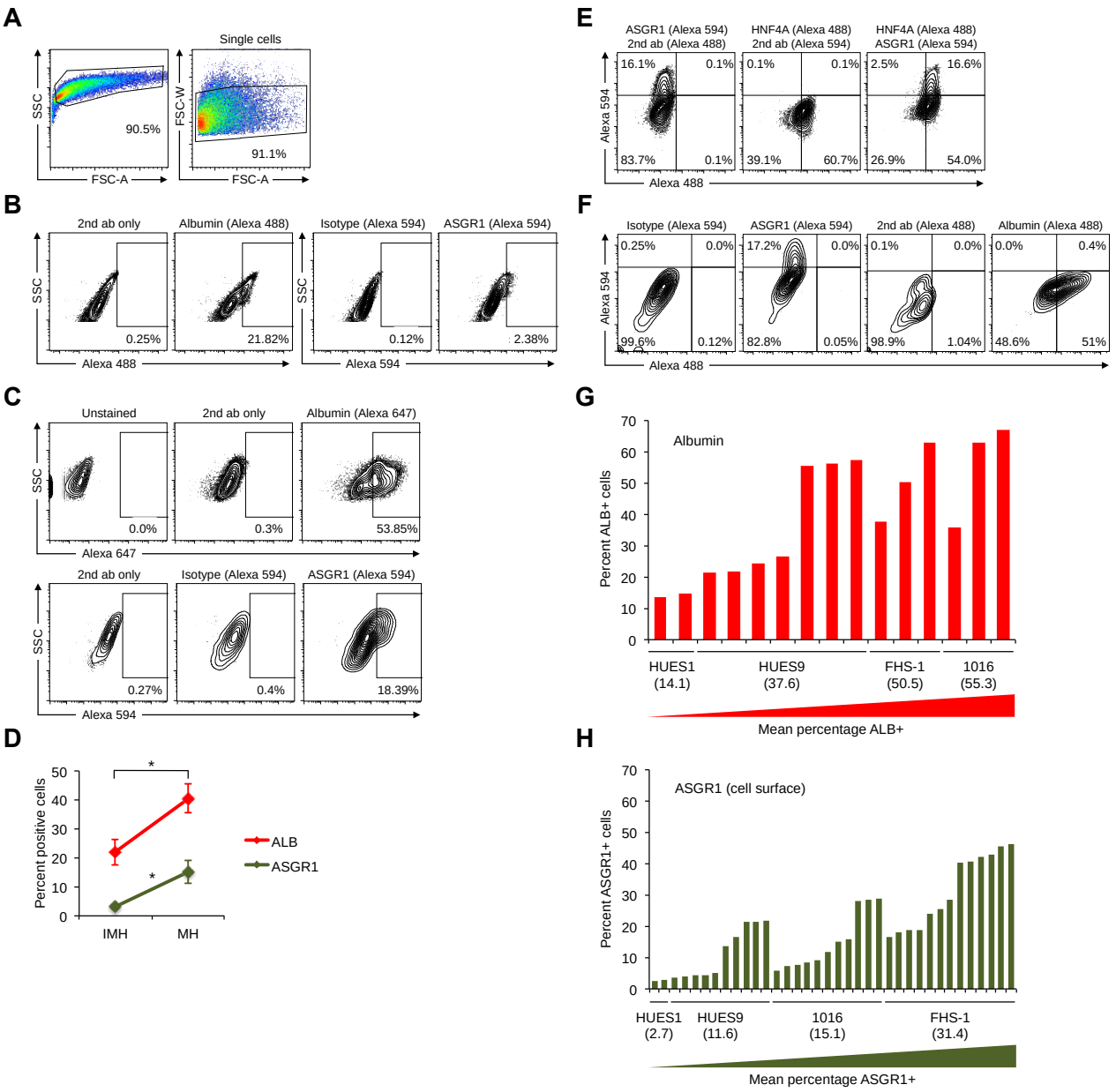

Figure S1. Flow cytometry analysis of HLC differentiation, related to Figure 1. (A) Gating used for flow cytometry analyses; single cells were defined by FSC-A/SSC (cell size and granularity) and FSC-W (cell width) to exclude debris, cell clumps, and doublets. (B) Albumin and ASGR1 expression after the IMH differentiation stage with corresponding staining controls. Alexa 488 and Alexa 594 conjugated secondary antibodies were used with albumin and ASGR1 primary antibodies respectively; gating for albumin and ASGR1 positive cells was performed using secondary-only and isotype-control staining conditions

