

Research Article

A Transcriptomic Signature of Mouse Liver Progenitor Cells

Adam M. Passman,^{1,2} Jasmine Low,^{1,3} Roslyn London,¹ Janina E. E. Tirnitz-Parker,^{4,5} Atsushi Miyajima,⁶ Minoru Tanaka,⁶ Helene Strick-Marchand,⁷ Gretchen J. Darlington,⁸ Megan Finch-Edmondson,^{9,10} Scott Ochsner,¹¹ Cornelia Zhu,^{1,2} James Whelan,^{3,12} Bernard A. Callus,^{1,2,13} and George C. T. Yeoh^{1,2}

¹ School of Chemistry and Biochemistry, The University of Western Australia, Crawley, WA 6009, Australia

² The Centre for Medical Research, Harry Perkins Institute of Medical Research, Nedlands, WA 6009, Australia

³ ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA 6009, Australia

⁴ School of Biomedical Sciences, Curtin Health Innovation Research Institute, Curtin University, Bentley, WA 6102, Australia

⁵ School of Medicine and Pharmacology, The University of Western Australia, Fremantle, WA 6160, Australia

⁶ Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo 113-8654, Japan

⁷ Unité de Génétique de la Différenciation, Institut Pasteur, 75015 Paris, France

⁸ Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA

⁹ Department of Physiology, NUS Yong Loo Lin School of Medicine, Singapore 117411

¹⁰ Mechanobiology Institute (MBI), National University of Singapore, Singapore 117411

¹¹ Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

¹² Department of Animal, Plant and Soil Sciences, La Trobe University, Melbourne, VIC 3086, Australia

¹³ School of Health Sciences, The University of Notre Dame Australia, Fremantle, WA 6959, Australia

Correspondence should be addressed to George C. T. Yeoh; george.yeoh@uwa.edu.au

Received 8 March 2016; Revised 4 August 2016; Accepted 14 August 2016

Academic Editor: James Adjaye

Copyright © 2016 Adam M. Passman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Liver progenitor cells (LPCs) can proliferate extensively, are able to differentiate into hepatocytes and cholangiocytes, and contribute to liver regeneration. The presence of LPCs, however, often accompanies liver disease and hepatocellular carcinoma (HCC), indicating that they may be a cancer stem cell. Understanding LPC biology and establishing a sensitive, rapid, and reliable method to detect their presence in the liver will assist diagnosis and facilitate monitoring of treatment outcomes in patients with liver pathologies. A transcriptomic meta-analysis of over 400 microarrays was undertaken to compare LPC lines against datasets of muscle and embryonic stem cell lines, embryonic and developed liver (DL), and HCC. Three gene clusters distinguishing LPCs from other liver cell types were identified. Pathways overrepresented in these clusters denote the proliferative nature of LPCs and their association with HCC. Our analysis also revealed 26 novel markers, LPC markers, including *Mcm2* and *Ltbp3*, and eight known LPC markers, including *M2pk* and *Ncam*. These markers specified the presence of LPCs in pathological liver tissue by qPCR and correlated with LPC abundance determined using immunohistochemistry. These results showcase the value of global transcript profiling to identify pathways and markers that may be used to detect LPCs in injured or diseased liver.

1. Introduction

Liver progenitor cells (LPCs) have garnered substantial interest in the field of liver biology due to their enormous regenerative capacity [1, 2] which positions LPCs as a strong candidate for cell therapy to treat liver disease. Whilst their

link to liver disease [3–5] and potential as a liver cancer stem cell constrain their utility, they afford a model to investigate the molecular and cellular mechanisms that underlie tumorigenic transformation.

In many chronic liver diseases, when proliferation of hepatocytes is limited, LPCs replicate and differentiate, providing

an alternate source of hepatocytes needed for regeneration [6]. This “LPC response” is observed in humans in cases of alcoholic liver disease, hemochromatosis, hepatitis C and B infection, and HCC [4, 5, 7–9]. Other liver pathologies may also involve expansion of LPCs; however, their role in these instances is undefined [10]. A rodent model of chronic liver disease, the choline-deficient, ethionine-supplemented (CDE) diet which results in fatty liver and subsequently fibrosis and cirrhosis, induces LPCs in rats [11] and mice [12] and has been used extensively for studying LPCs. This model was used to reveal Wnt and Notch-controlled signaling that specifies LPCs to the hepatocytic and cholangiocytic lineages, respectively [13]. Other strategies to induce LPCs include administration of 3,5-diethoxy-carbonyl-1,4-dihydrocollidine (DDC) [11] and acetyl aminofluorene to block hepatocyte proliferation followed by partial hepatectomy [14] and administration of alkylating agents, monocrotaline [15] or retrorsine [16].

The degree to which LPCs contribute to liver regeneration is controversial. In contrast to several publications that specify a role for LPCs in liver regeneration [6, 17, 18], a recent review [19] highlighted lineage-tracing studies [20, 21] in which LPCs did not. This discordance may be dependent on the method used to induce liver pathology. Nevertheless, using different approaches, several laboratories have established LPC lines that are capable of *in vitro* differentiation into cholangiocytes and hepatocytes [22–25].

LPCs are a heterogeneous cell population and whether LPCs induced by different etiologies and via various models are identical or even share similar genetic profiles is unknown. Comparing lines generated in different laboratories thus mandates caution when interpreting findings to reach generalized conclusions on the biology of LPCs. These issues are compounded as current LPC markers often stain cholangiocytes as well, highlighting a requirement for additional markers.

Proliferation of LPCs and development of liver cancer have been correlated. In mice, inducible liver-specific expression of c-mycelocytomatosis (c-Myc) oncogene stimulated proliferation of an immature cell population resulting in HCC [26]. Upon c-Myc inactivation, tumor cells differentiated into hepatocytes and cholangiocytes, suggesting the HCCs formed from LPC-like cells. Fibrolamellar HCCs are enriched with cancer stem cells resembling “biliary tree stem cells,” LPC precursors [27]. In another study, comprehensive gene array analysis of human HCC identified gene amplification of yes-associate protein (YAP) and cellular inhibitor of apoptosis (cIAP1) as potential oncogenes [28]. In the same study, overexpression of cIAP1 in a p53 null LPC cell line substantially reduced tumor onset time and increased tumor burden. Additional links between the Hippo pathway, LPCs, and HCC development have been identified [29–31]. High YAP expression marks progenitor cells within the liver and expression within hepatocytes is sufficient to differentiate them into LPC-like cell [32]. This may have implications for HCC development, especially for tumors described as poorly differentiated. It is therefore of interest to obtain a transcript profile of LPCs, both to aid discovery of markers for detection and to further study their association with HCC. We hypothesize

that LPCs will display a transcript profile that reflects some, if not many, of the “hallmarks” of cancer, in particular, dysregulated cell proliferation and resistance to cell death [33].

In this study, we performed transcriptome analysis of established and well-characterized LPC lines and pooled LPC transcriptome data from many laboratories to determine whether LPCs isolated by different groups using different approaches are similar or distinguishable based on their transcriptome. Additionally, we interrogated several repositories of transcriptome data for comparison with other cell and tissue types including datasets of muscle and embryonic stem cell lines, embryonic and developed liver, and HCC. We identified signaling pathways and promoters that are consistently active in LPC lines to provide greater insight into their biology. Lastly, we identified novel LPC markers and validated their utility to identify LPCs in different mouse models of liver pathology.

2. Materials and Methods

2.1. Capturing the Transcriptomes. The LPC lines were derived by different procedures in multiple laboratories. Bipotential murine embryonic liver (BMEL) cell lines were derived from CBA/J × C57Bl/6J cross embryonic livers at 14 days post coitum. Bipotential murine oval liver (BMOL) cell lines were derived from 7-week-old male C57Bl/6 strain wild-type mice livers subjected to 3 weeks of CDE diet. Tokyo-LPC (T-LPC) was isolated by selection of EPCAM+ cells from DDC-injured livers. The p53-immortalized liver (PIL) 1, 2, 3, 4, and 5 lines were derived from adult p53 null (C57Bl/6 strain) mice. These LPC lines have been characterized previously [22, 23, 25, 34, 35], by morphology, bipotentiality, proliferative ability, and expression of LPC markers. For microarrays generated in our laboratory, LPC RNA was isolated using TRIzol (Invitrogen) and RNeasy mini kit (Qiagen) with DNase I treatment following manufacturer’s recommendations. RNA quality was verified using an Agilent 2100 Bioanalyser. Synthesis and labeling of cDNA and cRNA, fragmentation, hybridization, washing, staining, and scanning of the Affymetrix Mouse Genome 430 2.0 GeneChips including quality control checks were performed according to the manufacturer’s instructions.

2.2. Normalization of Microarray Experiments. Microarray experiments (381) that fulfilled the requirements of having at least three replicates performed on the Affymetrix Mouse Genome 430 2.0 or 430A GeneChips platforms were downloaded from the National Centre for Biotechnology Information Gene Expression Omnibus (GEO) Repository. Datasets included those generated from LPC lines, C2C12 muscle cell lines, embryonic stem cells, DL, HCC, and embryonic liver. We performed a meta-analysis of 405 microarrays comprising the 381 publically available microarrays and 24 LPC microarrays generated by our laboratory, using the BMOL, T-LPC, and PIL1-5 LPC lines. The arrays generated from these seven LPC lines have been uploaded to the GEO (GSE85114). In addition to these seven LPC lines, our meta-analysis included three separate, publically

sourced BMEL lines. A summary of details pertaining to all arrays used in this manuscript is included in Supplementary Table 1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2016/5702873>.

Probe sets from all 405 arrays were filtered to include only present sets since the proportion of absent ones can considerably affect the data median. Presence for a sample was established following MAS (MicroArray Suite) 5.0 summarization (p value < 0.05) when the probe set was present in at least 75% of replicates. The median gene expression and Median Absolute Deviation (MAD) were determined for each microarray, as a representation of the spread of the data. Median standardization of data involved subtracting the median array expression from the expression of each probe set and then dividing the result by the MAD of each microarray. This produces a median of 0 for each microarray. Values were then multiplied by 500 to increase the spread and standardize to a common spread that is within the observed range of MAD for most of the microarrays. Finally, a constant of 750 was added to the values to increase the median to 750, also within the observed range of medians.

