Supplemental Information

Human Hepatic Organoids for the Analysis of Human Genetic Diseases

Yuan Guan, Dan Xu, Phillip M. Garfin, Ursula Ehmer, Melissa Hurwitz, Greg Enns, Sara Michie, Manhong Wu, Toshihiko Nishimura, Julien Sage and Gary Peltz

***Preparation and characterization of cultures*.** Although both ALGS subjects had severe liver disease, sequence analysis indicated that they had distinct mutations. ALGS1 had a heterozygous mutation that introduced a termination codon (C829X) into the *JAG1* coding sequence of exon 21. ALGS2 did not have mutations within an exon, but had a novel a G->A mutation at the border of exon 3, which is located within a donor spice site (**Fig. S3A**). This altered *JAG1* transcript splicing; ALGS2 *JAG1* mRNA lacks all 52 bp of exon 3 (**Fig. S3B-C**). Control and ALGS iPSC were indistinguishable in their pluripotent state. ALGS iPSCs retained their differentiation potential; teratomas containing cells derived from the three germ layers (endoderm, mesoderm, and ectoderm) developed after injection under the skin of immunocompromised mice (**Fig. S4**).

***Genome engineering of iPSC*.** CRISPR-associated protein 9 (Cas9) genome editing was used with the piggyBac transposon system to introduce and revert the disease causing mutation in control and ALGS iPS cells, respectively. The CRISPR-Cas9 system has been used to rapidly and efficiently modify endogenous genes in a wide variety of cell types and organisms, which were otherwise challenging to manipulate genetically (1). To minimize off-target activity, a double nicking strategy, which uses the D10A mutant Cas9 nickase (Cas9n) (2), was used to introduce double stranded breaks at the target site. The strategies shown in **Fig. S6A** were used to introduce or revert the ALGS1 mutation (C829X) in control or ALGS1 iPSCs, respectively. To modify the C829X locus in ALGS1, we designed two pairs of sgRNAs, and tested their efficiency in the iPS cells (**Fig. S6B**). After transposon-mediated selection, we obtained both heterozygous and homozygous clones after the first round of puromycin selection (**Fig. S7A, D**). Since ALGS is an autosomal dominant genetic disorder, only heterozygous clones were selected; and these were expanded for transposon excision using a second round of drug selection. The same strategy was used to revert the mutation present in ALGS1 iPSC (X829C). Homozygous revertant clones were selected, and these were expanded for a second

round of transposon excision by drug selection (**Fig. S7G**). Clones with the excised transposon were selected by growth in media containing 10 μM gancyclovir for 1~2 weeks. Clones with both piggyBac cassettes removed were selected for further testing (**Figs. S7B, E, H**). However, after the first round of selection, in the heterozygote clones the ALGS wild type allele was replaced by the piggyBac cassette, which left a TTAA sequence in the wild type allele; but the mutated allele was unaltered. The tagged clones were selected as genome engineered control iPSC to determine if an off-target mutation altered a response in the phenotypic assays (Fig. **S7I**). The alterations in all engineered iPSC were confirmed by sequencing (**Figs. S7C, F, J**).

Materials and Methods

Teratoma formation assay. iPSC colonies were generated on gelatin-coated plates in mTESR complete media (Stem Cell Technologies) and passaged with dispase treatment and gentle mechanical disruption to maintain clusters of cells. A portion of the cells was further disassociated to enable counting. iPS Cells were injected subcutaneously into either the flank or the mid-scapular region of anesthetized immune-compromised Nod-SCID-Il2Gamma (NSG) mice. The injected areas were examined 3 weeks after injection. iPS cells from Patient 1 formed teratomas after injection of 3,000,000 cells per injection, and iPS cells from Patient 2 formed teratomas after injection of 500,000 cells per injection. When tumors were detected, the mice were euthanized and the teratomas were dissected away from the surrounding tissues, and fixed in 4% formaldehyde solution at least overnight then dehydrated and embedded in paraffin. 5 μm sections of each teratoma were obtained and stained with hematoxylin or to detect germ-layer derivatives. Ectoderm and mesoderm were detected using DAB-immunohistochemistry for Cytokeratin 10 (Santa Cruz) and Smooth Muscle Actin (Sigma), respectively, and were counterstained with hematoxylin. Endoderm was detected using Alcian blue staining for mucin-producing goblet cells and counterstained with Safranin O. Photography was performed using a Leica DM2000 microscope with DFC500 camera and Leica LAS software.

Notes about the protocol for producing HO1. Since endodermal conversion is an extremely important step for hepatic lineage commitment, we initially compared the efficiency of conversion of iPSC into endoderm over a three day period in complete ABCFL medium (A, Activin; B, BMP4; C, CHIR99021, F, FGF2; L, LY294002) relative to that in media containing various other combinations of growth factors over a 3 day period (e.g. ABCFL-ABFL-ABF media). Since a higher percentage (>90%) of $SOX17^{+}/FOXA2^{+}$ endodermal cells were obtained in ABCFL medium over the 3 day period, this medium was used in our protocol. We also found

that FGF10 was superior to FGF4 (3) for inducing posterior foregut differentiation. Of note, we also tested several other small molecules that were reported to improve hepatic differentiation by measuring the expression level of TBX3 and HNF4A(4). Although retinoic acid and IWR-1 improved hepatic differentiation in the organoid cultures, they dramatically decreased cell viability. Therefore, these chemicals were not included in our differentiation protocol.

Drug metabolism. Two control organoids after 50 days in culture, and primary human hepatocyte cultures (24 hrs after plating at a density of $2x10^6$ cells) were incubated with clemizole (3.6 µg/ml) for 24 hrs. The amounts of two clemizole metabolites (M1, M6), which are produced primarily by a CYP3A4-dependent reaction, in the supernatants were measured by LC/MS analysis (5). Each bar is the average of 3 independent measurements, and was normalized relative to the total amount of cellular protein, which was determined by using the BCA Protein Assay (Pierce, Grand Island, NY).

Bile Acid measurement. Supernatants were collected after 48 hours of incubation from three control organoid cultures (day 50), and from 3 control iPSC lines. The supernatants (100 µl) were extracted with 4 volumes of cold acetonitrile; and the mixtures were incubated at -20 ^{0}C for 1 hour before centrifugation at 12,000 rpm for 10 minutes. The supernatants were dried and re-suspended in a 50/50 methanol/water solution. The extracts were then analyzed according to published methods (6, 7) on an Agilent Infinity 1290 UHPLC system with an Eclipse Plus RRHD C18 column (1.8µm, 2.1x100mm, Agilent, Santa Clara CA). The mobile phases were 0.2% acetic acid in water (mobile phase A) and 0.2% acetic acid (mobile phase B) in acetonitrile. The gradient was performed at the flow rate of 0.4ml/minute, with the gradient as follows: after the sample injection, held at 20% B for 0.5 minutes, then increased to 80% B in 25 minutes, final increased to 95% B for 3 minutes and back to 20% B to balance the column to initial condition. Mass spectrometry was performed using a QTOF 6520 (Agilent, Santa Clara USA), with an ESI source in negative mode, and only precursor ions of bile acids were monitored. The source parameters were set at: gas temperature 325 ^{0}C, gas flow 8.0 l/minute, and nebulizer 35 psi. The data was analyzed by Masshunter qualitative analysis software. A “find by formula” method was used to extract molecular features with a formula related to free bile acids, or their glycine or taurine conjugated derivatives. Feature abundance was represented by their peak height. The identification of the extracted molecular features was confirmed by their retention time and accurate mass within 10ppm relative to the following chemical standards (all from Steraloids, Inc. (Newport, RI)): Cholic Acid (CA), chenodeoxycholic acid (CDCA), Deoxycholic Acid (DCA),

Lithocholic Acid (LCA), Ursodeoxycholic acid (UDCA), glycocholic acid (GCA), glycochenodeoxycholic acid (GCDCA), glycodeoxycholic acid (GDCA), glycolithocholic acid (GLCA), glycoursodeoxycholic acid (GUDCA), taurocholic acid (TCA), taurochenodeoxycholic acid (TCDCA), taurodeoxycholic acid (TDCA), taurolithocholic acid (TLCA), and tauroursodeoxycholic acid sodium (TUDCA).

