Supplementary Appendix

Table of Contents

Detailed Material and Methods

1. Human studies

i) Study populations and phenotypic measurements:

Bariatric Cohort:

The *TM6SF2* rs58542926 variant was genotyped in 983 bariatric surgery patients at the Geisinger Health System in Danville, PA. The protocol was approved by the Geisinger Clinical Institutional Review Board, and all subjects provided written informed consent. Prior to surgery, patients were extensively phenotyped, which included a comprehensive medical history and physical examination providing anthropometry, fasting serum glucose and lipids, liver and kidney function and medication usage. Intra-operative liver biopsy specimens were processed for histocytochemistry and scored for steatosis grade (0=<5%, 1=5%-33%, 2=33%-66%, 3=>66%), lobular inflammation (0=no foci, 1=<2 foci per 200X field, 2=2-4 foci per 200X field, 3=>4 foci per 200X field) hepatocyte ballooning and perivenular fibrosis, as previously described (1, 2).

Old Order Amish of Lancaster County, PA cohort:

The Old Order Amish (OOA) of Lancaster County, PA is a founder population that our group at the University of Maryland School of Medicine has been studying since 1993. The current Lancaster OOA population numbers ~34,000 individuals, including ~12,000 adults. We estimate that these individuals are descendants of 554 founders, with 128 founders contributing to 95% of the present day gene pool (3). Through leadership of coauthor Dr. Alan Shuldiner, as of 1/20/2016 we had 7,236 Amish adults with DNA samples enrolled in one or more of our studies, where ~6,000 have extensive phenotype information in one or more of the following areas: glucose homeostasis/diabetes, cardiovascular health, bone health, longevity, and platelet function and response to anti-platelet agents. This rich collection of well-phenotyped subjects makes up the Amish Complex Disease Research Program (ACDRP). For our study, we selected 3,556 of these subjects for whom we had both *TM6SF2* genotypic and relevant phenotypic information. Specifically, the main data set included only subjects with non-missing values for body mass index (BMI), and serum LDL-cholesterol, HDL-cholesterol and triglycerides (TG) (3-5).

Standardized protocols and procedures were used for phenotyping across all studies. Height and weight were measured using a stadiometer and calibrated scale, and BMI (kg/m^2) calculated. Fasting serum TG, total cholesterol and HDL-cholesterol (HDL-C) were measured by Quest Diagnostics (Horsham, PA). LDL-cholesterol was calculated by the Friedewald equation where TG<400 mg/dl. Glucose concentrations were assayed with a glucose analyzer (Beckman Coulter, Brea, CA or YSI, Yellow Springs, OH). Insulin levels were determined by radioimmunoassay (Linco Research Inc., St. Charles, MO). HOMA-IR was calculated using the following formula: HOMA-IR = [Fasting glucose (in mg/dl)*Fasting insulin (in μU/ml)]/405.

Lipid sub-fraction data and triglyceride intervention data were measured in the Heredity and Phenotype Intervention (HAPI) study (5). HAPI is an ACDRP intervention study that evaluated short-term cardiovascular interventions. The current analysis includes baseline data and data after a high-fat challenge. The high-fat challenge consisted of a whipping cream milk shake

standardized to be 782 calories/m^2 body surface area and with the following composition: 77.6% fat, 19.2% carbohydrate, 3.1% protein. Blood was drawn before intervention (fasting) and at 1, 2, 3, 4 and 6 hours after intervention. Lipoprotein sub-fractions were measured by Vertical Auto Profile (VAP) technology (Athrotech, Birmingham, AL) at baseline (6). Plasma apoB-48 was assessed with an enzyme-linked immunosorbent assay (ELISA) (Biovendor R&D, Ashville NC). Hepatic fat content (determined via spleen:liver density ratio, where a higher ratio indicates more fat) was measured in a subset of subjects by thoracic electron-bean computerized tomography (EBCT) scans on an Imatron C-150 EBCT scanner. Two 1.0 cm^2 regions of interest (ROI) were measured for the liver and one was measured from the spleen in a manner that minimized measurements of vessels, focal lesions, artifacts or the edge of the organ. AccuView (Accuimage Diagnostics Corp., San Francisco, CA) software was used to calculate the attenuation coefficient in Hounsfield Units for each ROI (7). The average liver attenuation ratio was divided by the spleen measure for standardization (liver:spleen ratio). The liver to spleen ratio was inverse normally transformed for analysis. All scans were performed and analyzed blinded to genotype. Due to strict cultural and religious principles, alcohol consumption in the Amish population is frowned upon and very rare and thus we believe unlikely to confound our measurements of liver fat. In order to maintain cultural appropriateness, we do not ask participants about alcohol consumption.

ii) rs58542926 genotyping:

The bariatric surgery patients were recruited through the Geisinger Medical Center for Nutrition and Weight Management and were prospectively enrolled into a clinical research program in obesity. The research was approved by the Institutional Review Board of the Geisinger Clinic and all patients provided informed consent. Genomic DNA was isolated from patient whole-blood samples (8) and was then genotyped on the Illumina HumanOmniExpress-12 v1.0 DNA analysis bead chip. Only SNPs passing quality control were used for haplotype phasing and imputation, as detailed in the eMERGE platform manuscript (9). Accuracy of imputed results including rs58542926 was determined based on high concordance rates (>99.8%) in the masked analysis (9).

In the Amish cohort, Taqman (Applied Biosystems, Foster City, CA) genotyping was conducted using manufacturer recommended procedures. The variant conformed to the expectations of Hardy-Weinberg equilibrium (p=0.76), had a high call rate (99%) and 99.3% concordance rate was observed among duplicate samples.

iii) Analyses of quantitative traits:

In the Geisinger sample continuous variables (e.g., LDL-C/HDL-C) were modeled using linear regression and comparing carriers of the T allele to the CC genotype while simultaneously adjusting for age, sex, and current use of lipid medication. Liver histology measures were dichotomized and analyzed using logistic regression adjusting for age, sex and current use of lipid medication. Analyses were conducted using SAS (v9.3, Carey, NC).

Association analyses of quantitative traits in the OOA was performed using the measured genotype approach that models variation in the trait of interest as a function of measured environmental covariates, measured genotype and a polygenic component to account for phenotypic correlation due to relatedness. Sex and age were included as covariates and SNPs were coded using an additive model. The polygenic component was modeled using the relationship matrix derived from the complete 14-generation pedigree structure to properly

control for the relatedness of all subjects in the study. These analyses were carried out using the mixed model software MMAP developed in our group at the University of Maryland. The 25 traits examined include serum lipids (i.e. total cholesterol, HDL-cholesterol, LDL-cholesterol, TG and subfraction data), insulin and glucose after fasting, triglyceride excursion after high-fat feeding and hepatic fat content. Fasting insulin, glucose measures and HOMA-IR (skewed traits) were log transformed before analysis and untransformed means were reported for ease of interpretation. These association studies were not impacted by exclusion of carriers of *ApoB*, *ApoC3*, *PNPLA3* and *HSL* variants (data not shown).

2. Zebrafish studies:

i) Zebrafish husbandry, embryo culture, and adult gene expression analysis:

Adult wild-type Tubingen or foigr (*foie gras*) (10) zebrafish embryos were collected from natural matings and maintained at 28.5°C. Embryos were cultured in embryo medium (11) at 28.5°C until harvesting at time points between 1 and 7 days post-fertilization (dpf). Larval stages were verified by measurement of larval length (Supplementary Figure 8). Adult animals were isolated at 6 months of age and euthanized prior to dissection. Liver and intestine was separated from whole gastrointestinal tracts dissected out from adults (n=5). Liver tissue was verified by qRT-PCR using *fabp10a* as a marker, which was expressed at very high levels in liver samples and very low levels in intestine (data not shown).

ii) Quantitative RT-PCR:

cDNA was generated by isolation of whole embryo or extracted liver RNA and subsequent reverse transcription (RevertAid First Strand cDNA Synthesis Kit, Thermo Scientific, Waltham, MA). Samples and controls/standards were run in triplicates on a Roche LightCycler 480 instrument (Roche, Basel, Switzerland) and mRNA expression levels were quantified relative to β-actin. Experiments were repeated at least three times. Primer sequences available upon request.

iii) Morpholino injection and validation:

Antisense oligonucleotide morpholinos (Gene Tools, LLC, Philomath, OR), designed to individually target exon 3 or 4 of *tm6sf2,* were injected into wild-type embryos between the 1 and 2-cell stage (n>200). Each morpholino (MO) was injected at three different concentrations (n>200 per concentration) to determine dose-response as well as the optimal morpholino concentration for future studies. A non-specific control MO was utilized in all experiments. Morpholino sequences are as follows: *tm6sf2_e3*: 5'-ACCACTGGCCTGAAATACAAAACA-3'; *tm6sf2_e4*: 5'-CCTCATGTCATTACGCTAACCGTTT-3'; control MO: 5'-CCTCTTACCTCAGTTACAATTTATA-3'; *dgat2_e2*: 5'-CCGAGCCTACACAGGATTTAAAGGA-3'. Embryos injected with each MO were cultured at 28.5°C. To validate disruption of splicing and removal of a targeted exon, cDNA was generated from control, *tm6sf2_e3*, *tm6sf2_e4*, and *dgat2_e2* morphants, as outlined above, at each day of development starting at 24 hours post fertilization (hpf) through 5 dpf. A total of 20 embryos were pooled for RNA extraction at each

time point. Experiments were repeated three times. Primer sequences are available upon request.