respectively. (C) Albumin and ASGR1 expression after the MH differentiation stage with corresponding staining controls. Alexa 647 and Alexa 594 conjugated secondary antibodies were used with albumin and ASGR1 primary antibodies respectively; gating for albumin and ASGR1 positive cells was performed using secondary-only and isotype-control staining conditions respectively. A secondary-only staining control is also shown for ASGR1. (D) Kinetics of albumin and ASGR1 expression during the final stages of HLC differentiation as determined by intracellular flow cytometry. Shown are mean percent positive cells at the IMH and MH stages based on multiple independent differentiations (n = 5 - 15 replicates per marker, per stage). Error bars represent s.e.m. Asterisks indicate statistically significant differences in mean percentage of positive cells at the IMH and MH stages for each marker by Student's t-test. *, $P < 0.05$. (E) Flow cytometry analysis of ASGR1 and HNF4A co-expression with staining controls. Alexa 594 and Alexa 488 conjugated secondary antibodies were used for ASGR1 and HNF4A staining respectively. Gating was performed based on secondary-only (fluorescence minus one) staining conditions. (F) Flow cytometry analysis of ASGR1 and albumin co-expression with staining controls. Alexa 594 and Alexa 488 conjugated secondary antibodies were used for ASGR1 and albumin staining respectively. Gating for ASGR1 and albumin positive cells was performed based on isotype-control and secondary-only staining conditions respectively. (G-H) Efficiency of HLC differentiation with four different hPSC lines over multiple independent experiments. Shown are percent albumin-positive (G) and surface ASGR1-positive (H) cells determined by flow cytometry analysis after the MH differentiation stage, illustrating a range of differentiation efficiencies.