Rhodamine 123 transport assay. This assay measures multidrug resistance protein-1 (MDR1)-dependent transport by cholangiocytes of a substrate (rhodamine 123) into the ductal lumen. The MDR1 transporter is expressed on the apical luminal membrane of cholangiocytes. Day 25 HO1 generated from control or CRISPR-modified iPSCs were incubated in culture medium containing 100 μM rhodamine 123 (Sigma-Aldrich) for 10 min. The organoids were then washed with fresh medium three times. To inhibit MDR1-dependent transporter activity, some cultures were incubated with 20 μM verapamil (Sigma-Aldrich) for 30 min before addition of rhodamine 123. The fluorescent organoids were imaged with a Leica sp8 confocal microscope (Excitation 488 nm, Emission 529 nm). Ten pictures of each organoid were taken along the z-axis to stack the whole sphere. Then, the measured mean intraluminal fluorescence intensity of rhodamine 123 was normalized relative to the area of whole HO1. Each experiment was repeated in triplicate. The mean fluorescence intensity comparisons were performed using ANOVA analysis.

Measurement of organoid and vesicle formation. The formation of vesicles and organoids in day 30 HO1 cultures was determined by observing a well of a 6-well culture plate under a bright field microscope (TS100, Nikon, Melville, NY). The number of vesicles and intact organoids was assessed in 3 independent cultures of each type, and each had 100 or more countable spheres.

RNA-seq and data analysis. To obtain high-quality cDNA libraries for transcriptome analysis, we applied SMART-Seq v4 Ultra Low RNA Kit (Takara Bio USA, 635025) following the manufacturer's instruction. The size distribution and concentration of the cDNA was assessed on Agilent 2100 Bioanalyzer using a high sensitivity DNA Kit (Agilent Technologies, Santa Clara, CA, 5067). The sequencing libraries were constructed using the Nextera XT DNA Sample Preparation kit according to the protocol (Illumina, San Diego, CA, FC-131-1096) with dual barcoded indices (Illumina; FC-131-1002). Library cleanup and pooling was performed using AMPure XP beads (Agencourt Biosciences; A63880). The libraries were pooled and sequenced 150 bp paired-end on one lane of Illumina HiSeq 4000.

For analysis of the RNA-Seq data, the HO1 and HO2 data were separately analyzed; and then each dataset was compared with that obtained from iPSCs, PHHs and human liver tissue. This enabled us to better characterize the relationship between HO1 (or HO2) and the other samples without the confounding effects of the data from the HO2 (or HO1) samples. In each analysis, the original counts per million (CPM) values from the RNA-seq data were analyzed by Morpheus (software.broadinstitute.org/morpheus); and the software usually selects 500 genes with the largest median absolute deviation (**MAD**) values for subsequent analysis. The Morpheus software selected 2458 and 1510 genes for the HO1 and HO2 samples, respectively, due to ties in the MAD values. Of note, the MAD values are calculated using the data contained in all samples; so the 2458 (or 1510) genes were selected in an unbiased manner without reference to sample origin (i.e. irrespective of whether they were differentially expressed between cell types of interest). Hierarchical clustering of the cell types and genes were calculated using the Euclidean distance algorithm. The genes that made the highest contribution to these comparisons are identified by this method, and their expression levels were confirmed by RT-PCR analysis. The RNA-Seq data have been deposited in the NCBI's Gene Expression Omnibus, and are accessible through GEO Series accession number GSE101770 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE101770)

RT-PCR analyses. Total RNA was extracted using TRIzol® RNA Isolation Reagents (Ambion, Grand Island, NY), and 2 μg RNA was reverse-transcribed using the iScript™ Advanced cDNA Synthesis Kit (BIO-RAD, Hercules, CA) according to the manufacturer's guidelines. The following TaqMan primer sets were obtained from Life Technologies (Grand Island, NY) and used for gene expression analyses by PCR: SOX17, Hs00751752_s1; CCXR4, Hs00607978_s1; GATA4, Hs00171403_m1; FOXA2, Hs00232764_m1; AFP, Hs00173490_m1; albumin, Hs00910225_m1; SERPINA1 (A1AT), Hs01097800_m1; CYP2C9, Hs02383631_s1; CYP1A1, Hs01054797_g1; HNF1A, Hs01054797_g1; FOXA1, Hs04187555_m1; JAG1, Hs01070032_m1; HES1, Hs01070032_m1 ; HEY1, Hs01114113_m1; HNF4A, Hs00230853_m1; HHEX, Hs00242160_m1; HNF1B, Hs01001602_m1; SOX9, Hs01001602_m1.

CRISPR-mediated genome engineering. The sequences of the site-specific guide RNAs (sgRNAs) were selected using the CRISPR Design Tool (8). Oligonucleotides with these sequences were cloned into the pSpCas9n(BB)-2A-Puro (PX462) vector (Addgene, Cambridge,

MA), which uses the D10A mutant Cas9 nickase (Cas9n) to minimize off-target activity (2). For Homology-directed repair (HDR), the piggyBac transposon was used with puromycin followed by gancyclovir selection to obtain cells with desired genome modifications. The following plasmids were obtained from the Sanger Institute: pPB-R1R2_NeoPheS, pENTR-PGKpuroΔtk and pCMV-hyPBase (hyperactive mammalian-codon-optimized piggyBac transposase). To facilitate introduction of our designed sequences into the vectors, the piggyBac vector was modified by inverting the fragment, which is located between the original inverted terminal repeat sequences. To construct the mutation introducing and reversion targeting vectors, two 1 kb genomic fragments on both sides of the C829X mutation were PCR amplified and cloned into the piggyBac transposon vector using In-Fusion HD Cloning Kit (Clontech, Mountain View, CA). A TTAA site, which is used for the insertion of the piggyBac transposon, was created by introducing silent mutations near the intended modification site. TGA and TGT sequences were added to the forward primers in both the mutation inducing and reversion fragments, respectively. To confirm that the targeted mutations were introduced, genomic DNA was isolated from the engineered iPSC after they were cloned and exposed to two rounds of drug selection. PCR products from the targeted loci were analyzed using the T7E1 enzyme assay, and were then directly sequenced or cloned into a CloneJET PCR (Thermo Fisher Scientific, Grand Island, NY) cloning vector according to the manufacturer's instructions. Plasmid DNA was isolated from 10 bacterial clones and analyzed by Sanger sequencing (Sequetech, Mountain View, CA) to determine if bi-allelic gene alterations were introduced.

The strategy shown in **Fig. S6A** was used to introduce the ALGS1 mutation into control iPSCs. The structure of the *JAG1* gene, and the C829X mutation is located in exon 21. The sequences of sgRNA1 and sgRNA2 in the piggyBac targeting vector, and of the resulting mutated allele (after excision of the introduced piggyBac cassette) are shown. After the C829X mutation was introduced into control iPSC, the clones were first grown in the presence of puromycin to select for those with piggyBac transposons. PCR analysis was used to select clones with a heterozygous mutation. Clones with a heterozygous mutation were transfected with the PBase vector, and were grown in the presence of gancyclovir to select for clones from which the piggyBac transposon was removed. Briefly, ANML and GJ4 control iPSCs were pelleted and mixed with two CRISPR sgRNA expression vectors and piggyBac donor template in 100 ul of human ES cell solution 1 (Lonza) using program A23. After transposon-mediated selection in puromycin (1.0 ug/ml), we obtained both heterozygous and homozygous clones after the first round of puromycin selection. Since ALGS is an autosomal dominant genetic disorder, only

heterozygous clones were selected, and then expanded for the second round of transposon excision. The strategy also shown in **Fig. S6A** was used to revert the mutation present in ALGS1 iPSC. Only homozygous revertant clones were selected, and expanded for a second round of transposon excision. Clones with the excised transposon were selected by growth in media containing 10 μM gancyclovir for 1~2 weeks. Clones with both piggyBac cassettes removed and selected for further testing. However, after the first round of selection, in the heterozygote clones the ALGS wild type allele was replaced by the piggyBac cassette, which left a TTAA sequence in the wild type allele, while the mutated allele was unaltered. So those tagged clones were selected as genome engineered control iPSC to exclude an off-target effect in the phenotypic assays.

