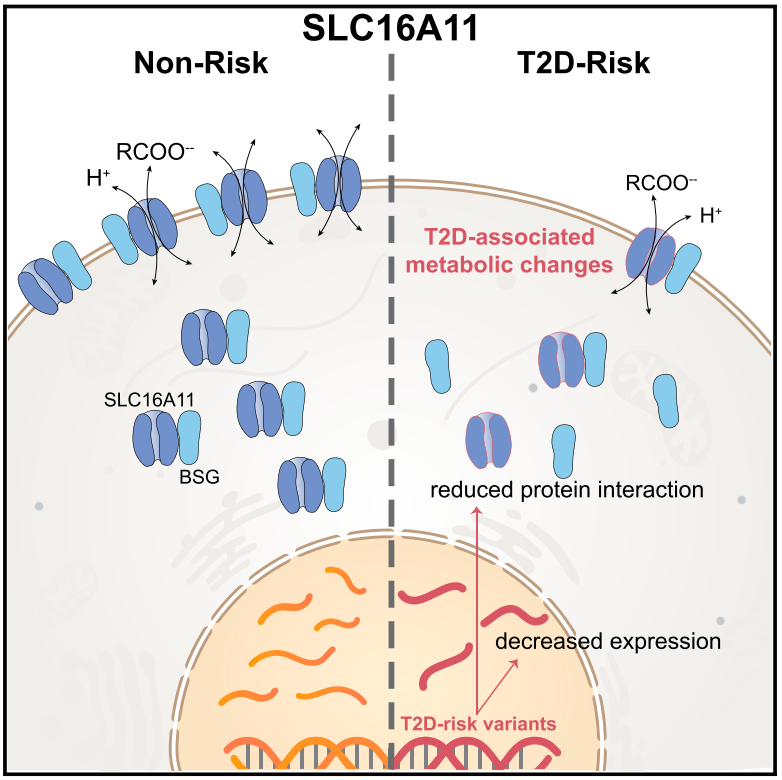
Article

Cell

Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms

Graphical Abstract



Highlights

- The T2D-risk haplotype contains a *cis*-eQTL for lower *SLC16A11* expression in liver
- T2D-risk variants disrupt a SLC16A11-BSG interaction and cell-surface localization
- Reduced SLC16A11 induces metabolic changes associated with increased T2D risk
- Therapeutics that enhance SLC16A11 levels or activity may be beneficial for T2D

Authors

Victor Rusu, Eitan Hoch, Josep M. Mercader, ..., Jose C. Florez, Suzanne B.R. Jacobs, Eric S. Lander

Correspondence

jcflorez@mgh.harvard.edu (J.C.F.), lander@broadinstitute.org (E.S.L.)

In Brief

Genetic variants associated with type 2 diabetes impair function of a monocarboxylate transporter that in turn impacts the metabolic state of the cell.

Rusu et al., 2017, Cell *170*, 199–212
June 29, 2017
http://dx.doi.org/10.1016/j.cell.2017.06.011

CellPress

Cell

Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms

Victor Rusu,[1,2,18,19] Eitan Hoch,[2,3,18] Josep M. Mercader,[2,4,5] Danielle E. Tenen,[6,7] Melissa Gymrek,[2,8,20] Christina R. Hartigan,[6] Michael DeRan,[6] Marcin von Grotthuss,[2] Pierre Fontanillas,[2,21] Alexandra Spooner,[2] Gaelen Guzman,[6] Amy A. Deik,[6] Kerry A. Pierce,[6] Courtney Dennis,[6] Clary B. Clish,[3,6] Steven A. Carr,[6] Bridget K. Wagner,[6] Monica Schenone,[6] Maggie C.Y. Ng,[9] Brian H. Chen,[10] MEDIA Consortium, SIGMA T2D Consortium, Federico Centeno-Cruz,[11] Carlos Zerrweck,[12] Lorena Orozco,[11] David M. Altshuler,[2,13,14,15,16,22] Stuart L. Schreiber,[6] Jose C. Florez,[2,3,4,15,*] Suzanne B.R. Jacobs,[2,3,4] and Eric S. Lander[6,16,17,23,*]

[1]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA 02115, USA
[2]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[3]Metabolism Program, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[4]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
[5]Barcelona Supercomputing Center (BSC), Joint BSC-CRG-IRB Research Program in Computational Biology, 08034 Barcelona, Spain
[6]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[7]Division of Endocrinology, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA
[8]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA
[9]Center for Genomics and Personalized Medicine Research, Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA
[10]Longitudinal Studies Section, Translational Gerontology Branch, Intramural Research Program, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA
[11]Instituto Nacional de Medicina Genómica, Tlalpan, 14610 Mexico City, Mexico
[12]The Obesity Clinic at Hospital General Tlahuac, 13250 Mexico City, Mexico
[13]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[14]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA
[15]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
[16]Department of Biology, MIT, Cambridge, MA 02139, USA
[17]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[18]These authors contributed equally
[19]Present address: Jnana Therapeutics, Boston, MA 02210, USA
[20]Present address: Department of Medicine, Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA 92093, USA
[21]Present address: 23andMe, Mountain View, CA 94041, USA
[22]Present address: Vertex Pharmaceuticals, Boston, MA 02210, USA
[23]Lead Contact
*Correspondence: jcflorez@mgh.harvard.edu (J.C.F.), lander@broadinstitute.org (E.S.L.)
http://dx.doi.org/10.1016/j.cell.2017.06.011

SUMMARY

Type 2 diabetes (T2D) affects Latinos at twice the rate seen in populations of European descent. We recently identified a risk haplotype spanning *SLC16A11* that explains ~20% of the increased T2D prevalence in Mexico. Here, through genetic fine-mapping, we define a set of tightly linked variants likely to contain the causal allele(s). We show that variants on the T2D-associated haplotype have two distinct effects: (1) decreasing *SLC16A11* expression in liver and (2) disrupting a key interaction with basigin, thereby reducing cell-surface localization. Both independent mechanisms reduce SLC16A11 function and suggest *SLC16A11* is the causal gene at this locus. To gain insight into how SLC16A11 disruption impacts T2D risk, we demonstrate that SLC16A11 is a proton-coupled monocarboxylate transporter and that genetic perturbation of SLC16A11 induces changes in fatty acid and lipid metabolism that are associated with increased T2D risk. Our findings suggest that increasing SLC16A11 function could be therapeutically beneficial for T2D.

INTRODUCTION

Type 2 diabetes (T2D) afflicts more than 415 million people and is a leading cause of morbidity and mortality worldwide. While T2D is influenced by environmental factors, it is also a highly heritable disorder (Prasad and Groop, 2015), with genetic variation contributing to a disparity in T2D prevalence both within and across populations (Diamond, 2003; Williams et al., 2014). For example, within American populations, the prevalence of T2D is nearly twice as high in individuals of Mexican or Latin American descent as compared to US non-Hispanic whites (Villalpando et al., 2010).

Understanding the genetic contributions to disease biology can help identify at-risk individuals, guide more effective personalized treatment approaches (Shepherd et al., 2009), and illuminate new targets and pathways for therapeutic development and intervention. To date, large-scale genetic studies have identified >100 genetic loci containing variants associated with altered risk of T2D, with variation at these loci increasing risk by ~10–20% (Flannick and Florez, 2016). In most cases, risk haplotypes span several genes, and the causal gene and variants have not yet been identified. While in some cases the locus contains a candidate gene related to processes implicated in T2D pathophysiology, such as insulin secretion and insulin sensitivity, in most cases there is no obvious candidate gene (Morris et al., 2012; Voight et al., 2010).

Most of the known T2D susceptibility loci were identified by studies in individuals of European descent. However, recent genetic studies in diverse populations and population isolates have proven successful in identifying novel genes and variants associated with T2D risk. These include a common Greenlandic population-specific variant at the *TBC1D4* locus that confers a 10-fold T2D risk to homozygous carriers (Moltke et al., 2014), a low frequency variant in *HNF1A* in Latinos that confers a 5-fold increased risk to heterozygous carriers (Estrada et al., 2014), and rare Finnish and Icelandic variants in *SLC30A8* that reduce T2D risk by 65% in heterozygous carriers (Flannick et al., 2014).

One of the largest genetic risk factors for T2D, located at 17p13, was first identified through genome-wide association studies (GWAS) investigating genetic influences on diabetes risk in Mexico and East Asia (Hara et al., 2014; Williams et al., 2014). The T2D-risk haplotype at this locus is common among individuals of Mexican or Latin American descent (allele frequency of ~30%), less common in East Asia (10%), and rare among individuals of European (<2%) and African (0%) descent. Each allele of the T2D-risk haplotype increases disease risk by ~25%, and the locus appears to explain ~20% of the increased T2D prevalence in Mexico (Williams et al., 2014).

The T2D-associated variants at 17p13 span two protein-coding genes, *SLC16A11* and *SLC16A13*. While most disease-associated common variants occur in noncoding regions (Maurano et al., 2012), the T2D-risk haplotype at 17p13 includes five coding variants in *SLC16A11*, comprising four missense mutations and one synonymous change (Williams et al., 2014). Notably, a distinct haplotype carrying two of these five coding variants occurs at high frequency (~36%) in Africa but is rare in other populations.

Both SLC16A11 and SLC16A13 are members of the SLC16 (or monocarboxylate transporter, MCT) family, a group of 14 solute carriers that is defined by two highly conserved sequences (Halestrap, 2013). Despite structural similarities, SLC16 family members mediate transport of distinct substrates, utilizing two different mechanisms. The first class of SLC16 members (category I) transport simple monocarboxylic acids, such as lactate, pyruvate, and ketone bodies, via a proton (H^+)-coupled mechanism (Halestrap, 2013). This class currently has four known members—SLC16A1, SLC16A3, SLC16A7, and SLC16A8—all of which have been shown to interact with basigin (BSG) and embigin (EMB), two chaperone proteins important for plasma-membrane localization of the transporters (Halestrap, 2013). By contrast, the second class of SLC16 transporters (category II) do not transport simple monocarboxylic acids nor do they use an H^+-coupled mechanism; instead, these proteins transport larger hydrophobic monocarboxylates, such as triiodothyronine (T_3) and thyroxine (T_4), through facilitated diffusion (Halestrap, 2013). This class currently has two known members: SLC16A2 and SLC16A10. Unlike category I members, SLC16A2 does not interact with BSG or EMB (Visser et al., 2009). Studies on two other family members—SLC16A6 and SLC16A9—are limited; however, SLC16A6 has been shown to transport ketone bodies (Hugo et al., 2012), suggesting it may belong to category I, and SLC16A9 has been shown to transport carnitine via an H^+-independent mechanism (Suhre et al., 2011), suggesting it may belong to category II. Along with their diverse transport functions, members of the SLC16 family have distinct, but overlapping, expression patterns (Halestrap, 2013). *SLC16A11* is expressed in relatively few tissues, with the highest levels detected in thyroid, liver, and salivary gland (Williams et al., 2014). The role of SLC16A11 in these tissues has not yet been characterized.

Here, we elucidate the functional basis of the T2D-risk haplotype at 17p13. We show that the genetic variants at this locus have two distinct actions on SLC16A11, which we show belongs to category I of SLC16 transporters. Specifically, some (presumably, noncoding regulatory) variants on the T2D-risk haplotype lead to decreased gene expression of *SLC16A11* in liver, while coding variants affect the interaction of the SLC16A11 protein with BSG, leading to reduced levels of the transporter at the cell surface. Importantly, we demonstrate that disruption of SLC16A11 in primary human hepatocytes leads to T2D-relevant changes in fatty acid and lipid metabolism. Together, these results implicate reduced SLC16A11 function in liver as a causal factor for T2D and suggest a novel therapeutic hypothesis.

RESULTS

Fine Mapping of the T2D Association at 17p13

To explore which variants are most likely causal for the T2D association at this locus, we began by analyzing the strength of association for all variants in the region to construct a "99% credible set"—that is, a set of variants that has a 99% probability of containing the causal variant(s). For this purpose, we used a dataset generated by integration of genotyping, whole-exome sequencing, and genotype imputation data from Mexican and Latin American populations (Estrada et al., 2014; Williams et al., 2014). We calculated the posterior probability of causality for each variant, resulting in a 99% credible set consisting of 18 variants (Figure 1; Table S1). In the Mexican population, these variants are in strong linkage disequilibrium with the top variant ($r^2 \geq 0.93$) and are associated with a ~30% increase in T2D risk ($p = 1.7 \times 10^{-11}$–8.1×10^{-13}; odds ratio = 1.28–1.30). The credible set includes 12 noncoding SNPs spanning *SLC16A11* and *SLC16A13*, one silent coding SNP in *SLC16A13*, and the four missense and one silent coding SNPs in *SLC16A11*. Among these, three noncoding variants proximal to the *SLC16A11* transcription start site (rs77086571 and rs74577409 in the *SLC16A11* proximal promoter region and rs2292351 in the 5′ UTR) rank as the most likely causal variants, with a collective

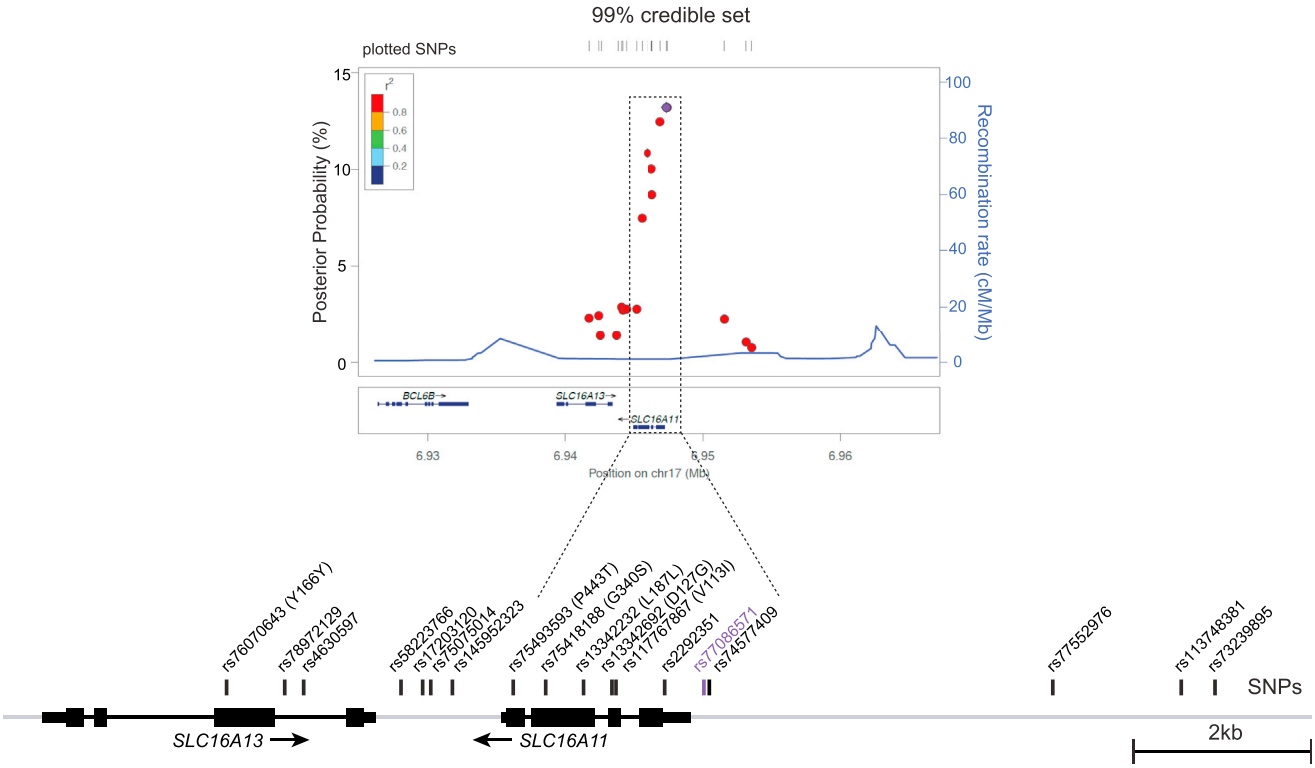


Figure 1. T2D-Risk Credible Set at 17p13

Regional signal plot representing variants in the 99% credible set for the T2D signal at 17p13. The T2D-risk credible set variants are depicted as points, and the colors indicate the R-squared with the top SNP (rs77086571), marked in purple. Depiction of the locus with credible set SNPs indicated is shown (bottom).

posterior probability of 39%, while the missense variants in *SLC16A11* show collective posterior probability of 29% (Figure 1; Table S1).

Variants on the T2D-Risk Haplotype at 17p13 Decrease *SLC16A11* Expression in Liver

To explore the possibility that one or more variants in the credible set act through effects on gene expression, we examined expression levels for the two genes spanned by these variants (*SLC16A11* and *SLC16A13*) and three additional genes in the vicinity of this region (*BCL6B*, *CLEC10A*, and *RNASEK*). We compared expression levels in T2D-relevant tissues, liver and visceral adipose, in individuals from Mexico who carry zero, one, or two copies of the T2D-risk haplotype. We quantified gene expression using droplet digital PCR (ddPCR) and tested associations of gene-expression levels with genotype using variant rs13342692 to tag the T2D-risk haplotype.

In liver, *SLC16A11* expression is significantly reduced in a dose-dependent manner in carriers of the T2D-risk haplotype ($P \approx 1.4 \times 10^{-4}$) (Figures 2A, S1A, and S1B). Relative to homozygotes for the non-risk haplotype, expressions levels were 42% lower in heterozygotes (SEM 9%) for the risk haplotype and 66% lower in homozygotes (SEM 6%) for the risk haplotype. None of the other four genes tested showed significant and reproducible genotype-dependent expression changes in liver (Figures 2A, S1A, and S1B). Because some tissue donors had been taking a medication that could interfere with metabolism (such as metformin) for several months prior to surgery, we repeated the analysis excluding these individuals to remove potential confounding effects from ongoing therapy. Although the smaller sample size provides reduced statistical power, we confirmed our initial finding that the T2D-risk allele is associated with a dose-dependent decrease in *SLC16A11* expression in liver ($P \approx 0.01$; Figure S1C). None of the five genes showed a genotype-dependent association to expression levels in visceral adipose (Figures S1D–S1F). Notably, our sensitive ddPCR assay was unable to detect any changes in *RNASEK* expression, for which the T2D-risk haplotype was recently reported to have an expression QTL (eQTL) effect in subcutaneous adipose, skeletal muscle, and whole blood cells (Traurig et al., 2016).

These results provide evidence that the T2D-risk haplotype contains an eQTL affecting *SLC16A11* expression in human liver, one of the tissues in which *SLC16A11* is most highly expressed (Williams et al., 2014), and a tissue in which disruption of metabolic processes is implicated in T2D pathophysiology (Perry et al., 2014).

