Molecular mechanisms regulating the establishment of hepatocyte polarity during human hepatic progenitor cell differentiation into a functional hepatocyte-like phenotype

Mingxi Hua*, Weitao Zhang*, Weihong Li, Xueyang Li, Baoqing Liu, Xin Lu and Haiyan Zhang‡

Department of Cell Biology, Municipal Laboratory for Liver Protection and Regulation of Regeneration, Capital Medical University, Beijing, 100069, China

*These authors contributed equally to this work
‡Author for correspondence (culture@ccmu.edu.cn)

Accepted 6 August 2012
Journal of Cell Science 125, 5800–5810

doi: 10.1242/jcs.110551

Summary

The correct functioning of hepatocytes requires the establishment and maintenance of hepatocyte polarity. However, the mechanisms regulating the generation of hepatocyte polarity are not completely understood. The differentiation of human fetal hepatic progenitor cells (hFHPCs) into functional hepatocytes provides a powerful *in vitro* model system for studying the molecular mechanisms governing hepatocyte development. In this study, we used a two-stage differentiation protocol to generate functional polarized hepatocyte-like cells (HLCs) from hFHPCs. Global gene expression profiling was performed on triplicate samples of hFHPCs, immature-HLCs and mature-HLCs. When the differential gene expression was compared based on the differentiation stage, a number of genes were identified that might be essential for establishing and maintaining hepatocyte polarity. These genes include those that encode actin filament-binding protein, protein tyrosine kinase activity molecules, and components of signaling pathways, such as *PTK7*, *PARD3*, *PRKCI* and *CDC42*. Based on known and predicted protein-protein interactions, the candidate genes were assigned to networks and clustered into functional categories. The expression of several of these genes was confirmed using real-time RT-PCR. By inactivating genes using small interfering RNA, we demonstrated that *PTK7* and *PARD3* promote hepatic polarity formation and affect F-actin organization. These results provide unique insight into the complex process of polarization during hepatocyte differentiation, indicating key genes and signaling molecules governing hepatocyte differentiation.

Key words: Hepatic progenitor cells, Differentiation, Hepatocyte polarity, Gene expression

Introduction

As for all epithelial cells, hepatocytes must be polarized to be functional in the adult liver. That is, the correct function of the liver is ensured by the establishment and maintenance of hepatocyte polarity. The polarization of hepatocytes involves the formation of functionally distinct apical and basolateral plasma membrane domains. The apical poles of front-facing and adjacent hepatocytes form a continuous network of bile canaliculi (BC), which is in contact with the external environment, into which bile is secreted. The basal membrane domain (sinusoidal), which is in contact with the blood, secretes various components into the circulation and is responsible for the uptake of recycled biliary salts. In polarized hepatocytes, the tight junctions (TJs) create the border between apical and lateral poles (Wang and Boyer, 2004; Decaens et al., 2008). Perturbation or loss of polarity is a hallmark of many hepatic diseases, e.g. cholestasis. A complete understanding of the molecular mechanisms involved in hepatocyte polarization is therefore of considerable significance to both liver cell biology and the pathogenesis of liver diseases. However, hepatocyte polarization is still poorly understood.

The establishment of hepatocyte polarity begins during liver embryogenesis. During hepatocyte differentiation, specific routes and mechanisms are defined for the delivery of proteins to the plasma membrane (Wang and Boyer, 2004; Lemaigre and Zaret, 2004; Lemaigre, 2009). Using microarray analysis, several studies have examined the key genes and pathways of potential importance for generating hepatocyte-like cells (HLCs). In these reports, the liver-specific gene expression was identified within the total, heterogeneous population of cells that differentiated from human embryonic stem cells (Chiao et al., 2008; Synnergren et al., 2010; Jozefczuk et al., 2011) and induced pluripotent stem cells (Jozefczuk et al., 2011), HepaRG liver progenitor cells (Parent and Beretta, 2008) or human adipose tissue-derived stromal cells (Bonora-Centelles et al., 2009; Saulnier et al., 2010). Although these systems have limitations inherent to their respective origins, they represent human models of hepatocyte differentiation. To gain the most specific insight possible into the molecular events driving the establishment of hepatocyte polarity, analyses should be performed on purified cells with specific lineage markers. Human fetal hepatic progenitor cells (hFHPCs) were isolated based on alpha-fetoprotein (AFP) promoter expression from fetus that were aborted during the first trimester (Wang et al., 2008). They express AFP, albumin (ALB), cytokeratin 19 (CK19), are able to

Journal of Cell Science

proliferate long-term *in vitro* (Wang et al., 2008) and can differentiate into functional HLCs *in vitro* (Zhang et al., 2012). Thus, hFHPCs are a potential source and a useful model system to study the mechanisms that regulate the process of hepatocyte differentiation.

In this study, we used a two-stage differentiation protocol to generate functional and polarized HLCs from hFHPCs and performed global gene expression profiling on undifferentiated hFHPCs, immature-HLCs, and mature-HLCs. By comparing differential gene expression profiles of different stages of differentiation, a common pool of genes that serve as regulators of hFHPCs differentiation were identified. Subsequently, the genes involved in cell morphogenesis were further investigated for their molecular functions and their roles in the interactive network of genes associated with hepatocyte polarity. We identified *PTK7*, *PARD3*, *PRKCI* and *CDC42*, known to play key roles in the establishment or maintenance of cell polarity, as genes that were regulated during hFHPCs differentiation. CDC42 may be a core factor in regulating hepatocyte polarization. The expression of several of these genes was confirmed using real-time RT-PCR. The function of two genes (*PTK7* and *PARD3*) in the formation of hepatic polarity was further explored by inactivating their expression using small interfering RNA (siRNA) technology.

Results

Differentiation of hFHPCs into functional hepatocyte-like cells

We previously established a protocol to isolate fetal hepatic progenitor cells based on AFP promoter expression from human aborted fetus, and produced hFHPCs using this protocol (Wang et al., 2008). hFHPCs express the hepatic stem cells/progenitor cells marker AFP, ALB, CK19, epithelial cell adhesion molecule (EpCAM) (Schmelzer et al., 2006; Inada et al., 2008), Delta-like protein (Dlk) (Yanai et al., 2010), and sal-like protein 4 (Sall4) (Oikawa et al., 2009) (supplementary material Fig. S1A). In this study, we used a two-stage differentiation protocol to differentiate of hFHPCs into functional HLCs. In the first stage, hFHPCs were induced to become immature hepatic cells by 5 days of HGF treatment. In the second stage, the immature hepatic cells were further matured by the combination of HGF, DEX and OSM treatment for another 5 days. The majority of the differentiated cells at day 5 showed an epithelial morphology with binucleated centers, expressed lower levels of AFP, and expressed higher levels of ALB and CK8 than at day 0. The differentiated cells at day 10 exhibited typical hepatocyte morphology with a polygonal shape, containing distinctly round nuclei with one or two prominent nucleoli and intercellular structures resembling hepatic canaliculi. They did not express AFP, but the levels of ALB and CK8 significantly increased over day 5 (Fig. 1A; supplementary material Fig. S1B).

To further characterize the cells, glycogen storage, ALB secretion and CYP450 enzyme activity were detected in cells at different time points. As shown in Fig. 1B, nearly all the cells at day 10 stained positive for periodic acid-Schiff (PAS), whereas no positive staining was visible in the progenitor cells and the HGF-induced cells. Similar to PAS staining, the ALB secretion rate from cells at day 5 and day 10 were significantly upregulated compared to that at day 0 (Fig. 1C). As expected, ~90±2.5% of the differentiated cells at day 10 exhibited CYP2B1/2 activity, compared to 10±4.3% of the cells at day 5 (Fig. 1D). The CYP1A1 and CYP1A2 activities in the cells at day 10 were greater than the equivalent activities in the cells at day 5 (Fig. 1E). These results suggest that this differentiation protocol gives rise to functional HLCs.

Differentiation of hFHPCs into polarized hepatocyte-like cells

In parallel to functional differentiation, hFHPCs undergo cell morphogenesis and may acquire polarity. To determine whether hepatocyte polarity is established in this differentiation process, the assembly of TJs was analyzed at different time points. The localization of TJ-associated protein ZO-1 and F-actin were

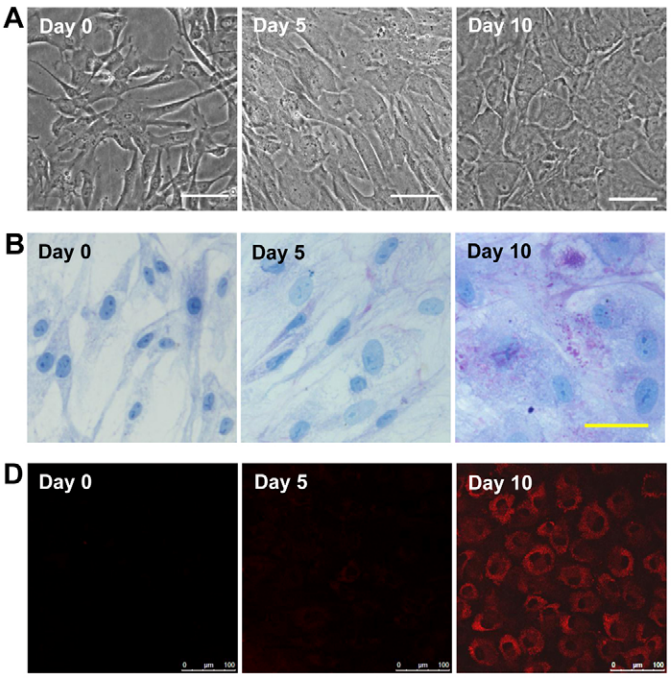

Fig. 1. Functional properties of hFHPCs and their progeny. (**A**) Morphology of cells under phase-contrast microscopy at different time points. (**B**) Intracellular glycogen in the hFHPCs was analyzed by PAS staining at different time points. (**C**) Albumin secretion was analyzed by ELISA. **P*<0.05 compared with day 0. (**D**) CYP2B1/2 enzyme activity in the differentiated hFHPCs was evaluated by PROD assay at different time points. (**E**) Activity of CYP1A1 and CYP1A2 was measured by EROD and MROD, respectively. **P*<0.05 compared with day 5. The data are presented as the mean ± s.d. (*n*=3). Scale bars: 50 μm unless labeled otherwise.

detected using anti-ZO-1 and FITC-conjugated phalloidin and visualized with confocal microscopy. As shown in Fig. 2A, we examined sections from day 0, day 5 and day 10 cultures in both the *x-z* and *x-y* planes. In *x-y* confocal images of hFHPCs (day 0), F-actin drew a network through the cells. In *x-z* cross-sections through the cells, small spots of ZO-1 were observed along an individual cell-cell border (Fig. 2A, left panel, arrow). After HGF stimulation for five days, F-actin was strongly concentrated at cell-cell boundaries, and in *x-z* confocal images, ZO-1 appeared as punctate spots at sites of cell-cell contact throughout the cell layer (Fig. 2A, middle panel, arrow). At day 10, thick F-actin bundles around the cells were increased in abundance and size; ZO-1 protein was concentrated along cell-cell borders in a linear pattern. At the same time, in *x-z* confocal images, ZO-1/F-actin double-staining spots increased at sites of cell-cell contact (Fig. 2A, right panel, arrow). The localization of another TJ-associated protein, junctional adhesion molecule A (JAM-A) (Braiterman et al., 2007; Paris et al., 2008), was determined later. Fig. 2B shows that JAM-A was predominantly expressed at the border of two adjacent cells at day 10. Moreover, ultra-thin sections displayed the TJ belt in hFHPC-derived HLCs (Fig. 2C). These results raise the possibility that TJs form between the hFHPC-derived HLCs during differentiation.

To test whether these morphogenetic features represent the acquisition of functional TJs, we evaluated the barrier function of TJs, the transepithelial electrical resistance (TER) and the paracellular fluxes. The values of TER in cells at day 5 and day 10 were significantly increased compared to the undifferentiated cells (day 0) (Fig. 2D). To determine the dynamic function of TJs, apical-to-basolateral FITC-dextran leakage across cultures was examined upon differentiation. In contrast to TER, which is an instantaneous measure of TJ functional integrity, FITC-dextran diffusion is measured over a period of 30 min and therefore may be a more sensitive measure of apical-to-basolateral leakage. The results reveal that permeability declined from day 6 after onward. Paracellular flux rates of FITC-dextran in differentiated cells at day 5, day 6, and day 8 were 101%, 81% and 69%, respectively, relative to that of the cells at day 0 (Fig. 2E).

To examine whether the hFHPC-derived HLCs formed BC-like structures, their ultra-structures were studied in cultures at day 10. Ultrathin sections displayed BC with microvilli (Fig. 3A, arrow) among the hFHPC-derived HLCs. To evaluate the functional activity of drug transporters in HLCs, 5(6)-carboxy-2,7-dichlorofluorescein diacetate (CDFDA) was added into the cultures at day 10 in the absence or presence of probenecid, an inhibitor of the multidrug resistance-associated protein (MRP) 2-mediated transport of 5 (and 6)-carboxy-2,7-dichlorofluorescein (CDF) in hepatocytes (Zamek-Gliszczynski et al., 2003). Fig. 3B shows that the fluorescent CDF was presented in the BC lumen among hFHPC-derived HLCs in the absence of probenecid. A significant decrease in CDF was observed in the BC lumen in the

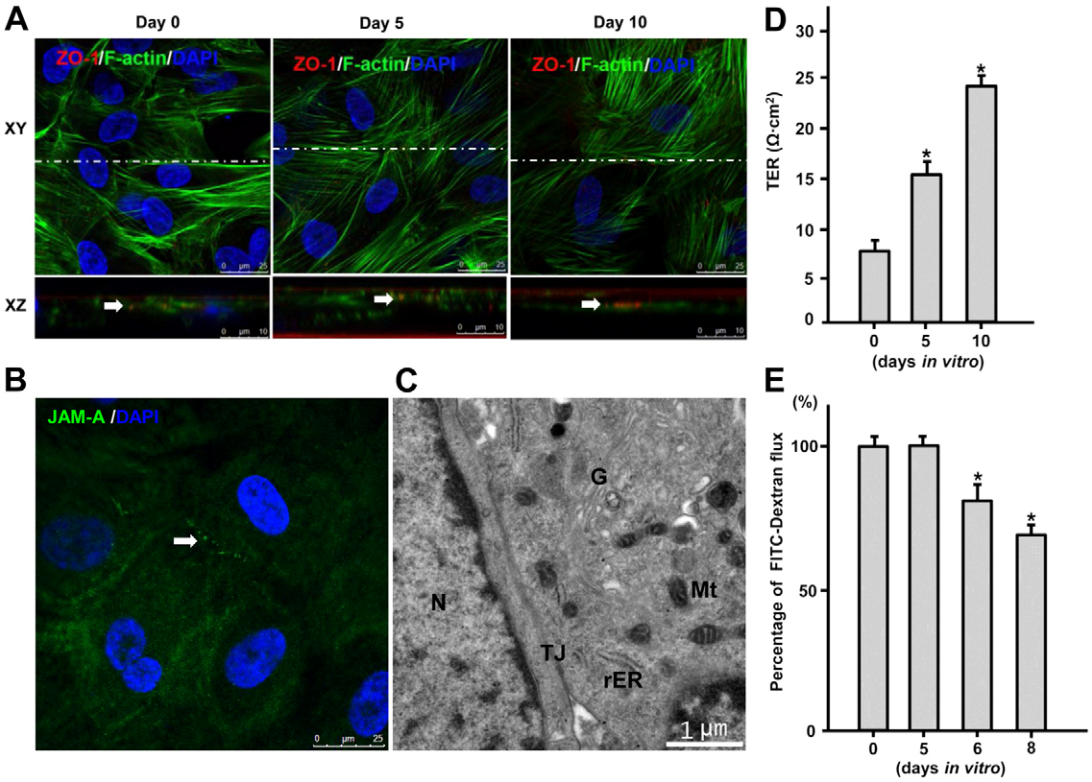

Fig. 2. TJ-associated structures and functions in hFHPCs and their progeny. (**A**) Cultured cells double-stained for ZO-1 and F-actin were imaged in the *x-y* plane and the *x-z* plane. The *x-z* planes taken as indicated (white line) on the corresponding *x-y* planes are shown. Arrows indicate the colocalization of ZO-1 and F-actin. (**B**) Cells at day 10 stained for JAM-A were imaged by confocal microscopy. Arrow indicates the localization of JAM-A. (**C**) Ultrathin section showing a TJ between the cells at day 10. N, nucleus; TJ, tight junction; rER, rough endoplasmic reticulum; G, Golgi; Mt, mitochondria. (**D**) Barrier function measured as TER in the differentiated hFHPCs at the indicated times. **P*<0.05 compared with day 0. (**E**). Barrier function measured as paracellular fluxes using FITC-dextran in the hFHPCs. **P*<0.05 compared with day 0. The data are presented as the mean ± s.d. (*n*=3).

