LIVERCELLPATH: A RESOURCE TO HELP EVALUATE HUMAN LIVER CELL LINES AS MODELS FOR HUMAN LIVER BIOLOGY

Brian D. Halligan[1], Yue Chen[1], Antonio Oliveri[1], Maurice Tohme[1], and Elizabeth K. Speliotes[1,2]
1Division of Gastroenterology Department of Internal Medicine,
2Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109

ABSTRACT

Primary hepatocytes are difficult to obtain and dedifferentiate quickly in cell culture making them challenging to use to model chronic liver disease. Human hepatocellular carcinoma (HCC) lines can be easily propagated but are often aneuploid and important genes or biological pathways of interest may not be expressed. We used RNAseq to measure gene expression and to determine exome genotypes from five different HCC lines and compared these to normal human primary hepatocytes. We also determined exonic variation in these cell lines. We created a new software application, **CorrPaths**, that reports gene expression and exonic variation across these cell lines to facilitate identifying cell lines suitable for modeling particular aspects of liver biology and diseases.

RESULTS

Examination of the Reactome XENOBIOTICS pathway (R-HSA-211981) shows that many of the P450 cytochromes expressed by primary hepatocytes (CYP2E1, CYP2C8, CYP2C9, CYP3A4, CYP2A6) are nearly completely absent in all cell lines. Analysis of the GLYCOGEN_SYNTHESIS pathway (R-HSA-3322077) shows that unlike primary hepatocytes and liver tissue, all cell lines express the muscle form of glycogen synthase (GYS1) rather than the liver form of glycogen synthase (GYS2). This result was confirmed by Western blotting of GYS1 and GYS2 in HuH-7 cells compared with liver tissue extracts. Genotyping HuH-7 and HepG2 shows that both HuH-7 and HepG2 cell lines are homozygous for the PNPLA3 SNP (rs738409; I148M) known to increase NAFLD making them good models for this disease. The liver specific output from **CorrPaths**, **LiverCellPath**, can be used online or offline. Its produces an Excel file that can be easily shared with other users, exported to HTML, and can be encrypted for use with sensitive data.

METHODS AND MATERIALS

RNAseq was performed on HepG2, HepG2-C3A, HuH-7, SNU-475 and THLE-1 cell lines and compared with RNAseq from six primary hepatocyte datasets from GEO. Gene expression levels for HCC lines were determined using TopHat and DESeq2 and compared to primary hepatocytes with **CorrPaths**. Exome sequences from HCC lines were mapped to reference human genome sequences using GATK. Some results were confirmed with Western blotting..

Figure 1 Liver Cell Models

Model System	Benefits and Limitations
Liver tissue	Autopsy or biopsy material Benefits: Medical relevance, diseased tissue available Limitations: Mixture of cell types, BBP concerns, cannot use experimental approaches
Primary hepatocytes	Benefits: Biologically pure gold standard Limitations: Difficult/expensive to procure, inter-individual variability, dedifferentiate in culture limits long term experiments, high cost
THLE	SV40 Large T transformed primary hepatocytes Benefits: Growth in culture and cryopreservation, normal chromosome count Limitations: complex culture requirements makes them more difficult to work with
HepaRG	Tumor cell line Benefits differentiate to PHH like + cholangiocyte like cells, more metabolic functions Limitations: need to differentiate before use, short life in culture after differentiation. High commercial cost.
HuH-7, HepG2, SNU475 +	Tumor cell lines Benefits: easy to grow in culture and cryopreserve, widely available, large literature Limitations: highly aneuploid, not all biochemical functions expressed