iv) CRISPR design and validation:

To confirm the results obtained by MO-injection, CRISPR-Cas9 mutants were generated for both *tm6sf2* exons (3 and 4). Briefly, guide RNAs composed of a 22 bp target sequence flanked by a 5' T7 promoter sequence (5'-TAATACGACTCACTATA-3') and a 3' overlap sequence (5'-GTTTTAGAGCTAGAAATAG-3') were annealed to a generic oligo (5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC-3'). The assembled oligos were transcribed *in vitro* using the MaxiScript Kit (Ambion, AM1314M). Cas9 transcript was generated using the mMESSAGE mMACHINE T3 kit (Life Technologies, AM1348). Approximately 300 single-celled wild-type (Tubingen) embryos were co-injected with 25 pg guide RNA and 300 pg Cas9 RNA. For each *tm6sf2* gRNA construct (exon 3 and exon 4), 50 embryos were fixed at 5 dpf, stained with oil Red O (see methods below), and imaged for lipid accumulation. Genomic DNA was extracted for T7 endonuclease-based confirmation of mismatch (12) following PCR amplification of the target region and sequencing. The remaining embryos were used to propagate *tm6sf2* exon3 and exon 4 mutant genetic lines.

v) Molecular cloning and transcript expression:

The coding sequence of human *TM6FS2* cDNA was amplified by PCR using Phusion DNA polymerase using forward primer 5'-AGATCTCATatggacatcccgccgctgg and reverse primer 5'-GTCGACtcaatgctgcttcttgtggag-3' from human liver cDNA. The PCR products were purified and subcloned into pCR2.1-TOPO vector (ThermoFisher, Waltham, MA). To generate the human *TM6SF2* E167K mutant construct, a DNA fragment was amplified from the wild-type human *TM6SF2* plasmid using a mutant forward primer containing a XmcI restriction enzyme sequence (5'-tcgccatgagcatcctggtgttccttacaggaaacattcttggcaaatacagctccaa-3') and reverse primer (5'-GTCGACtcaatgctgcttcttgtggag-3)'. The 3' end of the wild-type cDNA fragment from the XmcI cut site to the stop codon of the wild-type vector was replaced by mutant PCR fragment by T4 DNA ligase after digestion with XmcI and SalI. Each ORF was then subcloned into the pCS2+ vector (Clontech, MountainView, CA). Capped RNA was generated from each *TM6SF2*-pCS2+ construct using the mMESSAGE mMACHINE T7 Transcription Kit (Ambion, ThermoFisher Scientific, Waltham, MA) after linearization with NotI or SacII. *TM6SF2* RNA or mutated (E167K) *TM6SF2* was injected into embryos either alone or along with 6ng *tm6sf2_e4 MO* or 2ng *tm6sf2_e3 MO*.

All constructs were verified by restriction enzyme digestion and by DNA sequencing analysis.

vi) Oil red O (ORO) Staining:

5 dpf control and morphant larvae were fixed in 4% paraformaldehyde in 1x phosphate buffered saline and incubated overnight at 4°C. The following day, embryos were subjected to oil red O staining and imaged as previously described (14).

vii) Lipid Droplet Assessment and Quantification:

ORO stained embryos (wild-type, *foigr*, and morphant) were visualized under Brightfield and imaged on a glass slide within 3% methyl cellulose. Embryos were measured for length and size of swim bladder in order to adjust and identify any aging discrepancies; outliers were removed. Those embryos with appropriate 5 dpf sizing were then assessed for number of cytosolic lipid droplets in the liver (49 nm^2 square region). Specifically, cytosolic lipid droplets were quantified via manual counting within 5 separate 49 nm^2 liver regions while lipid droplet size (in the same five 49 nm^2 regions) was quantified using ImageJ software. Moreover, the number of morphant larvae affected versus not affected was also recorded to determine morpholino efficacy. A minimum of 50 embryos were subjected to quantification per injection group.

viii) Diets and feeding:

Starting at 5 dpf, larvae were fed a control diet (CD) [Zeigler AP100 (Aquatic Habitats, Inc., Cary, NC) which was replaced twice daily until 7 dpf (48 hours of feeding) (15). At 7 dpf, larvae were fixed in 4% paraformaldehyde in 1x phosphate buffered saline and incubated overnight at 4°C in preparation for ORO staining. For transmission electron microscopy (TEM), embryos were fed a high-fat diet (HFD) for 3 hours starting at 5 dpf and then processed for TEM (for TEM fixation and sectioning methods see "Transmission electron microscopy" below). Preparation of the 10% HFD is as follows: fresh organic egg yolk isolated, forced through a 28G1/2 needle, and then added to embryo medium to obtain a 10% egg yolk-embryo medium solution. The solution was then vortexed and forced pipetted for ~10 minutes until lipid micelles formed (sized at ~1 mm diameter via stage micrometer).

ix) LDL-cholesterol quantification:

Control diet-fed control MO and *tm6sf2_e4* MO embryo homogenates, comprised of 100 embryos per group, were utilized in the LDL/VLDL cholesterol assay kit (Cell Biolabs, Inc., San Diego, CA) following the protocol previously outlined (17). Exceptions to the protocol include utilization of de-livered embryos rather than whole embryos. Experiments were repeated four times. Values obtained by fluorimetric analysis were calculated relative to total protein concentration.

x) Transmission electron microscopy:

Three separate experiments were performed. TEM processing and imaging was performed at two locations, the University of Maryland and the Carnegie Institution for Sciences.

At 5 dpf, unfed, 3 hour HFD-fed, and 3 hour HFD/18- or 24-hour clearance control MO and *tm6sf2_e4* MO embryos were immediately placed in PIPES and incubated at 4°C overnight. The following morning, the head (anterior of the heart) and tail (posterior of the urogenital pore) portions were removed and the midsection was mounted and transversely sectioned to obtain both liver and anterior intestine. Silver colored ultrathin sections of 70 nm were mounted onto

grids and imaged using a FEI Tecnai T12 high-resolution transmission electron microscope (UMB Core Facility, Baltimore MD). Images were acquired with AMT-XR611 digital camera using AMTV600 software (Advanced Microscopy Techniques). Lipid (size, number, and morphology), ER dilation, and location (liver, intestine) were measured in at least 2 larvae per condition.

Addtionally, At 5 dpf, unfed, 3 hour HFD-fed, and 3 hour HFD/18- or 24-hour clearance control MO and *tm6sf2_e4* MO embryos were screened for "full" intestines (except for unfed embryos) and immediately fixed in a 3% glutaraldyhyde, 1% formaldehyde, 0.1 M cacadylate solution (pH 7.4) and incubated at 4°C overnight. Postfixation was done in 1% osmium tetroxide and En Bloc-stained with uranyl acetate in maleate (a detailed protocol is available upon request). Zebrafish were orientated in resin (Epon+Quetol (2:1), Spurr (3:1), 2%BDMA) blocks. Transverse sections were made to obtain liver and anterior intestinal sections using an ultramicrotome (Porter-Blum MT-2; Sorvall Instruments, Newton, CT), mounted on Formvar-coated grids, and stained with lead citrate. Images were captured with a Phillips Tecnai 12 microscope (Carnegie Institution for Science, Baltimore, MD) and recorded with a Gatan multiscan CCD camera using Digital Micrograph software. Lipid (size, number, and morphology), ER dilation, and location (liver, intestine) were measured in 4-8 larvae per condition.

Data from both instruments (UMB Core Facility and Carnegie Institution for Science) were combined for at least n=12 larvae per condition except for the 18 hr time point (n=4).

xi) Cytosolic lipid droplet and ER lumen quantification:

High-quality, high-magnification images were obtained using a FEI Tecnai T12 high-resolution transmission electron microscope or the Phillips Tecnai 12 microscope. Lipid droplet counts (size and number) were analyzed using ImageJ software (image magnification: 2-10 μm). Eight embryos were observed under TEM per condition (control MO vs. *tm6sf2_e4* MO; unfed, 3 hour HFD, 3 hour HFD/24 hour clearance). For ER lumen quantification, a total of 10 random smooth ER lumen widths were quantified per cell with three measurements taken per ER lumen. At least 20 cells were assessed per condition. ImageJ software was utilized (image magnification: 500 nm).

xii) Whole-mount *in situ* hybridization:

In situ hybridization was carried out as per standard protocols (Thisse & Thisse 2008 PMID:18193022). Riboprobes for *tm6sf2* were synthesized by *in vitro* transcription from vectors containing the open reading frame cloned from 5 dpf larval cDNA (primers available upon request). Transcription was carried out using the DIG RNA Labeling Kit (T7) (Roche, Indianapolis, IN) and precipitated with ammonium-acetate.

xiii) Drug treatments:

Injected larvae (control MO, *tm6sf2*_e4 MO, *tm6sf2*_e4 + *TM6SF2* WT RNA, *tm6sf2*_e4 + *TM6SF2* E167K RNA, *TM6SF2* WT RNA, or *TM6SF2* E167K RNA) were cultured until 96 hpf at which point embryos were incubated in one of three culture mediums for 48 hours: i) 1% DMSO

in embryo medium; ii) 1 µg/mL Tunicamycin (T7765-1MG; Sigma-Aldrich, St. Louis, MO) in embryo medium. At 5 dpf, livers from individual larvae were harvest and processed for RNA (TRIzol® reagent (Invitrogen, Grand Island, NY)) and cDNA synthesis (RevertAid First Strand cDNA Synthesis Kit, Thermo Scientific, Waltham, MA).

xiv) Statistical analyses:

A Student's *t*-test was used to determine significance. Bonferroni correction was used to adjust for multiple comparisons (at least four independent experiments in all cases).