The same strategy was used to introduce the TOF1 (*G274D*) mutation. Heterozygous clones were selected after the first round of puromycin selection; and these were expanded for transposon excision using a second round of drug selection. Clones with piggyBac cassettes removed were selected for further testing (**Fig. S9**).

Histology and immunostaining of liver tissue. Control and Alagille **l**iver tissues were kindly provided by Dr. Elizabeth B. Rand from Children's Hospital of Philadelphia (CHOP). The normal liver tissue was isolated from liver lobes that were resected due to liver cancer. Regions with normal appearing tissue, which were spatially distant from the cancerous areas, were collected. Liver tissues were collected from three different subjects with Alagille liver disease that were 3, 5, and 15 years of age. These individuals were followed in the clinic at CHOP for treatment of Alagille Syndrome. Liver tissues were flash frozen on dry ice in Tissue-Tek® O.C.T™ (Sakura Finetek U.S.A., Torrance, CA). Tissue blocks were sectioned using a Leica CM3050 S Cryostat into 10 μm sections, and serially obtained sections were stained. Liver tissues were fixed in 4% paraformaldehyde for 10 minutes, followed by permeabilization with 1% Triton X-100 (Sigma-Aldrich, St. Louis, MO), and blocking was performed with a solution containing 10% chicken serum (Jackson ImmunoResearch, Bar Harbor, MA) for 30 min. Then, the sections were incubated with the primary antibodies listed in **Table S1**. When needed, secondary staining was performed with the appropriately labeled anti-antibodies in 10% chicken serum.

Preparation of mouse embryo tissue. Embryonic tissues were obtained from female C57BL/6J mice with timed-pregnancies. For frozen section preparation, fresh tissue was directly embedded in OCT (Tissue-Tek), snap frozen on dry ice and cut at 10 μm thickness. For

immunostaining, frozen slides were thawed to room temperature and fixed with 4% PFA in PBS. The slides were then blocked with a solution containing 10% normal chicken serum in PBS buffer with 0.5-1%Triton-X for 10 minutes (Jackson ImmunoResearch, West Grove, PA). The primary antibodies are listed below.

References

1. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339(6121):819-23.
2. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 2013;154(6):1380-9.
3. Cai J, Zhao Y, Liu Y, Ye F, Song Z, Qin H, Meng S, Chen Y, Zhou R, Song X, et al. Directed differentiation of human embryonic stem cells into functional hepatic cells. *Hepatology.* 2007;45(5):1229-39.
4. Zorn AM, and Wells JM. Vertebrate endoderm development and organ formation. *Annu Rev Cell Dev Biol.* 2009;25(221-51.
5. Nishimura T, Hu Y, Wu M, Pham E, Suemizu H, Elazar M, Liu M, Idilman R, Yurdaydin C, Angus P, et al. Using Chimeric Mice with Humanized Livers to Predict Human Drug Metabolism and a Drug-Drug Interaction *Journal of Pharmacology And Experimental Therapeutics.* 2013;344(2):388-98.
6. Garcia-Canaveras JC, Donato MT, Castell JV, and Lahoz A. Targeted profiling of circulating and hepatic bile acids in human, mouse, and rat using a UPLC-MRM-MS-validated method. *J Lipid Res.* 2012;53(10):2231-41.
7. Alnouti Y, Csanaky IL, and Klaassen CD. Quantitative-profiling of bile acids and their conjugates in mouse liver, bile, plasma, and urine using LC-MS/MS. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2008;873(2):209-17.
8. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* 2013;31(9):827-32.

Supplemental Table

Table S1. The primary antibodies used for immunohistochemistry and their sources.

Antibody	Manufacturer	Catalogue
SOX17	BD Pharmingen™	561590
Anti-DLK antibody [MM0514-9D8]	Abcam	ab89908
CK18	DAKO	M 7010
Cytokeratin 7	DAKO	M701801-2
Cytokeratin 19	DAKO	M088801-2
Anti-E-Cadherin	BD	610182
pan-Cytokeratin Antibody (C11):	Santa Cruz	sc-8018
Purified anti-human CD326 (EpCAM) Antibody	Biolegend	324202
Anti-Tbx3 antibody	Abcam	ab89220
Anti-Glutamine Synthetase	Sigma	G2781
Monoclonal Anti-a Smooth Muscle Actin	Sigma	A2547
ZO-1	LIFE	339100
Anti-Fibroblasts Antibody, clone TE-7	Chemicon	CBL271
Jagged1	DSHB	TS1.15H
Notch2	DSHB	C651.6DbHN
Anti-Cytokeratin 8 antibody [EP1628Y]	Abcam	ab53280
Anti-Sox9 Antibody	Millipore	AB5535
α-1-Antitrypsin	DAKO	A001202
Non-phospho (Active) β-Catenin	Cell Signaling Technology	8814
β-Catenin (D10A8) XP®	Cell Signaling Technology	8480
ZO-1	life	402200
CFTR Antibody (H-182)	Santa Cruz	sc-10747
polyclonal Rabbit Anti-Human Alpha-1-Fetoprotein	DAKO	A000829-2
Anti-activated Notch1 antibody	Abcam	ab8925
Ki-67 Antibody (H-300):	Santa Cruz	sc-15402
Jagged1 Antibody (H-114):	Santa Cruz	sc-8303
Human Jagged 1 Affinity Purified Polyclonal Ab	R&D	AF1277-SP
FOXA2 HNF6b	SantaCruz	sc-6554
HNF-1β Antibody (C-20): sc-7411	SantaCruz	sc-7411
TBX3 (A-20)	SantaCruz	sc-17871
Human Albumin Antibody	Bethyl	A80-129A
Jagged1 Antibody (C-20):	SantaCruz	sc-6011
HNF4a (C-19)	SantaCruz	sc6556

