Fine Mapping and Functional Analysis Reveal a Role of SLC22A1 in Acylcarnitine Transport

Hye In Kim,[1] Johannes Raffler,[2] Wenyun Lu,[3] Jung-Jin Lee,[4] Deepti Abbey,[1] Danish Saleheen,[4] Joshua D. Rabinowitz,[3] Michael J. Bennett,[5] Nicholas J. Hand,[1] Christopher Brown,[1] and Daniel J. Rader[1,*]

Genome-wide association studies have identified a signal at the *SLC22A1* locus for serum acylcarnitines, intermediate metabolites of mitochondrial oxidation whose plasma levels associate with metabolic diseases. Here, we refined the association signal, performed conditional analyses, and examined the linkage structure to find coding variants of *SLC22A1* that mediate independent association signals at the locus. We also employed allele-specific expression analysis to find potential regulatory variants of *SLC22A1* and demonstrated the effect of one variant on the splicing of *SLC22A1*. *SLC22A1* encodes a hepatic plasma membrane transporter whose role in acylcarnitine physiology has not been described. By targeted metabolomics and isotope tracing experiments in loss- and gain-of-function cell and mouse models of *Slc22a1*, we uncovered a role of SLC22A1 in the efflux of acylcarnitines from the liver to the circulation. We further validated the impacts of human variants on SLC22A1-mediated acylcarnitine efflux *in vitro*, explaining their association with serum acylcarnitine levels. Our findings provide the detailed molecular mechanisms of the GWAS association for serum acylcarnitines at the *SLC22A1* locus by functionally validating the impact of *SLC22A1* and its variants on acylcarnitine transport.

Introduction

Acylcarnitines are intermediate metabolites of mitochondrial fatty acid and amino acid oxidation that shuttle acyl-moieties into and out of mitochondria. These mitochondrial metabolites are released from tissues and form a circulating pool of acylcarnitines. Plasma acylcarnitine levels have been found to associate with metabolic disease states such as obesity and diabetes.[1–12] Recent studies showed that acylcarnitines are not mere by-products of mitochondrial oxidation, but rather bioactive molecules that affect various metabolic and biological pathways including muscle energetics,[13] insulin secretion[14] and signaling,[15] and stress and inflammatory signaling.[1,15–18] Despite heightened interest in the emerging and diverse roles of acylcarnitines, the whole-body physiology of acylcarnitines is poorly understood.

Genome-wide association studies for blood metabolites have identified genetic loci that are significantly associated with circulating acylcarnitine levels.[19,20] Many of the associated loci contained genes that are involved in either carnitine metabolism (carnitine transporters) or fatty acid oxidation (carnitine acyltransferases and acyl-CoA dehydrogenases), both of which are known to affect plasma acylcarnitine levels. However, one of the strongly associated loci located at chromosome 6q25 harbors the *SLC22A1* gene (MIM: 602607), whose role in acylcarnitine biology has not been defined. Genetic variants at this locus are particularly associated with short-chain acylcarnitine species, suggesting a potential modulatory effect of SLC22A1 on this specific class of metabolites.

SLC22A1 is primarily expressed in the liver in humans[21] and encodes a plasma membrane transporter, also known as organic cation transporter 1, or OCT1. SLC22A1 is localized to the hepatocyte basolateral (sinusoidal) membrane[21] and transports its substrates between the liver and blood. Previously described substrates of SLC22A1 include a variety of endogenous and pharmacological molecules that commonly harbor a quaternary amine group,[22] a property also shared by carnitine and acylcarnitines.

In this study, we illustrate the molecular mechanisms of the association at the *SLC22A1* locus with serum acylcarnitine levels. We refined the association signal and performed allele-specific expression and conditional analyses to find independent causal variants of *SLC22A1* that mediate the association signal. Using loss- and gain-of-function cell and mouse models, we identified SLC22A1 as a cellular exporter of acylcarnitines in hepatocytes. We further demonstrated the effects of candidate variants on SLC22A1 protein and efflux function, which is directionally consistent with the observed genetic association.

Material and Methods

Serum Isobutyrylcarnitine Association in KORA F4 Cohort

The association study for serum metabolites in KORA F4 cohort was conducted as described previously.[19] The study was approved

[1]Department of Genetics, The Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA 19104, USA; [2]Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA; [4]Department of Biostatistics and Epidemiology, The Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA 19104, USA; [5]Department of Pathology and Laboratory Medicine, The Perelman School of Medicine of the University of Pennsylvania and Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
*Correspondence: rader@mail.med.upenn.edu
http://dx.doi.org/10.1016/j.ajhg.2017.08.008.
© 2017

by the local ethics committee, Bayerische Landesärztekammer, and written informed consent was obtained from all the participants. In brief, 1,768 European participants were genotyped on Affymetrix gene chips (Human SNP Array 6.0 and Axiom) and their fasting serum isobutyrylcarnitine levels were measured using the Metabolon LC/MS platform. Isobutyrylcarnitine levels were divided by the run-day median to correct for batch effects and were $\log_{10}$-scaled. Extreme data points (more than 4 SD away from the mean) were removed from the dataset to avoid spurious associations. In this study, genotypes were imputed using 1000 Genomes project data[23] (phase 1 v.3, March 2012 release, CEU), using SHAPEIT[24] (v.2) and IMPUTE[25] (v.2.3.0) for phasing and imputation, respectively. Association testing was performed using PLINK[26] (v.1.90) with linear regression models under the assumption of an additive genetic model and corrected for the effects of age and sex. Conditional analysis was performed using GCTA-COJO tool.[27] The association plot was generated using LocusZoom.[28] Linkage disequilibrium calculations were derived from 1000 Genomes project data,[23] using Haploview[29] (phase 1, CEU) or Haploreg[30] (v.4.1, phase 1, EUR).

Allele-Specific Expression Analysis

Allele-specific expression analysis was performed using the genotype and RNA-seq data of human liver samples obtained from two cohorts: 40 liver samples from the University of Pennsylvania (UPenn) and 32 liver samples from the Genotype-Tissue Expression (GTEx) consortium[31] (release v.4). The UPenn cohort was genotyped on the Illumina human 610 quad bead-chips.[32] The GTEx cohort was genotyped on the Illumina Human Omni 2.5 and 5.0 beadchips at the Broad Institute. RNA sequencing in the liver was performed using an Illumina HiSeq2500 and the reads were aligned to the human genome (GRCh37/hg19) using STAR in 2-pass mode. Reference mapping bias was controlled using WASP.[33] Reads harboring phased coding SNPs were assigned to either the major or the minor allele of the query SNP. Allele-specific expression was quantified by means of the combined haplotype test.[34]

CRISPR-Cas9 Gene Editing

Potential guide RNAs were screened using web-based MIT CRISPR Design Tool and the guide RNA that cuts nearest to the location of rs113569197 was selected (Figure S2A). The guide RNA was inserted to pSpCas9(BB)-2A-GFP vector following the standard protocol.[35] For homology directed repair templates, we used single-stranded DNA oligonucleotides (ssODNs) that consisted of 200 bp sequence flanking the targeted region with or without rs113569197 variant and mutated to harbor an EcoRI restriction enzyme site (Figure S2B). Huh7 hepatoma cells were transfected with guide RNA-Cas9 vector and ssODNs using Lipofectamine 3000 (Life Technologies) and sorted for GFP 48 hr after transfection. Colonies arising from single GFP-positive cells were screened by EcoRI digest and Sanger sequencing to find cells that are correctly edited. cDNAs from the correctly edited cells were obtained and the region between exons 7 and 8 was amplified by PCR and the PCR products were Sanger sequenced to examine the splicing pattern.

Haplotype Frequency and Association Analyses

We used SHAPEIT[24] (v.2) to phase the genotypes from KORA F4 cohort and estimate the haplotypes of each individual for rs12208357, rs202220802, and rs113569197. The frequency of each haplotype was calculated and only the haplotypes with frequency greater than 1% in the cohort were considered. Haplotype association for serum isobutyrylcarnitine levels was tested with linear regression models under the assumption of an additive genetic model and corrected for the effects of age and sex.

Slc22a1 Liver-Specific Knockout Mouse Model

Slc22a1 conditional knockout mice (*Slc22a1*$^{fl/fl}$) were from Merck and generated by a gene targeting approach. ES cells from C57BL/6 strain were transfected with a targeting vector to introduce LoxP sites and selection markers in areas flanking exons 2 and 3 of *Slc22a1* (Figure S3A). Targeted clones were selected and microinjected to blastocysts from BALB/c strain and transferred to surrogate females. The resulting chimeric mice were bred to C57BL/6 females to obtain *Slc22a1*$^{+/fl}$ on C57BL/6 background. *Slc22a1*$^{fl/fl}$ mice were subsequently crossed with Albumin-*Cre* transgenic mice (Jackson Laboratory) to generate *Slc22a1* liver-specific knockout mice (*Slc22a1*$^{fl/fl}$; Alb-*Cre*^{Tg}) and littermate controls (*Slc22a1*$^{fl/fl}$), referred in the text as *Slc22a1*$^{\Delta hep}$ and *Slc22a1*^{wt}, respectively. Mice were housed in a pathogen-free facility with controlled temperature and humidity and a 12-hour light/dark cycle. Mice were fed a standard chow diet (Rodent Diet 5010, LabDiet). All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania.

Immunofluorescence

Liver pieces fixed in 4% paraformaldehyde were prepared as paraffin sections by the CVI histology core at the University of Pennsylvania. Liver sections were blocked and incubated with SLC22A1 primary antibody (2C5, Novus Biologicals, NBP1-51684), ImmPRESS mouse secondary antibody, and TSA fluorescence. Subsequently, the sections were incubated with MDR primary antibody (C-19, Santa Cruz, sc-1517), ImmPRESS goat secondary antibody, and TSA fluorescence. ImmPRESS secondary antibodies were from Vector Laboratories and TSA fluorescence was from Perkin Elmer. Stained slides were examined by Leica TCS SP8 confocal microscope and images were acquired and processed using LAS X software and ImageJ software.