3. Cell culture studies

i) RNA extraction and qRT-PCR:

Total RNA of pooled human multiple tissue samples (control) were purchased (Clontech laboratory Inc., Mountain View, CA). Total RNA of from 24 liver biopsies of EE and EK carriers were obtained from bariatric surgery patients at the Geisinger Health System in Danville, PA. Six EE and EK carriers had been diagnosed with grade 0 or 1 steatosis but no fibrosis, 6 EE and EK with both steatosis (grade 1 and 2) and fibrosis. Patients were all female and morbidly obese. Multiple mouse tissues from adult C56/BL6J mice were homogenized in TRIzol® reagent (Invitrogen, Grand Island, NY). Total RNA was extracted according to manufacturer's instruction, and cDNA synthesized using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Basel, Switzerland). Quantitative real time PCR was performed in triplicate for each sample on a LightCycler 480 (Roche, Basel, Switzerland) using SYBR Green I PCR mixture (Roche Diagnostics Corporation, Indianapolis, IN). Relative fold changes of *TM6SF2* to *β-ACTIN* or 18S mRNA expression levels were calculated using the $2^{-\Delta CT}$ method. Primer sequences for SYBR Green assays are shown in Supplemental Table 7.

For Caco-2 cultured cells, total RNA was extracted using Qiagen RNease Kit (Qiagen, Valencia, CA), and cDNA synthesized using Transcriptor First Strand cDNA Synthesis Kit (Roche, Basel, Switzerland). Steady-state mRNA levels were determined by two-step quantitative real time PCR (qRT-PCR) using the LightCycler 480 (Roche, Basel Switzerland) and Taqman probe/primer sets (Applied Biosystems, Carlsbad, CA), with 18S as an internal control for normalization. Primer sequences for Taqman assays are shown in Supplemental Table 7.

ii) Molecular cloning:To generate the YFP fusion vector, cDNAs of *TM6SF2* were amplified from human cDNAs by PCR to include BglII/SalI restriction sites at the 5' and 3' ends; product was subsequently cloned in frame with monomeric pEYFP-C1 (13). See 2. Zebrafish studies: v) Molecular cloning and transcript expression for more information as to initial human *TM6SF2* vector construction, which was subsequently utilized to generate the above construct.

Two sets of DNA oligonucleotides encoding shRNA targeting *TM6SF2* (shRNA1: GCACTGTTCACTACCTCCTCT; shRNA2: GTGCTTGATTGCCCCACAGAT) and one set of scrambled shRNA (control shRNA GCACCAACATCGAAAGTGACT), which does not match any known human genes, were first cloned into the pEntra shuttle vector (Clontech, Mountain View, CA). The *TM6SF2* and control shRNA inserts were then subcloned into the lentiviral destination

vector pSMPUW-CMV-DEST (Cell Biolabs, San Diego, CA) under a CMV promoter using standard Gateway LR cloning protocol according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). All constructs were verified by restriction enzyme digestion and by DNA sequencing analysis.

iii) Caco-2 cells cultures and gene knockdown by shRNA:

Caco-2 cell cultures were obtained from American Type Culture Collection (Mannassas, VA) and grown in DMEM with 4.5 g/L glucose, 4 mM glutamine, 100 U/mL penicillin, 100 mg/mL streptomycin and 20% FBS in a 95% air-5% CO_2 atmosphere at 37°C, as previously described (18). The medium was changed every other day. Cells were plated at a density 1-2 10^4 cells/cm^2 in 25-cm^2 flasks and split with 0.25% trypsin-1 mM EDTA when they reached 70–90% confluence. For experiments, cells were plated at 10^7 cells/cm^2 in 12-well plates for gene expression studies or in confocal dishes for imaging. Lentiviral plasmid vectors containing shRNA inserts targeting human *TM6SF2* (GenBank accession NM_001001524) were obtained (see below). Caco-2 cells were transduced with an equal mix of lentiviral plasmid vector containing shRNA targeting human *TM6SF2* (GeneBank accession NM_001001524) or scrambled shRNA (see ii) Molecular cloning section above) that does not match any known human genes, with identical transduction procedures. After 48 hours, cells were analyzed for gene expression or by confocal image analysis for CLD content. Confocal imaging of live cells was performed at 37°C and 5% CO_2 using a Zeiss LSM510 microscope equipped with an S-M incubator (Carl Zeiss MicroImaging, Inc. Thornwood, NY), and controlled by a CTI temperature regulator along with humidification and an objective heater as previously described (19). For pulse/chase radiolabeling experiments, cells were differentiated as previously described (20) and lentiviral transduction performed 48 hours before the start of the pulse experiments (see below).

iv) Lentiviral vector production:

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% FBS, 100 µg/mL streptomycin and 100 U/mL penicillin. Cells were allowed to reach 90% confluence at which time they were transfected in the presence of DMEM with 1.2 µg of pSMPUW lentiviral expressing vector containing shRNA targeting *TM6SF2* or scrambled shRNA in (Cell Biolabs, San Diego, CA) (see ii) Molecular cloning section above) 1.2 µg of pCD/NL-BH*DDD (Addgene plasmid 17531; plasmid was generously made available to Addgene biobank by Dr. Jacob Reiser) and 0.2 µg of pVSVG (Cell Biolabs, San Diego, CA), for each well of a 6-well plate using LipoD293 (SignaGen Laboratories, Rockville, MD). Liposome/DNA complex-containing medium was removed and replaced with fresh medium 16 hours post transfection. The supernatant containing the viral vector particles was collected at 48 and 72 hours following initial transfection, combined and aliquoted for storage at −80°C. Lentiviral vectors used for all experiments had a minimum titer of 5×10^5 IFU/mL.

v) Detection lipid content by Bodipy 493/503 staining:

Lipid droplet visualization and cell counting (nuclei) via image analysis was performed as previously described (21).

vi) Cell incubation:

Radiolabeled [^{3}H]oleic acid was used to track synthesis and secretion of lipids; these experiments had two incubation periods designated pulse and washout (22). The experimental design was similar to a protocol previously described (22). During the pulse period, cells were incubated for 16 hours in culture medium containing 0.6 mM [^{3}H]oleic acid complexed to fatty acid-free bovine serum albumin (BSA) (Lot number 99, EDM Millipore, MA) at a 4:1 molar ratio. For the washout period, cells were incubated in culture medium containing 5 g fatty acid-free BSA per liter of medium for 1 hour, providing sufficient time for the secretion of lipoproteins assembled during the pulse period. Separate samples of cells and media were collected after the pulse and pulse/washout. Lipids were extracted by the Dole extraction method (23), and 12.5% of the total lipid from the cell samples and 25% from the media samples were analyzed by thin layer chromatography (24). [^{3}H]oleic acid incorporation into the TG fraction in the media and in the cells was calculated as picomoles and nanomoles, respectively.

vii) Statistical analysis:
Statistical significance was tested using either one-way analysis of variance or a two-tailed Student's *t*-test (GraphPad Software Inc., La Jolla, CA).

Supplemental Material and Methods references:

1. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, Ferrell LD, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313-1321.
2. Gerhard GS, Chokshi R, Still CD, Benotti P, Wood GC, Freedman-Weiss M, Rider C, et al. The influence of iron status and genetic polymorphisms in the HFE gene on the risk for postoperative complications after bariatric surgery: a prospective cohort study in 1,064 patients. Patient Saf Surg 2011;5:1.
3. Lee WJ, Pollin TI, O'Connell JR, Agarwala R, Schaffer AA. PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County. BMC Med Genet 2010;11:68.
4. Hsueh WC, Mitchell BD, Aburomia R, Pollin T, Sakul H, Gelder Ehm M, Michelsen BK, et al. Diabetes in the Old Order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. Diabetes Care 2000;23:595-601.
5. Mitchell BD, McArdle PF, Shen H, Rampersaud E, Pollin TI, Bielak LF, Jaquish C, et al. The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. Am Heart J 2008;155:823-828.
6. Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, Horenstein RB, Post W, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 2008;322:1702-1705.
7. Albert JS, Yerges-Armstrong LM, Horenstein RB, Pollin TI, Sreenivasan UT, Chai S, Blaner WS, et al. Null mutation in hormone-sensitive lipase gene and risk of type 2 diabetes. N Engl J Med 2014;370:2307-2315.
8. Chu X, Erdman R, Susek M, Gerst H, Derr K, Al-Agha M, Wood GC, et al. Association of morbid obesity with FTO and INSIG2 allelic variants. Arch Surg 2008;143:235-240; discussion 241.
9. Verma SS, de Andrade M, Tromp G, Kuivaniemi H, Pugh E, Namjou-Khales B, Mukherjee S, et al. Imputation and quality control steps for combining multiple genome-wide datasets. Front Genet 2014;5:370.
10. Cinaroglu A, Gao C, Imrie D, Sadler KC. Activating transcription factor 6 plays protective and pathological roles in steatosis due to endoplasmic reticulum stress in zebrafish. Hepatology 2011;54:495-508.
11. Westerfield M. The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio). 4 ed. Eugene, OR: University of Oregon Press, 2000.
12. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotech 2013;31:227-229.
13. Wang H, Bell M, Sreenivasan U, Hu H, Liu J, Dalen K, Londos C, et al. Unique regulation of adipose triglyceride lipase (ATGL) by perilipin 5, a lipid droplet-associated protein. J Biol Chem 2011;286:15707-15715.
14. O'Hare EA, Wang X, Montasser ME, Chang YP, Mitchell BD, Zaghloul NA. Disruption of ldlr causes increased LDL-c and vascular lipid accumulation in a zebrafish model of hypercholesterolemia. J Lipid Res 2014;55:2242-2253.
15. O'Hare EA, Wang X, Montasser ME, Chang Y-PC, Mitchell BD, Zaghloul NA. Disruption of ldlr causes increased LDL-cholesterol and vascular lipid accumulation in a zebrafish model of hypercholesterolemia. Journal of Lipid Research 2014.

