# Functional characterization of the iPSC generated hepatocytes using genome-wide transcriptomic analysis

Ashley Ikwuezunma[1] and Erica DeLeon2, Ana C. Leandro2, Joanne E. Curran2, John Blangero2, Satish Kumar2

[1] School of Medicine, University of Texas Rio Grande Valley, Edinburg TX, USA

[2] Department of Human Genetics & South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Edinburg and Brownsville, TX 78541, USA

**Faculty Mentor:**

Satish Kumar Ph.D.

Associate Professor/Research, Department of Human Genetics & South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, McAllen, TX 78504, USA

**Abstract:**

Advances in iPSC technologies now allow us to consider non-invasive large-scale *in-vitro* disease modeling experiments on disease appropriate cell types in human subjects to better understand human disease pathophysiology, disease genetics and to develop better diagnostic and therapeutic technologies. We performed differential gene expression and functional annotation analysis using genome wide mRNA sequencing data to evaluate the functional and disease modeling potential of iPSC generated hepatocytes. Following the criteria moderated *t* statistics FDR corrected *p-value* ≤ *0.05* and fold change-*absolute* ≥ 2.0, 7,246 genes/transcripts were significantly differentially expressed iPSCs and hepatocytes. The 3,791 of these DE genes/transcripts were significantly upregulated in the hepatocytes and accounted for about 55% of the hepatocytes total expressed transcriptome. The heatmap and principal component analysis suggests a discrete and uniform resetting of cellular transcriptome during iPSC to hepatocyte differentiation. The functional annotation analysis of the 3,791 significantly upregulated hepatocytes genes/transcripts showed significant enrichment hepatocytes associated cellular functions and canonical pathways. The gene known to be associated with various common human liver disorders and toxicities were also significantly enriched in hepatocytes upregulated 3,791 genes/transcripts. These results suggest that iPSC generated hepatocytes have a functional profile very similar to human primary hepatocytes and are suitable for in-vitro modeling of human liver disorders and toxicities.

**Introduction**:

Hepatocellular carcinoma is the fifth most common cancer worldwide and it occurs when the liver becomes cirrhotic or from other diseases affecting the liver. Despite the occurrence of hepatocellular carcinoma, the overall prognosis for the malignancy is poor. Considering the poor prognosis of hepatocellular carcinoma, prevention is key. However, it is difficult to study the underlying cause and risks due to the relative paucity of cases. The non-alcoholic fatty liver disease (NAFLD/hepatic steatosis) is a state of metabolic dysregulation characterized by excessive lipid accumulation into the hepatocytes, often progresses into nonalcoholic steatohepatitis (NASH) and liver cirrhosis and is a major determinant of hepatocellular carcinoma risk. Approximately, 29% of patients with NASH develop cirrhosis. Up to 27% of these further develop hepatocellular carcinoma (1). For NAFLD, one of the principle cell types is the

hepatocyte, the human hepatocellular carcinoma cell lines or immortalized primary hepatocytes, fails to recapitulate the functional profiles of primary hepatocytes and often accumulates genomic and hence functional aberrations (2), invasive liver biopsy of normal individuals to obtain primary liver cells is practically challenging and ethically questionable (3). However, advances in iPSC technologies now allow us to consider non-invasive large-scale deep cellular phenotyping efforts on disease appropriate cell types in human subjects. Robust derivation of iPSC lines and their differentiation into organ specific cell-types now is possible from easily obtained blood cells (4-6). In this project, we utilized existing genome wide mRNA sequencing data generated from de-identified iPSC lines and their differentiated hepatocytes (HEP) of 9 different individuals to evaluate the functional and disease modeling potential of iPSC generated hepatocytes.

**Objective**:

To perform functional characterization of the iPSC generated hepatocytes using genome-wide RNA sequencing based gene expression analysis to better understand the disease modeling potential of the generated hepatocytes.

**Material and Method**:

As indicated in the background, this project utilized existing genome wide mRNA sequencing data generated from 9 validated iPSC lines and their well characterized terminally differentiated hepatocytes. Briefly, the validated iPSC lines were generated using the method described in Kumar et al., (4-6) and then differentiated into mature hepatocytes. The generated hepatocytes analyzed by immunocytochemistry and microscopy, showed characteristic hepatocyte morphology (i.e. flat and polygonal in shape with distinct round nuclei and presence of bile canaliculi) and expressed hepatic markers alpha-fetoprotein (AFP), albumin (ALB), E-cadherin (CHD1), hepatocyte nuclear factor 4 alpha (HNF4α) and actin (Fig. 1). The validated iPSCs and their differentiated hepatocytes were genome wide RNA sequenced using Illumina TruSeq mRNA sample preparation kits and Illumina HiSeq 2500 next generation sequencing platform.