As the Mouse Genome 430 2.0 GeneChip contains more probes than the 430A GeneChip, the dataset was truncated to include only the 430A probes. \log_2 transformation and quantile normalization were performed to standardize the distribution of the probe set intensities to an appropriate scale.

2.3. Microarray Analysis. Analysis of Variance (ANOVA) was performed to identify genes with differing means between more than two groups of conditions [36]. Partek (Missouri, USA) was used to incorporate contrasts within ANOVA, prioritizing the comparison of LPCs to other cell types. Genes with p values less than 0.01 were considered differentially expressed.

Principal component analysis (PCA) [37] was performed using Partek to visualize the natural grouping of the arrays based on global gene expression data. The three principal components that captured most differences were displayed in three-dimensional scatter plots. Similar microarrays are grouped together and more diverse microarrays are spaced further apart.

Expression datasets were hierarchically clustered in Partek under the parameters of complete (>1000 rows) or average (<1000 rows) linkage, Euclidean distance, and agglomerative clustering. Unsupervised clustering enabled relationships between rows and columns to be organized into tree diagrams without imposing subjective bias on the number or size of the clusters. Clusters were systematically colored and selected based on the distances calculated and displayed in the tree diagram branch arms using the “color by cluster” function in Partek.

The coefficient of variation was determined for each gene by dividing the standard deviation of each gene expression pattern by its average expression value, as previously described [38], to identify gene expression profiles with minimal variation across the cell types analyzed. The coefficient of variations and Pearson correlation coefficients were calculated in Microsoft Excel.

2.4. Pathway and Promoter Analysis. Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to ascertain overrepresented pathways and perform gene identifier conversions. The pathway overrepresentation used an EASE score (a modified Fisher exact test) to calculate p values and determine if proportions of the categories differed [39, 40]. A significance threshold was set to $p < 0.05$. Only pathways that met a $p < 0.05$ threshold and contained 10 or more genes from the input cluster were included in our analysis.

Promoter Analysis and Interaction Network Toolset (PAINT) was used to search for overrepresented transcription factor binding elements in gene promoters [41]. The TransFac public database containing known transcription factor binding sites was used to search 2000 bases upstream of the transcriptional start sites of the input Entrez Gene lists. Other parameters selected were the match filter option set to minimize false positives, the core similarity threshold set to 1.0, and binding elements searched on both strands of DNA.

2.5. qPCR Analysis to Confirm Presence of Predicted LPC Markers in Liver Tissue. Mice were subjected to one of four regimes: 3,5-diethoxy-carbonyl-1,4-dihydrocollidine (DDC, $n = 4$) induced hepatotoxicity, CDE injury ($n = 6$), or one of two transgenic models of immune-mediated hepatitis, Met-Kb ($n = 3$) and 178.3 ($n = 3$). Experimental mice were subjected to the DDC regime for 28 days or 21 days for the remaining models. Controls ($n = 3, 8, 3, 3$, resp.) were used to contrast gene expression levels. All animal experiments complied with the guidelines specified by the National Health and Medical Research Council of Australia. Total RNA was extracted from the livers of the mice described above and cDNA was transcribed using Tetro RT™ reverse transcriptase (Bioline). TaqMan two-step, real-time quantitative PCR (qPCR) with hydrolysis probes from the Universal Probe Library and the Light Cycler 480 (LC480) Probe Mastermix (Roche) was used to quantify mRNA expression. Genes selected to predict the presence of LPCs in injured livers were muscle-restricted coiled-coil protein (*Murc*) from group α , latent transforming growth factor beta binding protein 3 (*Ltbp3*) and neural cell adhesion molecule 1 (*Ncam1*) from group β , and minichromosome maintenance deficient (*Mcm2*) and pyruvate kinase isozyme M2 (*M2pk*) from group γ . The expressions of known LPC markers, *Cd24a* antigen and *Sox9*, were also determined. Primer sequences are available in Supplementary Table 2. Data was analyzed using the LC480 Relative Quantification Software. A five-point standard curve was generated for each gene, using pooled cDNA of each sample. A calibrator sample was also included in every qPCR experiment, for which each qPCR run was replicated three times. Data was included when amplification efficiencies/correlation for the assay was $\geq 95\%$. Significant changes in gene expression were determined by comparing control to experimental samples using a two-tailed Student's t -test. Data are displayed relative to TATA-box associated factor 4A (*Taf4a*) mRNA expression and normalized to controls.

2.6. Quantifying Liver Progenitor Cell Numbers In Vivo. Sections of formalin-fixed, paraffin-embedded CDE, DDC, Met-Kb, and 178.3 mouse liver were dewaxed and stained with hematoxylin and eosin to ascertain liver morphology or were stained with anti-panCK to determine LPC response. PanCK staining was achieved by first performing antigen retrieval using proteinase K (Dako, VIC) before applying a 1 : 400 dilution of anti-panCK (#Z0622, Dako) overnight at 4°C. Staining was processed and visualized with DAB using the LSAB system (Dako) according to the manufacturer's instructions. Stained sections were scanned at 40x magnification using an Aperio ScanScope XT. The Positive Pixel Count v9.1 algorithm within the ImageScope software was used to determine LPC response a.k.a. "panCK positivity," the number of pixels positive for panCK staining as a percentage of total tissue pixels.

3. Results

3.1. Principal Component Analysis Confirms Successful Normalization of Microarrays. Since many of the microarray experiments were performed in different laboratories, non-biological variables resulting in differing gene expression and transcript abundance are expected. These factors include chip-to-chip variability, different culture conditions including medium additives (e.g., growth factors and serum), cell density at time of harvest, RNA quality/quantity, and differences in reagents, kits, and equipment. It was therefore necessary to normalize the data before comparisons were made. All raw microarray data were subject to identical linear scaling and normalization to reduce the impact of nonbiological factors.

To test the success of this normalization, PCA was performed and the three principal components that captured most differences (37.5% of the total variation) in the datasets were plotted (Figure 1). PC1, PC2, and PC3 captured 19.9%, 11.6%, and 6.0% of variation in the datasets, respectively, and are shown from three angles to aid in distinguishing each cell/tissue type cluster (Figures 1(a)–1(c)). Without adequate normalization, different arrays would likely appear scattered across the three-dimensional PCA plot. In this instance, the normalized arrays clearly grouped on the PCA plot according to tissue type, and differences between tissue types were greater than those within each grouping. For instance, LPC lines (orange nodes) clustered closely together and away from DL microarrays (red nodes). Of all cell/tissue groupings, the DL and HCC arrays clustered closest together, yet they still occupied distinct regions. This distinction is best visualized on the first principal component (PC1), which accounts for the greatest differences in the analysis (Figure 1(a)). Crucially all data was generated in an unsupervised manner, that is, without introducing bias from tissue type. These findings were demonstrative of successful normalization and justified subsequent analyses to identify a distinct LPC gene expression pattern.

3.2. Analysis of Hierarchical Clusters Reveals Distinct LPC Expression Profiles, Pathways, and Promoters. ANOVA was applied to the tissue type groupings to distinguish the LPC

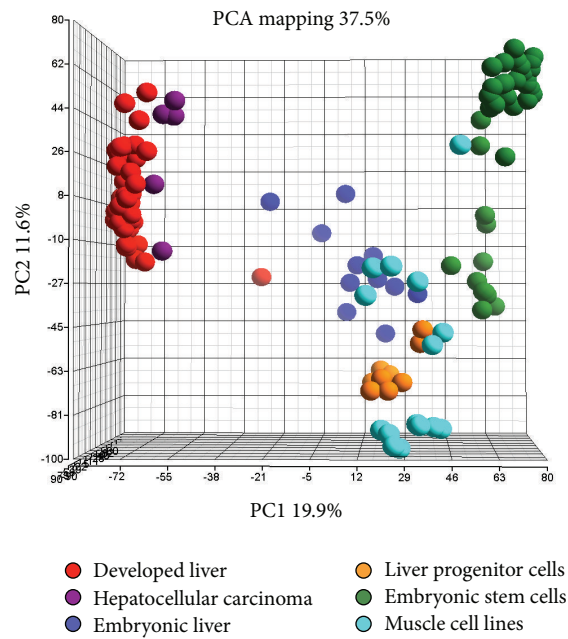
transcriptome from other cell types. Analysis identified 8623 differentially expressed probe sets when comparing LPCs to all other groups. From this dataset, five unique clusters were identified and named clusters A to E (Figure 2). Cluster A represents probe sets with equivalent or higher expression in LPCs compared to other cell types. Cluster B includes probes that display high variability within similar tissue types and thus is largely uninformative. Cluster C represents probe sets with high expression in "immortalized" cell lines including LPCs, muscle, embryonic stem cells, and embryonic liver whilst being poorly expressed in DL and HCC. Cluster D includes probes with higher expression in DL and HCC samples, and cluster E represents probe sets whose expressions were highly abundant in the majority of samples. Clusters A, C, and D presented the greatest differences between LPCs and other cell/tissue groups and thus were the focus of further analysis. Figure 2 data is provided as a supplementary spreadsheet (Figure 2 Matrix).

As evident from the scattered color within each cluster, not all probe sets closely aligned within each cluster. Stringent filters were applied to more precisely define the LPC transcriptome. The MAS 5.0 algorithm was applied to make present/absent calls for probe sets. For clusters A and C, only probe sets that were present in 90% of LPC arrays were retained. In contrast, only sets absent in 90% of LPC arrays were retained for cluster D, since they were largely absent in LPCs. The DAVID and KEGG pathway tools were subsequently applied to the truncated cluster A, C, and D lists to functionally categorize and identify patterns of biological significance within each. The resulting overrepresented pathways ($p < 0.05$), the number of cluster genes found (counts), and the total genes within each pathway (pathway total) are shown in Figure 3.