16. Walters James W, Anderson Jennifer L, Bittman R, Pack M, Farber Steven A. Visualization of Lipid Metabolism in the Zebrafish Intestine Reveals a Relationship between NPC1L1- Mediated Cholesterol Uptake and Dietary Fatty Acid. Chemistry & Biology 2012;19:913-925.
17. O'Hare EA, Wang X, Montasser ME, Chang Y-PC, Mitchell BD, Zaghloul NA. Disruption of ldlr causes increased LDL-c and vascular lipid accumulation in a zebrafish model of hypercholesterolemia. Journal of Lipid Research 2014;55:2242-2253.
18. Levy E, Mehran M, Seidman E. Caco-2 cells as a model for intestinal lipoprotein synthesis and secretion. FASEB J 1995;9:626-635.
19. Bell M, Wang H, Chen H, McLenithan JC, Gong DW, Yang RZ, Yu D, et al. Consequences of lipid droplet coat protein downregulation in liver cells: abnormal lipid droplet metabolism and induction of insulin resistance. Diabetes 2008;57:2037-2045.
20. Luchoomun J, Hussain MM. Assembly and secretion of chylomicrons by differentiated Caco-2 cells. Nascent triglycerides and preformed phospholipids are preferentially used for lipoprotein assembly. J Biol Chem 1999;274:19565-19572.
21. Beller M, Sztalryd C, Southall N, Bell M, Jackle H, Auld DS, Oliver B. COPI complex is a regulator of lipid homeostasis. PLoS Biol 2008;6:e292.
22. Caviglia JM, Sparks JD, Toraskar N, Brinker AM, Yin TC, Dixon JL, Brasaemle DL. ABHD5/CGI-58 facilitates the assembly and secretion of apolipoprotein B lipoproteins by McA RH7777 rat hepatoma cells. Biochim Biophys Acta 2009;1791:198-205.
23. Dole VP. Fractionation of plasma nonesterified fatty acids. Proc Soc Exp Biol Med 1956;93:532-533.
24. Sztalryd C, Bell M, Lu X, Mertz P, Hickenbottom S, Chang BH, Chan L, et al. Functional compensation for adipose differentiation-related protein (ADFP) by Tip47 in an ADFP null embryonic cell line. J Biol Chem 2006;281:34341-34348.

Supplemental Table 1. Demographic, Anthropometric and Lipid characteristics in the study participants enrolled in the Bariatric Surgery Program at the Geisinger Clinic Center for Nutrition and Weight Management According to *TM6SF2* T allele Genotype

	Mean (SE) by Genotype		P value
	CC N=853	**- / T N=130**	
Sex	M/F	M/F	
	169/684	24/101	
Ethnic	White/Caucasian	White/Caucasian	
Age	45.9 (0.4)	46.5 (1.0)	0.61
Body Mass Index* (kg/m^2)	49.3 (0.3)	49.8 (0.8)	0.55
Total Cholesterol* (mg/dL)	189.2 (1.4)	183.4 (3.6)	0.10
HDL-C* (mg/dL)	46.8 (0.4)	47.3 (0.9)	0.52
LDL-C* (mg/dL)	107.3 (1.2)	104.4 (3.1)	0.14
Triglycerides*,† (mg/dL)	182.6 (4.6)	160.9 (6.7)	0.20

*Age-, sex- and lipids medication adjusted mean and standard error (SE) are presented for each genotype group

† trait was log-transformed for analysis (p-value) and untransformed values are presented for genotype mean and SE values

& CT n=127 and TT n=3

Supplemental Table 2. Liver histology features in the study participants enrolled in the Bariatric Surgery Program at the Geisinger Clinic Center for Nutrition and Weight Management According to *TM6SF2* T allele Genotype

	Mean (SE) by Genotype		P value
	CC N=853	- / T N=130	
Steatosis *,§ (mean (SE))	1.14 (0.03)	1.40 (0.08)	**0.04**
Lobular inflammation¥ (% > grade 0)	32.4	43.8	**0.02**
Hepatocyte Ballooning¥ (% > grade 0)	28	33.3	0.26
Perivenular fibrosis¥ (% > grade 0)	17.5	26.6	**0.008**
Cirrhosis¥ (% > grade 0)	2.1	1.6	0.9

*Age-, sex- and lipids medication adjusted mean and standard error (SE) are presented for each genotype group

§ Steatosis was examined histologically and graded in severity from 0 (no steatosis) to 3 (severe steatosis)

¥Dichotomized grade 0 vs. >0 for inflammation, ballooning, fibrosis and cirrhosis for statistical analysis

Supplemental Table 3. Lipid traits in the study participants with a steatosis grade superior of 0 enrolled in the Bariatric Surgery Program at the Geisinger Clinic Center for Nutrition and Weight Management According to *TM6SF2* T allele Genotype

Steatosis grade > 0	**Mean (SE) by Genotype**		**Beta (SE)**	**p-value**
	CC N=584	**-/T N=103**		
Total Cholesterol (mg/dl)*	191.2 (1.7)	181.5 (4.1)	-11.24 (5.34)	**0.04**
HDL (mg/dl)*	45.5 (0.5)	46.5 (1.0)	1.41 (1.27)	0.27
LDL (mg/dl)*	107.7 (1.4)	103.0 (3.5)	-7.06 (4.57)	0.12
Triglycerides (mg/dl)*†	200.5 (6.2)	162.5 (7.2)	-0.15 (0.05)	**0.02**

SE=Standard Error

*adjusted for age, sex, lipid medication

† trait was log transformed for analysis (p value) and untransformed values are presented for genotype mean, beta and SE values

Supplemental Table 4. Demographic**,** Antropometric and Lipid Characteristics of Old Order Amish Study Participants (N=3556), According to *TM6SF2* T allele Genotype

Clinical Characteristics	Mean (SE) by Genotype			Beta (SE)	p-value
	CC N=2771	CT N=734	TT N=51		
Sex	M/F	M/F	M/F		
	1260/1511	341/393	16/35	0.003 (0.020)	0.90
Ethnic	White/Caucasian	White/Caucasian	White/Caucasian		
Age	46.2 (0.3)	47.2 (0.6)	48.3 (2.5)	0.95 (0.58)	0.10
Body Mass Index (kg/m^2)*	27 (0.1)	26.9 (0.2)	27.4 (0.6)	-0.11 (0.20)	0.59
Total Cholesterol (mg/dL)*	210.3.0 (0.9)	207.5 (1.7)	178.1 (7.8)	-8.54 (1.84)	**3.6x10^{-6}**
HDL-C (mg/dL)*	57.4 (0.3)	58.8 (0.6)	64.8 (2.6)	2.48 (0.61)	**5.0x10^{-5}**
LDL-C (mg/dL)*	136.8 (0.8)	133.9 (1.6)	101.0 (6.9)	-9.45 (1.68)	**2.1x10^{-8}**
Triglycerides (mg/dL)*,†	80.3 (1.0)	73.9 (1.7)	61.7 (6.6)	-0.10 (0.02)	**7.1x10^{-7}**

SE=Standard Error

*adjusted for age, sex, study

† trait was log transformed for analysis (p value) and untransformed values are presented for genotype mean, beta and SE values

Supplemental Table 5. Metabolic Characteristics of Old Older Amish Study Participants (N=2059), According to *TM6SF2* T allele Genotype

	Mean (SE) by Genotype			Beta (SE)	p-value
	CC N=1597	**CT N=439**	**TT N=23**		
Glucose (mg/dl)*	90.5 (0.5)	88.9 (0.8)	101.6 (10.4)	0.88 (0.8)	0.29
Insulin (mu/ml)*	10.9 (0.2)	10.6 (0.3)	10.8 (0.8)	0.01 (0.02)	0.57
HOMA-IR*	2.5 (0.1)	2.4 (0.1)	2.8 (0.4)	0.02 (0.03)	0.49

SE=Standard Error

*adjusted for age, sex, study and diabetes status

Supplemental Table 6. Effect of *TM6SF2* T allele Genotype on the Lipoprotein Subclass Profile in Old Order Amish Study Participants (N=833)