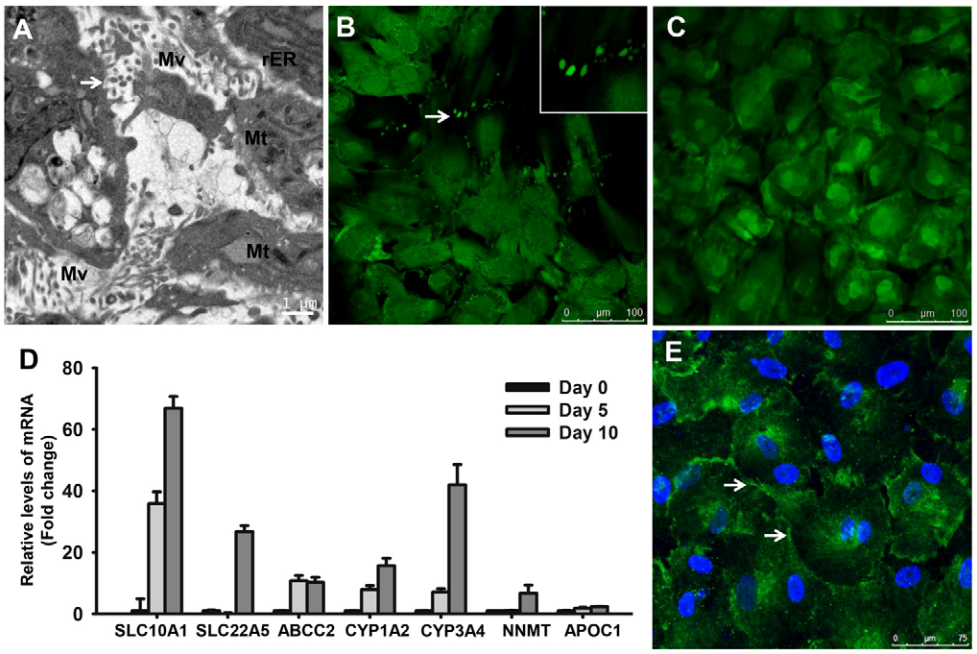

Fig. 3. Gene expression and functional polarization in hFHPC-derived HLCs. (**A**) Ultrathin section showing a BC among the cells. Arrows indicate the microvillus. Mv, microvillus; rER, rough endoplasmic reticulum; Mt, mitochondria. CDFDA is internalised by hFHPC-derived functional HLCs, cleaved by intracellular esterases and excreted into BC as fluorescent CDF without (**B**) and with (**C**) 4 mM probenecid. Arrows indicate the fluorescent CDF transported into BC. (**D**) Quantitative comparison of the transcription of phase I and phase II enzymes as well as drug transporters in hFHPCs and their progenies. SLC10A1, solute carrier family 10 (sodium/bile acid co-transporter family), member 1; SLC22A5, solute carrier family 22 (organic cation/carnitine transporter), member 5; ABCC2, ATP-binding cassette, sub-family C (CFTR/MRP), member 2; CYP1A2, cytochrome P450 1A2; CYP3A4, cytochrome P450 3A4; NNMT, nicotinamide *N*-methyltransferase; APOC1, apolipoprotein C-I. (**E**) hFHPC-derived HLCs stained for OATP were imaged by confocal microscopy. Arrows indicate OATP-positive staining.

presence of probenecid, suggesting that the excretion of CDF in hFHPC-derived HLCs was mediated by MRP2, a transporter in the apical pore of hepatocytes that functions canalicular secretion (Fig. 3C). Quantitative comparisons of gene expression reveal that the mRNA level of phase I/II enzymes as well as that of drug transporters in hFHPC-derived HLCs were significantly higher than in hFHPCs (Fig. 3D). Fig. 3E demonstrates the basal surface staining of the organic anion transporting polypeptide (OATP) (König et al., 2000) in hFHPC-derived HLCs. These results suggest that this differentiation protocol gives rise to functional polarized HLCs in a way that resembles natural hepatocyte development.

Differential gene expression and functional annotation analysis during hFHPC differentiation

To obtain an initial perspective on global gene expression changes, we performed a pairwise comparison of gene expression microarray data based on the stage of hFHPCs differentiation. A Volcano Plot filtering identified two major transitions in the gene expression patterns (supplementary material Fig. S2): undifferentiated hFHPCs to immature HLCs (stage I) and immature HLCs to mature HLCs (stage II). The microarray revealed that 1780 genes were significantly affected (814 upregulated and 966 downregulated) in stage I; 1835 genes were significantly affected (970 upregulated and 865 downregulated) in stage II. The regulated genes among the lists derived from the two stages were compared by a Venn diagram, which led to the classification of three gene sets: 514 significantly regulated genes were shared by the two stages; 1266 genes were unique to stage I; 1321 regulated genes were unique to stage II (supplementary material Fig. S3A). The 514 overlapping genes were classified into four gene subsets by the hierarchical clustering (supplementary material Fig. S3B; Table S2).

Functional gene annotation analysis of differentially expressed genes was performed according to the gene subsets to explore the functions of the regulated genes in hFHPCs differentiation. Analysis of the significantly upregulated genes found in stage I led to the identification of functional groups, such as ‘cell adhesion’, ‘extracellular matrix organization’, and ‘anatomical structure morphogenesis’ (Fig. 4A; supplementary material Table S3-1). The downregulated gene functional groups in stage I included ‘cell cycle’, ‘organelle organization’, and ‘cellular metabolic process’ (Fig. 4B; supplementary material Table S3-2). A similar analysis was performed for regulated genes found in stage II, and the upregulated gene functional groups included: ‘alcohol metabolic process’, ‘glycolysis’, and ‘cellular amino acid metabolic process’ (Fig. 4C; supplementary material Table S4-1). The functional classification of the downregulated included groups: ‘cell-cell signalling’, ‘anatomical structure morphogenesis’, and ‘cellular developmental process’ (Fig. 4D; supplementary material Table S4-2). The ten most significantly overrepresented annotations within the functional classification of the regulated genes shared by the two stages are listed in supplementary material Table S5.

Taken together, these results suggest that the gene expression patterns of undifferentiated hFHPCs, immature HLCs and mature HLCs are in keeping with the changes in cell phenotypes, and

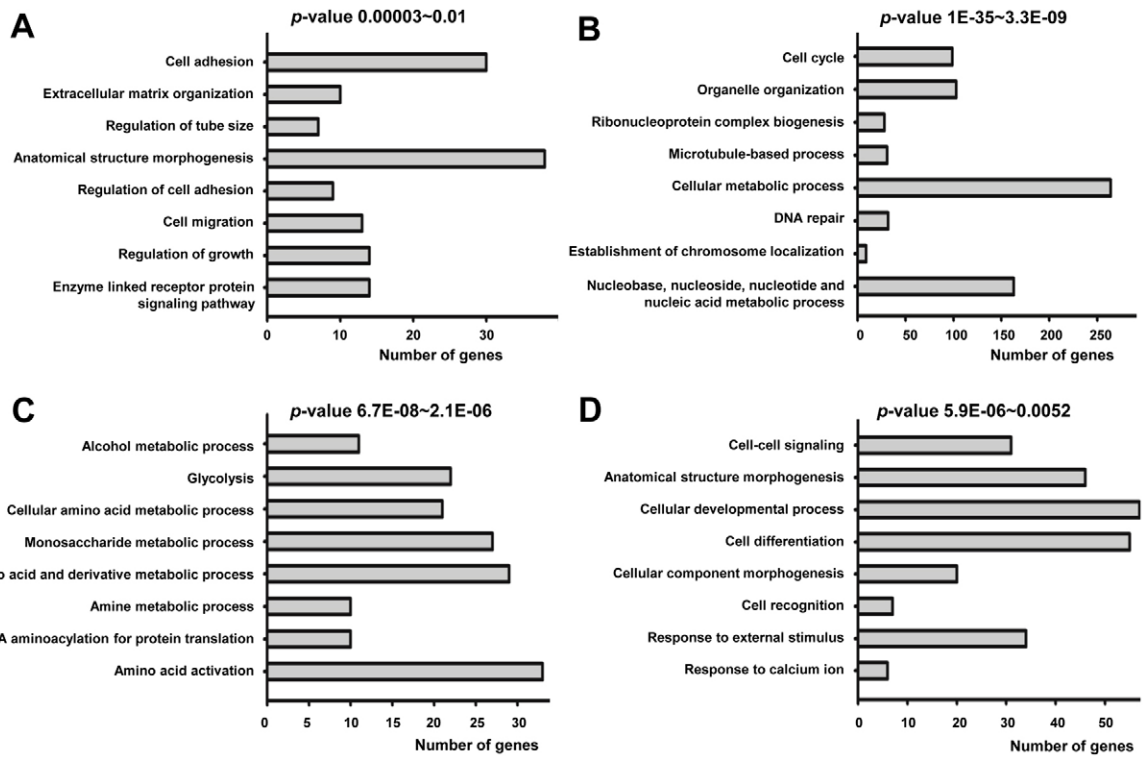

Fig. 4. Most enriched GO biological process terms for genes expressed differentially during hFHPC differentiation. (**A**) Most significant biological process (BP) terms for upregulated genes unique to stage I. (**B**) Most significant BP terms for downregulated genes unique to stage I. (**C**) Most significant BP terms for upregulated genes unique to stage II. (**D**) Most significant BP terms for downregulated genes unique to stage II.

reflecting the specific effects of regulators on hepatocyte differentiation.

Functional network analysis of the differentially expressed genes involved in cell morphogenesis

To gain a better understanding of the molecular mechanisms that regulate hepatocyte polarity generation during hFHPCs differentiation, the differentially expressed genes within the defined GO categories (GO 0000902: cell morphogenesis; GO 0022604: regulation of cell morphogenesis; GO 0010769: regulation of cell morphogenesis involved in differentiation) were selected for further examination. We found 63 genes (40 upregulated, 23 downregulated) in stage I and 69 genes (29 upregulated, 40 downregulated) in stage II, and 19 genes were common between the two stages (supplementary material Table S6). Among the common genes, six genes (*ADM*, *APOE*, *CDC42EP3*, *EFNB1*, *FBLIM1* and *THY1*) were upregulated, whereas *IL6*, *NOX4* and *RAC2* were downregulated, in both stages; eight genes (*ANTXR1*, *DFNB31*, *GDF7*, *ITGB4*, *LIMA1*, *MBP*, *MYH10* and *TGFB3*) were upregulated in stage I and downregulated in stage II. In contrast, *SPG3A* and *VEGF* were downregulated in stage I and upregulated in stage II (supplementary material Fig. S4).

Establishment of epithelial cell polarity is part of cell morphogenesis, which is driven by the forces that cells produce both internally and exert on neighboring cells. To examine the correlations between the candidate genes and hepatocyte polarization, the candidate genes were further categorized using GO analysis. The candidate genes were involved in 'establishment or maintenance of cell polarity', 'membrane organization', 'cell junction organization', and 'cytoskeleton organization' (Table 1). Interestingly, three gene subsets involved in the establishment or maintenance of cell polarity were identified. *PTK7*, *PARD3* and *PRKCI* were upregulated in stage I, with *PTK7* and *PRKCI* responsible for establishing or maintaining epithelial cell apical/basal polarity. *KNTC2*, *CDC42*, *CENPA* and *CCDC99* were downregulated in stage I, with *KNTC2*, *CENPA* and *CCDC99* responsible for establishing mitotic spindle orientation; *CAP2*, *DOCK2* and *NDE1* were downregulated in stage II, all of which are involved in cytoskeleton organization. One gene subset involved in membrane organization (*DAB2*, *APOE*, *MYH10*, *PRKCI* and *ULK1*) was upregulated in stage I, all of which are involved in 'vesicle-mediated transport'. Among the genes involved in cell junction organization, *THY1*, *GJA1*, *PRKCI* and *TGFB3* were upregulated in stage I; *SMAD3*, *SMAD7*, *CDC42* and *DLC1* were downregulated in stage I; and, *SHROOM2* and *TGFB3* were downregulated in stage II. *CDC42* and *PRKCI* are involved in the formation of epithelial TJs. Thus, the GO analysis showed that the candidate genes were assigned to the key biological processes required for cell polarization.

To better visualize the functional interactions at the network level, the candidate genes were submitted to functional network reconstruction using STRING. As shown in Fig. 5, 16 genes (*FYN*, *RHOJ*, *PLXNB1*, *DLC1*, *CDC42*, *NTN1*, *CDC42EP3*, *PARD3*, *EPHB3*, *EFNB1*, *PVRL1*, *PRKCI*, *TPM1*, *RAC2*, *LIMA1*, and *MYH10*) (Fig. 5A) were clustered together. Within this functional cluster, *CDC42*, *MYH10*, *PARD3* and *PRKCI* are involved in TJs, and *FYN*, *CDC42*, *PARD3* and *PVRL1* are involved in adherens junctions. *CLASP1*, *KNTC2* (*NDC80*), *CENPA*, *CCDC99* and *NDE1* participate in maintenance of

Table 1. Selected differential express genes involved in the HLCs polarity formation

Upregulated genes in Stage I
1. Establishment or maintenance of cell polarity (including apical/basal cell polarity): *PTK7*, *PARD3* and *PRKCI*
2. Regulation of cell shape: *CDC42EP3*, *FYN*, *FBLIM1*, *MYH10* and *RHOJ*
3. Regulation of cell size: *APBB1*, *DAB2*, *TGFB3*, *TGFBR3* and *ULK1*
4. Regulation of cellular component size: *LIMA1*, *APBB1*, *DAB2*, *TGFB3*, *TGFBR3* and *ULK1*
5. Cell junction organization: *THY1*,*GJA1*, *PRKCI* and *TGFB3*
6. Membrane organization: *DAB2*, *APOE*, *MYH10*, *PRKCI* and *ULK1*
7. Cellular component organization: *SMO*, *SLITRK5*, *SEMA5A*, *S100A4*, *RXRA*, *ROBO3*, *LAMB1*, *ITGA1*, *GDF7*, *EFNB1*, *DFNB31* and *ADM*
8. Regulation cellular component organization: *RUFY3*, *MBP*, *PLXNB2* and *PLXNB1*
9. Cellular component assembly: *BBS1*, *LIMA1*, *THY1*, *GJA1*, *APOE*, *ITGB4*, *MYH10*, *OFD1*, *PARD3* and *TTLL3*
10. Cytoskeleton organization: *BBS1*, *LIMA1*, *THY1*, *ANTXR1*, *DMD*, *APOE*, *MYH10*, *OFD1*, *PRKCI*, *RHOJ* and *TTLL3*
11. Actin cytoskeleton organization: *LIMA1*, *ANTXR1*, *MYH10*, *PRKCI*, *RHOJ* and *TTLL3*

Downregulated genes in Stage I
1. Establishment or maintenance of cell polarity: *KNTC2*, *CDC42*, *CENPA* and *CCDC99*
2. Regulation of cell shape: *DLC1* and *VEGF*
3. Cell junction organization: *SMAD3*, *SMAD7*, *CDC42* and *DLC1*
4. Cellular component organization: *TTL*, *CDC42*, *CCDC99*, *KNTC2*, *DLC1*, *SMAD7*, *SMAD3*, *RAC2*, *CENPA*, *VEGF*, *UCHL1*, *TRAPPC4*, *SPG3A*, *SH2B*, *NOX4*, *KIAA1893*, *IL6*, *HPRT1*, *HES1*, *GLI2*, *EGR2* and *CDH4*
5. Cytoskeleton organization: *KNTC2*, *CDC42*, *CENPA*, *CCDC99*, *DLC1*, *RAC2* and *TTL*
6. Actin cytoskeleton organization: *CDC42*, *DLC1* and *RAC2*
7. Regulation of actin filament bundle formation: *SMAD3* and *DLC1*

Upregulated genes in Stage II
1. Establishment or maintenance of cell polarity: *CLASP1*
2. Regulation of cell shape: *PALM2-AKAP2*, *CDC42EP3*, *FBLIM1* and *VEGF*
3. Cellular component organization: *ARL6*, *BCL6*, *DKFZP586P0123*, *JAK2*, *KLF7*, *L1CAM*, *THY1*, *ADM*, *SPG3A*, *COL4A3BP*, *CLASP1*, *EFNB1*, *APOE*, *HIF1A*, *IL7R*, *KIF3A*, *NFATC1*, *NUMBL*, *STK4*, *SLC1A3*, *STXBP1*, *UNC5B* and *CSPG2*
4. Regulation of cellular component organization: *PALM2-AKAP2*, *CDC42EP3*, *THY1*, *CLASP1*, *FBLIM1*, *APOE*, *LRRC4C*, *NUMBL*, *STXBP1* and *VEGF*
5. Cytoskeleton organization: *BCL6*, *THY1*, *CLASP1* and *APOE*

Downregulated genes in Stage II
1. Establishment or maintenance of cell polarity: *CAP2*, *DOCK2* and *NDE1*
2. Regulation of cell size: *NGFB*, *NTN1*, *PDGFB*, *SEMA3A* and *TGFB3*
3. Cell-cell junction organization: *SHROOM2* and *TGFB3*
4. Cellular component organization: *EPHB3*, *GDF7*, *CABP4*, *WNT3A*, *COL18A1*, *NOX4*, *SPON2*, *NRP2*, *PVRL1*, *PTPRF*, *GAP43*, *ETV4*, *BDNF*, *IL6*, *EGFR*, *CHRNB2* and *DFNB31*
5. Regulation of cellular component organization: *CHRNB2*, *MBP*, *PVRL1*, *PTPRF* and *TTC3*
6. Cellular component assembly: *FGD6*, *LIMA1*, *GAS7*, *ITGB4*, *MYH10*, *NRXN3*, *ONECUT2*, *PDGFB*, *RAC2*, *TPM1* and *VANGL2*
7. Regulation of cellular component size: *LIMA1*, *NGFB*, *NTN1*, *PDGFB*, *SEMA3A*, *SHROOM2* and *TGFB3*
8. Cytoskeleton organization: *CAP2*, *FGD6*, *LIMA1*, *ANTXR1*, *DOCK2*, *GAS7*, *MYH10*, *NDE1*, *PDGFB*, *RAC2* and *TPM1*
9. Actin cytoskeleton organization: *FGD6*, *LIMA1*, *ANTXR1*, *DOCK2*, *GAS7*, *MYH10*, *PDGFB*, *RAC2* and *TPM1*
10. Actin filament bundle formation: *LIMA1* and *GAS7*
11. Regulation of actin cytoskeleton organization: *LIMA1*, *SHROOM2* and *TPM1*

kinetochore integrity and kinetochore-microtubule attachments (Fig. 5B).