SLC16A11 Expression and H3K27ac Marks Are Skewed in Heterozygous Carriers of the T2D-Risk Haplotype

In principle, the T2D-risk variants could reduce *SLC16A11* gene expression by acting directly in *cis* (affecting expression of the *SLC16A11* allele carried on the risk haplotype) or indirectly in

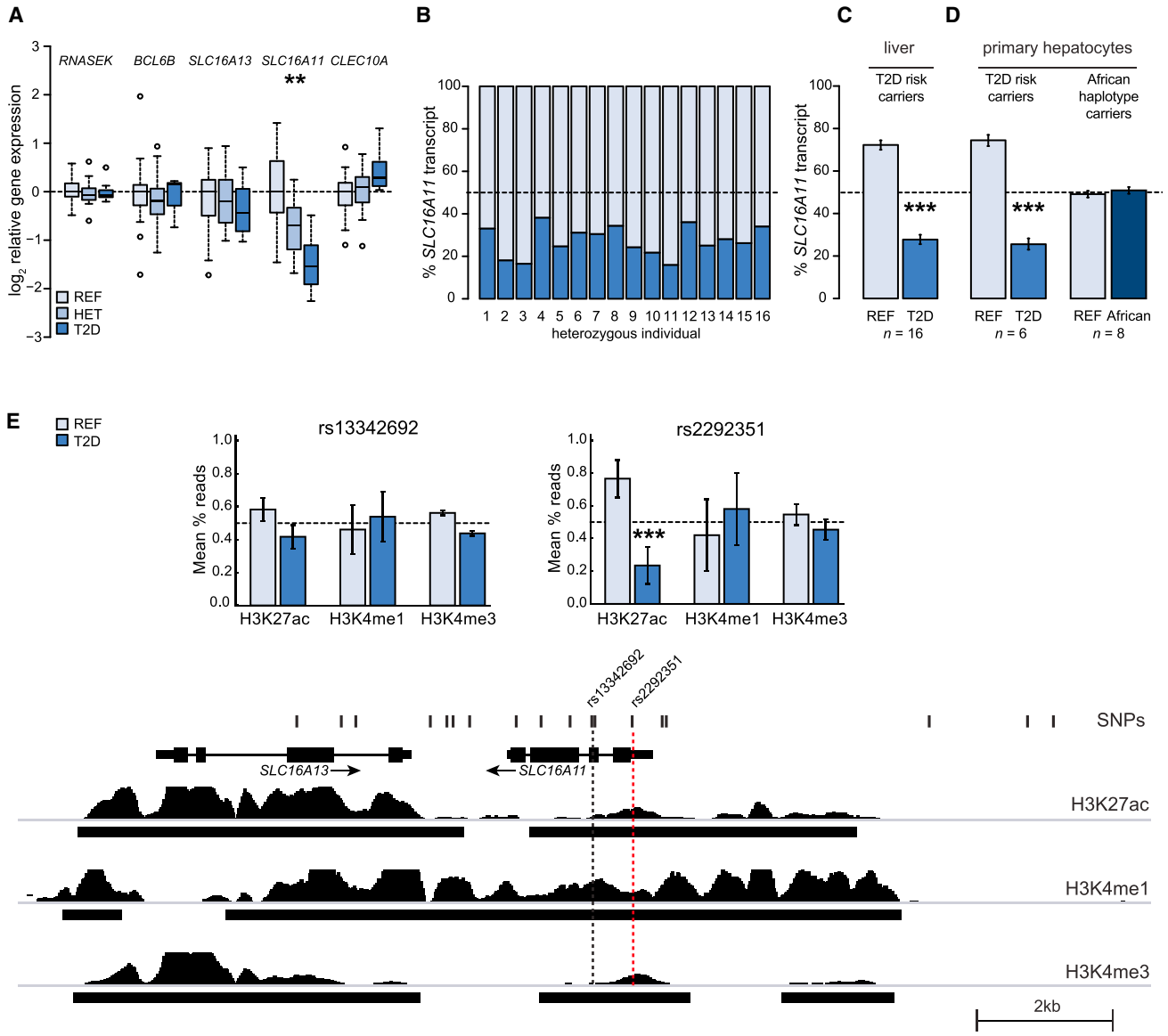


Figure 2. The T2D-Risk Haplotype Contains a *Cis*-eQTL for *SLC16A11* in Liver

(A) Expression-QTL (eQTL) analyses in liver. Boxplots depict the log2 of the relative expression level for *RNASEK*, *BCL6B*, *SLC16A11*, *SLC16A13*, and *CLEC10A* according to genotype at rs13342692. n = 21 homozygous reference (REF), 16 heterozygous (HET), and 10 homozygous T2D risk (T2D). See also Figure S1.

(B–D) Allele-specific expression analyses in (B) and (C) HET livers (n = 16) and (D) primary human hepatocytes heterozygous for either the T2D risk (n = 6) or African (n = 8) haplotypes. Bar plots depict the estimated allelic proportion and 95% confidence interval of *SLC16A11* transcript originating from each allele. See also Figure S2.

(E) ChIP-seq for H3K27ac, H3K4me1, and H3K4me3 histone modifications in human hepatocytes from three individuals heterozygous for the T2D-risk haplotype. Tracks overlapping variants in the T2D-risk credible set are shown. Bar plots depict allelic proportions ± SEM at rs13342692 and rs2292351. See also Figure S2. Asterisks indicate significance after Bonferroni correction for multiple hypothesis testing: **$p < 1 \times 10^{-3}$, ***$p < 1 \times 10^{-5}$.

trans (affecting other cellular processes that feedback on *SLC16A11* expression on both haplotypes). To distinguish between these possibilities, we carried out allelic-expression-imbalance studies in heterozygous individuals; this analysis has the added benefit of controlling for inter-individual confounders that might influence total gene-expression levels (Locke et al., 2015). We compared the expression levels from the T2D-risk and reference (non-risk) haplotypes in liver samples from 16 heterozygous individuals. Expression from the two haplotypes was measured by ddPCR, with probes that distinguish between the reference and risk allele at rs13342692. The results provide strong support for a *cis*-effect: expression from the risk allele is 62% lower than from the non-risk haplotype ($p = 2 \times 10^{-73}$; Figures 2B and 2C), which is consistent with the 66% lower expression level seen in homozygotes for the risk haplotype than the non-risk haplotype.

Table 1. Categorization of SLC16 family members

Category	Family member	Primary substrates	Mechanism	Ancillary proteins
I	SLC16A1 SLC16A3 SLC16A7	Pyruvate, Lactate, Ketone bodies	H^+-coupled	Basigin (BSG) Embigin (EMB)
	SLC16A8	Lactate		
II	SLC16A2	T3, T4 hormones	Facilitated diffusion	No interaction
	SLC16A10	Aromatic amino acids		-
Uncategorized	SLC16A6	β-hydroxybutyrate	-	-
	SLC16A9	Carnitine	Not H^+-coupled	
	SLC16A4 SLC16A5 **SLC16A11** SLC16A12 SLC16A13 SLC16A14	-	-	

Categorization of SLC16 family members with transport substrates, mechanism, and ancillary proteins indicated (Halestrap, 2013; Hugo et al., 2012; Suhre et al., 2011). SLC16A11 (bold) is an uncharacterized family member.

We observed a similar allelic skew in primary hepatocytes, cultured ex vivo from 6 heterozygous individuals: expression from the risk haplotype is 66% lower than from the non-risk haplotype ($p = 2 \times 10^{-63}$; Figures 2D and S2A), thus replicating the allelic skew we observed in human liver from heterozygous individuals. Notably, we found no allelic skew in primary hepatocytes from individuals heterozygous for the African haplotype (Figures 2D and S2B). Using the largest available African American genome-wide association meta-analysis from the MEDIA consortium, involving 8,284 cases and 15,543 control individuals (Ng et al., 2014), we also did not find a significant association with T2D risk: the two-coding-variant haplotype in African Americans has an odds ratio of ~1.06 ($P \sim 0.08$), whereas the five-coding-variant haplotype in Mexico has an odds ratio of ~1.29 ($p = 1.7 \times 10^{-11}$–8.1×10^{-13}) in the Mexican study. When comparing the effect sizes between the two studies, we saw that there was statistically significant heterogeneity ($P_{HET} \approx 4.1 \times 10^{-5}$). These analyses suggest that the two coding variants present at high frequency in Africa, rs13342692 and rs13342232 (Table S1), are not alone sufficient to cause the association with T2D or confer skewed gene expression of *SLC16A11* and that the *SLC16A11 cis*-eQTL in human liver is not present in these individuals of African descent.

We next sought to determine whether the lower expression of *SLC16A11* from the T2D-risk haplotype is associated with chromatin structure in *cis*, indicative of altered transcription at the gene. We examined the chromatin landscape on each haplotype, in hepatocytes, from three heterozygous individuals by performing ChIP-seq for several histone modifications (Figures 2E and S2C), including H3K27ac (active enhancer and promoter mark), H3K4me1 (associated with enhancers), and H3K4me3 (associated with promoters and transcription start sites) (ENCODE Project Consortium, 2012). Consistent with reduced *SLC16A11* promoter activity from the T2D-risk allele, we saw the expected allelic skew (69% lower from the T2D-risk allele) for the activating mark H3K27ac at a variant located near the 5′ end of *SLC16A11* (rs2292351; $p = 2 \times 10^{-14}$). We observed no additional significant allelic skews in chromatin marks near other variants in the T2D-risk credible set (Figure S2D).

Together, these data demonstrate the presence of a *SLC16A11 cis*-eQTL in liver carried on the risk haplotype, providing support for *SLC16A11* as a causal gene at this locus and suggesting decreased SLC16A11 function as the disease-relevant direction of effect.

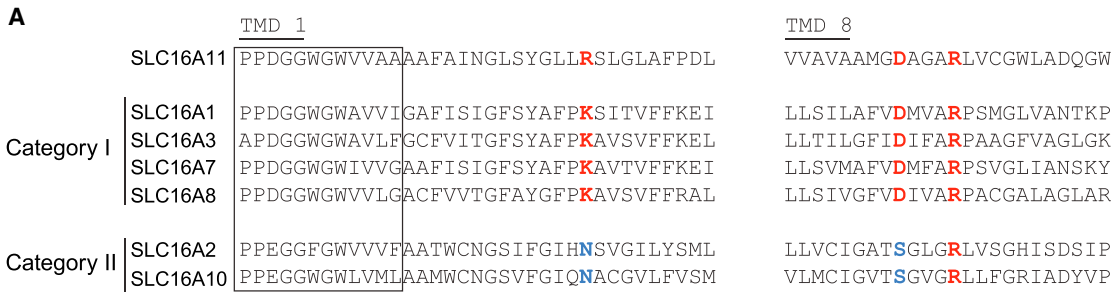
SLC16A11 Is an H^+-Coupled Monocarboxylate Transporter

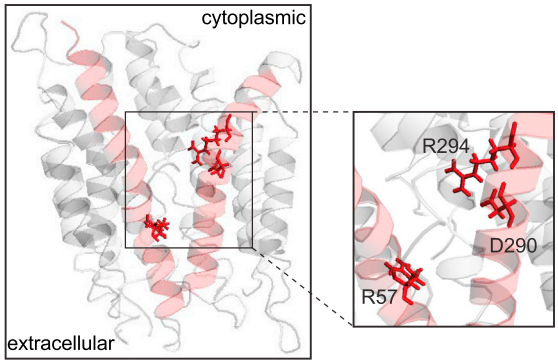
We next investigated the function of the previously uncharacterized SLC16A11 protein—in particular, whether it is a chaperone-dependent, H^+-coupled monocarboxylate transporter (category I) or a facilitator of hydrophobic monocarboxylate diffusion (category II) (Table 1).

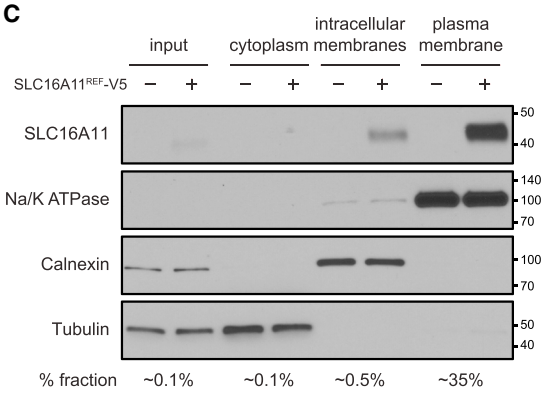
We began by comparing the protein sequences of SLC16 family members belonging to categories I and II to identify features that distinguish the two classes. Two transmembrane domains (TMDs) in SLC16A1, numbered 1 and 8, have previously been suggested to play key roles in mediating monocarboxylate transport, with three charged residues—K38 in TMD 1 and D309 and R313 in TMD 8—highlighted as likely to be functionally important (Manoharan et al., 2006). We found that R313 is present in all six members of the two categories. In contrast, K38 and D309 are present in the four known members of category I but replaced by non-charged residues in the two members of category II (Figures 3A and S3A).

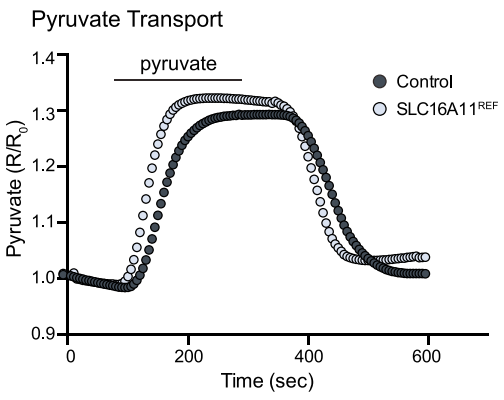
In SLC16A11, charges are present at the three corresponding positions (R57, D290, and R294). Three-dimensional homology modeling of SLC16A11 based on the solved structure of the bacterial glycerol-3-phosphate transporter (GlPT) (Lemieux et al., 2003) indicates that these three residues are found in the inner pore of the protein (Figure 3B), in a similar fashion as modeled for SLC16A1 (Manoharan et al., 2006). These structural analyses suggest SLC16A11 belongs to category I, raising the hypothesis that it might transport monocarboxylates via an H^+-coupled mechanism. We thus sought to study the transport properties of SLC16A11. We studied both the protein encoded by the non-risk haplotype, denoted SLC16A11 (or, where helpful, $SLC16A11^{REF}$), and the risk haplotype—which contains all five coding variants found on the T2D risk allele—denoted $SLC16A11^{T2D}$.

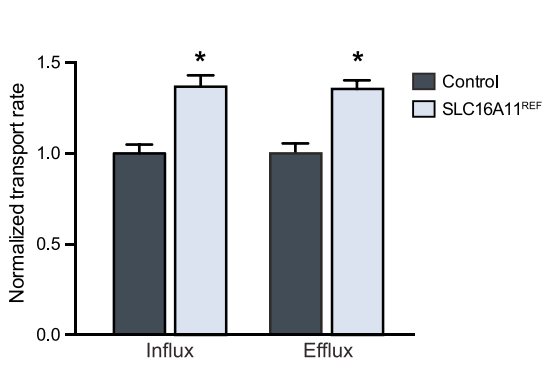
As a first step, we studied the subcellular localization of SLC16A11. Based on immunofluorescent imaging, we previously reported that SLC16A11 tagged with a C-terminal V5 epitope (SLC16A11-V5) is present in the endoplasmic reticulum (Williams et al., 2014). Because other SLC16 family members have been reported to localize to both intracellular membranes and the plasma membrane, we explored whether $SLC16A11^{REF}$-V5 is also present at the cell surface. Toward this end, we performed membrane extraction assays in

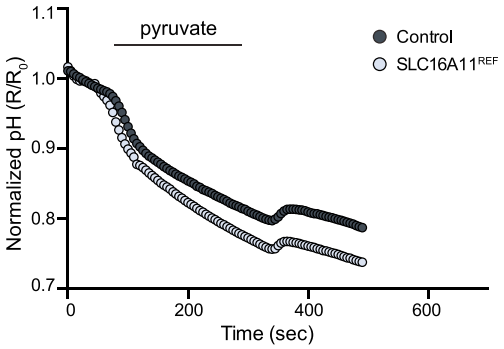


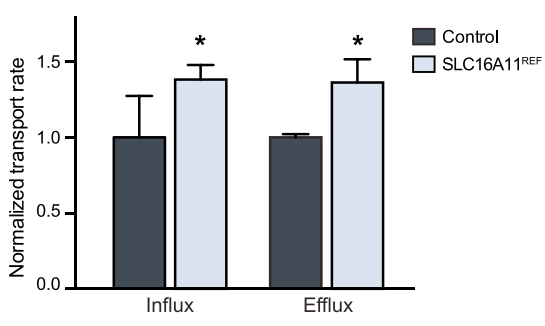












(legend on next page)

HEK293T cells expressing SLC16A11. Consistent with our previous findings, the majority of SLC16A11^{REF}-V5 is associated with intracellular membranes, with a portion (~5%) localized to the plasma membrane (Figure 3C).

We then tested whether SLC16A11 can mediate monocarboxylate transport across the plasma membrane. We focused on pyruvate as a high-affinity substrate of SLC16 category I members and because intracellular pyruvate levels can be readily measured in real time using a genetically encoded pyruvate FRET sensor, pyronic (San Martín et al., 2014).

HEK293T cells co-transfected with pyronic and either empty vector control or SLC16A11^{REF}-V5 (Figure S3B) were exposed to pyruvate, resulting in an increase of the fluorescence signal indicative of pyruvate uptake (Figure 3D). This signal is completely eliminated upon pre-incubation of the cells with a chemical inhibitor of SLC16-mediated transport, AR-C155858 (Figure S3C) (Ovens et al., 2010), consistent with the signal in control cells resulting from endogenous SLC16-dependent pyruvate transport.

We compared the influx (rising phase) and efflux (falling phase) rates of pyruvate transport in SLC16A11^{REF}-expressing and control cells. Rates in both directions were found to be ~45% higher in cells expressing SLC16A11^{REF} (Figure 3E), supporting the idea that SLC16A11 is a category I transporter. Similar patterns of bi-directional transport have been previously reported for other SLC16 family members (Halestrap, 2013). We note that a difference between SLC16A11^{REF}-expressing and control cells was only observed at neutral pH, likely due to activation of other, endogenous SLC16 family members at acidic pH (Figure S3D).

To assess whether SLC16A11 utilizes an H^+-coupled transport mechanism, we monitored pH changes in SLC16A11^{REF}-expressing and control cells using a pH-sensitive fluorescent probe, BCECF-AM (Rink et al., 1982). Following addition and withdrawal of pyruvate, we observed a ~40% increase in the rates of acidification and alkalization, with both phases reciprocal to influx and efflux of pyruvate. These results support an H^+-coupled mechanism underlying SLC16A11 transport (Figures 3F and 3G).

In summary, bioinformatic analyses and transport assays together establish SLC16A11 as an H^+-coupled monocarboxylate transporter belonging to category I.

T2D-Risk Coding Variants Are Associated with Lower SLC16A11 Transport Activity

We next investigated whether the T2D-associated coding variants in SLC16A11—four missense variants (V113I, D127G, G340S, and P443T) and one silent variant (L187L)—affect the amount of SLC16A11 transport activity. We expressed both the SLC16A11^{REF} and SLC16A11^{T2D} coding regions from a cytomegalovirus (CMV) promoter to remove any potential confounding effects due to the non-coding variants (Figure S3E). Using the pyronic assay, we found that rates of both pyruvate influx and efflux were ~50% lower in cells expressing SLC16A11^{T2D} than in cells expressing SLC16A11^{REF} (Figures 4A and 4B). Similar reductions in proton-transport rates were observed (Figures 4C and 4D). Together, these data demonstrate that the T2D-associated coding variants in SLC16A11 result in decreased SLC16A11 transport activity. The T2D haplotype is thus associated with two distinct effects: decreased gene expression and decreased transport activity of SLC16A11.

SLC16A11 Interacts with SLC16 Category I Ancillary Proteins

The T2D-risk coding variants seem unlikely to directly interfere with substrate binding or release because they are far from the inner pore in the SLC16A11 homology model (Figure 5A). An alternative hypothesis for their observed effects is that they disrupt protein-protein interactions responsible for correct SLC16A11 function. To explore this possibility, we immunoprecipitated SLC16A11 from HEK293T cells expressing either REF or T2D-risk SLC16A11-V5 or empty vector control and used quantitative mass spectrometry to define the SLC16A11 interactome.

These proteomic studies required robust expression of SLC16A11; however, we observed that SLC16A11 protein levels remain low, even after transient overexpression (Figure S4A). We noted that the SLC16A11 protein is rapidly degraded following inhibition of protein synthesis and stabilized by proteasome inhibition (Figures S4B and S4C). We therefore explored ways to stabilize SLC16A11 against proteasome-mediated degradation. While mutation of the sole lysine residue in SLC16A11 had no effect, we found that a proline to aspartic acid substitution (P2D)—at a potential site of regulation through an N-terminal ubiquitination pathway—increased SLC16A11 protein levels (Figure S4D). Therefore, we utilized SLC16A11 proteins containing the P2D mutation in our interaction screen, resulting in successful enrichment of tagged SLC16A11 and associated proteins in our immunoprecipitations (Figure 5B).

Among the ~50 most highly enriched SLC16A11^{REF}-interacting proteins (top 10% of interactors with a Blandt-Altman adjusted $p < 0.05$; Figure 5B; Table S2), two stood out: BSG and EMB. As noted above, BSG and EMB act as chaperones

Figure 3. SLC16A11 Is a Proton-Coupled Monocarboxylate Transporter

(A) Sequence alignment of transmembrane domains (TMDs) 1 and 8 from SLC16A11 and SLC16 category I and II family members. Box indicates the SLC16 consensus sequence. Residues conserved in SLC16 category I members are indicated in red. Residues conserved in SLC16 category II members are indicated in blue.