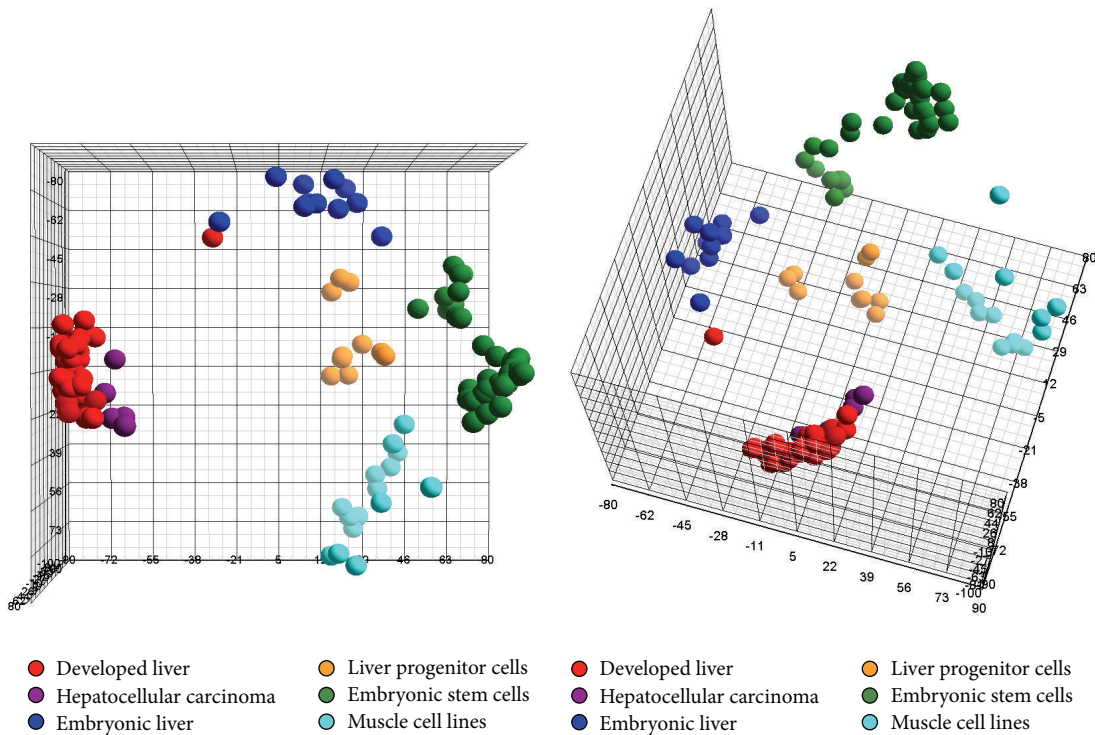
DAVID was able to annotate 254 genes to KEGG pathways from the filtered cluster A probe set list. In the overrepresented mitogen-activated protein kinase (MAPK) signaling pathway there were 25 counts from a MAPK pathway total. Other cluster A pathways included GnRH signaling, cell cycle, transforming growth factor- (TGF-) beta signaling, and epidermal growth factor (EGF) receptor ErbB signaling. Within this cluster, PAIN analysis identified nuclear factor (erythroid-derived 2)-like 2 (*Nrf2*), upstream stimulatory factor (*Usf1*), cyclic-AMP response element binding protein (*Creb*), and activating transcription factor (*Atf1*) as predicted binding elements playing an important role in the coregulation of LPC associated genes.

Using the truncated cluster C probe set list, DAVID annotated 426 genes to KEGG pathways. Overrepresented pathways included cell cycle, DNA replication, pyrimidine and purine metabolism, and p53 signaling pathways (Figure 3). There were also several overrepresented predicted binding elements for cluster C including growth factor independent 1 transcription repressor (*Gfi1*), E2F transcription factor 1 (*E2f1*), melanocyte-inhibiting factor (*Mif*), and cartilage oligomeric matrix protein (*Comp*).

Clusters A and C are both highly expressed in LPCs whereas probe sets in cluster D are lower in LPCs compared to DL and HCC samples (Figure 2). The 90% absent filtered list for cluster D was annotated to 359 genes and



(a)



(b)

(c)

FIGURE 1: Principal component analysis (PCA) displays transcriptome clustering sorted by tissue type. Each node represents the average of at least 3 microarray replicates and is color-coded according to tissue type. The three principal components that captured most differences in the datasets are plotted with three different three-dimensional views of the same PCA visualization shown in (a), (b), and (c).

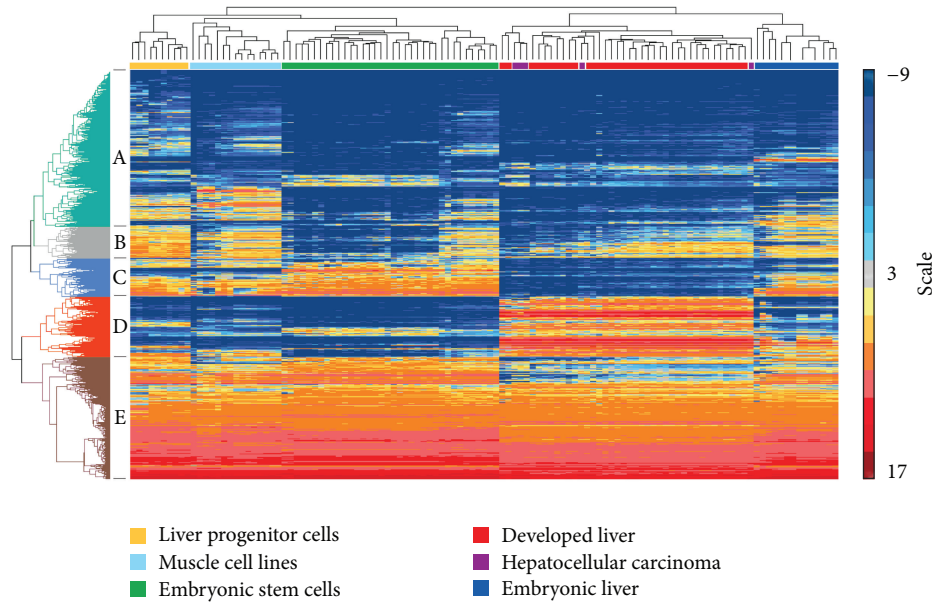


FIGURE 2: Hierarchical clustering identifies 5 clusters (A, B, C, D, and E) of distinct gene expression data for liver progenitor cells. ANOVA was performed to compare liver progenitor cells to all other cell/tissue type groups. The 8,623 probe sets that displayed expression levels significantly different to LPCs ($p < 0.01$) were subsequently clustered according to the parameters of Euclidean distance and complete linkage.

overrepresented KEGG pathways were involved in fundamental liver functions, including immunity, detoxification, and lipid metabolism. Some of these pathways were complement and coagulation cascades, peroxisome proliferator-activated receptor (PPAR) signaling, linoleic acid, and multiple metabolic pathways. Following PAINT analysis, overrepresented promoter elements discovered include jun oncogene, hepatocyte nuclear factors 1 and 4 (*Hnfl* and *Hnf4*), and nuclear factor of kappa light polypeptide gene enhancer in B cells (*NfκB*), many of which support liver function.

The probe set IDs (ordered according to pathway), gene names, symbols, and the fold change between LPC and DL for clusters A, C, and D are displayed in Supplementary Tables 3–5. A summary of the promoter elements for these clusters is available as Supplementary Table 6.

3.3. The LPC Transcriptomic Profile Reveals Known and Novel Marker Genes. To identify individual genes representative of LPCs, probe sets corresponding to known or widely used LPC markers were selected. These included A6 actin-binding protein, alpha fetoprotein, albumin, CD34 antigen, keratins 7, 8, 14, 18, and 19, c-kit oncogene, Cx32 and Cx43 gap junction proteins, Delta-like 1 homolog, gamma glutamyl transferase 1, pyruvate kinase isoenzyme type M2, glutathione S-transferase, ataxin 1, epithelial cadherin, and neural cell adhesion molecule [25, 42–45]. Pearson correlation coefficients were calculated for comparisons between the above probe sets and each of those in Figure 2. This provided a statistical relationship between the known LPC markers and the total list of differentially expressed probe sets in LPCs and produced a list of 355 sets with a correlation >0.9 .

These probe sets were hierarchically clustered to visualize expression patterns in various tissue types (Figure 4(a)). Figure 4(a) data is provided as a supplementary spreadsheet (Figure 4A Matrix). Although all probe sets correlated with LPC markers, not all are highly expressed in LPCs. This can be explained as LPC markers include a mixture of hepatocyte and cholangiocyte markers, which may not be expressed by all types of LPCs. Figure 4(b) is a magnification of an area of interest highlighted in yellow in Figure 4(a). We focused on LPC and DL tissues since diagnostically it is most beneficial for LPCs to be genetically distinguished from DL. As expected from the PCA, HCC samples are interspersed within the DL arrays (Figure 4(b)). Interestingly LPCs share some common expression patterns with HCC samples, perhaps indicative of their tumorigenic potential. To further investigate this, probe sets with low expression in DL but at least two-fold greater expression in HCC and LPC categories were identified. These were then compared to genes displaying a similar expression pattern within hepatocytes, fibrolamellar HCCs, and hepatoblasts by Oikawa et al. [27] (Supplementary Table 7). We designated the distinct clustered groupings in Figure 4(b): α , β , and γ . Group α probe sets have low expression in LPCs, whilst those in groups β and γ are highly expressed.

Lastly, to generate a list of LPC-expressed genes, the 355 probe sets of Figure 4(a) were refined by retaining only those that were present in at least 9 of the 10 averaged LPC microarrays. This filtering yielded 40 probe sets annotated to 34 unique genes, shown in Table 1 including 8 genes that have been previously associated with LPCs.