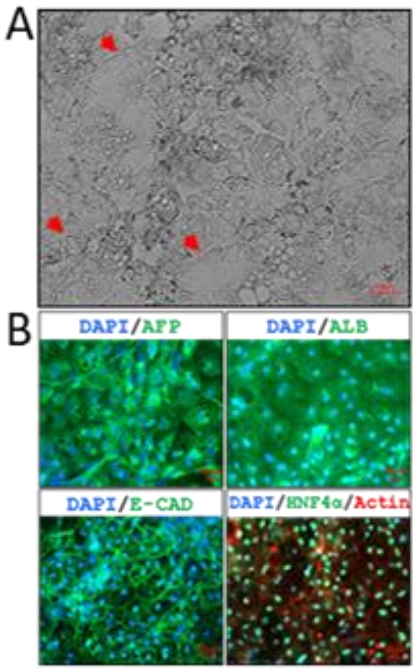

Fig. 1: Morphological and ICC analysis of human iPSC generated hepatocytes. (A) Brightfield image showing hepatocyte morphology and bile canaliculi (red arrows). (B) Showing expression of hepatic markers AFP, ALB, E-cad, HNF4α and Actin.

To further characterize these iPSC generated mature hepatocytes and to identify their disease modeling potential the exiting genome wide mRNA sequencing data generated from the iPSCs and their differentiated hepatocytes were analyzed for differential gene expression between iPSCs and mature hepatocytes. After the pre-alignment quality check, the sequencing reads were aligned to human genome build 38 (hg38) and mapped to Refseq transcripts using StrandNGS software v3.3 (Strand Genomics Inc.). After log transformation and "DESeq" normalization only known mRNAs having a normalized read count (*NRC*) ≥ *20* in all iPSCs and/or all hepatocytes were considered expressed. Only expressed transcripts/genes were then used in differential gene expression analysis. Following the criteria moderated *t* statistics FDR corrected *p-value* ≤ *0.05* and Fold Change absolute (*FC-abs*) ≥ *2.0*, differentially expressed (DE) genes were identified. The DE genes that are significantly upregulated in hepatocytes were then used in functional annotation

and pathway analysis using online tools and data bases using Ingenuity Pathway Analysis (IPA) online platform (Qiagen).

**Results and Discussion**:

Following the aforesaid criteria of *NRC* ≥ *20* in all iPSCs and/or all hepatocytes, a total of 14,682 genes/transcripts were found expressed in iPSCs and their differentiated hepatocytes. The 7,246 of these transcripts were differentially expressed (moderated *t* statistics FDR corrected *p-value* ≤ *0.05* and *FC-abs* ≥ 2.0) between iPSCs and hepatocytes. The 3,791 of these DE genes/transcripts were significantly upregulated in the hepatocytes and accounted for about 55% of the hepatocytes total expressed transcriptome.

The heatmap and principal component analysis (Fig. 2) suggests a discrete and uniform resetting of cellular transcriptome during iPSC to hepatocyte differentiation. The first principal component (Component 1), which represents the expression variance due to iPSC to hepatocyte differentiation (i.e., expression change as a function of iPSC and hepatocyte cellular identities), accounts for 87.06% of the variance observed in DE mRNAs.