Confirmation of the differential expression of genes by real-time RT-PCR

To validate the accuracy of the gene indexes calculated from the microarray, six differentially expressed genes selected based on their roles in hepatocyte function (*ABCA1*, *CPS1* and *PCK2*) or cell morphogenesis (*APOE*, *CDC42* and *PTK7*) were confirmed by real-time RT-PCR. Comparison of the microarray and real-time RT-PCR results with their correlation coefficients are presented in Fig. 6. The data reveal that our gene index calculations from the microarray data accurately represented the expression of genes across a wide range of expression levels.

Functional analysis of candidate genes involved in the hepatocyte polarity

As a proof of principle, *PTK7* and *PARD3* were selected from the differentially expressed genes involved in hepatocyte polarity based on their unique cellular function. By reducing their expression level using siRNA, their effects on the formation of hepatocyte polarity during hFHPC differentiation were assessed. As shown in Fig. 7A, we successfully reduced the *PTK7* and *PARD3* expression level by 80–90% compared to the transfection control. The reduction of *PTK7* and *PARD3* significantly decreased the TER and increased the paracellular fluxes over a period of 30–120 min at 6 days post-transfection (Fig. 7B,C).

To identify morphology changes, hFHPCs with reduced expression of *PTK7* and *PARD3* were challenged to undergo differentiation into hepatocytes using our two-stage differentiation protocol. As shown in Fig. 7D, cells with inactivated *PTK7* and *PARD3* were unable to acquire the typical epithelial morphology, which exhibited the effects on F-actin assembly and apical actin concentration. Furthermore, the Golgi apparatus was randomly distributed in the cytoplasm in *PTK7* and *PARD3* knockdown cells, while the Golgi apparatus was localized between the nucleus and apical pole in control cells. These data demonstrate that reduction of either *PTK7* or *PARD3* may inhibit hepatocyte polarity formation during hFHPC differentiation.

Discussion

Generation of membrane polarity in hepatocytes is key developmental program and a multistage process that requires cell-cell interactions and the specific organization of proteins and lipids on the cell surface and in cell interior, which leads to modeling the complex architecture during hepatocyte

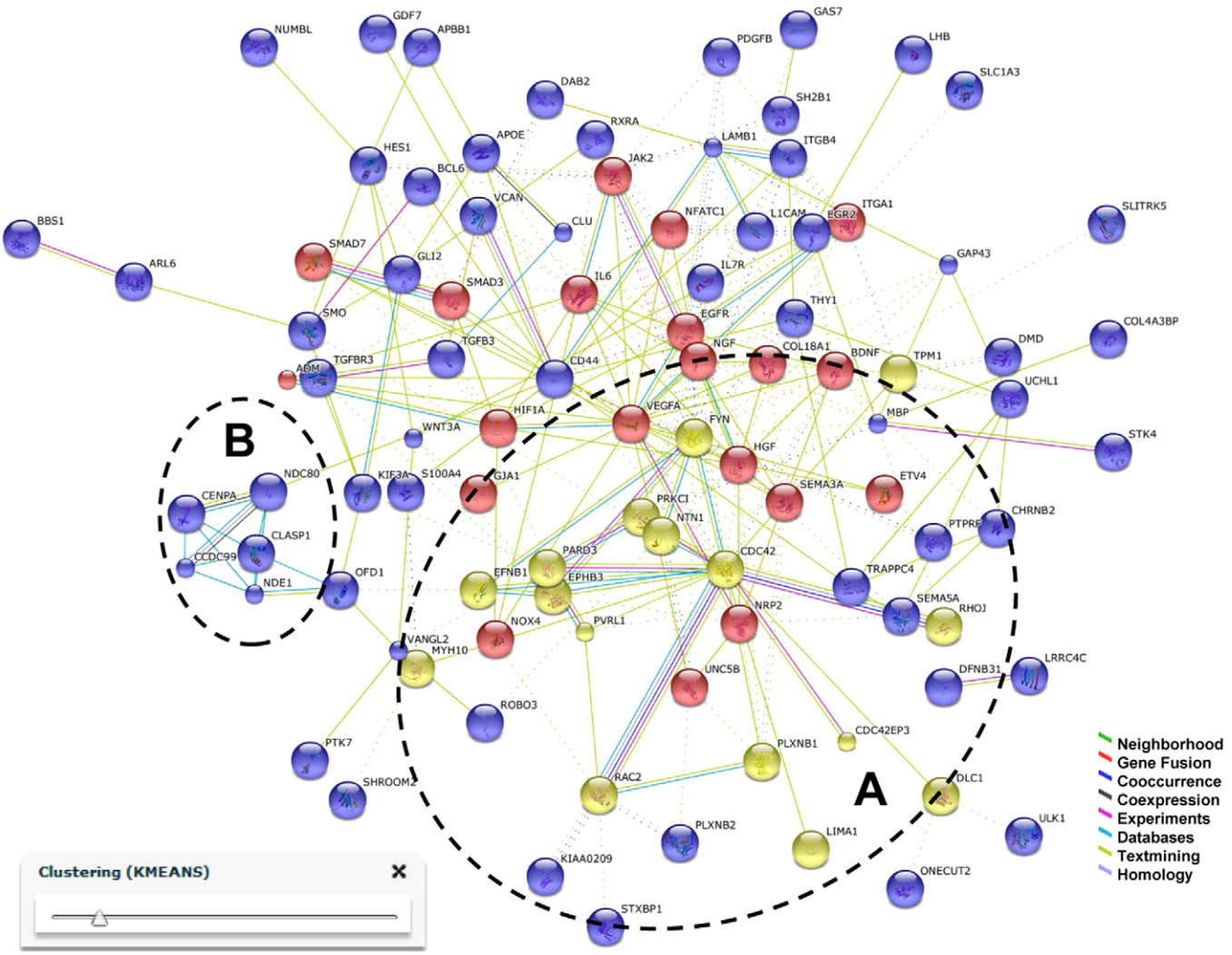

Fig. 5. Predicted protein network visualization with STRING. The network view predicted the associations between proteins from the regulated genes involved in cell morphogenesis during hFHPC differentiation. The network nodes are proteins. The edges represent the predicted functional associations with seven differently colored lines. These proteins were clustered using KMEANS clustering algorithms, depicted in the panel (arrow). Every color of node corresponds to a cluster. A and B indicate the identified clusters.

differentiation (Wang and Boyer, 2004; Takashi et al., 2007). Although information has recently been accumulated suggesting specific routes and mechanisms for this process (Matsui et al., 2002; Michalopoulos et al., 2003; Imamura et al., 2007; Chiao et al., 2008; Bonora-Centelles et al., 2009; Kidambi et al., 2009; Saulnier et al., 2010; Synnergren et al., 2010; Jozefczuk et al., 2011; Fu et al., 2011), the exact mechanism for hepatocyte polarity generation is unclear partly because only a few *in vitro* models are available for developing complex human hepatocyte polarity (Decaens et al., 2008). The goal of this study was to generate functional polarized hepatocytes from the hFHPCs *in vitro* and to explore the molecular mechanisms involved in hepatocyte polarization.

First, we found that sequential treatment of hFHPCs with HGF, OSM, and DEX is efficient to generate functional hepatocyte polarity in hFHPC-derived hepatocytes, which mimics hepatocyte differentiation both morphologically and functionally and resembles the natural hepatocyte development (Kinoshita et al., 1999; Kamiya et al., 2001; Zorn, 2008). We observed that hFHPCs were unpolarised with spindle-shaped morphology and possessed higher AFP and lower ALB expression under the initial culture conditions. After HGF treatment, the cells developed a simple polar phenotype with the establishment of functional TJs at the points of cell-cell contact along the cell perimeters, as evidenced by the zonular assembly of F-actin, ZO-1 and JAM-A through the function of TJs (TER and paracellular flux). These results suggest that HGF stimulates a morphogenesis of tight and functional integrity in the early stage of hFHPC differentiation, which is in line with previous reports (Pollack et al., 2004). In stage II, the immature cells were induced with HGF, OSM and DEX, and the multipolar hepatic phenotype gradually manifested itself during this period. Parallel to the observed morphological changes, ALB secretion, bile acid synthesis, glycogen storage and CYP450 enzyme activity accelerated, suggesting that the hepatocytes polarized rapidly. The formation of the BC architecture and the establishment of the hepatobiliary transport system in hFHPC-derived HLCs was further confirmed by electron microscopy and hepatocyte uptake assays. Taken together, our data suggest that a

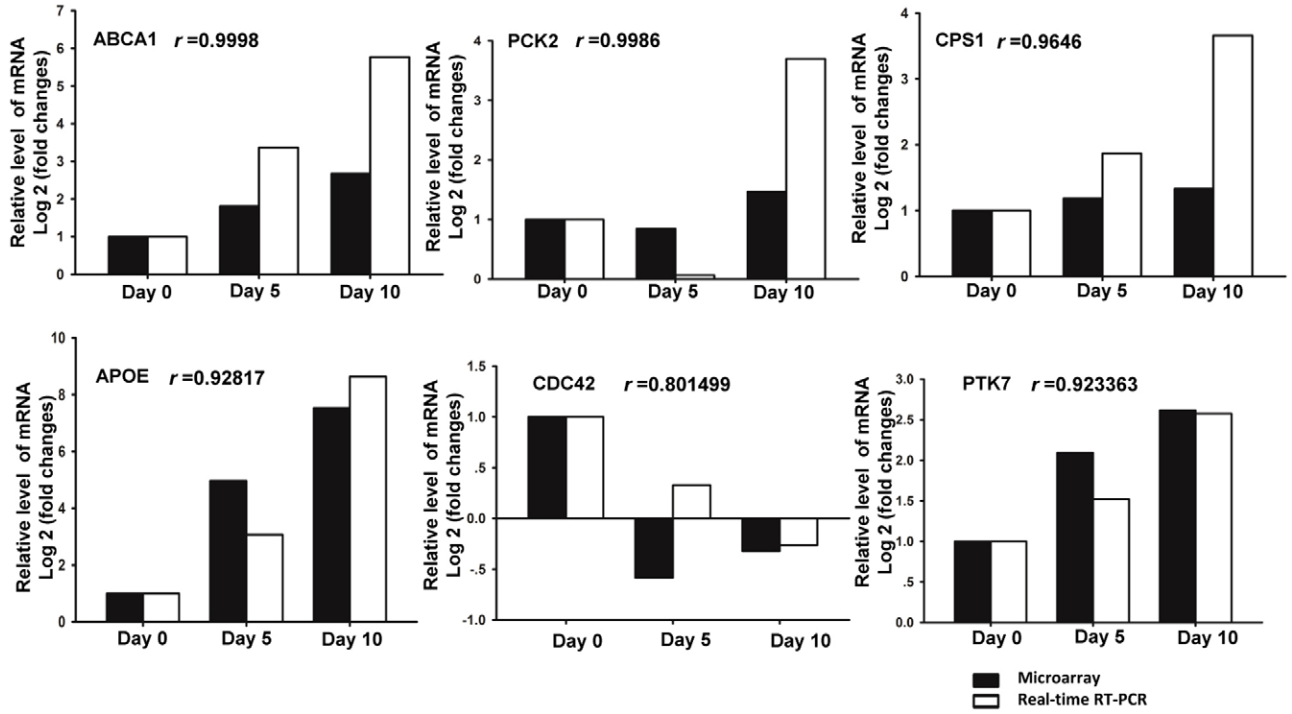

Fig. 6. Gene indexes from the microarray validated by real-time RT-PCR. The expression levels of six genes of interest in the hFHPCs were verified by real-time RT-PCR. The relative expression of each gene was normalized against 18S rRNA. Shown is a $\log_2$ plot of the fold changes, *r* indicates the correlation coefficients. ABCA1, ATP-binding cassette, sub-family A (ABC1), member 1; PCK2, phosphoenolpyruvate carboxykinase 2; CPS1, carbamoyl-phosphate synthase 1, mitochondrial; APOE, apolipoprotein E; CDC42, cell division cycle 42 (GTP binding protein); PTK7, PTK7 protein tyrosine kinase 7.

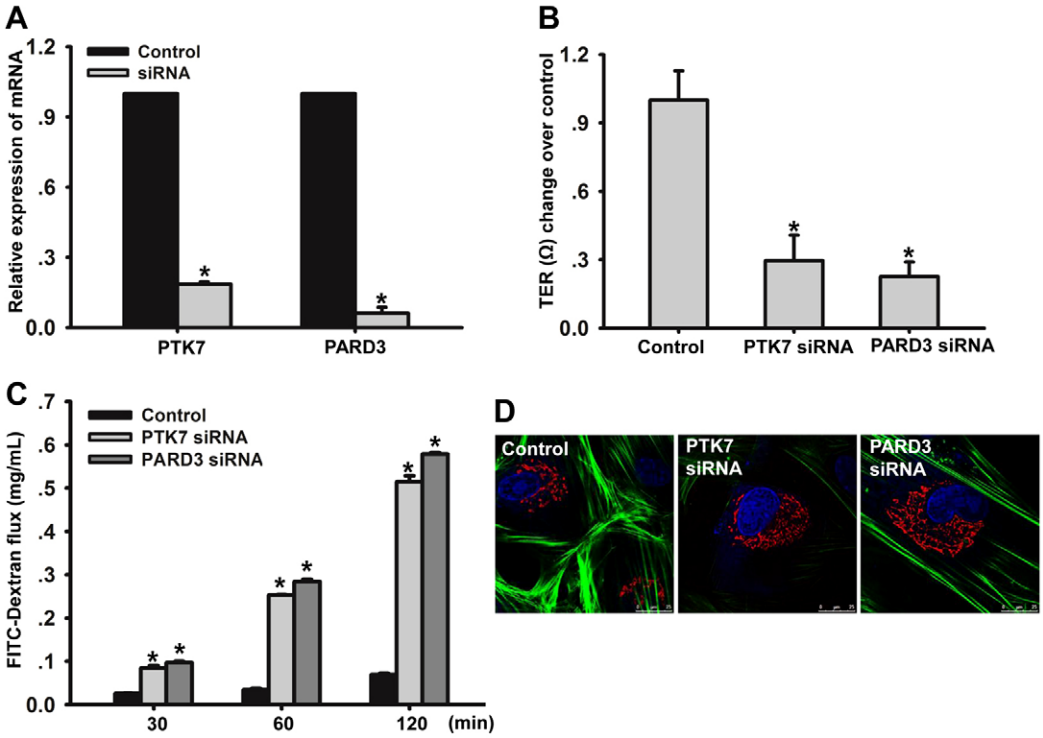

Fig. 7. Effect of gene reduction by siRNA transfection on hepatocyte polarity during hFHPC differentiation. (**A**) Real-time RT-PCR analysis of the transcript levels of the *PTK7* and *PARD3* genes 24 hours after siRNA transfection. Non-targeting siRNAs were used as a control. (**B**) Barrier function was measured as TER 24 hours after siRNA transfection in hFHPCs cultured in the presence of induction medium with HGF alone for 5 days. (**C**) Barrier function was measured as paracellular fluxes using FITC-dextran 24 hours post-siRNA transfection in hFHPCs cultured in the presence of induction medium with HGF alone for 5 days. (**D**) Localization of F-actin and Golgi 24 hours after siRNA transfection in hFHPCs cultured in the presence of induction medium with HGF for 5 days and HGF, OSM and DEX for another 5 days. Green, F-actin; red, Golgi. The data are presented as the mean ± s.d. ($n=3$); *$P<0.05$ compared with control.

human hepatocyte polarity was gradually generated during hFHPC differentiation *in vitro*. Therefore, this system may be a useful platform to explore the mechanisms of hepatocyte differentiation and hepatic polarity establishment.