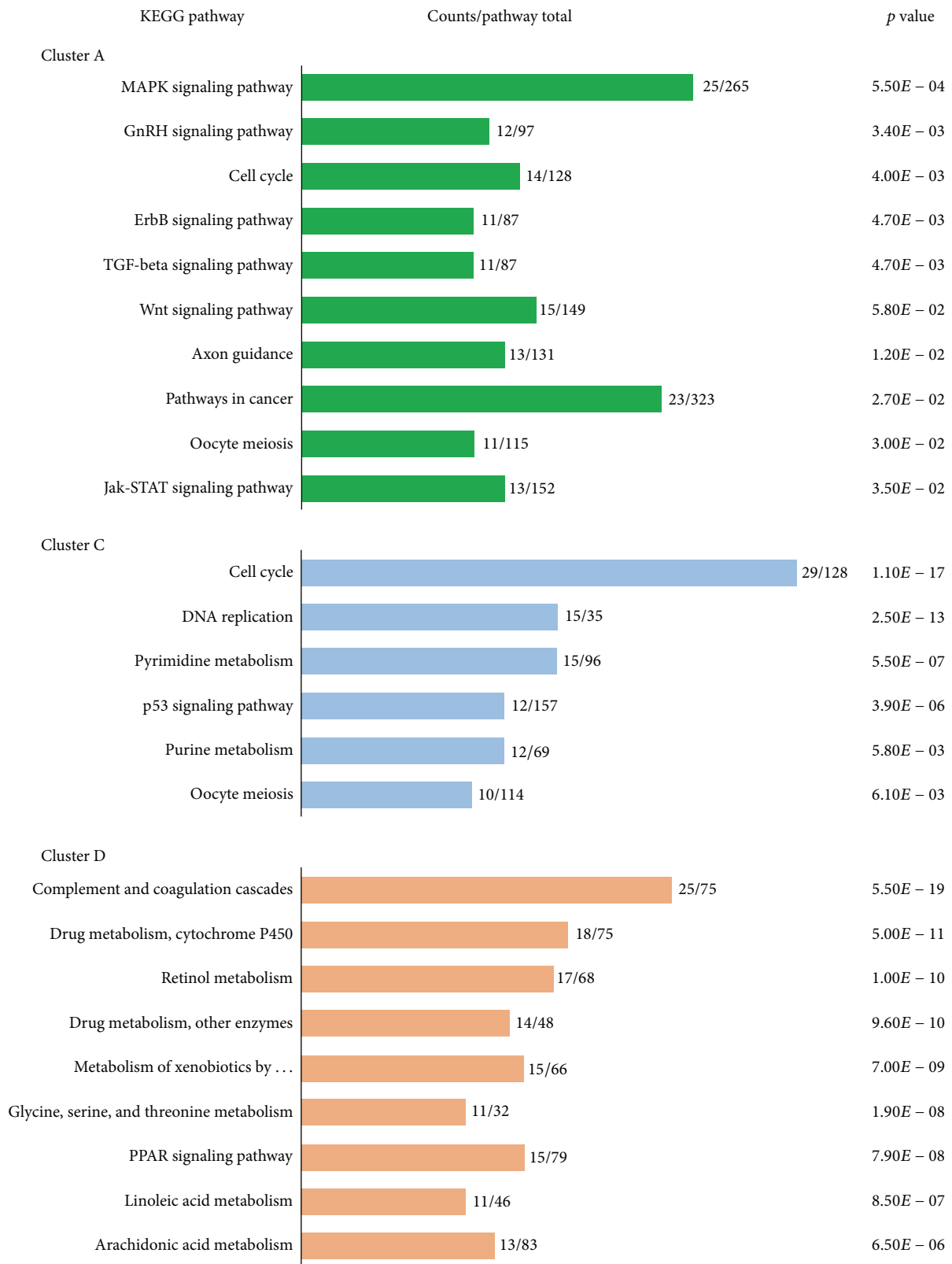


FIGURE 3: Probe set clusters A, C, and D contain pathways that are overrepresented in LPCs and liver tissue. Clusters A and C contain probe sets that are upregulated in liver progenitor cells (LPCs) and cluster D contains those that are upregulated in developed liver whilst being low in LPCs. The Database for Annotation, Visualization and Integrated Discovery online tool was used to identify overrepresented pathways ($p < 0.05$) within each of these clusters. Pathways are displayed together with the number of genes from the cluster list (counts), the total that belongs to each pathway (pathway total) and corresponding p values.

TABLE 1: Markers for liver progenitor cells (LPCs): expression profiles that correlate closely with known LPC markers were mined for potential marker genes. Gene symbol, title, and transcript abundance fold change in LPCs relative to developed liver (DL). References to studies that identified genes as LPC markers are included where applicable.

Gene symbol	Gene title	Fold change (LPC/DL)	References
Abcc1	ATP-binding cassette, subfamily C (CFTR/MRP) 1	152	[73]
Aldh18a1	Aldehyde dehydrogenase 18 family, member A1	66.1	
Capg	Capping protein (actin filament), gelsolin-like	82.6	
Cdc20	Cell division cycle 20 homolog (<i>S. cerevisiae</i>)	16.5	
Enah	Enabled homolog (<i>Drosophila</i>)	Not expressed in DL	[68, 76]
Fhl2	Four and a half LIM domains 2	Not expressed in DL	[68]
Gdpd1	Glycerophosphodiester phosphodiesterase domain 1	140	
Gnb5	Guanine nucleotide binding protein (G protein), beta 5	92.3	
Gpr177	G protein-coupled receptor 177	107	
Hs2st1	Heparan sulfate 2-O-sulfotransferase 1	Not expressed in DL	
Ift57	Intraflagellar transport 57 homolog (<i>Chlamydomonas</i>)	29.8	
Klf5	Kruppel-like factor 5	506	[68]
Ltbp3	Latent transforming growth factor beta binding protein 3	Not expressed in DL	
Luzp1	Leucine zipper protein 1	77.7	
Map4k4	Mitogen-activated protein kinase kinase kinase kinase 4	34.2	[77]
Mcm2	Minichromosome maintenance deficient 2	8.53	
Ncam1	Neural cell adhesion molecule 1	Not expressed in DL	[78]
Ndr4	N-myc downstream regulated gene 4	146	
Npnt	Nephronectin	83.4	
Nup93	Nucleoporin 93	10.3	
Nxn	Nucleoredoxin	16.9	[76]
Panx1	Pannexin 1	Not expressed in DL	
Pfkfb	Phosphofructokinase, platelet	Not expressed in DL	
Phc1	Polyhomeotic-like 1 (<i>Drosophila</i>)	Not expressed in DL	
Pkia	Protein kinase inhibitor, alpha	Not expressed in DL	
Pkm2	Pyruvate kinase, muscle	115	[79]
Plod2	Procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	Not expressed in DL	
Pmm1	Phosphomannomutase 1	68.2	
Prkg2	Protein kinase, cGMP-dependent, type II	Not expressed in DL	
Sfxn3	Sideroflexin 3	92.5	
Tnnt2	Troponin T2, cardiac	Not expressed in DL	
Tubb2b	Tubulin, beta 2b	818	
Tulp3	Tubby-like protein 3	Not expressed in DL	
Wisp1	WNT1 inducible signaling pathway protein 1	69.0	

3.4. *Identified Markers Detect LPCs in Injured Liver Tissue and Their Expression Levels Correlate with LPC Numbers.* For detection of LPCs within a liver sample, genes from group α should display low expression, whilst genes from groups β and γ should exhibit high expression, relative to a reference gene. An ideal reference gene should have little to no variability across a variety of tissue types, and thus it should have a low coefficient of variation. *Taf4a* on the Affymetrix Mouse Genome 430 2.0 chip fits this criterion.

To determine whether a combination of selected genes from groups α , β , and γ could accurately detect the presence and abundance of LPCs in liver tissue of varying pathologies we obtained the livers of mice subjected to the CDE diet, the immune-mediated Met-Kb and 178.3 transgenic mouse models, and DDC induced hepatotoxicity. The histology

of these livers showed varying degrees of disease severity (Figure 5). CDE liver samples displayed steatosis (Figure 5(b), arrowheads) and numerous small basophilic cells with oval-shaped nuclei (Figure 5(b), arrows) indicative of LPCs. The DDC liver featured porphyrin accumulations (Figure 5(d), arrowheads) and ductular reactions (Figure 5(d), arrows), whilst the 178.3 liver (Figure 5(f)) displayed normal architecture and Met-Kb liver showed increased ductal structures surrounded by many small basophilic cells (Figure 5(h), arrows). The degree of LPC induction was determined by staining and quantification of panCK positivity (Figures 5(a), 5(c), 5(e), and 5(g)), found to be 3.3%, 2.5%, 0.09%, and 0.11% for CDE, DDC, 178.3, and Met-Kb models, respectively.

From our LPC gene signature *Ltbp3* and *Ncam1* from group β and *Mcm2* and *M2pk* from group γ displayed the

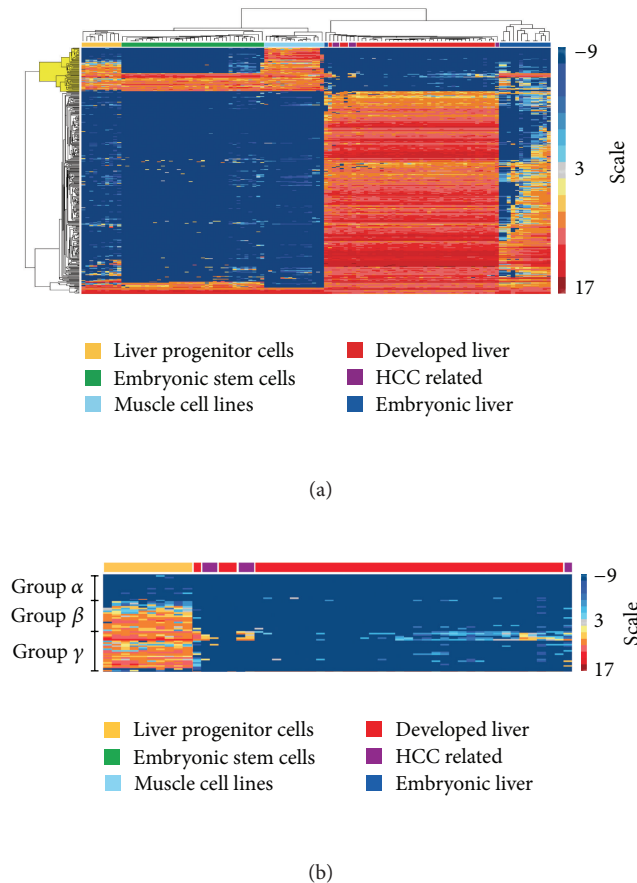


FIGURE 4: Distinct expression profiles for liver progenitor cells identify three gene expression groups. (a) Hierarchical clustering of expression profiles that correlate with known liver progenitor cell markers. (b) Close view of the yellow highlighted rows from (a), with particular focus on the LPC and developed liver groupings. Three distinct groups of gene expression patterns (α , β , and γ) are defined.

highest abundance in livers of mice on diets yielding the greatest proportion of panCK positivity (Figure 6). Protein expression of NCAM1 isoform 140 and M2PK was increased in 75% CDE diet livers relative to controls, as determined by Western blot (Supplementary Figure 1). No clear pattern was observed within NCAM1 isoform 180. The group α gene *Murc* that is expressed at low levels in LPCs and DL showed decreased expression in DDC, Met-Kb, and 178.3 liver but surprisingly was increased in CDE liver (Figure 6).