(B) Three-dimensional homology modeling of SLC16A11.

(C) Membrane fractionation of HEK293T cells expressing SLC16A11^{REF}-V5. Equal quantities of protein from each fraction were loaded; the percentage of each fraction loaded is indicated (bottom). Note the higher proportion of plasma-membrane fraction loaded. Fraction markers include Na/K ATPase (plasma membrane), calnexin (endoplasmic reticulum), and tubulin (cytoplasm). Molecular weight markers (kDa) are indicated.

(D–G) Assessment of pyruvate (pyronic) and proton (BCECF-AM) flux in HEK293T cells expressing either SLC16A11^{REF} (light blue) or empty vector control (gray). Pyruvate (0.4 mM) was added and removed, as indicated. (D) Representative traces and (E) bar plots depicting normalized rates of pyruvate influx and efflux ± SEM. *$p < 0.05$, $n = 11$. (F) Corresponding representative traces and (G) bar plots depicting normalized rates of proton influx and efflux ± SEM. *$p < 0.05$, $n = 11$. See also Figure S3.

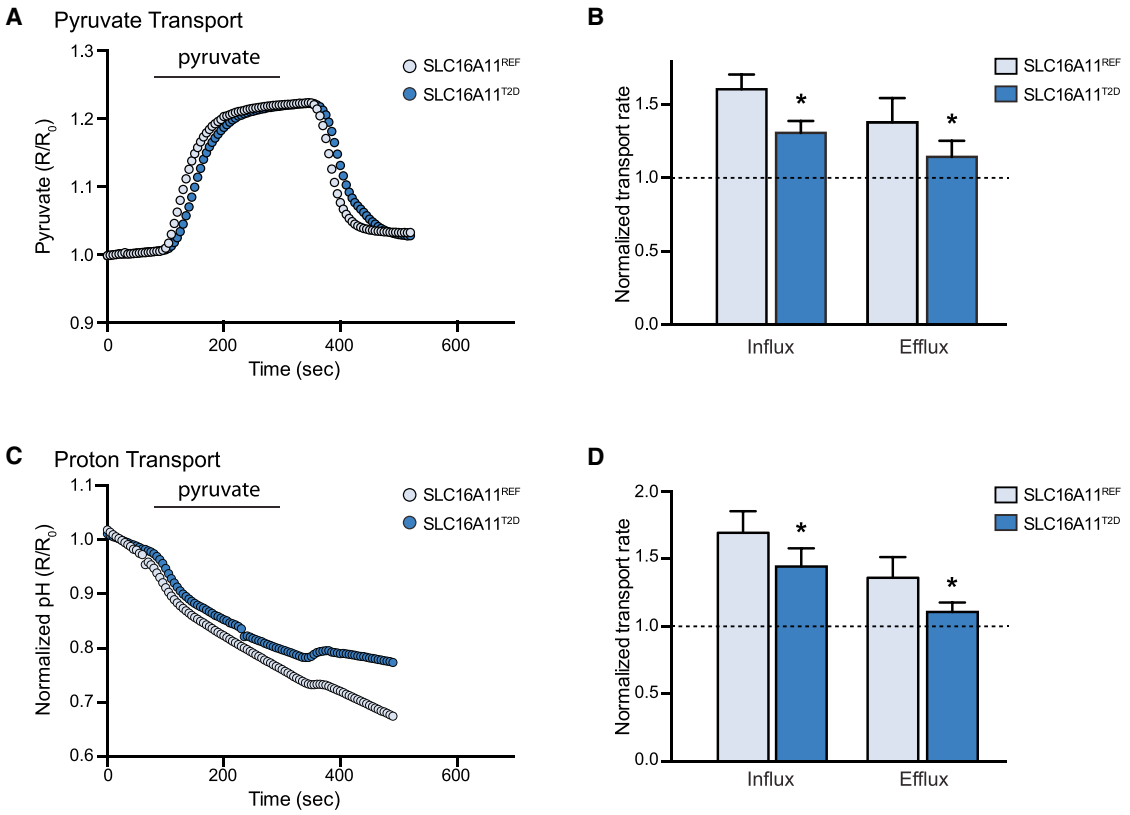


Figure 4. T2D-Risk-Associated Coding Variants Abrogate SLC16A11 Activity

(A–D) Assessment of pyruvate (pyronic) and proton (BCECF-AM) flux in HEK293T cells expressing either SLC16A11^{REF} (light blue) or SLC16A11^{T2D} (dark blue). Pyruvate (0.4 mM) was added and removed, as indicated. (A) Representative traces and (B) bar plots depicting normalized rates of pyruvate influx and efflux ± SEM. The rate of transport is normalized to empty vector control, which is indicated by the dashed line. *$p < 0.05$, $n = 11$. (C) Corresponding representative traces and (D) bar plots depicting normalized rates of proton influx and efflux ± SEM. *$p < 0.05$, $n = 11$. See also Figure S3.

that promote cell-surface localization of SLC16 proteins in category I. The interactions between SLC16A11 and BSG and EMB provide further support that SLC16A11 is a member of category I.

We also analyzed the proteins in the SLC16A11^{REF} interactome using the pathway-based analysis tool GeNets (http://apps.broadinstitute.org/genets) (Figure S4E). This analysis revealed that SLC16A11 interacts with a cluster of proteasome components, which suggests an explanation for the low levels of SLC16A11 found in our transient expression experiments—namely, that the protein undergoes proteasome-mediated degradation (Figures S4B and S4C). The data thus suggest that SLC16A11 protein levels may be tightly regulated.

T2D-Risk Coding Variants Reduce SLC16A11 Localization to the Plasma Membrane by Disrupting an Interaction with BSG

We next identified proteins that interact with SLC16A11^{T2D}. By comparing the results of our mass-spectrometry experiment for the reference version of SLC16A11 with the results for the T2D-risk version of the protein, we found a single protein interaction that was disrupted in the latter case: the interaction with BSG (Figures 5C and S4F; Table S2). Using co-immunoprecipitation assays (performed in the absence of the P2D stabilizing mutation), we confirmed both the interaction between BSG and SLC16A11^{REF}, as well as the reduction of this interaction with SLC16A11^{T2D} (Figures 5D–5F).

BSG plays a key role in plasma-membrane localization of other SLC16 category I family members (Halestrap, 2013). To determine whether BSG performs a similar function for SLC16A11, we used the plasma-membrane extraction assay to examine the localization of SLC16A11 in BSG-knockout HEK293T cell lines. We found that BSG knockout significantly reduces localization of SLC16A11^{REF} to the plasma membrane by ~80% (Figure 5G). We confirmed these results using an orthogonal plasma-membrane localization assay based on a split β-galactosidase reporter (Figures S5A and S5B).

The BSG-dependence of SLC16A11 plasma-membrane localization and the finding that the T2D-risk-associated coding variants disrupt the interaction between SLC16A11 and BSG suggest that these coding variants might lead to reduced localization of SLC16A11 to the plasma membrane. We confirmed that this is indeed the case, as we observe an ~60% reduction in plasma-membrane localization of SLC16A11^{T2D} with respect to SLC16A11^{REF} (Figures 5H and S5C).

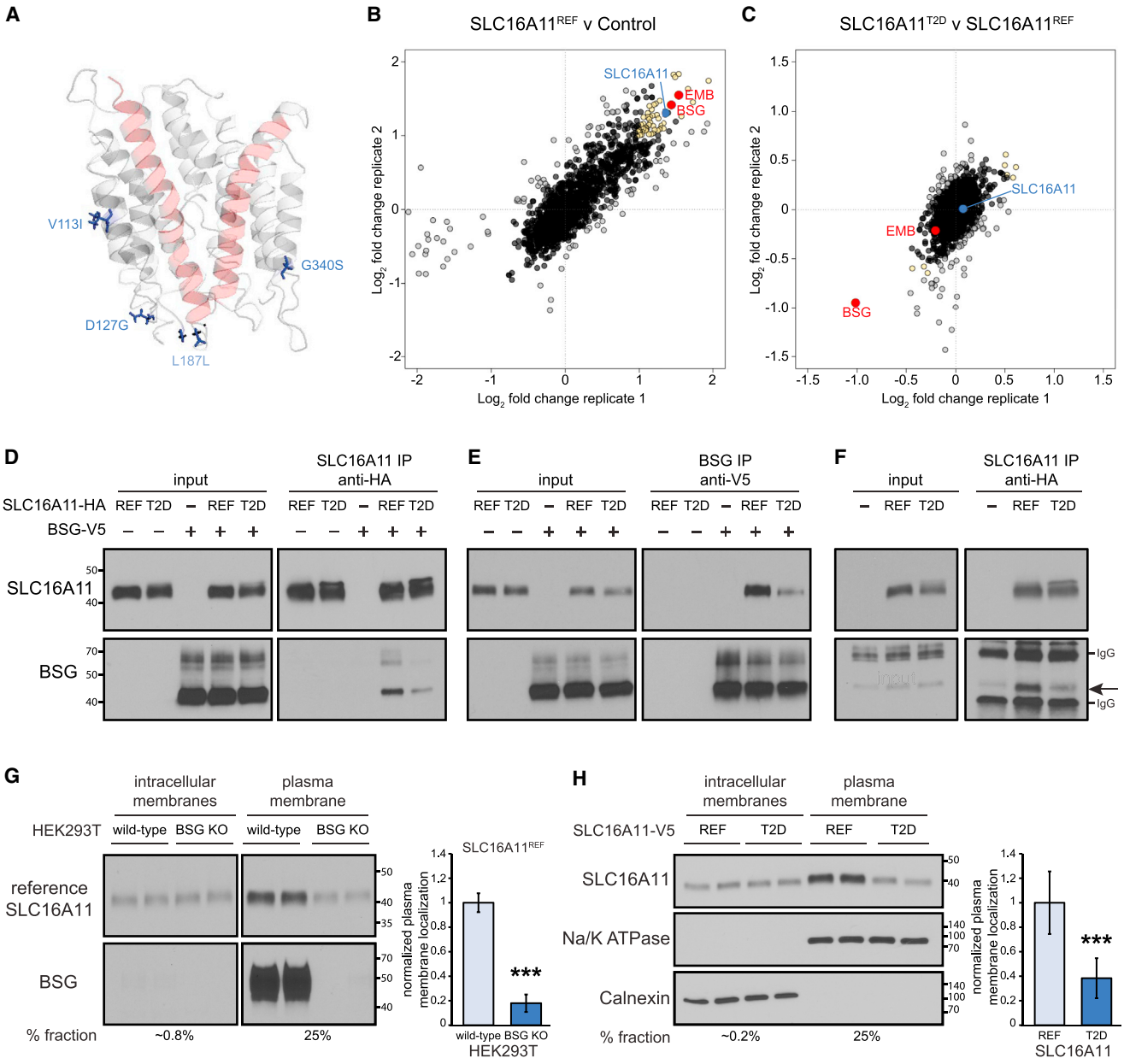


Figure 5. T2D-Risk Coding Variants Reduce Plasma-Membrane Localization by Disrupting an Interaction between SLC16A11 and BSG

(A) Homology model of SLC16A11 with T2D-risk coding variants indicated. P443T is located on an unstructured cytosolic tail and is not included in our model.

(B) Scatterplot showing enrichment of proteins immunoprecipitated from HEK293T cells expressing P2D SLC16A11^{REF}-V5 compared to cells expressing empty vector control. SLC16A11 is shown in blue. BSG and EMB are shown in red along with other highly enriched proteins (top 10% with a Blandt-Altman-adjusted $p < 0.05$) in yellow. Biological replicates from two independent experiments are plotted on the different axes. See also Table S2.

(C) Scatterplot showing relative interaction of proteins with P2D SLC16A11^{T2D}-V5 compared to P2D SLC16A11^{REF}-V5. See also Table S2.

(D–F) Co-immunoprecipitation of SLC16A11 (in the absence of P2D) and BSG. (D) Interaction of BSG-V5 with immunoprecipitated SLC16A11-HA. (E) Interaction of SLC16A11-HA with immunoprecipitated BSG-V5. (F) Interaction of endogenous BSG with immunoprecipitated SLC16A11-HA.

(G) Representative membrane fractionation in wild-type (WT) and BSG-knockout HEK293T cells. Equal quantities of protein from each fraction were loaded. The percentage of each fraction loaded is indicated (bottom). Bar plots depict the relative fraction of SLC16A11 at plasma membrane ± SD (n = 8, ***$p = 3 \times 10^{-12}$).

(H) Representative membrane fractionation in HEK293T cells expressing either SLC16A11^{REF}-V5 or SLC16A11^{T2D}-V5. Equal quantities of protein from each fraction were loaded. The percentage of each fraction loaded is indicated (bottom). Fraction markers include Na/K ATPase (plasma membrane) and calnexin (endoplasmic reticulum). Bar plots depict the relative fraction of SLC16A11 at the plasma membrane ± SD (n = 15, ***$p = 4 \times 10^{-8}$). See also Figures S4 and S5.

Together, our experiments establish that the T2D-risk variants reduce SLC16A11 activity in two distinct ways: (1) the variants decrease *SLC16A11* gene expression, and (2) the coding variants alter the protein in a manner that disrupts its interaction with BSG, decreasing the amount of SLC16A11 protein at the plasma membrane. Under a simple model whereby T2D-risk variants reduce both *SLC16A11* gene expression and plasma-membrane localization each by ~60% per copy of the T2D-risk allele, we estimate that homozygous carriers may have up to ~85% less SLC16A11 at the cell surface. Importantly, these findings suggest the hypothesis that diminished levels of SLC16A11 at the plasma membrane is the causal mechanism of increased T2D risk associated with variants at this locus.

Disruption of *SLC16A11* in Primary Human Hepatocytes Leads to Metabolic Changes Associated with Increased T2D Risk

Given our discovery that T2D-risk variants reduce SLC16A11 function in liver, a key question is how diminished SLC16A11 activity might lead to increased risk of T2D. To begin to elucidate the mechanism, we investigated the cellular metabolic changes induced by decreased SLC16A11 function in hepatocytes. We used pooled small interfering RNAs (siRNAs) to knock down *SLC16A11* expression in primary human hepatocytes and achieved a ~90% reduction in *SLC16A11* with no effect on expression of *SLC16A13*, other category I SLC16 family members (*SLC16A1*, *SLC16A3*, *SLC16A7*, *SLC16A8*), or *BSG* (Figure 6A).

We measured levels of ~350 known polar and lipid metabolites and used enrichment analysis to identify metabolic changes induced by knockdown of *SLC16A11*. Steady-state levels of intracellular acylcarnitines, diacylglycerols (DAGs), and triacylglycerols (TAGs) are significantly increased in human hepatocytes treated with siRNAs targeting *SLC16A11* compared to negative-control siRNAs (false discovery rate [FDR] < 0.05; Figures 6B and S6; Tables S3 and S4). Also increased are *extracellular* levels of TAGs (FDR < 0.05; Figures 6C and S7; Tables S3 and S4), which are secreted by hepatocytes in the form of VLDL (very low density lipoprotein). Together, these changes demonstrate that SLC16A11 affects cellular fatty acid and lipid metabolism, with the increase in acylcarnitines indicating an effect on fatty acid β-oxidation by the mitochondria and with the increases in DAGs and TAGs indicating a shift toward energy storage in the form of glycerolipids (Rui, 2014; Wajner and Amaral, 2015).

Strikingly, the metabolic changes match those seen in the pathophysiology of insulin resistance and T2D: (1) acylcarnitines are elevated in the plasma of people with type 2 diabetes (Adams et al., 2009; Mihalik et al., 2010), (2) DAGs are associated with hepatic and skeletal muscle insulin resistance (Erion and Shulman, 2010), and (3) TAG accumulation in liver and plasma is associated with diabetic insulin resistance and T2D (Rhee et al., 2011; Samuel and Shulman, 2012; Seymour and Byrne, 1993).

In summary, variants on the T2D risk haplotype disrupt SLC16A11 function through two distinct molecular mechanisms and result in changes in cellular metabolism consistent with those seen in insulin resistance and T2D in humans (Figure 6D).

DISCUSSION

Human genetics has the potential to illuminate causal disease mechanisms and lay the foundation for development of therapies that target underlying disease biology (Plenge et al., 2013). Notably, relying on human genetic support for the role of a drug target in disease increases the success rate during clinical development (Nelson et al., 2015). In recent years, numerous new treatments based on genetic insights have been developed for diseases ranging from monogenic disorders, such as cystic fibrosis, to common diseases, such as hypercholesterolemia, coronary artery disease, and T2D (Cohen et al., 2006; Nauck, 2014; Robinson et al., 2015; Sabatine et al., 2015; Wainwright et al., 2015). While the benefit of using genetics to identify new drug targets is substantial, the path from genetic association to causal variant, disease mechanism, and clinical impact is challenging, requiring fine mapping and in-depth investigation into the molecular, cellular, and physiological functions of the gene in question. Here, we present two mechanisms through which T2D-risk-associated variants at 17p13 lead to lower levels and activity of SLC16A11: (1) a *cis*-eQTL reducing *SLC16A11* expression in liver, and (2) a disrupted interaction with a chaperone protein, resulting in less SLC16A11 at the plasma membrane. Both mechanisms point to reduced SLC16A11 as the disease-relevant direction of effect. Our work converges on the therapeutic hypothesis that increasing SLC16A11 function might be beneficial for treating T2D.

Identifying mechanisms through which the T2D-associated variants at 17p13 might affect SLC16A11 function is an important step toward understanding how these variants influence T2D risk; however, much remains unknown. First, we do not yet know which regulatory variant(s) leads to reduced *SLC16A11* expression or which coding variant(s) decreases plasma-membrane localization. Furthermore, it is not known whether the same variant mediates both effects or whether multiple causal variants exist on the T2D-risk haplotype. Second, it is not clear why two distinct mechanisms for disrupting SLC16A11 function have converged on the same haplotype—it is unusual for a disease-associated haplotype to disrupt a gene in two distinct ways. One possible explanation is that decreased function of SLC16A11 conferred a benefit in an ancestral population, resulting in positive selection; this is reminiscent of the thrifty-gene hypothesis (Neel, 1962).

The third—and most important—unanswered question is how decreased SLC16A11 activity increases T2D risk. Our data point to reduced SLC16A11 activity in liver as the causal effect of variants at this locus on T2D risk, and we demonstrate that disruption of *SLC16A11* in primary human hepatocytes induces T2D-relevant metabolic changes in fatty acid and lipid metabolism—results that are in line with our previous observation that ectopic expression of SLC16A11 alters lipid metabolism in HeLa cells (Williams et al., 2014). However, the biochemical mechanism connecting SLC16A11 function to these metabolic processes remains unknown. Moreover, other related family members, including SLC16A1 and SLC16A13, are also expressed in liver (Williams et al., 2014), raising the question as to why SLC16A11 function is distinct. One possibility is that SLC16A11 has a different affinity for shared substrates, suggesting each member of the family is dominant under different conditions.