Supplemental Figure Legends

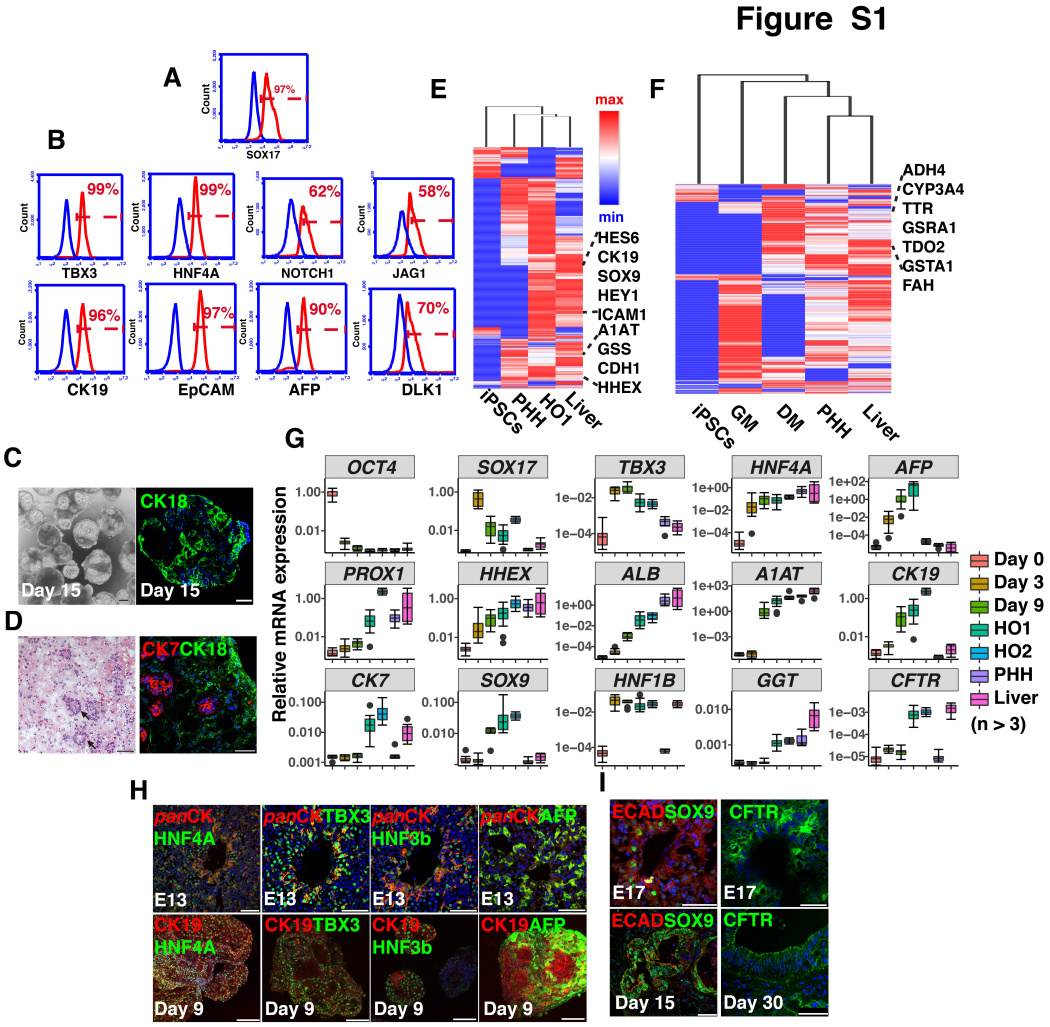


Figure S1. (A) Flow cytometric analyses showing the proportion of SOX17^{+} cells in HO1 on day 3. The red and blue lines show the cells incubated with the anti-SOX17 and isotype control antibodies, respectively. These results are representative of 3 technical replicates performed on four separately prepared samples. **(B)** Flow cytometric analyses showing the proportion of TBX3, HNF4A, NOTCH1, JAG1, CK19^{+}, AFP^{+}, EpCAM^{+}, and DLK^{+} cells present in control HO1 on day 9. The red and blue lines show the cells incubated with the indicated antibody or with an isotype control antibody, respectively. These results are representative of 3 technical replicates performed using four separately prepared samples. **(C)** Bright field and immunostained frozen sections prepared from control HO1 on day 15, which is when they express the hepatocyte marker CK18. **(D)** H&E staining shows ductal structure emerging at early day 15 HO1. The arrows indicate where duct-like structures are located. Immunofluorescent images of HO1 after 15 of differentiation indicate that they express hepatocyte (CK18) and cholangiocyte (CK7) markers. **(E)** Whole transcriptome analysis of HO1 cultures. The heatmap shows the results of a Euclidean hierarchical clustering analysis of 2458 genes that are highly expressed in liver tissue, PHH or HO1. The essential genes for bile duct function (upper dash lines, CK19 and SOX9) are only present in HO1 and in liver tissue, but not in PHH or iPSCs. **(F)** Whole transcriptome analysis of iPSCs, human liver and of HO2s cultured in GM and DM. The heatmap indicates the results of a Euclidean hierarchical clustering analysis of genes that are

highly expressed by HO2 in DM. This cluster contains essential genes for mature liver function; and their expression is increased after HO2 are cultured in DM (indicated between the dashed lines. The blue and red colors indicate genes that are down- or up-regulated, respectively. **(G**) Quantitative RT-PCR analyses showing the dynamic changes in the expression level for 15 mRNAs during HO development. The level of mRNA expression was measured in day 0 (iPSCs), day 3 (EN), day 9 (HB), day 20 (HO1), and day 62 (HO2) cultures; and in PHH and in human liver tissue. The expression levels are shown relative to *GAPDH*. The values represent the mean ± SD of three experiments performed using organoids produced from three different control lines, PHH or liver tissue. **(H, I)** Immunofluorescent images comparing the antigens expressed by day 9 HO1 and embryonic day 13 (E13) mouse liver; and those expressed by day 15 and day 30 HO1 cultures and E17 liver. The scale bars are 50 μm in all panels.

Figure S2

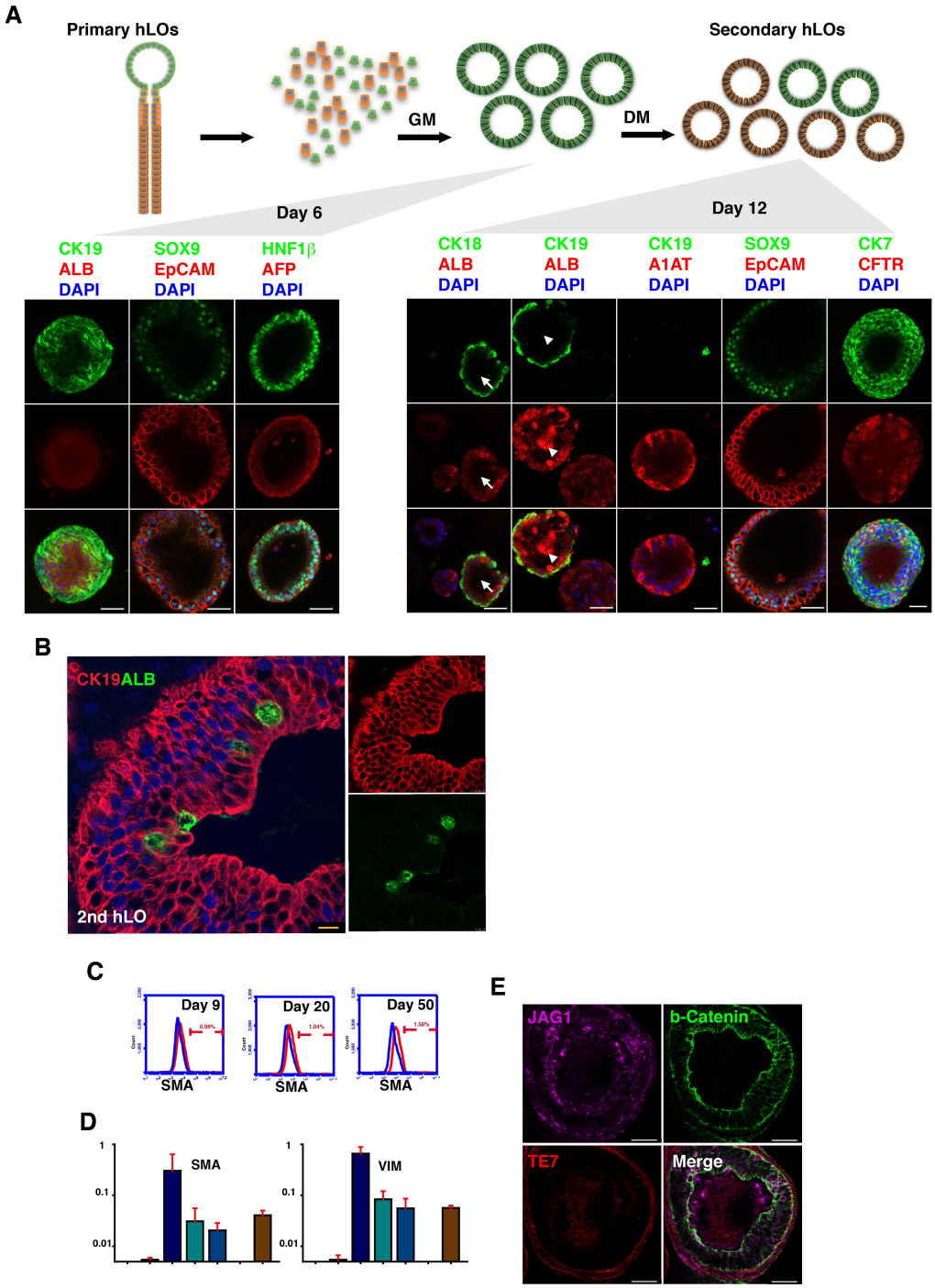
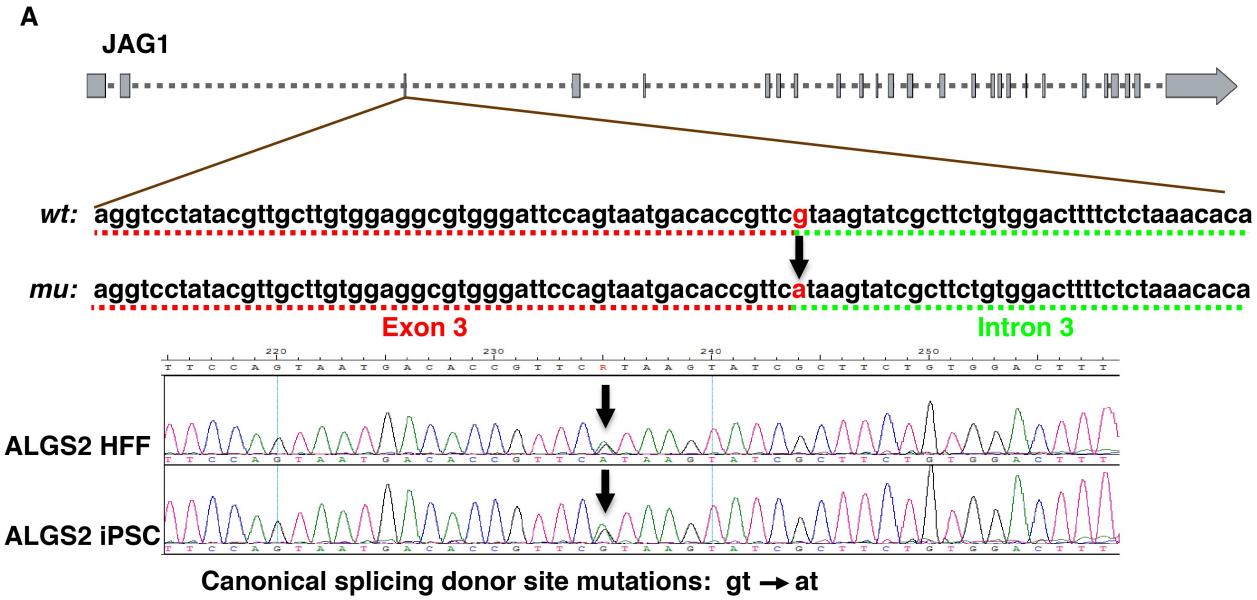
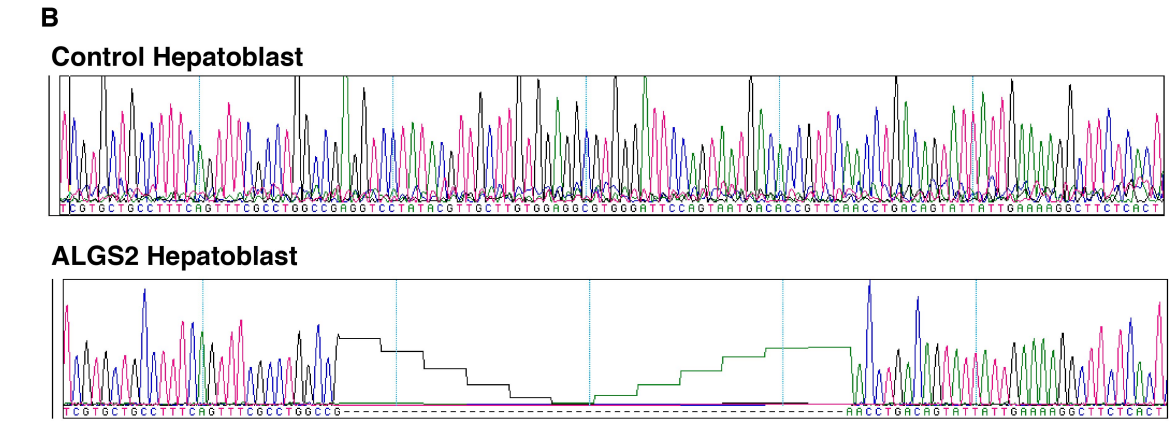