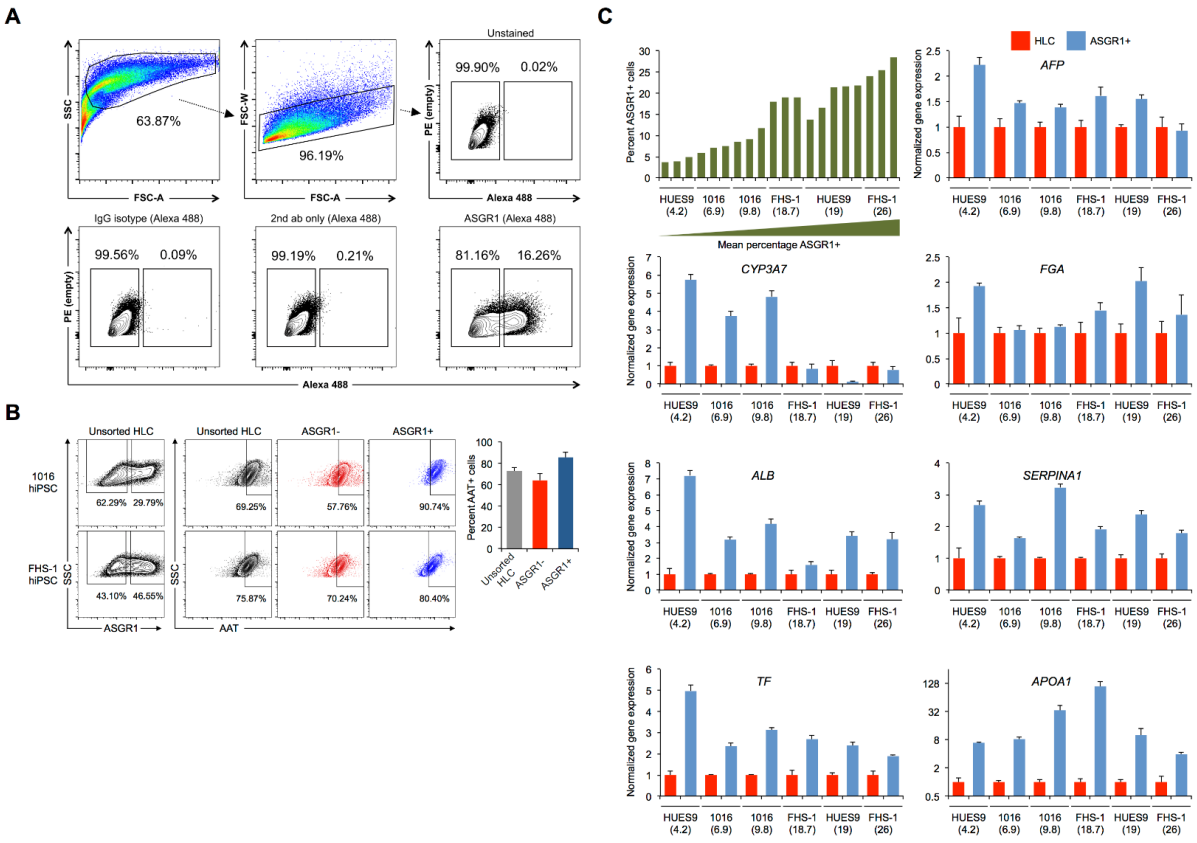

Figure S2. Detailed results of ASGR1 FACS and hepatocyte marker gene expression analysis, related to Figure 2. (A) Strategy and staining controls for FACS isolation of ASGR1+ HLCs. Shown are results of a representative differentiation and FACS experiment. (B) Left: two different hPSC lines were differentiated to HLCs. The percentage of cells expressing the hepatocyte marker alpha-1 antitrypsin (AAT) among unsorted HLCs, surface ASGR1-negative cells, and surface ASGR1-positive cells was quantified by intracellular flow cytometry. Right: mean percent AAT-positive cells by flow cytometry, among unsorted HLCs, surface ASGR1-negative cells, and surface ASGR1-positive cells (*n* = 2 differentiations). Error bars represent s.e.m. (C) ASGR1 FACS and hepatocyte marker gene expression analysis by qRT-PCR. Gene expression data is displayed as a heatmap in Figure 2C. Three different hPSC lines were differentiated to HLCs, with two independent differentiations performed per cell line (differentiations performed in triplicate or greater, *n* = 6 – 8 biological replicates per cell line). Data are arranged according to mean percent ASGR1+ cells obtained in each differentiation. Green bars: percentage surface ASGR1+ cells after the MH

differentiation stage. All other graphs show qRT-PCR gene expression analysis in unsorted HLCs ("HLC," red bars) and ASGR1+ cells ("ASGR1+," blue bars) isolated by FACS. Expression levels are relative to *RPLP0* expression; gene expression levels in ASGR1+ cells were normalized to level in unsorted HLCs. Shown are normalized mean expression levels for each differentiation (n = 3 – 5 paired biological replicates per differentiation). Error bars represent s.e.m.

