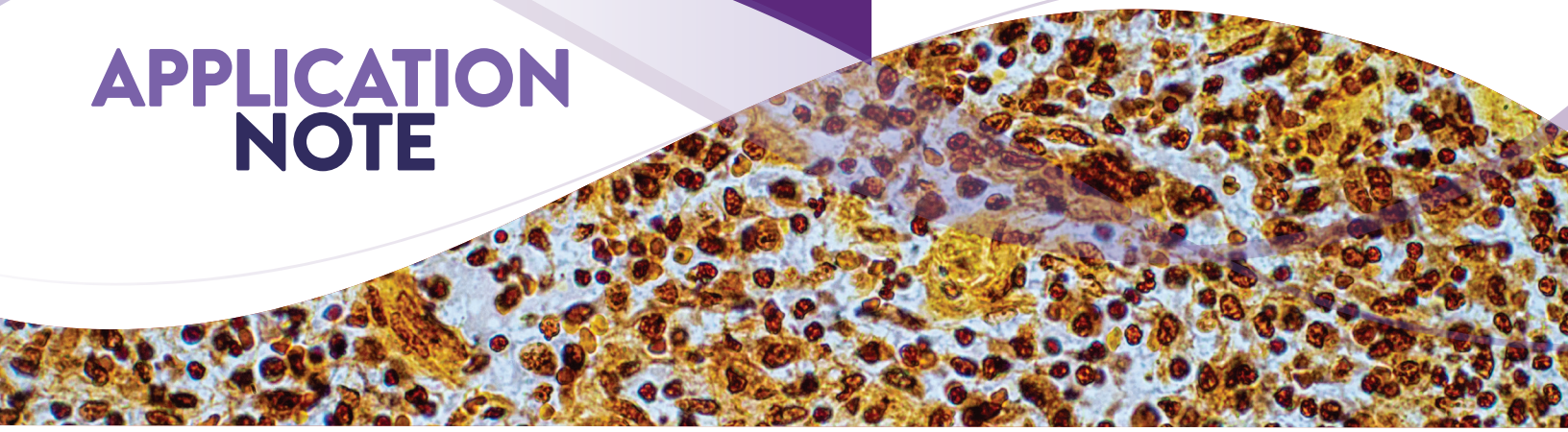
ATCC®

APPLICATION NOTE



RNA-SEQ REVEALS COMPREHENSIVE GENE EXPRESSION LANDSCAPE IN PRIMARY HUMAN HEPATOCYTES

Ajeet P. Singh, PhD; Emma Todd, BS; Sujoy Lahiri, PhD; Robert Marlow, BS Jonathan Jacobs, PhD
ATCC, Manassas, VA 20110

ABSTRACT

Primary human hepatocytes (PHHs) are essential for modeling liver-specific functions, including drug metabolism, detoxification, and transport. In this study, RNA sequencing (RNA-seq) was performed on PHHs derived from four individual donors to characterize transcriptomic profiles and assess interdonor variability. High-throughput sequencing and rigorous bioinformatic analyses revealed consistent expression of core hepatic genes, along with notable differences in cytochrome P450 enzymes, UDP-glucuronosyltransferases, and ABC transporters. Principal component analysis and hierarchical clustering confirmed strong intradonor reproducibility and distinct interdonor transcriptional signatures. These findings underscore the importance of accounting for donor-specific variability in hepatocyte-based research and support the use of standardized RNA-seq workflows for advancing liver biology and pharmacogenomics.

INTRODUCTION

Primary human hepatocytes (PHHs) are widely recognized as the gold standard for in vitro liver models due to their metabolic competence and physiological relevance. These cells play a central role in xenobiotic metabolism, lipid regulation, and protein synthesis, making them indispensable models for drug development, toxicology, and disease modeling.[1] However, their limited proliferative capacity and rapid dedifferentiation in vitro hinder long-term experimentation and compromise reproducibility.

RNA sequencing (RNA-seq) has revolutionized transcriptomic profiling by enabling high-resolution, genome-wide analysis of gene expression. Unlike microarrays, RNA-seq does not rely on predefined probes, enabling the detection of novel transcripts, splice variants, and subtle expression changes. In hepatocyte research, RNA-seq has been instrumental in uncovering liver-specific gene signatures, identifying regulatory networks, and characterizing responses to metabolic and inflammatory stimuli.[1]

Recent studies have integrated RNA-seq with chromatin accessibility assays (e.g., ATAC-seq) to explore the epigenetic landscape of hepatocytes during differentiation and aging.[1-3] For instance, Ding et al. demonstrated that long-term cultured hepatocyte-derived progenitor-like cells exhibit transcriptional and epigenetic signatures of cellular aging, driven by inflammatory pathways regulated by FOSL2.[1] These findings underscore the value of combining transcriptomic and epigenomic data to understand hepatocyte biology and improve culture protocols.