Adeno-Associated Virus (AAV)

Mouse *Slc22a1* cDNA (Origene) was subcloned to an AAV vector plasmid provided by the Penn Vector core of the University of Pennsylvania. The cDNA was placed under thyroxine binding globulin promoter to allow liver-specific expression. The cDNA-containing or empty AAV plasmids were packaged into AAV viral particles (serotype 2/8) by the Penn Vector core. 1×10^{12} particles were administered per mouse by intraperitoneal injection.

Gene Expression Analysis

RNA was extracted from tissues and cells using Trizol (Invitrogen) following the manufacturer's protocol. cDNA was generated from ~1 μg of total RNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Real-time PCR analysis was performed on QuantStudio 7 Real-Time PCR System using SYBR Green master mix or Taqman master mix (Life Technologies). The primers used were: for mouse *Slc22a1*, forward 5′-AGGCT GATGGAAGTTTGGCA-3′ and reverse 5′-GTGGGGATTTGCCTGTT TGG-3′; for mouse *Slc22a2*, forward 5′-TGGCATCGTCACAC CTTTCC-3′ and reverse 5′-AGCTGGACACATCAGTGCAA-3′; for mouse *Slc22a3*, forward 5′-TCAGAGTTGTACCCAACGACATT-3′ and reverse 5′-TCTGCCACACTGATGCAACT-3′; and for mouse

ActB, forward 5′-TTGGGTATGGAATCCTGTGG-3′ and reverse 5′-CTTCTGCATCCTGTCAGCAA-3′. Taqman primers and probes (Thermo Fisher) were used for human *SLC22A1* (Hs00427550_m1) and human *ACTB* (Hs01060665_g1). Expression levels were normalized to mouse *ActB* or human *ACTB* mRNA levels.

Carnitine Measurement

Carnitine levels in the plasma and tissues were measured by a plate assay. Tissues were homogenized in PBS and centrifuged to remove insoluble materials. Tissue homogenates were deproteinized using Microcon 10 kDa centrifugal filters (EMD Millipore). Filtered homogenates were used to measure carnitine contents using L-carnitine assay kit (Sigma). Protein contents in the liver homogenates were measured by Pierce BCA protein assay kit (Thermo Scientific) and used for normalization.

LC-MS for Carnitine and Acylcarnitine Measurement

Liver, plasma, and bile were collected after 24 hr fasting. Spot urine samples were collected in the fed state. Livers were homogenized in PBS and centrifuged to remove insoluble materials. Protein contents in the liver homogenates were used for normalization. Creatinine contents in the urine were measured by Creatinine Companion Kit (Glycadia/Exocell) and used for normalization. Liver and plasma acylcarnitines were measured as their butylated derivatives using tandem mass spectrometry. Stable isotope-labeled internal standards (Cambridge Isotope Laboratories) were added to plasma and liver homogenates. Ethanol was added and the samples were dried under a stream of nitrogen at 60°C. After adding butanolic hydrochloric acid (Regis Technologies), the samples were heated to 65°C for 15 min and dried under a stream of nitrogen. The dried samples were reconstituted with acetonitrile:water (80:20) and injected into a Xevo TQ-S tandem mass spectrometer (Waters Corporation). Data were acquired to collect the parent compounds of mass m/z 85. Quantitation was against the nearest chain-length stable isotope labeled internal standard. Bile and urine metabolites were extracted as follows. Pre-cooled methanol was added and samples were incubated at −20°C for 20 min. Samples were centrifuged and the supernatants were collected as the first extracts. The remaining pellets were reconstituted in pre-cooled methanol:acetonitrile:water (40:40:20) solution and incubated on ice for 10 min. Samples were centrifuged and the supernatants were collected and combined with the first extracts. The combined extracts were dried using Vacufuge (Eppendorf). Dried extracts were dissolved in LC-grade water and analyzed on a Q Exactive Plus mass spectrometer coupled to Vanquish UHPLC system (Thermo Fisher). The mass spectrometer was operated in positive ion mode with resolving power of 140,000 at m/z 200, scanning range being m/z 140–600. The LC separation was achieved on an Agilent Poroshell 120 Bonus-RP column (150 × 2.1 mm, 2.7 μm particle size). The gradient was 0 min, 50 μL/min, 0% B; 6 min, 50 μL/min, 0% B; 12 min, 200 μL/min, 70% B; 14 min, 200 μL/min, 100% B; 18 min, 200 μL/min, 100% B; 19 min, 200 μL/min, 0% B; 24 min, 200 μL/min, 0% B; 25 min, 50 μL/min, 0% B. Solvent A is 10 mM ammonium acetate with 0.1% acetic acid in water:acetonitrile (98:2) and solvent B is acetonitrile. Acylcarnitine species were detected based on accurate mass with a 5 ppm mass window and confirmed by authentic standards.

Transient Transfection of HEK293T Cells

Mouse *Slc22a1* (Origene) and human *SLC22A1* (DNASU[36]) were subcloned to pcDNA3.1 vector with V5-His tag using TOPO TA cloning kit (Invitrogen). HEK293T cells were cultured in DMEM supplemented with 10% FBS and 1% antibiotics. Prior to transfection, HEK293T cells were seeded on collagen-coated plates (Corning) in antibiotics-free culture media and allowed to attach overnight. Lipofectamine 3000 (Life Technologies) was used to induce transient overexpression. Media was changed to antibiotics-containing media 6 hr after transection. Experiments were performed 48 hr after transfection.

Mouse Primary Hepatocyte

Mouse primary hepatocytes were isolated by a standard collagenase method. Briefly, mice were anesthetized with a mixture of ketamine and xylazine. Livers were perfused with HBSS without Ca^{2+} and Mg^{2+} and subsequently with DMEM containing collagenase. Hepatocytes were freed by gentle tearing, washed, and plated on collagen-coated plates. Experiments were performed immediately after cells attached to the plates around 3–4 hr after plating.

In Vitro Transport Assays using Radiolabeled Substrate

Transport assays were performed on transiently transfected HEK293T cells and mouse primary hepatocytes. Uptake media was made of HBSS and 20 mM HEPES with respective transport substrates. For 1-methyl-4-phenylpyridinium (MPP^+) uptake, 10 μM MPP^+ (Sigma) and 0.2 μCi [^{3}H]-MPP+ (PerkinElmer) were added. For carnitine uptake, 20 μM carnitine (Sigma) and 0.5 μCi [^{3}H]-carnitine (American Radiolabeled Chemicals) were added. Cells were washed once with PBS and incubated in the uptake media for 3 min for MPP^+ uptake and 10 min for carnitine uptake. Cells were washed 2–3 times with PBS and lysed in 0.1N NaOH. The radioactivity in the cell lysate was measured by scintillation counter (Beckman Coulter) and normalized by the protein content in the cell lysate. For efflux assay, cells were washed once with PBS and first incubated in the uptake media containing [^{3}H]-carnitine for 30 min to label the cellular pool of carnitine and acylcarnitines. Cells were washed 2–3 times with PBS and subsequently incubated in the efflux media made of HBSS and 20 mM HEPES with 20 μM carnitine. At indicated time points, media was collected and cells were lysed in 0.1N NaOH. The radioactivity in the cell lysate and media was measured and normalized by the protein content in the cell lysate. Percent efflux was calculated as the percentage of the [^{3}H] count in the media out of the combined counts in the media and cell. In transport assays involving human SLC22A1, 50 mM L-valine (Sigma) was added to the transport media to drive the production of isobutyrylcarnitine, which showed the strongest association signal. SLC22A1-specific efflux activity was determined as the efflux activity in *SLC22A1*-expressing cells minus that in *GFP*-expressing control cells. For assessing Na^+ dependence of the transport activity, we used media composed of 125 mM NaCl, 4.8 mM KCl, 5.6 mM D-glucose, 1.2 mM $CaCl_2$, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$, and 25 mM HEPES (pH 7.5).[37] In Na^+-free media, NaCl was replaced with KCl. Cells were pre-incubated in Na^+-containing or Na^+-free media for 20 min before the uptake experiment.

Immunoblotting

Tissues and cells were lysed in RIPA buffer containing protease inhibitors (Roche). Tissue homogenates were centrifuged to remove insoluble materials. 50 μg of total protein was separated in Bis/Tris NuPage gels (Invitrogen) and transferred to nitrocellulose membranes. Membranes were blocked in 5% milk solution and blotted with primary antibodies and subsequently with HRP-conjugated

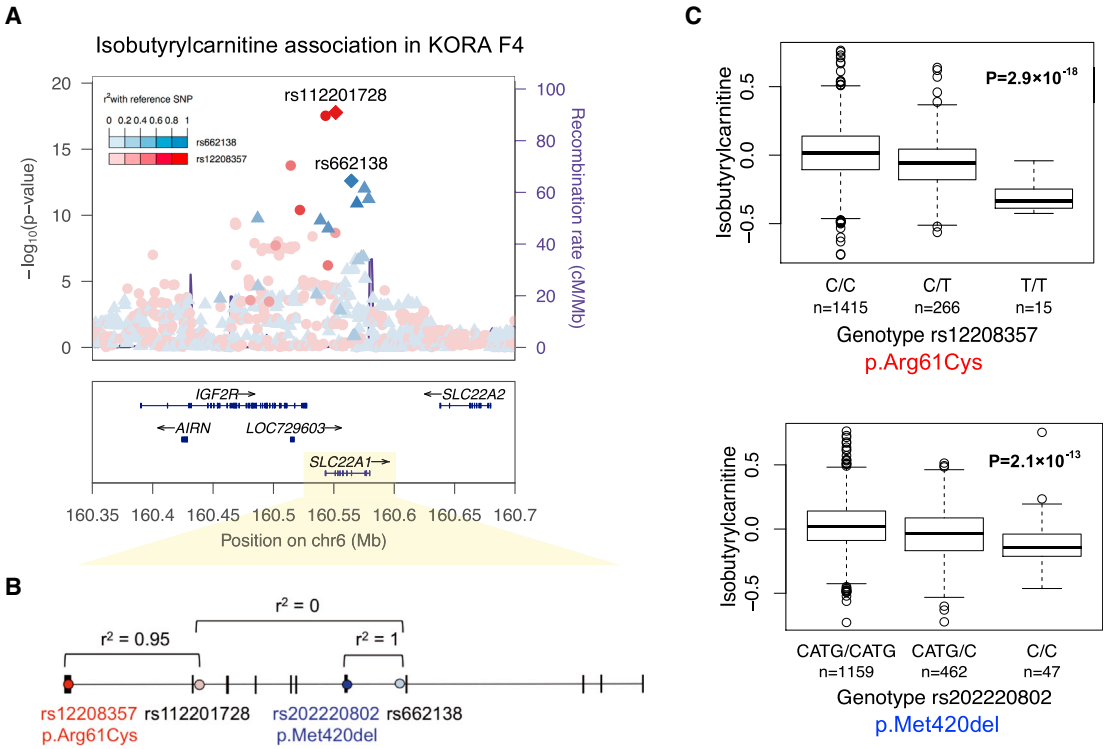

Figure 1. Refinement of the Association Signal at the *SLC22A1* Locus Reveals Two Coding Variants of *SLC22A1* to Be Associated with Serum Acylcarnitine Levels

(A) Serum isobutyrylcarnitine association data at the *SLC22A1* locus from KORA F4 cohort were refined using 1000 Genomes project data for imputation. The two strongest signals are shown in red circles and blue triangles, with the shade indicating the degree of linkage to the sentinel variants rs112201728 and rs662138.