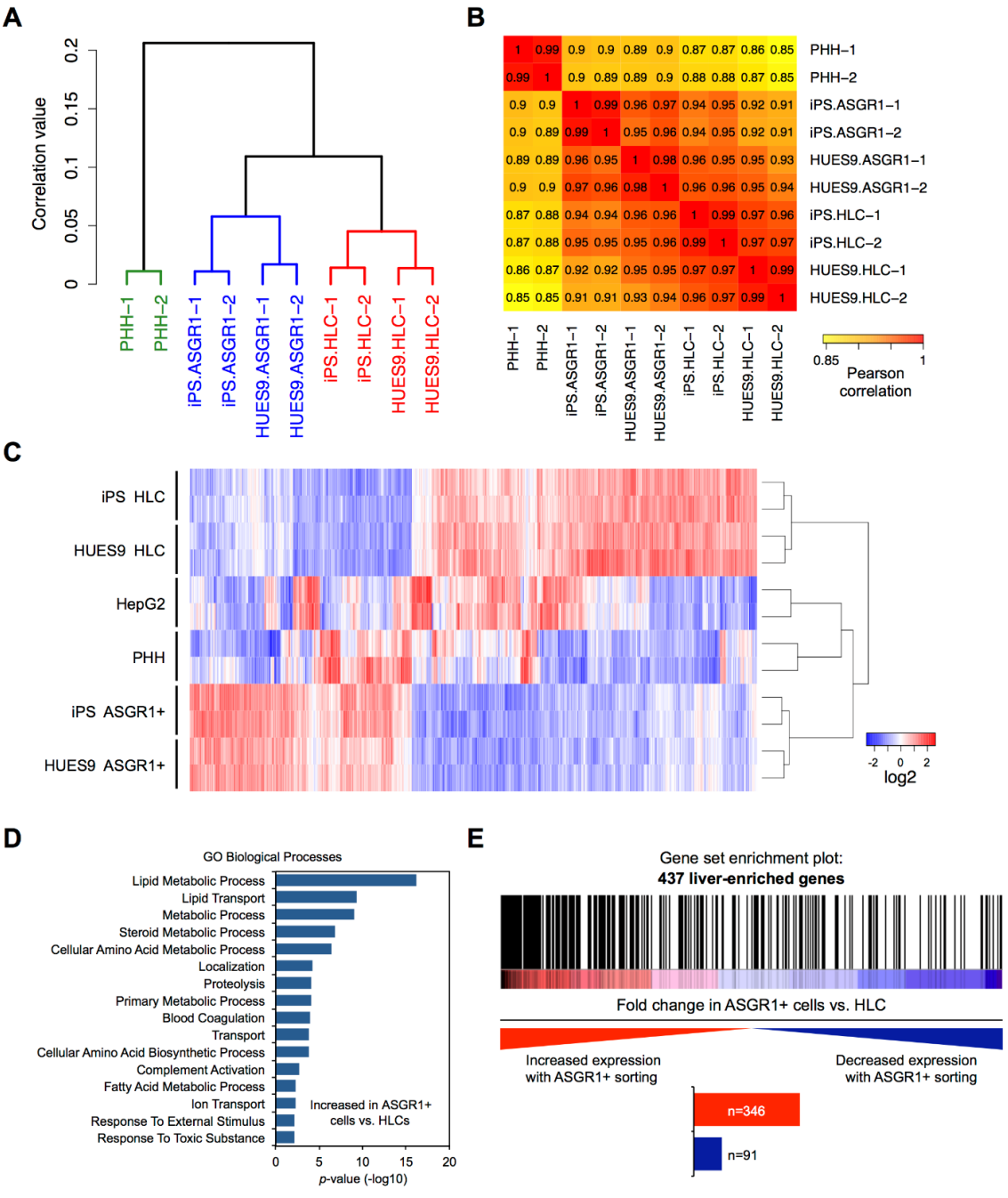

Figure S3. Additional analyses of microarray gene expression profiling data from ASGR1-positive cells, unsorted HLCs, primary human hepatocytes, and HepG2 hepatoma cells, related to figure 3. (A) Hierarchical clustering of ASGR1-positive, HLC, and PHH samples based on all genes measured by transcriptional microarray and expressed above background. (B) Heatmap showing pairwise Pearson correlation values for ASGR1-

positive, HLC, and PHH samples based on the same expression data used in part A. Yellow, orange, and red color denotes lower, intermediate, and higher correlation respectively. (C) Heatmap of hierarchical clustering performed on all genes differentially expressed between ASGR1-positive cells and HLCs at a 5% FDR. Blue, below average expression, red, above average expression. 813 probesets differentially expressed between ASGR1+ and unsorted HLC at 5% FDR; 318 upregulated and 495 downregulated in ASGR1+ vs. unsorted HLCs respectively. (D) Functional enrichment analysis of genes differentially expressed in ASGR1-positive cells relative to HLCs. (E) Gene set enrichment analysis (GSEA). Microarray gene expression data was arranged based on greater average expression in ASGR1-positive cells (red color) or unsorted HLCs (blue color). Vertical black bars represent genes within the liver-enriched gene set.

