9-2019

Rare DEGS1 variant significantly alters de novo ceramide synthesis pathway

Nicholas B. Blackburn
The University of Texas Rio Grande Valley, nicholas.blackburn@utrgv.edu

Joanne E. Curran
The University of Texas Rio Grande Valley, joanne.curran@utrgv.edu

Laura F. Michael
Eli Lilly

Peter J. Meikle
Baker Heart and Diabetes Institute

Juan M. Peralta
The University of Texas Rio Grande Valley

See next page for additional authors

Authors
Nicholas B. Blackburn, Joanne E. Curran, Laura F. Michael, Peter J. Meikle, Juan M. Peralta, Marian Mosio, Scott McAhern, Hai H. Bui, Melissa A. Bellinger, Corey Giles, Satish Kumar, Ana C. Leandro, Marcio Alameida, Jacquelyn M. Weir, Michael C. Mahaney, Thomas D. Dyer, Laura Almasy, John L. VandeBerg, Sarah Williams-Blangero, David C. Glahn, Ravindranath Duggirala, Mark Kowola, and John Blangero

Rare *DEGS1* variant significantly alters *de novo* ceramide synthesis pathway

Nicholas B. Blackburn,[1] Laura F. Michael,[2] Peter J. Meikle,[3] Juan M. Peralta,[1,4] Marian Mosior,[2] Scott McAhren,[2] Hai H. Bui,[2] Melissa A. Bellinger,[2] Corey Giles,[3] Satish Kumar,[1] Ana C. Leandro,[1] Marcio Almeida,[1] Jacquelyn M. Weir,[3] Michael C. Mahaney,[1] Thomas D. Dyer,[1] Laura Almasy,[5] John L. VandeBerg,[1] Sarah Williams-Blangero,[1] David C. Glahn,[6,7] Ravindranath Duggirala,[1] Mark Kowala,[2] John Blangero,[1*] Joanne E. Curran,[1*]

1. South Texas Diabetes and Obesity Institute and Department of Human Genetics, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX, USA.
2. Eli Lilly and Company, Lilly Research Laboratories, Indianapolis, IN, 46225, USA.
3. Baker Heart and Diabetes Institute, Melbourne, VIC, 3004, Australia.
4. Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, 7000, Australia
5. Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia and Department of Human Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA, 19104, USA.
6. Department of Psychiatry, Boston Children's Hospital and Harvard Medical School, Boston, MA, 02115, USA
7. Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, Hartford, CT, 06106, USA.

*These authors contributed equally to this work.

To whom correspondence should be addressed:

Dr. Nicholas B. Blackburn, Ph.D.

South Texas Diabetes and Obesity Institute and Department of Human Genetics, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX, USA.

Phone: +1 956 882 7523

Email: nicholas.blackburn@utrgv.edu

Dr. Joanne E. Curran, Ph.D.

South Texas Diabetes and Obesity Institute and Department of Human Genetics, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX, USA.

Phone: +1 956 882 7532

Email: joanne.curran@utrgv.edu

Running title

Rare variant in *DEGS1* affects ceramide synthesis pathway

Abbreviations

1KGP: 1000 genomes project; AMR: Admixed American; AFR: African; CADD: Combined Annotation Dependent Depletion; DEGS1: delta 4-desaturase, sphingolipid 1; EUR: European; SAFHS: San Antonio Family Heart Study; SDU: standard deviation unit; SNV: single nucleotide variant; T2D: type 2 diabetes; VCF: variant call format; WGS: whole genome sequencing

Abstract

The *de novo* ceramide synthesis pathway is essential to human biology and health but genetic influences remain unexplored. The core function of this pathway is the generation of biologically active ceramide from its precursor, dihydroceramide. Dihydroceramides have diverse, often protective, biological roles; conversely, increased ceramide levels are biomarkers of complex disease. To explore the genetics of the ceramide synthesis pathway, we searched for deleterious nonsynonymous variants in the genomes of 1,020 Mexican Americans from extended pedigrees. We identified a Hispanic ancestry−specific rare functional variant, L175Q, in DEGS1, a key enzyme in the pathway that converts dihydroceramide to ceramide. This amino acid change was significantly associated with large increases in plasma dihydroceramides. Indexes of DEGS1 enzymatic activity were dramatically reduced in heterozygotes. CRISPR/Cas9 genome editing of HepG2 cells confirmed that the L175Q variant results in a partial loss of function for the DEGS1 enzyme. Understanding the biological role of *DEGS1* variants, such as L175Q, in ceramide synthesis may improve the understanding of metabolic-related disorders, and spur ongoing research of drug targets along this pathway.

Keywords: Lipidomics, Ceramides, Sphingolipids, Genomics, Genetics

Introduction

The classical lipid parameters (such as high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), triglycerides (TG) and total cholesterol) that are most commonly examined in relation to disease risk are themselves complex entities composed of multiple lipid and protein components. The lipidome, defined as the total lipid complement of a given biological system, contains many thousands of these individual lipid species and represents a wealth of phenotypes that might serve as better predictors of disease than more complex lipoproteins. The biologically simpler nature of individual lipid species suggests that their phenotypic variation lies closer to the genomic level and ultimately the causal action of genes, making them valuable endophenotypes (1, 2). Our studies have shown that specific lipoproteins and indeed their constituent lipids are as heritable as the classic lipoprotein measures that have dominated disease epidemiology and are important factors in the development of disease (3-9). The field of lipidomics has received increasing attention in recent years with several studies highlighting the relationships between individual lipid species and health outcomes in both disease based cohorts and in healthy population cohort studies (10, 11). In particular, circulating levels of particular ceramide species are being considered as clinical biomarkers for cardiometabolic disease (12).

Ceramides are a family of lipid molecules that consist of a sphingosine covalently linked to a fatty acid acyl chain, with species differing by length, hydroxylation, and saturation of the sphingoid base and fatty acid moieties. Ceramide production occurs through three major pathways: *de novo* ceramide synthesis, sphingomyelin synthesis and the ceramide salvage pathway. Both the sphingomyelin synthesis and ceramide salvage pathways require ceramide initially for activation; with the sphingomyelin pathway generating ceramide through the hydrolysis of sphingomyelin, and the salvage pathway through the catabolism of other sphingolipids. In contrast, the *de novo* synthesis pathway depends on the availability of palmitate and serine to synthesize dihydroceramide (which is converted to ceramide) through a series of enzymatic reactions (13).

Levels of individual dihydroceramide and ceramide species, and ratios between species, have been implicated in cardiovascular and metabolic disease. First thought to be biologically inert dihydroceramides have now been shown to have diverse roles in autophagy, apoptosis, infection and metabolic disease, with increases in dihydroceramide levels generally protective (reviewed in Siddique et al. (2015) (14)), however Wigger et al. have shown in two human cohorts that increased dihydroceramide levels are associated with future development of type 2 diabetes (T2D), challenging the protective role assigned to dihydroceramides (15). Evidence through previous work from our group also challenges the protective role of dihydroceramides. We developed a lipidomic risk score that identified, with two other lipid species, increased levels of the dihydroceramide species Cer(d18:0/18:0) as a predictive factor for future T2D development (9). The addition of a double bond to the chemical structure of dihydroceramides, by the enzyme DEGS1, produces ceramides. In contrast to dihydroceramides, increases in levels of ceramides is repeatedly shown to be linked to cardiovascular disease, particular ceramide species Cer(d18:1/16:0), Cer(d18:1/20:0) and Cer(d18:1/24:1) (16, 17). As there is a clear link to disease demonstrated in both disease based and healthy cohorts a new avenue to understand the relationship between ceramide biology and disease is to identify genetic variation underlying variation in levels of individual ceramide biology lipid species.

Here, in the first human genomic study of ceramide synthesis, using targeted lipidomic profiles and whole genome sequence (WGS) in 1,020 Mexican Americans from 46 large extended pedigrees of the San Antonio Family Heart Study (SAFHS), we searched for nonsynonymous variants, predicted to be deleterious, influencing ceramides and related lipid species. Within our lipidomic profiles, (previously described (5-9)) we have measures of six dihydroceramide species and six ceramide species. We use these 12 lipid species as a gateway to investigate the genetics of the ceramide synthesis pathway.