To confirm that transcript abundance of genes in groups β and γ correlates with LPC number, we quantified the LPC response in day 0 ($n = 7$) and day 21 ($n = 11$) mouse CDE livers (Figure 7). LPC quantitation for all samples ranged between 0.0036% (Figure 7(a)) and 3.7% (Figure 7(b)). Predictably, mRNA levels for an established LPC marker *Cd24a* strongly correlated ($r^2 = 0.8810$) with panCK positivity (Figure 7(e)). Importantly, expression of *Ncam1* (group β) and *Mcm2* (group γ) also correlated well ($r^2 = 0.6952$ and 0.8128 , resp.) with panCK staining (Figures 7(c) and 7(d)). Collectively these results validate our approach to identify novel LPC markers that can be used for their detection *in vivo* and

their utility to assess liver pathology with respect to the extent of LPC induction.

4. Discussion

In this study we demonstrate that it is possible to collect, analyze, and compare publicly available transcriptome data to derive useful information relating to a particular cell type of interest, in this instance, LPCs. This is possible so long as appropriate criteria for their inclusion are adopted, and it is important to use related microarray platforms with a minimum of three highly correlated replicates and consistent data normalization. Adherence to these principles benefited this study as it integrated more than 400 microarrays to highlight a unique transcriptomic signature of LPCs.

PCA successfully separated LPCs from other stem cell types and more importantly from other liver cell and tissue samples. Notably, LPC lines from different research groups, including array data from a BMEL line derived in Paris and analyzed in Houston [46], T-LPC isolated in Tokyo [22] and analyzed in Perth, and other lines isolated and analyzed in Perth, were more similar to each other than to other tissue or cell types. Spatially, the LPC cluster is distant from DL, not surprising considering their different functions and degree of differentiation. Given the association between LPCs and HCC, closer clustering of these two groups could be expected. However, the LPC and HCC clusters are distant, indicating these HCCs may not be derived from LPCs. Indeed one of the five HCC analyses used in this study mentions that the tumors are mostly well-differentiated [47]. In this case we would expect the HCC group to cluster closer to the DL group than the LPC one, as reported in this study. Unfortunately, there is limited or no information available regarding the differentiated status of the remaining four HCCs used [48, 49]. Histologically, poorly differentiated HCCs often carry LPC and/or stem markers [50, 51] and may be more indicative of LPC-derived HCC. Use of these types of tumors may have offered further insights into the relationship between LPCs and LPC-derived HCC. Moreover LPCs cluster closely with embryonic liver due to their shared plasticity and origin. In fact the BMEL LPC line used was derived from embryonic liver. The proximity of LPCs to muscle is unexpected, though it may be a result of their shared immortalized progenitor cell status. LPCs and C2C12 myoblasts used in this study represent immature cells of their respective lineages and are both involved with tissue regeneration. Comparison of primary LPCs and muscle cells may have prevented this association. Despite this, our use of immortalized cell lines yielded a transcriptomic signature consistent with LPC biology and genetics.

Notably, overrepresented pathways in cluster D, which were poorly expressed in LPCs but highly expressed in developed liver, were consistent with liver functions including drug, vitamin, and fatty acid metabolism. Furthermore upregulated LPC pathways in clusters A and C are consistent with transcriptomic analyses of rat oval cells [52]. These LPC pathways included the Wnt, ErbB, p53, MAPK, and TGF-beta signaling pathways. Wnt/ β -catenin is a mediator of LPC activation and expansion [53]. Wnt induced β -catenin

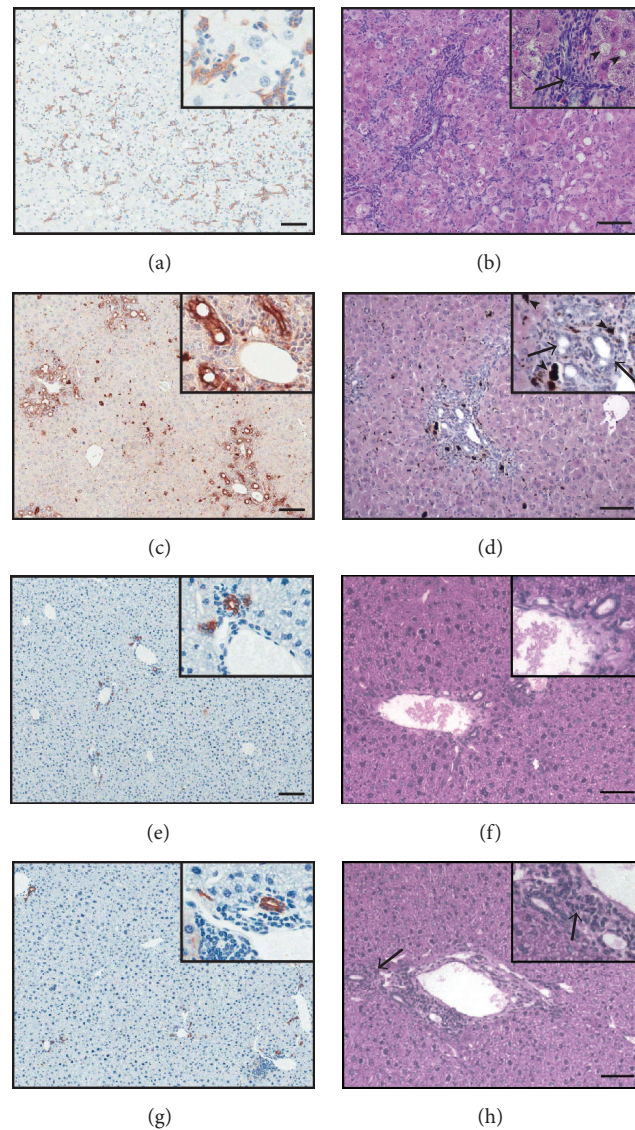


FIGURE 5: Four liver injury models display different degrees of pathology and liver progenitor cell (LPC) response. Liver sections from wild-type mice subjected to a choline-deficient, ethionine-supplemented (CDE; (a) and (b)) or 3,5-diethoxy-carbonyl-1,4-dihydrocollidine (DDC; (c) and (d)) diet, or transgenic 178.3 ((e) and (f)) or Met-Kb ((g) and (h)) mice were stained with panCK ((a), (c), (e), and (g)) or H&E ((b), (d), (f), and (h)). Arrows in panels (b) and (d) and (h) indicate LPCs, ductular reactions, and small basophilic cells, respectively. Arrowheads in panels (b) and (d) indicate steatosis and porphyrin accumulations, respectively. Scale bars represent 100 μ m.

phosphorylation targets the gene promoters of EGF, cyclin D (*Ccnd1*), fos-like antigen 1 (*Fosl1*), and c-Myc, to drive cellular proliferation [54, 55]. Consistently, in cluster A *Ccnd1* and *Fosl1* were upregulated in LPCs. Furthermore, activation of cancer-related processes through *Myc* and *Ccnd1* and matrix metalloproteinases have been identified [56] and increased β -catenin expression has been reported in HCC [57] and aggressive hepatoblastomas [58].

The ErbB signaling pathway affects diverse cellular functions including proliferation, angiogenesis, migration, and differentiation. An *ErbB2* overexpressing transgenic mouse exhibits increased numbers of LPC-like cells in the liver [59], and increased expression of ErbB family members is reported in a high proportion of human HCCs and correlates

with cancer progression [60]. These findings support the integral role of the ErbB family in liver carcinogenesis. It is widely accepted that aberrant p53 signaling participates in the development of HCC [61]. Moreover the MAPK signaling pathway is associated with tumorigenesis via regulation of proliferation, differentiation, survival, and migration [62]. Finally, increased TGF-beta signaling is critical in inhibition of hepatocyte-mediated regeneration [63], promoting LPC-mediated regeneration instead [64]. These results highlight LPCs as a regenerative cell source for the liver but also as a putative cancer stem cell that may give rise to HCC [53, 65].