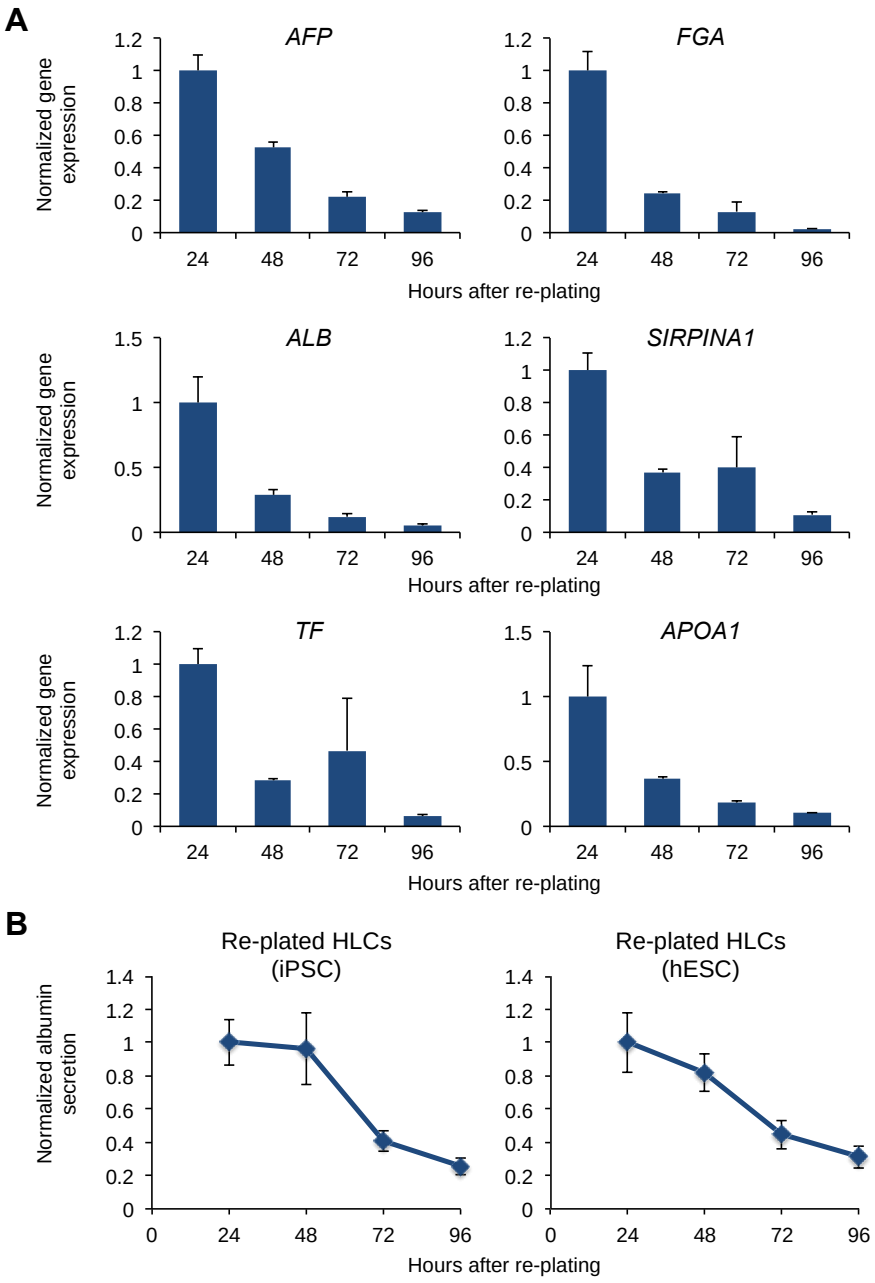

Figure S4. Hepatocyte marker gene expression and albumin secretion declines as anticipated following re-plating of ASGR1 MACS-enriched HLCs. (A) Changes in hepatocyte marker gene expression over time were determined by qRT-PCR after re-plating HLCs differentiated from a representative hESC line. Gene expression levels were calculated relative to *RPLP0* expression and normalized to gene expression level at 24 hours post re-plating for each gene (*n* = 3 differentiation wells per time point). Error bars represent s.e.m. (B)

Quantification of albumin secretion by ELISA following re-plating of HLCs differentiated from two hPSC lines. Albumin concentration at the indicated time points were calculated using a standard curve and normalized to the level at 24 hours post re-plating for each cell line (n = 3 differentiation wells per cell line, per time point). Medium was collected after 24 hours in culture for each time point. Error bars represent s.e.m.

Table S1. Distribution of liver-enriched genes in gene set enrichment analysis (GSEA) comparing ASGR1+ cells and unsorted HLCs, related to Figure S3E.