Materials and Methods

Study Participants

The San Antonio Family Heart Study (SAFHS) began in 1991 and was initially designed to investigate the genetics of CVD in Mexican American individuals. The SAFHS enrolled large, extended Mexican American families residing in San Antonio and ascertainment occurred by way of the random selection of an adult Mexican American proband, without regard to disease. The enrollment procedures, inclusion and exclusion criteria, and phenotypic assessments of the study participants have been described in detail previously (18, 19). This is an ongoing investigation which has had four phases of data collection over a 25 year period. The data and samples used in this study were collected during the first phase of data collection, occurring between 1992 and 1996. For the 1,020 SAFHS participants specifically used in this study 61.8% were female, and the sample was aged between 15 and 94 (mean=39.8). Informed consent was obtained from all participants before collection of samples. The study conformed to the Declaration of Helsinki and the Institutional Review Boards of the University of Texas Health Sciences Center at San Antonio and the University of Texas Rio Grande Valley have approved this study.

Lipidomic Profiling

Plasma concentrations of 319 lipid species (representing 23 lipid classes) were measured in 1,212 participants from the first visit of SAFHS using electrospray ionization-tandem mass spectrometry (ESI-MS/MS) quantification. Lipidomic profiling was performed in the Metabolomics Laboratory, Baker Heart and Diabetes Institute, Melbourne, Australia. The experimental protocols used have been described elsewhere (5, 20). In the current study 1,020 individuals have both lipidome profiles and whole genome sequence data.

Genome Sequencing and SNP array genotyping

In the currently available data from the original San Antonio Family Heart Study, 1,514 individuals have whole genome sequence available for analysis. Whole genome sequencing for 496 individuals in this study was performed by Complete Genomics at 50x coverage (21) as part of the T2D-GENES consortium. Additionally, WGS of another 1,018 individuals at 30x coverage was obtained through Illumina, 406 through the Illumina CASAVA pipeline and 612 through the Illumina Isaac pipeline, all sequences were aligned to the hg19 reference genome. For samples sequenced by Complete Genomics, genotypes were available per individual in the Complete Genomics proprietary genotype format. Using Complete Genomics Analysis Tools 1.8.0 genotypes were converted to VCF format. For samples sequenced by Illumina, genotypes were available for each individual in VCF format. Samples were first merged into three multi-sample VCF files corresponding to their major platform (Complete Genomics, Illumina CASAVA, Illumina ISAAC). Metadata associated with each sample/variant differed across each platform and was incompatible for merging VCF files between platforms, accordingly each platform VCF file was converted to PLINK format and then combined in PLINK (version 1.90b3m) (22, 23). Variants were assumed to be homozygous reference in samples that did not have a genotype call in the sample VCF file for that variant. Variants were restricted to biallelic autosomal SNVs. Indel variation was excluded from analysis on the basis of high variability in calling and nomenclature across platforms causing incompatibilities in the merging of indels. In total 37,819,860 autosomal SNVs were identified and 19,247,876 (51%) were present in the entire dataset with five or more allele copies. Samples sequenced by Illumina were also genotyped by Illumina on Omni2.5 microarrays. Individual level data was merged using GenomeStudio v2011.1, Genotyping module v1.9.4 to combine each version of the Omni2.5 microarrays. PLINK (22, 23) was then used to merge variants across array versions and subset to autosomal variants only. The final set of Omni2.5 array variants was then extracted from the Complete Genomics WGS to create a harmonized set of 1,146,843 Omni2.5 array based variants.

Identification of deleterious nonsynonymous SNVs

Using ANNOVAR (24) to annotate biallelic SNVs with RefSeq annotations (25) 72,694 nonsynonymous, stop gain or stop loss autosomal SNVs were identified. Of these 36,262 variants had CADD (version 1.3) deleteriousness prediction scores ≥ 15 (26). Considering only the subset of 1,020 WGS individuals with lipidome data used in this research, there were 28,067 SNVs present with 5 or more allele copies that were not monomorphic. These 28,067 nonsynonymous potential deleterious variants were scored as the number of minor allele copies for use in measured genotype association testing in SOLAR.

Admixture calculation

The Omni2.5 array variant set was used to calculate population admixture in our cohort using ADMIXTURE (27) (version 1.3.0). Variants from our cohort were overlapped with samples from phase 3 of the 1000 Genomes Project (1KGP) in the EUR, AFR and AMR super populations, creating a merged sample set with 1,038,229 variants. Using known pedigree relationships in our SAFHS cohort we subset the sample to contain only SAFHS founders and unrelated 1KGP samples. Using PLINK, variants were pruned for nominal Hardy-Weinberg equilibrium ($P<0.05$) and LD pruned (--indep-pairwise 50 10 0.1) to a set of 153,388 variants. ADMIXTURE with K=3, under the assumption of three predominant ancestral population groups was conducted on this sample set. Admixture components for the remaining SAFHS samples were projected onto the population structure learnt from unsupervised clustering of the 1KGP and SAFHS founders. This resulted in three clear population groups, corresponding to Caucasian, African and Hispanic admixture. The African and Hispanic components were used in all analyses as covariates to control for population substructure.

Empirical kinship coefficient estimation

For empirical kinship calculation, 100,000 variants from the Omni2.5 array variant set were selected at random using PLINK. Empirical kinship coefficients were estimated from this variant set following a similar approach to that used by Peralta et al. (28), using the IBDLD method (29) (version 3.33). In the derivation of an empirical genetic relationship matrix the propagation of inherent measurement

uncertainties into the kinship approximations can led to singular genetic relationship matrices (a brief list of other possible causes can be found in VanRaden (2007) (30)). We used our own implementation of the alternating projections method described by Higham (2002) (31) to derive the nearest positive semidefinite empirical genetic relationship matrix. This procedure yielded genetic relationship matrices that had the expected covariance matrix properties.

Statistical analyses

All statistical analyses were performed in SOLAR (32) (version 8.1.1) and R version 3.4.3. Heritability estimates and variant association testing of inverse normalized lipid measurements were calculated in the variance components framework of SOLAR including in each analyses the covariates of age, age^2, sex, age×sex, Hispanic admixture and African admixture. The non-independence between individuals due to their genetic relatedness was accounted for in analysis in SOLAR by use of the calculated empirical kinship coefficients. Likelihood ratio tests were obtained for genetic variant associations.

siRNA transfection of HepG2 cells

HepG2 cells were maintained on 3/1 Dulbecco's Modified Eagle Medium/Nutrient Mixture F12 media (Invitrogen, Carlsbad CA) + 10% fetal bovine serum. One day prior to transfection, the cells were trypsinized, counted in a Coulter Z1 counter (Beckman Coulter, Brea CA) and seeded at a density of 1×10^6 cells/well in 6 well polyD lysine plates (BD Biocoat 354413). Human *DEGS1* ON-TARGETplus SMART pool (Thermo Scientific Dharmacon L-006675-00), Non-Targeting ON-TARGETplus Control pool (Thermo Scientific Dharmacon D-001810-05), and Human *GAPDH* ON-TARGETplus Control pool (Thermo Scientific Dharmcon T-2004-03) were transfected into cells at a final concentration of 200 nM using Dharmafect 4 Transfection Reagent. After various hours post-transfection (as indicated), cells were collected by scraping into phosphate buffered saline (PBS) and pelleted by centrifugation at 1400 rpm for 10min; PBS was aspirated and the cell pellets were frozen at -80C. Cell pellets were analyzed for DEGS1 protein by western blot.

Western blot analysis

HepG2 cells from siRNA *DEGS1* knockdown and non-targeting siRNA control treatments were harvested for western blot analysis. Protein analysis was performed on cellular lysates utilizing a Pierce BCA protein kit (Thermo Fisher Scientific, San Jose CA). Abcam rabbit anti-Degs1 antibody (ab80654) was used as primary antibody, and Alexa-Flur 680 anti-Rabbit IgG was used as secondary antibody (Abcam, Cambridge MA). Protein expression was visualized using LI-COR infra-red imaging system (LI-COR, Lincoln NE).

CRISPR/Cas9 plasmid construction

DEGS1(523-541bp)-pCRISPR-OFP construct was generated by inserting the double-stranded oligonucleotide, which was created by annealing forward strand oligonucleotide (GTTTGGGGTT-GATGAACAGGTTTT) and reverse strand oligonucleotide (CTGTTCATCAACCCCAAACCGGTG), into GeneArt™ CRISPR Nuclease Vector pCRISPR-OFP according to the manufacture's instruction (Thermo Fisher Scientific, San Jose CA). *DEGS1*(L175Q) donor-pUC57 construct was created through inserting synthetic DNA containing *DEGS1*(L175Q) donor sequence into pUC57 vector by GenScript. Codon modification of the *DEGS1*(L175Q) donor sequence was completed to create resistance to the *DEGS1* targeting CRISPR plasmid. Both constructs were confirmed by DNA sequencing.