	Mean (SE) by Genotype			Beta (SE)	p-value
	CC N=631	**CT N=191**	**TT N=11**		
Total Cholesterol*	210.7 (1.8)	205.1 (3.6)	157.6 (12.8)	-17.7 (3.4)	**2.5×10^{-7}**
HDL-C*	56.1 (0.5)	57.9 (1.0)	52.6 (3.7)	1.7 (1)	0.09
HDL2-C*,†	14.7 (0.2)	15.5 (0.5)	13.2 (1.9)	0.88 (0.48)	0.10
HDL3-C*	41.4 (0.3)	42.3 (0.6)	39.5 (1.9)	0.84 (0.58)	0.15
Total Non-HDL-C*	160.6 (1.7)	153.9 (3.3)	112.0 (9.7)	-19.2 (3.2)	**2.7×10^{-9}**
IDL-C*,†	10.2 (0.4)	8.5 (0.6)	5.5 (1.4)	-2.93 (0.66)	**1.9×10^{-6}**
LDL-C*	126.2 (1.4)	121.7 (2.7)	82.7 (8.3)	-14.8 (2.7)	**3.5×10^{-8}**
Total VLDL-C*,†	17.4 (0.2)	16.7 (0.3)	16.3 (1.3)	-1.4 (0.39)	**2.4×10^{-4}**
VLDL3-C*,†	9.5 (0.1)	9.1 (0.2)	8.7 (0.6)	-0.83 (0.23)	**2.2×10^{-4}**
Remnant Lipoprotein-C*,†	19.4 (0.5)	17.3 (0.8)	14.0 (2.0)	-3.9 (0.87)	**9.8×10^{-6}**
Triglycerides*,†	69.5 (1.7)	64.1 (2.7)	64.8 (13.7)	-9.2 (3.0)	**8.7×10^{-4}**

SE=Standard Error

*adjusted for age, sex, study

† trait was log transformed for analysis (p value) and untransformed values are presented for genotype mean, beta and SE values

Supplemental Table 7. Sequences of human and mouse primers used for qRT-PCR measurements

Gene	Forward primer	Reverse primer
Taqman		
Human 18s	CTCAACACGGGAAACCTCAC	CGCTCCACCAACTAAGAACG
Human *GRP78*	CAGATGAAGCTGTAGCGTATGG	ACATACATGAAGCAGTACCAGGTC
Human *CHOP*	CAGAGCTGGAACCTGAGGAG	TGGATCAGTCTGGAAAAGCA
Human *ATF4*	GGTCAGTCCCTCCAACAACA	CTATACCCAACAGGGCATCC
Human *ATF6*	CTTTTAGCCCGGGACTCTTT	TCAGCAAAGAGAGCAGAATCC
Human *TM6SF2*	TTCCTTACAGGAAACATTCTTGG	GCAGGTAGGGGATGGTGAG
SYBR Green		
Mouse *B-actin*	CAGCTTCTTTGCAGCTCCTT	CACGATGGAGGGAATACAG
Mouse *tm6sf2*	GGTATTTGCTGGAGCCATTG	AGGTAGCCCAGGTGTCCTCT
Human *B-ACTIN*	AGAAAATCTGGCACCACACC	GGGTGTTGAAGGTCTCAAA
Human *TM6SF2*	GCTGCCTATGCTCTCACCTT	ACACGGTAGGTGAAGGGTGT
Human Spliced *XBP1*	CTGAGTCCGAATCAGGTGCAG	ATCCATGGGGAGATGTTCTGG
Human Unspliced *XBP1*	CAGCACTCAGACTACGTGCA	ATCCATGGGGAGATGTTCTGG

Supplemental Figure Legends

Supplemental Figure 1 Expression of *tm6sf2* in larval zebrafish. (A) Wholemount *in situ* hybridization in 4 dpf larvae using antisense riboprobe for *tm6sf2.* Robust expression in posterior intestine (dashed box, magnified below). (B) qRT-PCR expression of either *beta-actin* or *tm6sf2,* relative to *beta-actin,* in cDNA generated from mRNA extracted from 5 dpf larval liver. (C) qRT-PCR expression of either *beta-actin* or *tm6sf2,* relative to *beta-actin,* in cDNA generated from mRNA extracted from adult zebrafish liver or intestine.

Supplemental Figure 2. Validation of efficacy and specificity of *tm6sf2* MOs in zebrafish. (A**)** qRT-PCR quantification of *tm6sf2* transcript levels in embryos injected with control MO, *tm6sf2_e3* MO or *tm6sf2_e4* MO at 1, 3, and 5 dpf. Data represent average expression from homogenates of 20 pooled embryos from each of two experiments per treatment. (B) Gel electrophoresis results of RT-PCR from embryo homogenates at 1, 3 or 5 dpf treated with *tm6sf2_e3* MO, *tm6sf4_e4* MO or control MO. A 261 bp from the *tm6sf2_e3* MO and a 271 bp from the *tm6sf2_e4* MO was detected through 5 dpf, in addition to the 365 bp wild-type transcript. (C**)** qRT-PCR quantification of ∆113 p53 expression, diagnostic marker for off-target MO toxicity. (D) Gel electrophoresis results of RT-PCR from embryo homogenates at 1, 3 or 5 dpf injected with *dgat2* MO and control MO. A spliced 373 bp from the *dgat2* MO was detected through 5 dpf, in addition to the 502 bp wild-type transcript. ** p≤0.005. (E) mRNA expression levels for dgat2, qPCR analysis (n = 3 for each group, data are means + SEM

Supplemental Figure 3. Representative 5 dpf larvae used to assess lipid accumulation. Representative image of a control MO (A-A") and *tm6sf2_e4* MO (B-B') zebrafish embryos at 5 dpf stained for neutral lipids using oil red O staining. Features used for assessment of phenotypes: yellow line, liver; red box, area imaged under higher magnification in lateral view liver images; green box, 49 nm^2 area used for quantification of total lipid droplet number and size.

Supplemental Figure 4. Targeting *tm6sf2* by CRISPR/Cas9 recapitulates hepatic steatosis phenotype. (A-C) Representative lateral images of liver in oil red O stained 5 dpf zebrafish F_0 larvae targeted by CRISPR/Cas9 against *tm6sf2* exons 3 or 4. (D) The severity of hepatic lipid accumulation by targeted exon was similar to that elicited by MO against each exon. (E-F) Gel electrophoresis images of PCR products amplified from the targeted region of genomic DNA extracted from individual embryos and digested by T7 endonuclease to detect mismatch. (G-H) Sequencing results of individual mutations imparted in each exon targeted by CRISPR/Cas9.

Supplemental Figure 5. TM6SF2 is predominately expressed in liver and enterocytes, is regulated by dietary feeding and is mainly located in ER. (A) Expression of *TM6SF2* mRNA in human tissues**.** Expression of *TM6SF2* was normalized to that of *ACTB.* (B) Expression of *Tm6sf2* mRNA in mouse tissues (C57/BL strain). (C) Caco-2 cells were co-transfected with KDEL-DsRed (ER retention peptide) and *TM6SF2*-GFP and imaged by confocal microscopy after 24 hours. Micrographs depict one representative cell of ten observed in two experiments. Bar: 10 μm.