Second, using high-density oligonucleotide microarrays, the global transcriptional profile of hFHPCs and their derivatives was investigated based on the stage of differentiation. We found 1780 genes in stage I and 1835 genes in stage II that were significantly regulated. From comparison of the transcriptional profiles between the two differentiation regimes, three main subsets of regulated genes were identified. Within each subset, the functional profiling of the upregulated and downregulated genes was further investigated using GO analysis. The upregulated genes in stage I were mainly involved in 'cell adhesion', 'extracellular matrix organization', and 'anatomical structure morphogenesis'. The upregulated genes in stage II were mainly involved in 'alcohol metabolic process', 'glycolysis', and 'cellular amino acid metabolic process'. Results revealed a dramatic switch in functional profile as proliferating hFHPCs differentiated into HLCs, which was in accordance with the cell phenotypes and previous reports on liver development (Lemaigre, 2009; Duncan, 2003; Zorn, 2008). These results present a detailed characterization of a unique set of genes, which can be used to explore the mechanisms of hepatocyte differentiation.

To examine which molecules govern hepatocyte morphogenesis, including polarization during differentiation, a subset of regulated genes (which included 113 genes) annotated as being associated with 'cell morphogenesis' in the two stages was selected (supplementary material Table S6) to explore the core regulated genes that are potentially involved in hepatocyte polarization. As expected, our data (Table 1) indicated that these core genes were involved in the key process of polarization, such as in apical/basal cell polarity and junction organization. Among them, *PTK7* (Lhoumeau et al., 2011; Puppo et al., 2011), *PARD3* (Ooshio et al., 2007; Achilleos et al., 2010; Ishiuchi and Takeichi, 2011), *PRKCI* (Eder et al., 2005), *CDC42* (Harris and Tepass, 2010), and *CLASP1* (Miller et al., 2009) have recently been identified to have roles in cell polarity. Our results demonstrated that *PTK7* and *PARD3* promote the formation of hepatic polarity and affect F-actin organization. The network analysis of the candidate genes, based on known and predicted protein-protein interactions, further illustrated the interactions of the molecules that regulate hepatocyte polarization and clustered them into functional categories. Sixteen genes (*FYN*, *RHOJ*, *PLXNB1*, *DLC1*, *CDC42*, *NTN1*, *CDC42EP3*, *PARD3*, *EPHB3*, *EFNB1*, *PVRL1*, *PRKCI*, *TPM1*, *RAC2*, *LIMA1* and *MYH10*) may be the functionally regulated modules for hepatocyte polarity generation.

In conclusion, our results provide insights into understanding hepatocytes polarization at the transcriptional level and reveal novel key genes and a functional network of potential importance for future efforts in exploring hepatic differentiation.

Materials and Methods

Culture and differentiation of hFHPCs

hFHPCs were prepared as described previously (Wang et al., 2008) with the informed patient consent and under the approval of the Ethics Committee of Capital Medical University (Beijing, China) for this project. hFHPCs were maintained in MEM/NEAA (Hyclone, Logan, UT) containing 5% fetal bovine serum (FBS, Hyclone), 100 U/mL penicillin, 100 μg/mL streptomycin, 20 ng/mL epidermal growth factor (EGF, Peprotech, Rocky Hill, NJ) and 10 μg/mL insulin (Sigma-Aldrich, St Louis, MO). Cells from passages nine to eleven were used in the study.

To induce of hepatocyte maturation, the cells were plated at 2×10^4 cells/cm^2 in culture plates and dishes, in expansion medium. Once cells reached 95% confluence, they were washed twice with PBS and differentiated in basal medium [MEM/NEAA, supplemented with 0.5 mg/mL albumin fraction V (Sigma-Aldrich), 100 U/mL penicillin, 100 μg/mL streptomycin, and 1% insulin-transferrin selenium (ITS, Sigma-Aldrich)] containing 20 ng/mL hepatocyte growth factor (HGF, Peprotech) for 5 days. Then, the differentiated cells were further matured in basal medium containing 20 ng/mL HGF, 10 ng/mL oncostatin M (OSM, Peprotech) and 10^{-7} M dexamethasone (DEX, Sigma-Aldrich) for another 5 days. Differentiation media were changed every 2 days.

Immunofluorescence staining

The cells were fixed with 4% paraformaldehyde for 20 min at room temperature, followed by permeabilize with 0.3% Triton X-100 in PBS for 5 min and then, rinsed and blocked with 20% goat serum (Zsgb-Bio, China) for 60 min at room temperature. The cells were then incubated with rabbit anti-ZO-1 at 1:200 (Invitrogen, Carlsbad, CA), mouse anti-JAM-A at 1:100 (F11R, BD Biosciences, San Jose, CA), or goat anti-OATP at 1:100 (Santa Cruz Biotechnology, Santa Cruz, CA); rabbit anti-AFP at 1:200 (Product Number A0008, Dako); mouse anti-ALB at 1:500 (clone HAS-11, Product Number A 6684, Sigma-Aldrich); mouse anti-CK8 at 1:200 (Product Number A, C5301, Sigma-Aldrich); mouse anti-EpCAM at 1:400 (Cell Signaling, MA); rabbit anti-Dlk at 1:500 (Abcam, Hong Kong) and rabbit anti-Sall4 at 1:50 (Abgent, San Diego, CA) at 4℃ overnight. Following three 5-min washes in PBS with gentle agitation, an Alexa Fluor-conjugated secondary antibody (1:500; Invitrogen) was added, and the samples were incubated for 60 min at 37℃. Afterwards, the cells were incubated with a mixture of FITC-labeled phalloidin (Sigma-Aldrich) and the nuclear stain DAPI (Sigma-Aldrich) at room temperature for 20 min. The cells were then washed and examined under a Leica TCS SP5 Confocal Microscope (Leica, Wetzlar, Germany). Photomultiplier settings were adjusted to the brightest signal and maintained during data acquisition. Samples were imaged either in the *x-y* plane or in the *x-z* plane.

Assay of Golgi apparatus localization

To mark the Golgi apparatus, a plasmid containing the sequence of the Golgi-resident enzyme N-acetylgalactosaminyltransferase 2 (GalNAc-T2) (Röttger et al., 1998) fused to red fluorescent protein (RFP) was transduced into cells using CellLight® Golgi-RFP *BacMam 2.0* (C10593, Invitrogen) (Kost et al., 2005) according to the manufacturers' instructions. Then, the labeled cells were fixed with 4% formaldehyde in PBS for 20 min at room temperature, followed by permeabilization with 0.3% Triton X-100 in PBS for 5 min at room temperature.

Transmission electron microscopy

For ultra-structural analysis, the differentiated cells were fixed in 2.5% glutaraldehyde in 0.1 M phosphate buffer, pH 7.2, for 60 min at 4℃. They were postfixed in 1% osmium tetroxide in 0.1 M phosphate buffer and embedded with the Spurr embedding kit. The sections were examined using an H-7650 electron microscope (Hitachi, Japan).

Glycogen assay

Intracellular glycogen was analyzed by PAS staining. Cultures were fixed in carnoy fixative at room temperature for 10 min and oxidised in 1% periodic acid for 15 min. After oxidation, the cultures were rinsed three times with distilled water and then treated with Schiff's reagent (Sigma-Aldrich) for 30 min. After the cells were rinsed with distilled water for 5 min, the cell nuclei were counterstained with Mayer's hematoxylin. The PAS-positive cells and whole cells were examined under a Leica light microscope (Leica).

Albumin analysis

The concentration of ALB secreted into the culture media was analyzed as previously described (Zhang et al., 2012).

Cytochrome P450 activity

To evaluate the activities of CYP1A1 and CYP1A2, ethoxyresorufin-O-deethylase (EROD) and methoxyresorufin-O-deethylase (MROD) assays were performed, respectively, as previously described (Donato et al., 1993; Kidambi et al., 2009). To evaluate the activity of CYP2B1/2, the penthoxyresorufin-O-deethylase (PROD) assay was performed. The differentiated cells were treated with 1 μM pentoxyresorufin and assessed under a Leica TCS SP5 Confocal Microscope (Leica). The PROD activity-positive cells were counted with Image-Pro Plus software (Media Cybernetics Inc., MD).

TER and permeability assay

To determine TER, cells were cultured to confluence in 12-mm Transwells with 0.4-μm pore size filters (Millipore, Billerica, MA). TER was measured in culture medium using a MILLICELL-ERS (Millipore) with 'chopstick' electrodes at different time points. For calculation, the resistance of blank filters was subtracted

from that of filters covered with cells to give the net resistance, which was multiplied by the membrane area to give the resistance in area-corrected units ($\Omega \cdot cm^2$). The data presented as the mean ± s.d.

To determine the paracellular flux, cells were cultured to confluence on 12-mm Transwells with 0.4-μm pore size filters (Millipore) and 2.5 mg/mL FITC-dextran (FD-70s, Sigma-Aldrich) was added to the inner chamber. The apparatus was then placed in a CO_2 incubator at 37℃ for 30, 60 and 120 min, respectively. Samples were collected from the outer chamber, and the fluorescent signals were measured at an excitation wavelength of 485 nm and emission wavelength of 535 nm using a spectrophotometer (Thermo Fisher Scientific Inc.). The results are expressed as the concentration of FITC-dextran in the outer chamber (mg/mL). The paracellular FITC-dextran flux rate was calculated as the measured value of the treated group relative to that of the control group.

At least three independent determinations of each parameter were compared among the treatment groups by one-way ANOVA using the statistical software SPSS 11.5. Differences were considered significant if $P<0.05$.

BC analysis

To determine the BC function, cells were incubated with 10 μM CDFDA (Sigma-Aldrich) in the absence or presence of 4 mM probenecid (Sigma-Aldrich) at 37℃ for 10 min to allow its internalization and subsequent translocation into the BC lumen by MRP 2 (Zamek-Gliszczynski et al., 2003; Turncliff et al., 2006). After extensive washes, the capacity of BC to contain the fluorescent CDF was analyzed under a Leica TCS SP5 Confocal Microscope (Leica).

Microarray experiments

Total RNA was isolated from undifferentiated hFHPCs and their derivatives in triplicate for each time point using TRIzol (Invitrogen) and the RNeasy kit (Qiagen, Hilden, Germany) according to the manufacturers' instructions, including a DNase digestion step. RNA quantity and purity were determined using the ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) and denaturing gel electrophoresis. Approximately 5 μg total RNA of each sample was used for labeling with a NimbleGen one-color DNA labeling kit, array hybridization using the NimbleGen Hybridization System (NimbleGen 12×135K microarrays, Roche NimbleGen) and array scanning using the Axon GenePix 4000B microarray scanner (Molecular Devices Corporation, Sunnyvale, CA).

Data analysis

The raw data were extracted and normalized using NimbleScan v2.5 Software. The gene summary files were imported into Agilent GeneSpring Software (version 11.0) for further analysis. Genes that had values greater than or equal to the lower cutoff of 50.0 in all samples were chosen for further analysis. Genes that were differentially expressed with statistical significance were identified through Volcano Plot filtering (fold change ≥2.0 and P-value ≤0.05).

Similarities and differences among the differentially expressed genes at different stages were analyzed by Gene Venn (Hulsen et al., 2008). Differentially expressed genes in these lists were analyzed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (Huang et al., 2008) according to Gene Ontology (GO) roles. To examine potential interactions of genes involved in cell morphogenesis during the hFHPCs differentiation, we applied the Search Tool for the Retrieval of Interacting Genes (STRING) to predict the protein-protein interaction networks (Szklarczyk et al., 2011).

Real-time RT-PCR

Total cellular RNA was extracted from 1×10^6 cells with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. For PCR analysis, 2 μg RNA was reverse-transcribed to cDNA using Superscript II reverse transcriptase and random hexamer primers (Invitrogen). Real-time PCR analysis was performed on an ABI Prism 7300 Sequence Detection System using the SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA). The reaction consisted of 10 μL of SYBR Green PCR Master Mix, 1 μL of a 5 μM mix of forward and reverse primers, 8 μL water, and 1 μL template cDNA in a total volume of 20 μL. Cycling was performed using the default conditions of the ABI 7300 SDS Software 1.3.1. The relative expression of each gene was normalized against 18S rRNA. The primers used are shown in supplementary material Table S1. The data are presented as the mean ± s.d.

siRNA Transfection

hFHPCs were plated at 2×10^4 cells/cm^2 in antibiotic-free basal medium 24 hours prior to transfection. siRNA transfection was performed following the manufacturer's protocol. Briefly, ON-TARGET SMARTpool siRNAs directed against *PTK7* (L-003167-00-0005, Dharmacon, Lafayette) or *PARD3* (L-015602-00-0005) or non-targeting siRNAs (D-001810-10-05) were mixed with Transfection DharmaFECT 4 (Dharmacon), respectively. After a 20-min incubation at room temperature, the complexes were added to the cells at a final siRNA concentration of 50 nM. The medium was replenished with medium containing antibiotic 24 hours post-transfection. Culture medium was then changed every 3 days for the duration of the experiment. hFHPCs transfected with non-targeting siRNAs were used as experimental control.

Statistics

Experiments were independently repeated three times with replicate samples. Statistical analysis was performed using one-way ANOVA tests using SPSS 11.5 statistics software. Differences were considered significant if $P<0.05$.

Acknowledgements

Microarray experiments were performed by KangChen Bio-tech, Shanghai, China.

Funding

This work was supported by grants from the National Natural Science Foundation of China [grant numbers 30971472, 31171310 to H. Z. and 81030009 to Liying Li]; and the Beijing Natural Science Foundation [grant number 5102013 to H. Z.].

Supplementary material available online at
http://jcs.biologists.org/lookup/suppl/doi:10.1242/jcs.110551/-/DC1

References

Achilleos, A., Wehman, A. M. and Nance, J. (2010). PAR-3 mediates the initial clustering and apical localization of junction and polarity proteins during *C. elegans* intestinal epithelial cell polarization. *Development* **137**, 1833-1842.

Bonora-Centelles, A., Jover, R., Mirabet, V., Lahoz, A., Carbonell, F., Castell, J. V. and Gómez-Lechón, M. J. (2009). Sequential hepatogenic transdifferentiation of adipose tissue-derived stem cells: relevance of different extracellular signaling molecules, transcription factors involved, and expression of new key marker genes. *Cell Transplant* **18**, 1319-1340.

Braiterman, L. T., Heffernan, S., Nyasae, L., Johns, D., See, A. P., Yutzy, R., McNickle, A., Herman, M., Sharma, A., Naik, U. P. et al. (2007). JAM-A is both essential and inhibitory to development of hepatic polarity in WIF-B cells. *Am. J. Physiol. Gastrointest. Liver. Physiol.* **294**, G576-G588.

Chiao, E., Elazar, M., Xing, Y., Xiong, A., Kmet, M., Millan, M. T., Glenn, J. S., Wong, W. H. and Baker, J. (2008). Isolation and transcriptional profiling of purified hepatic cells derived from human embryonic stem cells. *Stem Cells* **26**, 2032-2041.

Decaens, C., Durand, M., Grosse, B. and Cassio, D. (2008). Which *in vitro* models could be best used to study hepatocyte polarity? *Biol. Cell* **100**, 387-398.

Donato, M. T., Gómez-Lechón, M. J. and Castell, J. V. (1993). A microassay for measuring cytochrome P450IA1 and P450IIB1 activities in intact human and rat hepatocytes cultured on 96-well plates. *Anal. Biochem.* **213**, 29-33.

Duncan, S. A. (2003). Mechanisms controlling early development of the liver. *Mech. Dev.* **120**, 19-33.