CRISPR/Cas9 transfection of HepG2 cells

HepG2 cells were maintained on 3/1 (DMEM/F12) media (Invitrogen, Carlsbad CA) + 10% FBS. One day prior to transfection, the cells were trypsinized and passed through a 2µM filter basket. The cells were counted in a Coulter Z1 counter (Beckman Coulter, Brea CA) and seeded at a density of 5×10^6 cells/10 cm dish in growth media. After overnight incubation the cells were transfected using 6.0µg *DEGS1*(523-541bp)-pCRISPR-OFP and 6.0µg *DEGS1*(L175Q) donor-pUC57 and Lipofectamine 2000 according to manufacturer's specifications (Invitrogen, Carlsbad CA). The following day, the cells were trypsinized and passed through a 2µM mesh. The cells were sorted for expression of OFP and 10 x 96 well plates were

seeded at 1 cell/well. Forty clones were grown, and twenty were expanded for sanger sequencing confirmation of L175Q target alteration. Of these twenty, four clones with successful integration were selected for lipidomic analysis, ten replicates per clone. One additional clone for which the L175Q alteration was not successful was selected as the 'mock' control for this experiment with ten replicates analyzed.

Cell lipidomics analysis

HepG2 cells were trypsinized and washed 3X in PBS. After counting, the cell density was adjusted to $2x10^6$ cells/ml in 2.0mls of PBS. The cells were pelleted at 1500 X g for 10 minutes at 40C. The PBS was aspirated and the cell pellet frozen at -80°C until mass spectrometry analysis. LC/ESI/MS/MS analysis of dihydroceramides and ceramides was performed using a TSQ Quantum Ultra-triple quadrupole mass spectrometer (Thermo Fisher, San Jose CA) interfaced with an Agilent 1100 HPLC (Agilent Technologies, Wilmington, DE) and an Xbridge C8 column (2.1 × 30 mm; Waters, Milford, MA). Lipids from serum were extracted using one phase extraction (methanol-dichloromethane) with internal standards. Quantification was performed using the ratio of analyte to internal standards relative to sphingolipid calibration curves.

Results

Variation in ceramide synthesis pathway lipid levels is heritable

Using variance component modelling in SOLAR (32) we calculated the heritability of each of these dihydroceramide and ceramide species, accounting for age, sex, their interactions and measures of Amerindian and African ancestry to control for population substructure. Previously we have used known pedigrees to construct genetic relationship matrices for use in SOLAR to appropriately handle the non-independence of related samples. Here we instead estimate empirical kinship coefficients using the IBDLD method (identity by descent linkage disequilibrium) between all pairs of individuals to identify distant or unknown connections between individuals in the study (29). As previously shown with pedigree-based kinship in our complete lipidome sample set of 1,212, (5) here in the subset of 1,020 with WGS (and thus empirical kinship estimates) currently available in our cohort we confirm that all 12 lipid species are significantly heritable indicating a genetic component to the variation observed (Table S1), prompting further investigation of the underlying genetic contribution to this variation.

A rare potentially functional amino acid change variant, DEGS1 L175Q, is associated with dihydroceramide levels

To detect potentially functional genetic variation we focused our search on 28,067 missense and nonsense mutations predicted to be highly deleterious using a CADD (26) phred-scaled score threshold of ≥ 15 and which occurred 5 or more times in our sample. Using measured genotype association testing in SOLAR (32) we identified a highly significant association ($p=3.33\times10^{-13}$) with a rare variant (L175Q, rs191144864, chr1:224377720 T>A) in the *DEGS1* gene on chromosome 1 influencing the dihydroceramide species Cer(d18:0/24:0). This variant is associated with a statistically and biologically significant increase (β_{SNP} = 1.35 standard deviation units (SDU)) of the mean in heterozygotes. Figure 1 shows the genome-wide results for the Cer(d18:0/24:0) associations with these functional variants, adjusted for age, sex, their interactions

and Amerindian and African admixture components. Here we have captured enough copies (n = 28 heterozygotes) of this globally rare variant to achieve a significant association. In our cohort the L175Q variant shows additional genome-wide significant associations with increases in other dihydroceramide species including Cer(d18:0/22:0) levels ($p = 3.13\times10^{-9}$, $\beta_{SNP} = 1.10$ SDU), Cer(d18:0/24:1) levels ($p = 1.77\times10^{-7}$, $\beta_{SNP} = 0.97$ SDU) and total dihydroceramide levels ($p = 6.66\times10^{-10}$, $\beta_{SNP} = 1.14$ SDU) (Table 1, see Figures S2-S4 for Manhattan plots). Nominal L175Q associations, not genome-wide significant, with increases in the smaller chain dihydroceramides Cer(d18:0/16:0) ($p = 8.38\times10^{-4}$, $\beta_{SNP} = 0.61$ SDU), Cer(d18:0/18:0) ($p = 0.007$, $\beta_{SNP} = 0.50$ SDU), Cer(d18:0/20:0) ($p = 6.81\times10^{-4}$, $\beta_{SNP} = 0.64$ SDU) and decreased smaller chain ceramide Cer(d18:1/16:0) ($p = 0.039$, $\beta_{SNP} = -0.36$ SDU) were also detected (Table 1).

The DEGS1 L175Q variant has its strongest influence in the de novo synthesis of ceramides pathway

With the clear effect identified in dihydroceramide levels we accordingly hypothesized that variants affecting the function of DEGS1, such as the L175Q variant, would show associations with multiple lipid species across pathways containing ceramide. As dihydroceramide is an essential component of the *de novo* ceramide synthesis pathway, we posited that we would observe associations between the L175Q variant and levels of lipid species in pathways containing ceramide, including the sphingomyelin pathways and the glycosphingolipid pathway in which ceramide is broken down. Figure 2 shows the effect size of L175Q on each lipid species from our data in these three ceramide biology pathways. Overall the effect of L175Q is most influential in the *de novo* ceramide synthesis pathway, but also shows nominal associations with decreases in total dihexosylceramides, dihexosylceramide 16:0, trihexosylceramide 16:0 and GM_3 ganglioside 18:0 in the glycosphingolipid pathway and sphingomyelin 34:1 in the sphingomyelin pathway. These findings are detailed in Table S2.

The DEGS1 L175Q variant affects disease related ceramide ratios and biological indexes of DEGS1 activity

We then examined the L175Q variant in the context of established cardiovascular disease related ceramide ratios Cer(d18:1/16:0) to Cer(d18:1/24:0), Cer(d18:1/20:0) to Cer(d18:1/24:0), Cer(d18:1/24:1) to Cer(d18:1/24:0) (16, 17) and possible biological indexes of DEGS1 activity by calculating the ratios of total ceramide to total dihydroceramide, Cer(d18:1/24:0) to Cer(d18:0/24:0) and apoB to Cer(d18:0/24:0). Ratios of ceramide to dihydroceramide are an appropriate index of DEGS1 enzymatic activity because of the known role of the DEGS1 enzyme in converting dihydroceramide to ceramide. Similarly, our choice of plasma apoB as a relevant numerator for a DEGS1 activity index was driven by a desire to introduce an independent assay method into the index and the fact that most ceramides are found on apoB containing lipoproteins (33). Each ceramide ratio and DEGS1 activity index showed significant associations with the L175Q variant (Table 2). For the ceramide ratios, the L175Q variant is modestly associated with decreases

in ceramide levels. For the DEGS1 activity indexes, each ratio showed a highly significant association with the L175Q variant (Table 2). Corresponding to a decrease in DEGS1 enzymatic activity, biologically large decreases were observed in heterozygotes for each index (β = -1.68 SDU for apoB/dihydroceramide, β = -1.41 SDU for the Cer(d18:1/24:0)/Cer(d18:0/24:0) ratio, β = -1.28 for the total ceramide : total dihydroceramide ratio) further supporting functional effect of L175Q. A decrease in DEGS1 function, as a result of L175Q, causes a buildup of dihydroceramide and a decrease in ceramide production. Figure 3 shows the increase in the dihydroceramide Cer(d18:0/24:0) and the decreases in DEGS1 activity indexes in carriers of the L175Q variant.