Figure S2. (A) A schematic representation of the organoid regeneration system used to form HO2 is shown. The cells in HO1 are dissociated, and then grown for 6 days in growth media (GM). Then, the cells are grown as clonal populations for 6 more days in matrigel and differentiation media (DM) to form HO2. The bottom panels show immuno-fluorescence staining of HO2 cultures with the indicated antibodies on day 6 (left) and day 12 (right). The cultures were counter-stained with DAPI to indicate the nuclei. Each column shows a different organoid, the bottom panels are the merged images, and the scale bars are 50 μM. The cells in HO2 on

day 6 are $CK19^+SOX9^+HNF1\beta^{-}$ and only weakly express Albumin. On day 12, the HO2 developed a mixed population of cells that consisted of: hepatocytes ($ALB^+/A1AT^+$) and cholangiocytes ($CK7^+/CFTR^+$), along with a population of $SOX9^+/EpCAM^+$ hepatic progenitor cells. In the day 12 images, the arrow indicates a $CK18^+/ALB^+$ hepatocyte; and the arrowhead indicates $CK19^+/ALB^+$ progenitor cells. (**B**) Immunostaining of day 12 HO2 indicates that ALB^+ cells are located within a ductal structure with $CK19^+$ cells. Scale bar, 10 μm. (**C**) Flow cytometric analyses showing the proportion of SMA^+ cells present in day 9, day 20 and day 50 control HO1. The lines show the profile for cells that were incubated with anti-SMA (red) or isotype control (blue) antibodies. These results are representative of 3 technical replicates, which were performed using four separately prepared samples. (**D**) Quantitative RT-PCR analyses showing the dynamic changes in the expression level of mesenchymal cell antigens (*smooth muscle actin (SMA)* or vimentin (*VIM*) in day 0, 3, 9, 20 and 50 organoid cultures. (**E**) Immunofluorescent images showing the expression of JAG1, TE7 and β-Catenin in day 50 control HO1.

Figure S3





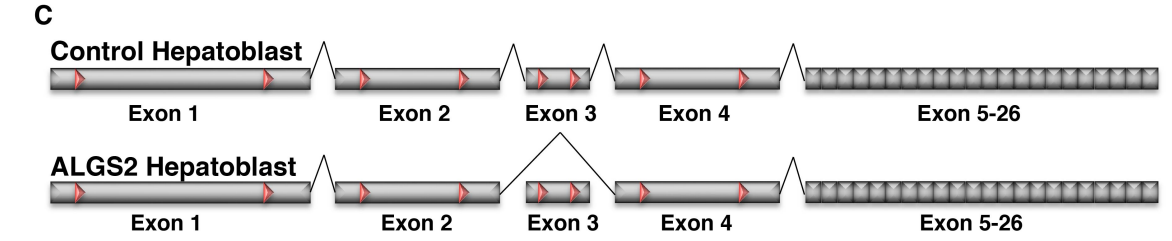


Figure S3. (**A**) Sequencing of *JAG1* genomic DNA in ALGS2 identifies a de novo mutation at the boundary between exon 3 and intron 3. The conversion of GT in controls to AT in ALGS2 alters a canonical splicing donor site. **(B,C)** Sequencing of *JAG1* mRNA in ALGS2 iPSC reveals that all of exon 3 was deleted. A mutation at the donor splice led to the production of a *JAG1* mRNA transcript in ALGS2 that lacked exon 3, but had all of the other exons.