In the following study, we present a comprehensive RNA-seq analysis of PHHs under defined conditions. Our goal is to benchmark transcriptomic profiles, assess reproducibility, and visualize key pathways relevant to liver function and drug metabolism. This is particularly important because hepatocytes serve as a central model for studying liver physiology, toxicology, and pharmacokinetics. However, variability in culture conditions, donor sources, and analytical pipelines often limit the comparability and reliability of transcriptomic data across studies. By establishing standardized workflows and integrating rigorous quality metrics, we aim to enhance the reproducibility and interpretability of hepatocyte-based transcriptomic analyses. Ultimately, this contributes to the development of robust platforms for both academic research and industrial applications, including drug screening, disease modeling, and regulatory science.

MATERIALS AND METHODS

CELL CULTURE

HepatoXcell™ Normal Human Hepatocytes (ATCC® PCS-450-010™, ATCC® PCS-450-011™, and ATCC® PCS-450-012™) were stored in the vapor-phase of liquid nitrogen (< -130°C) until use. Cells were thawed in a 37°C water bath for 1–2 minutes and transferred to 19 mL of pre-warmed HepatoXcell™ Primary Hepatocyte Thawing Medium 1x (ATCC® PCS-450-032™). Residual cells were recovered by rinsing the vial with 1 mL of thawing medium.

After gentle mixing, cells were centrifuged at 100 × g for 10 minutes, resuspended in 1–2 mL of HepatoXcell™ Primary Hepatocyte Plating Media 1x (ATCC® PCS-450-038™), counted, and diluted to 800,000 cells/mL. 500 µL of cell suspension was seeded per well in Collagen I-coated 24-well plates and then incubated at 36°C ± 2°C with 5% CO_2.

Plates were gently shaken at 1 hour to remove dead cells, and media was changed at 2 hours and again at 4–6 hours with cold HepatoXcell™ Primary Hepatocyte Maintenance Media 1x (ATCC® PCS-450-034™) plus 0.3 mg/mL Cell Basement Membrane (ATCC® ACS-3035™).

RNA EXTRACTION AND QUALITY CONTROL

Twenty-four hours after seeding, hepatocyte cultures were terminated. Total RNA was extracted from hepatocytes using TRIzol reagent (Thermo Fisher Scientific), followed by purification with the PureLink RNA Mini Kit (Thermo Fisher Scientific). Cells were lysed in 0.5 mL TRIzol per well, and phase separation was performed with chloroform. The aqueous phase was mixed with 70% ethanol and loaded onto spin cartridges for binding, washing, and elution according to the kit instructions. RNA concentration was measured using Qubit (Thermo Fisher Scientific), and purity (260/280 ratio) was assessed by Nanodrop (Thermo Fisher Scientific).

RNA-SEQ LIBRARY PREPARATION AND SEQUENCING

Automated RNA-seq next-generation sequencing (NGS) library preparation was performed on an Eppendorf epMotion 5075 Liquid Handler using the Illumina Stranded mRNA Prep, Ligation kit. Prepared NGS libraries were assessed using Invitrogen Qubit dsDNA High Sensitivity Assay Kit and an Agilent 4200 TapeStation and D5000 ScreenTape System. Libraries were prepared using an Illumina P3 200-cycle Reagent kit and sequenced on the NextSeq 2000 platform.

RNA-SEQ DATA ANALYSIS

Our data analysis pipeline included quality control, read trimming, alignment to the reference transcriptome, and quantification of gene expression. Utilizing the CLC Genomics Workbench v25 (QIAGEN Digital Insights), an end-to-end pipeline was created that, briefly, entailed the following steps. First, raw paired-end Illumina reads were trimmed and filtered to a minimum quality of Q30 and a maximum of 2 ambiguous bases. Potential "read-through" adapter sequences and 3' polyG sequences (due to using the 2-color NextSeq platform) were also automatically identified and trimmed. Reads below 50 bp were then discarded. Next, reads were mapped to the human genome hg38 reference genome (obtained from Ensemble) using the default settings for bulk-RNA-seq experiments: mismatch cost, 2; InDel cost, 3; length fraction, 0.8; similarity fraction, 0.8; maximum hits per reads, 10; reversed strand-specific mapping, 1; and ignore broken read pairs, 1. A minimum of 18M mapped reads per library were required for each biological replicate, and trimmed mean of M-values (TMM) normalization was carried out for each library.[4] Statistical comparisons between groups were conducted using a Wald test, and fold-change values calculated from the GLM (Generalized Linear Model) model.[5] Outliers were down weighted and iteratively re-fit to the GLM model.[6] Low expression genes were filtered prior to FDR correction and calculation.[7]

RESULTS

KEY GENES DRIVING HEPATOCYTE FUNCTION AND DRUG METABOLISM

Primary human hepatocytes (PHHs) serve as a foundational model for studying liver-specific drug metabolism and pharmacokinetics. To assess the activity of key metabolic and transport pathways, we conducted whole-transcriptome analyses on four human donor-derived hepatocyte samples obtained from ATCC: PCS-450-010™ lot 015, PCS-450-011™ lot 0012, PCS-450-012™ lot 0018, and PCS-450-012™ lot 009a. Heatmaps were generated to visualize the expression patterns of genes involved in hepatic metabolism and transport (Figure 1). Across the dataset, most genes showed consistent enrichment in PCS-450-011™ lot 0012, PCS-450-012™ lot 0018, and PCS-450-012™ lot 009a, indicating robust metabolic activity in these donors. In contrast, PCS-450-010™ lot 015 exhibited generally lower expression across most genes analyzed. A subset of genes, however, was uniformly expressed across all four donors, suggesting a core set of hepatic functions maintained irrespective of donor variability. However, key metabolic genes *CYP2E1, UGT2B4, ABCC2,* and *SLCO1B1* exhibited donor-specific distinctive expression patterns (Table 1, Table 2). These findings highlight interdonor differences in gene expression and emphasize the need to account for individual variability when interpreting hepatocyte-based pharmacogenomic data.