Genetic variants in DEGS1 broadly impact ceramide pathway biology

Further, we assessed all observed single nucleotide variants (SNVs) in the *DEGS1* gene for association with lipid species in ceramide pathways, including levels of dihydroceramides, ceramides, monohexosylceramide, dihexosylceramide, trihexosylceramide, GM_3 gangliosides and sphingomyelins, with the L175Q variant included as a covariate in the analysis. In total there were 101 SNVs identified in the *DEGS1* gene in this cohort. Measured genotype association analysis in SOLAR identified two other *DEGS1* variants associated with increases in levels of lipids from the *de novo* ceramide synthesis and sphingomyelin pathways each surpassing the Bonferroni p-value threshold for these 101 variants. Results showed that an intronic variant, chr1:224379287 A>C rs965873262, with 6 heterozygotes seen in this cohort, was associated with large increases in dihydroceramides (Cer(d18:0/24:0), $p = 1.35\times10^{-6}$, $\beta_{SNP} = 2.01$), while a synonymous variant, chr1:224377955 C>A T253T rs759023173, with 8 heterozygotes in this cohort, was associated with large increases in sphingomyelins (SM(32:2), $p = 6.10\times10^{-5}$, $\beta_{SNP} = 1.22$) (Table 3). The functionality of these variants is unknown however they did not occur in individuals with the L175Q variant and this indicates that variation in *DEGS1* more broadly impacts ceramide pathway biology and therefore warrants further investigation and replication in other cohorts.

The three *DEGS1* variants that show associations with the *de novo* ceramide synthesis pathway in this cohort are each rare variants in publicly available population frequency databases with two (L175Q rs191144864 and T253T rs759023173) of likely Hispanic ancestry origin. For the L175Q variant, the largest frequency database, gnomAD (v2.1.1) (34, 35), contains 382 copies of the L175Q variant in 17,720 individuals (Latino MAF=0.01087) from the Latino ancestry population and 3 heterozygotes in the 'Other' ancestry population. This variant is not detected in any other gnomAD population, with the overall total MAF in gnomAD = 0.001361. Here in our Mexican American family cohort we detect 28 copies of this variant in 1,020 whole genome sequences. For the intronic chr1:224379287 A>C rs965873262 *DEGS1* variant we detect 6 heterozygotes in our cohort while only 4 heterozygotes are observed in 13,388 individuals in gnomAD, all of whom are of African ancestry, however as this variant is intronic it is not well measured by the gnomAD Latino population samples which are primarily from whole exome sequence data. Finally, the synonymous chr1:224377955 C>A rs759023173 *DEGS1* variant, for which there are 8 heterozygotes in our cohort, also appears to be Hispanic ancestry-specific with 38 heterozygotes occurring in 10,155 Latino individuals from gnomAD (MAF=0.001871). These three variants are independent in our cohort, the carriers do not also have a copy of any of the other variants.

In vitro confirmation of siRNA knockdown effects of DEGS1 on lipidomics in HepG2 cells

To confirm *in vitro* that an impairment in the function of DEGS1 affects DEGS1 enzymatic activity, as measured through the total ceramide : total dihydroceramide ratio, we designed an siRNA experiment to knockdown *DEGS1* gene expression in HepG2 cells, used as a model of hepatocyte function. Western blot analysis after transfection of siRNAs showed a clear decrease in DEGS1 protein levels in siRNA knock down cells across a time course of 24 to 96 hours in comparison to non-targeting siRNA control cells (Figure S5). We harvested cell lysate at 96 hours and conducted cell based lipidome analysis. Comparison of the total ceramide to dihydroceramide ratio in control cells versus *DEGS1* siRNA knockdown cells shows

a large decrease in DEGS1 activity (Figure 4, p = 0.007), a directional effect that is consistent with that observed for L175Q heterozygotes measured in our cohort.

CRISPR/Cas9 funtional confirmation of lipidomic effects of DEGS1 L175Q variant in HepG2 cells

To assess whether the DEGS1 L175Q variant encodes a functional mutation, we performed a CRISPR/Cas9 genome editing experiment in HepG2 cells to introduce the L175Q variant in the heterozygous state to identify lipidomic effects, through the total ceramide : total dihydroceramide ratio, in a cell culture model of hepatocyte function. A western blot confirmed equivalent DEGS1 protein expression levels in each modified cell line and the mock control cell line (data not shown). We compared the total ceramide : total dihydroceramide ratio, as a measure of DEGS1 activity, between four CRISPR/Cas9 L175Q cell lines (10 replicates each) and one mock HepG2 cell line that underwent the CRISPR/Cas9 transfection protocol but for which the L175Q alternation did not integrate. Figure 5 shows the significant reduction (p = 0.037) in DEGS1 activity in clones carrying the L175Q mutation (mean total ceramide : total dihydroceramide = 17.46) versus the mock control clone (mean total ceramide : total dihydroceramide = 22.27), equating to approximately a 22% reduction (0.43 SDU decrease) in DEGS1 activity. These data indicate that the DEGS1 L175Q mutation is associated with a partial loss-of-function of the protein in a HepG2 cell CRISPR/Cas9 model.

Discussion

Our investigation of the genetic factors underlying phenotypic variation in the levels of ceramide pathway lipids has identified a rare protein coding mutation, L175Q, of apparent Hispanic origin, in *DEGS1*. This variant has a large biological consequence on levels of several dihydroceramide species and has a functional effect on DEGS1 activity, confirmed through the introduction of the L175Q mutation into HepG2 cells via CRISPR/Cas9 genome editing, and subsequent lipidomic analysis. This is the first rare variant identified through WGS to have a large biological impact on the ceramide synthesis pathway. This work highlights the power of pedigree-based cohort designs for WGS studies and the potential for genetic studies in minority populations.

DEGS1 is the key gene within the *de novo* ceramide synthesis pathway, encoding the fatty acid desaturase enzyme responsible for the generation of ceramide from dihydroceramide, by addition of a double bond, in the last step of the pathway. Both dihydroceramides and ceramides are also precursor molecules to complex sphingolipids, including sphingomyelins (36). Ruangsiriluk et al (2012) found that when *DEGS1* was silenced in Huh7 cells, a dramatic increase in dihydroceramide levels was observed along with a significant reduction in ceramide levels (37). Gene expression analysis of the *DEGS1* knockdown cells revealed significantly altered cellular functions, including cell replication, cell-to-cell adhesion, organelle migration and autophagy. An increase in cholesterol influx and a reduction in cholesterol efflux is also observed in cells lacking DEGS1 activity. Ablation of *DEGS1* in *in vitro* and *in vivo* studies shows that loss of DEGS1 function impairs both adipogenesis and lipogenesis, and increases oxidative stress (38). Additionally, the presence of completely functional *DEGS1* is required for the full differentiation of preadipocytes to mature adipocytes (38). Interestingly, Holland et al (2007) showed that heterozygous *DEGS1* knockout mice have normal glucose tolerance but enhanced insulin sensitivity (39), providing a phenotypic link between *DEGS1* function and disease related mechanisms. Each of these findings are the result of the targeted ablation or silencing of *DEGS1*, with no loss of function mutations identified or studied in this gene to date. This study,

in contrast to the previous is the first to identify a *DEGS1* coding mutation and its effect on the *de novo* ceramide synthesis pathway.

While we did not detect a direct disease based association for carriers of the L175Q variant, likely as a result of small sample numbers, we did identify that the L175Q variant had a modest association with decreases in three established ceramide ratios connected to cardiovascular disease, suggesting a protective effect of this variant for cardiovascular disease. Recently, two publications have reported overt neurodegenerative disorders in humans that are compound heterozygous or homozygous for pathogenic mutations in *DEGS1* (40, 41). While the L175Q variant described here does not appear in either of these studies, and the partial loss of function of the L175Q variant has not led to development of these disorders in the heterozygous carriers in our cohort, it is interesting to note the extent that dysfunction in sphingolipid metabolism can have on human health.

Taken together, and as expected, we see increases in dihydroceramide levels as a result of the L175Q variant causing a loss of function of DEGS1 enzymatic activity. This is supported by indices of DEGS1 activity, in both human samples and HepG2 cell models, as well as decreases in downstream lipid species that are direct and indirect products of the DEGS1 conversion of dihydroceramide to ceramide; further supporting those studies investigating the loss of DEGS1 function. In terms of biological effect sizes due to specific genetic variants, these observed estimates are moderate to large depending upon their proximity to gene action and are consistent with substantial alterations in the lipidomic profiles of L175Q carriers. The L175Q variant also shows associations with established cardiovascular risk ceramide ratios, decreasing these ratios in line with a functional effect of the L175Q variant.