Figure S4

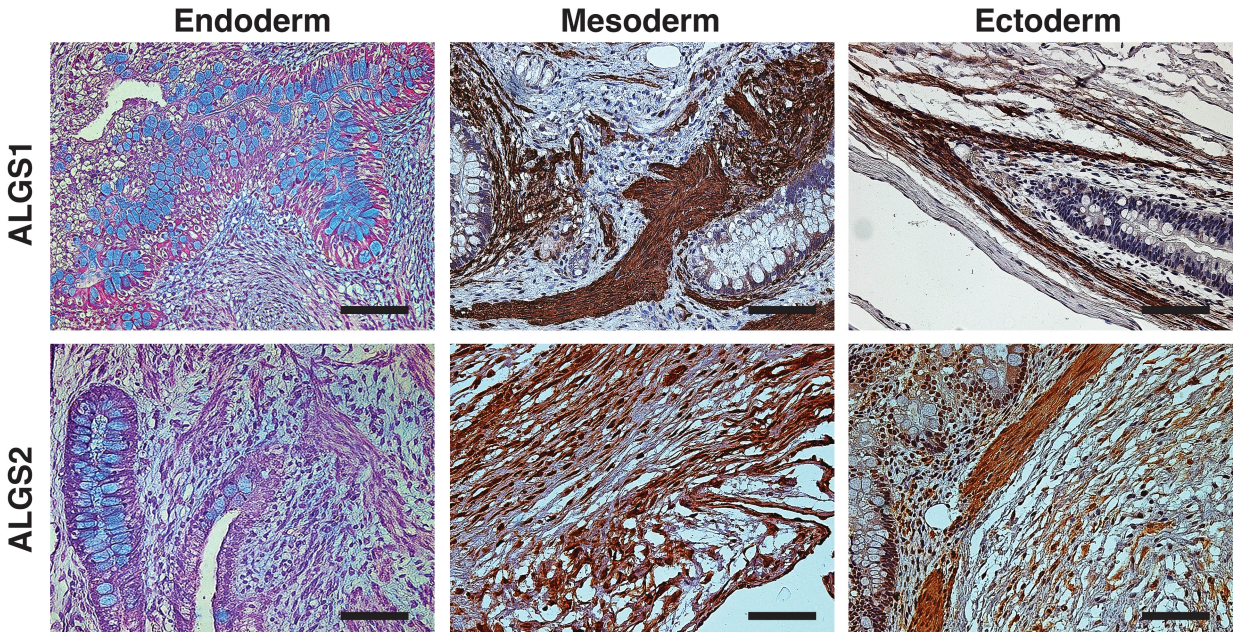


Figure S4. ALGS iPSCs are competent to form teratomas. The images obtained from mice injected with ALGS1 and ALGS2 iPSC are shown. The teratomas formed three weeks after ALGS iPSC were injected under the skin of immunocompromised SCID mice contained cells derived from three lineages (endoderm, mesoderm, and ectoderm), which indicates that the iPSC retained their differentiation potential. Endodermal differentiation was visualized by Alcian blue staining of mucin-producing goblet cells (in left-most column). Immunohistochemical detection of smooth muscle actin and cytokeratin 10 was used to detect the presence of mesodermal (middle column) and ectodermal (right-most column) tissue, respectively. Representative images are shown. Scale bars = 100 μm.

Figure S5

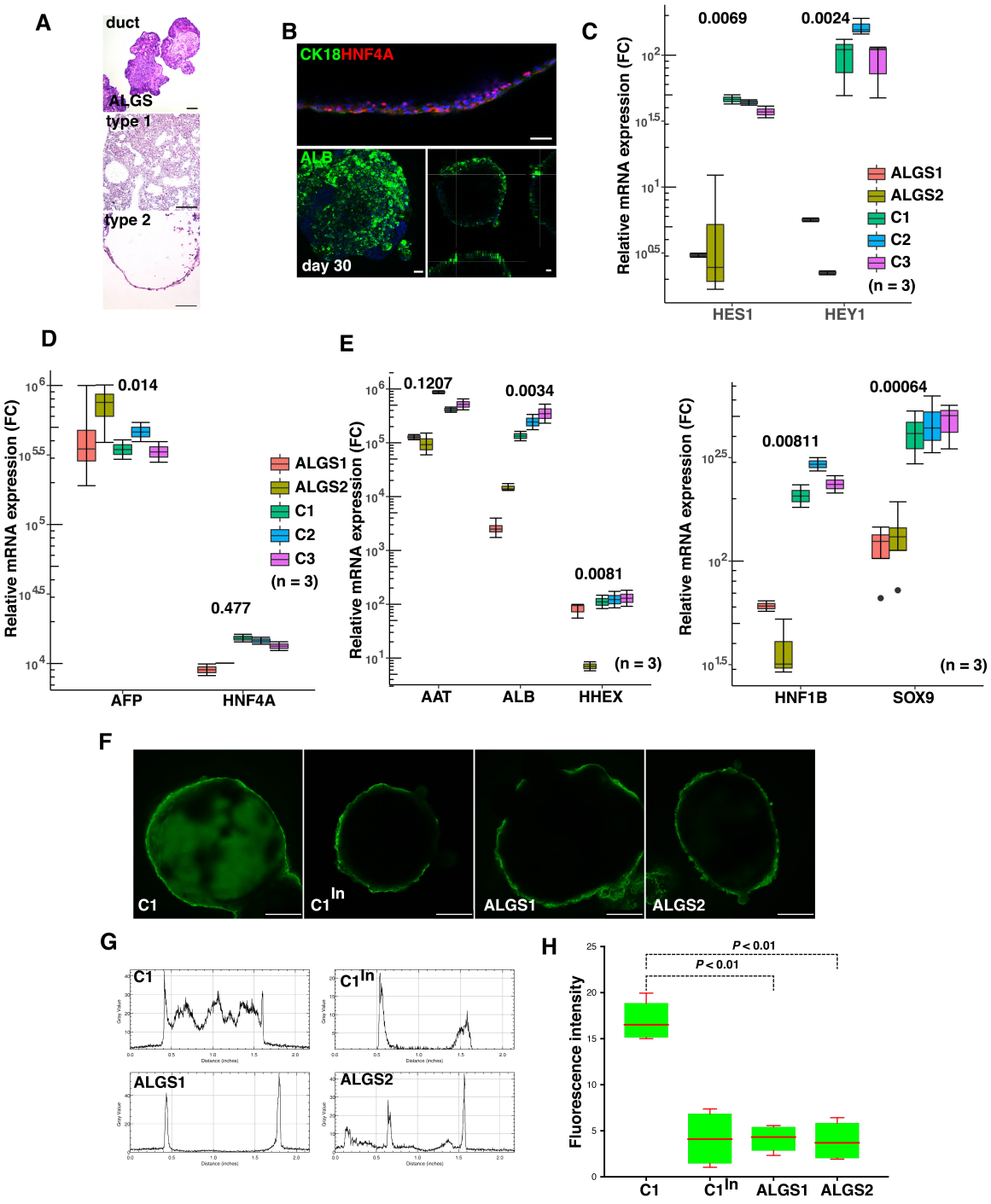


Figure S5. (A) H&E staining shows the structures of the vesicles formed in ALGS HO1 cultures. **(B)** Immunostaining shows that the cells surrounding the fluid-filled vesicles in ALGS HO1 expressed HNF4a, CK18 (top panel) and albumin (bottom panel). The stacked confocal image (bottom left) shows that albumin is expressed throughout the cystic structure. The scale bars are 50 μm. **(B)** Quantitative RT-PCR analyses showing that the expression of mRNAs for

canonical NOTCH target genes (*HEY1*, *HES1*) on day 20 in control HO1 is down-regulated in ALGS HO1 (day 20). **(C)** In contrast, the level of expression of mRNAs for hepatoblast markers (*HNF4A*, *alpha-fetoprotein*) after 9 days of differentiation did not differ between control and ALGS cells. Each bar is the average (± SD) of 3 independent measurements, and the values are normalized relative to those measured on day 0. **(D)** The level of expression of HHEX, Albumin, and AAT mRNAs in day 20 ALGS HO1 are significantly reduced relative to control HO1. The level of expression of *HNF1B* and *SOX9* mRNAs in day 20 ALGS HO1 are 5.2 fold and 3.6 fold reduced, respectively, relative to that in control HO1. Each bar is the average of 3 independent measurements, and the values are normalized relative to their expression level on day 0. Asterisks indicate that the level of expression of the indicated gene is significantly (*, $p<0.05$; **, $p<0.01$) different between control and ALGS cells; and the results are indicative of those obtained in 3 independent experiments. (**E**) Fluorescence images of day 25 HO1 generated from normal control (C1), ALGS1 or ALGS2 iPSCs were obtained 10 minutes after addition of 100 μM rhodamine 123 (green). Rhodamine 123 uptake was also measured in HO1 ($C1^{In}$) after incubation with an MDR1 transport inhibitor (20 μM verapamil) for 30 min before addition of rhodamine 123. Scale bars, 250 μm. (**F**) The fluorescence intensity was measured across the center of the indicated HO1. (**G**) The fluorescence intensity within the lumina was normalized relative to the area. The red line shows the average of 3 independent measurements, and the box indicates the first and third quartiles. Inhibition by verapamil indicates that Rhodamine 123 uptake was carrier dependent. ALGS HO1 could not transport the dye into the lumen.