Eder, A. M., Sui, X., Rosen, D. G., Nolden, L. K., Cheng, K. W., Lahad, J. P., Kango-Singh, M., Lu, K. H., Warneke, C. L., Atkinson, E. N. et al. (2005). Atypical PKCiota contributes to poor prognosis through loss of apical-basal polarity and cyclin E overexpression in ovarian cancer. *Proc. Natl. Acad. Sci. USA* **102**, 12519-12524.

Fu, D., Wakabayashi, Y., Lippincott-Schwartz, J. and Arias, I. M. (2011). Bile acid stimulates hepatocyte polarization through a cAMP-Epac-MEK-LKB1-AMPK pathway. *Proc. Natl. Acad. Sci. USA* **108**, 1403-1408.

Harris, K. P. and Tepass, U. (2010). Cdc42 and vesicle trafficking in polarized cells. *Traffic* **11**, 1272-1279.

Huang, da. W., Sherman, B. T. and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44-57.

Hulsen, T., de Vlieg, J. and Alkema, W. (2008). BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488-494.

Imamura, M., Kojima, T., Lan, M., Son, S., Murata, M., Osanai, M., Chiba, H., Hirata, K. and Sawada, N. (2007). Oncostatin M induces upregulation of claudin-2 in rodent hepatocytes coinciding with changes in morphology and function of tight junctions. *Exp. Cell Res.* **313**, 1951-1962.

Inada, M., Follenzi, A., Cheng, K., Surana, M., Joseph, B., Benten, D., Bandi, S., Qian, H. and Gupta, S. (2008). Phenotype reversion in fetal human liver epithelial cells identifies the role of an intermediate meso-endodermal stage before hepatic maturation. *J. Cell Sci.* **121**, 1002-1013.

Ishiuchi, T. and Takeichi, M. (2011). Willin and Par3 cooperatively regulate epithelial apical constriction through aPKC-mediated ROCK phosphorylation. *Nat. Cell Biol.* **13**, 860-866.

Jozefczuk, J., Prigione, A., Chavez, L. and Adjaye, J. (2011). Comparative analysis of human embryonic stem cell and induced pluripotent stem cell-derived hepatocyte-like cells reveals current drawbacks and possible strategies for improved differentiation. *Stem Cells Dev.* **20**, 1259-1275.

Kamiya, A., Kinoshita, T. and Miyajima, A. (2001). Oncostatin M and hepatocyte growth factor induce hepatic maturation via distinct signaling pathways. *FEBS Lett.* **492**, 90-94.

Kidambi, S., Yarmush, R. S., Novik, E., Chao, P., Yarmush, M. L. and Nahmias, Y. (2009). Oxygen-mediated enhancement of primary hepatocyte metabolism, functional polarization, gene expression, and drug clearance. *Proc. Natl. Acad. Sci. USA* **106**, 15714-15719.

Kinoshita, T., Sekiguchi, T., Xu, M. J., Ito, Y., Kamiya, A., Tsuji, K., Nakahata, T. and Miyajima, A. (1999). Hepatic differentiation induced by oncostatin M attenuates fetal liver hematopoiesis. *Proc. Natl. Acad. Sci. USA* **96**, 7265-7270.

König, J., Cui, Y., Nies, A. T. and Keppler, D. (2000). Localization and genomic organization of a new hepatocellular organic anion transporting polypeptide. *J. Biol. Chem.* **275**, 23161-23168.

Kost, T. A., Condreay, J. P. and Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat. Biotechnol.* **23**, 567-575.

Lemaigre, F. P. (2009). Mechanisms of liver development: concepts for understanding liver disorders and design of novel therapies. *Gastroenterology* **137**, 62-79.

Lemaigre, F. and Zaret, K. S. (2004). Liver development update: new embryo models, cell lineage control, and morphogenesis. *Curr. Opin. Genet. Dev.* **14**, 582-590.

Lhoumeau, A. C., Puppo, F., Prébet, T., Kodjabachian, L. and Borg, J. P. (2011). PTK7: a cell polarity receptor with multiple facets. *Cell Cycle* **10**, 1233-1236.

Matsui, T., Kinoshita, T., Morikawa, Y., Tohya, K., Katsuki, M., Ito, Y., Kamiya, A. and Miyajima, A. (2002). K-Ras mediates cytokine-induced formation of E-cadherin-based adherens junctions during liver development. *EMBO J.* **21**, 1021-1030.

Michalopoulos, G. K., Bowen, W. C., Mulè, K. and Luo, J. (2003). HGF-, EGF-, and dexamethasone-induced gene expression patterns during formation of tissue in hepatic organoid cultures. *Gene Expr.* **11**, 55-75.

Miller, P. M., Folkmann, A. W., Maia, A. R., Efimova, N., Efimov, A. and Kaverina, I. (2009). Golgi-derived CLASP-dependent microtubules control Golgi organization and polarized trafficking in motile cells. *Nat. Cell Biol.* **11**, 1069-1080.

Oikawa, T., Kamiya, A., Kakinuma, S., Zeniya, M., Nishinakamura, R., Tajiri, H. and Nakauchi, H. (2009). Sall4 regulates cell fate decision in fetal hepatic stem/progenitor cells. *Gastroenterology* **136**, 1000-1011.

Ooshio, T., Fujita, N., Yamada, A., Sato, T., Kitagawa, Y., Okamoto, R., Nakata, S., Miki, A., Irie, K. and Takai, Y. (2007). Cooperative roles of Par-3 and afadin in the formation of adherens and tight junctions. *J. Cell Sci.* **120**, 2352-2365.

Parent, R. and Beretta, L. (2008). Translational control plays a prominent role in the hepatocytic differentiation of HepaRG liver progenitor cells. *Genome Biol.* **9**, R19.

Paris, L., Tonutti, L., Vannini, C. and Bazzoni, G. (2008). Structural organization of the tight junctions. *Biochim. Biophys. Acta* **1778**, 646-659.

Pollack, A. L., Apodaca, G. and Mostov, K. E. (2004). Hepatocyte growth factor induces MDCK cell morphogenesis without causing loss of tight junction functional integrity. *Am. J. Physiol. Cell Physiol.* **286**, C482-C494.

Puppo, F., Thomé, V., Lhoumeau, A. C., Cibois, M., Gangar, A., Lembo, F., Belotti, E., Marchetto, S., Lécine, P., Prébet, T. et al. (2011). Protein tyrosine kinase 7 has a conserved role in Wnt/β-catenin canonical signalling. *EMBO Rep.* **12**, 43-49.

Röttger, S., White, J., Wandall, H. H., Olivo, J. C., Stark, A., Bennett, E. P., Whitehouse, C., Berger, E. G., Clausen, H. and Nilsson, T. (1998). Localization of three human polypeptide GalNAc-transferases in HeLa cells suggests initiation of O-linked glycosylation throughout the Golgi apparatus. *J. Cell Sci.* **111**, 45-60.

Saulnier, N., Piscaglia, A. C., Puglisi, M. A., Barba, M., Arena, V., Pani, G., Alfieri, S. and Gasbarrini, A. (2010). Molecular mechanisms underlying human adipose tissue-derived stromal cells differentiation into a hepatocyte-like phenotype. *Dig. Liver Dis.* **42**, 895-901.

Schmelzer, E., Wauthier, E. and Reid, L. M. (2006). The phenotypes of pluripotent human hepatic progenitors. *Stem Cells* **24**, 1852-1858.

Synnergren, J., Heins, N., Brolén, G., Eriksson, G., Lindahl, A., Hyllner, J., Olsson, B., Sartipy, P. and Björquist, P. (2010). Transcriptional profiling of human embryonic stem cells differentiating to definitive and primitive endoderm and further toward the hepatic lineage. *Stem Cells Dev.* **19**, 961-978.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P. et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **39**, D561-D568.

Takashi, H., Katsumi, M. and Toshihiro, A. (2007). Hepatocytes maintain their function on basement membrane formed by epithelial cells. *Biochem. Biophys. Res. Commun.* **359**, 151-156.

Turncliff, R. Z., Tian, X. and Brouwer, K. L. (2006). Effect of culture conditions on the expression and function of Bsep, Mrp2, and Mdr1a/b in sandwich-cultured rat hepatocytes. *Biochem. Pharmacol.* **71**, 1520-1529.

Wang, L. and Boyer, J. L. (2004). The maintenance and generation of membrane polarity in hepatocytes. *Hepatology* **39**, 892-899.

Wang, P., Zhang, H., Li, W., Zhao, Y. and An, W. (2008). Promoter-defined isolation and identification of hepatic progenitor cells from the human fetal liver. *Histochem. Cell Biol.* **130**, 375-385.

Yanai, H., Nakamura, K., Hijioka, S., Kamei, A., Ikari, T., Ishikawa, Y., Shinozaki, E., Mizunuma, N., Hatake, K. and Miyajima, A. (2010). Dlk-1, a cell surface antigen on foetal hepatic stem/progenitor cells, is expressed in hepatocellular, colon, pancreas and breast carcinomas at a high frequency. *J. Biochem.* **148**, 85-92.

Zamek-Gliszczynski, M. J., Xiong, H., Patel, N. J., Turncliff, R. Z., Pollack, G. M. and Brouwer, K. L. (2003). Pharmacokinetics of 5 (and 6)-carboxy-2′,7′-dichlorofluorescein and its diacetate promoiety in the liver. *J. Pharmacol. Exp. Ther.* **304**, 801-809.

Zhang, W., Li, W., Liu, B., Wang, P., Li, W. and Zhang, H. (2012). Efficient generation of functional hepatocyte-like cells from human fetal hepatic progenitor cells *in vitro*. *J. Cell. Physiol.* **227**, 2051-2058.

Zorn, A. M. (2008). Liver development. In *StemBook* (ed. The Stem Cell Research Community, StemBook). pp. doi/10.3824/stembook.1.25.1. http://www.stembook.org/node/512

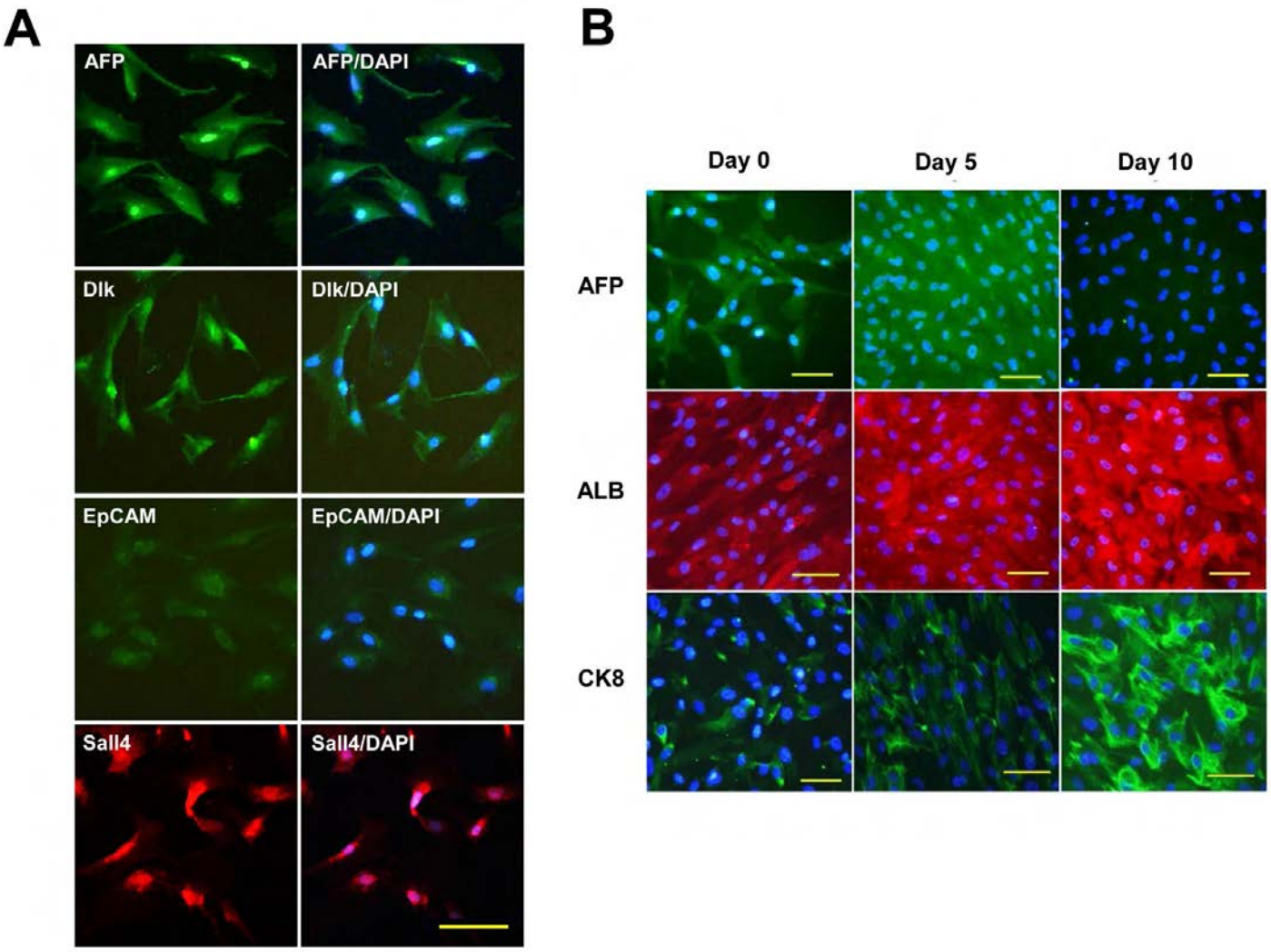

Fig. S1. Properties of hFHPCs and their progeny. (A) Expression of AFP, Dlk, EpCAM and Sall4 detected by immunofluorescence staining of the undifferentiated hFHPCs. (B)Expression of AFP, ALB and CK8 detected by immunofluorescence staining of the hFHPCs in Day 0 (undifferentiated hFHPCs), Day 5 (immature HLCs) and Day 10 (mature HLCs) are presented. Nuclei were counter-stained by DAPI. Scale Bars: 50 μm.

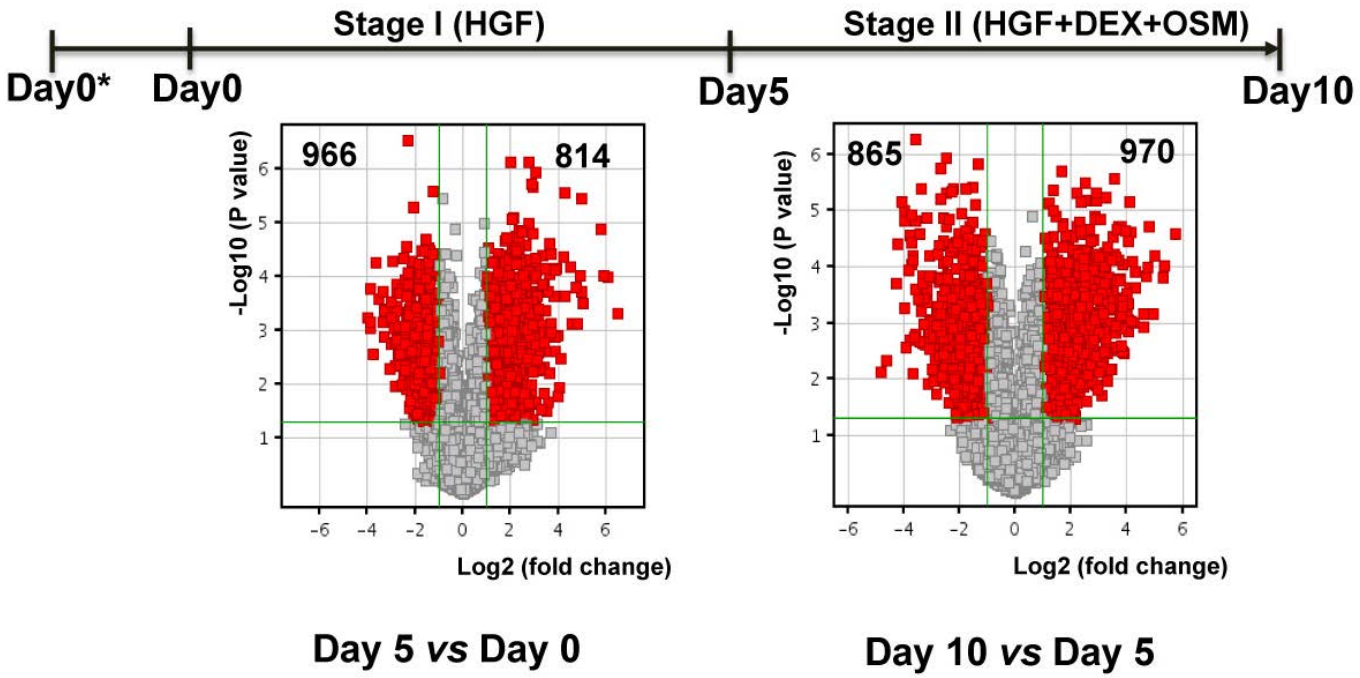

Fig. S2. Differential gene expression analysis during the hFHPCs differentiation. A Volcano Plot filtering identified two major transitions in the gene expression patterns: undifferentiation hFHPCs to immature HLCs (stage I) and immature HLCs to mature HLCs (stage II). Results revealed that 814 genes were up-regulated and 966 genes were down-regulated in stage I; 970 genes were up-regulated and 865 genes were down-regulated in Stage II. The horizontal lines correspond to $\log_2$ plot of the fold changes up and down, respectively, and the vertical line represents a $\log_{10}$ plot of the p-value. The red points in the plot represent the differentially expressed genes with statistically significance. Day 0*, hFHPCs were plated at 2×10^4 cells/cm^2 in expansion medium; Day 0, cells reached 95% confluence.