Studies focusing on specific processes such as metabolic stress, oxidative stress, apoptosis and cancer have shown that silencing *DEGS1* in cells results in a decrease in ceramide and sphingomyelin synthesis, and increasing levels of dihydroceramide and sphinganine species. Additionally, *DEGS1* knockdown results in down-regulation in almost all metabolic biosynthesis pathways, including cell-cycle growth (37, 42). Our results support these findings. This is the first example of a rare coding variant affecting the *de novo* ceramide synthesis pathway that has been identified through the combination of WGS and lipidomic profiling. Given the importance of dihydroceramide, ceramide and sphingomyelin metabolism in normal cellular functions, *DEGS1* may prove to be an important gene in a range of disease contexts.

Acknowledgments

The authors thank and acknowledge the participants of the San Antonio Family Heart Study for their continued involvement in our research programs. We thank Katherine Truax and Marcelo Leandro from the South Texas Diabetes and Obesity Institute for laboratory assistance.

Funding

This work was supported in part by National Institutes of Health (NIH) grants P01 HL045522 (SAFHS data collection), R01 HL113323 (whole genome sequencing and genotyping), R01 HL140681 (lipid and CVD analysis), R37 MH059490 (analytical methods and software used) and T2D-GENES Consortium grants [U01 DK085524 (San Antonio Mexican American Family Studies), U01 DK085584, U01 DK085501, U01 DK085526, and U01 DK085545 (whole genome sequencing)]. This work was conducted in part in facilities constructed under the support of NIH grant C06 RR020547. P.J.M. is supported by a senior research fellowship from the National Health and Medical Research Council of Australia (APP1042095). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

J.B., J.E.C., L.F.M., and M.K., designed and conceptualized the study; N.B.B., J.E.C. and L.F.M., and J.B. wrote the manuscript; P.J.M., J.M.W., A.C.L., S.K., J.E.C., J.L.V., H.H.B., M.B. and M.M. conducted experiments and laboratory analyses; N.B.B., J.B., M.A., J.M.P., C.G., and S.M. performed statistical analysis; T.D.D., L.A., M.C.M., and S.W.B. provided technical support and conceptual advice; J.B., J.E.C., D.C.G., R.D., L.A., M.C.M., J.L.V. and S.W.B., participated in data collection; L.F.M., M.M. performed functional analyses. All authors reviewed and approved the manuscript. J.E.C and J.B. are the guarantors for this study.

Competing interests: L.F.M., M.K., M.M., S.M., H.H.B. and M.B. are employees of Eli Lilly and Company. No additional financial interests to declare.

Data and materials availability: Data used in this paper are publicly available through dbGaP (accession numbers: phs000462.v2.p1, phs001215.v1.p1) except for the San Antonio Family Heart Study lipidomic data. The lipidomic data for the San Antonio Family Heart Study can be made available to researchers from Dr. Joanne E. Curran (joanne.curran@utrgv.edu) via a material transfer agreement for work consistent with the informed consent.

Web Resources

SOLAR available from http://solar-eclipse-genetics.org/

References

1. Curran, J. E., P. J. Meikle, and J. Blangero. 2011. New approaches for the discovery of lipid-related genes. *Clinical Lipidology*.

2. Wong, G., C. K. Barlow, J. M. Weir, J. B. M. Jowett, D. J. Magliano, P. Zimmet, J. Shaw, and P. J. Meikle. 2013. Inclusion of plasma lipid species improves classification of individuals at risk of type 2 diabetes. *PLoS ONE*. **8**: e76577.

3. Weiss, L. A., L. Pan, M. Abney, and C. Ober. 2006. The sex-specific genetic architecture of quantitative traits in humans. *Nat. Genet.* **38**: 218–222.

4. Kathiresan, S., A. K. Manning, S. Demissie, R. B. D'Agostino, A. Surti, C. Guiducci, L. Gianniny, N. P. Burtt, O. Melander, M. Orho-Melander, D. K. Arnett, G. M. Peloso, J. M. Ordovas, and L. A. Cupples. 2007. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med. Genet.* **8 Suppl 1**: S17.

5. Bellis, C., H. Kulkarni, M. Mamtani, J. W. Kent, G. Wong, J. M. Weir, C. K. Barlow, V. Diego, M. Almeida, T. D. Dyer, H. H. H. Göring, L. Almasy, M. C. Mahaney, A. G. Comuzzie, S. Williams-Blangero, P. J. Meikle, J. Blangero, and J. E. Curran. 2014. Human plasma lipidome is pleiotropically associated with cardiovascular risk factors and death. *Circ Cardiovasc Genet*. **7**: 854–863.

6. Kulkarni, H., P. J. Meikle, M. Mamtani, J. M. Weir, C. K. Barlow, J. B. Jowett, C. Bellis, T. D. Dyer, M. P. Johnson, D. L. Rainwater, L. Almasy, M. C. Mahaney, A. G. Comuzzie, J. Blangero, and J. E.

Curran. 2013. Plasma lipidomic profile signature of hypertension in Mexican American families: specific role of diacylglycerols. *Hypertension*. **62**: 621–626.

7. Kulkarni, H., P. J. Meikle, M. Mamtani, J. M. Weir, M. Almeida, V. Diego, J. M. Peralta, C. K. Barlow, C. Bellis, T. D. Dyer, L. Almasy, M. C. Mahaney, A. G. Comuzzie, H. H. H. Göring, J. E. Curran, and J. Blangero. 2014. Plasma lipidome is independently associated with variability in metabolic syndrome in Mexican American families. *J. Lipid Res.* **55**: 939–946.

8. Mamtani, M., H. Kulkarni, T. D. Dyer, L. Almasy, M. C. Mahaney, R. Duggirala, A. G. Comuzzie, P. B. Samollow, J. Blangero, and J. E. Curran. 2014. Increased waist circumference is independently associated with hypothyroidism in Mexican Americans: replicative evidence from two large, population-based studies. *BMC Endocr Disord*. **14**: 46.

9. Mamtani, M., H. Kulkarni, G. Wong, J. M. Weir, C. K. Barlow, T. D. Dyer, L. Almasy, M. C. Mahaney, A. G. Comuzzie, D. C. Glahn, D. J. Magliano, P. Zimmet, J. Shaw, S. Williams-Blangero, R. Duggirala, J. Blangero, P. J. Meikle, and J. E. Curran. 2016. Lipidomic risk score independently and cost-effectively predicts risk of future type 2 diabetes: results from diverse cohorts. *Lipids Health Dis*. **15**: 67.

10. Lemaitre, R. N., C. Yu, A. Hoofnagle, N. Hari, P. N. Jensen, A. M. Fretts, J. G. Umans, B. V. Howard, C. M. Sitlani, D. S. Siscovick, I. B. King, N. Sotoodehnia, and B. McKnight. 2018. Circulating Sphingolipids, Insulin, HOMA-IR, and HOMA-B: The Strong Heart Family Study. *Diabetes*. **67**: 1663–1672.

11. Huynh, K., C. K. Barlow, K. S. Jayawardana, J. M. Weir, N. A. Mellett, M. Cinel, D. J. Magliano, J. E. Shaw, B. G. Drew, and P. J. Meikle. 2018. High-Throughput Plasma Lipidomics: Detailed Mapping of the Associations with Cardiometabolic Risk Factors. *Cell Chem Biol*.

12. Summers, S. A. 2018. Could Ceramides Become the New Cholesterol? *Cell Metab.* **27**: 276–280.

13. Bikman, B. T., and S. A. Summers. 2011. Ceramides as modulators of cellular and whole-body metabolism. *J. Clin. Invest.* **121**: 4222–4230.

14. Siddique, M. M., Y. Li, B. Chaurasia, V. A. Kaddai, and S. A. Summers. 2015. Dihydroceramides: From Bit Players to Lead Actors. *J. Biol. Chem.* **290**: 15371–15379.

15. Wigger, L., C. Cruciani-Guglielmacci, A. Nicolas, J. Denom, N. Fernandez, F. Fumeron, P. Marques-Vidal, A. Ktorza, W. Kramer, A. Schulte, H. Le Stunff, R. Liechti, I. Xenarios, P. Vollenweider, G. Waeber, I. Uphues, R. Roussel, C. Magnan, M. Ibberson, and B. Thorens. 2017. Plasma Dihydroceramides Are Diabetes Susceptibility Biomarker Candidates in Mice and Humans. *Cell Rep*. **18**: 2269–2279.

16. Anroedh, S., M. Hilvo, K. M. Akkerhuis, D. Kauhanen, K. Koistinen, R. Oemrawsingh, P. Serruys, R.-J. van Geuns, E. Boersma, R. Laaksonen, and I. Kardys. 2018. Plasma concentrations of molecular lipid species predict long-term clinical outcome in coronary artery disease patients. *J. Lipid Res.* **59**: 1729–1737.