The individual LPC genes identified in this study (Table 1) also reflect LPC function. *Mcm2* is expressed in a variety of regenerative cells including neural stem cells and hepatocytes

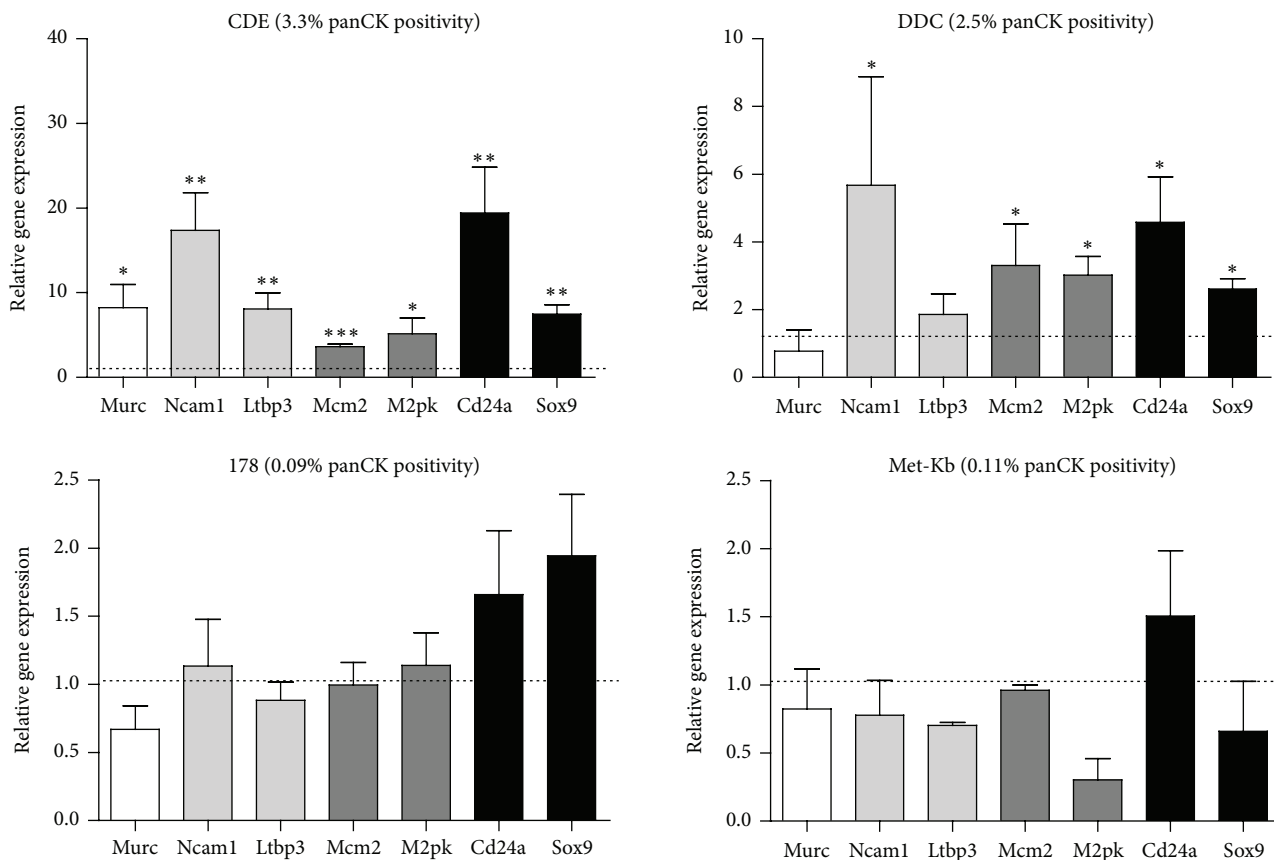


FIGURE 6: A liver progenitor cell (LPC) transcriptomic signature can identify liver injury models with LPC induction. qPCR-generated mRNA expression levels of known (*Cd24a* and *Sox9*) and identified (group α , *Murc*; β , *Ncam1* and *Ltpb3*; and γ , *Mcm2* and *M2pk*) LPC markers in four different liver injury models. Data are normalized to gene expression in the appropriate control livers (indicated by the dotted line) and are relative to the *Taf4a* housekeeping gene. Data represents mean + SEM of 3 separate qPCR assays with significance determined by Student's *t*-test (* $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$).

[66, 67]. *Fhl2*, *Klf5*, and *Enah* are highly expressed in rat embryonic hepatoblasts relative to adult hepatocytes [68] and *Fhl2* is a marker of progenitor cells of myeloid [69], myocardial [70], and mesenchymal [71] origin. *Ndr4* is a gene involved in cell cycle progression and survival [72]. *Abcc1* is highly expressed in stem-like cells within HCC cell lines and confers chemotherapeutic resistance, a property of HCC [73].

In this study we were able to use genes from our LPC signature to successfully reflect the presence or absence of LPCs in a number of liver injury models. The unexpected expression of the group α gene, *Murc*, in CDE liver may be explained by the presence of activated hepatic stellates in this model [74], since they are known to express *Murc* [75]. Consequently, we found expression of individual genes from either group β or group γ was more reliable for detection of LPCs in tissue samples than a signature encompassing all three groups.

Our approach has identified many novel LPC markers, useful for detection of these cells *in vivo*. Markers that are expressed on the cell surface have the potential for isolating LPCs. According to their gene ontology, the novel LPC markers GNB5, PANX1, PRKG2, and TULP3 from Table 1 localize to the plasma membrane; however this would need to

be confirmed using fluorescence microscopy and subcellular fractionation.

5. Conclusion

Using publicly available gene expression/microarray data, we compare the LPC transcriptome against other cell and tissue types and identify signaling pathways that support their ability to rapidly proliferate and differentiate. We find pathways which are consistent with their putative role as a cancer stem cell in HCC. Importantly we also identify many novel LPC markers, several of which localize to the cell membrane. This is beneficial for the field as there is a scarcity of specific markers that can both identify and purify LPCs. Finally, since LPCs may reflect a precancerous liver condition as well as the severity for a range of liver pathologies, our LPC signature could ideally be used as a novel early indicator of liver disease in patients.

Abbreviations

BMEL: Bipotential murine embryonic liver
 BMOL: Bipotential murine oval liver

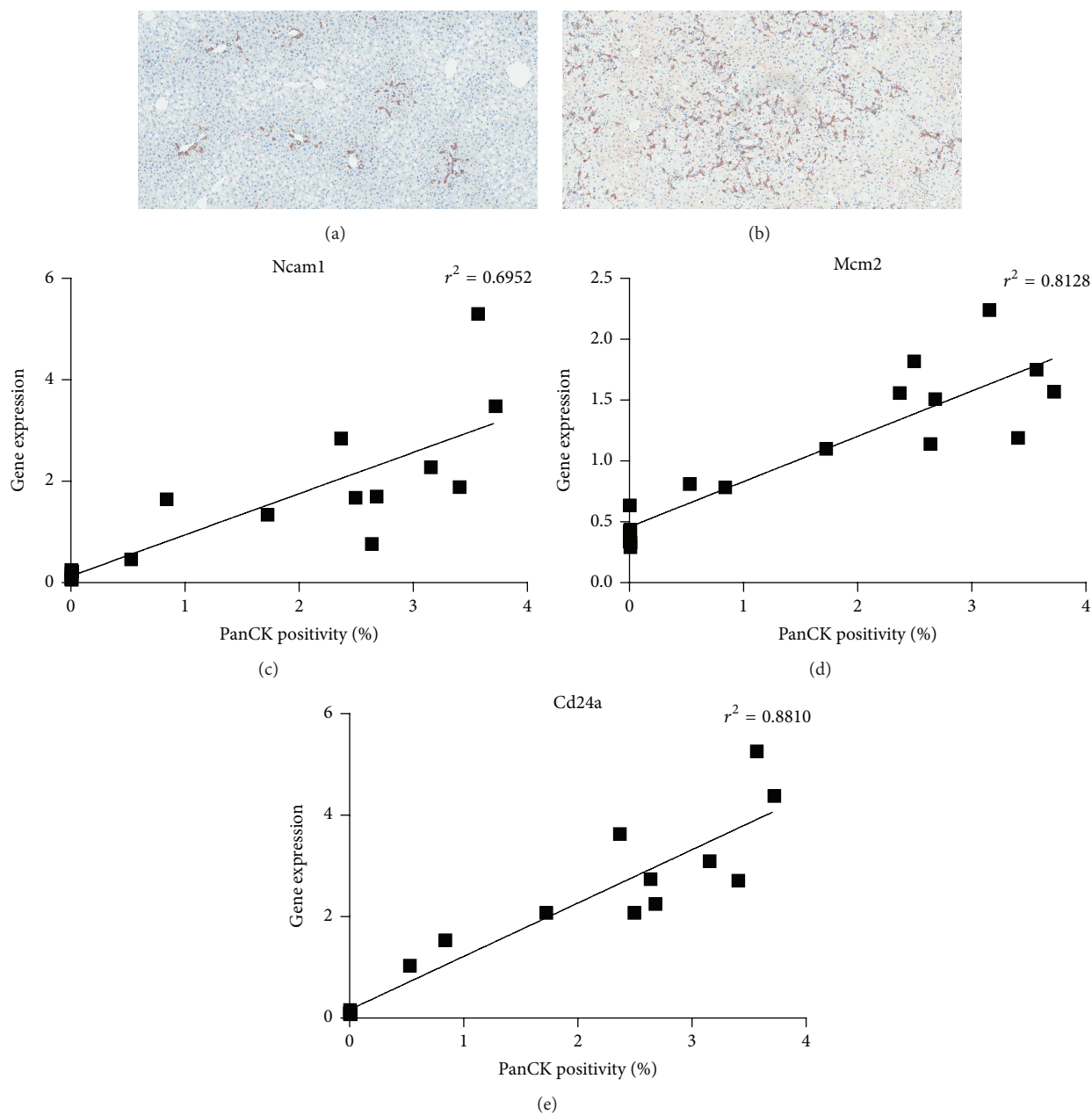


FIGURE 7: Gene expression of group β and γ genes, *Ncam1* and *Mcm2*, and the known marker *Cd24a* correlate with panCK staining *in vivo*. Representative images of (a) control and (b) choline-deficient, ethionine-supplemented (CDE) mouse liver sections stained for panCK. Expressions of (c) *Ncam1*, (d) *Mcm2*, and (e) *Cd24a* relative to *Taf4a* in a variety of CDE and control liver tissues are plotted against their corresponding panCK positivity values.