(B) rs112201728, the sentinel variant of the red signal, is tightly linked to rs12208357 coding variant, which encodes p.Arg61Cys substitution. rs662138, the sentinel variant of the blue signal, is tightly linked to rs202220802 coding variant, which encodes p.Met420del substitution.

(C) Normalized serum isobutyrylcarnitine levels were $\log_{10}$-transformed and plotted according to the genotypes. The minor alleles of rs12208357 and rs202220802 are associated with lower serum isobutyrylcarnitine levels.

secondary antibodies. Primary antibodies used were anti-SLC22A5 (LifeSpan Biosciences, LS-C313281), anti-V5 tag (Thermo Fisher, R960-25), and anti-GAPDH (V-18, Santa Cruz, sc-20357). Membranes were incubated in ECL (Amersham) or Crescendo (EMD Millipore) reagents and detected by ImageQuant system (GE Healthcare Life Sciences).

Statistics

Two-tailed t test or two-way ANOVA were used for statistical analysis. *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$.

Results

Refinement and Conditional Analyses of the Association Signal Reveals Two Independent Coding Variants at the *SLC22A1* Locus Associated with Acylcarnitines

We first sought to refine the association signal at the *SLC22A1* locus using the genotype and serum metabolite data derived from KORA F4 cohort consisting of 1,768 European subjects.[19] While the original GWAS imputed genotypes based on the HapMap project data,[38] we imputed a denser set of genotypes based on the 1000 Genomes project data.[23] We specifically examined the association for isobutyrylcarnitine since it was the most strongly associated acylcarnitine species. Two strong signals were found, each of which consisted of SNPs in linkage disequilibrium (LD) (Figure 1A): in addition to the association signal with rs662138 that was previously found in the original GWAS,[19] we discovered an additional association signal with rs112201728 that could be imputed based on the 1000 Genomes project data. The sentinel SNPs of these signals, rs112201728 and rs662138, are not in LD ($r^2 = 0$), suggesting that these two signals are independent of each other.

To confirm independent signals at the locus, we performed conditional analyses where the association of a variant is analyzed while accounting for the effect of another variant. When corrected for the effect of rs112201728, rs662138 remained significantly associated (Figures S1A and S1B and Table S1). When corrected for the effect of rs662138, rs112201728 remained significantly associated (Figures S1A and S1C and Table S1). These results show that the signals represented by rs112201728 and rs662138 are mutually independent. When conditioned on both rs112201728 and rs662138, there was no remaining association signal that reached genome-wide

significance ($p < 10^{-8}$) (Figure S1D). Therefore, rs112201728 and rs662138 represent two independent association signals, each with genome-wide statistical significance.

As rs112201728 and rs662138 are intronic variants with no known or predicted *cis*-regulatory effects, we suspected that they may tag true causal variants that are linked to them. We found that rs112201728 is in strong LD ($r^2 = 0.95$) with a coding variant, rs12208357, which encodes p.Arg61Cys (c.181C>T) substitution in SLC22A1 protein (Figure 1B). The minor allele of rs12208357 encoding the cysteine residue is significantly associated with lower serum isobutyrylcarnitine levels ($p = 2.9 \times 10^{-18}$ and beta = −0.102; Figure 1C). rs662138, on the other hand, is in perfect LD ($r^2 = 1$) with a 3-bp deletion coding variant, rs202220802, that encodes p.Met420del (c.1258_1260delATG) substitution in SLC22A1 protein (Figure 1B). Using 1000 Genomes project data,[23] we imputed the rs202220802 genotype and found that the minor allele encoding the deletion is significantly associated with lower serum isobutyrylcarnitine levels ($p = 2.1 \times 10^{-13}$ and beta = −0.066; Figure 1C). As rs12208357 (p.Arg61Cys) and rs202220802 (p.Met420del) variants are known to impair the transport function of SLC22A1 protein,[39,40] these data strongly suggest *SLC22A1* as the candidate gene for the association signal at this locus.

Liver Allele-Specific Expression Analysis Reveals a Splicing Variant of *SLC22A1*

We noted that, after conditioning on the two sentinel variants, there still were some variants that remained associated at sub-threshold levels of significance ($p < 10^{-4}$) (Figure S1D). We hypothesized that there might be additional causal variants, potentially non-coding variants that affect *SLC22A1* at the transcript level. To find potential regulatory variants, we performed an allele-specific expression (ASE) analysis using RNA-seq and genotype data from 72 human liver samples. This analysis identified variants that are significantly associated with allele-specific transcript levels of *SLC22A1*, with rs594709 being the most strongly associated variant (Figure 2A). When the allele-specific reads of each rs594709 heterozygote subject were plotted, the data points clearly skewed from 1:1 ratio, indicating ASE (Figure 2B). The fraction of the major "A" allele-specific reads out of the total reads was 60.8%, while that of the minor "G" allele-specific reads was 39.2% (Figure 2C). The minor "G" allele-specific reads were, in average, 34.1% lower than the major "A" allele-specific reads in heterozygous subjects.

In KORA F4 cohort, the minor allele of rs594709 was associated with lower serum isobutyrylcarnitine levels with p value of 6.3×10^{-4}. We tried to examine whether the association of rs594709 is independent of the two coding variants described above, rs12208357 and rs20220802, but found that these variants co-segregate with the minor and the major allele of rs594709 variant ($D' = 1$, but $r^2 < 0.2$ due to difference in frequency),[23,30] respectively. When conditioned on rs112201728 (a proxy for rs12208357), the association of rs594709 became weaker (p value from 6.3×10^{-4} to 3.3×10^{-1}) (Figures S1A and S1B and Table S1), indicating that the co-segregating minor alleles had additive effects on acylcarnitines. In contrast, when conditioned on rs662138 (a proxy for rs202220802), the association of rs594709 became stronger (p value from 6.3×10^{-4} to 1.4×10^{-10}) (Figures S1A and S1C and Table S1), indicating that the co-segregating alleles (minor allele of rs202220802 and major allele of rs594709) had opposing effects on acylcarnitines. When conditioned on both rs112201728 and rs662138, the association of rs594709 was similar to the strength of the original association ($p = 7.6 \times 10^{-5}$) (Figure S1D and Table S1). Conversely, when conditioned on rs594709, the association of rs112201728 diminished (p value from 2.9×10^{-18} to 1.2×10^{-15}), while the association of rs662138 improved (p from 2.5×10^{-13} to 1.2×10^{-19}) (Figures S1A and S1E and Table S1). These results indicate that the association of rs594709 with acylcarnitines is confounded by linkage with the coding variants (and vice versa), but it remains independent of the coding variants.

To find the causal variant(s) for the top ASE signal, we examined the variants in strong LD ($r^2 > 0.8$) with rs594709 for their potential effects on *SLC22A1* transcript (Figure 2D). We found that rs113569197, which is in tight LD ($r^2 = 0.91$) with rs594709, resides near the 3′ junction of exon 7 and could affect the splicing between exons 7 and 8. rs113569197 is an 8-bp insertion variant that duplicates the splicing motif 8 bp downstream of the original splicing motif (Figure 2E). If the splicing happens at the duplicated site, it will introduce 8-nucleotide (nt) intronic sequence between exons 7 and 8, resulting in frameshift and premature stop codon. We hypothesized that rs113569197 drives the ASE signal by generating an aberrantly spliced transcript that would be targeted for nonsense-mediated decay, resulting in lower transcript levels.

To directly test whether rs113569197 causes incorrect splicing, we used CRISPR-Cas9 gene editing to introduce rs113569197 into Huh7 hepatoma cells and generated cells that are wild-type, heterozygous, or homozygous for rs113569197 genotype (Figure S2C). We Sanger sequenced the cDNA sequence at the junction between exons 7 and 8 in these cells and detected the abnormally spliced *SLC22A1* transcript that includes 8-nt intronic sequence in heterozygous and homozygous cells (Figures 2F and S2D). This result suggests that rs113569197 is the causal variant that mediates the strongest ASE signal for *SLC22A1* by inducing the abnormally spliced transcript that would be degraded.