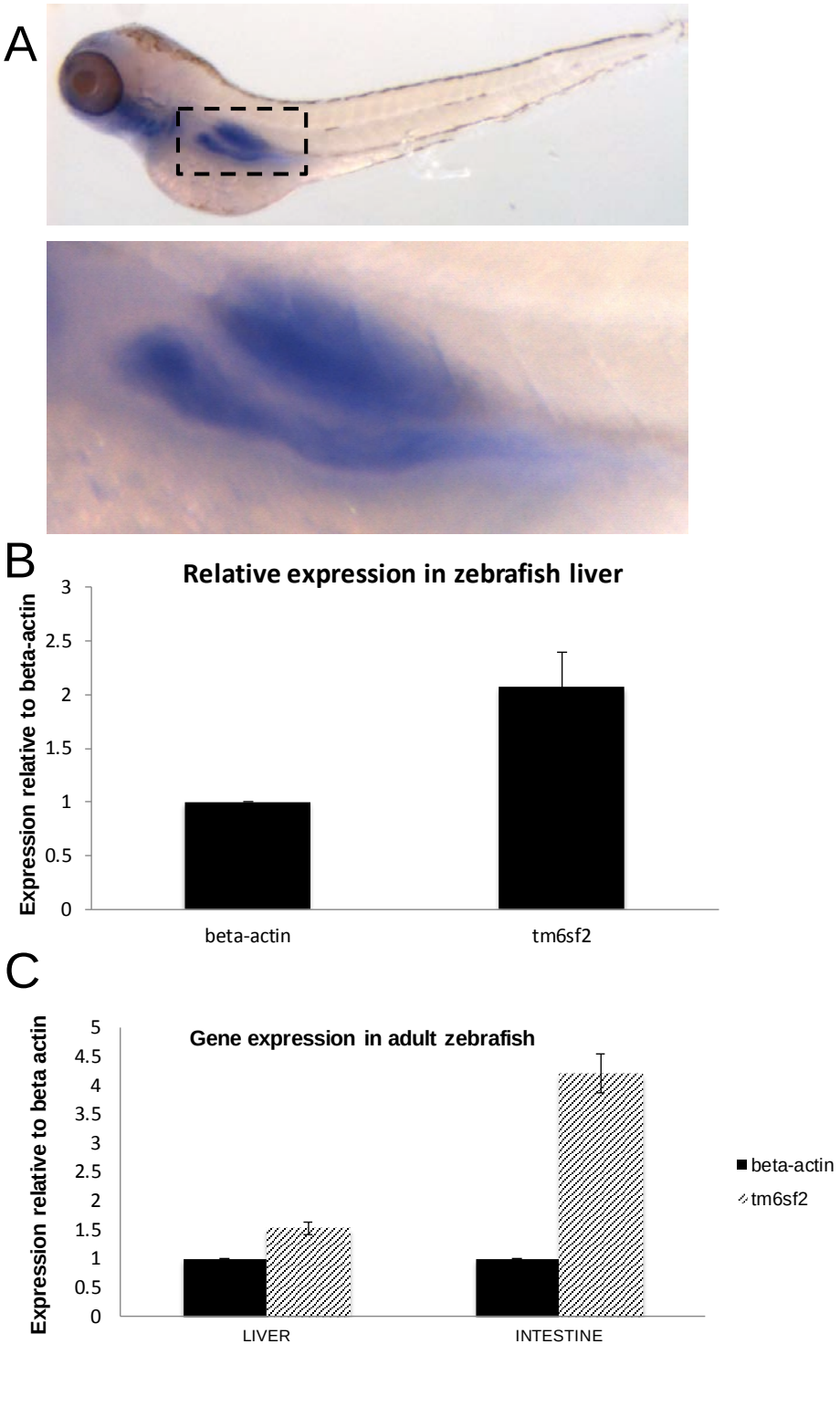
Supplemental Figure 6. Morphological characterization of hepatic liver stress in tm6sf2 morphants. (A) Representative TEM images of larval livers from unfed control MO or *tm6sf2_e4 MO* larvae. (B) Quantification of average ER lumen width measured from. Data represent average width of 10 randomly selected smooth ER lumens per cell, three measurements per lumen, n=15 cells per treatment. (E) qRT-PCR quantification of gene expression for markers of ER stress from livers dissected from 5 dpf control MO or *tm6sf2_e4 MO* larvae. Data represent average of triplicated values from 2 pooled groups of 40 larval livers each per experiment for two separate experiments. *p<0.05; **p<0.005; *** $p \leq 3x10^{-64}$

Supplemental Figure 7. ER stress markers measured in zebrafish liver after MO-mediated gene knockdown or human TM6SF2 overexpression and Drug Treatment. (A) Expression of ER stress markers in larvae treated 1mg/ml Tunicamycin for 48 hours, (B) Expression of ER stress markers in larvae treated with DMSO for 48 hours, Significant expression differences were observed in *tm6sf2*_e4 MO, overexpression for *chop*, *bip*, and *xbp1*. Values relative to control MO; gene expression normalized to *β-actin*. N=40 larval livers * p≤0.01; ** p≤0.001

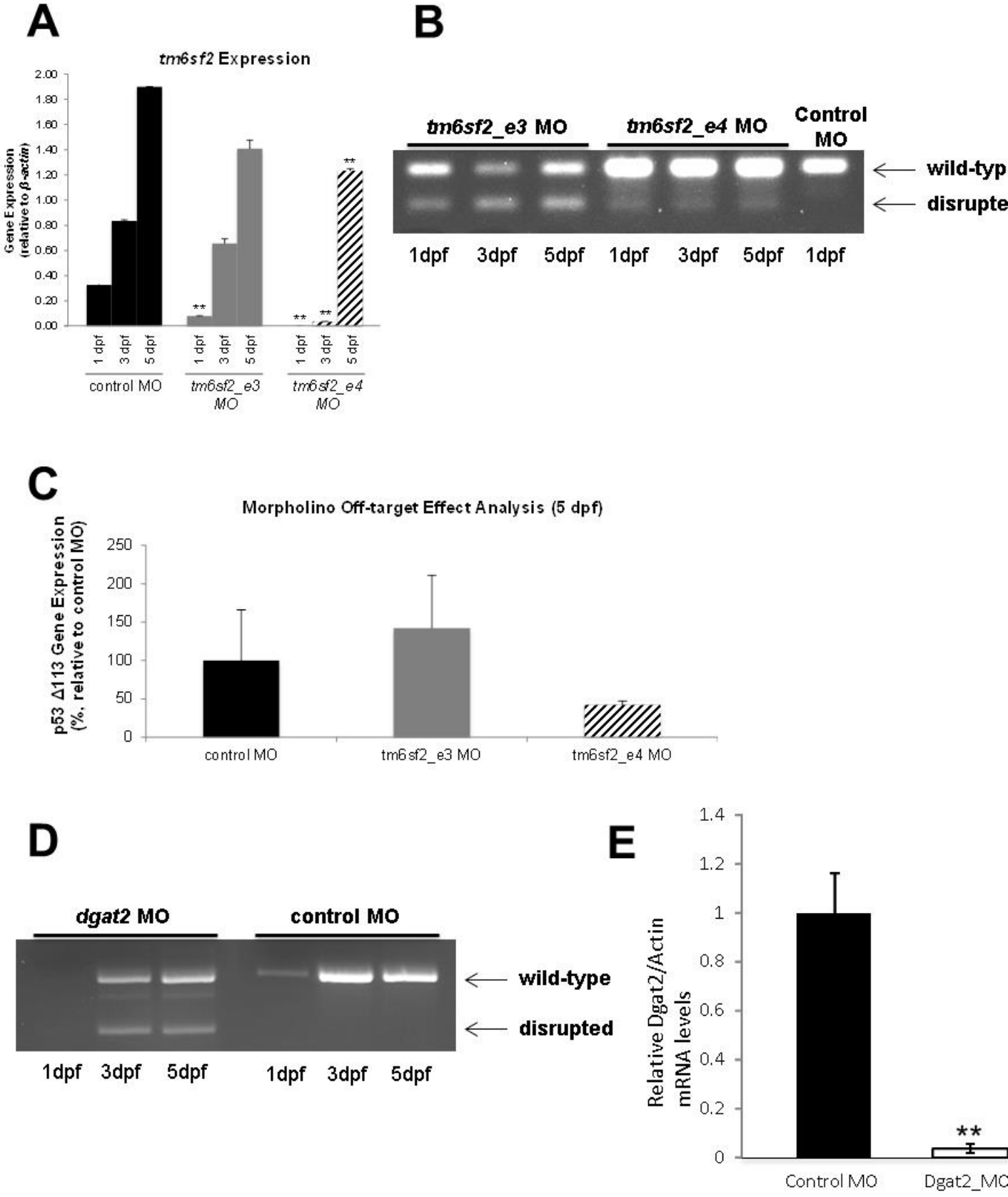
Supplemental Figure 8. Validation of human TM6SF2 over expression in livers extracted from tunicamycin treated zebrafish larvae. Expression of human TM6SF2 mRNA in larvae with or without co-injection *tm6sf2_e4* MO and treated with 1mg/ml Tunicamycin for 48 hours. Significantly increased expression was observed in zebrafish injected with WT mRNA compared to control MO (control), gene expression normalized to *β-actin*. N=20 pooled larvae, triplicated., ** p≤0.001

Supplementary Figure 9. Larval length measurement. Average length of 5 dpf larvae injected with control MO, *tm6sf2*_e4 MO, *Cas9* mRNA only, or *Cas9* mRNA plus gRNA targeted against *tm6sf2* exon 4 (tm6sf2 exon4 CRISPR).