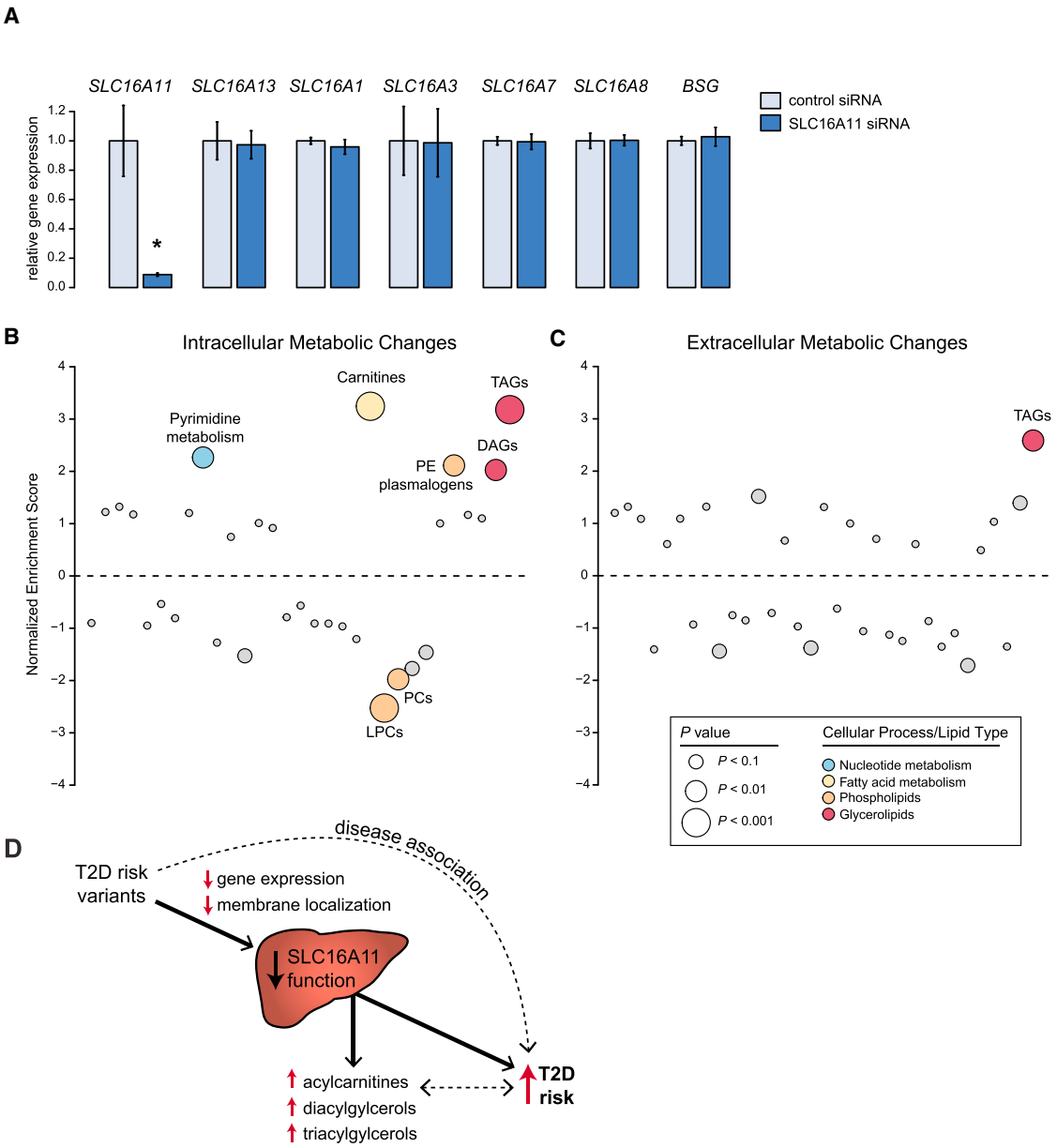


Figure 6. Knockdown of *SLC16A11* in Primary Human Hepatocytes Alters Metabolites Associated with Insulin Resistance and T2D

(A) Gene expression in primary human hepatocytes treated with siRNAs targeting *SLC16A11* or negative control siRNAs. Bar plots depict relative gene expression ± SD using *TBP* for normalization. * $P \approx 4.8 \times 10^{-3}$

(B and C) Enrichment analysis of (B) intracellular and (C) extracellular metabolic pathway changes following *SLC16A11* knockdown in primary human hepatocytes. Each dot represents a different metabolic pathway or metabolite class. *P* values are indicated by dot size. Significantly altered pathways (false discovery rate [FDR] < 0.05) are labeled, with non-significant pathways shown in gray. LPCs, lysophosphatidylcholines; PCs, phosphatidylcholines; PE, phosphatidylethanolamine; DAGs, diacylglycerols; TAGs, triacylglycerols. See also Figures S6 and S7 and Tables S3 and S4.

(D) Depiction summarizing the effects of T2D-associated variants at 17p13 on T2D risk. The T2D disease association at 17p13 is driven by variants that disrupt SLC16A11 function, which itself leads to changes in fatty acid and lipid metabolism that are associated with increased risk of T2D. The causality of the associations between increased acylcarnitines, DAGs, and TAGs and disease are uncertain.

Another possibility is that SLC16A11 may transport an as-yet-unidentified substrate not recognized by other SLC16 family members. While we utilized pyruvate to establish that SLC16A11 is a category I monocarboxylate transporter, we have not cataloged the full range of substrates transported by SLC16A11 nor established which transported substrate is relevant to increased T2D risk in humans. Yet another possibility is that SLC16A11 functions in specific cellular contexts in which metabolic function is challenged. Further studies to elucidate the identity of substrates transported by SLC16A11 and the mechanisms that regulate

SLC16A11 levels and activity, together with investigations into the physiological consequences of Slc16a11 perturbation in animal models and clinical studies of risk carriers, will be needed to provide insight into the cellular role of this transporter, the specific biochemical processes impacted by its disruption, and the resulting impact on human physiology and disease.

Although detailed knowledge of the biochemical, cellular, and physiological mechanisms by which perturbation of SLC16A11 disrupts metabolism will be needed to enable eventual drug development, the therapeutic hypothesis that emerges from the current study already provides a rationale for initiating small-molecule screening to increase SLC16A11 activity. The mechanisms of variant action described here suggest several possible therapeutic approaches, including increasing transcript or protein levels, correcting the SLC16A11-BSG interaction, augmenting plasma-membrane localization, or enhancing transport activity. We note that discovery of agents that mitigate the deleterious effect of this genetic risk factor might be generally beneficial regardless of genotype. While in cases such as cystic fibrosis, corrector and potentiator combination therapy is targeted toward individuals affected by specific mutations (Van Goor et al., 2014; Wainwright et al., 2015), the possibility exists that modulating SLC16A11 could be a more generalizable treatment approach that extends to the broader at-risk population and not just risk-variant carriers. A precedent for this is illustrated by the efficacy of sulfonylureas in treating both individuals with neonatal diabetes who carry activating mutations in the sulfonylurea receptor complex as well as the general population of people with T2D (Gloyn et al., 2004; Inzucchi, 2002; Pearson et al., 2003). In conclusion, this study exemplifies the power of an unbiased genetic approach to illuminate new mechanisms of disease and reveal actionable therapeutic hypotheses that may potentially benefit not only risk-variant carriers but also the global population at risk of developing diabetes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
  - Cell Lines
  - Primary Cell Cultures
- METHOD DETAILS
  - Genotyping
  - Visceral Adipose and Liver Tissue Collection
  - Droplet Digital PCR
  - ChIP-Seq
  - Generation of cell lines
  - Plasmids
  - Transient Transfections
  - Lentiviral Transduction
  - Sequence Alignment and Homology Modeling
  - Live Cell Imaging
  - Analysis of SLC16 Protein Levels
  - Membrane Fractionation
  - PathHunter MEM-EA Pharmacotrafficking Assay
  - Immunoprecipitations for Protein Interactions
  - Western Blot Analyses
  - Proteomic Methods for SLC16A11 Interactors
  - *SLC16A11* Knockdown for Metabolite Profiling
  - Metabolite Profiling
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data Integration and Credible Set Analysis
  - Association Analysis in MEDIA
  - eQTL and Allele-Specific Expression Analyses
  - ChIP-Seq Analysis
  - Metabolomic Analyses
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2017.06.011.

A video abstract is available at http://dx.doi.org/10.1016/j.cell.2017.06.011#mmc7.

CONSORTIA

MEDIA Consortium: Maggie C.Y. Ng, Daniel Shriner, Brian H. Chen, Jiang Li, Wei-Min Chen, Xiuqing Guo, Jiankang Liu, Suzette J. Bielinski, Lisa R. Yanek, Michael A. Nalls, Mary E. Comeau, Laura J. Rasmussen-Torvik, Richard A. Jensen, Daniel S. Evans, Yan V. Sun, Ping An, Sanjay R. Patel, Yingchang Lu, Jirong Long, Loren L. Armstrong, Lynne Wagenknecht, Lingyao Yang, Beverly M. Snively, Nicholette D. Palmer, Poorva Mudgal, Carl D. Langefeld, Keith L. Keene, Barry I. Freedman, Josyf C. Mychaleckyj, Uma Nayak, Leslie J. Raffel, Mark O. Goodarzi, Y-D Ida Chen, Herman A. Taylor Jr., Adolfo Correa, Mario Sims, David Couper, James S. Pankow, Eric Boerwinkle, Adebowale Adeyemo, Ayo Doumatey, Guanjie Chen, Rasika A. Mathias, Dhananjay Vaidya, Andrew B. Singleton, Alan B. Zonderman, Robert P. Igo Jr., John R. Sedor, the FIND Consortium, Edmond K. Kabagambe, David S. Siscovick, Barbara McKnight, Kenneth Rice, Yongmei Liu, Wen-Chi Hsueh, Wei Zhao, Lawrence F. Bielak, Aldi Kraja, Michael A. Province, Erwin P. Bottinger, Omri Gottesman, Qiuyin Cai, Wei Zheng, William J. Blot, William L. Lowe, Jennifer A. Pacheco, Dana C. Crawford, the eMERGE Consortium, the DIAGRAM Consortium, Elin Grundberg, the MuTHER Consortium, Stephen S. Rich, M. Geoffrey Hayes, Xiao-Ou Shu, Ruth J.F. Loos, Ingrid B. Borecki, Patricia A. Peyser, Steven R. Cummings, Bruce M. Psaty, Myriam Fornage, Sudha K. Iyengar, Michele K. Evans, Diane M. Becker, W.H. Linda Kao, James G. Wilson, Jerome I. Rotter, Michèle M. Sale, Simin Liu, Charles N. Rotimi, Donald W. Bowden

SIGMA T2D Consortium: Josep M. Mercader, Alicia Huerta-Chagoya, Hortensia Moreno-Macías, Humberto García-Ortiz, Alisa Manning, Lizz Caulkins, Jason Flannick, Nick Patterson, Angélica Martínez-Hernández, Federico Centeno-Cruz, Francisco Martin Barajas-Olmos, Carlos Zerrweck, Cecilia Contreras-Cubas, Elvia Mendoza-Caamal, Cristina Revilla-Monsalve, Sergio Islas-Andrade, Emilio Córdova, Xavier Soberón, María Elena González-Villalpando, Lynne Wilkens, Loic Le Marchand, Kristine Monroe, Laurence Kolonel, Olimpia Arellano-Campos, Maria L. Ordóñez-Sánchez, Maribel Rodríguez-Torres, Yayoi Segura-Kato, Rosario Rodríguez-Guillén, Ivette Cruz-Bautista, Linda Liliana Muñoz-Hernandez, Tamara Sáenz, Donají Gómez, Ulices Alvirde, Paloma Almeda-Valdés, Noël P. Burtt, Maria L. Cortes, David M. Altshuler, Christopher A. Haiman, Carlos A. Aguilar-Salinas, Clicerio González-Villalpando, Lorena Orozco, Teresa Tusié-Luna, Jose C. Florez

AUTHOR CONTRIBUTIONS

V.R., E.H., D.M.A., S.L.S., J.C.F., S.B.R.J., and E.S.L. conceived, planned, and oversaw the study. V.R. and E.H. designed, performed, and analyzed most experiments. S.B.R.J assisted with experimental design and analyses. D.E.T.

and V.R. performed ChIP-seq, which M.G. analyzed. C.R.H., G.G., M.S., and S.A.C conducted the proteomic methods. A.A.D., K.A.P., C.D., and C.B.C. carried out metabolite profiling methods. V.R. and M.D. performed split β-galactosidase reporter assays, with supervision from B.K.W. J.M.M. and P.F. performed genetic analyses. M.v.G. performed homology modeling analyses. A.S. provided technical support. C.Z., F.C.-C., and L.O. provided tissue samples. M.C.Y.N., B.H.C., and the MEDIA and SIGMA T2D consortia provided genetic data. V.R., E.H., S.L.S., J.C.F., S.B.R.J., and E.S.L. analyzed the overall data and wrote the manuscript. Consortia member affiliations can be found in Document S6 of the Supplemental Information.

ACKNOWLEDGMENTS

The authors thank Jason Wright, Jessica Gasser, Jamie Marshall, and Daniel O'Connell for many helpful discussions; Bryan MacDonald for discussions and critical reading of the manuscript; Kasper Lage for guidance on GeNets analysis; Zach Dymek for lab management support; and Lior Friedman and Leslie Gaffney for graphic design input. We also acknowledge the many discussions with the Lander, Schreiber, Altshuler, and Florez laboratories and the Diabetes Research Group at the Broad Institute. The Meta-analysis of Type 2 Diabetes in African Americans (MEDIA) Consortium was partly supported by NIH R01 DK066358 to Donald Bowden. J.M.M. was supported by Beatriu de Pinós fellowship from the Agency for Management of University and Research Grants (2014 BP-B 00227). E.H. is a Klarman Family Foundation fellow. E.S.L. is supported in part by NIH UM1HG008895. This work was conducted as part of the Slim Initiative for Genomic Medicine, a project funded by the Carlos Slim Foundation in Mexico. S.L.S. is a cofounder and a scientific advisor to Jnana Therapeutics. V.R. is now an employee of Jnana Therapeutics. E.S.L. has no financial conflict of interest related to work on SLC16A11 or type 2 diabetes.

Received: October 6, 2016
Revised: March 16, 2017
Accepted: June 8, 2017
Published: June 29, 2017

REFERENCES

Adams, S.H., Hoppel, C.L., Lok, K.H., Zhao, L., Wong, S.W., Minkler, P.E., Hwang, D.H., Newman, J.W., and Garvey, W.T. (2009). Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. J. Nutr. *139*, 1073–1081.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169–181.

Cohen, J.C., Boerwinkle, E., Mosley, T.H., Jr., and Hobbs, H.H. (2006). Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N. Engl. J. Med. *354*, 1264–1272.

Delaneau, O., Zagury, J.F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. Nat. Methods *10*, 5–6.

Diamond, J. (2003). The double puzzle of diabetes. Nature *423*, 599–602.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Erion, D.M., and Shulman, G.I. (2010). Diacylglycerol-mediated insulin resistance. Nat. Med. *16*, 400–402.

Estrada, K., Aukrust, I., Bjørkhaug, L., Burtt, N.P., Mercader, J.M., García-Ortiz, H., Huerta-Chagoya, A., Moreno-Macías, H., Walford, G., Flannick, J., et al.; SIGMA Type 2 Diabetes Consortium (2014). Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. JAMA *311*, 2305–2314.

Flannick, J., and Florez, J.C. (2016). Type 2 diabetes: genetic data sharing to advance complex disease research. Nat. Rev. Genet. *17*, 535–549.

Flannick, J., Thorleifsson, G., Beer, N.L., Jacobs, S.B., Grarup, N., Burtt, N.P., Mahajan, A., Fuchsberger, C., Atzmon, G., Benediktsson, R., et al.; Go-T2D Consortium; T2D-GENES Consortium (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. Nat. Genet. *46*, 357–363.

Gloyn, A.L., Pearson, E.R., Antcliff, J.F., Proks, P., Bruining, G.J., Slingerland, A.S., Howard, N., Srinivasan, S., Silva, J.M., Molnes, J., et al. (2004). Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. N. Engl. J. Med. *350*, 1838–1849.

Halestrap, A.P. (2013). The SLC16 gene family - structure, role and regulation in health and disease. Mol. Aspects Med. *34*, 337–349.

Hara, K., Fujita, H., Johnson, T.A., Yamauchi, T., Yasuda, K., Horikoshi, M., Peng, C., Hu, C., Ma, R.C., Imamura, M., et al.; DIAGRAM consortium (2014). Genome-wide association study identifies three novel loci for type 2 diabetes. Hum. Mol. Genet. *23*, 239–246.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hugo, S.E., Cruz-Garcia, L., Karanth, S., Anderson, R.M., Stainier, D.Y., and Schlegel, A. (2012). A monocarboxylate transporter required for hepatocyte secretion of ketone bodies during fasting. Genes Dev. *26*, 282–293.

Inzucchi, S.E. (2002). Oral antihyperglycemic therapy for type 2 diabetes: scientific review. JAMA *287*, 360–372.

Kurowski, M.A., and Bujnicki, J.M. (2003). GeneSilico protein structure prediction meta-server. Nucleic Acids Res. *31*, 3305–3307.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lemieux, M.J., Song, J., Kim, M.J., Huang, Y., Villa, A., Auer, M., Li, X.D., and Wang, D.N. (2003). Three-dimensional crystallization of the Escherichia coli glycerol-3-phosphate transporter: a member of the major facilitator superfamily. Protein Sci. *12*, 2748–2756.

Locke, J.M., Hysenaj, G., Wood, A.R., Weedon, M.N., and Harries, L.W. (2015). Targeted allelic expression profiling in human islets identifies cis-regulatory effects for multiple variants identified by type 2 diabetes genome-wide association studies. Diabetes *64*, 1484–1491.

Manoharan, C., Wilson, M.C., Sessions, R.B., and Halestrap, A.P. (2006). The role of charged residues in the transmembrane helices of monocarboxylate transporter 1 and its ancillary protein basigin in determining plasma membrane expression and catalytic activity. Mol. Membr. Biol. *23*, 486–498.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic localization of common disease-associated variation in regulatory DNA. Science *337*, 1190–1195.

Mihalik, S.J., Goodpaster, B.H., Kelley, D.E., Chace, D.H., Vockley, J., Toledo, F.G., and DeLany, J.P. (2010). Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity (Silver Spring) *18*, 1695–1700.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Moltke, I., Grarup, N., Jørgensen, M.E., Bjerregaard, P., Treebak, J.T., Fumagalli, M., Korneliussen, T.S., Andersen, M.A., Nielsen, T.S., Krarup, N.T., et al. (2014). A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. Nature *512*, 190–193.

Morris, A.P., Voight, B.F., Teslovich, T.M., Ferreira, T., Segrè, A.V., Steinthorsdottir, V., Strawbridge, R.J., Khan, H., Grallert, H., Mahajan, A., et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (2012). Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat. Genet. *44*, 981–990.

Nauck, M.A. (2014). Update on developments with SGLT2 inhibitors in the management of type 2 diabetes. Drug Des. Devel. Ther. *8*, 1335–1380.

Neel, J.V. (1962). Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am. J. Hum. Genet. *14*, 353–362.

Nelson, M.R., Tipney, H., Painter, J.L., Shen, J., Nicoletti, P., Shen, Y., Floratos, A., Sham, P.C., Li, M.J., Wang, J., et al. (2015). The support of human genetic evidence for approved drug indications. Nat. Genet. *47*, 856–860.

Ng, M.C., Shriner, D., Chen, B.H., Li, J., Chen, W.M., Guo, X., Liu, J., Bielinski, S.J., Yanek, L.R., Nalls, M.A., et al.; FIND Consortium; eMERGE Consortium; DIAGRAM Consortium; MuTHER Consortium; MEta-analysis of type 2 DIabetes in African Americans Consortium (2014). Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. PLoS Genet. *10*, e1004517.

Ovens, M.J., Davies, A.J., Wilson, M.C., Murray, C.M., and Halestrap, A.P. (2010). AR-C155858 is a potent inhibitor of monocarboxylate transporters MCT1 and MCT2 that binds to an intracellular site involving transmembrane helices 7-10. Biochem. J. *425*, 523–530.

Pearson, E.R., Starkey, B.J., Powell, R.J., Gribble, F.M., Clark, P.M., and Hattersley, A.T. (2003). Genetic cause of hyperglycaemia and response to treatment in diabetes. Lancet *362*, 1275–1281.

Perry, R.J., Samuel, V.T., Petersen, K.F., and Shulman, G.I. (2014). The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes. Nature *510*, 84–91.

Plenge, R.M., Scolnick, E.M., and Altshuler, D. (2013). Validating therapeutic targets through human genetics. Nat. Rev. Drug Discov. *12*, 581–594.

Prasad, R.B., and Groop, L. (2015). Genetics of type 2 diabetes-pitfalls and possibilities. Genes (Basel) *6*, 87–123.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Rhee, E.P., Cheng, S., Larson, M.G., Walford, G.A., Lewis, G.D., McCabe, E., Yang, E., Farrell, L., Fox, C.S., O'Donnell, C.J., et al. (2011). Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. J. Clin. Invest. *121*, 1402–1411.

Rink, T.J., Tsien, R.Y., and Pozzan, T. (1982). Cytoplasmic pH and free Mg2+ in lymphocytes. J. Cell Biol. *95*, 189–196.