17. Tippetts, T. S., W. L. Holland, and S. A. Summers. 2018. The ceramide ratio: a predictor of cardiometabolic risk. *J. Lipid Res.* **59**: 1549–1550.

18. Mitchell, B. D., C. M. Kammerer, J. Blangero, M. C. Mahaney, D. L. Rainwater, B. Dyke, J. E. Hixson, R. D. Henkel, R. M. Sharp, A. G. Comuzzie, J. L. Vandeberg, M. P. Stern, and J. W. MacCluer. 1996. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. *Circulation*. **94**: 2159–2170.

19. Olvera, R. L., C. E. Bearden, D. I. Velligan, L. Almasy, M. A. Carless, J. E. Curran, D. E. Williamson, R. Duggirala, J. Blangero, and D. C. Glahn. 2011. Common genetic influences on depression, alcohol, and substance use disorders in Mexican-American families. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **156B**: 561–568.

20. Weir, J. M., G. Wong, C. K. Barlow, M. A. Greeve, A. Kowalczyk, L. Almasy, A. G. Comuzzie, M. C. Mahaney, J. B. M. Jowett, J. Shaw, J. E. Curran, J. Blangero, and P. J. Meikle. 2013. Plasma lipid profiling in a large population-based cohort. *J. Lipid Res.* **54**: 2898–2908.

21. Drmanac, R., A. B. Sparks, M. J. Callow, A. L. Halpern, N. L. Burns, B. G. Kermani, P. Carnevali, I. Nazarenko, G. B. Nilsen, G. Yeung, F. Dahl, A. Fernandez, B. Staker, K. P. Pant, J. Baccash, A. P. Borcherding, A. Brownley, R. Cedeno, L. Chen, D. Chernikoff, A. Cheung, R. Chirita, B. Curson, J. C. Ebert, C. R. Hacker, R. Hartlage, B. Hauser, S. Huang, Y. Jiang, V. Karpinchyk, M. Koenig, C. Kong, T. Landers, C. Le, J. Liu, C. E. McBride, M. Morenzoni, R. E. Morey, K. Mutch, H. Perazich, K. Perry, B. A. Peters, J. Peterson, C. L. Pethiyagoda, K. Pothuraju, C. Richter, A. M. Rosenbaum, S. Roy, J. Shafto, U. Sharanhovich, et al. 2010. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*. **327**: 78–81.

22. Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. **4**: 7.

23. Purcell, S. M., and C. C. Chang. PLINK v1.90b3m. [online] www.cog-genomics.org/plink/1.9/.

24. Wang, K., M. Li, and H. Hakonarson. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**: e164–e164.

25. Pruitt, K. D., G. R. Brown, S. M. Hiatt, F. Thibaud-Nissen, A. Astashyn, O. Ermolaeva, C. M. Farrell, J. Hart, M. J. Landrum, K. M. McGarvey, M. R. Murphy, N. A. O'Leary, S. Pujar, B. Rajput, S. H. Rangwala, L. D. Riddick, A. Shkeda, H. Sun, P. Tamez, R. E. Tully, C. Wallin, D. Webb, J. Weber, W. Wu, M. DiCuccio, P. Kitts, D. R. Maglott, T. D. Murphy, and J. M. Ostell. 2014. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* **42**: D756–63.

26. Kircher, M., D. M. Witten, P. Jain, B. J. O'Roak, G. M. Cooper, and J. Shendure. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**: 310–315.

27. Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**: 1655–1664.

28. Peralta, J. M., N. B. Blackburn, A. Porto, J. Blangero, and J. Charlesworth. 2018. Genome-wide linkage scan for loci influencing plasma triglyceride levels. *BMC Proc.* **12**: 52.

29. Han, L., and M. Abney. 2011. Identity by descent estimation with dense genome-wide genotype data. *Genet. Epidemiol.* **35**: 557–567.

30. VanRaden, P. M. 2007. Genomic measures of relationship and inbreeding. *Interbull Bull.* **25**: 111–114.

31. Higham, N. J. 2002. Computing the nearest correlation matrix - a problem from finance. *Ima Journal of Numerical Analysis*. **22**: 329–343.

32. Almasy, L., and J. Blangero. 1998. Multipoint quantitative-trait linkage analysis in general pedigrees. *The American Journal of Human Genetics*. **62**: 1198–1211.

33. Iqbal, J., M. T. Walsh, S. M. Hammad, and M. M. Hussain. 2017. Sphingolipids and Lipoproteins in Health and Metabolic Disorders. *Trends Endocrinol. Metab.* **28**: 506–518.

34. Lek, M., K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, Ben Weisburd, H.-H. Won, D. Yu, D. M. Altshuler, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. **536**: 285–291.

35. Karczewski, K. J., L. C. Francioli, G. Tiao, B. B. Cummings, J. Alfoldi, Q. Wang, R. L. Collins, K. M. Laricchia, A. Ganna, D. P. Birnbaum, L. D. Gauthier, H. Brand, M. Solomonson, N. A. Watts, D. Rhodes, M. Singer-Berk, E. G. Seaby, J. A. Kosmicki, R. K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J. X. Chong, K. E. Samocha, E. Pierce-Hoffman, Z. Zappala, A. H. O'Donnell-Luria, E. Vallabh Minikel, B. Weisburd, M. Lek, J. S. Ware, C. Vittal, I. M. Armean, L. Bergelson, K. Cibulskis, K. M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, et al. 2019. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes: Supplementary Information. *bioRxiv*. 1–39.

36. Kok, J. W., M. Nikolova-Karakashian, K. Klappe, C. Alexander, and A. H. Merrill. 1997. Dihydroceramide biology - Structure-specific metabolism and intracellular localization. *J. Biol. Chem.* **272**: 21128–21136.

37. Ruangsiriluk, W., S. E. Grosskurth, D. Ziemek, M. Kuhn, S. G. des Etages, and O. L. Francone. 2012. Silencing of enzymes involved in ceramide biosynthesis causes distinct global alterations of lipid homeostasis and gene expression. *J. Lipid Res.* **53**: 1459–1471.

38. Barbarroja, N., S. Rodriguez-Cuenca, H. Nygren, A. Camargo, A. Pirraco, J. Relat, I. Cuadrado, V. Pellegrinelli, G. Medina-Gomez, C. Lopez-Pedrera, F. J. Tinahones, J. D. Symons, S. A. Summers, M. Oresic, and A. Vidal-Puig. 2015. Increased dihydroceramide/ceramide ratio mediated by defective expression of degs1 impairs adipocyte differentiation and function. *Diabetes*. **64**: 1180–1192.

39. Holland, W. L., J. T. Brozinick, L.-P. Wang, E. D. Hawkins, K. M. Sargent, Y. Liu, K. Narra, K. L. Hoehn, T. A. Knotts, A. Siesky, D. H. Nelson, S. K. Karathanasis, G. K. Fontenot, M. J. Birnbaum, and S. A. Summers. 2007. Inhibition of ceramide synthesis ameliorates glucocorticoid-, saturated-fat-, and obesity-induced insulin resistance. *Cell Metab.* **5**: 167–179.

40. Karsai, G., F. Kraft, N. Haag, G. C. Korenke, B. Hänisch, A. Othman, S. Suriyanarayanan, R. Steiner, C. Knopp, M. Mull, M. Bergmann, J. M. Schröder, J. Weis, M. Elbracht, M. Begemann, T.

Hornemann, and I. Kurth. 2019. DEGS1-associated aberrant sphingolipid metabolism impairs nervous system function in humans. *J. Clin. Invest.* **129**: 1229–1239.

41. Pant, D. C., I. Dorboz, A. Schluter, S. Fourcade, N. Launay, J. Joya, S. Aguilera-Albesa, M. E. Yoldi, C. Casasnovas, M. J. Willis, M. Ruiz, D. Ville, G. Lesca, K. Siquier-Pernet, I. Desguerre, H. Yan, J. Wang, M. Burmeister, L. Brady, M. Tarnopolsky, C. Cornet, D. Rubbini, J. Terriente, K. N. James, D. Musaev, M. S. Zaki, M. C. Patterson, B. C. Lanpher, E. W. Klee, F. Pinto e Vairo, E. Wohler, N. L. de M. Sobreira, J. S. Cohen, R. Maroofian, H. Galehdari, N. Mazaheri, G. Shariati, L. Colleaux, D. Rodriguez, J. G. Gleeson, C. Pujades, A. Fatemi, O. Boespflug-Tanguy, and A. Pujol. 2019. Loss of the sphingolipid desaturase DEGS1 causes hypomyelinating leukodystrophy. *J. Clin. Invest.* **129**: 1240–1256.