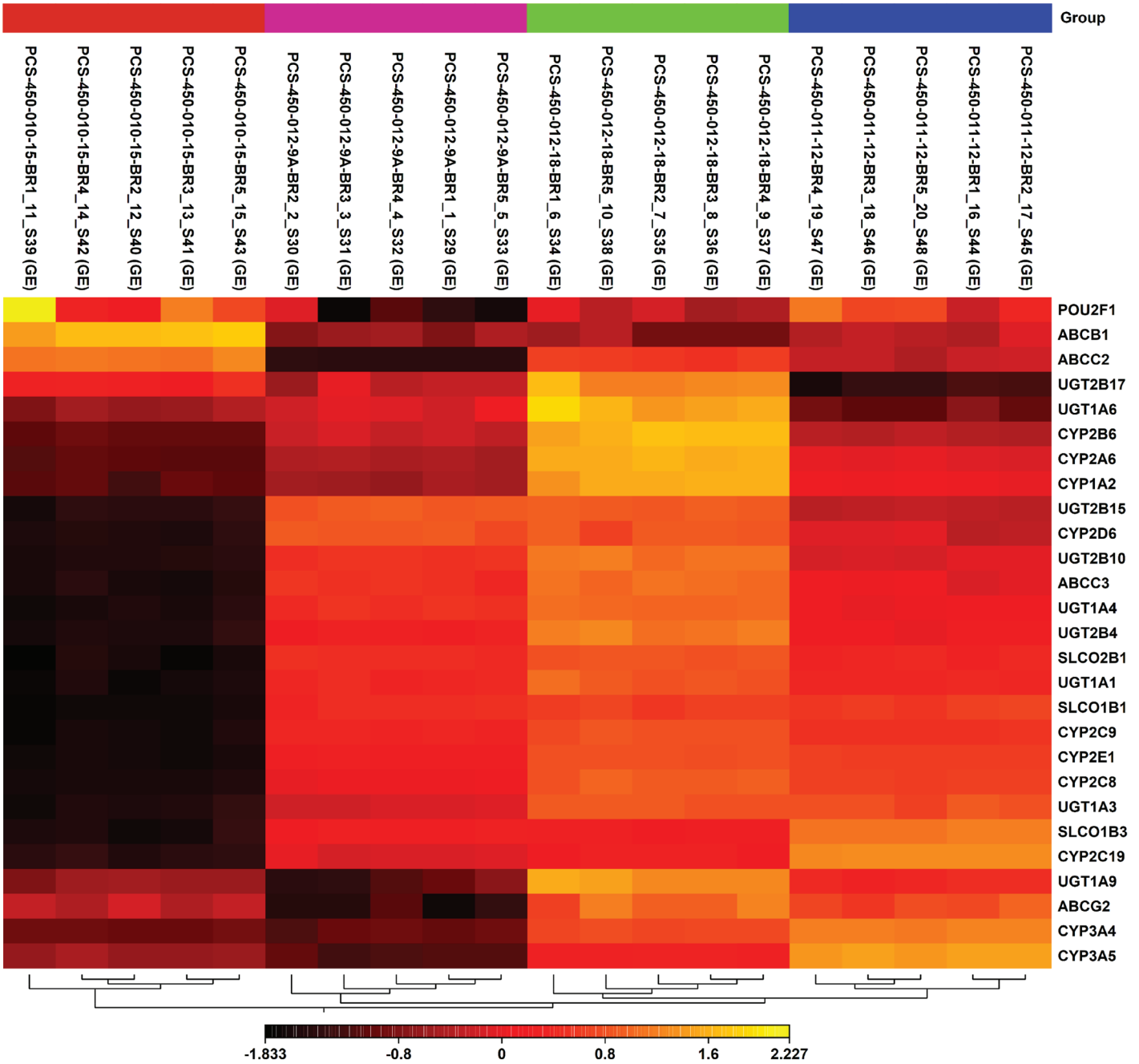


Figure 1: Gene expression heatmap of key metabolic and transport genes across primary human hepatocyte samples. Heatmap displays transcript levels of genes involved in hepatic drug metabolism and transport across multiple donor-derived hepatocyte samples. Columns represent individual replicates from donors PCS-450-010™ lot 015, PCS-450-011™ lot 0012, PCS-450-012™ lot 0018, and PCS-450-012™ lot 009a, grouped by color-coded categories indicating donor identity. Each donor is represented by a minimum of five biological replicates, each cultivated and harvested independently from separate flasks to ensure reproducibility and minimize batch effects. Rows correspond to genes including cytochrome P450 enzymes (*CYP1A2, CYP2B6, CYP2D6, CYP3A4, CYP3A5*), UDP-glucuronosyltransferases (*UGT1A1, UGT2B17, UGT2B15*), solute carriers (*SLCO1B1, SLCO2B1*), and ABC transporters (*ABCB1, ABCC2, ABCG2*), among others. Expression levels are color-coded from low (dark black, -1.833) to high (bright yellow, 2.227). The heatmap reveals consistent clustering among replicates from the same donor and highlights interdonor variability in gene expression, reflecting differences in hepatic metabolic capacity and transporter activity.