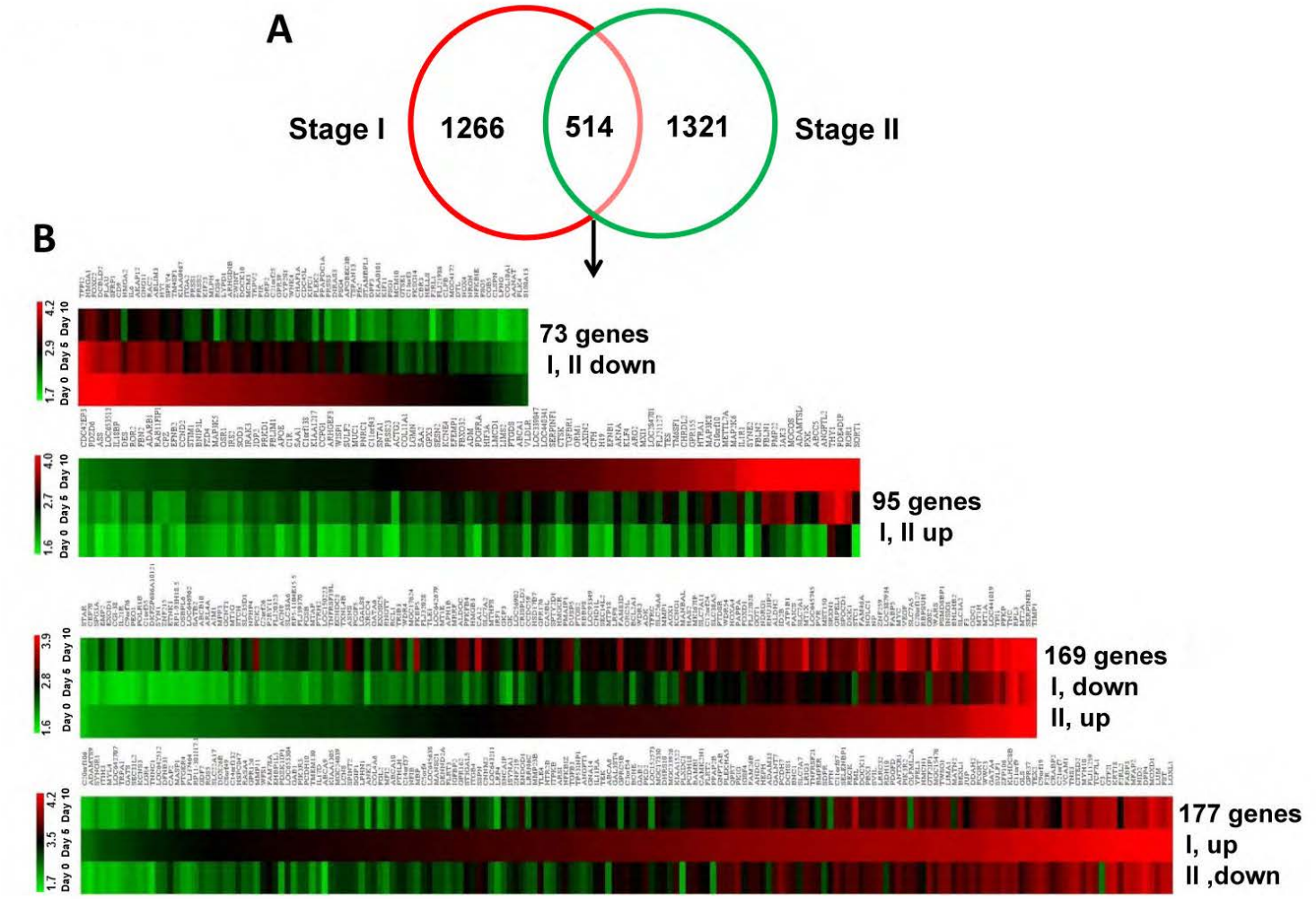

Fig. S3. Comparative analysis of differential expression genes during the hFHPCs differentiation. (A) Venn diagrams showing the number of genes that are shared between or are unique to stage I (Day 5 *vs* Day 0) and stage II (Day 10 *vs* Day 5). (B) Heatmap generated by the common genes shared between two stages based on different gene subsets using MeV v4.8 (http://www.tm4.org/mev/). In the heat map, high expression is depicted in red and low expression in green. The columns are $\log_{10}$ plot of the normalized intensity of triplicate samples for each group.

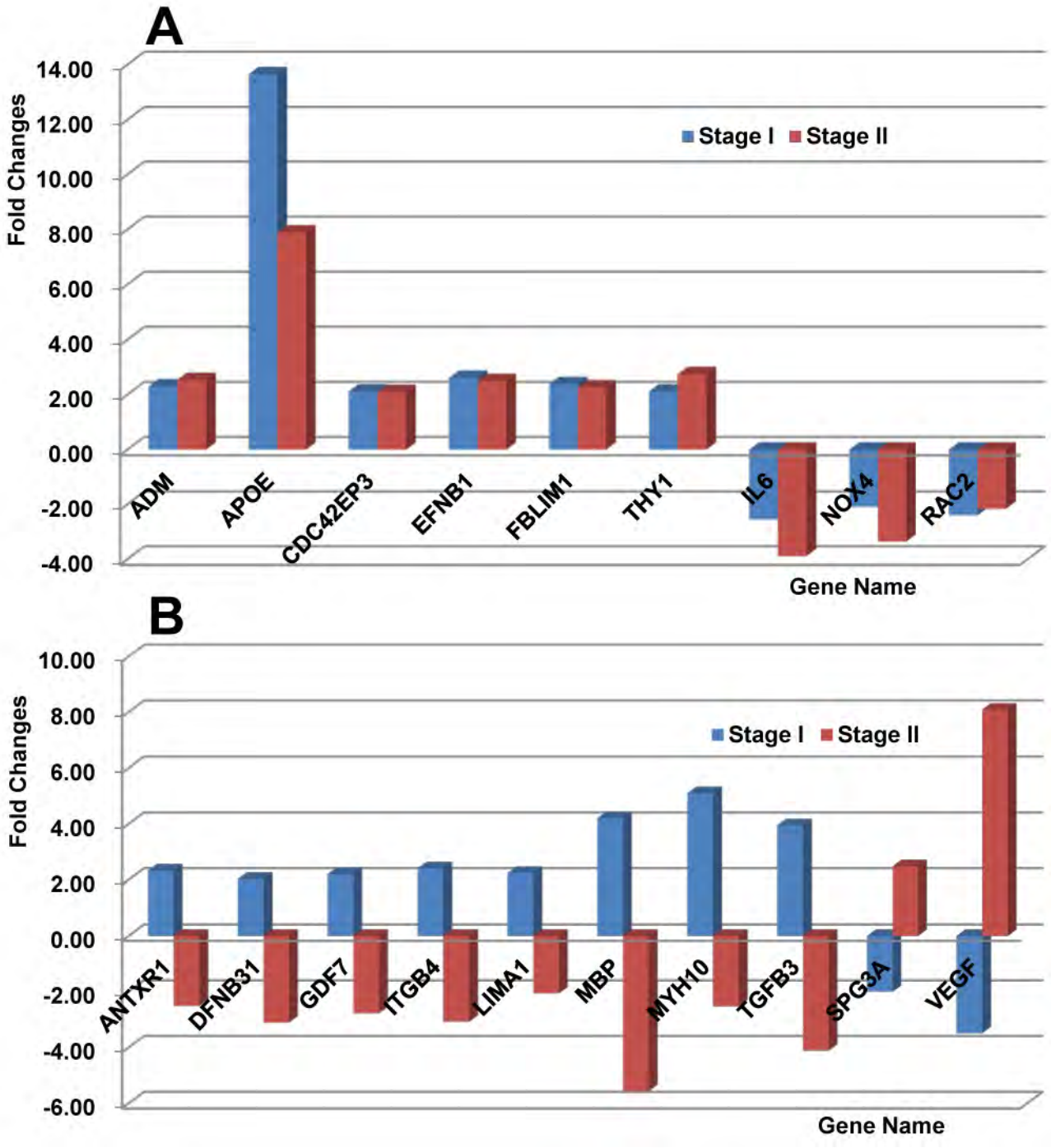

Fig. S4. Changes on the expression of the common genes shared by two differentiated stages involved in cell morphogenesis. (A) The selected genes were regulated consistently in stage I and stage II. (B) The selected genes were regulated inconsistently in stage I and stage II. The columns are fold changes for each gene. Expression in stage I is depicted in blue and stage II in red.

Table S1. Primers for Real-Time RT-PCR

Accession number	Name	5'-Sequence-3'
NM_005502.3	ABCA1	F: CCCTGTGGAATGTACCTATGTG R: GAGGTGTCCCAAAGATGCAA
NM_000392.3	ABCC2	F: AGCAGGTATTCGTTGGTT R: AGGTAGAGGCTTGGATTG
NM_001645.3	APOC1	F: GGTCCTGGTGGTGGTTCT R: TGTTTGATGCGGCTGATG
NM_000041.2	APOE	F: GGTCGCTTTTGGGATTACCT R: CTCAGTTCCTGGGTGACCTG
NM_044472.2	CDC42	F: CTTTCTTGCTTGTTGGGACT R: TAGGCTTCTGTTTGTTCTTGG
NM_001875.4	CPS1	F: CTGACCCTGCCTACAAAG R: CACCAGCAAACCTGAAAC
NM_000761.3	CYP1A2	F: AGTCTGTTCCCTTCTCGG R: TGGCTCTGGTGGACTTTT
NM_001202855.2	CYP3A4	F: GGCGGATGTTGAAGTGAG R: GTTGGGTGTTGAGGATGG
NM_006169	NNMT	F: GAGATCGTCGTCACTGAC R: ATCACACTTCAGCACCTG
NM_001184794.1	PARD3	F: TGAGCCTTCTGGTCTTTC G R: TTCCCAAATCTGCGTGGT
NM_001018073.1	PCK2	F: TTCCCCACCGCACATACC R: CCACCACCAATCCCAACG
NM_002821	PTK7	F: TGGTAGTAGCGAGGTATGAGG R: TGCGGTTAGTGATGGGAGT
NM_003049.3	SLC10A1	F: GCTTTCTGCTGGGTTATGTTC R: CATCCAGTCTCCATGCTGACA
NM_003060.3	SLC22A5	F: GTGAGGACGACTGGAAGG R: CAAGGACAA ACAGCACGAC
	18S	F: GTAACCCGTTGA ACCCCATT R: CCATCCAATCGGTAGTAGCG

Table S2 The identified overlapping significant genes shared by two stages

S2-1 95 genes up-regulated in two stages

SEQ ID	GENE NAME	DESCRIPTION
NM_001854	COL11A1	spastic ataxia of Charlevoix-Saguenay (sacsin)
NM_001014447	CPZ	similar to CG31151-PB, isoform B
NM_003102	SOD3	hydroxysteroid (17-beta) dehydrogenase 7
BC053636	H19	smoothened homolog (Drosophila)
NM_019032	ADAMTSL4	interleukin enhancer binding factor 3, 90kDa
NM_000050	ASS	tumor protein p53 inducible nuclear protein 1
NM_006813	PNRC1	ribosomal protein L22-like 1
NM_015641	TES	G protein-coupled receptor 39
NM_000041	APOE	promyelocytic leukemia
BC054503	MAP3K5	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 2
NM_001733	C1R	ring finger and WD repeat domain 3
NM_002615	SERPINF1	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform
BC004492	METTL7A	glutamate-cysteine ligase, modifier subunit
NM_000954	PTGDS	potassium voltage-gated channel, Shal-related subfamily, member 3
NM_000186	CFH	ATP-binding cassette, sub-family A (ABC1), member 6
NM_001927	DES	pannexin 2
BC011786	C11orf43	calcium regulated heat stable protein 1, 24kDa
NM_017980	LIMS2	wingless-type MMTV integration site family, member 5A
NM_017947	MOCOS	uridine phosphorylase 1
NM_002084	GPX3	zinc finger protein 195
NM_006206	PDGFRA	high mobility group AT-hook 1
NM_002959	SORT1	thioredoxin-like 5
NM_000304	PMP22	decorin
NM_001039348	EFEMP1	hypothetical gene supported by BC062774
AL833280	KIAA1217	EGF-containing fibulin-like extracellular matrix protein 2
NM_005012	ROR1	hypothetical protein HSPC111
NM_001018056	VLDLR	similar to phosphodiesterase 4D interacting protein isoform 2
NM_001002233	RAB11FIP1	juxtaposed with another zinc finger gene 1
NM_018837	SULF2	Ras and Rab interactor 1
NM_130469	JDP2	thioredoxin-like 4B
NM_000877	IL1R1	NOD9 protein
NM_014583	LMCD1	TAF9 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 32kDa

NM_001999	FBN2	D4, zinc and double PHD fingers, family 3
NM_058229	FBXO32	coenzyme Q3 homolog, methyltransferase (S. cerevisiae)
NM_006288	THY1	early growth response 1
NM_004748	CCPG1	mixed lineage kinase domain-like
NM_030754	SAA2	extra spindle poles like 1 (S. cerevisiae)
NM_001018016	MUC1	similar to Complement C1r subcomponent precursor (Complement component 1, r subcomponent)
BC012907	MXI1	Homo sapiens mRNA; cDNA DKFZp686O2421 (from clone DKFZp686O2421).
NM_001025493	C1orf138	complement component 1, r subcomponent
NM_199161	SAA1	isopentenyl-diphosphate delta isomerase 2
NM_005502	ABCA1	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)
NM_001039659	IL18BP	plasminogen-like A1
NM_001124	ADM	calcium channel, voltage-dependent, L type, alpha 1C subunit
NM_007021	C10orf10	vaccinia related kinase 1
NM_002742	PRKD1	adenosine kinase
BC019270	CDC42EP3	FLJ41170 protein
AY279092	CHRDL2	hypothetical protein FLJ11259
NM_004093	EFNB2	tyrosyl-DNA phosphodiesterase 1
NM_002775	HTRA1	family with sequence similarity 64, member A
NM_001615	ACTG2	coagulation factor III (thromboplastin, tissue factor)
CR749324	SYNE2	anthrax toxin receptor 1
XM_927826	LOC653513	apoptosis-inducing, TAF9-like domain 1
NM_001004019	FBLN2	hypothetical protein FLJ33706
NM_001002810	PDE4DIP	opposite strand transcription unit to STAG3
NM_080671	KCNE4	hypothetical protein FLJ20186
NM_007173	PRSS23	arginine-rich, mutated in early stage tumors
NM_003098	SNTA1	metallothionein 2A
BC099715	ARHGEF3	family with sequence similarity 26, member B
NM_005204	MAP3K8	asp (abnormal spindle)-like, microcephaly associated (Drosophila)
NM_003156	STIM1	chromosome 22 open reading frame 18
NM_012098	ANGPTL2	full-length cDNA clone CS0DI043YJ07 of Placenta Cot 25-normalized of Homo sapiens (human).
NM_003882	WISP1	aurora kinase B
BC001559	BNIP3L	EH-domain containing 3
NM_004612	TGFBR1	IMP1 inner mitochondrial membrane

		peptidase-like (S. cerevisiae)
NM_022462	HIF3A	S100 calcium binding protein A13
BC019895	FBLIM1	integrin, alpha 11
NM_145260	OSR1	transmembrane, prostate androgen induced RNA
NM_001759	CCND2	chromosome 16 open reading frame 60
BC028068	JAK3	adenosine deaminase, RNA-specific, B1 (RED1 homolog rat)
NM_015693	PDZD6	carbonyl reductase 1
NM_012193	FZD4	hypothetical protein LOC150223
AK056381	GPR155	Homo sapiens cDNA FLJ46862 fis, clone UTERU3011398, highly similar to Collagen alpha 2(VI) chain precursor.
XM_932036	LOC440341	hypothetical LOC401884
NM_001112	ADARB1	similar to tousled-like kinase 2
NM_002758	MAP2K6	kinesin family member 22
NM_000607	ORM1	hypothetical protein MGC26963
NM_017771	PXK	stanniocalcin 1
NM_001008530	LGMN	N-acyl-phosphatidylethanolamine-hydrolyzing phospholipase D
BC032007	TM6SF1	carboxypeptidase A1 (pancreatic)
NM_030767	AKNA	kinesin family member 4A
NM_003749	IRS2	chromosome 14 open reading frame 149
BC033697	ROR2	inositol 1,4,5-trisphosphate 3-kinase B
BC008464	ARG2	RAB40A, member RAS oncogene family
NM_000396	CTSK	peptidylglycine alpha-amidating monooxygenase
BC030811	KLF4	cyclin D2
BC101533	AXIN2	G protein-coupled receptor, family C, group 5, member B
NM_007199	IRAK3	elongation factor Tu GTP binding domain containing 1
BC046145	LOC339047	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)
XM_934034	LOC284701	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 12 (CHL1-like helicase homolog, S. cerevisiae)
BC052979	EFNB1	lectin, galactoside-binding, soluble, 3 binding protein
AK024780	FLJ21127	ADAM metallopeptidase domain 33
NM_031459	SESN2	mitochondrial ribosomal protein L42
AY196484	ABCC5	dysbindin (dystrobrevin binding protein 1) domain containing 2
NM_006485	FBLN1	hypothetical protein FLJ14834