Novel Features

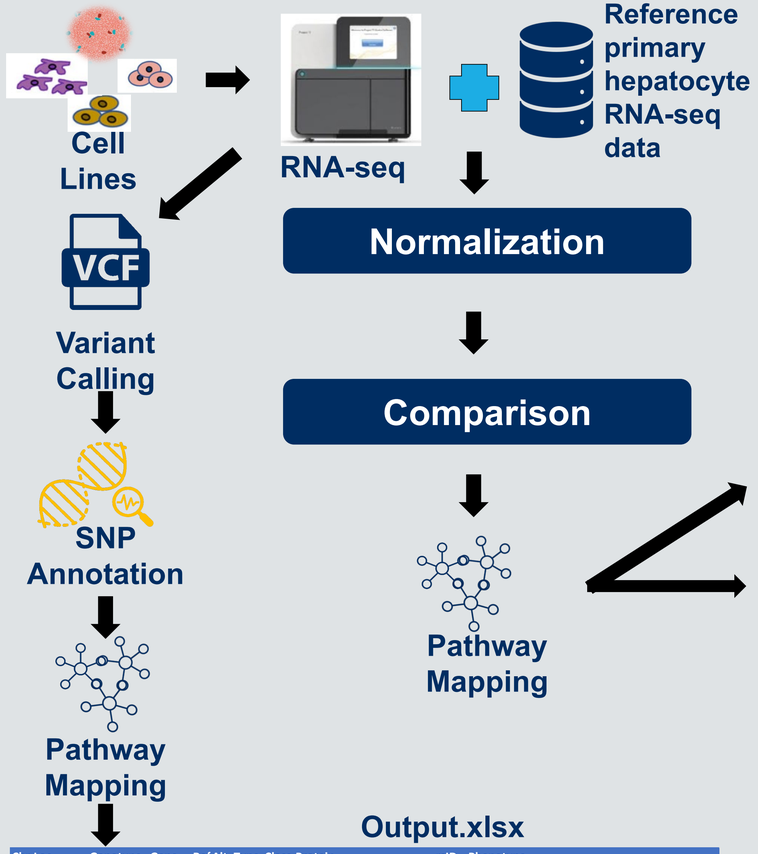
- **CorrPaths** is an experimental design tool that allows investigators to examine the suitability of specific cell line models for the analysis of particular biological pathways
- **CorrPaths**, unlike GSEA, focuses on single genes not whole pathways
- **CorrPaths** compares cell line gene expression to a primary cell standard to identify pathway specific differences in a cell type specific manner
- **CorrPaths** allows users to use annotated pathways from multiple sources, combine pathways that involve a gene of interest or define their own gene lists for analysis.
- **CorrPaths** combines both gene expression and gene variation in one application and allows for the identification of cell line models carrying specific variations or lack of gene express as disease models.

Conclusions

CorrPaths can help in the design of functional experiments by identifying gene expression or variant issues that can interfere with the study of specific biological pathways of interest

Future

CorrPaths approach can be applied to other cell and tissue types to make specific version for other biological systems using different RNAseq data



Output.xlsx

Chr	Loc	Genotype	Gene	Ref	Alt	Type	Class	Protein	rsID	Phenotype
22	44324727	2	PNPLA3	C	G	exoni	nonsy	PNPLA3:NM_025225:exon	rs738	HDL cholesterol:Total cholesterol:Nonalcoholic fatty liver disease:Total cholesterol:HDL cholesterol:LDL cholesterol:HDL cholesterol:SAMM50 cis expression in subcutaneous fat:SAMM50 cis expression in omental fat:CT hepatic steatosis:Nonalcoholic fatty liver disease:LDL cholesterol:Soluble intercellular adhesion molecule 1 (ICAM-1):Alanine aminotransferase (ALT) in plasma:Hepatic steatosis assessed by CT:Non-alcoholic fatty liver disease:Alanine aminotransferase (ALT):Non-alcoholic fatty liver disease:Fibrosis (brunt grade 1/2/3) in Nonalcoholic fatty liver disease:Fasting blood glucose in Nonalcoholic fatty liver disease:Fatty liver:Non-alcoholic fatty liver disease (Matteoni
22	44324730	2	PNPLA3	C	T	exoni	synor	PNPLA3:NM_025225:exon	rs738	HDL cholesterol:Total cholesterol:Total cholesterol:HDL cholesterol:LDL cholesterol:Fatty liver
22	44342116	2	PNPLA3	A	G	exoni	nonsy	PNPLA3:NM_025225:exon	rs229	Stabilized warfarin dose:HDL cholesterol:Fasting blood glucose:Total cholesterol:Mitral annular calcium

Figure 2 Genetic Variants PNPLA3 in HUH-7 cell line
Genomic Data Workbook contains genomic variant data for each of the cell lines. Genomic variants from HuH-7 cells for the PNPAL3 gene. This shows that the HUH-7 cells are homozygous for the most significant MASLD causing human variant, RS738409.