Table 1: Donor-Specific Expression of Key Genes

Donor ID	CYP Genes (High)	UGT Genes (High)	ABC Transporters (High)	General Expression Profile
PCS-450-010™ lot 015	CYP2E1 (Low)	UGT2B4 (Low)	ABCC2 (High), SLCO1B1 (Low)	Generally lower across most genes
PCS-450-011™ lot 0012	CYP2E1 (High)	UGT2B4 (High)	ABCC2, SLCO1B1 (High)	Robust metabolic activity
PCS-450-012™ lot 0018	CYP2E1 (High)	UGT2B4 (High)	ABCC2, SLCO1B1 (High)	Robust metabolic activity
PCS-450-012™ lot 009a	CYP2E1 (High)	UGT2B4 (High)	ABCC2 (low), SLCO1B1 (High)	Robust metabolic activity

Table 2: Comparative Expression Levels of CYP Genes (CPM)

CYP Gene	PCS-450-010™ lot 015	PCS-450-011™ lot 0012	PCS-450-012™ lot 0018	PCS-450-012™ lot 009a
CYP2E1	Moderate	High (~8000)	High (~8000)	High (~8000)
CYP2C9	Moderate	High	High	High
CYP2C8	Moderate	High	High	High
CYP3A5	Low-Moderate	High	Low-Moderate	Low-Moderate
CYP3A4	Low	Moderate	Moderate	Low
CYP2A6	Low	Low-Moderate	Moderate	Low
CYP2B6	Low	Low	Moderate	Low
CYP2D6	Low	Low	Moderate	Moderate
CYP2C19	Very Low	Moderate	Very Low	Very Low
CYP1A2	Very Low	Very Low	Low-Moderate	Very Low

CYTOCHROME P450 (CYP) SUPERFAMILY GENE EXPRESSION

To assess the metabolic activity of hepatocytes, we began by analyzing the expression levels of key cytochrome P450 (CYP) enzymes involved in drug metabolism.[8] Expression levels were quantified in counts per million (CPM) for ten CYP genes: *CYP3A5, CYP3A4, CYP2C19, CYP2C9, CYP2C8, CYP2E1, CYP1A2, CYP2A6, CYP2B6*, and *CYP2D6*.

Among these, *CYP2E1* exhibited the highest expression across all samples except PCS-450-010™ lot 0015, with CPM values approaching 8000, consistent with its established role in hepatic drug metabolism. *CYP2C9* and *CYP2C8* showed moderate to high expression (1000–4000 CPM), while *CYP3A5, CYP3A4, CYP2A6, CYP2B6*, and *CYP2D6* were expressed at lower levels (100–1000 CPM) (Figure 2). *CYP2C19*, and *CYP3A2* were the least expressed genes, with CPM values below 300 in most samples. Notably, PCS-450-010™ lot 0015 consistently exhibited reduced expression across all CYP genes compared to other donors.

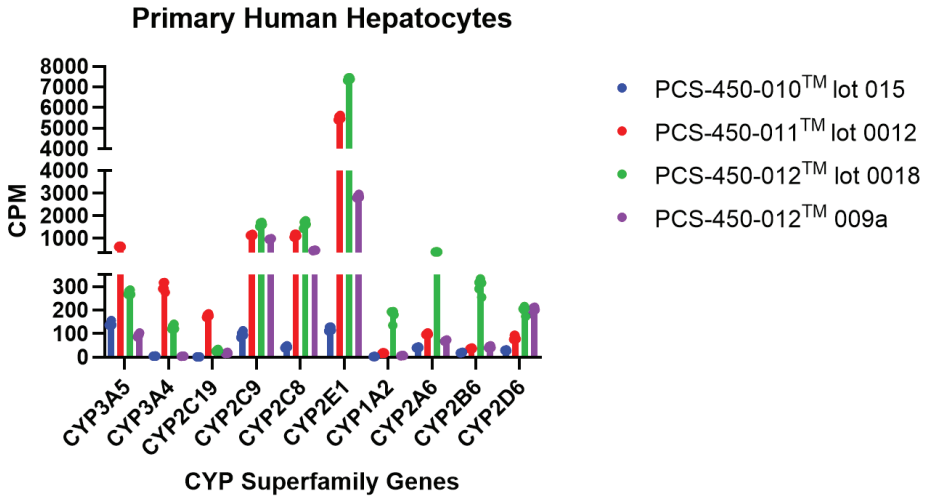


Figure 2: Comparative expression of cytochrome P450 genes across four primary human hepatocyte samples. The graph illustrates transcript levels in counts per million (CPM) of cytochrome P450 (CYP) genes in four donor-derived hepatocyte samples. The expression levels for respective donor samples are shown for each gene, with error bars indicating variability across replicates. The data highlight donor-specific differences in CYP gene expression, reflecting variation in hepatic metabolic capacity.