CDE: Choline-deficient, ethionine-supplemented
 DAVID: Database for Annotation, Visualization and Integrated Discovery
 DDC: 3,5-Diethoxy-carbonyl-1,4-dihydrocollidine
 DL: Developed liver
 GEO: Gene Expression Omnibus
 HCC: Hepatocellular carcinoma

KEGG: Kyoto Encyclopedia of Genes and Genomes
 LPC: Liver progenitor cell
 MAD: Median Absolute Deviation
 MAS 5.0: MicroArray Suite 5.0
 PAINT: Promoter Analysis and Interaction Network Toolset
 PCA: Principal component analysis
 PIL: p53-immortalized liver
 T-LPC: Tokyo-LPC.

Disclosure

Adam M. Passman and Jasmine Low are joint first authors.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Australian Research Council Centre of Excellence Grant (CE056149), the Cancer Council of Western Australia (IDs 634424, APP1010749, and APP1067515), the National Health and Medical Research Council Project Grant (APP1042368), and the Association of International Cancer Research (no. 09-0084).

References

- [1] X. Wang, M. Foster, M. Al-Dhalimy, E. Lagasse, M. Finegold, and M. Grompe, "The origin and liver repopulating capacity of murine oval cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, supplement 1, pp. 11881–11888, 2003.
- [2] O. Yasui, N. Miura, K. Terada, Y. Kawarada, K. Koyama, and T. Sugiyama, "Isolation of oval cells from Long-Evans Cinnamon rats and their transformation into hepatocytes in vivo in the rat liver," *Hepatology*, vol. 25, no. 2, pp. 329–334, 1997.
- [3] L. Libbrecht and T. Roskams, "Hepatic progenitor cells in human liver diseases," *Seminars in Cell and Developmental Biology*, vol. 13, no. 6, pp. 389–396, 2002.
- [4] K. N. Lowes, B. A. Brennan, G. C. Yeoh, and J. K. Olynyk, "Oval cell numbers in human chronic liver diseases are directly related to disease severity," *The American Journal of Pathology*, vol. 154, no. 2, pp. 537–541, 1999.
- [5] L. Mishra, T. Banker, J. Murray et al., "Liver stem cells and hepatocellular carcinoma," *Hepatology*, vol. 49, no. 1, pp. 318–329, 2009.
- [6] S. Shin, N. Upadhyay, L. E. Greenbaum, and K. H. Kaestner, "Ablation of Foxl1-Cre-labeled hepatic progenitor cells and their descendants impairs recovery of mice from liver injury," *Gastroenterology*, vol. 148, no. 1, pp. 192–202.e3, 2015.
- [7] S. C. Ward, S. N. Thung, K. H. Lim et al., "Hepatic progenitor cells in liver cancers from Asian children," *Liver International*, vol. 30, no. 1, pp. 102–111, 2010.
- [8] T. Yamashita, J. Ji, A. Budhu et al., "EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features," *Gastroenterology*, vol. 136, no. 3, pp. 1012.e4–1024.e4, 2009.
- [9] J.-S. Lee, J. Heo, L. Libbrecht et al., "A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells," *Nature Medicine*, vol. 12, no. 4, pp. 410–416, 2006.
- [10] Y.-W. Zheng and H. Taniguchi, "Diversity of hepatic stem cells in the fetal and adult liver," *Seminars in Liver Disease*, vol. 23, no. 4, pp. 337–348, 2003.
- [11] H. Shinozuka, B. Lombardi, S. Sell, and R. M. Iammarino, "Enhancement of DL-ethionine-induced liver carcinogenesis in rats fed a choline-devoid diet," *Journal of the National Cancer Institute*, vol. 61, no. 3, pp. 813–817, 1978.
- [12] A. M. Passman, R. P. Strauss, S. B. Mcspadden et al., "A modified choline-deficient, ethionine-supplemented diet reduces morbidity and retains a liver progenitor cell response in mice," *Disease Models and Mechanisms*, vol. 8, no. 12, pp. 1635–1641, 2015.
- [13] L. Boulter, O. Govaere, T. G. Bird et al., "Macrophage-derived Wnt opposes Notch signaling to specify hepatic progenitor cell fate in chronic liver disease," *Nature Medicine*, vol. 18, no. 4, pp. 572–579, 2012.
- [14] M. Tatematsu, R. H. Ho, T. Kaku, J. K. Ekem, and E. Farber, "Studies on the proliferation and fate of oval cells in the liver of rats treated with 2-acetylaminofluorene and partial hepatectomy," *The American Journal of Pathology*, vol. 114, no. 3, pp. 418–430, 1984.
- [15] R. P. Witek, S. H. Fisher, and B. E. Petersen, "Monocrotaline, an alternative to retrorsine-based hepatocyte transplantation in rodents," *Cell Transplantation*, vol. 14, no. 1, pp. 41–47, 2005.
- [16] E. Laconi, R. Oren, D. K. Mukhopadhyay et al., "Long-term, near-total liver replacement by transplantation of isolated hepatocytes in rats treated with retrorsine," *The American Journal of Pathology*, vol. 153, no. 1, pp. 319–329, 1998.
- [17] R. Español-Suñer, R. Carpentier, N. Van Hul et al., "Liver progenitor cells yield functional hepatocytes in response to chronic liver injury in mice," *Gastroenterology*, vol. 143, no. 6, pp. 1564–1575.e7, 2012.
- [18] D. Rodrigo-Torres, S. Affo, M. Coll et al., "The biliary epithelium gives rise to liver progenitor cells," *Hepatology*, vol. 60, no. 4, pp. 1367–1377, 2014.
- [19] M. Grompe, "Liver stem cells, where art thou?" *Cell Stem Cell*, vol. 15, no. 3, pp. 257–258, 2014.
- [20] J. R. Schaub, Y. Malato, C. Gormond, and H. Willenbring, "Evidence against a stem cell origin of new hepatocytes in a common mouse model of chronic liver injury," *Cell Reports*, vol. 8, no. 4, pp. 933–939, 2014.
- [21] K. Yanger, D. Knigin, Y. Zong et al., "Adult hepatocytes are generated by self-duplication rather than stem cell differentiation," *Cell Stem Cell*, vol. 15, no. 3, pp. 340–349, 2014.
- [22] M. Okabe, Y. Tsukahara, M. Tanaka et al., "Potential hepatic stem cells reside in EpCAM+ cells of normal and injured mouse liver," *Development*, vol. 136, no. 11, pp. 1951–1960, 2009.
- [23] H. Strick-Marchand and M. C. Weiss, "Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos," *Hepatology*, vol. 36, no. 4, pp. 794–804, 2002.
- [24] N. Tanimizu, H. Saito, K. Mostov, and A. Miyajima, "Long-term culture of hepatic progenitors derived from mouse Dlk⁺ hepatoblasts," *Journal of Cell Science*, vol. 117, no. 26, pp. 6425–6434, 2004.
- [25] J. E. E. Tirnitz-Parker, J. N. Tonkin, B. Knight, J. K. Olynyk, and G. C. T. Yeoh, "Isolation, culture and immortalisation of hepatic oval cells from adult mice fed a choline-deficient, ethionine-supplemented diet," *International Journal of Biochemistry and Cell Biology*, vol. 39, no. 12, pp. 2226–2239, 2007.
- [26] C. M. Shachaf, A. M. Kopelman, C. Arvanitis et al., "MYC inactivation uncovers pluripotent differentiation and tumour dormancy in hepatocellular cancer," *Nature*, vol. 431, no. 7012, pp. 1112–1117, 2004.
- [27] T. Oikawa, E. Wauthier, T. A. Dinh et al., "Model of fibrolamellar hepatocellular carcinomas reveals striking enrichment in cancer stem cells," *Nature Communications*, vol. 6, article 8070, 2015.