Figure 3 Pathway Index Sheet
Index pathway worksheet allows direct access to each of the pathways in the workbook. Clicking on pathway ID link opens browser window with pathway details.

	Pathway ID	Pathway Description	gene count
71	R-HSA-6814848 t	REACTOME_GLYCEROPHOSPHOLIPID_CATABOLISM	7
72	R-HSA-8982491 t	REACTOME_GLYCOGEN_METABOLISM	27
73	R-HSA-3322077 t	REACTOME_GLYCOGEN_SYNTHESIS	16
74	R-HSA-70171 tar	REACTOME_GLYCOLYSIS	72
75	R-HSA-1630316 t	REACTOME_GLYCOSAMINOGLYCAN_METABOLISM	124
76	R-HSA-1660662 t	REACTOME_GLYCOSPHINGOLIPID_METABOLISM	45
77	R-HSA-389661 ta	REACTOME_GLYOXYLATE_METABOLISM_AND_GLYCINE_DEGRADATION	31
78	R-HSA-189451 ta	REACTOME_HEME_BIOSYNTHESIS	14
79	R-HSA-189483 ta	REACTOME_HEME_DEGRADATION	15
80	R-HSA-70921 tar	REACTOME_HISTIDINE_CATABOLISM	8
81	R-HSA-2022928 t	REACTOME_HS_GAG_BIOSYNTHESIS	31
82	R-HSA-2024096 t	REACTOME_HS_GAG_DEGRADATION	22
83	R-HSA-2142850 t	REACTOME_HYALURONAN_BIOSYNTHESIS_AND_EXPORT	5
84	R-HSA-2142845 t	REACTOME_HYALURONAN_METABOLISM	17
85	R-HSA-2160916 t	REACTOME_HYALURONAN_UPTAKE_AND_DEGRADATION	12
86	R-HSA-1483115 t	REACTOME_HYDROLYSIS_OF_LPC	9
87	R-HSA-1483249 t	REACTOME_INOSITOL_PHOSPHATE_METABOLISM	48
88	R-HSA-163685 ta	REACTOME_INTEGRATION_OF_ENERGY_METABOLISM	108

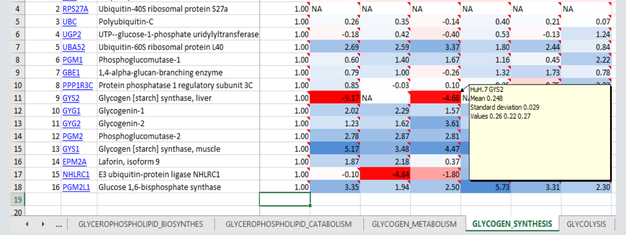


Figure 4 Glycogen Synthesis Pathway
Individual pathway worksheets show the genes in the pathway ranked by their expression in the reference primary hepatocytes. For each cell line, the $\log_2$ fold ratio with respect to the reference primary hepatocyte of gene expression is shown. Red=decreased Blue=increased relative to reference. NA indicates no gene expression was observed. ToolTips allow viewing of raw gene expression data. Clicking on gene name opens UniProt page for the protein product of the gene.

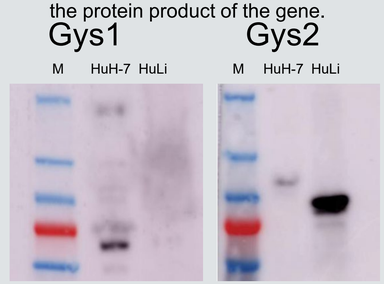


Figure 5 Western blot of Gys1 and Gys2
Western blot of protein extracts from HuH-7 cells and normal human liver tissue (Hu Li) with anti-Gys1 and anti-Gys2 antibodies. Results show that although the liver tissue expresses the liver form of glycogen synthase (Gys2) and does not express the muscle form of glycogen synthase (Gys1), consistent with the RNAseq gene expression data, in HuH-7 cells, the Gys1 form is expressed and the Gys2 form is not. M=marker.