S2-2 73 genes down-regulated in two stages

SEQ_ID	GENE_NAME	DESCRIPTION
NM_002203	ITGA2	ribosomal protein L3
NM_016426	GTSE1	v-myc myelocytomatosis viral oncogene homolog (avian)
NM_004675	DIRAS3	CDC6 cell division cycle 6 homolog (S. cerevisiae)
BC030538	FLJ21986	basonuclin 1
NM_004900	APOBEC3B	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
NM_003504	CDC45L	epithelial membrane protein 2
BC006257	CLPB	fibulin 2
NM_005251	FOXC2	early growth response 2 (Krox-20 homolog, Drosophila)
NM_024308	MGC4172	chaperonin containing TCP1, subunit 2 (beta)
NM_001015886	HMGA2	hypothetical protein LOC283970
NM_080927	DCBLD2	chromosome 9 open reading frame 6
NM_014689	DOCK10	hypothetical LOC550631
NM_001236	CBR3	chaperonin containing TCP1, subunit 6A (zeta 1)
NM_004523	KIF11	popeye domain containing 3
NM_016445	PLEK2	origin recognition complex, subunit 5-like (yeast)
NM_006528	TFPI2	pituitary tumor-transforming 1
NM_020799	STAMBPL1	GLI-Kruppel family member GLI2
BC001665	ABLIM3	KIAA1602
NM_002771	PRSS3	disrupter of silencing 10
NM_018063	HELLS	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5
NM_001005413	ZWINT	PH domain and leucine rich repeat protein phosphatase
NM_001029989	KIAA0101	hypothetical protein LOC643591
NM_002388	MCM3	TIMP metallopeptidase inhibitor 1
NM_000493	COL10A1	hypothetical protein MGC24039
NM_016113	TRPV2	high-mobility group box 2
NM_138555	KIF23	nudix (nucleoside diphosphate linked moiety X)-type motif 15
BC006140	HYI	phytanoyl-CoA dioxygenase domain containing 1
NM_001018109	PIR	similar to Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keraton-8) (K8)
NM_016095	Pfs2	H2A histone family, member Z
NM_030964	SPRY4	hyaluronan-mediated motility receptor (RHAMM)
NM_001769	CD9	chromosome 10 open reading frame 18
NM_001088	AANAT	chromosome 3 open reading frame 55
BC103998	PRSS1	cyclin M3

NM_145061	C13orf3	RAD18 homolog (S. cerevisiae)
AK126972	KIAA0467	F11 receptor
NM_000600	IL6	hypothetical protein MGC33926
NM_002781	PSG5	pleckstrin homology-like domain, family A, member 2
BC000737	RGS4	NIMA (never in mitosis gene a)- related kinase 11
BC009108	MCM10	calcyphosine
NM_004126	GNG11	ubiquitin-conjugating enzyme E2C
NM_005100	AKAP12	ataxin 3
NM_002658	PLAU	solute carrier family 19 (folate transporter), member 1
BC040046	LYPD1	stromal interaction molecule 1
NM_002770	PRSS2	CD44 molecule (Indian blood group)
NM_002263	KIFC1	lin-9 homolog (C. elegans)
NM_032387	WNK4	BTB (POZ) domain containing 3
NM_022145	FKSG14	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)
NM_024101	MLPH	chaperonin containing TCP1, subunit 4 (delta)
AK098668	PPAPDC1A	ectodysplasin A2 receptor
NM_003012	SFRP1	cyclin B1
NM_199050	C21orf25	polyribonucleotide nucleotidyltransferase 1
NM_030622	CYP2S1	TPX2, microtubule-associated, homolog (Xenopus laevis)
NM_005242	F2RL1	UL16 binding protein 2
NM_001939	DRP2	nucleoside phosphorylase
NM_012074	DPF3	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3D (putative)
NM_005483	CHAF1A	aryl hydrocarbon receptor nuclear translocator-like
U91616	NFKBIE	retinol dehydrogenase 13 (all-trans and 9-cis)
NM_001508	GPR39	solute carrier family 26, member 6
AK027651	DTL	ring finger protein 44
NM_014220	TM4SF1	ZW10 interactor
BC040105	NOX4	supervillin
NM_014399	TSPAN13	baculoviral IAP repeat-containing 5 (survivin)
NM_002304	LFNG	replication factor C (activator 1) 3, 38kDa
NM_002872	RAC2	thioredoxin-like 2
NM_014264	PLK4	thymine-DNA glycosylase
NM_213633	PSG4	chromosome 20 open reading frame 127
AK075079	PSG1	RAB GTPase activating protein 1
NM_001024211	S100A13	similar to Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keraton-8) (K8)
NM_002131	HMGA1	Sequence 201 from Patent WO03048202.
BC103970	CGB5	CDV3 homolog (mouse)

NM_022111	CLSPN	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)
NM_001175	ARHGDIB	SEC14-like 2 (S. cerevisiae)
NM_006176	NRGN	thioredoxin domain containing 9

S2-3 177 genes up-regulated in stage I but down-regulated in stage II

SEQ_ID	GENE_NAME	DESCRIPTION
NM_080282	ABCA10	mitochondrial ribosomal protein L39
NM_080284	ABCA6	family with sequence similarity 29, member A
NM_080283	ABCA9	downstream neighbor of SON
NM_025220	ADAM33	breast cancer 1, early onset
NM_182920	ADAMTS9	chromosome 3 open reading frame 26
NM_001029882	AHDC1	PREDICTED: Homo sapiens neuroblastoma breakpoint family, member 10, transcript variant 76 (NBPF10), mRNA.
NM_001146	ANGPT1	protein tyrosine phosphatase, receptor type, K
NM_001149	ANK3	heat shock 105kDa/110kDa protein 1
NM_018153	ANTXR1	RuvB-like 1 (E. coli)
NM_001012301	ARSI	mucin 1, cell surface associated
NM_012342	BAMBI	collagen, type VII, alpha 1 (epidermolysis bullosa, dystrophic, dominant and recessive)
NM_000055	BCHE	hypothetical protein MGC4655
AK000959	BEXL1	matrix-remodelling associated 8
BC060868	BMPER	bone morphogenetic protein 1
NM_001717	BNC1	male sterility domain containing 1
NM_024684	C11orf67	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18
NM_013279	C11orf9	interleukin 8
NM_020215	C14orf132	leupaxin
NM_001001872	C14orf37	kelch domain containing 8B
NM_001012971	C20orf106	endonuclease domain containing 1
NM_020152	C21orf7	DIRAS family, GTP-binding RAS-like 3
NM_000064	C3	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5
NM_203370	C3orf54	G protein-regulated inducer of neurite outgrowth 1
NM_016348	C5orf4	leukemia inhibitory factor (cholinergic differentiation factor)
NM_001031839	C8orf49	similar to nuclear pore complex interacting protein
NM_022343	C9orf19	mitochondrial intermediate peptidase

NM_018584	CAMK2N1	pituitary tumor-transforming 1
BC008481	CAP2	keratin associated protein 9-3
NM_006079	CITED2	zinc finger protein 313
AK023479	CNNM2	Homo sapiens cDNA FLJ42850 fis, clone BRHIP2005600.
NM_001847	COL4A6	pregnancy specific beta-1-glycoprotein 1
NM_001878	CRABP2	transmembrane protein 48
NM_003737	DCHS1	hypothetical protein FLJ20850
NM_013974	DDAH2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
NM_182540	DDX26B	mitochondrial ribosomal protein S5
NM_015689	DENND2A	ATPase, Na+/K+ transporting, beta 1 polypeptide
NM_015404	DFNB31	adenosine kinase
NM_016246	DHRS10	insulin-like growth factor binding protein 6
NM_144658	DOCK11	hypothetical protein FLJ20232
NM_001935	DPP4	Homo sapiens cDNA FLJ39878 fis, clone SPLEN2016045, moderately similar to Homo sapiens mRNA for Hrs.
XM_934333	EHBP1L1	family with sequence similarity 39, member A
XM_290546	ENDOD1	peptidylprolyl isomerase (cyclophilin)-like 5
NM_001992	F2R	GATA zinc finger domain containing 2A
NM_004101	F2RL2	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)
NM_004102	FABP3	lymphocyte adaptor protein
NM_015916	FAM26B	cell division cycle 25A
BC037559	FAM78A	pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)
NM_004107	FCGRT	ATPase, H+ transporting V0 subunit E2-like (rat)
BC033820	FGL2	laminin, alpha 4
NM_018370	FLJ11259	flap structure-specific endonuclease 1
NM_173815	FLJ37464	defective in sister chromatid cohesion homolog 1 (S. cerevisiae)
NM_013231	FLRT2	nucleoside phosphorylase
BC020870	FLRT3	granulin
NM_006350	FST	MCM8 minichromosome maintenance deficient 8 (S. cerevisiae)
NM_002039	GAB1	WD repeat domain 69
NM_012296	GAB2	PAS domain containing serine/threonine kinase

NM_024637	GAL3ST4	heat shock 105kDa/110kDa protein 1
NM_017423	GALNT7	proteasome (prosome, macropain) subunit, alpha type, 3
NM_002052	GATA4	Homo sapiens cDNA FLJ46181 fis, clone TESTI4004210.
NM_178831	GATS	transmembrane protein 22
NM_182828	GDF7	calcium/calmodulin-dependent protein kinase kinase 1, alpha
NM_014905	GLS	chromosome 4 open reading frame 13
NM_004297	GNA14	F-box protein 5
NM_024312	GNPTAB	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)
NM_001032394	GPR126	hypothetical protein LOC339047
NM_014449	GPR162	mitochondrial ribosomal protein L22
NM_005302	GPR37	polo-like kinase 4 (Drosophila)
NM_016235	GPRC5B	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
NM_001518	GTF2I	centrosomal protein 70kDa
XM_932298	GTF2IP1	similar to polymerase (RNA) III (DNA directed) polypeptide K
NM_014799	HEPH	gephyrin
NM_031935	HMCN1	eukaryotic translation initiation factor 2B, subunit 2 beta, 39kDa
BC035371	HSPC047	myosin, light polypeptide 4, alkali; atrial, embryonic
NM_000867	HTR2B	G protein-coupled receptor 161
NM_000597	IGFBP2	solute carrier family 43, member 2
NM_014333	IGSF4	hypothetical protein LOC199953
NM_004512	IL11RA	hypothetical protein FLJ22028
NM_138284	IL17D	chromosome 6 open reading frame 139
NM_016368	ISYNA1	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae)
NM_000213	ITGB4	chromosome 20 open reading frame 175
NM_002217	ITIH3	Homo sapiens PDZ and LIM domain 5 (PDLIM5) mRNA, complete cds.
NM_002221	ITPKB	PDZ domain containing 11
NM_002230	JUP	melanophilin
NM_002252	KCNS3	cytidine deaminase
BC008219	KIAA1305	par-3 partitioning defective 3 homolog (C. elegans)
NM_020888	KIAA1522	high-mobility group box 3
NM_173546	KLHDC8B	spermatogenesis associated 5-like 1
BC000180	KRT18	ligand of numb-protein X 1

AK090647	LCN8	zinc finger protein 106 homolog (mouse)
NM_018490	LGR4	synaptotagmin VII
AB209512	LIMA1	par-3 partitioning defective 3 homolog (C. elegans)
XM_496688	LOC152573	mitochondrial ribosomal protein L52
XM_926008	LOC642512	ubiquitin-conjugating enzyme E2C
XM_926154	LOC642707	zinc finger, AN1-type domain 2A
XM_926575	LOC643211	cAMP responsive element modulator
XM_928653	LOC645638	multidrug resistance-related protein
XM_932024	LOC653304	family with sequence similarity 39, member A
NM_005576	LOXL1	hypothetical protein FLJ14803
NM_001008701	LPHN1	potassium channel regulator
NM_153377	LRIG3	interleukin 21 receptor
NM_002334	LRP4	chromosome 1 open reading frame 135
BC052210	LRRC32	signal transducer and activator of transcription 2, 113kDa
NM_152570	LRRN6C	zinc finger and BTB domain containing 2
NM_002345	LUM	glutamic-oxaloacetic transaminase 1, soluble (aspartate aminotransferase 1)
NM_018050	MANSC1	X-ray repair complementing defective repair in Chinese hamster cells 4
NM_139125	MASP1	PDZ binding kinase
NM_002380	MATN2	immediate early response 3
NM_001025100	MBP	similar to golgi autoantigen, golgin subfamily a-like
NM_002403	MFAP2	splicing factor, arginine/serine-rich 7, 35kDa
NM_033316	MFI2	M-phase phosphoprotein 1
NM_145056	MGC15476	putatative 28 kDa protein
NM_052880	MGC17330	pleckstrin 2
NM_144973	MGC24039	polymerase (RNA) I polypeptide B, 128kDa
NM_152390	MGC33926	disrupter of silencing 10
NM_005940	MMP11	prenyl (decaprenyl) diphosphate synthase, subunit 1
AB010962	MMP23B	membrane associated guanylate kinase, WW and PDZ domain containing 2
NM_001031699	MOXD1	similar to nuclear pore complex interacting protein
NM_005964	MYH10	wingless-type MMTV integration site family, member 5A

NM_002476	MYL4	hematological and neurological expressed 1
NM_005596	NFIB	uridine phosphorylase 1
AK091782	NFKBIZ	Sequence 1 from Patent WO02059284.
NM_007361	NID2	thioredoxin
NM_182487	OLFML2A	glutamate-rich WD repeat containing 1
NM_006195	PBX3	metallothionein 1L
NM_032961	PCDH10	mitochondrial ribosomal protein L2
NM_019035	PCDH18	kinesin family member 11
NM_032457	PCDH7	KIAA0179
NM_025208	PDGFD	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6)
NM_005764	PDZK1IP1	tyrosinase-related protein 1
NM_005027	PIK3R2	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 10
NM_003629	PIK3R3	steroidogenic acute regulator
BC112261	PKD2	KIAA1522
NM_181804	PKIG	cysteine-rich secretory protein LCCL domain containing 2
NM_019012	PLEKHA5	glycine cleavage system protein H (aminomethyl carrier)
AK127539	PLXDC1	ribonuclease P 25kDa subunit
NM_033238	PML	CGI-115 protein
NM_003713	PPAP2B	protein disulfide isomerase family A, member 4
NM_000958	PTGER4	colony stimulating factor 1 (macrophage)
NM_002820	PTHLH	chromosome 1 open reading frame 41
NM_002825	PTN	replication protein A3, 14kDa
NM_014298	QPRT	nucleoporin 35kDa
NM_006989	RASA4	polymerase (RNA) I polypeptide B, 128kDa
NM_172037	RDH10	SRY (sex determining region Y)-box 17
NM_021111	RECK	Kruppel-like factor 10
BC030059	RGS5	protein arginine methyltransferase 2
BC024178	RP11-301I17.1	AT hook, DNA binding motif, containing 1
BC016475	SDPR	aminomethyltransferase (glycine cleavage system protein T)
NM_015490	SEC31L2	cyclin H
NM_003944	SELENBP1	integrin beta 3 binding protein (beta3-endonexin)
NM_032291	SGIP1	KIAA0101