Figure S6

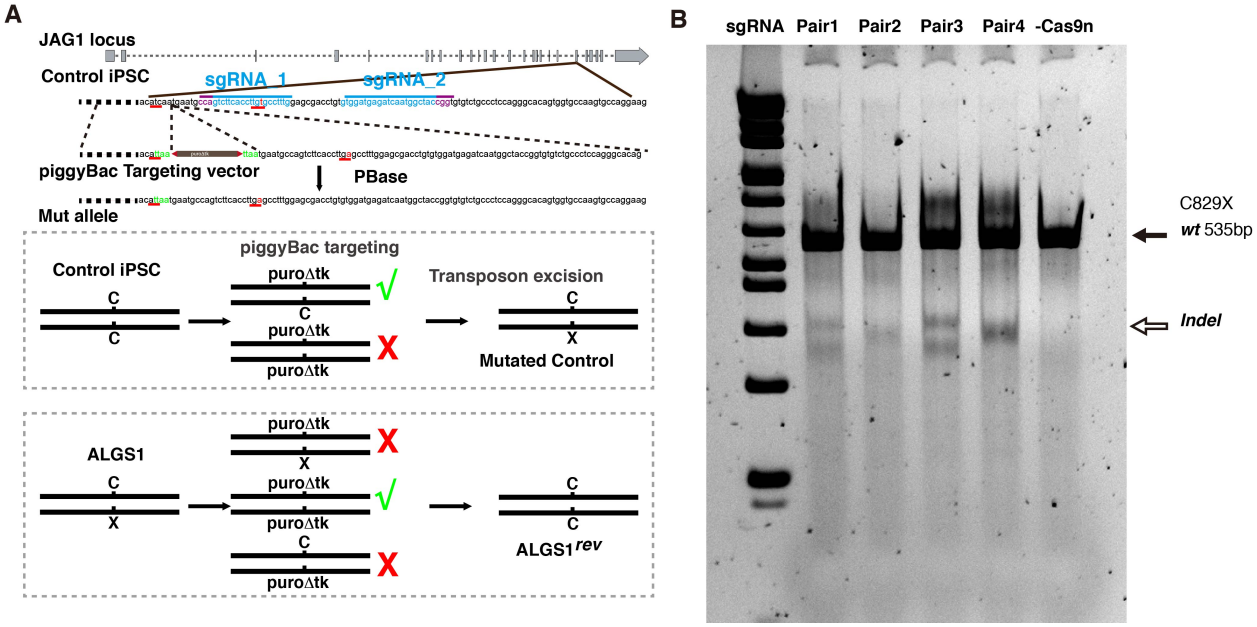


Figure S6. (**A**) A diagram of the strategies used to introduce or revert the ALGS1 mutation (C829X) in control or ALGS1 iPSCs, respectively. (**B**) The T7E1 enzyme assay is used to examine the pattern of genomic changes in iPSC after CRISPR engineering with different sgRNA pairs. The results obtained with two different pairs of sgRNAs (pair 1 and pair 2), which target the same site but have two different (PAM) sites in *JAG1,* are shown in the first two lanes. Of note, different cleavage patterns in *JAG1* genomic DNA are produced by these two different pairs of sgRNAs. We also tested two other pairs of sgRNAs (pair 3 and pair 4) that targeted another site in *JAG1* (C829X) (but had different PAM sites) in this assay. Pairs 3 and 4 produced the same cleavage patterns that were obtained with pairs 1 and 3. Hence, although different patterns are obtained with different pairs of sgRNAs, the cleavage patterns are reproducible.

Figure S7

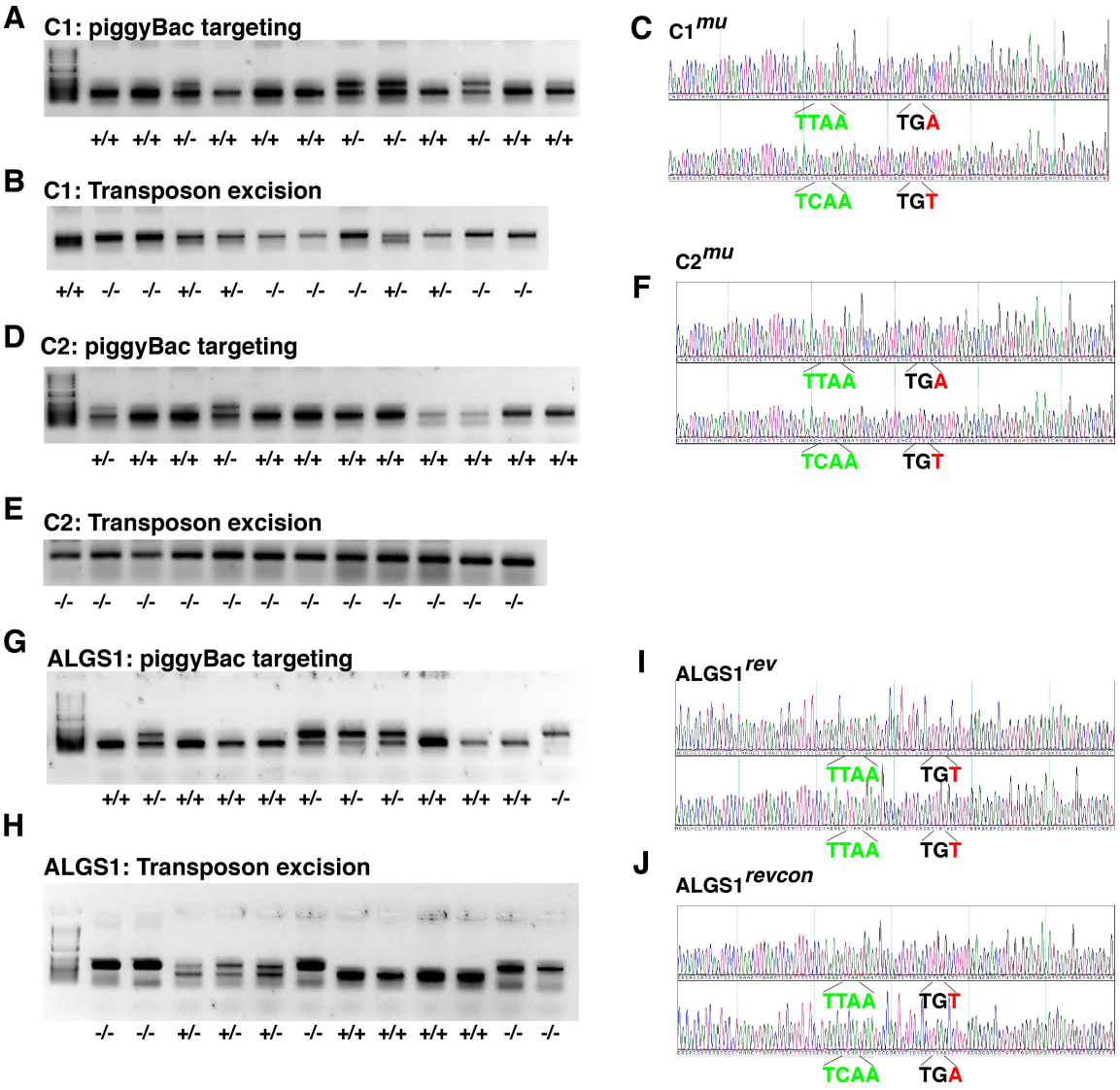


Figure S7. CRISPR-mediated genomic engineering of control and ALGS1 iPSC. (**A, D**) PCR analyses showing whether the PB cassette was inserted into lines prepared from two control iPSC lines (C1 and C2). After this analysis, only clones with a heterozygous insertion of PB (+/-) were picked for expansion. (**B, E**) PCR analyses showing the PB cassette was excised from the heterozygous C1 and C2 iPSC lines. Only clones with transposons removed from both alleles (-/-) were picked for sequencing. (**C, F**) Sequence analyses revealed the C829X mutation was introduced into the *JAG1* gene C1 and C2 iPSCs, and there was seamless excision of the piggyBac transposon. Note that only one allele was modified to create the heterozygous genotype in these iPSC lines. (**G**) PCR analyses showing the insertion of the PB cassette in iPSC lines prepared from ALGS patient 1 (ALGS1), and only homozygous clones were picked for subsequent expansion. (**H**) PCR analyses showing the excision of the PB cassette from homozygous iPSC line derived from ALGS1. Only clones with transposons removed from both alleles were picked for sequencing. (**I, J**) Sequence analyses revealed that C829X mutation was

reverted in ALGS1^{rev} iPSCs, and that there was seamless excision of the piggyBac transposon. Note that both alleles were modified to create the homogenous genotype in the revertant line. Moreover, after the first round of selection, the wild-type allele was replaced by the piggyBac cassette in the heterozygote clones, which leaves a TTAA sequence in the genome. However, the mutated allele is remained unaltered. Hence, the tagged clones, which were referred to as ALGS1^{revcon}, were used as controls for the genome engineering process.