42. Siddique, M. M., B. T. Bikman, L. Wang, L. Ying, E. Reinhardt, G. Shui, M. R. Wenk, and S. A. Summers. 2012. Ablation of dihydroceramide desaturase confers resistance to etoposide-induced apoptosis in vitro. *PLoS ONE*. **7**: e44042.

Tables

Table 1: Summary of significant DEGS1 L175Q associations with *de novo* ceramide synthesis pathway lipid species			
Lipid species	**chi**	**p(SNP)**	**β_{SNP} (SDU)**
Cer(d18:0/24:0)	53.00	3.33×10^{-13}	1.35
total dihydroceramide levels	38.18	6.66×10^{-10}	1.14
Cer(d18:0/22:0)	35.10	3.13×10^{-9}	1.10
Cer(d18:0/24:1)	27.27	1.77×10^{-7}	0.97
Cer(d18:0/20:0)	11.54	0.000681	0.64
Cer(d18:0/16:0)	11.16	0.000838	0.61
Cer(d18:0/18:0)	7.41	0.006494	0.50
Cer(d18:1/16:0)	4.26	0.039059	-0.36

Table 2. Summary of L175Q associations with ceramide ratios and biological indexes of DEGS1 activity.

Lipid species	chi	p(SNP)	β_{SNP} (SDU)
Cer(d18:1/16:0) : Cer(d18:1/24:0)	9.89	0.002	-0.59
Cer(d18:1/20:0) : Cer(d18:1/24:0)	5.54	0.019	-0.44
Cer(d18:1/24:1) : Cer(d18:1/24:0)	4.61	0.032	-0.41
ApoB : Cer(d18:0/24:0) (N=987)	69.09	9.39×10^{-17}	-1.65
Cer(d18:1/24:0) : Cer(d18:0/24:0)	55.73	8.32×10^{-14}	-1.41
Total ceramide : total dihydroceramide	46.18	1.08×10^{-11}	-1.28

Table 3. Additional *DEGS1* variants associated with lipids in ceramide pathways.				
Variant	**Lipid species**	**chi**	**p(SNP)**	**β_{SNP} (SDU)**
chr1:224379287 A>C rs965873262	Total dihydroceramide	20.35	6.45×10^{-6}	1.89
chr1:224379287 A>C rs965873262	Cer(d18:0/22:0)	23.03	1.60×10^{-6}	2.04
chr1:224379287 A>C rs965873262	Cer(d18:0/24:0)	23.36	1.36×10^{-6}	2.01
chr1:224379287 A>C rs965873262	Cer(d18:0/24:1)	14.51	1.39×10^{-4}	1.64
chr1:224377955 C>A rs759023173	Total sphingomyelin	13.22	2.77×10^{-4}	1.25
chr1:224377955 C>A rs759023173	SM(32:2)	16.08	6.10×10^{-5}	1.22
chr1:224377955 C>A rs759023173	SM(34:1)	16.03	6.20×10^{-5}	1.36
chr1:224377955 C>A rs759023173	SM(41:2)	13.14	2.90×10^{-4}	1.25

Figure titles and legends

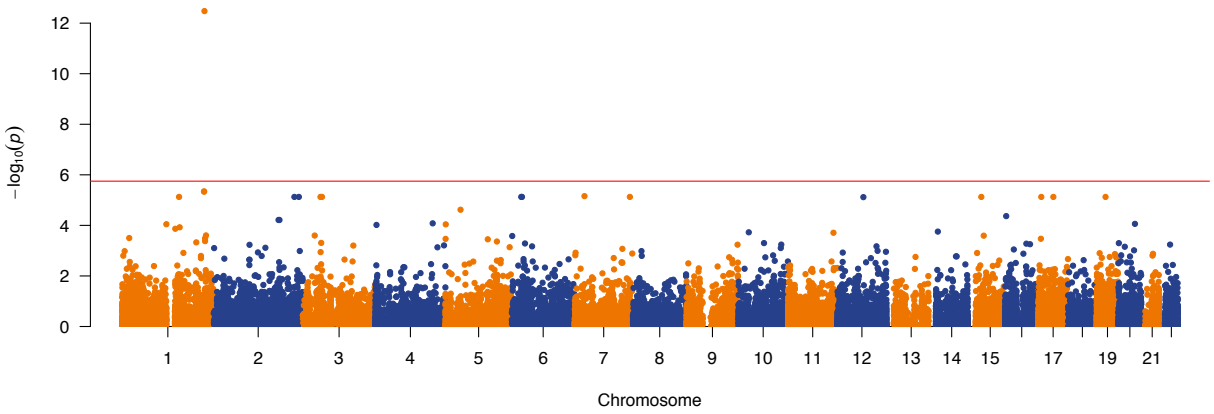

Figure 1. Genome-wide scan for Cer(d18:0/24:0) levels

Measured genotype association testing in SOLAR of 28,067 nonsynonymous, likely deleterious SNVs for Cer(d18:0/24:0) levels identifies a single genome-wide significant (Bonferroni significance level = 1.781×10^{-6}) association on chromosome 1 in *DEGS1* (L175Q, rs191144864, chr1:224377720 T>A). See Figure S1 for QQ plot.

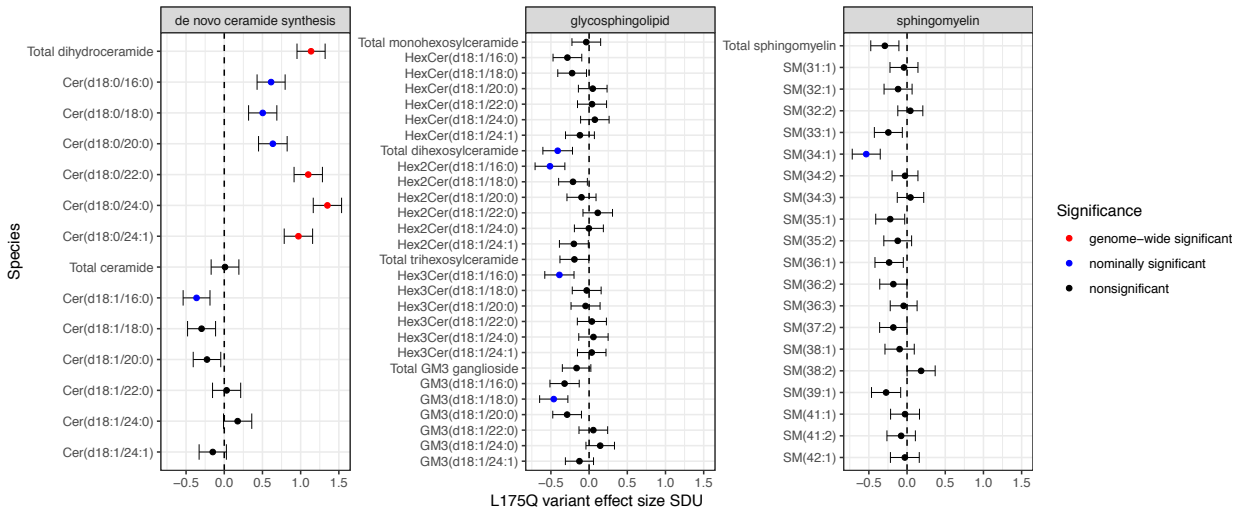

Figure 2. Effect (β_{SNP} SDU) of L175Q variant on lipid species in the *de novo* ceramide synthesis, glycosphingolipid and sphingomyelin pathways measured in this cohort

Bars indicate the standard error of β_{SNP}. Table S2 reports the exact values from this figure. Cer(d18:0/X:X) indicates dihydroceramide species; Cer(d18:1/X:X) ceramide; HexCer, monohexosylceramide; Hex2Cer, dihexosylceramide; Hex3Cer, trihexosylceramide; GM3, G_{M3} ganglioside; SM, sphingomyelin.