Subset of liver-enriched genes with greater expression in ASGR1+ cells. 205 genes contributed most strongly to the enrichement result (core enrichment). 346 of 437 genes had greater expression in ASGR1+ cells vs. unsorted HLCs.						Liver-enriched genes with lower expression in ASGR1+ cells. 91 of 437 genes had lower expression in ASGR1+ cells vs. unsorted HLCs		
AADAC	*ASGR1*	*FBP1*	*LEPR*	*SEPHS2*	*TTR*	*ABCB4*	*FMO4*	*SEC14L2*
ABCG8	*ASGR2*	*FETUB*	*LIPC*	*SERPINA10*	*UGT2B4*	*ABCG5*	*GCGR*	*SEC14L4*
ACAA1	*ASS1*	*FGA*	*MAT1A*	*SERPINA4*	*UGT3A1*	*ACAT2*	*GHR*	*SLCO1B1*
ACAA2	*ATF5*	*FGB*	*METTL7B*	*SERPINA6*	*VTN*	*ACOT12*	*GLS2*	*SOD1*
ACADSB	*AZGP1*	*FGFR4*	*MTHFS*	*SERPINA7*	*ZNHIT1*	*ANXA10*	*GLT1D1*	*SPP2*
ACAT1	*BAAT*	*FGG*	*MTTP*	*SERPINC1*		*APOC4*	*GNE*	*STEAP3*
ACMSD	*BPHL*	*FGL1*	*MUT*	*SERPIND1*		*AQP9*	*GNPNAT1*	*THPO*
ACOX1	*BRP44*	*FMO5*	*NAT8*	*SERPINF2*		*AS3MT*	*GPR126*	*TNFSF14*
ACOX2	*C1S*	*G6PC*	*NEK6*	*SERPING1*		*ATF7IP2*	*GRHPR*	*TUBB1*
ACSL1	*C5*	*GATM*	*NIPSNAP1*	*SHMT1*		*BCO2*	*GSTZ1*	*UROC1*
ACY1	*C6*	*GC*	*NIT2*	*SHMT2*		*C4BPA*	*GYS2*	*ZNF281*
ADH1A	*CD302*	*GCHFR*	*OIT3*	*SLC10A1*		*C4BPB*	*HAGH*	
ADH6	*CDO1*	*GGH*	*ORM1*	*SLC13A5*		*CCL16*	*HAMP*	
ADK	*CFH*	*GSTO1*	*OTC*	*SLC22A7*		*CD5L*	*HLF*	
AFF4	*CFI*	*HABP2*	*PAH*	*SLC22A9*		*CDC37L1*	*HPS3*	
AFM	*CIDEB*	*HMGCL*	*PCBD1*	*SLC25A13*		*CES1*	*HSD11B1*	
AGTR1	*CLDN1*	*HMGCS2*	*PCCB*	*SLC27A2*		*CES2*	*IBTK*	
AGXT	*CPB2*	*HNF4A*	*PCK2*	*SLC2A2*		*CFP*	*IGFALS*	
AGXT2	*CPN1*	*HP*	*PEBP1*	*SLC30A1*		*CLEC4M*	*ITIH4*	
AHSG	*CPN2*	*HPD*	*PECR*	*SLC35D1*		*COLEC11*	*KMO*	
AIG1	*CPS1*	*HPN*	*PEMT*	*SLC38A3*		*CP*	*LARP4*	
AKR1A1	*CREB3L3*	*HPX*	*PGRMC1*	*SLC38A4*		*CRP*	*LECT2*	
AKR1D1	*CRYL1*	*HRSP12*	*PHYH*	*SLC39A14*		*CTH*	*LPIN2*	
ALAS1	*CYB5A*	*HSD17B4*	*PIPOX*	*SLC41A2*		*CYP1A2*	*MAMDC4*	
ALB	*CYP3A7*	*HSD3B7*	*PKLR*	*SLC43A1*		*CYP26A1*	*MASP1*	
ALDH6A1	*CYP4A11*	*HSPE1*	*PLA2G12B*	*SLCO2B1*		*CYP2A6*	*MASP2*	
ALDH8A1	*CYP4V2*	*ID2*	*PLG*	*SORD*		*CYP2B7P1*	*MBL2*	
ALDOB	*CYP8B1*	*IGFBP1*	*PNPLA3*	*SPRYD4*		*CYP2C8*	*MCL1*	
AMBP	*DCXR*	*IL1RAP*	*PRAP1*	*ST6GAL1*		*CYP2D6*	*MTHFD1*	
ANG	*DDT*	*INHBE*	*PRDX4*	*SULT2A1*		*CYP2E1*	*MYO1B*	
ANGPTL3	*DECR1*	*INSIG1*	*PROC*	*TAT*		*CYP39A1*	*N4BP2L1*	
APOA1	*EPHX1*	*ITIH1*	*PROS1*	*TDO2*		*CYP4A22*	*PHLDA1*	
APOA2	*ETFB*	*ITIH2*	*PROX1*	*TFR2*		*DEFB123*	*PLGLB2*	
APOB	*F10*	*KHK*	*PROZ*	*TM4SF4*		*DEPDC7*	*PON1*	
APOC2	*F13B*	*KLB*	*PXMP2*	*TM4SF5*		*DHRS1*	*PON3*	
APOC3	*F2*	*KLKB1*	*RBP4*	*TMEM176A*		*DMGDH*	*PPAPDC2*	
APOE	*F5*	*KNG1*	*RCL1*	*TMEM176B*		*ERRFI1*	*PZP*	
APOH	*F7*	*LASS2*	*SAA4*	*TMEM56*		*ETFDH*	*RTP3*	
APOM	*F9*	*LBP*	*SC5DL*	*TMPRSS6*		*FAM167B*	*SCUBE3*	
ARG1	*FAM96A*	*LCAT*	*SCP2*	*TP53INP1*		*FMO3*	*SDS*	