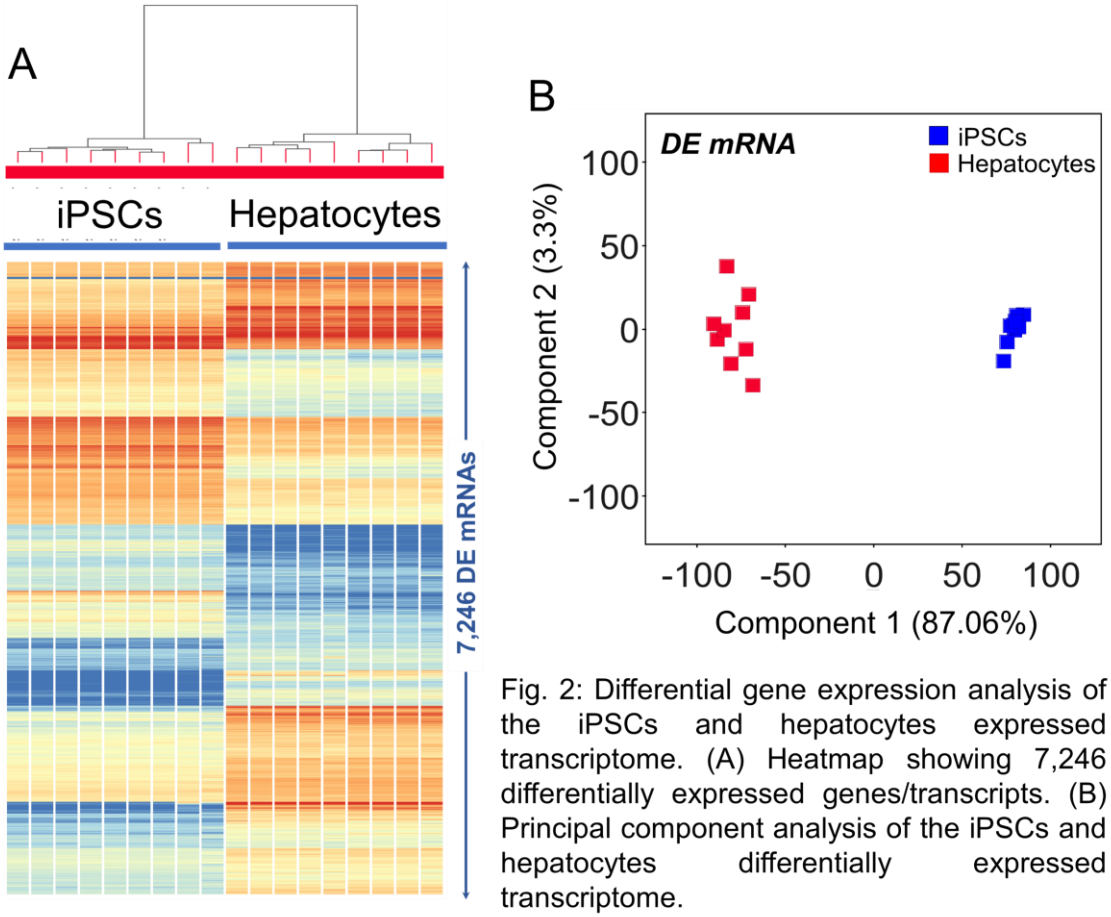


Fig. 2: Differential gene expression analysis of the iPSCs and hepatocytes expressed transcriptome. (A) Heatmap showing 7,246 differentially expressed genes/transcripts. (B) Principal component analysis of the iPSCs and hepatocytes differentially expressed transcriptome.

The functional annotation analysis of the 3,791 significantly upregulated hepatocytes genes/transcripts showed significant enrichment of lipid metabolism (*p-value range: 1.53x10$^{-12}$ to 6.17x10$^{-51}$*), molecular transport (*p-value range: 3.49x10$^{-13}$ to 6.17x10$^{-51}$*), small molecule biochemistry (*p-value range: 1.52x10$^{-12}$ to 6.17x10$^{-51}$*), cellular movement (*p-value range: 3.32x10$^{-13}$ to 5.3x10$^{-46}$*), cell death and survival (*p-value range: 1.12x10$^{-12}$ to 9.19x10$^{-34}$*) functional categories. The top ten highly enriched lipid metabolism functions listed in table 1 suggests that generated cells have a functional profile of human hepatocytes.

Table 1: Top ten highly enriched lipid metabolism function in iPSC generated hepatocytes.

| Functions Annotation | p-value | Molecules Enriched |
|---|---|---|
| Concentration of lipid | 6.17E-51 | 400 |
| Synthesis of lipid | 1.28E-46 | 383 |
| Fatty acid metabolism | 5.83E-35 | 274 |
| Quantity of steroid | 1.75E-34 | 232 |
| Metabolism of membrane lipid derivative | 5.80E-29 | 216 |
| Synthesis of steroid | 1.05E-25 | 140 |
| Concentration of acylglycerol | 1.35E-25 | 170 |
| Concentration of sterol | 7.56E-25 | 153 |
| Synthesis of terpenoid | 1.67E-24 | 147 |
| Steroid metabolism | 4.16E-24 | 131 |