Robinson, J.G., Farnier, M., Krempf, M., Bergeron, J., Luc, G., Averna, M., Stroes, E.S., Langslet, G., Raal, F.J., El Shahawy, M., et al.; ODYSSEY LONG TERM Investigators (2015). Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N. Engl. J. Med. *372*, 1489–1499.

Rui, L. (2014). Energy metabolism in the liver. Compr. Physiol. *4*, 177–197.

Rychlewski, L., Jaroszewski, L., Li, W., and Godzik, A. (2000). Comparison of sequence profiles. Strategies for structural predictions using sequence information. Protein Sci. *9*, 232–241.

Sabatine, M.S., Giugliano, R.P., Wiviott, S.D., Raal, F.J., Blom, D.J., Robinson, J., Ballantyne, C.M., Somaratne, R., Legg, J., Wasserman, S.M., et al.; Open-Label Study of Long-Term Evaluation against LDL Cholesterol (OSLER) Investigators (2015). Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. N. Engl. J. Med. *372*, 1500–1509.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Samuel, V.T., and Shulman, G.I. (2012). Mechanisms for insulin resistance: common threads and missing links. Cell *148*, 852–871.

San Martín, A., Ceballo, S., Baeza-Lehnert, F., Lerchundi, R., Valdebenito, R., Contreras-Baeza, Y., Alegría, K., and Barros, L.F. (2014). Imaging mitochondrial flux in single cells with a FRET sensor for pyruvate. PLoS ONE *9*, e85780.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat. Methods *11*, 783–784.

Seymour, C.A., and Byrne, C.D. (1993). Triglycerides and disease. Postgrad. Med. J. *69*, 679–695.

Shepherd, M., Shields, B., Ellard, S., Rubio-Cabezas, O., and Hattersley, A.T. (2009). A genetic diagnosis of HNF1A diabetes alters treatment and improves glycaemic control in the majority of insulin-treated patients. Diabet. Med. *26*, 437–441.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Suhre, K., Shin, S.Y., Petersen, A.K., Mohney, R.P., Meredith, D., Wägele, B., Altmaier, E., Deloukas, P., Erdmann, J., Grundberg, E., et al.; CARDIoGRAM (2011). Human metabolic individuality in biomedical and pharmaceutical research. Nature *477*, 54–60.

Traurig, M., Hanson, R.L., Marinelarena, A., Kobes, S., Piaggi, P., Cole, S., Curran, J.E., Blangero, J., Göring, H., Kumar, S., et al. (2016). Analysis of SLC16A11 Variants in 12,811 American Indians: Genotype-Obesity Interaction for Type 2 Diabetes and an Association With RNASEK Expression. Diabetes *65*, 510–519.

van de Geijn, B., McVicker, G., Gilad, Y., and Pritchard, J.K. (2015). WASP: allele-specific software for robust molecular quantitative trait locus discovery. Nat. Methods *12*, 1061–1063.

Van Goor, F., Yu, H., Burton, B., and Hoffman, B.J. (2014). Effect of ivacaftor on CFTR forms with missense mutations associated with defects in protein processing or function. J. Cyst. Fibros. *13*, 29–36.

Villalpando, S., de la Cruz, V., Rojas, R., Shamah-Levy, T., Avila, M.A., Gaona, B., Rebollar, R., and Hernández, L. (2010). Prevalence and distribution of type 2 diabetes mellitus in Mexican adult population: a probabilistic survey. Salud Publica Mex. *52*(*Suppl 1*), S19–S26.

Visser, W.E., Philp, N.J., van Dijk, T.B., Klootwijk, W., Friesema, E.C., Jansen, J., Beesley, P.W., Ianculescu, A.G., and Visser, T.J. (2009). Evidence for a homodimeric structure of human monocarboxylate transporter 8. Endocrinology *150*, 5163–5170.

Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al.; MAGIC investigators; GIANT Consortium (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. *42*, 579–589.

von Grotthuss, M., Pas, J., Wyrwicz, L., Ginalski, K., and Rychlewski, L. (2003). Application of 3D-Jury, GRDB, and Verify3D in fold recognition. Proteins *53*(*Suppl 6*), 418–423.

Wainwright, C.E., Elborn, J.S., Ramsey, B.W., Marigowda, G., Huang, X., Cipolli, M., Colombo, C., Davies, J.C., De Boeck, K., Flume, P.A., et al.; TRAFFIC Study Group; TRANSPORT Study Group (2015). Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. N. Engl. J. Med. *373*, 220–231.

Wajner, M., and Amaral, A.U. (2015). Mitochondrial dysfunction in fatty acid oxidation disorders: insights from human and animal studies. Biosci. Rep. *36*, e00281.

Wakefield, J. (2007). A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am. J. Hum. Genet. *81*, 208–227.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.

Williams, A.L., Jacobs, S.B., Moreno-Macías, H., Huerta-Chagoya, A., Churchhouse, C., Márquez-Luna, C., García-Ortíz, H., Gómez-Vázquez, M.J., Burtt, N.P., Aguilar-Salinas, C.A., et al.; SIGMA Type 2 Diabetes Consortium (2014). Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. Nature *506*, 97–101.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-H3K27ac	Active Motif	Cat#39133; RRID: AB_2561016
Rabbit anti-H3K4me1	Cell Signaling Technologies	Cat#5326BF; RRID: AB_2616017
Rabbit anti-H3K4me3 (D1A9)	Cell Signaling Technologies	Cat#9751BF; RRID: AB_2616028
Rabbit anti-V5 (D3H8Q)	Cell Signaling Technologies	Cat#13202
Rabbit anti-HA (C29F4)	Cell Signaling Technologies	Cat#3724; RRID: AB_1549585
Rabbit anti-BSG	Cell Signaling Technologies	Cat#12314
Mouse anti-Na/KATPase	Abcam	Cat#ab7671; RRID: AB_306023
Rabbit anti-Calnexin	Cell Signaling Technologies	Cat#2433; RRID: AB_2243887
Rabbit anti-Tubulin	Abcam	Cat#ab21058; RRID: AB_727045
Rabbit anti-Vinculin	Abcam	Cat#ab129002; RRID: AB_11144129
Anti-V5 Agarose Affinity Gel	Sigma	Cat#A7345; RRID: AB_10062721
Anti-HA Agarose Affinity Gel	Sigma	Cat#A2095; RRID: AB_257974
Biological Samples		
Primary human hepatocytes	BioreclamationIVT	See Table S5
Liver and visceral adipose tissue samples	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
BCECF-AM	Invitrogen	Cat#B1150
Cycloheximide	Cayman Chemical	Cat#14126
MG132	Cayman Chemical	Cat#10012628
AR-C155858	Tocris	Cat#4960
Sodium pyruvate	Sigma	Cat#P5280
NuPAGE LDS Sample Buffer	Thermo Fisher Scientific	Cat#NP0008
NuPAGE Sample Reducing Agent	Thermo Fisher Scientific	Cat#NP0004
Critical Commercial Assays		
miRNeasy Mini Kit	QIAGEN	Cat#217004
Gentra Puregene	QIAGEN	Cat#158467
RNeasy Mini Kit	QIAGEN	Cat#74106
High-Capacity RNA-to-cDNA kit	Thermo Fisher Scientific	Cat#4387406
ddPCR Supermix for Probes (no dUTP)	BioRad	Cat#186-3024
End-It DNA End-Repair Kit	Epicenter	Cat#ER0720
Klenow Fragment (3′-5′ exo-)	New England BioLabs	Cat#M0212L
DNA ligase	New England BioLabs	Cat#M2200S
PfuUltra II Hotstart PCR Master Mix	Agilent Technologies	Cat#600850
NextSeq 500/550 High Output Kit v2 (300 cycles)	Illumina	Cat#FC-404-2004
Plasma Membrane Protein Extraction Kit	Abcam	Cat#ab65400
PathHunter Detection kit	DiscoverX	Cat# 93-0001
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific	Cat#23225
Pierce ECL Western Blotting Substrate	Thermo Fisher Scientific	Cat#32106
SuperSignal West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific	Cat#34080
Deposited Data		
Chromatin Immunoprecipitation-Sequencing	NCBI GEO	GEO: GSE99301
Protein Interaction Dataset	MassIVE: http://massive.ucsd.edu/ProteoSAFe/static/massive.jsp	Identifier: MSV000081105
SIGMA Genetic Data	Type 2 Diabetes Knowledge Portal	http://www.type2diabetesgenetics.org

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
HEK293T	Broad Genetic Perturbation Platform	N/A
BSG-knockout HEK293T #1	This paper	N/A
BSG-knockout HEK293T #2	This paper	N/A
HuH7	Cancer Cell Line Encyclopedia (CCLE)	N/A
HuH7 cells stably expressing SLC16A11^{REF}-V5	This paper	N/A
U2OS MEM-EA cells	DiscoverX	Cat#93-1101C3
Oligonucleotides		
rs13342692 TaqMan SNP Genotyping Assay	Applied Biosystems	Assay ID: C__25760519_10
rs13342232 TaqMan SNP Genotyping Assay	Applied Biosystems	Assay ID: C__31793671_10
ddPCR probes	See Table S6	N/A
BSG guide RNA #1 5′-TGGATGTTGGCCGTGCCCAT-3′	This paper	N/A
BSG guide RNA #2 5′-CACCTGTCACTGACTGGGCC-3′	This paper	N/A
Accell Human *SLC16A11* siRNA SMARTpool	GE Dharmacon	Cat#E-007404-00-0050
Accell Non-targeting Pool	GE Dharmacon	Cat#D-001910-10-50
Recombinant DNA		
pLX304-empty	This paper	N/A
pLX304-SLC16A11^{REF}-V5	This paper; Williams et al., 2014	N/A
pLX304-SLC16A11^{T2D}-V5	This paper	N/A
pLX304-K7A SLC16A11-V5	This paper	N/A
pLX304-P2D SLC16A11^{REF}-V5	This paper	N/A
pLX304-P2D SLC16A11^{T2D}-V5	This paper	N/A
pLX304-SLC16A13-V5	Williams et al., 2014	Clone ID: ccsbBroad304_09813
pLX304-BSG-V5	This paper	Clone ID: ccsbBroad304_05904
pLX304-SLC16A1-V5	Williams et al., 2014	N/A
pLX304-SLC16A11-HA	This paper	N/A
Pyronic	San Martín et al., 2014	Addgene plasmid #51308
LentiCRISPRv2	Sanjana et al., 2014	Addgene plasmid #52961
pCMV-ProLink2	DiscoverX	Cat#93-0491
SLC16A11^{REF}-V5-ProLink2	This paper	N/A
SLC16A11^{T2D}-V5-ProLink2	This paper	N/A
Software and Algorithms		
Birdsuite	Broad Institute	https://www.broadinstitute.org/birdsuite/birdsuite
SHAPEITv2	Delaneau et al., 2013	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html
Metal	Willer et al., 2010	http://genome.sph.umich.edu/wiki/METAL_Program
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
HOMER	Heinz et al., 2010	http://homer.ucsd.edu/homer/
Allele-specific ChIP-sequencing analysis pipeline	This paper	https://github.com/mgymrek/mgymrek-sigma-hetchip
WASP	van de Geijn et al., 2015	https://github.com/bmvdgeijn/WASP
Clustal Omega	Sievers et al., 2011	http://www.ebi.ac.uk/Tools/msa/clustalo/
GeneSilico Metaserver	Kurowski and Bujnicki, 2003	https://genesilico.pl/meta2/
MODELER	Sali and Blundell, 1993	https://salilab.org/modeller/
TraceFinder 3.3	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30493

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Progenesis QI	Nonlinear Dynamics	http://www.nonlinear.com/progenesis/qi/
MultiQuant 2.1.1	Sciex	https://sciex.com/products/software/multiquant-software
Kaleidagraph	Synergy Software	N/A
ImageJ	NIH	https://imagej.nih.gov/ij/download.html
GSEA Tool	Subramanian et al., 2005	http://software.broadinstitute.org/gsea/index.jsp
R	R Project	https://www.r-project.org/
Bioconductor	Limma package	http://www.bioconductor.org/
Spectrum Mill	Agilent technologies	v6.0 pre-release
Other		
Dynabeads Protein A	Thermo Fisher Scientific	Cat#1002D
Dynabeads Protein G	Thermo Fisher Scientific	Cat#1004D
AMPure XP beads	Beckman Coulter	Cat#A63881
iTRAQ reagents	AB Sciex	Cat# 4352135
1.9 μm ReproSil-Pur C18-AQ medium	Dr. Maisch GmbH	Size 1.9 μm
ACQUITY UPLC BEH C8 Column, 1.7 μm, 2.1 mm X 100 mm	Waters	Cat#186002878
Atlantis Silica HILIC Column, 100Å, 3 μm, 2.1 mm X 150 mm	Waters	Cat#186002015
Luna 5 μm NH2 100 Å, LC Column 150 × 2 mm	Phenomenex	Cat#00F-4378-B0

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by Suzanne Jacobs (sjacobs@broadinstitute.org). Certain materials are shared with academic and non-profit research organizations for research and educational purposes only under an MTA to be discussed in good faith with the recipient.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Study participants have been described previously (Estrada et al., 2014; Williams et al., 2014). All contributing studies were approved by their respective local ethics committees.

Cell Lines

HEK293T (human embryonic kidney), HuH7 (human hepatoma), and U2OS MEM-EA cells (DiscoverX) cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS), 100 μg/ml Streptomycin and 100 U/ml Penicillin. Cells were grown in a humidified CO2 incubator, at 37°C.

Primary Cell Cultures

Primary human hepatocytes were purchased from BioreclamationIVT. Lots heterozygous for the T2D risk haplotype include ACB, BEB, DSX, NQA, PAA, and QSK. Lots heterozygous for the African haplotype include AIH, FRY, GEB, JLP, KDD, NRE, ZBG, and ZXO. Genotyping at rs13342232 and rs75493593 from BioreclamationIVT was used to infer heterozygosity for the T2D haplotype (heterozygous for the alternative allele at both SNPs) and the African haplotype (heterozygous for the alternative allele at rs13342232 and homozygous reference at rs75493593). Cells were thawed and immediately resuspended in CP media (BioreclamationIVT). Cell concentration and viability were assessed prior to use.

METHOD DETAILS

Genotyping

Genotyping of study participants using the Illumina OMNI2.5 array (Williams et al., 2014) and exome-sequencing (Estrada et al., 2014) have been described previously. The Genomics Platform at the Broad Institute (Cambridge, MA) received, quality controlled, and

tracked DNA samples, and carried out exome array processing. The samples were plated into 96-well plates that included a quality control sample for processing on the Illumina HumanExome BeadChip (Illumina, Inc. San Diego, CA) using manufacturer's protocols. The arrays were scanned using Illumina iScans.

Genotypes were called using Birdsuite (https://www.broadinstitute.org/birdsuite/birdsuite). Clusters were fit using the Birdseed algorithm to each genotyping plate independently. Genotypes with confidence below 99.9% were excluded from analysis (e.g., considered ''missing'' or ''no-call'' genotypes). Samples with low numbers of non-reference alleles (< ~20,000, depending on the cohort), low call rate (< 99.3%) or unusually high heterozygosity (> ~0.05, depending on the cohort) were removed from subsequent analysis; thresholds were chosen based on visual inspection of the sample distributions. Variants with low call rate (< 99.2%) or mean confidence for alternative genotype calls (< 99%) were also excluded from subsequent analysis.

Visceral Adipose and Liver Tissue Collection

Visceral adipose (VAT) and liver samples were collected from subjects undergoing bariatric surgery for severe obesity (BMI greater than 40 kg/m2, or greater than 35 kg/m2 with comorbid entities) or elective surgery in non-obese patients. Patients were selected for bariatric surgery after 6 months of rigorous lifestyle intervention, and all were free of medication 24 hr prior to surgery. All individuals were Mexican Mestizos older than 18 years, carefully selected from the Integral Clinic of Surgery for Obesity and Metabolic Diseases or General Surgery Department at the Tláhuac Hospital in Mexico City. Tissue samples were obtained at the beginning of the surgery with harmonic scalpel in all cases as follow: visceral fat was obtained from the greater omentum at the middle of the greater curvature of the stomach. Liver biopsy was obtained at the distal end of the left hepatic lobe, just above the spleen. VAT and liver samples were frozen immediately after removal. The protocol for collecting VAT and liver samples was approved by the respective local research and ethics committees and all patients signed an informed consent form. The Broad Genomics Platform extracted RNA from frozen tissue samples using the miRNeasy Mini Kit (QIAGEN).

For genotyping, genomic DNA was purified from whole blood samples using a modified salting-out precipitation method (Gentra Puregene, QIAGEN Systems, Inc., Valencia, CA, USA). Genotyping of variants rs13342692 (Assay ID: C__25760519_10) and rs13342232 (Assay ID: C__31793671_10) were performed using TaqMan SNP Genotyping Assay (Applied Biosystems, Foster City, CA, USA). Five previously genotyped samples were added in all plates as positive controls.

Droplet Digital PCR

Total RNA was extracted using miRNeasy Mini Kits or RNeasy Mini Kits (QIAGEN). RNA was DNase treated and converted into cDNA using the High-Capacity RNA-to-cDNA kit (Thermo Fisher Scientific). The following FAM-labeled, TaqMan Real-Time PCR assays (Thermo Fisher Scientific) were used to quantify gene expression: *RNASEK* (Hs00947009_m1 and Hs00947010_g1), *BCL6B* (Hs00394655_m1 and Hs00960914_g1), *SLC16A13* (Hs00416832_m1 and Hs00914030_m1), *SLC16A11* (Hs00601062_g1 and Hs01558330_g1), *CLEC10A* (Hs00924864_g1 and Hs00197107_m1), *SLC16A1* (Hs01560299_m1), *SLC16A3* (Hs00358829_m1), *SLC16A7* (Hs00940851_m1), *SLC16A8* (Hs00895133_g1), *BSG* (Hs00936295_m1), and *EMB* (Hs00904660_m1). A VIC-labeled *TBP* probe (Hs00427620_m1) was used for normalization. For allele-specific expression experiments, droplet digital PCR (ddPCR) assays (Fwd primer: 5′-AGGCAGCCAGCCC-3′; Rev primer: 5′-CCGAGGTAGAGATGCAG-3′) that distinguish the *SLC16A11* reference (5′-TTTCGCCAGCGATCTG-3′; HEX-labeled probe) and T2D risk (5′-TCGCCAGCGGTCTG-3′; FAM-labeled probe) alleles at rs13342692 were custom designed by BioRad. Droplets were generated and analyzed using a QX200 Droplet Generator and Reader system (BioRad). Data was extracted using QuantaSoft (BioRad) and analyzed using Microsoft Excel.

ChIP-Seq

Chromatin immunoprecipitation (ChIP)-sequencing for H3K27ac, H3K4me1, and H3K4me3 were performed on primary human hepatocytes (lots ACB, DSX, and QSK). For ChIP-sequencing, 1×10^6 cells were cross-linked with 1% formaldehyde for 10 min at 37°C prior to freezing in liquid nitrogen. Cross-linked pellets were lysed for 10 min on ice and chromatin fragmented using a Branson 250 digital sonifier. Each ChIP was performed as described previously (Bernstein et al., 2005) with 1 μg of antibody, incubated overnight at 4°C. The following antibodies were used for ChIP: H3K27ac (Active Motif #39133), H3K4me1 (Cell Signaling Technologies #5326BF), and H3K4me3 (Cell Signaling Technologies #9751BF). A 50/50 slurry of protein A and protein G Dynabeads was used to capture enriched chromatin, which was then washed before reverse-crosslinking and proteinase K digestion at 65°C. AMPure XP beads were used to clean up and isolate ChIP DNA for subsequent library construction. Illumina sequencing library construction was performed as previously described (Mikkelsen et al., 2007) and sequenced on an Illumina NextSeq500, 150 bp paired end.

Generation of cell lines

Two independent *BSG*-knockout HEK293T cell lines were generated through CRISPR/Cas9-mediated genome editing by transduction with lentiviruses carrying SpCas9 and guides targeting human *BSG*. HuH7 cell lines stably expressing $SLC16A11^{REF}$ were generated by lentiviral transduction.