- [28] L. Zender, M. S. Spector, W. Xue et al., "Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach," *Cell*, vol. 125, no. 7, pp. 1253–1267, 2006.
- [29] L. Lu, Y. Li, S. M. Kim et al., "Hippo signaling is a potent in vivo growth and tumor suppressor pathway in the mammalian liver," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 4, pp. 1437–1442, 2010.
- [30] K.-P. Lee, J.-H. Lee, T.-S. Kim et al., "The Hippo-Salvador pathway restrains hepatic oval cell proliferation, liver size, and liver tumorigenesis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 18, pp. 8248–8253, 2010.
- [31] F. D. Camargo, S. Gokhale, J. B. Johnnidis et al., "YAP1 increases organ size and expands undifferentiated progenitor cells," *Current Biology*, vol. 17, no. 23, pp. 2054–2060, 2007.
- [32] D. Yimlamai, C. Christodoulou, G. G. Galli et al., "Hippo pathway activity influences liver cell fate," *Cell*, vol. 157, no. 6, pp. 1324–1338, 2014.
- [33] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," *Cell*, vol. 100, no. 1, pp. 57–70, 2000.
- [34] H. Strick-Marchand, S. Morosan, P. Charneau, D. Kremsdorf, and M. C. Weiss, "Bipotential mouse embryonic liver stem cell lines contribute to liver regeneration and differentiate as bile ducts and hepatocytes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 22, pp. 8360–8365, 2004.
- [35] M. L. Dumble, E. J. Croager, G. C. T. Yeoh, and E. A. Quail, "Generation and characterization of p53 null transformed hepatic progenitor cells: oval cells give rise to hepatocellular carcinoma," *Carcinogenesis*, vol. 23, no. 3, pp. 435–445, 2002.
- [36] G. A. Churchill, "Using ANOVA to analyze microarray data," *BioTechniques*, vol. 37, no. 2, pp. 173–177, 2004.
- [37] L. E. Peterson, "Partitioning large-sample microarray-based gene expression profiles using principal components analysis," *Computer Methods and Programs in Biomedicine*, vol. 70, no. 2, pp. 107–119, 2003.
- [38] J. Vandesompele, K. De Preter, F. Pattyn et al., "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes," *Genome Biology*, vol. 3, no. 7, 2002.
- [39] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
- [40] G. Dennis Jr., B. T. Sherman, D. A. Hosack et al., "DAVID: database for annotation, visualization, and integrated discovery," *Genome Biology*, vol. 4, article P3, 2003.
- [41] R. Vadigepalli, P. Chakravarthula, D. E. Zak, J. S. Schwaber, and G. E. Gonye, "PAINT: a promoter analysis and interaction network generation tool for gene regulatory network identification," *Omics A Journal of Integrative Biology*, vol. 7, no. 3, pp. 235–252, 2003.
- [42] B. Knight, V. B. Matthews, J. K. Olynyk, and G. C. Yeoh, "Jekyll and Hyde: evolving perspectives on the function and potential of the adult liver progenitor (oval) cell," *BioEssays*, vol. 27, no. 11, pp. 1192–1202, 2005.
- [43] H. A. Crosby, D. A. Kelly, and A. J. Strain, "Human hepatic stem-like cells isolated using c-kit or CD34 can differentiate into biliary epithelium," *Gastroenterology*, vol. 120, no. 2, pp. 534–544, 2001.
- [44] G. Nowak, B.-G. Ericzon, S. Nava, M. Jaksch, M. Westgren, and S. Sumitran-Holgersson, "Identification of expandable human hepatic progenitors which differentiate into mature hepatic cells in vivo," *Gut*, vol. 54, no. 7, pp. 972–979, 2005.
- [45] E. Schmelzer, L. Zhang, A. Bruce et al., "Human hepatic stem cells from fetal and postnatal donors," *Journal of Experimental Medicine*, vol. 204, no. 8, pp. 1973–1987, 2007.
- [46] S. A. Ochsner, H. Strick-Marchand, Q. Qiu et al., "Transcriptional profiling of bipotential embryonic liver cells to identify liver progenitor cell surface markers," *STEM CELLS*, vol. 25, no. 10, pp. 2476–2487, 2007.
- [47] M. Katzenellenbogen, L. Mizrahi, O. Pappo et al., "Molecular mechanisms of liver carcinogenesis in the Mdr2-knockout mice," *Molecular Cancer Research*, vol. 5, no. 11, pp. 1159–1170, 2007.
- [48] K. Khetchoumian, M. Teletin, J. Tisserand et al., "Loss of Trim24 (Tif1 α) gene function confers oncogenic activity to retinoic acid receptor alpha," *Nature Genetics*, vol. 39, no. 12, pp. 1500–1506, 2007.
- [49] S. S. Sheth, J. S. Bodnar, A. Ghazalpour et al., "Hepatocellular carcinoma in Txnip-deficient mice," *Oncogene*, vol. 25, no. 25, pp. 3528–3536, 2006.
- [50] H. Kim, G. H. Choi, D. C. Na et al., "Human hepatocellular carcinomas with 'Stemness'-related marker expression: keratin 19 expression and a poor prognosis," *Hepatology*, vol. 54, no. 5, pp. 1707–1717, 2011.
- [51] A. Villanueva, Y. Hoshida, C. Battiston et al., "Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma," *Gastroenterology*, vol. 140, no. 5, pp. 1501–1512.e2, 2011.
- [52] V. Cimica, D. Batusic, Y. Chen, T. Hollemann, T. Pieler, and G. Ramadori, "Transcriptome analysis of rat liver regeneration in a model of oval hepatic stem cells," *Genomics*, vol. 86, no. 3, pp. 352–364, 2005.
- [53] U. Apte, M. D. Thompson, S. Cui, B. Liu, B. Cieply, and S. P. S. Monga, "Wnt/ β -catenin signaling mediates oval cell response in rodents," *Hepatology*, vol. 47, no. 1, pp. 288–295, 2008.
- [54] X. Tan, U. Apte, A. Micsenyi et al., "Epidermal growth factor receptor: a novel target of the Wnt/ β -catenin pathway in liver," *Gastroenterology*, vol. 129, no. 1, pp. 285–302, 2005.
- [55] W. Yang, H.-X. Yan, L. Chen et al., "Wnt/ β -catenin signaling contributes to activation of normal and tumorigenic liver progenitor cells," *Cancer Research*, vol. 68, no. 11, pp. 4287–4295, 2008.
- [56] J. Borlak, T. Meier, R. Halter, R. Spanel, and K. Spanel-Borowski, "Epidermal growth factor-induced hepatocellular carcinoma: gene expression profiles in precursor lesions, early stage and solitary tumours," *Oncogene*, vol. 24, no. 11, pp. 1809–1819, 2005.
- [57] Y. Ishizaki, S. Ikeda, M. Fujimori et al., "Immunohistochemical analysis and mutational analyses of beta-catenin, Axin family and APC genes in hepatocellular carcinomas," *International Journal of Oncology*, vol. 24, no. 5, pp. 1077–1083, 2004.
- [58] S. Cairo, C. Armengol, A. De Reyniès et al., "Hepatic stem-like phenotype and interplay of Wnt/ β -catenin and myc signaling in aggressive childhood liver cancer," *Cancer Cell*, vol. 14, no. 6, pp. 471–484, 2008.
- [59] I. Freitas, S. Fracchiolla, G. Baronzio et al., "Stem cell recruitment and liver de-differentiation in MMTV-neu (ErbB-2) transgenic mice," *Anticancer Research*, vol. 23, no. 5, pp. 3783–3794, 2003.

- [60] Y. Ito, T. Takeda, M. Sakon et al., "Expression and clinical significance of erb-B receptor family in hepatocellular carcinoma," *British Journal of Cancer*, vol. 84, no. 10, pp. 1377–1383, 2001.
- [61] P. A. Farazi and R. A. DePinho, "Hepatocellular carcinoma pathogenesis: from genes to environment," *Nature Reviews Cancer*, vol. 6, no. 9, pp. 674–687, 2006.
- [62] E. F. Wagner and Á. R. Nebreda, "Signal integration by JNK and p38 MAPK pathways in cancer development," *Nature Reviews Cancer*, vol. 9, no. 8, pp. 537–549, 2009.
- [63] W. E. Russell, R. J. Coffey Jr., A. J. Ouellette, and H. L. Moses, "Type beta transforming growth factor reversibly inhibits the early proliferative response to partial hepatectomy in the rat," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, no. 14, pp. 5126–5130, 1988.
- [64] L.-I. Kang, W. M. Mars, and G. K. Michalopoulos, "Signals and cells involved in regulating liver regeneration," *Cells*, vol. 1, no. 4, pp. 1261–1292, 2012.
- [65] A. Gregorieff and H. Clevers, "Wnt signaling in the intestinal epithelium: from endoderm to cancer," *Genes & Development*, vol. 19, no. 8, pp. 877–890, 2005.
- [66] A. Mohan, M. Kandalam, H. L. Ramkumar, L. Gopal, and S. Krishnakumar, "Stem cell markers: ABCG2 and MCM2 expression in retinoblastoma," *British Journal of Ophthalmology*, vol. 90, no. 7, pp. 889–893, 2006.
- [67] A. Freeman, S. Hamid, L. Morris et al., "Improved detection of hepatocyte proliferation using antibody to the pre-replication complex: an association with hepatic fibrosis and viral replication in chronic hepatitis C virus infection," *Journal of Viral Hepatitis*, vol. 10, no. 5, pp. 345–350, 2003.
- [68] P. M. Petkov, J. Zavadil, D. Goetz et al., "Gene expression pattern in hepatic stem/progenitor cells during rat fetal development using complementary DNA microarrays," *Hepatology*, vol. 39, no. 3, pp. 617–627, 2004.
- [69] Z. Qian, L. Mao, A. A. Fernald et al., "Enhanced expression of FHL2 leads to abnormal myelopoiesis *in vivo*," *Leukemia*, vol. 23, no. 9, pp. 1650–1657, 2009.
- [70] Y. Kong, J. M. Shelton, B. Rothermel et al., "Cardiac-specific LIM protein FHL2 modifies the hypertrophic response to β -adrenergic stimulation," *Circulation*, vol. 103, no. 22, pp. 2731–2738, 2001.
- [71] Z. Hamidouche, E. Haÿ, P. Vaudin et al., "FHL2 mediates dexamethasone-induced mesenchymal cell differentiation into osteoblasts by activating Wnt/ β -catenin signaling-dependent Runx2 expression," *The FASEB Journal*, vol. 22, no. 11, pp. 3813–3822, 2008.
- [72] S. H. Schilling, A. B. Hjelmeland, D. R. Radloff et al., "NDRG4 is required for cell cycle progression and survival in glioblastoma cells," *The Journal of Biological Chemistry*, vol. 284, no. 37, pp. 25160–25169, 2009.
- [73] Z. Zhu, X. Hao, M. Yan et al., "Cancer stem/progenitor cells are highly enriched in CD133⁺CD44⁺ population in hepatocellular carcinoma," *International Journal of Cancer*, vol. 126, no. 9, pp. 2067–2078, 2010.
- [74] R. G. Ruddell, B. Knight, J. E. E. Tirnitz-Parker et al., "Lymphotoxin- β receptor signaling regulates hepatic stellate cell function and wound healing in a murine model of chronic liver injury," *Hepatology*, vol. 49, no. 1, pp. 227–239, 2009.
- [75] V. Krizhanovsky, M. Yon, R. A. Dickins et al., "Senescence of activated stellate cells limits liver fibrosis," *Cell*, vol. 134, no. 4, pp. 657–667, 2008.
- [76] H. Wang, Y. Gao, X. Jin, and J. Xiao, "Expression of contactin associated protein-like 2 in a subset of hepatic progenitor cell compartment identified by gene expression profiling in hepatitis B virus-positive cirrhosis," *Liver International*, vol. 30, no. 1, pp. 126–138, 2010.
- [77] M. K. Behnke, *Growth gone awry: exploring the role of embryonic liver development genes in HCV induced cirrhosis and hepatocellular carcinoma [Ph.D. thesis]*, 2012.
- [78] C. Dorrell, L. Erker, J. Schug et al., "Prospective isolation of a bipotential clonogenic liver progenitor cell in adult mice," *Genes & Development*, vol. 25, no. 11, pp. 1193–1203, 2011.
- [79] B. Akhurst, E. J. Croager, C. A. Farley-Roche et al., "A modified choline-deficient, ethionine-supplemented diet protocol effectively induces oval cells in mouse liver," *Hepatology*, vol. 34, no. 3, pp. 519–522, 2001.