The top canonical pathways that were significantly enriched in the 3,791 significantly upregulated hepatocytes genes included acute phase response signaling (45.8% overlap, $p=3.08x10^{-22}$), FXR/RXR activation (51.6% overlap, $p=2.2x10^{-21}$), LXR/RXR activation (48.8% overlap, $p=5.18x10^{-18}$) and hepatic fibrosis/hepatic stellate activation (37.6% overlap, $p=8.22x10^{-14}$). The acute phase response signaling has a role in immune responses through inflammatory cytokine signaling and restoring tissue homeostasis. Farnesoid X Receptor (FXR) controls numerous metabolic pathways and is activated by bile acid and its intermediates. Thus, FXR plays a crucial role in bile acid regulation and its functions. The retinoid X receptors (RXRs) are also nuclear receptors that facilitate the effects of retinoids through its involvement in retinoic acid-mediated gene activation. Liver X receptor (LXR) is activated by oxysterol ligands and can form a heterodimer with RXR. The LXR/RXR complex is involved in regulation of lipid metabolism, inflammation, and catabolic processes in the liver.

Further we evaluated the liver disease modeling potential of the iPSC generated hepatocytes. The gene known to be associated with various common human liver disorders and toxicities were also significantly enriched in the iPSC generated hepatocytes upregulated 3,791 genes/transcripts (Table 2).

Table 2: Top highly enriched human liver disorders and toxicities functions

| Human Liver Disorders and Toxicities Function | p-value range | Molecules Enriched |
|---|---|---|
| Liver Steatosis | 3.93E-01-2.86E-20 | 143 |
| Liver Hyperplasia/Hyperproliferation | 1.00E00-4.63E-20 | 1401 |
| Hepatocellular Carcinoma | 4.82E-01-6.08E-19 | 321 |
| Liver Necrosis/Cell Death | 3.93E-01-1.38E-17 | 110 |
| Liver Proliferation | 3.93E-01-5.75E-17 | 92 |

## **Conclusion**:

Overall, these results suggest that our iPSC generated hepatocytes have a functional profile very similar to human primary hepatocytes, therefore these iPSC generate hepatocytes are suitable for in-vitro modeling of human liver disorders and toxicities including NAFLD and NAFLD induced liver fibrosis. These iPSC generated hepatocytes are specifically relevant in disease gene discovery, drug screening, and in the development of diagnostic and therapeutic technologies.

## **Acknowledgments**:

The authors are grateful to the STDOI's Mexican American family study participants, whose de-identified iPSC and hepatocytes mRNA sequencing data were used in this study and to STDOI's

Stem Cell Laboratory for providing the data. Part of the RNA sequencing work was conducted in facility constructed through the use of NIH grant 1C06RR020547.

**References**:

1. Cohen JC, Horton JD, Hobbs HH. Human fatty liver disease: old questions and new insights. Science. 2011;332(6037):1519-23. Epub 2011/06/28. doi: 10.1126/science.1204265. PubMed PMID: 21700865; PMCID: PMC3229276.
2. Graffmann N, Ring S, Kawala MA, Wruck W, Ncube A, Trompeter HI, Adjaye J. Modeling Nonalcoholic Fatty Liver Disease with Human Pluripotent Stem Cell-Derived Immature Hepatocyte-Like Cells Reveals Activation of PLIN2 and Confirms Regulatory Functions of Peroxisome Proliferator-Activated Receptor Alpha. Stem Cells Dev. 2016;25(15):1119-33. Epub 2016/06/17. doi: 10.1089/scd.2015.0383. PubMed PMID: 27308945; PMCID: PMC4971413.
3. Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. Hepatology. 2012;55(6):2005-23. Epub 2012/04/11. doi: 10.1002/hep.25762. PubMed PMID: 22488764.
4. Kumar S, Blangero J, Curran JE. Induced Pluripotent Stem Cells in Disease Modeling and Gene Identification. Methods Mol Biol. 2018;1706:17-38. Epub 2018/02/10. doi: 10.1007/978-1-4939-7471-9_2. PubMed PMID: 29423791.
5. Kumar S, Curran JE, Espinosa EC, Glahn DC, Blangero J. Highly efficient induced pluripotent stem cell reprogramming of cryopreserved lymphoblastoid cell lines. J Biol Methods. 2020;7(1):e124. Epub 2020/01/25. doi: 10.14440/jbm.2020.296. PubMed PMID: 31976351; PMCID: PMC6974695.
6. Kumar S, Curran JE, Glahn DC, Blangero J. Utility of Lymphoblastoid Cell Lines for Induced Pluripotent Stem Cell Generation. Stem Cells Int. 2016;2016:2349261. doi: 10.1155/2016/2349261. PubMed PMID: 27375745; PMCID: PMC4914736.