Plasmids

Plasmids encoding C terminus, V5-tagged human wild-type (reference) and T2D risk SLC16A11 ($SLC16A11^{REF}$ and $SLC16A11^{T2D}$, respectively) were generated through synthesis of the open reading frames and subcloning into the pLX304 lentiviral vector

(Genscript). K7A and P2D variants of SLC16A11 were introduced through site-directed mutagenesis (Agilent). SLC16A13 and BSG pLX304 expression plasmids were obtained from the Genetic Perturbation Platform at the Broad Institute. SLC16A1 pLX304 has been described previously (Williams et al., 2014). Empty vector control pLX304 was generated by Gateway recombination-mediated replacement of the *ccdB* gene with a multiple cloning site. Plasmids encoding SLC16A11 tagged with a C-terminus HA epitope in pLX304 were generated using standard cloning techniques by inserting a HA tag and a stop codon before the V5 tag encoded by pLX304. Pyronic was a gift from Luis Felipe Barros (Addgene plasmid #51308) (San Martín et al., 2014). For CRISPR/Cas9-mediated genome editing, two guide sequences targeting human *BSG* (5′-TGGATGTTGGCCGTGCCCAT-3′ and 5′-CACCTGTCACTGAC TGGGCC-3′) were cloned into LentiCRISPRv2 (a gift from Feng Zhang), as described (Sanjana et al., 2014).

Transient Transfections

Plasmid transfections were performed using Lipofectamine 2000 (for HEK293T cells) and Lipofectamine 3000 (for HuH7 cells) according to manufacturer's guidelines (Thermo Fisher Scientific).

Lentiviral Transduction

The Genetic Perturbation Platform at the Broad Institute generated lentivirus carrying SLC16A11 variants from pLX304 plasmids. To generate lentivirus for CRISPR/Cas9-mediated knockout, HEK293T cells were plated at a density of 3×10^6 cells per 10 cm plate. The next day, cells were transfected with 12 μg LentiCRISPRv2 plasmid carrying guides targeting human *BSG*, 9 μg PAX2, and 3 μg VSVG using TransIT (Mirus Bio). After 24 hr, media was replaced with media supplemented with 30% FBS in phenol-free DMEM. Twenty-four hours later, lentivirus was harvested and cellular debris removed by filtration through a 0.45 μm filter (Millipore). Lentivirus was concentrated with a 100 kDa spin column (Millipore).

For viral transduction, HEK293T, HuH7, or U2OS MEM-EA cells were spin-infected for 1 hr at 800 g at 31°C with lentivirus and 8 μg/mL polybrene. Transduced cells were selected with either 3 μg/mL puromycin or 5 μg/mL of blasticidin (Thermo Fisher Scientific) for at least 7 days to establish stable cell lines.

Sequence Alignment and Homology Modeling

Sequences of SLC16 family members (obtained from Uniprot) were aligned using the multiple sequence alignment program Clustal Omega (Sievers et al., 2011). We generated a three-dimensional structural model of SLC16A11 by homology modeling (von Grotthuss et al., 2003). A template for the modeling was obtained using the GeneSilico Metaserver (Kurowski and Bujnicki, 2003). The crystal structure of bacterial Glycerol-3-phosphate transporter, GlpT (1PW4) was ranked first by several different fold recognition methods, and was therefore chosen as a high confidence homolog. The SLC16A11 model was created using MODELER (Sali and Blundell, 1993) based on the alignment provided by the profile-profile FFAS method (Rychlewski et al., 2000).

Live Cell Imaging

Cells were transferred to glass coverslips in 60 mm tissue culture dishes 72 hr prior to the experiment. For pyruvate transport assays, HEK293T cells were co-transfected 24 hr after plating with 2 μg pyronic and 2 μg empty pLX304, $SLC16A11^{REF}$-pLX304 or $SLC16A11^{T2D}$-pLX304. Ten minutes prior to imaging, cells were equilibrated in physiological Ringer's solution (140 mM NaCl, 2 mM KCl, 1.5 mM Na2HPO4, 1 mM MgSO4, 2 mM CaCl2, and 10 mM D-glucose in 10 mM HEPES buffer; pH 7.4), following which cells were mounted onto a Ludin chamber (Life Imaging Services) for imaging. Pyruvate uptake was initiated by switching the buffer in the chamber to Ringer's solution containing 0.4 mM pyruvate at the indicated time point. The slopes immediately following addition and withdrawal of pyruvate indicate rates of pyruvate influx and efflux. Cytoplasmic pH changes, indicative of H^+ transport, were determined in cells loaded with 1 μM BCECF-AM (Rink et al., 1982), a pH sensitive dye. Proton uptake was initiated by adding 0.4 mM pyruvate. Rates of proton influx and efflux are calculated as the slopes immediately following addition and withdrawal of pyruvate, after correcting for baseline drift. Baseline drift due to fluorescence bleaching is calculated during the initial 50 s of each trace prior to addition of pyruvate. Linear regression of initial rates was calculated in Kaleidagraph (Synergy Software) and normalized to empty vector controls.

The imaging system consisted of a Zeiss Cell Observer microscope, an X-Cite 120LED illumination system (Lumen Dynamics) and a Hamamatsu Orca-Flash4.0 digital CMOS camera (Hamamatsu). Pyronic was excited using a 436/20 nm band-pass filter; emission was collected through a 540/40 band-pass filter nm with a 510 nm dichroic mirror for venus, and a 480/40 band-pass filter with a 455 nm dichroic mirror for mTFP. BCECF-AM was excited using a 436/20 nm band-pass filter for λ1, and a 495/10 nm band-pass filter for λ2; emission was collected through a 540/40 band-pass filter nm with a 510 nm dichroic mirror. Fluorescent images measurements were acquired every 5 s with the ZEN software (Zeiss).

Analysis of SLC16 Protein Levels

For evaluation of SLC16 family member and SLC16A11 variant levels by western blot analysis, cell lysates were collected 48 hr after transfection. For protein stability experiments, HEK293T cells transfected with SLC16 family members were treated with 10 μg/mL cycloheximide (Cayman Chemical, 14126) for 10 min, 30 min, 1 hr, 2 hr, and 4 hr prior to collection of protein lysates. For evaluation of proteasomal degradation, HuH7 cells stably transduced with $SLC16A11^{REF}$-V5 were treated with 5 μM MG132 (Cayman Chemical, 10012628) for 3 hr prior to collection of protein lysates. Protein lysates were collected in lysis buffer consisting of 1% NP40, 0.1%

SDC, 150 mM NaCl, 50 mM TrisHCl pH 7.5, and 1 mM EDTA with protease inhibitors (Roche). Lysates were rotated for 30 min at 4°C. Supernatant was collected following centrifugation for 10 min at 14,000 g at 4°C.

Membrane Fractionation

Plasma membranes were separated from intracellular membranes using a Plasma Membrane Protein Extraction Kit (Abcam ab65400) with minor modifications to the manufacturer's protocol. Briefly, a 15 cm plate of HEK293T cells transfected with $SLC16A11^{REF}$ or $SLC16A11^{T2D}$ was harvested in 1 mL homogenization buffer prior to Dounce homogenization. The plasma membrane fraction was resuspended in 33 μL lysis buffer. The cytoplasmic and intracellular membrane fractions were diluted by adding 75 μL lysis buffer to 25 μL of each fraction prior to quantification. Equal quantities of protein from each fraction were used in western blot analyses.

PathHunter MEM-EA Pharmacotrafficking Assay

$SLC16A11^{REF}$ or $SLC16A11^{T2D}$ with a C-terminal V5 tag were cloned into pCMV-ProLink2 (DiscoverX) for the PathHunter MEM-EA Pharmacotrafficking assays (DiscoverX) using standard cloning techniques. U2OS MEM-EA cells (DiscoverX) were plated into 96-well plates at a density of 8,000 per well. After 24 hr, cells were transfected with 200 ng of the appropriate SLC16A11 construct using Lipofectamine 2000. Media was changed 24 hr after transfection. Luminescence, indicative of enzyme complementation and activity upon SLC16A11 localization to the cell surface, was measured 24 hr later using the PathHunter Detection kit (DiscoverX) and an EnVision plate reader according to the manufacturer's protocol. Data was analyzed in Excel. Background from empty vector controls was subtracted from SLC16A11 signal. To combine data across different experiments, data within an experiment was normalized to the average value of the variants in that experiment. Statistical significance was assessed by a t test across the combined data.

Immunoprecipitations for Protein Interactions

HEK293T cells plated onto 15 cm plates were transfected with 40 μg plasmid encoding $SLC16A11^{REF}$ or $SLC16A11^{T2D}$ and/or BSG or empty vector control. After 48 hr, protein lysates were harvested in lysis buffer consisting of 1% NP40, 0.1% SDC, 150 mM NaCl, 50 mM TrisHCl pH 7.5, and 1 mM EDTA with protease inhibitors (Roche). Lysates were passed through a 20 G syringe and rotated for 30 min at 4°C. Supernatant was collected following a 10 min spin at 14,000 g at 4°C, and total protein concentration was measured using a BCA assay (Thermo Fisher Scientific).

For proteomics analysis, 20 mg total protein lysate was incubated with 100 μL of anti-V5 conjugated agarose beads (SIGMA, A7345) for 3 hr at 4°C. Beads were then washed once with lysis buffer and three times with wash buffer (150 mM NaCl, 50 mM TrisHCl pH 7.5, and 1 mM EDTA with protease inhibitors). After the last wash, residual wash buffer was removed, 10 μL fresh wash buffer was added, and beads were stored at −80°C until analysis by the Proteomics Platform at the Broad Institute, as detailed in the "Proteomic Methods for SLC16A11 Interactors" below.

For co-immunoprecipitation analysis, 5-10 mg total protein lysates were incubated with 50 μL anti-V5 or anti-HA conjugated agarose beads (SIGMA, A7345 and A2095, respectively). Beads were washed three times with lysis buffer and once with wash buffer. Immunoprecipitated proteins were eluted off beads with 50 μL LDS sample buffer (Thermo Fisher Scientific) diluted in wash buffer and either NuPage reducing reagent (Thermo Fisher Scientific) or 5% beta-mercaptoethanol. Samples were incubated at 42°C for 15 min and then 70°C for 1 min.

Western Blot Analyses

Total protein concentration was measured using a BCA assay (Thermo Fisher Scientific) and equal quantities of protein were loaded on each gel. Lysates were prepared for western blotting by adding LDS sample buffer (Thermo Fisher Scientific) and either NuPage reducing reagent (Thermo Fisher Scientific) or 5% beta-mercaptoethanol. Samples were denatured by incubation at 42°C for 10 min prior to analysis by SDS-PAGE and immunoblotting. Nitrocellulose membranes were blocked for 15 min with 5% milk in TBST and then incubated with primary antibody in blocking solution overnight at 4°C. Primary antibodies are detailed below. Higher antibody dilutions were used to detect immunoprecipitated proteins. Signals were detected with HRP-conjugated secondary antibodies, followed by chemiluminescent detection and autoradiography.

Primary Antibodies used in Western Blot Analysis

Protein	Dilution	Product Number	Company
V5 epitope	1:1,000-1:15,000	#13202 D3H8Q	Cell Signaling
HA epitope	1:1,000-1:20,000	#3724 C29F4	Cell Signaling
BSG	1:750-1:2,500	#12314	Cell Signaling
Na/KATPase	1:20,000	ab7671	Abcam
Calenxin	1:10,000	#2433	Cell Signaling
Tubulin	1:10,000	ab21058	Abcam
Vinculin	1:10,000	ab129002	Abcam

For membrane fractionation experiments, autoradiographs were scanned and densitometry quantified with ImageJ. Data was exported to Excel and SLC16A11-V5 levels in plasma membrane fractions were normalized to SLC16A11-V5 levels in the intracellular membrane fraction. Data from independent experiments was combined by dividing each normalized value in a given experiment by the average of the normalized values obtained for $SLC16A11^{REF}$ in that experiment. The relative amount of SLC16A11 at the plasma membrane was then calculated by averaging across these values and statistical significance was assessed with a t test.

Proteomic Methods for SLC16A11 Interactors

On-bead digest

The beads from immunopurification were washed once with lysis buffer, then three times with PBS, the three different lysates of each replicate were resuspended in 90 μL digestion buffer (2 M Urea, 50 mM Tris HCl), 2 μg of sequencing grade trypsin added, 1 hr shaking at 700 rpm. The supernatant was removed and placed in a fresh tube. The beads were then washed twice with 50 μL digestion buffer and combined with the supernatant. The combined supernatants were reduced (2 μL 500 mM DTT, 30 min, RT), alkylated (4 μL 500 mM IAA, 45 min, dark) and a longer overnight digestion was performed: 2 μg (4 μL) trypsin, shake o/n, The samples were then quenched with 20 μL 10% FA and desalted on 10 mg SepPak columns.

iTRAQ labeling of peptides and strong cation exchange (scx) fractionation

Desalted peptides were labeled with iTRAQ reagents according to the manufacturer's instructions (AB Sciex, Foster City, CA). Peptides were dissolved in 30 μl of 0.5 M TEAB pH 8.5 solution and labeling reagent was added in 70 μL of ethanol. After 1 hr incubation the reaction was stopped with 50 mM Tris/HCl pH 7.5. Differentially labeled peptides were mixed and subsequently desalted on 10 mg SepPak columns.

	iTRAQ labeling			
	114	115	116	117
Rep1	Reference (WT)	T2Drisk (QNT)	Empty vector Control 1	Empty vector Control 2
Rep2	Reference(WT)	T2Drisk (QNT)	Empty vector Control 1	Empty vector Control 2

SCX fractionation of the differentially labeled and combined peptides was done as described (Rappsilber et al., 2007), with 6 pH steps (buffers- all contain 25% acetonitrile) as below:

1: ammonium acetate 50 mM pH 4.5,
2: ammonium acetate 50 mM pH 5.5,
3: ammonium acetate 50 mM pH 6.5,
4: ammonium bicarbonate 50 mM pH 8,
5: ammonium hydroxide 0.1% pH 9,
6: ammonium hydroxide 0.1% pH 11.

Empore SCX disk used to make StageTips are as described in the paper.

MS analysis

Reconstituted peptides were separated on an online nanoflow EASY-nLC 1000 UHPLC system (Thermo Fisher Scientific) and analyzed on a benchtop Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). The peptide samples were injected onto a capillary column (Picofrit with 10 μm tip opening / 75 μm diameter, New Objective, PF360-75-10-N-5) packed in-house with 20 cm C18 silica material (1.9 μm ReproSil-Pur C18-AQ medium, Dr. Maisch GmbH, r119.aq). The UHPLC setup was connected with a custom-fit microadapting tee (360 μm, IDEX Health & Science, UH-753), and capillary columns were heated to 50°C in column heater sleeves (Phoenix-ST) to reduce backpressure during UHPLC separation. Injected peptides were separated at a flow rate of 200 nL/min with a linear 80 min gradient from 100% solvent A (3% acetonitrile, 0.1% formic acid) to 30% solvent B (90% acetonitrile, 0.1% formic acid), followed by a linear 6 min gradient from 30% solvent B to 90% solvent B. Each sample was run for 120 min, including sample loading and column equilibration times. The Q Exactive instrument was operated in the data-dependent mode acquiring HCD MS/MS scans (R = 17,500) after each MS1 scan (R = 70,000) on the 12 top most abundant ions using an MS1 ion target of $3x\ 10^6$ ions and an MS2 target of $5x10^4$ ions. The maximum ion time utilized for the MS/MS scans was 120 ms; the HCD-normalized collision energy was set to 27; the dynamic exclusion time was set to 20 s, and the peptide match and isotope exclusion functions were enabled.

Quantification and identification of peptides and proteins

All mass spectra were processed using the Spectrum Mill software package v6.0 pre-release (Agilent Technologies) which includes modules developed by us for iTRAQ -based quantification. Precursor ion quantification was done using extracted ion chromatograms (XIC's) for each precursor ion. The peak area for the XIC of each precursor ion subjected to MS/MS was calculated automatically by the Spectrum Mill software in the intervening high-resolution MS1 scans of the LC-MS/MS runs using narrow windows around each individual member of the isotope cluster. Peak widths in both the time and *m/z* domains were dynamically determined

based on MS scan resolution, precursor charge and *m/z,* subject to quality metrics on the relative distribution of the peaks in the isotope cluster versus theoretical. Similar MS/MS spectra acquired on the same precursor m/z in the same dissociation mode within ± 60 s were merged. MS/MS spectra with precursor charge > 7 and poor quality MS/MS spectra, which failed the quality filter by not having a sequence tag length > 1 (i.e., minimum of 3 masses separated by the in-chain mass of an amino acid) were excluded from searching.

For peptide identification MS/MS spectra were searched against human Uniprot database to which a set of common laboratory contaminant proteins was appended as well as the sequence for V5-tagged SLC16A11^{REF} (also called WT, wild-type, through-out the proteomics datasets) and SLC16A11^{T2D} (also called QNT through-out the proteomics datasets). Search parameters included: ESI-QEXACTIVE-HCD scoring parameters, trypsin enzyme specificity with a maximum of two missed cleavages, 40% minimum matched peak intensity, +/− 20 ppm precursor mass tolerance, +/− 20 ppm product mass tolerance, and carbamidomethylation of cysteines and iTRAQ labeling of lysines and peptide n-termini as fixed modifications. Allowed variable modifications were oxidation of methionine, N-terminal acetylation, Pyroglutamic acid (N-termQ), Deamidated (N), Pyro Carbamidomethyl Cys (N-termC), with a precursor MH+ shift range of −18 to 64 Da. Identities interpreted for individual spectra were automatically designated as valid by optimizing score and delta rank1-rank2 score thresholds separately for each precursor charge state in each LC-MS/MS while allowing a maximum target-decoy-based false-discovery rate (FDR) of 1.0% at the spectrum level.

In calculating scores at the protein level and reporting the identified proteins, redundancy is addressed in the following manner: the protein score is the sum of the scores of distinct peptides. A distinct peptide is the single highest scoring instance of a peptide detected through an MS/MS spectrum. MS/MS spectra for a particular peptide may have been recorded multiple times, (i.e., as different precursor charge states, isolated from adjacent SCX fractions, modified by oxidation of Met) but are still counted as a single distinct peptide. When a peptide sequence > 8 residues long is contained in multiple protein entries in the sequence database, the proteins are grouped together and the highest scoring one and its accession number are reported. In some cases when the protein sequences are grouped in this manner there are distinct peptides which uniquely represent a lower scoring member of the group (isoforms or family members). Each of these instances spawns a subgroup and multiple subgroups are reported and counted toward the total number of proteins. iTRAQ ratios were obtained from the protein-comparisons export table in Spectrum Mill. To obtain iTRAQ protein ratios the median was calculated over all distinct peptides assigned to a protein subgroup in each replicate. To assign interacting proteins we used the Limma package in the R environment to calculate moderated t test *p* and added Blandt-Altman testing to filter out proteins for which the CI for reproducibility was below 95%.

SLC16A11 Knockdown for Metabolite Profiling

The following siRNAs were ordered from GE Dharmacon: Accell Human *SLC16A11* siRNA SMARTpool (E-007404-00-0050) and Accell Non-targeting Pool (D-001910-10-50). siRNAs were reconstituted in PBS at 100 μM. Hepatocytes were plated into collagen coated 24 well plates (Corning, 354408) at a density of 350,000 cells per well. After 4 hr, cells were washed once with HI media (BioreclamationIVT) and siRNAs in HI media were added at a final concentration of 1 μM. After 24 hr, fresh CP media was added. Hepatocytes were grown for an additional 24 hr. The samples consisted of 8 or 12 biological replicates of hepatocytes treated either with pooled siRNAs targeting *SLC16A11* or negative controls across 2 or 3 replicate experiments. Lysates were collected for profiling 48 hr after siRNA knockdown. Due to the media change, differences in extracellular metabolites reflect changes accumulated during the 24 hr period prior to sample collection.