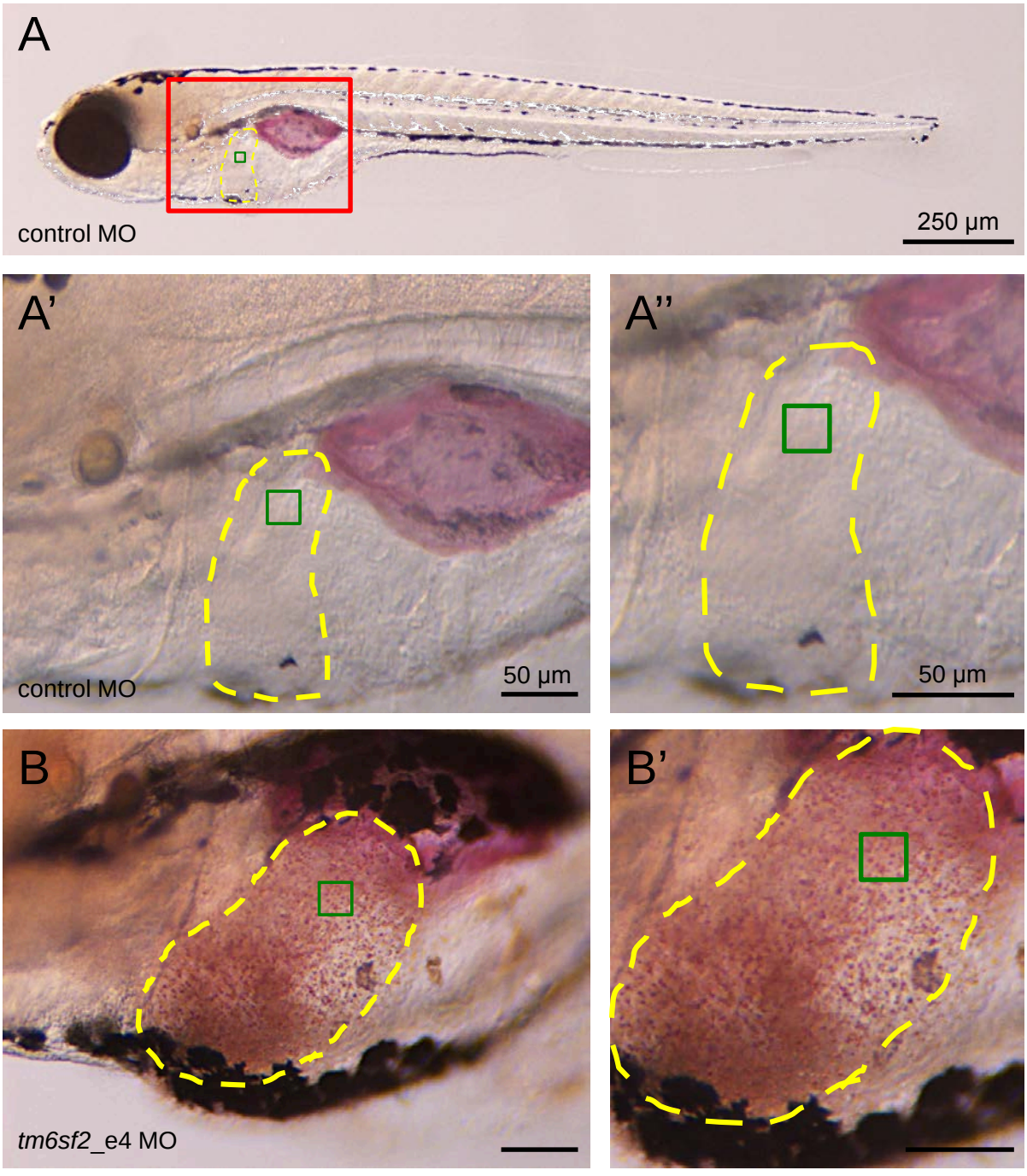
Supplemental Figure 1



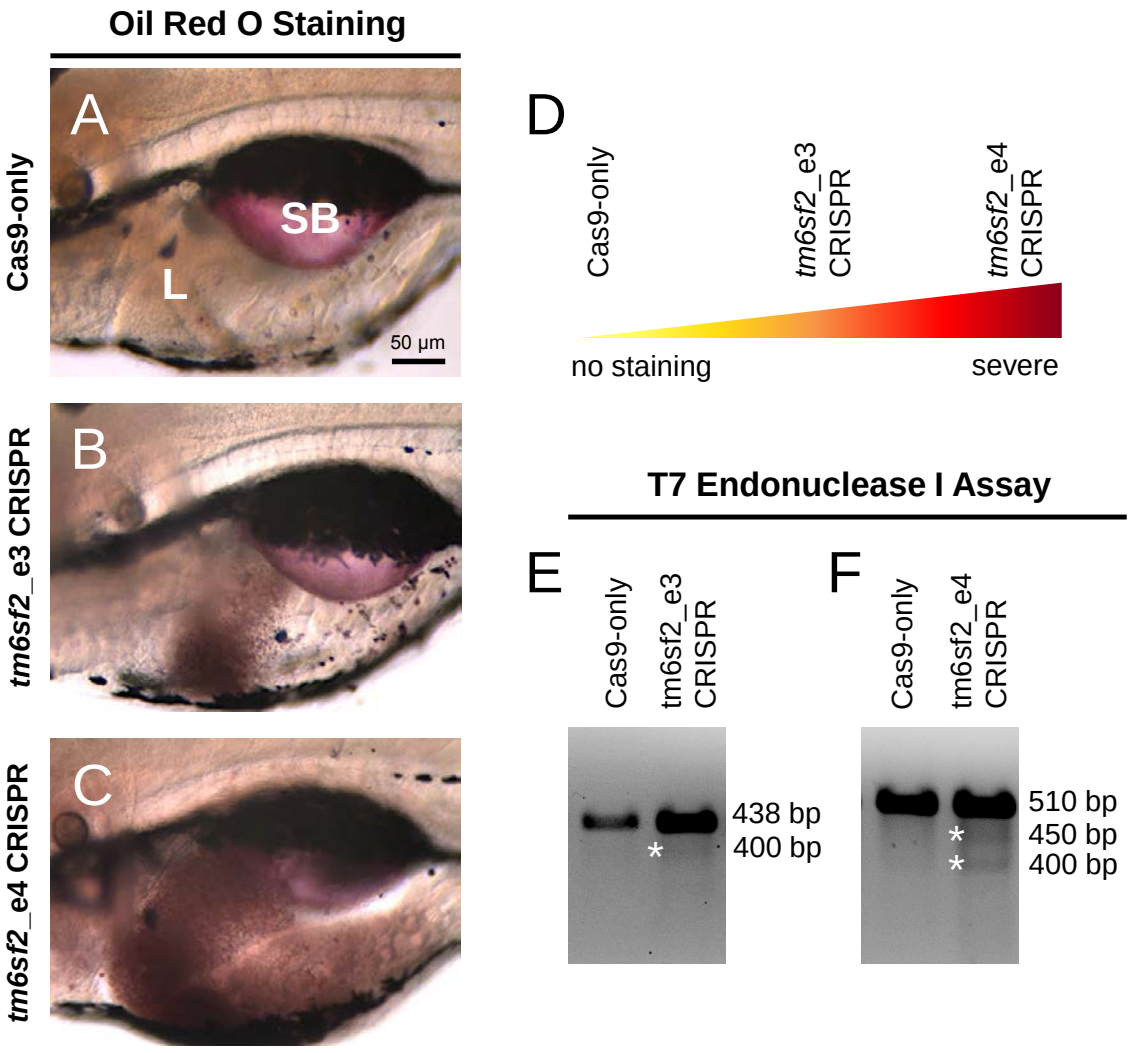
Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4



G

intron 3

```
tm6sf2 wild-type >aaatttgctctaatatttactattgattattgttttgtattttcagGCCAGTGGTTATTCTTGAA>
tm6sf2_e3 CRISPR >aaatttgctctaatattt----------------------------------------------->

tm6sf2 wild-type >ATTGGCGTGGCTGTTCTTGTGGCAGTTTTCCTCCTGGTCTACCTCGCTACGCGCTTCAATCCTCC>
tm6sf2_e3 CRISPR >------------------------------------------------------------------>
```

intron 4

```
tm6sf2 wild-type >TAAGGATCCTTTGTTCTATGgtatgatttgattggttgattcttaaaaacactgaatttatccac>
tm6sf2_e3 CRISPR >----------------------------gttgattcttaaaaacactgaatttatccac>
```

H

intron 4

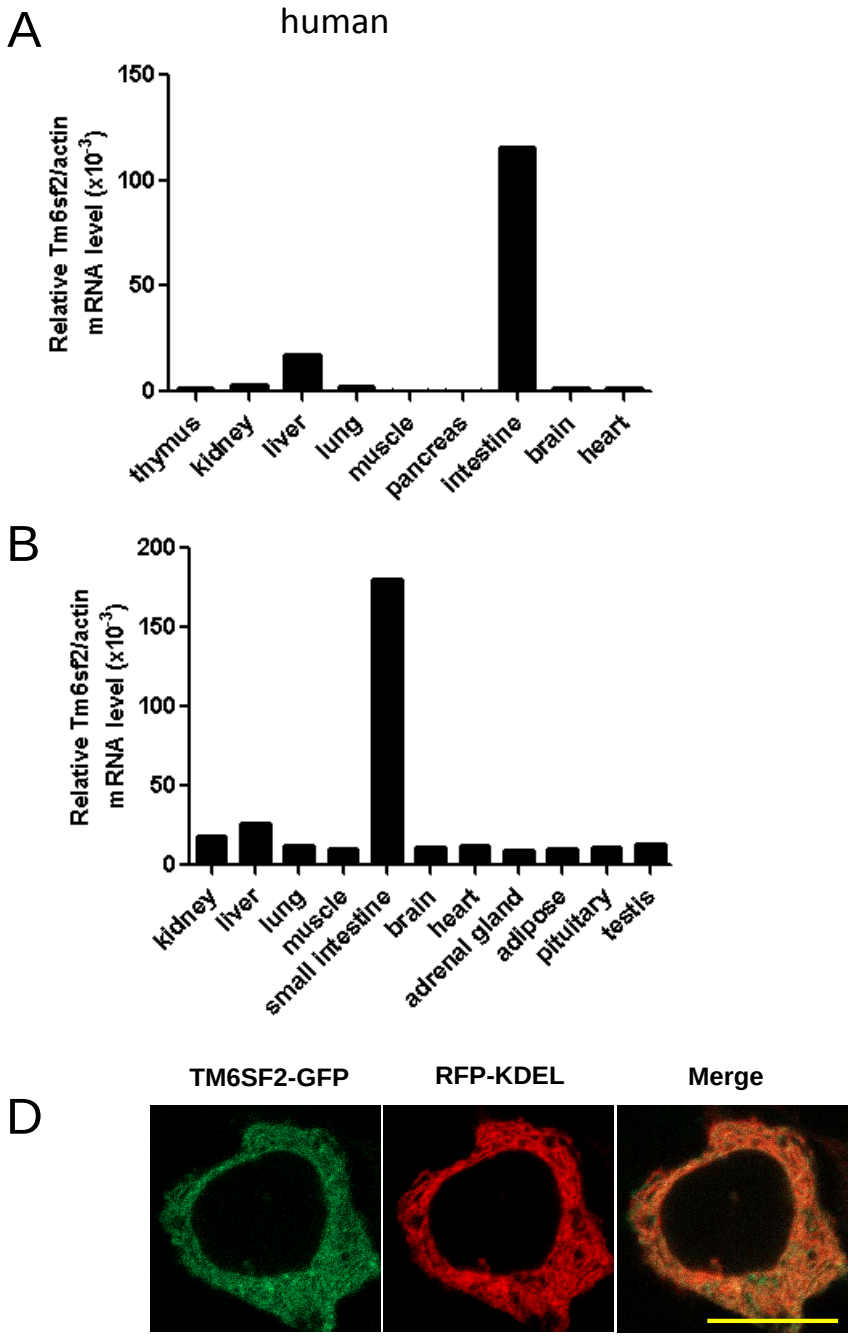
```
tm6sf2 wild-type >taaacaagtattactgtattgactttgtttccagTTTTTGCAGAGTTTTCCTTCACGTGTGTCAT>
tm6sf2_e4 CRISPR >taaacaagtattactgtattgactttgtttccagTTTTTGCAGAGTTTTCCTTCACGTGTGTCAT>

tm6sf2 wild-type >TGACCTCACGAGTGCCCTAGAATATGATGGCTTCGCCTCTGGGTTCATGGAGTTCTACCAGAAAA>
tm6sf2-e4 CRISPR >TGACCTCACGAGTGCCCTAGAATATGATGGCTTCGCCTCTGGGCTCATGGAGTTCTACCAGAAAA>
```