Haplotype Association Analyses Corroborate the Independent Association of the Coding Variants

We proposed that rs12208357 (p.Arg61Cys), rs202220802 (p.Met420del), and rs113569197 (splicing) variants mediate independent association with serum acylcarnitine

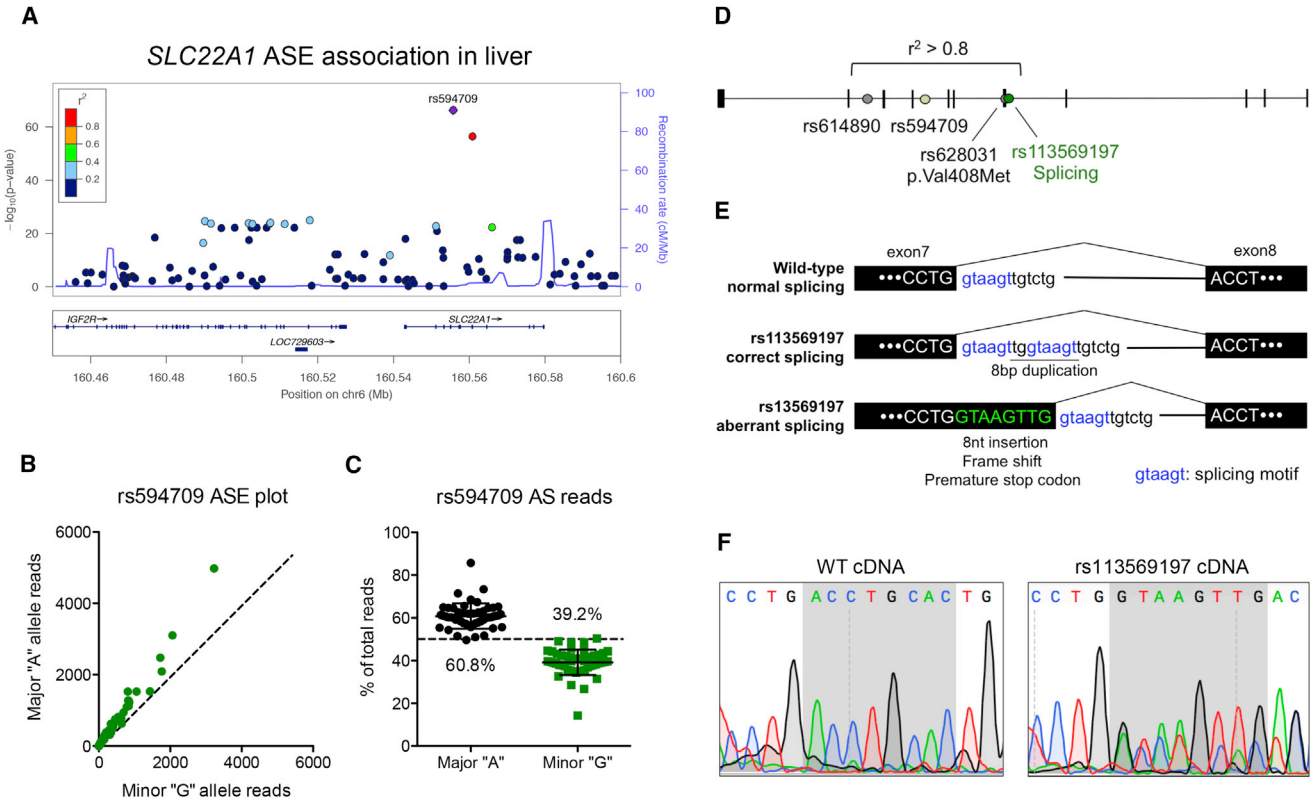

Figure 2. Allele-Specific Expression Analysis Reveals a Splicing Variant of *SLC22A1*
(A) Locus-wide allele-specific expression analysis identified variants that are associated with allele-specific *SLC22A1* transcript abundance.
(B) Allele-specific reads in subjects heterozygous for rs594709 are plotted. The dotted line indicates the expected ratio when no ASE is present.
(C) The fraction of allele-specific reads out of total reads is indicated according to rs594709 alleles. Data are presented as mean ± SD.
(D) Variants that are in strong LD with rs594709 sentinel SNP are shown along *SLC22A1*.
(E) rs113569197 duplicates a conserved splicing motif, creating an alternative splicing site 8 bp downstream of the original splicing site. This aberrant splicing leads to the inclusion of 8-nt sequence at the end of exon 7, resulting in frameshift and premature stop codon.
(F) rs113569197 was introduced to Huh7 hepatoma cell line by CRISPR-Cas9 editing. Sanger sequencing at the junction between exons 7 and 8 revealed the presence of incorrectly spliced transcript in cells homozygous for rs113569197.

levels based on conditional analyses. To examine the relationship among the three variants further, we constructed the haplotypes of the three variants in KORA F4 cohort and found four haplotypes that exist at frequency over 1% (Figure 3A). Consistent with the D' values, the minor allele of rs12208357 was strictly phased with the minor allele of rs113569197, while the minor allele of rs202220802 was strictly phased with the major allele of rs113569197. Next, we performed haplotype association tests where serum isobutyrylcarnitine levels were compared among individuals with haplotypes that differ by only one test variant to assess the effect of the variant while controlling for the other two variants (Figure 3B). The two coding variants were strongly and independently associated with lower serum isobutyrylcarnitine levels; however, the strength of the association was substantially diminished compared to the original single variant association presumably due to the reduction in sample size. We did not find evidence for the independent association of rs113569197 variant in this analysis, which we suspect is largely due to the reduced sample size that rendered limited statistical power to detect association with small effect sizes.

Slc22a1 Liver-Specific Knockout Mice Display Systemically Altered Acylcarnitine Profiles

To interrogate the function of SLC22A1 in acylcarnitine biology, we generated *Slc22a1* liver-specific knockout mice (*Slc22a1*$^{\Delta hep}$) by crossing *Slc22a1*$^{fl/fl}$ mice to Albumin-*Cre* transgenic mice (Figure S3A). Reduced *Slc22a1* mRNA and SLC22A1 protein levels were confirmed in the livers of *Slc22a1*$^{\Delta hep}$ mice by qPCR and immunohistochemistry, respectively (Figures S3B–S3D). Consistent with reports on the human SLC22A1 protein,[21] mouse SLC22A1 protein in hepatocyte is localized to the basolateral membrane facing the blood vessels and is excluded from the apical membrane facing the biliary tract. This indicates that SLC22A1 mediates transport between liver and blood and not between liver and bile.

We performed targeted metabolomics to measure acylcarnitine levels in *Slc22a1*$^{\Delta hep}$ and wild-type mice. Levels relative to wild-type mice are shown in Figure 4, while

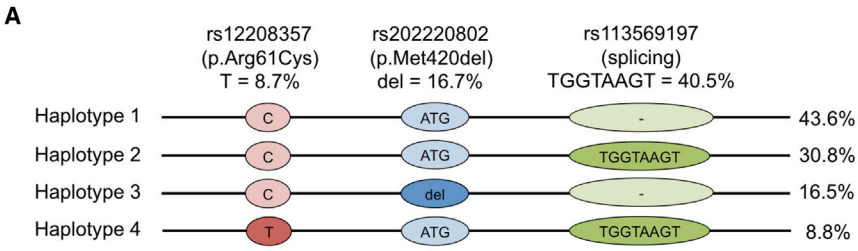

Figure 3. Haplotype Association Analyses Confirm the Independent Association of the Coding Variants

(A) Four haplotypes were observed at frequency over 1% for the three variants, rs12208357, rs202220802, and rs113569197, in KORA F4 cohort.

(B) Serum isobutyrylcarnitine levels were compared among individuals with haplotypes that differ only by one test variant to assess the association of the variant.

B

Test variant	Haplotypes	Effect estimate	P-value	Standard error
rs12208357	Hap2/Hap4	-0.113	5.12×10^{-8}	0.020
rs202220802	Hap1/Hap3	-0.068	4.21×10^{-9}	0.011
rs113569197	Hap1/Hap2	-0.013	1.44×10^{-1}	0.008

absolute levels are shown in Figure S4. Livers from *Slc22a1*$^{\Delta hep}$ mice had greatly elevated levels of carnitine and acylcarnitines compared to wild-type livers (Figure 4A). Interestingly, while the increased hepatic levels of medium (C8-12)- and long (C14-18)-chain acylcarnitines were reflected by corresponding increases in the plasma of *Slc22a1*$^{\Delta hep}$ mice, plasma carnitine and short (C2-C6)-chain acylcarnitine levels were unaltered despite elevated levels in the liver (Figure 4B). This led to significantly lower plasma-to-liver ratios for carnitine and short-chain acylcarnitines in *Slc22a1*$^{\Delta hep}$ mice (Figure 4C), suggesting that SLC22A1 normally exports carnitine and short-chain acylcarnitines from the liver to blood. Supporting this notion, the urinary excretion of carnitine and short-chain acylcarnitine was significantly decreased in *Slc22a1*$^{\Delta hep}$ mice (Figure 4D), presumably due to reduced hepatic output to the circulation. Carnitine and acylcarnitines in the liver can be alternatively removed to bile. Bile from *Slc22a1*$^{\Delta hep}$ mice showed significantly elevated carnitine and acylcarnitine levels (Figure 4E), suggesting a compensatory increase in the biliary secretion of hepatic carnitine species in the absence of SLC22A1. Collectively, acylcarnitine profiles in *Slc22a1*$^{\Delta hep}$ mice suggest that SLC22A1 transports carnitine and short-chain acylcarnitines from the liver to blood.

To demonstrate that hepatic *Slc22a1* is responsible for the altered acylcarnitine profiles in *Slc22a1*$^{\Delta hep}$ mice, we performed a rescue experiment by injecting *Slc22a1*$^{\Delta hep}$ mice with AAV-expressing mouse *Slc22a1* cDNA under the control of thyroxine binding globulin promoter. We confirmed that *Slc22a1* expression is reconstituted in the liver of *Slc22a1*$^{\Delta hep}$ mice 10 weeks after AAV injection (Figure 4F). Reconstitution of hepatic *Slc22a1* expression in *Slc22a1*$^{\Delta hep}$ mice indeed normalized or even reversed acylcarnitine profiles in the liver and plasma (Figures 4G and 4H).

SLC22A1 Exports Carnitine Species from Cells to Media

To directly test whether SLC22A1 transports carnitine and/or acylcarnitines, we performed *in vitro* transport assays. We first used HEK293T cells transiently transfected to overexpress *GFP* control or mouse *Slc22a1*. Cells overexpressing mouse *Slc22a1* showed markedly increased uptake of 1-methyl-4-phenylpyridinium (MPP^+), a known substrate of SLC22A1 (Figure 5A). Overexpression of *Slc22a1* did not affect carnitine uptake, indicating that SLC22A1 does not transport carnitine from the outside to the inside of the cells (Figure 5B). Next, we tested whether SLC22A1 transports carnitine and acylcarnitines from the inside to the outside of the cells. Cells were first incubated in a media containing [^{3}H]-carnitine to label the cellular pool of carnitine and acylcarnitines (referred to collectively as "carnitine species" hereafter) and washed and subsequently incubated in a fresh media without radiolabel to allow the detection of effluxed [^{3}H]-labeled carnitine species. The amount of [^{3}H] in the media and cells were measured and the ratio of [^{3}H] in the media to the total [^{3}H] (media and cell combined) was calculated to determine the overall efflux activity for carnitine species. *Slc22a1* overexpression markedly increased the efflux of [^{3}H], indicating that SLC22A1 transports carnitine species from the inside to the outside of cells (Figure 5C).