Primary human hepatocytes were removed from the incubator and immediately placed on ice. 500 μL of media was collected from the hepatocytes and put aside. Hepatocytes were then washed with 1 mL PBS. Lipids were extracted from hepatocytes by scraping in 250 μL of isopropanol (HPLC Grade; Honeywell) containing 1,2-didodecanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids; Alabaster, AL). Polar metabolites were extracted by scraping in 250 μL of 80% methanol (VWR) containing 0.05 ng/μL inosine-15N4, 0.05 ng/μL thymine-d4, and 0.1 ng/μL glycocholate-d4 as internal standards (Cambridge Isotope Laboratories). Media was spun at 600 g at 4°C for 5 min to remove any debris. Media metabolites for lipid analyses were precipitated by taking 10 μL and adding 190 μL isopropanol. Media metabolites for HLIC-neg were precipitated with 30 μL media + 120 μL extraction solution (0.05 ng/μL Inosine-15N4, 0.05 ng/μL Thymine-d4, 0.1 ng/μL Glycocholate-d4 in 80% Methanol). Finally, media metabolites for HILIC-pos were precipitated with 10 μL media + 90 μL extraction solution consisting of Acetonitrile:Methanol:Formic acid (75:25:0.2 vol:vol:vol) with a nominal concentration of 0.2 μg/mL valine-d8 (Sigma) and phenylalaine-d8 (Cambridge Isotopes Laboratories).

Metabolite Profiling

Analyses of lipids were conducted using an LC-MS system comprised of a Shimadzu Nexera X2 U-HPLC (Shimadzu Corp.; Marlborough, MA) coupled to an Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific; Waltham, MA). Lipid extracts were injected onto an ACQUITY BEH C8 column (100 × 2.1 mm, 1.7 μm; Waters, Milford, MA). The column was eluted isocratically with 80% mobile phase A (95:5:0.1 vol/vol/vol 10 mM ammonium acetate/methanol/formic acid) for 1 min followed by a linear gradient to 80% mobile-phase B (99.9:0.1 vol/vol methanol/formic acid) over 2 min, a linear gradient to 100% mobile phase B over 7 min, then 3 min at 100% mobile-phase B. MS data were acquired using electrospray ionization in the positive ion mode over 200–1100 *m/z* and at 70,000 resolution. Other MS settings were: sheath gas 50, in source CID 5 eV, sweep gas 5, spray voltage 3 kV, capillary temperature 300°C, S-lens RF 60, heater temperature 300°C, microscans 1, automatic gain control target 1e6, and maximum ion time 100 ms. Raw data were processed using TraceFinder 3.3 (Thermo Fisher Scientific; Waltham, MA) and Progenesis QI (Nonlinear

Dynamics; Newcastle upon Tyne, UK) software for detection and integration of LC-MS peaks. Lipid identities were determined based on comparison to reference standards and reference plasma extracts and are denoted by total number of carbons in the lipid acyl chain(s) and total number of double bonds in the lipid acyl chain(s). HILIC (hydrophilic interaction chromatography) methods in the positive and negative modes were run as described previously (Williams et al., 2014).

QUANTIFICATION AND STATISTICAL ANALYSIS

Data Integration and Credible Set Analysis

For the credible set analysis we first built two datasets. One dataset was comprised of 4,478 samples that had been genotyped by exome chip and OMNI 2.5 (Dataset 1) (Williams et al., 2014). The other dataset comprised another subset of 3,732 samples genotyped by exome chip, OMNI2.5, and whole-exome sequencing (Dataset 2) (Estrada et al., 2014). We kept all the variants with MAF higher than 0.001 for both datasets. We phased both datasets with SHAPEIT2 (Delaneau et al., 2013) (version 2.5) and then imputed the 1000G (phase 3, release June 2014) into both datasets separately. We also imputed whole-exome variants that were not imputable using 1000G phase 3 into samples that had not been ascertained by whole-exome sequencing (Dataset 1). We removed variants with *impute 2 information score* < 0.8 as a post-imputation quality control. We then performed the association analysis separately in each cohort using SNPtest adjusting for BMI, age, sex and the first two principal components to adjust for population stratification. We then meta-analyzed both results using Metal (Willer et al., 2010).

The credible set was constructed as previously described (Wakefield, 2007). Briefly, for each association meta-analysis results, we computed an approximate Bayes factor (abf) for each variant with an r-squared greater than 0.1 with the top variant at the *SLC16A11* locus,

$$abf = \frac{\sqrt{(1-r)}}{\exp\left(-r \times \frac{z^2}{2}\right)}$$

where $z = beta/se$ and $r = (0.04/se^2 + 0.04)$ under the assumption that the prior on beta is Gaussian with variance 0.04. The posterior probability for each variant was then computed by dividing the abf by the total number of variants in the region. All variants were ranked by posterior probability and the minimal set of variants that resulted in a cumulative posterior probability of 0.99 was deemed the 99% credible set.

Association Analysis in MEDIA

The MEDIA dataset consists of meta-analysis results from 17 T2D studies (ARIC, CARDIA, CFS, CHS, FamHS, GeneSTAR, GENOA, HANDLS, Health ABC, HUFS, JHS, MESA, MESA Family, SIGNET-REGARDS, WFSM, FIND, and WHI) with up to 23,827 African American subjects (8,284 cases and 15,543 controls) (Ng et al., 2014). The results of each study were imputed using the HapMap reference panel and meta-analyzed using inverse-variance fixed-effects meta-analysis using METAL.

eQTL and Allele-Specific Expression Analyses

For the eQTL analyses, normalized ddPCR gene expression counts were compared between individuals who carried 0, 1, or 2 copies of the T2D risk haplotype. Within each genotype, the mean count for each gene was calculated. Percent change and standard error mean were calculated for each pairwise comparison between genotypes. Statistical significance was assessed by linear regression assuming an additive model for the SNP, and adjusting by age, sex, BMI, and T2D status. We accounted for multiple hypothesis testing using a Bonferroni correction for the number of genes tested.

Association between genotype and quantitative traits was assessed by linear regression assuming an additive model for the SNP and adjusting by age, sex, BMI, and T2D status. Statistical significance of quantitative trait data of tissue data was assessed by linear regression assuming an additive model for the SNP, and adjusting by age, sex, BMI, and T2D status. We accounted for multiple hypothesis testing using a Bonferroni correction for the number of traits tested.

For each heterozygote individual in the allele-specific expression analyses, we computed the proportion of total *SLC16A11* gene expression counts that originate from the T2D risk allele versus the reference allele. We calculated the overall proportion and 95% confidence intervals from all the samples by meta-analyzing with inverse variance method after logit transformation using the 'meta' package (version 4.5-0). We then computed the statistical significance of the differences between the pooled proportion and the expected proportion under the null (0.5) by computing the Z-statistic from which the two-tailed p value was derived.

ChIP-Seq Analysis

Reads for immunoprecipitated H3K27ac, H3K4me1, H3K4me3, and input DNA were mapped to GRCh37 using Bowtie2 (Langmead and Salzberg, 2012). Peaks were called using HOMER (Heinz et al., 2010). Alignments were processed using WASP to adjust for reference-mapping bias (van de Geijn et al., 2015). We did not detect evidence of the risk SNP at rs4630597, rs78972129, and rs76070643 in lot QSK and, consequently, removed these variants in this donor from further analyses. For each variant in each sample, we performed a binomial test in order to detect a skew in reference versus alternate (T2D risk) allele read counts. We combined p values across samples using Fisher's method and accounted for multiple hypothesis testing using a Bonferroni correction for the

number of variants and histone modifications tested. We considered a histone modification at a variant significantly skewed if it met our Bonferroni-corrected significance threshold and if the direction of the allelic imbalance was consistent across all donors.

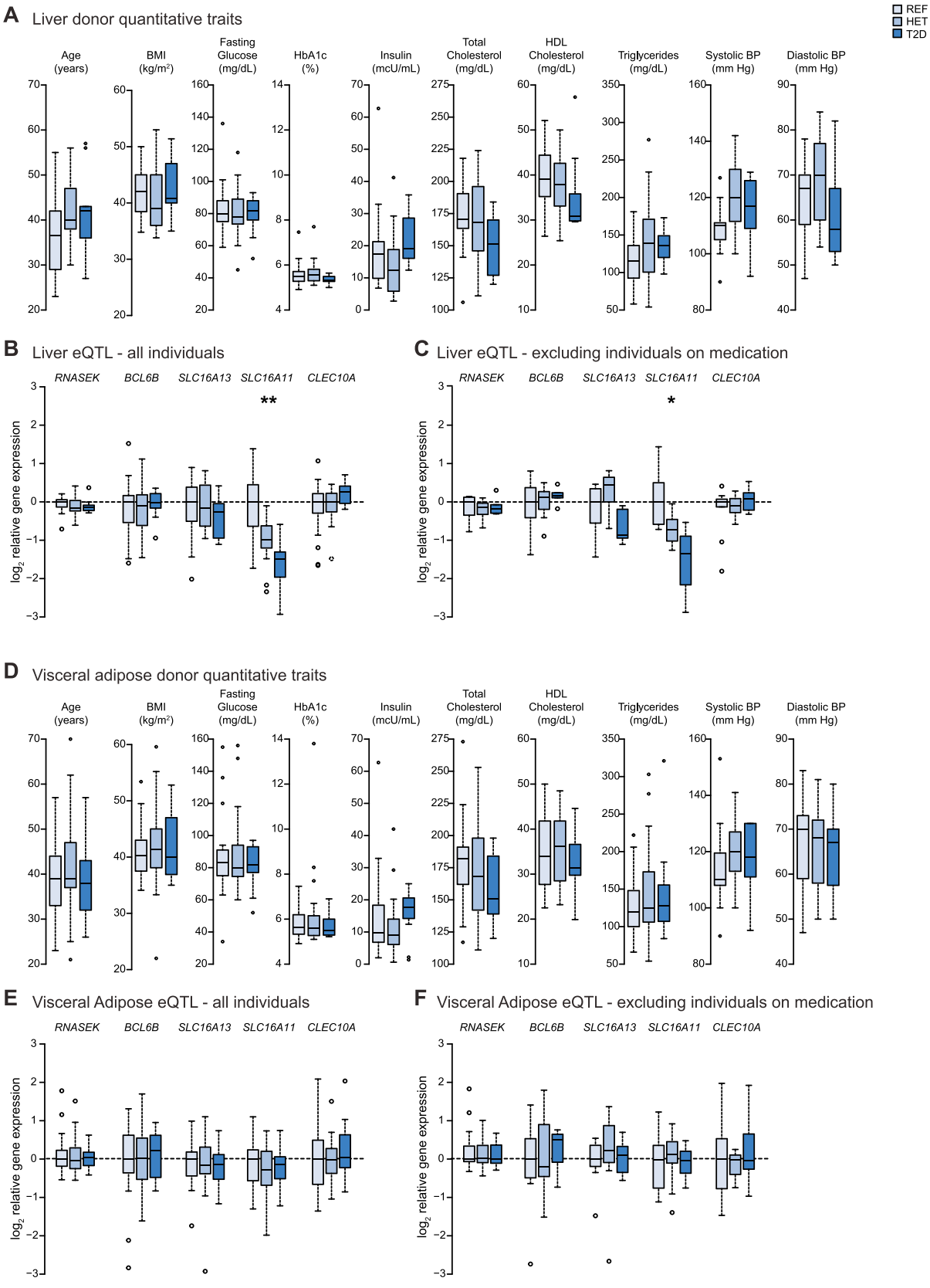
Metabolomic Analyses

Within each experiment, metabolite-profiling data was first total signal normalized. Individual metabolite values were flagged as outliers and removed if they were more than 2 standard deviations from the mean value within a sample type (where sample types are *SLC16A11* siRNA or negative control siRNA treatment). For each experiment, metabolite values were normalized by the mean value for that metabolite within the negative control siRNA treatments to enable aggregation of data across all 3 experiments. The additional data in the combined dataset allowed for improved ability to detect outliers. Outlier values were identified and removed if they were more than 2 standard deviations from the mean value within a sample type. Fold-changes for each metabolite were calculated by dividing the median value of the *SLC16A11* siRNA treatments by the median value of the negative control siRNA treatments. Significance was computed using a Wilcoxon test between the two sample types.

To identify metabolite changes at the pathway level, we applied a strategy that is commonly used for analysis of gene expression data: gene-set enrichment analysis (GSEA) (Subramanian et al., 2005). Pathway enrichment was computed using the GSEA PreRanked tool, as implemented at http://software.broadinstitute.org/gsea/index.jsp, using an unweighted enrichment score and 1,000 permutations. The log2 transformed fold-changes between SLC16A11 knockdown and control were used as input, along with curated sets of KEGG pathways from the human reference set and 15 additional classes of metabolites covering lipid sub-types and carnitines. Only metabolite pathways and classes with at least 5 members measured in our dataset were considered. Normalized enrichment scores quantify the concordance of individual metabolite fold-changes within a given metabolic pathway or class, with a positive score indicating enrichment and a negative score corresponding to depletion.

DATA AND SOFTWARE AVAILABILITY

The original mass spectra from the SLC16A11 protein interaction screen may be downloaded from MassIVE (http://www.type2diabetesgenetics.org/) using the identifier: MSV000081105. The data is directly accessible via ftp://massive.ucsd.edu/MSV000081105. The accession number for the ChIP-sequencing data reported in this paper is NCBI GEO: GSE99301. SIGMA genetic data is available at http://www.type2diabetesgenetics.org.



(legend on next page)

Figure S1. A *SLC16A11 Cis*-eQTL Is Present in Liver but Not Visceral Adipose Tissue, Related to Figure 2

(A) Boxplots depict quantitative trait analyses in all donors for liver expression QTL (eQTL) analyses. n = 21 homozygous reference (REF), 16 heterozygous (HET), and 10 homozygous T2D risk (RISK).

(B) eQTL analyses in all liver donors using a second probe for each gene. Boxplots depict the log2 of the relative expression level for *RNASEK, BCL6B*, *SLC16A11*, *SLC16A13*, and *CLEC10A* according to genotype at rs13342692.

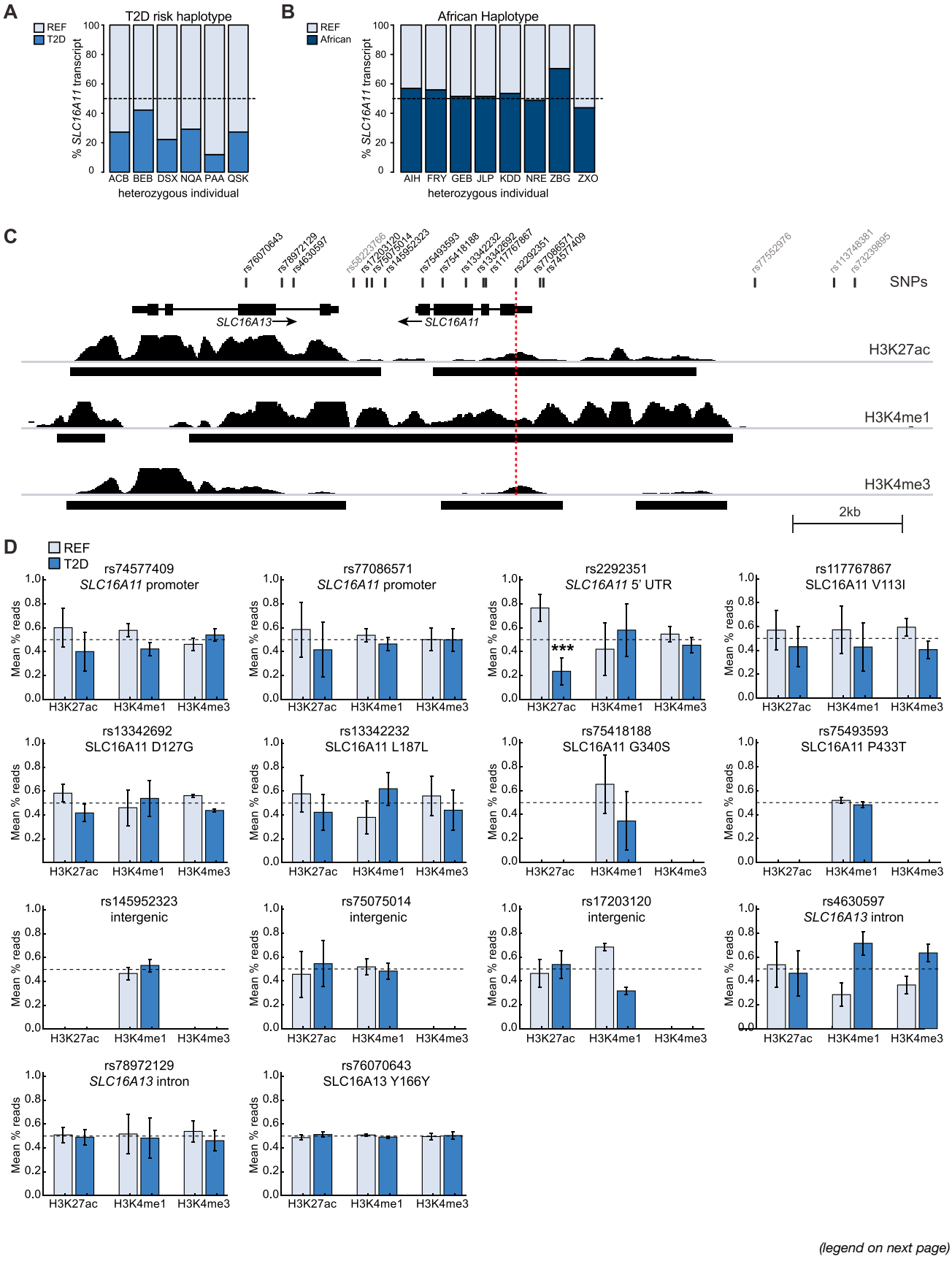
(C) eQTL analyses in liver, excluding individuals taking medication prior to the 24 hr before tissue collection. n = 12 REF, 10 HET, and 5 RISK.

(D) Boxplots depict quantitative trait analyses in all donors for visceral adipose eQTL analysis. n = 33 REF, 27 HET, and 18 RISK.

(E) eQTL analyses in all visceral adipose donors. Boxplots depict the log2 of the relative expression level for *RNASEK, BCL6B*, *SLC16A11*, *SLC16A13*, and *CLEC10A* according to genotype at rs13342692.

(F) eQTL analyses in adipose, excluding individuals taking medication prior to the 24 hr before tissue collection. n = 15 homozygous REF, 19 HET, and 13 RISK.

*$p < 0.05$, **$p < 1 \times 10^{-3}$.



(legend on next page)

Figure S2. Allelic Imbalance of *SLC16A11* Expression and Histone Modifications in Carriers of the T2D-Risk Haplotype, Related to Figure 2

(A and B) Allele-specific expression analyses in primary human hepatocytes heterozygous for the (A) T2D risk (n = 6) and (B) African (n = 8) haplotypes. The stacked bar plots depict the relative proportion of *SLC16A11* transcript originating from each allele in each sample.

(C and D) ChIP-sequencing for H3K27ac, H3K4me1, and H3K4me3 histone modifications in primary human hepatocytes from three individuals heterozygous for the T2D risk haplotype (lots ACB, DSX, and QSK). (C) ChIP-sequencing peaks overlapping variants in the T2D risk credible set are shown. (D) Bar plots depict allelic proportions ± SEM at indicated credible set SNPs. Asterisk indicates significance after Bonferroni correction for multiple hypothesis testing: ***$p < 1 \times 10^{-5}$.