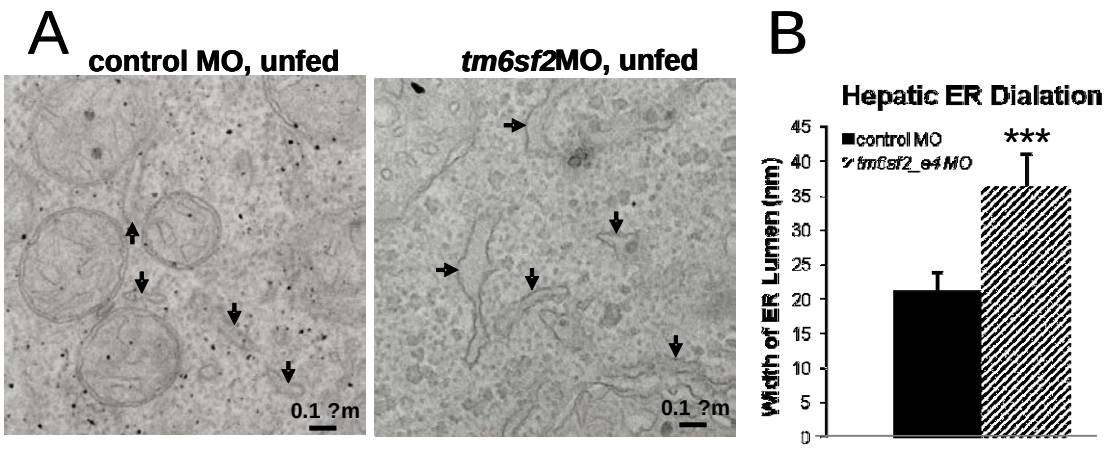
intron 5

```
tm6sf2 wild-type >CGgttagcgtaatgacatgaggagaaagacaagttgatcttatcgattgcagtataaatcttccc>
tm6sf2_e4 CRISPR >CGgttagcgtaatgacatgaggagaaagacaagttgatcttatcgattgcagtataaatcttccc>
```

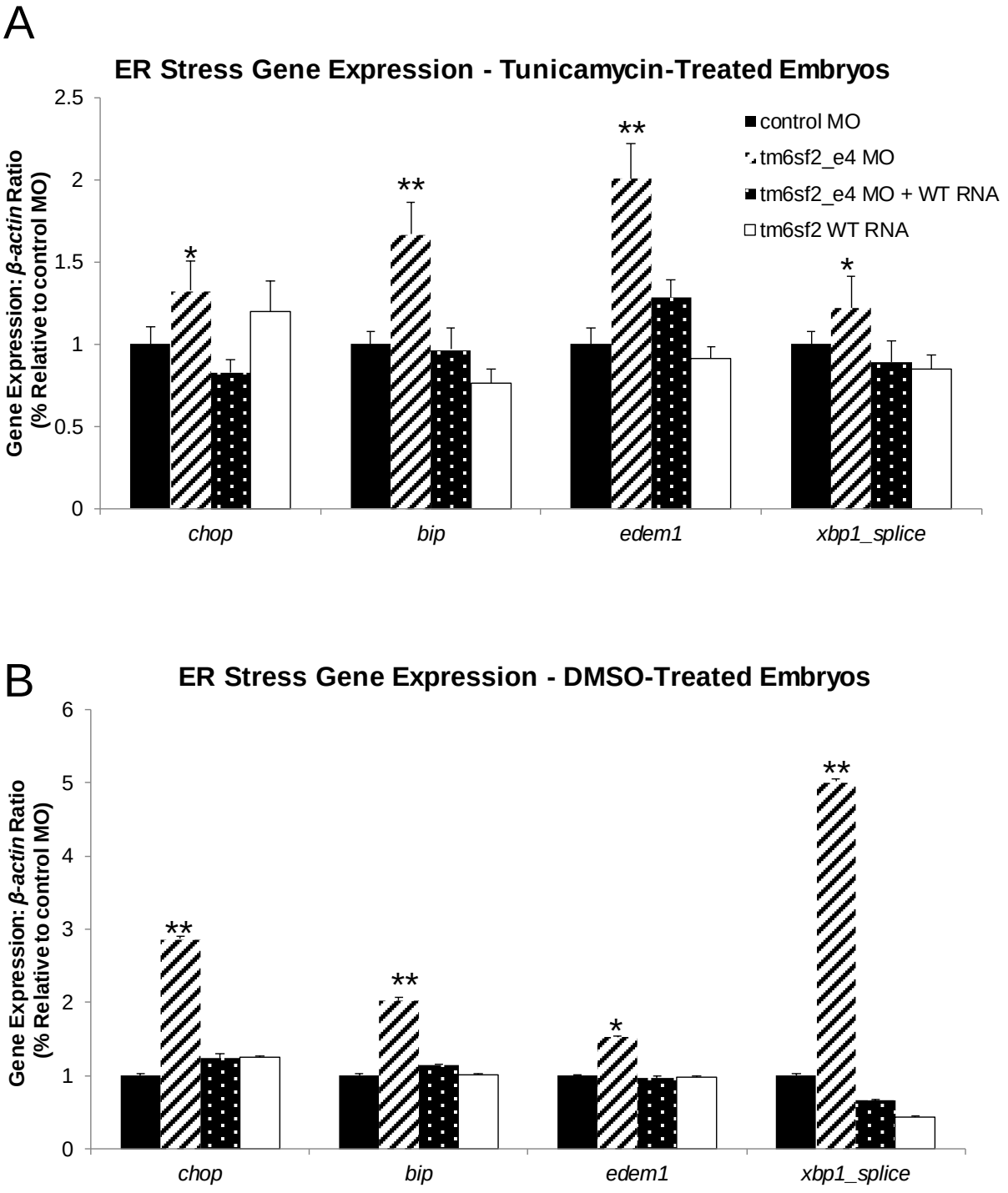
Supplemental Figure 5



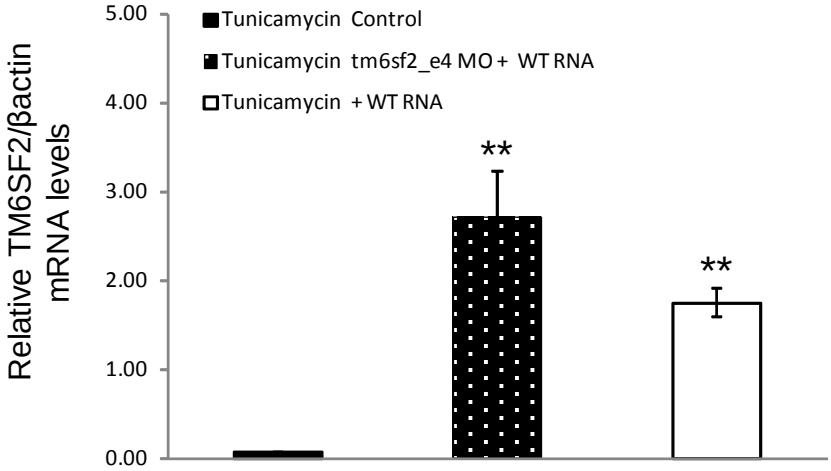
Supplemental Figure 6



Supplemental Figure 7



Supplemental Figure 8



Supplemental Figure 9