UDP-GLUCURONOSYLTRANSFERASE (UGT) GENE EXPRESSION

Next, we analyzed the expression of UDP-glucuronosyltransferase (UGT) enzymes, which play a critical role in detoxification and conjugation reactions.[9,10] The genes examined included *UGT1A9, UGT1A6, UGT1A4, UGT1A3, UGT1A1, UGT2B17, UGT2B15, UGT2B10*, and *UGT2B4*.

UGT2B4 showed the highest expression across most samples, with CPM values exceeding 1000, indicating its prominent role in hepatic glucuronidation. *UGT1A9, UGT1A6, UGT1A4, UGT1A1, UGT2B15*, and *UGT2B10* exhibited moderate expression (100–1000 CPM), while *UGT1A3* and *UGT2B17* were expressed at lower levels (<100 CPM) (Figure 3). Like the CYP profile, PCS-450-010™ lot 015 showed consistently lower expression across all UGT genes.

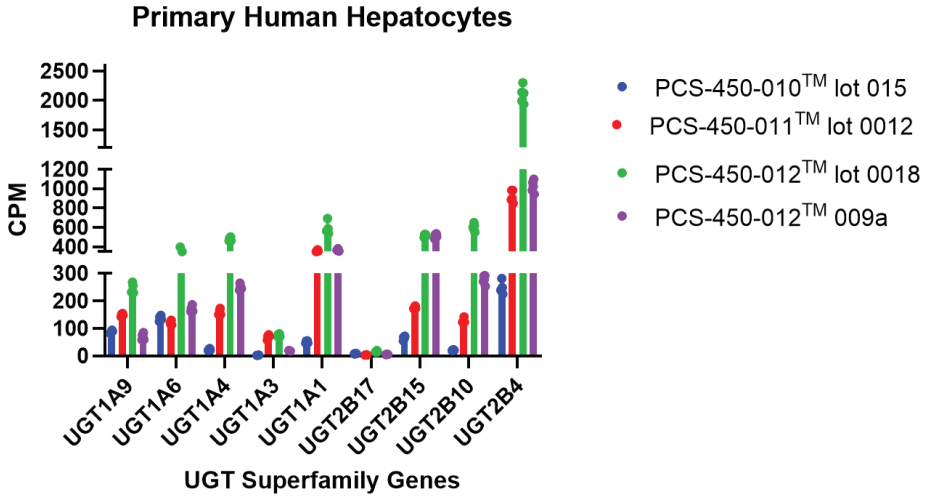


Figure 3: Expression levels of UDP-glucuronosyltransferase superfamily genes in PHHs. The graph illustrates the transcript abundance in counts per million (CPM) of selected UDP-glucuronosyltransferase (UGT) genes in four human donor-derived primary hepatocyte samples. The expression levels for respective donor samples are shown for each gene, with error bars indicating variability across replicates. The data highlights interdonor variability in UGT gene expression, reflecting differences in hepatic conjugation capacity.

ABC TRANSPORTER AND SOLUTE CARRIER GENE EXPRESSION

Finally, we assessed the expression of genes involved in hepatic transport, including members of the ATP-binding cassette (ABC) transporter superfamily and solute carrier organic anion transporters.[11] The genes analyzed were *POU2F1, ABCG2, ABCB1, ABCC2, SLCO2B1, SLCO1B3, SLCO1B1*, and *ABCC3*.

ABCC2 and *SLCO1B1* demonstrated the highest expression levels across most samples, with CPM values approaching or exceeding 500, suggesting their significant role in hepatic transport processes. *ABCB1, SLCO2B1, SLCO1B3*, and *ABCC3* showed moderate expression (100–300 CPM), while *POU2F1* and *ABCG2* were expressed at low levels (<50 CPM) (Figure 4).

Across all gene families analyzed (CYP, UGT, and ABC) interindividual variability was evident, particularly in sample PCS-450-010™ lot 015, which consistently showed reduced expression across most genes. These findings demonstrate the significance of accounting for donor-specific differences in hepatic gene expression and support the use of PHHs as a robust model for pharmacogenomic and toxicological studies.