As carnitine species exist at much higher concentrations inside the tissues than in the blood, the efflux of carnitine species from tissues to blood is expected to be passive and concentration dependent. This contrasts with Na^+-dependent active carnitine uptake from the blood into the tissues against the concentration gradient.[37] Indeed, while carnitine uptake was significantly reduced (Figure S5A) in the absence of Na^+ in the media, SLC22A1-mediated efflux of carnitine species was not impaired (Figure S5B).

To test whether endogenous SLC22A1 effluxes carnitine species from hepatocytes, we next performed transport assays in the primary hepatocytes isolated from *Slc22a1*$^{\Delta hep}$ and wild-type mice. Hepatocytes from *Slc22a1*$^{\Delta hep}$ mice showed decreased MPP^+ uptake, confirming reduced SLC22A1 activity (Figure 5D). Unexpectedly,

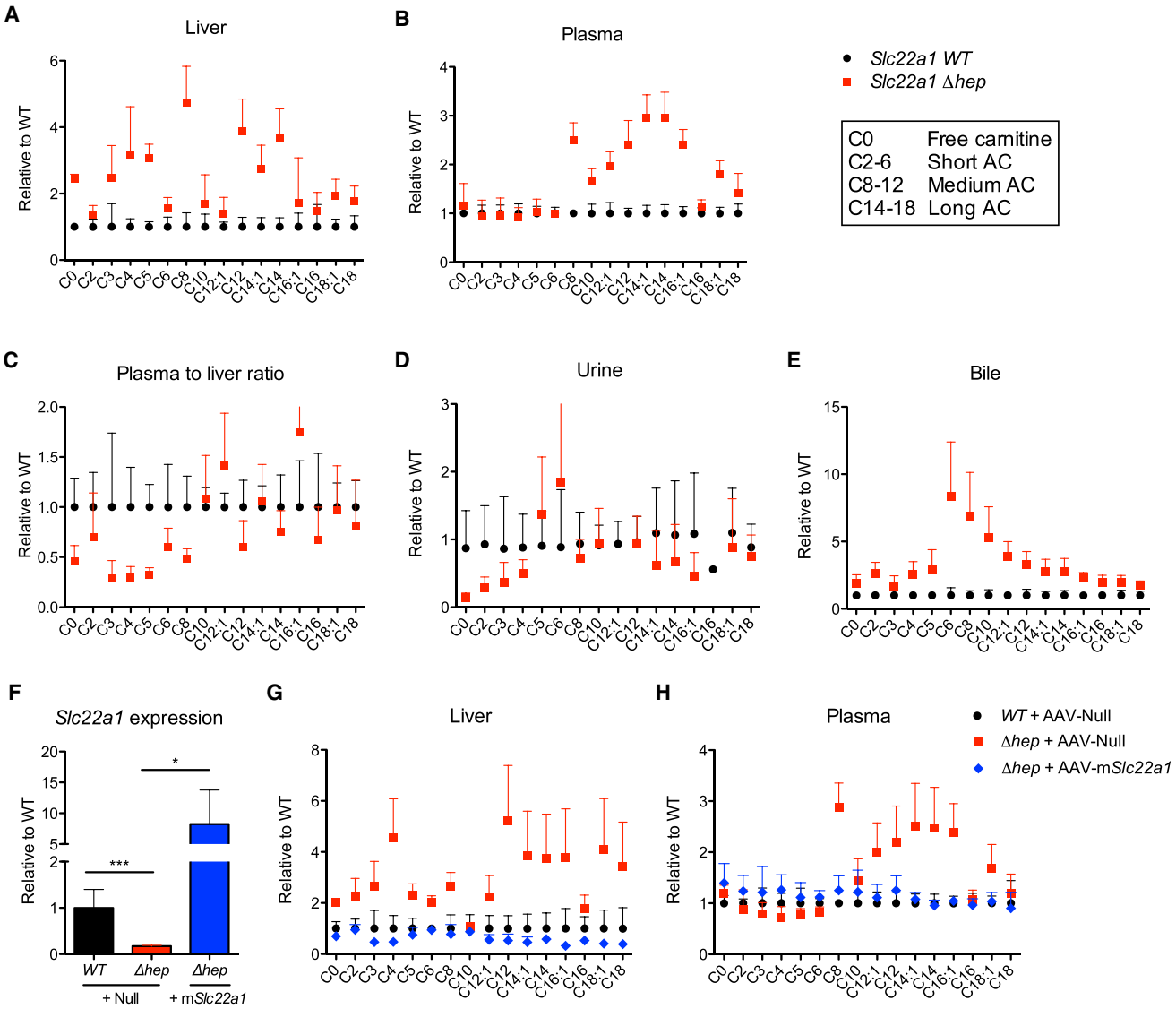

Figure 4. *Slc22a1*$^{\Delta hep}$ Mice Display Systemically Altered Acylcarnitine Profiles
(A–E) Acylcarnitine levels were measured in the liver (A), plasma (B), bile (E), and urine (D) (n = 3–7, males). Relative levels to the wild-type levels are indicated.
(F–H) *Slc22a1*^{wt} and *Slc22a1*$^{\Delta hep}$ mice were injected with AAV Null or AAV-m*Slc22a1* (n = 4, males). 10 weeks after AAV injection, mRNA expression and acylcarnitine levels were measured.
(F) Hepatic expression of *Slc22a1* is restored in *Slc22a1*$^{\Delta hep}$ mice that are injected with AAV-m*Slc22a1*.
(G and H) Hepatic reconstitution of *Slc22a1* expression normalizes liver and plasma acylcarnitine profiles in *Slc22a1*$^{\Delta hep}$ mice. Liver levels were normalized by the amount of protein. Urine levels were normalized by the amount of creatinine.
Data are presented as mean + SD. In (F), two-tailed t test was used for statistical analysis. *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$.

SLC22A1-deficient hepatocytes showed reduced carnitine uptake (Figure 5E) although SLC22A1 does not mediate carnitine uptake (Figure 5B). We suspected that this could be due to a compensatory change in SLC22A5 protein, a canonical carnitine uptake transporter. We confirmed that SLC22A5 protein levels were reduced in the livers of *Slc22a1*$^{\Delta hep}$ mice (Figure S6), which may have been downregulated to offset carnitine accumulation in the liver. As expected, SLC22A1-deficient hepatocytes showed reduced [^{3}H] efflux to the media (Figure 5F), indicating that SLC22A1 normally effluxes carnitine species from hepatocytes.

Variants of *SLC22A1* Impair SLC22A1-Mediated Efflux of Carnitine Species

We earlier showed that two coding variants, rs12208357 (p.Arg61Cys) and rs202220802 (p.Met420del), and one splicing variant, rs113569197, of *SLC22A1* are associated with serum acylcarnitine levels, suggesting that they impact the efflux function of SLC22A1 for acylcarnitines. We generated human *SLC22A1* cDNA clones harboring the two coding mutations and transfected them into HEK293T cells. While *SLC22A1* mRNA levels were comparable (Figure 6A), both mutations led to significantly reduced SLC22A1 protein levels (Figure 6B). Consistent

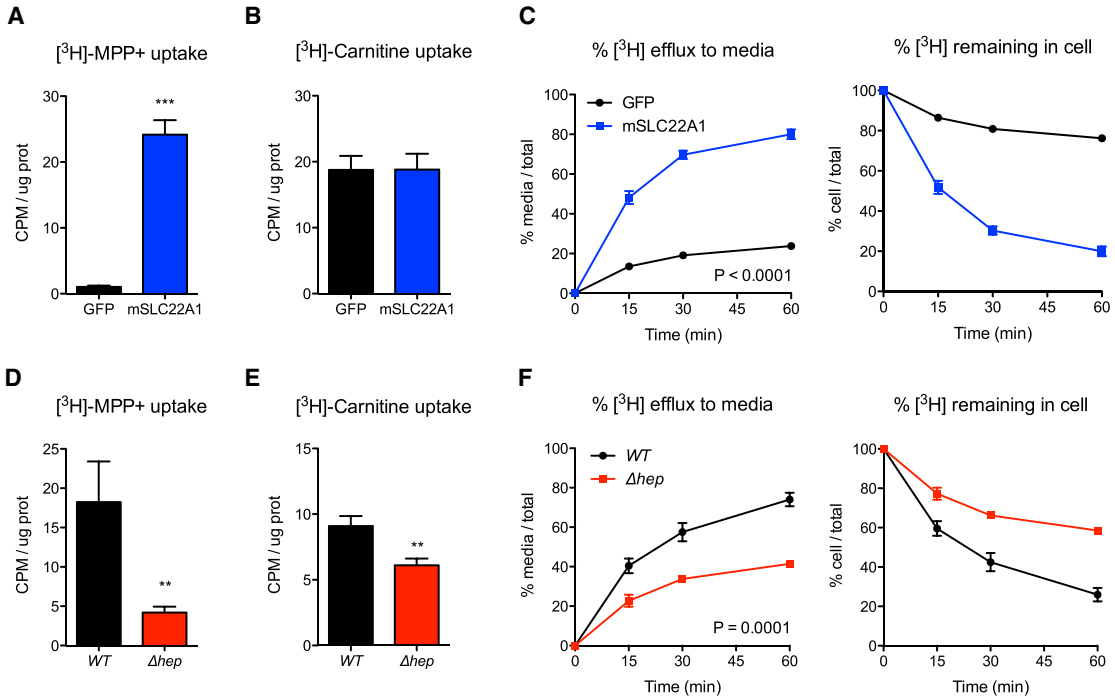

Figure 5. SLC22A1 Effluxes Carnitine Species from Cells to Media
(A–C) Experiments in HEK293T cells overexpressing *GFP* control or mouse *Slc22a1* are shown. A representative result from three technical replicates is presented.
(A) Mouse *Slc22a1* overexpression increases the uptake of MPP^+, a canonical substrate of SLC22A1 protein.
(B) SLC22A1 does not affect carnitine uptake.
(C) SLC22A1 increases the efflux of [^{3}H]-labeled carnitine species.
(D–F) Experiments in primary hepatocytes isolated from wild-type and *Slc22a1*$^{\Delta hep}$ mice (n = 3, females) are shown.
(D) Hepatocytes from *Slc22a1*$^{\Delta hep}$ mice have reduced MPP^+ uptake, confirming the lack of SLC22A1.
(E) Carnitine uptake is decreased in SLC22A1-deficient hepatocytes.
(F) Lack of SLC22A1 reduces the efflux of [^{3}H]-labeled carnitine species.
Data are presented as mean ± SD. Two-tailed t test was used for statistical analysis in (A), (B), (D), and (E). *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$. Two-way repeated-measures ANOVA test was used in (C) and (F).