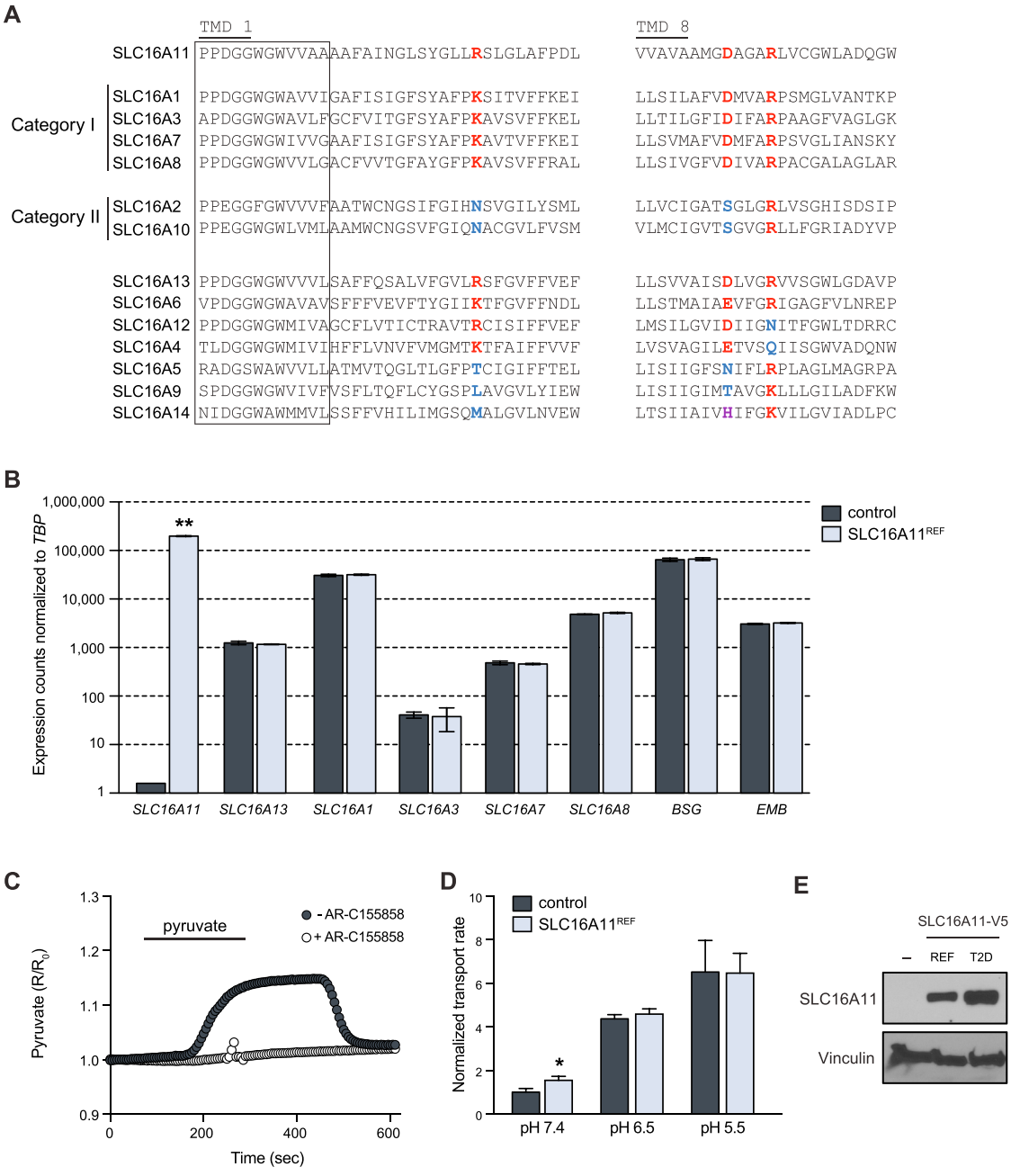


Figure S3. Pyruvate Transport Measured by Pyronic Is SLC16-Dependent, Related to Figures 3 and 4

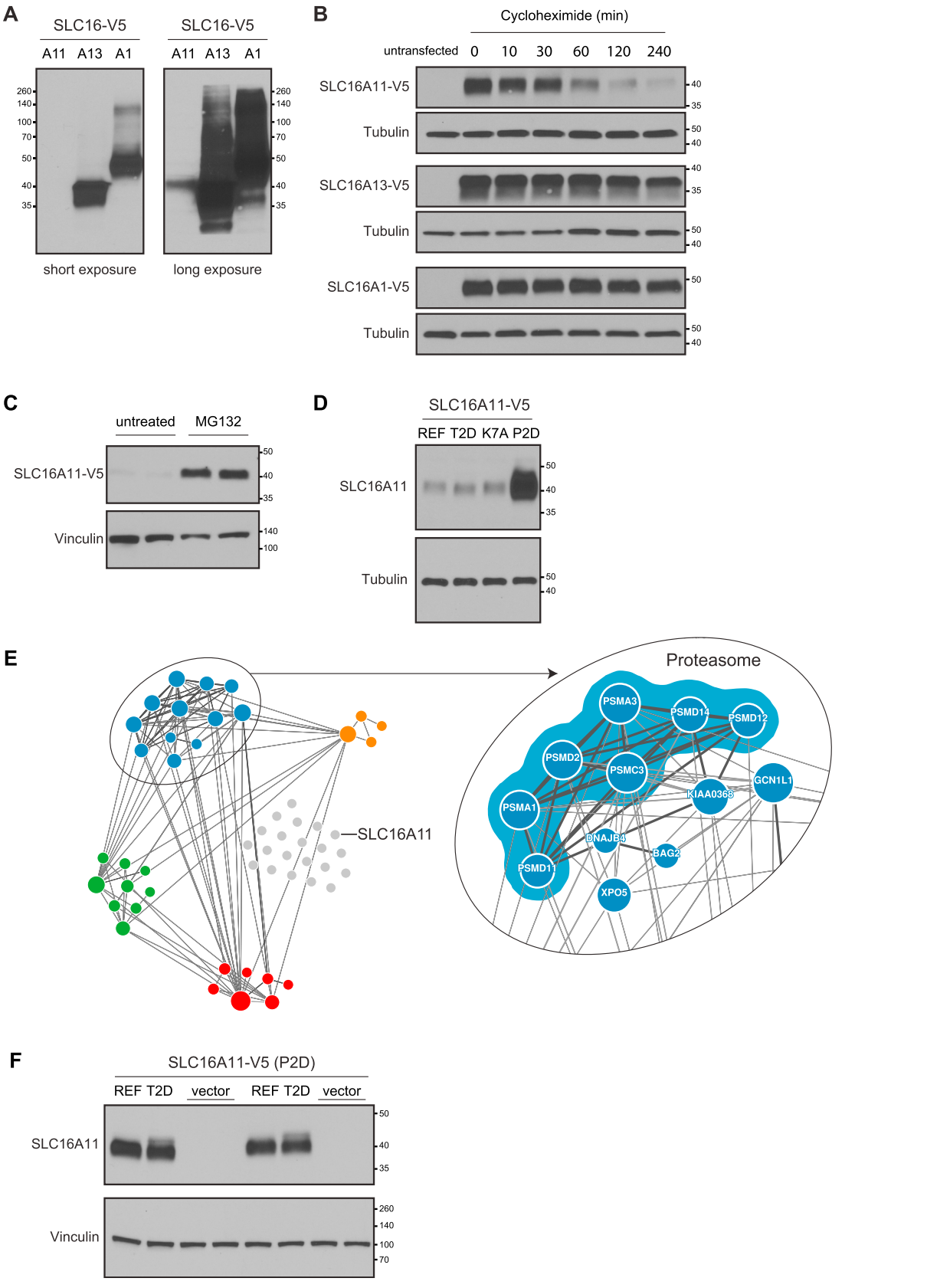
(A) Sequence alignment of transmembrane domains (TMDs) 1 and 8 from all members of the SLC16 family. Charged residues at the putative transport motif are indicated in red. Non-charged residues are indicated in blue. Opposite charge is indicated in purple.

(B) Gene expression counts of the indicated category I SLC16 family members, *BSG*, and *EMB*, normalized to the housekeeper gene *TBP* in HEK293T cells transfected with empty vector control or SLC16A11^{REF}. **$p < 1 \times 10^{-3}$

(C) Traces of pyruvate flux, in HEK293T cells overexpressing the pyruvate-specific FRET sensor, pyronic (San Martín et al., 2014). Cells were pre-treated with the SLC16/MCT inhibitor AR-C155858 (Ovens et al., 2010). Pyruvate (0.4 mM) was added and removed, as indicated.

(D) Bar plots depicting normalized rates of pyruvate influx ± SEM at different pH levels in control and SLC16A11^{REF}-expressing cells, as measured with the pyruvate-specific FRET sensor, pyronic. Pyruvate (0.4 mM) was added at the indicated pH. *$p < 0.05$

(E) Protein levels of V5-tagged SLC16A11^{REF} and SLC16A11^{T2D} in HEK293T cells.



(legend on next page)

Figure S4. SLC16A11 Protein Levels Are Tightly Regulated, Related to Figure 5

(A) Protein levels of C-terminal, V5-tagged SLC16A11^{REF}, SLC16A13, and SLC16A1 expressed in HuH7 cells. Note low levels of SLC16A11 relative to other SLC16 family members.

(B) SLC16A11 protein is rapidly degraded. HEK293T cells expressing either V5-tagged SLC16A11^{REF}, SLC16A13, or SLC16A1 were treated with cycloheximide (10 μg/mL) for the indicated times.

(C) SLC16A11 protein levels are increased by proteasome inhibition. HuH7 cells stably expressing SLC16A11^{REF}-V5 were treated with MG132 (5 μM) for 3 hr.

(D) Protein levels of REF, T2D, K7A, and P2D SLC16A11-V5 in HEK293T cells. The P2D variant increases SLC16A11 protein levels.

(E) GeNETs network analysis (http://apps.broadinstitute.org/genets) highlights previously reported protein-protein interactions between proteins in the SLC16A11-interactome. Pathway analysis (using the InWeb3 database in GeNETs) was performed on proteins highly enriched for interaction with SLC16A11 (top 10% with a Blandt-Altman-adjusted $p < 0.05$). Components of the proteasome are enriched among proteins interacting with SLC16A11.

(F) Protein levels of P2D SLC16A11^{REF} and P2D SLC16A11^{T2D} in input samples for the protein-protein interaction screen. Approximate locations of molecular weight markers (kDa) are indicated.

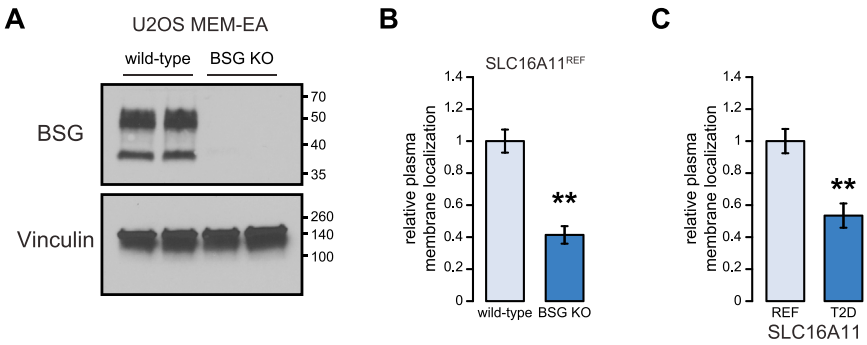


Figure S5. SLC16A11 Cell-Surface Localization Is BSG-Dependent and Reduced by the T2D-Risk Coding Variants, Related to Figure 5

(A) Western blot analysis of BSG levels in wild-type and BSG-knockout U2OS MEM-EA cells. Two different BSG knockout lines were generated using two different sgRNAs and CRISPR/Cas9. Approximate locations of molecular weight markers (kDa) are indicated.

(B) Split β-galactosidase reporter assay in WT and BSG-knockout U2OS MEM-EA cells demonstrates $SLC16A11^{REF}$ plasma membrane localization is reduced in BSG-knockout cells. Bar plots depict the relative amount of $SLC16A11^{REF}$ at plasma membrane ± SD (n = 4 for WT; n = 8 for BSG knockout (4 each per BSG-knockout U2OS MEM-EA cell line); **$p = 4 \times 10^{-5}$).

(C) Split β-galactosidase reporter assay in U2OS MEM-EA cells demonstrates a reduction in $SLC16A11^{T2D}$ at the cell surface relative to $SLC16A11^{REF}$. Bar plots depict the relative amount of SLC16A11 at the cell surface ± SD (n = 5; **$p = 1 \times 10^{-5}$).

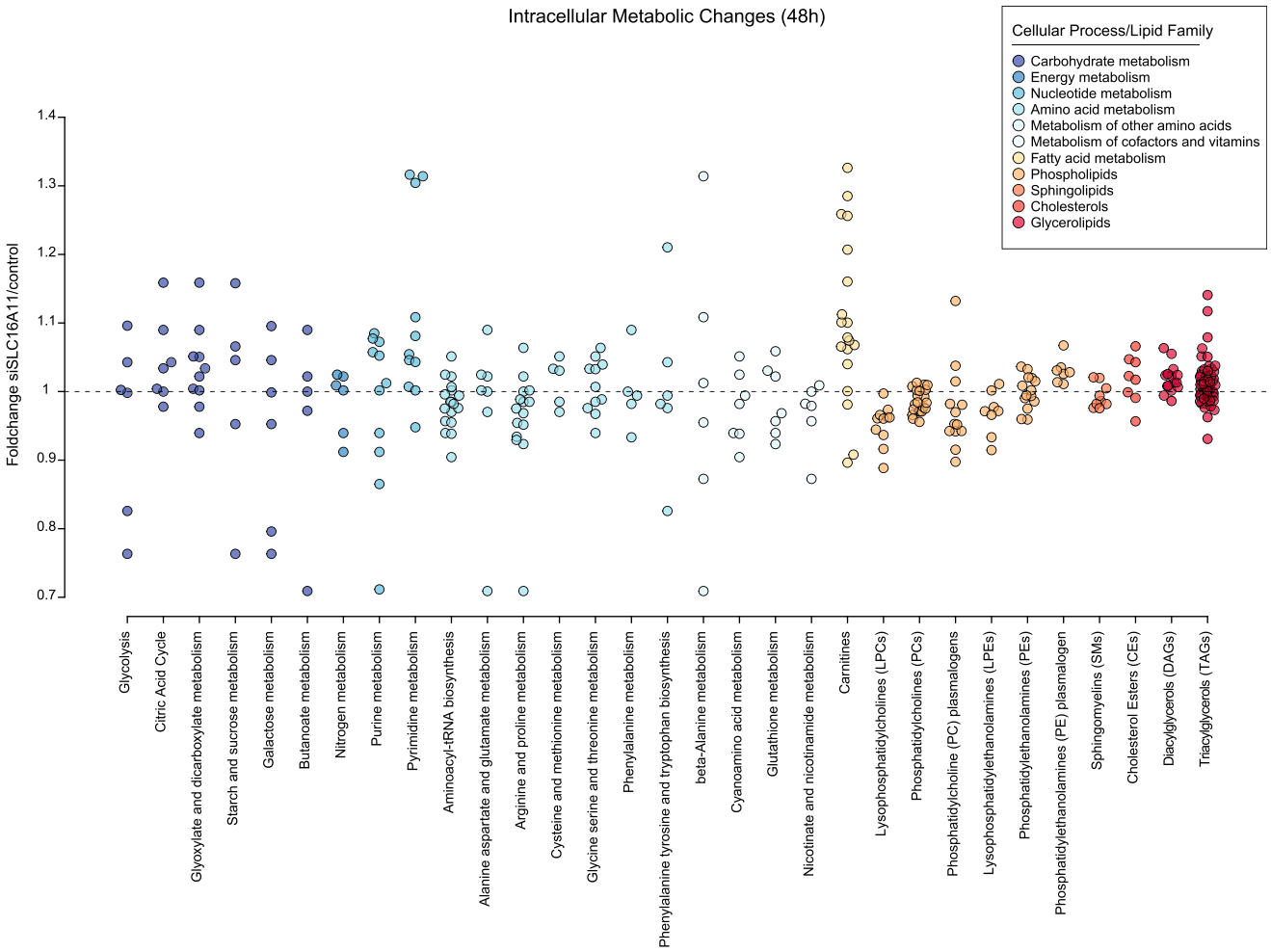


Figure S6. Intracellular Metabolic Pathways Altered by Knockdown of *SLC16A11* in Primary Human Hepatocytes, Related to Figure 6

Changes in intracellular levels of individual metabolites in primary human hepatocytes treated with *SLC16A11* siRNAs (versus control) for 48 hr are plotted in groups according to metabolic pathway or class. Pathways shown include all KEGG pathways from the human reference set for which at least five metabolites were measured as well as additional classes of metabolites covering carnitines and lipid sub-types. Each point within a pathway or class shows the fold-change of a single metabolite within that pathway or class. Pathways are colored according to cellular processes or lipid family as indicated in the legend. See also Table S4 for details on individual metabolite changes.

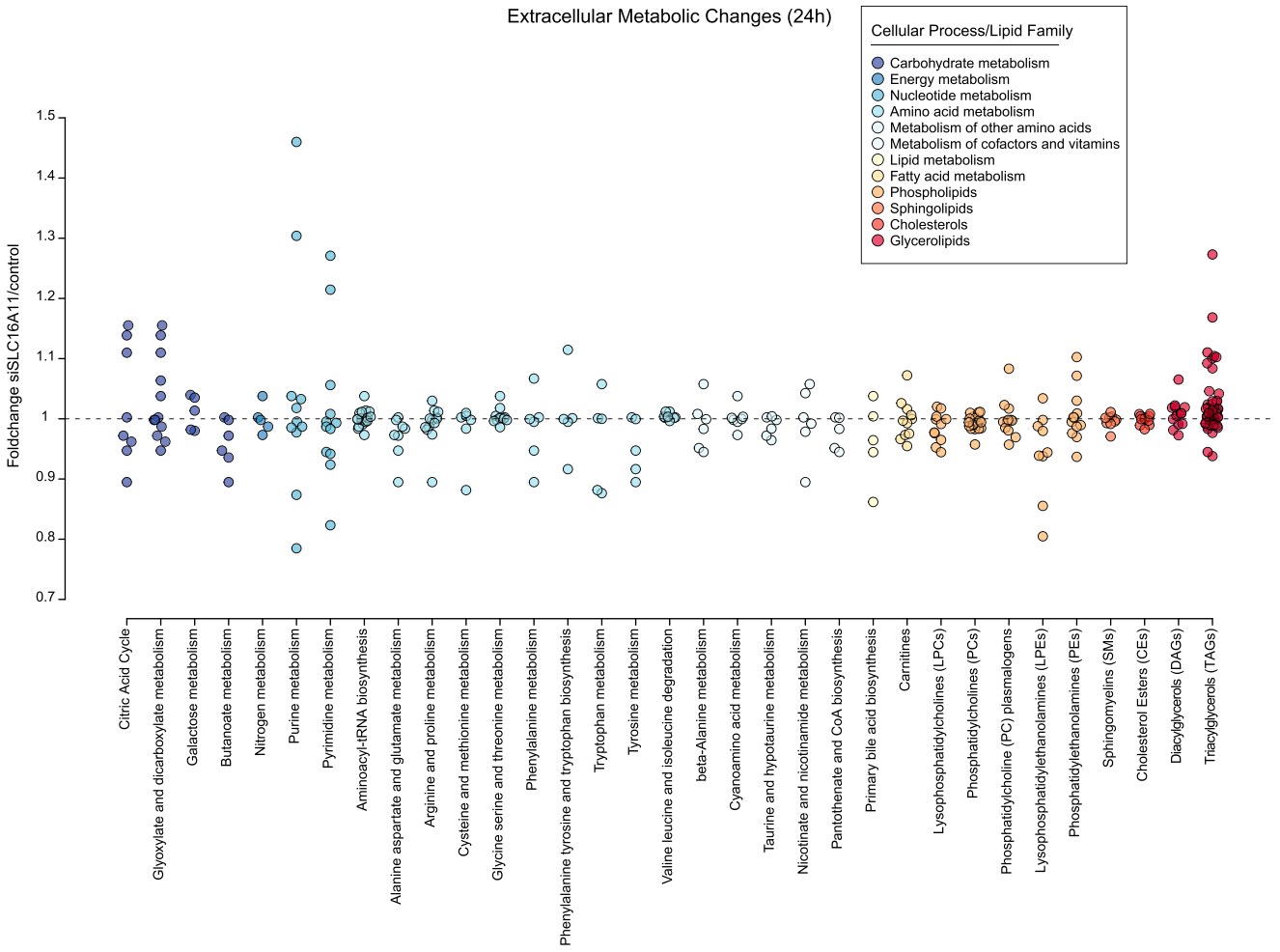


Figure S7. Extracellular Metabolic Pathways Altered by Knockdown of *SLC16A11* in Primary Human Hepatocytes, Related to Figure 6

Changes in extracellular levels of individual metabolites in primary human hepatocytes treated with *SLC16A11* siRNAs compared to control cells are plotted in groups according to metabolic pathway or class. Differences in extracellular metabolites reflect changes accumulated during the 24 hr period prior to sample collection. Pathways shown include all KEGG pathways from the human reference set for which at least five metabolites were measured as well as additional classes of metabolites covering carnitines and lipid sub-types. Each point within a pathway or class shows the fold-change of a single metabolite within that pathway or class. Pathways are colored according to cellular processes or lipid family as indicated in the legend. See also Table S4 for details on individual metabolite changes.