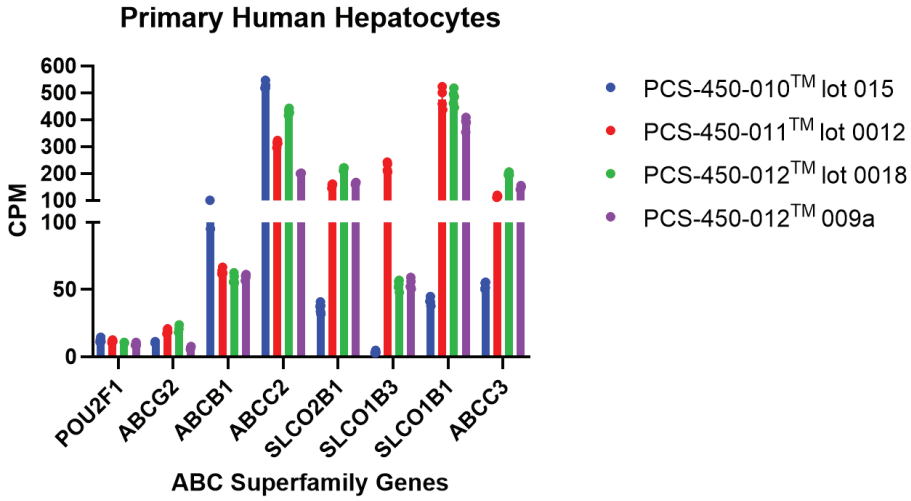


Figure 4: Expression levels of ATP-binding cassette superfamily and related transporter genes in PHHs. The graph shows transcript abundance in counts per million of selected ABC transporters and solute carrier genes across the four hepatocyte samples. The expression levels for respective donor samples are shown for each gene, with error bars indicating variability across replicates. The data illustrates interdonor variability in transporter gene expressions, relevant to hepatic uptake and efflux functions.

TOP GENES DIFFERENTIALLY EXPRESSED ACROSS DONORS

To examine gene expression patterns across donors without imposing prior assumptions, we generated a heatmap of the top 50 differentially expressed genes. Each column corresponds to an individual donor, while each row represents a gene (Figure 5). Expression levels are visualized using a color gradient from black (low expression) to yellow (high expression), spanning a range from < -0.8 to > 1.934. Donors PCS-450-011™ lot 0012 (Blue bar), PCS-450-012™ lot 0018 (Green bar), and PCS-450-012™ lot 009a (Purple bar) exhibit broadly similar gene expression profiles, with only a small subset of different genes showing reduced expression in PCS-450-012™ lot 0018 (Green bar) and PCS-450-012™ lot 009a (Purple bar), respectively. In contrast, PCS-450-010™ lot 015 (Red bar) displays a more balanced distribution, with approximately over two thirds of the analyzed genes showing upregulation and the others vary from low expression to some extent of downregulation.

The data reveals distinct expression profiles across donors, with several mitochondrial genes, including *MT-ND3, MT-CYB, MT-ATP6*, and *MT-CO2*, showing coordinated changes that may reflect differences in mitochondrial activity (Table 3). Genes associated with protein folding and cellular stress responses, such as *HSPA8, HSPA6, HSP90AA1*, and *HSPD1*, vary across donors, suggesting heterogeneity in stress-related pathways (Table 3). The acute-phase and inflammatory markers *SAA1, FGA, SERPINA1, SERPINA3*, and *HP* also show donor-specific expression, potentially indicating variation in immune or inflammatory states. In contrast, structural and translational genes like *ACTB, ACTG1, EEF2*, and *RPL13A* exhibit relatively stable expression, consistent with their roles in maintaining core cellular functions. Notably, genes such as *MYH9, AHNAK*, and *TXNRD1* are more highly expressed in certain donor subsets, while others like *GALDH* and *APOB* are reduced, pointing to transcriptional differences that may be linked to underlying biological or clinical variation.

Table 3: Summary of Differentially Expressed Genes Across Donors

Gene Category	Genes (Examples)	Donor(s) with High Expression	Biological Implication
Mitochondrial	MT-ND3, MT-CYB, MT-ATP6, MT-CO2	PCS-450-011™ lot 0012, PCS-450-012™ lot 009a	Mitochondrial activity differences
Stress Response	HSPA8, HSPA6, HSP90AA1, HSPD1	Variable	Heterogeneity in stress pathways
Acute Phase/ Inflammation	SAA1, FGA, SERPINA1, SERPINA3, HP	Variable	Immune/inflammatory state variation
Structural/Translat	ACTB, ACTG1, EEF2, RPL13A	Stable most donors	Core cellular function maintained
Other	MYH9, AHNAK, TXNRD1, GALDH, APOB	Variable	Linked to biological/clinical variation