with previous reports,[41,42] SLC22A1 protein was detected at two different molecular weights depending on its glycosylation status. To confirm that these mutations affect SLC22A1-mediated efflux of carnitine species, we performed radioactive efflux assays in these cells. SLC22A1-specific efflux activity was calculated by subtracting the efflux activity in *GFP*-expressing cells from the efflux activity in *SLC22A1*-expressing cells. As predicted, p.Arg61Cys and p.Met420del significantly impaired SLC22A1-specific efflux of carnitine species from cells (Figure 6C).

rs113569197 induces incorrect splicing of *SLC22A1* transcript, which would accelerate the turn-over of the resulting transcript due to nonsense-mediated decay. The protein that is translated from the aberrant transcript prior to its decay is predicted to have a frameshift starting at amino acid 426, followed by a premature stop codon at 433. To verify the effect of the incorrect splicing on SLC22A1 protein, we generated *SLC22A1* cDNA clone encoding the aberrantly spliced transcript (termed p.Asp426fs [c.1276_1277insGTAAGTTG]) and overexpressed it in HEK293T cells. While the mutant clone was expressed comparably to the wild-type clone (Figure 6D), the resulting truncated protein was rapidly turned over and barely detectable in the cells (Figure 6E). As expected, the cells overexpressing p.Asp426fs clone did not show SLC22A1-specific efflux activity for carnitine species (Figure 6F).

Discussion

Acylcarnitines are secreted mitochondrial metabolites that have gained substantial attention due to their implications in metabolic diseases; however, it has been largely unknown which genes and proteins export these molecules from cells to the circulation. We identified *SLC22A1* as a candidate gene solely based on its association with serum acylcarnitine levels in metabolite GWASs[19,20] and experimentally validated that SLC22A1 exports hepatic acylcarnitines to the circulation and affects whole-body acylcarnitine profiles.

Acylcarnitines are best known for their roles in mitochondrial oxidation. Although the main site of action of acylcarnitines is considered to be inside the cell, specifically in mitochondria, they are secreted from cells and are found in body fluids including blood, urine, and bile.

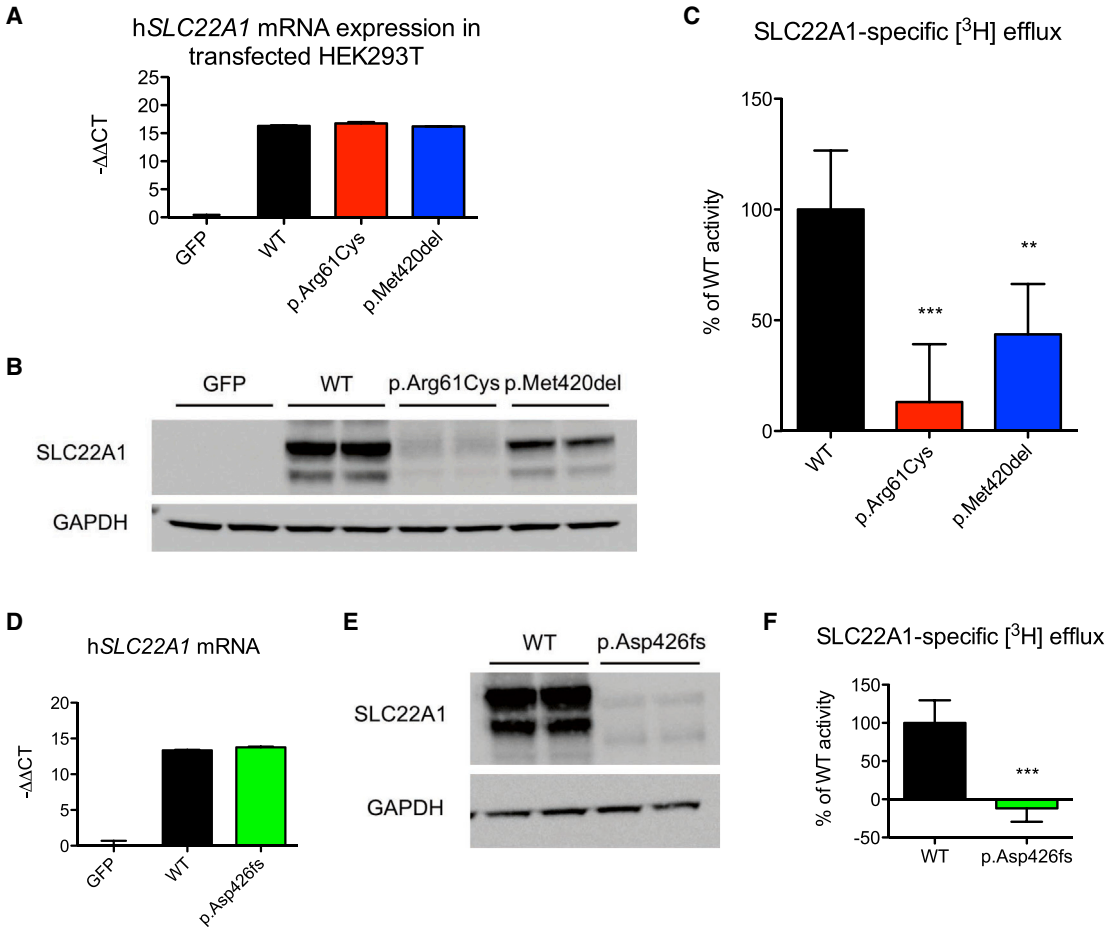

Figure 6. Coding and Splicing Variants of *SLC22A1* Impair the Efflux Function of SLC22A1 for Carnitine Species
(A) Human *SLC22A1* cDNA clones harboring the coding mutations of interest were generated and transfected into HEK293T cells. *SLC22A1* mRNA expression was comparable among wild-type and mutant clones.
(B) p.Arg61Cys and p.Met420del variants reduced SLC22A1 protein levels.
(C) p.Arg61Cys and p.Met420del variants reduced SLC22A1-specific efflux of [^{3}H]-labeled carnitine species.
(D) *SLC22A1* cDNA clone encoding the aberrantly spliced transcript (p.Asp426fs) was expressed at a comparable level to the wild-type clone in HEK293T cells.
(E) The resulting protein product with frameshift and premature truncation was barely detectable.
(F) SLC22A1-specific efflux activity was absent in HEK293T cells that are transfected with p.Asp426fs cDNA clone.
Data are presented as mean + SD. Two-tailed t tests were used for statistical analysis. *$p < 0.05$; **$p < 0.01$; ***$p < 0.005$.

One potential reason acylcarnitines are secreted from cells is to remove products that either are incompletely oxidized or cannot be oxidized by the cell, which might otherwise be cytotoxic. Alternatively, acylcarnitines could be secreted to the circulation to be exchanged among different tissues or act as signaling molecules. Recent studies suggested novel roles of plasma acylcarnitines in muscle energetics,[13] insulin secretion in pancreatic β-cells,[14] or inflammation in monocytes and macrophages.[1,16,18] Since liver is the major source of circulating acylcarnitines,[43] SLC22A1 may have a broad metabolic or biological impact on various tissues by regulating the systemic availability of hepatic acylcarnitines.

GWAS is a powerful approach to find candidate genes and variants that modulate biological traits and diseases of interest; however, identifying and validating the genes and variants that mediate the association is challenging. In our study, the association for serum isobutyrylcarnitine spanned an LD block that contains *IGF2R* (MIM: 147280) and *SLC22A1*. The sentinel variants were non-coding variants with no known or predicted effects on either *IGF2R* or *SLC22A1*. By close examination of the LD structure followed by experimental validation, we showed that the underlying causal variants were in fact the coding variants of *SLC22A1*. Similarly, the strongest association for *SLC22A1* ASE was captured by an intronic rs594709 variant that has no known effect on *SLC22A1* transcription. Instead, we proposed the rs113569197 variant, which is tightly linked to rs594709 variant, as the causal variant by verifying its effect on the splicing of *SLC22A1* transcript. The affected transcript would be targeted for nonsense-mediated decay, which would explain the allele-specific difference in

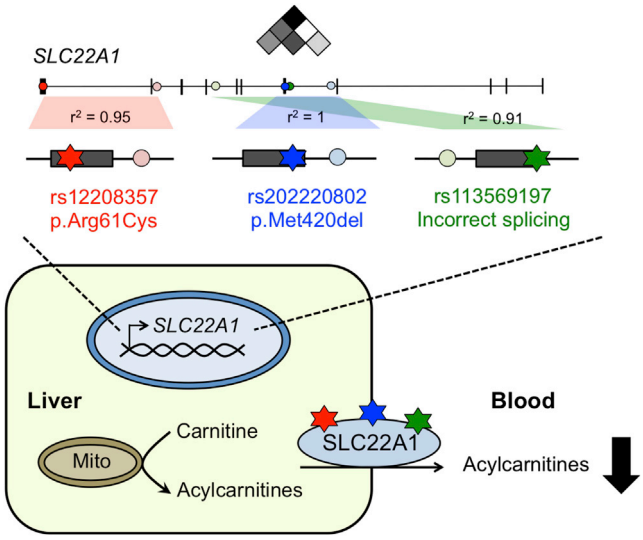

Figure 7. Discovery and Functional Validation of the Gene and Variants that Mediate Serum Acylcarnitine Association at the *SLC22A1* Locus

Cell and mouse models of SLC22A1 revealed its role in the efflux of hepatic acylcarnitines to the blood. Refinement of the GWAS association signal and close examination of the linkage structure identified rs12208357 (p.Arg61Cys) and rs202220802 (p.Met420del) coding variants as the causal variants for the association at the locus. Allele-specific expression analysis in human livers revealed rs113569197 variant that affects the splicing of *SLC22A1* transcript. These variants were shown to impair the expression and function of SLC22A1 protein *in vitro*. In humans, these variants are predicted to reduce the efflux of hepatic acylcarnitines to the blood, which would explain their association with lower serum acylcarnitine levels.

transcript level. These results indicate the importance of taking local LD structure into consideration for the identification of causal variants in GWASs.