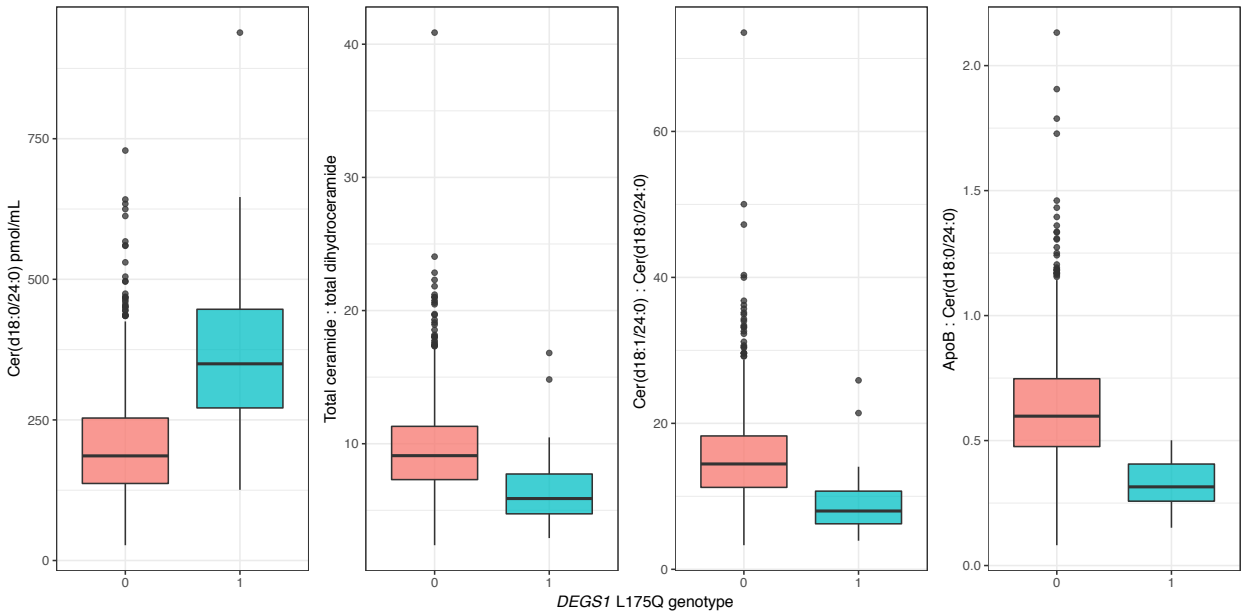

Figure 3. Changes in lipid levels in carriers of the DEGS1 L175Q variant

Box plots showing A) unadjusted Cer(d18:0/24:0) levels (pmol/mL), B) unadjusted ratio of total ceramide : total dihydroceramide, and C) unadjusted ratio of Cer(d18:1/24:0) to Cer(d18:0/24:0), and D) unadjusted ratio of ApoB to Cer(d18:0/24:0).

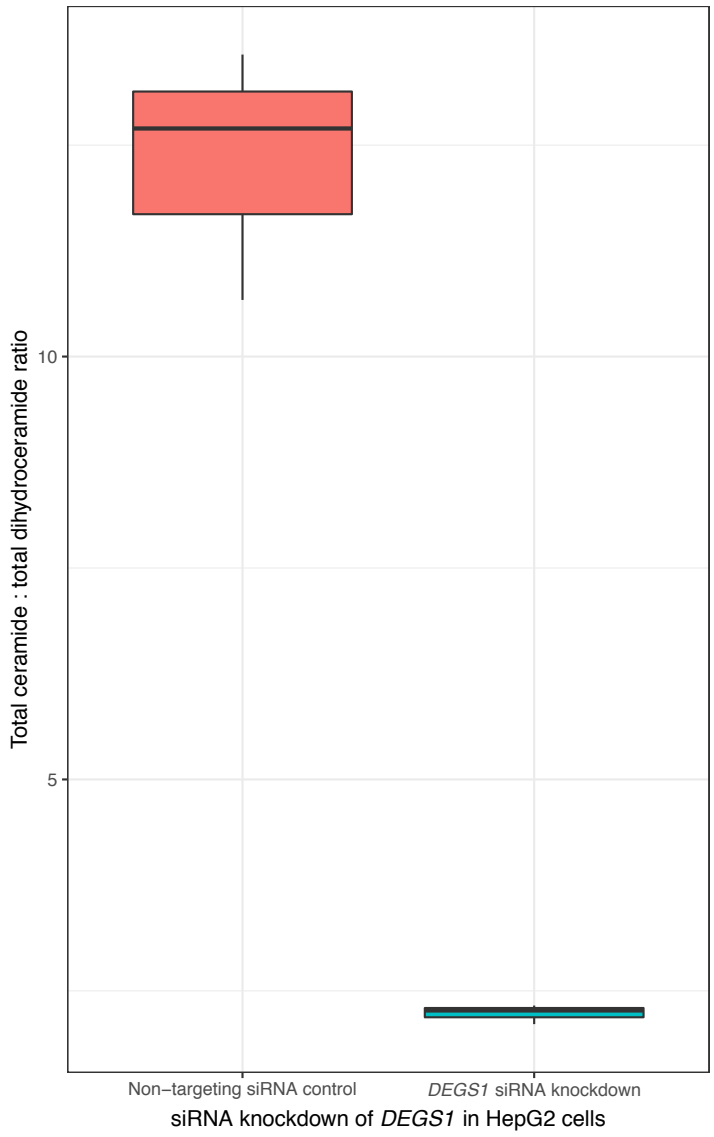

Figure 4. siRNA knockdown of *DEGS1 in vitro* in HepG2 cells

Box plots showing the siRNA knockdown of *DEGS1* in HepG2 cells shows a decrease in DEGS1 activity, measured as the ratio of total ceramide to total dihydroceramide levels. N = 3 for each experimental group, Welch two sample t-test p = 0.007.

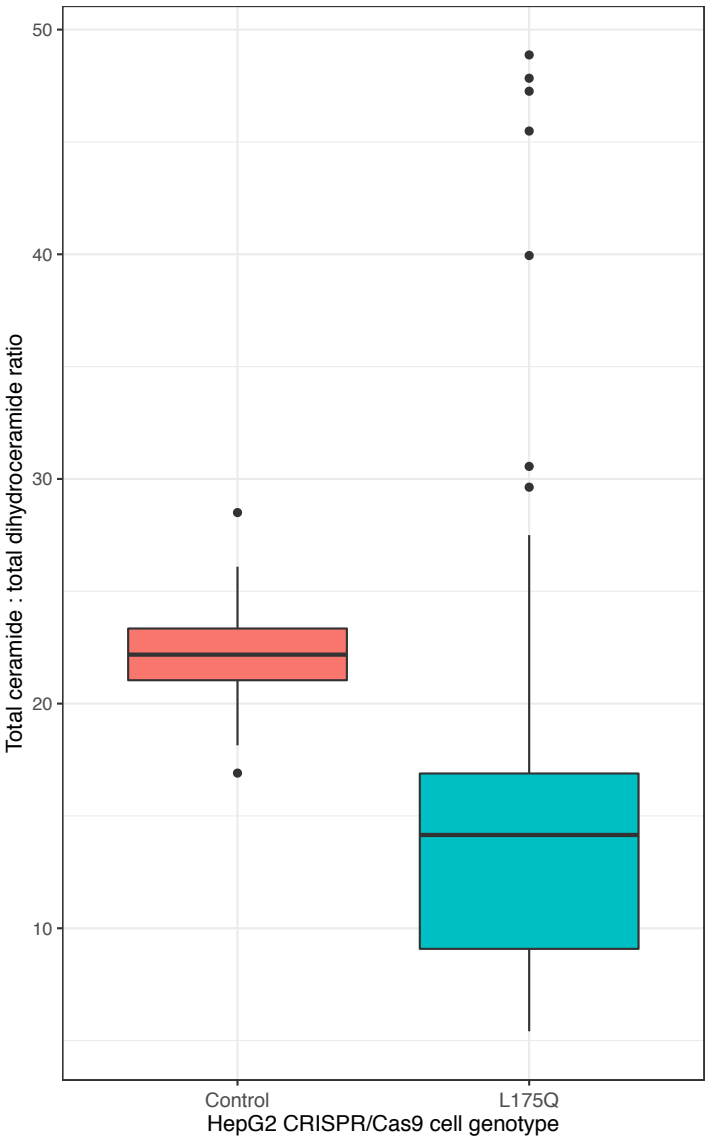

Figure 5. CRISPR/Cas9 *DEGS1* L175Q genome editing of HepG2 cells

Box plots showing the loss-of-function effect of the *DEGS1* L175Q variant after the CRISPR/Cas9 genome editing of HepG2 cell lines. Control cells are ten replicates from one HepG2 clone that underwent the full CRISPR/Cas9 transfection protocol, without mutation integration. The L175Q HepG2 clones are from four cell lines where L175Q mutation integration was successful, with ten replicates of each clone. Welch two sample t-test p = 0.037.