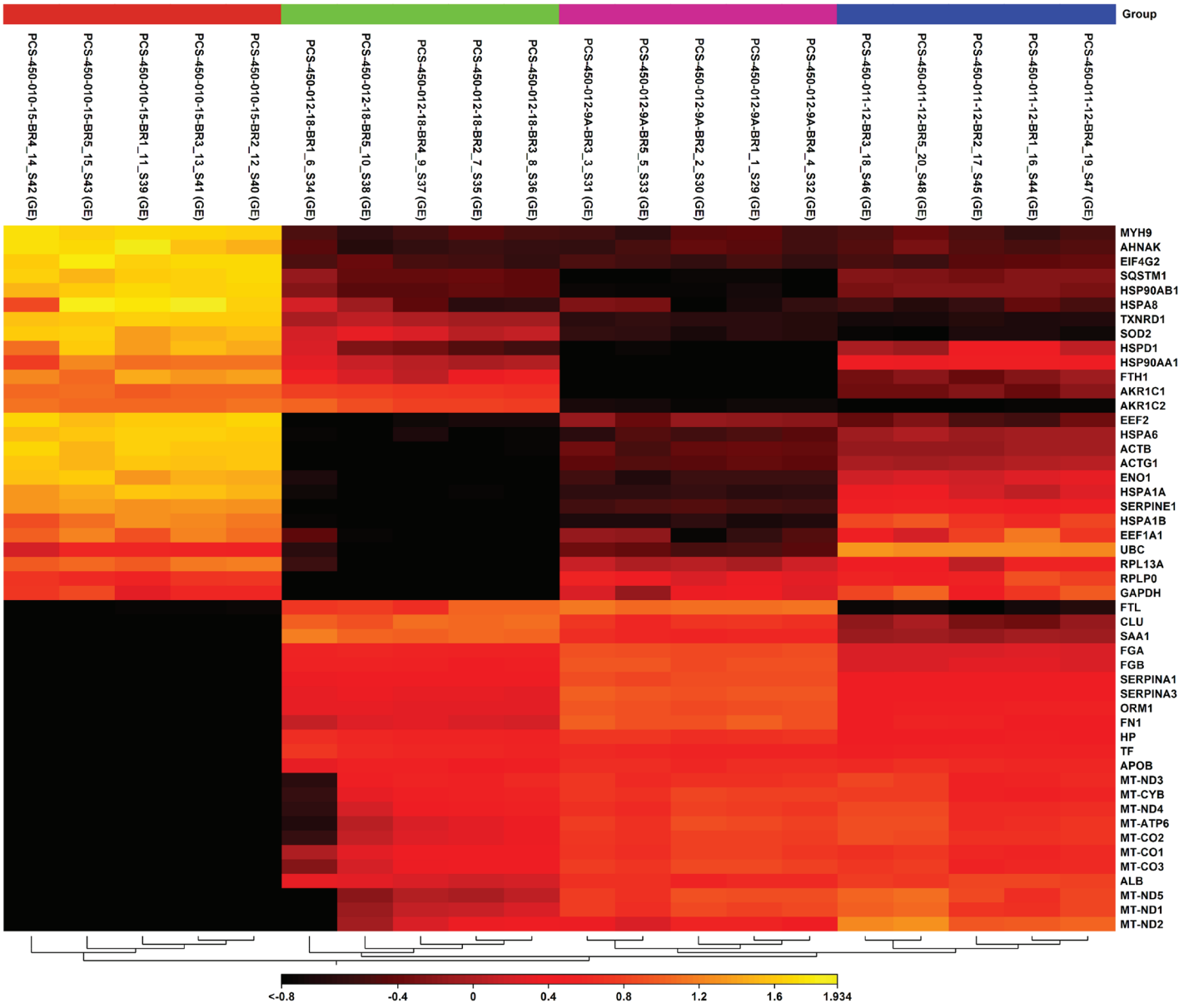


Figure 5: Heatmap shows expression patterns of the top differentially expressed genes across donor hepatocytes. Columns represent individual replicates from grouped by color-coded categories indicating donor identity. Rows correspond to genes. Expression levels are color-coded from low (dark black, <-0.8) to high (bright yellow, 1.934).

HIGH-LEVEL VIEW OF VARIATION IN HEPATOCYTE SAMPLES

To assess the biological consistency among sample replicates and to explore transcriptomic variability across donors, we performed principal component analysis (PCA) and hierarchical clustering (Figure 6). PCA revealed tight clustering of replicates derived from individual donors, confirming technical reproducibility (Table 4). In contrast, samples from different donors formed distinct clusters, indicating donor-specific differences in gene expression profiles. This separation suggests underlying biological variation in hepatocyte function and metabolic capacity. Complementing the PCA, hierarchical clustering heatmaps further demonstrated strong intradonor similarity and clear interdonor divergence (Table 4). Together, these analyses provide a high-level overview of transcriptional heterogeneity in PHHs and emphasize the importance of accounting for donor variability in studies of liver biology and drug metabolism.