When causal variants are not linked and inherited independently, the association of one causal variant is not affected by other causal variants. For example, we demonstrated by conditional analyses that the association of rs112201728 (tagging rs12208357, p.Arg61Cys) and the association of rs662138 (tagging rs202220802, p.Met420del) for serum isobutyrylcarnitine are independent of each other. In contrast, when causal variants are linked and not inherited independently, the association of one causal variant can be confounded by other causal variants that are linked to it. For instance, we showed that the association of rs594709 ASE variant (tagging rs113569197 splicing variant) for serum isobutyrylcarnitine was greatly affected by rs12208357 (p.Arg61Cys) and rs202220802 (p.Met420del) variants that co-segregate with rs594709. Confounding variants may overestimate or underestimate the association of the linked variant depending on how they are phased and the direction of their association. To accurately assess the association of a given variant in the absence of confounding effects by the linked variants, association testing could be done between haplotypes that differ only by the study variant to determine its association in isolation as exemplified in our study (Figure 3); however, haplotype association test has reduced statistical power compared to single-variant association since many individuals with confounding variants are excluded from the analysis, resulting in smaller sample size.

Our study used allele-specific expression (ASE) analysis to find potential regulatory variants of *SLC22A1*. The advantage of ASE method compared to the conventional expression quantitative locus (eQTL) analysis is that the allele-specific expression is compared within each heterozygous individual; therefore, it avoids confounding effects that stem from differences in genetic and environmental factors. Hepatic *SLC22A1* expression exhibits great inter-individual variation,[21,44] suggesting that it is heavily influenced by genetic background and environmental exposure. While many variants at the *SLC22A1* locus show significant ASE association, they do not have corresponding eQTL association when examined in the Genotype-Tissue Expression (GTEx) database.[31] For example, the rs594709 variant, which has the strongest ASE association ($p = 5.7 \times 10^{-67}$), has only a nominal eQTL association ($p = 1.5 \times 10^{-3}$). This result demonstrates that ASE analysis is a more robust and sensitive method than eQTL analysis for assessing the effects of regulatory variants. In this study, we identified the causal variant of the strongest ASE signal and demonstrated its impact on the *SLC22A1* transcript. Close examination of other variants with ASE association is anticipated to reveal other regulatory variants of *SLC22A1*.

Our results are in agreement with and provide additional insights to the pharmacogenetics of drugs that are transported by SLC22A1, notably metformin and imatinib, common anti-diabetic and chemotherapeutic agents, respectively. Coding variants of SLC22A1, including p.Arg61Cys and p.Met420del variants, are known to impair metformin uptake [39,40] and alter metformin pharmacokinetics[45,46] and are associated with diminished response[39,47] or side effects[48] upon metformin treatment. SLC22A1 coding variants are also associated with the response to imatinib,[49] although it remains unclear whether imatinib is directly transported by SLC22A1.[41] In contrast to coding variants of SLC22A1, the presence and potential impact of regulatory variants of *SLC22A1* on drug efficacy have not been closely explored. Of note, the rs113569197 splice variant was associated with poorer prognosis in patients with chronic myeloid leukemia receiving imatinib treatment,[50] indicating the possibility that the effect of rs113569197 on *SLC22A1* splicing and SLC22A1 protein ultimately led to reduced imatinib response. rs113569197 variant is in strong LD ($r^2 = 0.95$) with the rs628031 coding variant that encodes p.Val408Met (c.1222G>A) substitution in SLC22A1 protein. While p.Val408Met substitution was shown to have no effect on the transport function of SLC22A1,[39,40] rs628031 has been associated with metformin side effects[51,52] and imatinib outcome.[53–55] Our results suggest that rs113569197 may be the true causal variant underlying the observed associations with rs628031.

In conclusion, our study elucidates the molecular mechanisms of the genetic association at the *SLC22A1* locus for serum acylcarnitine levels by functionally validating *SLC22A1* and its variants for hepatic acylcarnitine efflux (Figure 7). While we have found three independent causal variants at the locus in the current cohort, we anticipate that GWASs in larger cohorts with enhanced statistical power might discover additional variants that impact *SLC22A1* through different mechanisms. Our work illustrates that genome-wide association studies, combined with molecular and functional validation, can lead to novel biological insights and better understanding of the impact of natural genetic variation on human physiology.

Supplemental Data

Supplemental Data include six figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2017.08.008.

Acknowledgments

The present study is supported by the grants from NIH RC2HL101864 (D.J.R.) and the Leducq Foundation 10CVD03 (D.J.R.). We thank Merck for the provision of *Slc22a1* conditional knockout mice, Penn Vector Core for the generation of AAVs, Penn CVI histology core for the help in liver histology, and Penn Diabetes Research Center for the use of the Metabolomics Core (P30-DK19525), and members of the Rader lab, especially Dr. Jeffrey T. Billheimer, for helpful discussions. We thank Matthias Arnold (Helmholtz Zentrum München) for his generous help in haplotype analyses. We appreciate Michael D. Gallagher for his kind help and input. We acknowledge Kwanjeong Educational Foundation for their financial support to H.K. and NIH Research Specialist Award (R50CA211437) for W.L.

Received: April 28, 2017
Accepted: August 7, 2017
Published: September 21, 2017

Web Resources

GTEx Portal, http://www.gtexportal.org/home/
HaploReg, http://www.broadinstitute.org/mammals/haploreg/haploreg.php
MIT CRISPR Design Tool, http://crispr.mit.edu/
OMIM, http://www.omim.org/
PLINK, http://www.cog-genomics.org/plink/1.9/

References

1. 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.
2. Genuth, S.M., and Hoppel, C.L. (1981). Acute hormonal effects on carnitine metabolism in thin and obese subjects: responses to somatostatin, glucagon, and insulin. Metabolism *30*, 393–401.
3. Mai, M., Tönjes, A., Kovacs, P., Stumvoll, M., Fiedler, G.M., and Leichtle, A.B. (2013). Serum levels of acylcarnitines are altered in prediabetic conditions. PLoS ONE *8*, e82459.
4. 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.
5. Newgard, C.B., An, J., Bain, J.R., Muehlbauer, M.J., Stevens, R.D., Lien, L.F., Haqq, A.M., Shah, S.H., Arlotto, M., Slentz, C.A., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab. *9*, 311–326.
6. Coker, M., Coker, C., Darcan, S., Can, S., Orbak, Z., and Gökşen, D. (2002). Carnitine metabolism in diabetes mellitus. J. Pediatr. Endocrinol. Metab. *15*, 841–849.
7. Inokuchi, T., Imamura, K., Nomura, K., Nomoto, K., and Isogai, S. (1995). Changes in carnitine metabolism with ketone body production in obese glucose-intolerant patients. Diabetes Res. Clin. Pract. *30*, 1–7.
8. Mihalik, S.J., Michaliszyn, S.F., de las Heras, J., Bacha, F., Lee, S., Chace, D.H., DeJesus, V.R., Vockley, J., and Arslanian, S.A. (2012). Metabolomic profiling of fatty acid and amino acid metabolism in youth with obesity and type 2 diabetes: evidence for enhanced mitochondrial oxidation. Diabetes Care *35*, 605–611.
9. Bene, J., Márton, M., Mohás, M., Bagosi, Z., Bujtor, Z., Oroszlán, T., Gasztonyi, B., Wittmann, I., and Melegh, B. (2013). Similarities in serum acylcarnitine patterns in type 1 and type 2 diabetes mellitus and in metabolic syndrome. Ann. Nutr. Metab. *62*, 80–85.
10. Wang-Sattler, R., Yu, Z., Herder, C., Messias, A.C., Floegel, A., He, Y., Heim, K., Campillos, M., Holzapfel, C., Thorand, B., et al. (2012). Novel biomarkers for pre-diabetes identified by metabolomics. Mol. Syst. Biol. *8*, 615.
11. Villarreal-Pérez, J.Z., Villarreal-Martínez, J.Z., Lavalle-González, F.J., Torres-Sepúlveda, Mdel.R., Ruiz-Herrera, C., Cerda-Flores, R.M., Castillo-García, E.R., Rodríguez-Sánchez, I.P., and Martínez de Villarreal, L.E. (2014). Plasma and urine metabolic profiles are reflective of altered beta-oxidation in non-diabetic obese subjects and patients with type 2 diabetes mellitus. Diabetol. Metab. Syndr. *6*, 129.
12. Butte, N.F., Liu, Y., Zakeri, I.F., Mohney, R.P., Mehta, N., Voruganti, V.S., Göring, H., Cole, S.A., and Comuzzie, A.G. (2015). Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am. J. Clin. Nutr. *102*, 256–267.
13. Seiler, S.E., Koves, T.R., Gooding, J.R., Wong, K.E., Stevens, R.D., Ilkayeva, O.R., Wittmann, A.H., DeBalsi, K.L., Davies, M.N., Lindeboom, L., et al. (2015). Carnitine acetyltransferase mitigates metabolic inertia and muscle fatigue during exercise. Cell Metab. *22*, 65–76.
14. Soni, M.S., Rabaglia, M.E., Bhatnagar, S., Shang, J., Ilkayeva, O., Mynatt, R., Zhou, Y.P., Schadt, E.E., Thornberry, N.A., Muoio, D.M., et al. (2014). Downregulation of carnitine acyl-carnitine translocase by miRNAs 132 and 212 amplifies glucose-stimulated insulin secretion. Diabetes *63*, 3805–3814.
15. Aguer, C., McCoin, C.S., Knotts, T.A., Thrush, A.B., Ono-Moore, K., McPherson, R., Dent, R., Hwang, D.H., Adams, S.H., and Harper, M.E. (2015). Acylcarnitines: potential