Figure S8

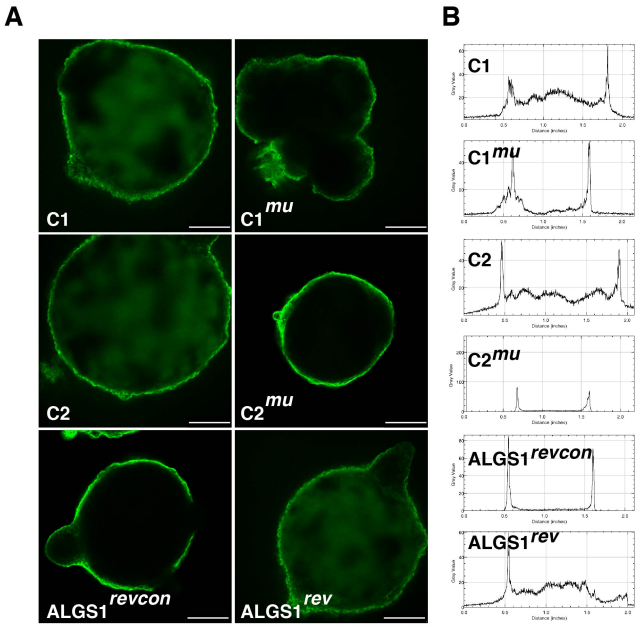


Figure S8. (**A**) Fluorescence images of day 25 HO1 generated from normal control (C1, C2) iPSCs, from CRISPR modified iPSCs with the ALGS1 ($C1^{mu}$) or ALGS2 ($C2^{mu}$) mutations, or $ALGS1^{rev}$ or $ALGS1^{revcon}$ iPSCs. These images were obtained 10 minutes after incubation with 100 ɨ M rhodamine 123 (green). Scale bars, 250 μm. (**B**) The fluorescence intensity was measured across the center of the indicated HO1.

Figure S9

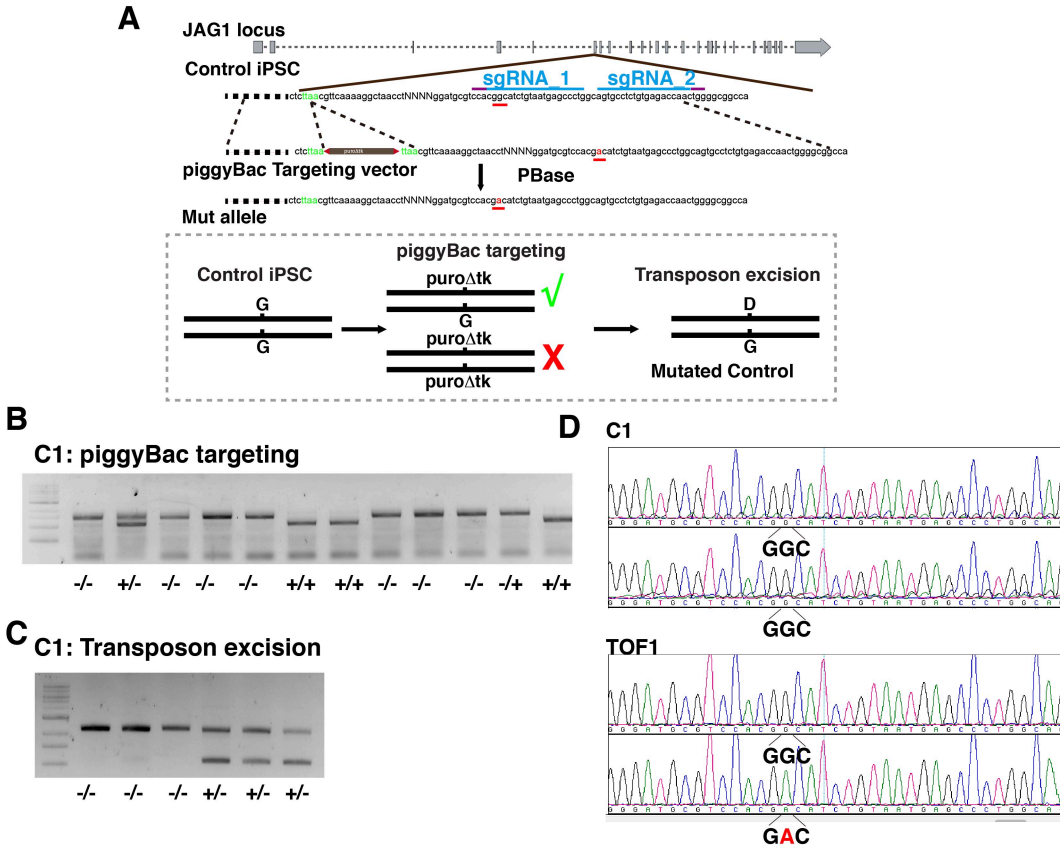


Figure S9. Generation of the TOF1-iPSC line. (**A**) The strategy used to introduce a heterozygous *G274D JAG1* mutation into a control (C1) iPSC line by CRISPR-mediated genomic engineering. (**B**) PCR analyses were performed to determine whether the PB cassette was inserted into the control IPSC (C1) line. Then, clones with a heterozygous PB insertion (+/-) were selected for expansion. (**C**) PCR analyses were performed to determine whether the PB cassette was excised from the heterozygous C1 iPSC lines. Only clones with complete removal of the transposon (-/-) were selected for sequencing. (**D**) Sequencing of genomic DNA from these iPSCs revealed whether the *G274D* mutation was introduced into the *JAG1* gene, and that there was a seamless excision of the piggyBac transposon.

Figure S10

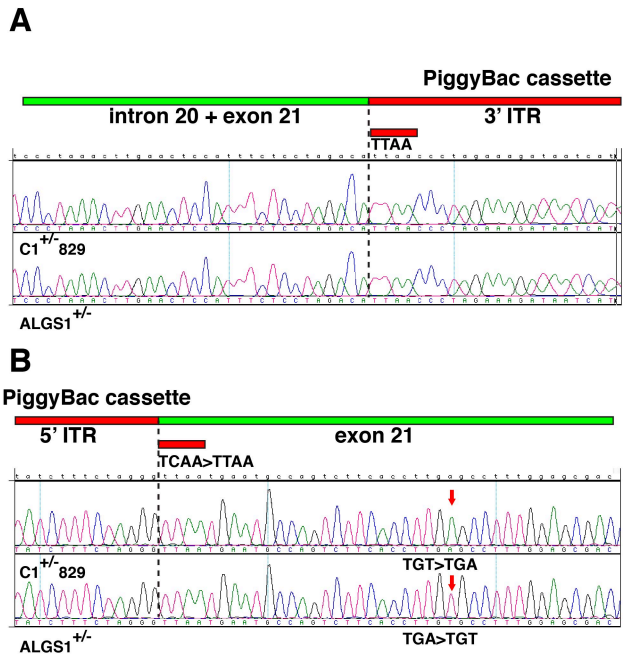


Figure S10. Genomic DNA obtained from two iPSC lines (C1+/-829 and ALGS1+/-) with engineered heterozygous JAG1 gene knockouts was sequenced to determine if the PiggyBac transposon was correctly inserted into the JAG1 gene. The amplicon sequences for the (A) left and (B) right ends of the PiggyBac insert are shown for each cell line. The red lines indicate the sites of the 3' and 5' inverted terminal repeats (ITR) of the PiggyBac insert; and the green lines indicate the regions of the genomic DNA with intron 20 and exon 21 sequences. The sequence data indicates that the PiggyBac transposon was correctly inserted at the border of intron 20 and exon 21, which is 25 bp upstream of the C829 site of the ALGS1 mutation. The TCAA to TTAA conversion shown in the sequence was introduced by a silent mutation (ATC to ATT) within the PiggyBac insertion, which is a sequence that is required for subsequent removal of the transposon.