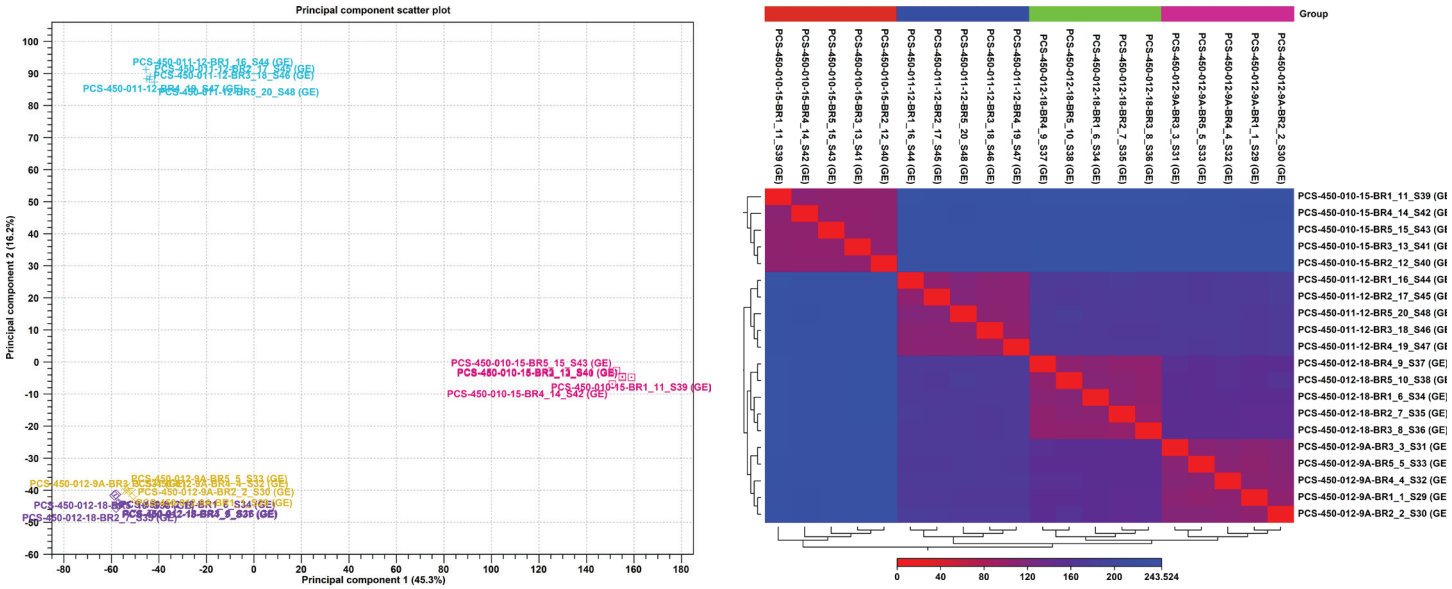


Figure 6: Assessment of across donor variability. (A) Principal Component Analysis (PCA) of PHHs. Donor samples are grouped based on gene expression variance. (B) Hierarchical clustering heatmap of gene expression intensities across human hepatocyte samples.

Table 4: Summary of Technical Reproducibility

Analysis Type	Finding
PCA	Tight clustering within donor replicates
Hierarchical Clustering	Strong intradonor similarity, clear interdonor divergence
Overall Conclusion	High technical reproducibility, biological variability between donors

CONCLUSIONS

This study presents a comprehensive transcriptomic analysis of PHHs, highlighting donor-specific variability in gene expression across key metabolic, conjugation, and transport pathways. The RNA-seq data reveals consistent intradonor reproducibility and distinct inter-donor differences, particularly in the expression of cytochrome P450 enzymes, UDP-glucuronosyltransferases, and ABC transporters. These findings underscore the importance of accounting for individual variability when interpreting hepatocyte-based pharmacogenomic and toxicological data. The integration of quality-controlled workflows and robust analytical pipelines establishes a valuable framework for future studies in liver biology, drug metabolism, and personalized medicine.

REFERENCES

1. Ding M, et al. Integration of ATAC-Seq and RNA-Seq reveals FOSL2 drives human liver progenitor-like cell aging by regulating inflammatory factors. BMC Genomics 24(1): 260, 2023. PubMed: 37173651
2. Wu H, et al. Multi-omics analysis reveals distinct gene regulatory mechanisms between primary and organoid-derived human hepatocytes. Dis Model Mech 18(1): dmm050883, 2025. PubMed: 39878507
3. Liu H, Zhang Y, Ning S. ScRNA-seq combined with ATAC-seq analysis to explore the metabolic balance mechanism of CCl4-induced liver inflammatory injury. Front Immunol 16: 1600685, 2025. PubMed: 40589742
4. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11(3): R25, 2010. PubMed: 20196867
5. Robinson MD, McCarthy DJ, Smyth GK. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1): 139–140, 2010. PubMed: 19910308
6. Zhou X, Lindsay H, Robinson MD. Robustly detecting differential expression in RNA sequencing data using observation weights. Nucleic Acids Res 42(11): e91, 2014. PubMed: 24753412
7. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12): 550, 2014. PubMed: 25516281
8. Zhou Y, Lauschke VM. The genetic landscape of major drug metabolizing cytochrome P450 genes-an updated analysis of population-scale sequencing data. Pharmacogenomics J 22(5-6): 284–293, 2022. PubMed: 36068297
9. Meech R, et al. The UDP-Glycosyltransferase (UGT) Superfamily: New Members, New Functions, and Novel Paradigms. Physiol Rev 99(2): 1153–1222, 2019. PubMed: 30724669
10. Liu W, et al. The Uridine diphosphate (UDP)-glycosyltransferases (UGTs) superfamily: the role in tumor cell metabolism. Front Oncol 12: 1088458, 2023. PubMed: 36741721
11. Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. J Lipid Res 42(7): 1007–1017, 2001. PubMed: 11441126

10801 University Boulevard
Manassas, Virginia 20110-2209

703.365.2700 703.365.2701 sales@atcc.org www.atcc.org

RNA-APN-102025-v01

©2025 American Type Culture Collection. The ATCC trademark and trade name, and any other trademarks listed in this publication are trademarks owned by the American Type Culture Collection unless indicated otherwise. PureLink, Nanodrop, Qubit, Invitrogen, and Thermo Fisher Scientific are trademarks or registered trademarks of Thermo Fisher Scientific, Inc. epMotion and Eppendorf are registered trademarks of Eppendorf SE. Illumina and NextSeq are a registered trademarks of Illumina, Inc. Agilent, TapeStation, and ScreenTape are registered trademarks of Agilent Technologies, Inc. QIAGEN is a registered trademark of QIAGEN GmbH.

These products are for laboratory use only. Not for human or diagnostic use. ATCC products may not be resold, modified for resale, used to provide commercial services or to manufacture commercial products without prior ATCC written approval.