implications for skeletal muscle insulin resistance. FASEB J. *29*, 336–345.
16. Sampey, B.P., Freemerman, A.J., Zhang, J., Kuan, P.F., Galanko, J.A., O'Connell, T.M., Ilkayeva, O.R., Muehlbauer, M.J., Stevens, R.D., Newgard, C.B., et al. (2012). Metabolomic profiling reveals mitochondrial-derived lipid biomarkers that drive obesity-associated inflammation. PLoS ONE *7*, e38812.
17. McCoin, C.S., Knotts, T.A., Ono-Moore, K.D., Oort, P.J., and Adams, S.H. (2015). Long-chain acylcarnitines activate cell stress and myokine release in C2C12 myotubes: calcium-dependent and -independent effects. Am. J. Physiol. Endocrinol. Metab. *308*, E990–E1000.
18. Rutkowsky, J.M., Knotts, T.A., Ono-Moore, K.D., McCoin, C.S., Huang, S., Schneider, D., Singh, S., Adams, S.H., and Hwang, D.H. (2014). Acylcarnitines activate proinflammatory signaling pathways. Am. J. Physiol. Endocrinol. Metab. *306*, E1378–E1387.
19. 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.
20. Shin, S.Y., Fauman, E.B., Petersen, A.K., Krumsiek, J., Santos, R., Huang, J., Arnold, M., Erte, I., Forgetta, V., Yang, T.P., et al.; Multiple Tissue Human Expression Resource (MuTHER) Consortium (2014). An atlas of genetic influences on human blood metabolites. Nat. Genet. *46*, 543–550.
21. Nies, A.T., Koepsell, H., Winter, S., Burk, O., Klein, K., Kerb, R., Zanger, U.M., Keppler, D., Schwab, M., and Schaeffeler, E. (2009). Expression of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) is affected by genetic factors and cholestasis in human liver. Hepatology *50*, 1227–1240.
22. Lozano, E., Herraez, E., Briz, O., Robledo, V.S., Hernandez-Iglesias, J., Gonzalez-Hernandez, A., and Marin, J.J. (2013). Role of the plasma membrane transporter of organic cations OCT1 and its genetic variants in modern liver pharmacology. BioMed Res. Int. *2013*, 692071.
23. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R.; and 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature *526*, 68–74.
24. Delaneau, O., Marchini, J.; and 1000 Genomes Project Consortium (2014). Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat. Commun. *5*, 3934.
25. Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. *5*, e1000529.
26. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
27. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Weedon, M.N., Loos, R.J., et al.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; and DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat. Genet. *44*, 369–375, S1–S3.
28. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics *26*, 2336–2337.
29. Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics *21*, 263–265.
30. Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. *40*, D930–D934.
31. Consortium, G.T.; and GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.
32. Innocenti, F., Cooper, G.M., Stanaway, I.B., Gamazon, E.R., Smith, J.D., Mirkov, S., Ramirez, J., Liu, W., Lin, Y.S., Moloney, C., et al. (2011). Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. PLoS Genet. *7*, e1002078.
33. 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.
34. McVicker, G., van de Geijn, B., Degner, J.F., Cain, C.E., Banovich, N.E., Raj, A., Lewellen, N., Myrthil, M., Gilad, Y., and Pritchard, J.K. (2013). Identification of genetic variants that affect histone modifications in human cells. Science *342*, 747–749.
35. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.
36. Seiler, C.Y., Park, J.G., Sharma, A., Hunter, P., Surapaneni, P., Sedillo, C., Field, J., Algar, R., Price, A., Steel, J., et al. (2014). DNASU plasmid and PSI:Biology-Materials repositories: resources to accelerate biological research. Nucleic Acids Res. *42*, D1253–D1260.
37. Tamai, I., Ohashi, R., Nezu, J., Yabuuchi, H., Oku, A., Shimane, M., Sai, Y., and Tsuji, A. (1998). Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2. J. Biol. Chem. *273*, 20378–20382.
38. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., et al.; International HapMap Consortium (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851–861.
39. Shu, Y., Sheardown, S.A., Brown, C., Owen, R.P., Zhang, S., Castro, R.A., Ianculescu, A.G., Yue, L., Lo, J.C., Burchard, E.G., et al. (2007). Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. J. Clin. Invest. *117*, 1422–1431.
40. Ahlin, G., Chen, L., Lazorova, L., Chen, Y., Ianculescu, A.G., Davis, R.L., Giacomini, K.M., and Artursson, P. (2011). Genotype-dependent effects of inhibitors of the organic cation transporter, OCT1: predictions of metformin interactions. Pharmacogenomics J. *11*, 400–411.
41. Nies, A.T., Schaeffeler, E., van der Kuip, H., Cascorbi, I., Bruhn, O., Kneba, M., Pott, C., Hofmann, U., Volk, C., Hu, S., et al. (2014). Cellular uptake of imatinib into leukemic cells is

independent of human organic cation transporter 1 (OCT1). Clin. Cancer Res. *20*, 985–994.
42. Seitz, T., Stalmann, R., Dalila, N., Chen, J., Pojar, S., Dos Santos Pereira, J.N., Krätzner, R., Brockmöller, J., and Tzvetkov, M.V. (2015). Global genetic analyses reveal strong inter-ethnic variability in the loss of activity of the organic cation transporter OCT1. Genome Med. *7*, 56.
43. Schooneman, M.G., Ten Have, G.A., van Vlies, N., Houten, S.M., Deutz, N.E., and Soeters, M.R. (2015). Transorgan fluxes in a porcine model reveal a central role for liver in acylcarnitine metabolism. Am. J. Physiol. Endocrinol. Metab. *309*, E256–E264.
44. Denk, G.U., Soroka, C.J., Mennone, A., Koepsell, H., Beuers, U., and Boyer, J.L. (2004). Down-regulation of the organic cation transporter 1 of rat liver in obstructive cholestasis. Hepatology *39*, 1382–1389.
45. Shu, Y., Brown, C., Castro, R.A., Shi, R.J., Lin, E.T., Owen, R.P., Sheardown, S.A., Yue, L., Burchard, E.G., Brett, C.M., and Giacomini, K.M. (2008). Effect of genetic variation in the organic cation transporter 1, OCT1, on metformin pharmacokinetics. Clin. Pharmacol. Ther. *83*, 273–280.
46. Tzvetkov, M.V., Vormfelde, S.V., Balen, D., Meineke, I., Schmidt, T., Sehrt, D., Sabolić, I., Koepsell, H., and Brockmöller, J. (2009). The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. Clin. Pharmacol. Ther. *86*, 299–306.
47. Christensen, M.M., Brasch-Andersen, C., Green, H., Nielsen, F., Damkier, P., Beck-Nielsen, H., and Brosen, K. (2011). The pharmacogenetics of metformin and its impact on plasma metformin steady-state levels and glycosylated hemoglobin A1c. Pharmacogenet. Genomics *21*, 837–850.
48. Dujic, T., Causevic, A., Bego, T., Malenica, M., Velija-Asimi, Z., Pearson, E.R., and Semiz, S. (2016). Organic cation transporter 1 variants and gastrointestinal side effects of metformin in patients with type 2 diabetes. Diabet. Med. *33*, 511–514.
49. Giannoudis, A., Wang, L., Jorgensen, A.L., Xinarianos, G., Davies, A., Pushpakom, S., Liloglou, T., Zhang, J.E., Austin, G., Holyoake, T.L., et al. (2013). The hOCT1 SNPs M420del and M408V alter imatinib uptake and M420del modifies clinical outcome in imatinib-treated chronic myeloid leukemia. Blood *121*, 628–637.
50. Grinfeld, J., Gerrard, G., Alikian, M., Alonso-Dominguez, J., Ale, S., Valgañon, M., Nteliopoulos, G., White, D., Marin, D., Hedgley, C., et al. (2013). A common novel splice variant of SLC22A1 (OCT1) is associated with impaired responses to imatinib in patients with chronic myeloid leukaemia. Br. J. Haematol. *163*, 631–639.
51. Tarasova, L., Kalnina, I., Geldnere, K., Bumbure, A., Ritenberga, R., Nikitina-Zake, L., Fridmanis, D., Vaivade, I., Pirags, V., and Klovins, J. (2012). Association of genetic variation in the organic cation transporters OCT1, OCT2 and multidrug and toxin extrusion 1 transporter protein genes with the gastrointestinal side effects and lower BMI in metformin-treated type 2 diabetes patients. Pharmacogenet. Genomics *22*, 659–666.
52. Klen, J., Goričar, K., Janež, A., and Dolžan, V. (2014). The role of genetic factors and kidney and liver function in glycemic control in type 2 diabetes patients on long-term metformin and sulphonylurea cotreatment. BioMed Res. Int. *2014*, 934729.
53. Koren-Michowitz, M., Buzaglo, Z., Ribakovsky, E., Schwarz, M., Pessach, I., Shimoni, A., Beider, K., Amariglio, N., le Coutre, P., and Nagler, A. (2014). OCT1 genetic variants are associated with long term outcomes in imatinib treated chronic myeloid leukemia patients. Eur. J. Haematol. *92*, 283–288.
54. Takahashi, N., Miura, M., Scott, S.A., Kagaya, H., Kameoka, Y., Tagawa, H., Saitoh, H., Fujishima, N., Yoshioka, T., Hirokawa, M., and Sawada, K. (2010). Influence of CYP3A5 and drug transporter polymorphisms on imatinib trough concentration and clinical response among patients with chronic phase chronic myeloid leukemia. J. Hum. Genet. *55*, 731–737.
55. Vaidya, S., Ghosh, K., Shanmukhaiah, C., and Vundinti, B.R. (2015). Genetic variations of hOCT1 gene and CYP3A4/A5 genes and their association with imatinib response in chronic myeloid leukemia. Eur. J. Pharmacol. *765*, 124–130.