NM_020372	SLC22A17	polo-like kinase 1 (Drosophila)
NM_003982	SLC7A7	translocase of outer mitochondrial membrane 34
NM_005460	SNCAIP	NF-kappaB repressing factor
NM_021102	SPINT2	laminin, gamma 2
BC062299	SSPN	bone morphogenetic protein 1
AY152815	ST3GAL5	transmembrane protein 76
NM_015170	SULF1	ADP-ribosylation factor-like 4A
NM_003174	SVIL	uridine monophosphate synthetase (orotate phosphoribosyl transferase and orotidine-5'-decarboxylase)
NM_004711	SYNGR1	membrane protein, palmitoylated 3 (MAGUK p55 subfamily member 3)
NM_031283	TCF7L1	PSMC3 interacting protein
NM_000459	TEK	nexilin (F actin binding protein)
NM_018469	TEX2	tryptophanyl-tRNA synthetase
NM_003239	TGFB3	Homo sapiens cDNA clone IMAGE:3351130, complete cds.
BC045650	TLE4	stomatin
BC037895	TMEM130	RAB5B, member RAS oncogene family
NM_014452	TNFRSF21	peptidyl-tRNA hydrolase 2
NM_003280	TNNC1	major histocompatibility complex, class II, DP beta 1
NM_022748	TNS3	heparan sulfate (glucosamine) 3-O-sulfotransferase 3A1
NM_033285	TP53INP1	ankyrin repeat and SOCS box-containing 3
NM_007332	TRPA1	hypothetical protein FLJ13611
NM_020648	TWSG1	glutamine-fructose-6-phosphate transaminase 2
NM_001078	VCAM1	hypothetical protein LOC647184
BC030130	WFS1	actin, alpha 2, smooth muscle, aorta
NM_015691	WWC3	core-binding factor, beta subunit
NM_031477	YPEL3	chromosome 6 open reading frame 66
NM_022473	ZFP106	T-box 2
NM_016423	ZNF219	nudix (nucleoside diphosphate linked moiety X)-type motif 1

S2-4 169 genes down-regulated in stage I but up-regulated in stage II

SEQ_ID	GENE_NAME	DESCRIPTION
NM_012089	ABCB10	exosome component 2
NM_020186	ACN9	nucleolar and coiled-body phosphoprotein 1
NM_001123	ADK	phosphoinositide-3-kinase, class 2, beta

		polypeptide
NM_001624	AIM1	metallothionein 1E (functional)
NM_000690	ALDH2	FYN oncogene related to SRC, FGR, YES
NM_005165	ALDOC	ectonucleoside triphosphate diphosphohydrolase 4
NM_001159	AOX1	RNA guanylyltransferase and 5'-phosphatase
AB189172	APH1B	ATP-binding cassette, sub-family A (ABC1), member 1
NM_001037164	ARL4A	similar to Keratin, type II cytoskeletal 8 (Cytokeratin-8) (CK-8) (Keraton-8) (K8)
NM_001673	ASNS	heat shock protein 90kDa alpha (cytosolic), class A member 1
NM_001677	ATP1B1	DEP domain containing 1
NM_004049	BCL2A1	immature colon carcinoma transcript 1
NM_003670	BHLHB2	coagulation factor III (thromboplastin, tissue factor)
BC020967	C12orf24	hypothetical LOC441242
BX648319	C1orf55	FAT tumor suppressor homolog 4 (Drosophila)
NM_080757	C20orf127	coiled-coil domain containing 99
NM_023016	C2orf26	deleted in liver cancer 1
BC064937	C9orf76	A kinase (PRKA) anchor protein 13
NM_001218	CA12	chromosome 21 open reading frame 25
XM_291346	CA5BL	hypothetical protein FLJ39370
NM_014167	CCDC59	TWIST neighbor
AF289495	CEP70	breast carcinoma amplified sequence 3
NM_015964	CGI-38	CTP synthase
NM_004284	CHD1L	gem (nuclear organelle) associated protein 6
NM_031476	CRISPLD2	PSMD8 binding protein 1
NM_033107	DKFZP686A10121	WD repeat domain 50
NM_012242	DKK1	SCL/TAL1 interrupting locus
NM_004419	DUSP5	metallothionein 1G
NM_024693	ECHDC3	WD repeat domain 3
BC009687	EMP2	uronyl-2-sulfotransferase
NM_001039481	ETNK1	hypothetical protein LOC646881
AB040937	EXOD1	synuclein, alpha interacting protein (synphilin)
AF285785	EXOSC5	brain expressed X-linked-like 1
NM_001993	F3	hypothetical protein MGC13005
NM_001444	FABP5	TPX2, microtubule-associated, homolog (Xenopus laevis)

XM_165511	FABP5L3	ubiquitin-like domain containing CTD phosphatase 1
AF350451	FAM46A	pleckstrin homology domain containing, family H (with MyTH4 domain) member 3
NM_030919	FAM83D	SEC14-like 2 (S. cerevisiae)
BC042605	FKBP5	ADAM metallopeptidase domain 15 (metargidin)
NM_019005	FLJ20323	phenylalanine-tRNA synthetase-like, alpha subunit
BC021662	FLJ22028	hypothetical gene supported by AK126863
NM_152680	FLJ32028	phorbol-12-myristate-13-acetate-induced protein 1
NM_152400	FLJ39370	neuron navigator 1
NM_006732	FOSB	protease, serine, 1 (trypsin 1)
NM_004472	FOXD1	T cell receptor alpha locus
NM_015714	G0S2	CD9 molecule
X95701	GATA6	EH-domain containing 4
BC074885	GCNT1	EH-domain containing 3
NM_000167	GK	ATP-binding cassette, sub-family A (ABC1), member 10
BC066960	GKP3	leiomodin 1 (smooth muscle)
NM_002079	GOT1	chromosome 16 open reading frame 59
BC067106	GPR176	RAB GTPase activating protein 1
NM_025196	GRPEL1	citron (rho-interacting, serine/threonine kinase 21)
NM_005328	HAS2	ribosomal protein L3
NM_005342	HMGB3	coiled-coil domain containing 52
NM_000859	HMGCR	podocan
NM_002141	HOXA4	kinesin family member 2C
NM_016371	HSD17B7	matrix metallopeptidase 1 (interstitial collagenase)
XM_929750	HSUP1	bolA-like 3 (E. coli)
NM_001039082	ID2B	KIAA1462
NM_021798	IL21R	teashirt family zinc finger 1
NM_005542	INSIG1	FK506 binding protein 5
NM_002198	IRF1	BCL2-related protein A1
NM_002237	KCNG1	pregnancy-associated plasma protein A, pappalysin 1
AY037304	LGALS8	midkine (neurite growth-promoting factor 2)
CR614195	LOC150223	AT-hook transcription factor
XM_498063	LOC222699	replication factor C (activator 1) 3, 38kDa
XM_370729	LOC387934	cell division cycle associated 2
XM_371261	LOC388642	hypothetical protein FLJ32028

XM_373914	LOC388796	lin-9 homolog (C. elegans)
XM_498969	LOC441019	quiescin Q6-like 1
NM_020143	LOC56902	tetratricopeptide repeat domain 4
XM_926131	LOC642679	ubiquitin-conjugating enzyme E2C
XM_926338	LOC642956	cAMP responsive element modulator
XM_932383	LOC643366	similar to TBP-associated factor 9L
NM_001039954	LOC645745	hypothetical protein LOC647064
NM_001039792	LOC646962	hypothetical protein LOC283970
XM_925841	LOC653068	lines homolog 1 (Drosophila)
NM_138402	LOC93349	KIAA1794
NM_001018054	LRP8	hypothetical protein LOC644230
BC016822	MANBAL	growth differentiation factor 5 (cartilage-derived morphogenetic protein-1)
NM_206967	MGC17624	BRCA2 and CDKN1A interacting protein
NM_032390	MKI67IP	centrosomal protein 152kDa
NM_002421	MMP1	RNA terminal phosphate cyclase-like 1
BC025743	MPP3	Rho-related BTB domain containing 1
BC002814	MRRF	hypothetical protein LOC283130
NM_032947	MST150	CGI-09 protein
NM_005946	MT1A	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 5
NM_175617	MT1E	elongation factor Tu GTP binding domain containing 1
BC020757	MT1G	RNA binding protein with multiple splicing
NM_005951	MT1H	zinc finger protein 415
NM_175622	MT1JP	queuine tRNA-ribosyltransferase domain containing 1
NM_176870	MT1M	SHC SH2-domain binding protein 1
NM_005952	MT1X	family with sequence similarity 107, member B
NM_005953	MT2A	PSMD8 binding protein 1
NM_002451	MTAP	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18
AF249277	MTHFS	development and differentiation enhancing factor-like 1
NM_001003704	MTP18	olfactomedin-like 2A
NM_002467	MYC	uridine phosphorylase 1
AK095649	NDRG1	mucin 1, cell surface associated
NM_004741	NOLC1	kelch-like 26 (Drosophila)
NM_000270	NP	collagen and calcium binding EGF domains 1
AK127333	NPHP4	G protein-coupled receptor 162
NM_002539	ODC1	actin binding LIM protein family, member 3
NM_002553	ORC5L	chromosome 10 open reading frame 18

NM_032523	OSBPL6	KIAA0056 protein
NM_002566	P2RY11	ribonucleotide reductase M2 polypeptide
BC019255	PAICS	CD200 molecule
NM_002581	PAPPA	exosome component 3
BC007377	PCGF5	DIP2 disco-interacting protein 2 homolog C (Drosophila)
NM_004563	PCK2	zinc finger, CCHC domain containing 10
NM_021830	PEO1	M-phase phosphoprotein 6
NM_004567	PFKFB4	SET domain containing 6
NM_002627	PFKP	centromere protein A, 17kDa
NM_021127	PMAIP1	phenylalanine-tRNA synthetase-like, beta subunit
NM_019014	POLR1B	isopentenyl-diphosphate delta isomerase 1
NM_014062	PSMD8BP1	exosome component 3
NM_015167	PTDSR	ribonucleotide reductase M2 polypeptide
NM_000963	PTGS2	colony stimulating factor 1 (macrophage)
NM_001015509	PTRH2	hypothetical LOC401093
NM_006505	PVR	MCM10 minichromosome maintenance deficient 10 (S. cerevisiae)
BC020832	RABL3	arylsulfatase family, member I
NM_002894	RBBP8	geminin, DNA replication inhibitor
NM_005349	RBPSUH	glycerol kinase
NM_005772	RCL1	zinc finger protein 175
NM_003800	RNGTT	Fas apoptotic inhibitory molecule
NM_004704	RNU3IP2	cell division cycle associated 4
XM_927086	RP11-98I6.3	chromosome 18 open reading frame 37
NM_001010919	RP1-93H18.5	glycogenin 2
NM_019008	RP5-1104E15.5	forkhead box D1
AY320405	RPL3	EGF-containing fibulin-like extracellular matrix protein 1
NM_015265	SATB2	protein arginine methyltransferase 5
NM_012429	SEC14L2	laminin, beta 3
NM_000602	SERPINE1	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 19
NM_022911	SLC26A6	coronin, actin binding protein, 2B
NM_015139	SLC35D1	small trans-membrane and glycosylated protein
NM_033518	SLC38A5	nucleolar and spindle associated protein 1
NM_153811	SLC38A6	CTF18, chromosome transmission fidelity factor 18 homolog (S. cerevisiae)
NM_001012661	SLC3A2	KIAA1462
NM_003045	SLC7A1	nucleolar and coiled-body phosphoprotein 1

NM_014331	SLC7A11	zinc finger protein 655
NM_001008539	SLC7A2	serum amyloid A1
NM_003486	SLC7A5	CTP synthase
NM_015915	SPG3A	cell division cycle 2, G1 to S and G2 to M
NM_144569	SPOCD1	WD repeat domain 4
BC056261	SPTY2D1	O-acyltransferase (membrane bound) domain containing 5
NM_080725	SRXN1	DEP domain containing 1
BC010550	STAR	Rho guanine nucleotide exchange factor (GEF) 3
NM_003155	STC1	polycomb group ring finger 5
NM_006948	STCH	La ribonucleoprotein domain family, member 2
NM_006950	SYN1	T cell receptor alpha locus
NM_003234	TFRC	Cbl-interacting protein Sts-1
NM_003254	TIMP1	mitochondrial ribosomal protein S30
NM_005077	TLE1	denticleless homolog (Drosophila)
BC015729	TLE3	glucosidase, alpha; acid (Pompe disease, glycogen storage disease type II)
BX641111	TNC	phosphodiesterase 4D interacting protein (myomegalin)
BC051810	TNFRSF19L	granzyme M (lymphocyte met-ase 1)
BC007086	TPI1	microfibrillar-associated protein 4
NM_021158	TRIB3	kinesin family member 20A
NM_017853	TXNL4B	TTK protein kinase
NM_001025366	VEGF	similar to TBC1 domain family member 3 (Rab GTPase-activating protein PRC17) (Prostate cancer gene 17 protein) (TRE17 alpha protein)
NM_004184	WARS	mannosidase, beta A, lysosomal-like
NM_006784	WDR3	nucleoside phosphorylase
NM_018669	WDR4	phosphatidylinositol glycan, class L
XM_087089	WDR43	WD repeat domain 66
NM_032118	WDR54	origin recognition complex, subunit 6 homolog-like (yeast)
NM_003401	XRCC4	target of EGR1, member 1 (nuclear)
NM_013250	ZNF215	uridine monophosphate synthetase (orotate phosphoribosyl transferase and orotidine-5'-decarboxylase)
NM_003904	ZNF259	complement component 1, q subcomponent binding protein

Table S3 Functional gene annotation analysis of significantly regulated genes in stage I alone

S3-1 GO terms BP categories of 542 up-regulated genes.

Go Term-BP ALL	Count	P-Value
cell adhesion	30	0.00003
biological adhesion	30	0.000031
extracellular matrix organization	10	0.00012
blood circulation	13	0.00016
circulatory system process	13	0.00016
regulation of cell migration	12	0.00027
regulation of blood vessel size	7	0.00039
regulation of tube size	7	0.00039
cell-substrate adhesion	9	0.00041
vascular process in circulatory system	7	0.00063
regulation of locomotion	12	0.00079
extracellular structure organization	11	0.00081
regulation of cell motion	12	0.00083
complement activation	6	0.00094
anatomical structure morphogenesis	38	0.001
activation of plasma proteins involved in acute inflammatory response	6	0.0011
cell-matrix adhesion	8	0.0012
regulation of cell-substrate adhesion	6	0.0014
acute inflammatory response	8	0.0021
positive regulation of biological process	55	0.0024
anatomical structure development	65	0.0028
humoral immune response	7	0.0031
regulation of cell adhesion	9	0.0036
developmental process	77	0.0037
multicellular organismal development	71	0.0042
cell migration	13	0.0047
organ development	47	0.0056
system development	59	0.0067
activation of immune response	7	0.0074
regulation of growth	14	0.0098
enzyme linked receptor protein signaling pathway	14	0.01
positive regulation of cellular process	48	0.01

S3-2 GO terms BP categories of 724 down-regulated genes.

Go Term-BP ALL	Count	P-Value
M phase	71	4.7E-40
cell cycle	99	1E-35
cell cycle phase	73	4.6E-35

M phase of mitotic cell cycle	56	6E-35
nuclear division	55	2.6E-34
mitosis	55	2.6E-34
cell cycle process	83	4.6E-34
organelle fission	55	2.4E-33
mitotic cell cycle	64	4E-30
cell division	50	4.8E-23
organelle organization	103	1.3E-19
chromosome segregation	25	6.8E-18
DNA metabolic process	57	1.2E-17
DNA replication	32	1.1E-14
mitotic sister chromatid segregation	16	3.6E-14
regulation of cell cycle	41	4.8E-14
sister chromatid segregation	16	5.9E-14
cellular process	378	9.2E-14
cellular component organization	134	2.6E-12
ribonucleoprotein complex biogenesis	28	4.7E-12
ribosome biogenesis	23	1.1E-11
nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	163	4.4E-11
cellular nitrogen compound metabolic process	172	4.4E-11
microtubule cytoskeleton organization	24	7.7E-11
microtubule-based process	31	1.3E-10
nitrogen compound metabolic process	174	1.4E-10
cellular metabolic process	264	2E-10
organelle localization	19	2.1E-10
cellular macromolecule metabolic process	220	2.6E-10
DNA repair	32	5.1E-10
regulation of cell cycle process	20	1.2E-09
ncRNA metabolic process	28	1.4E-09
establishment of organelle localization	16	1.5E-09
ncRNA processing	25	1.9E-09
establishment of chromosome localization	9	3.3E-09
chromosome localization	9	3.3E-09
spindle organization	13	5.4E-09
metaphase plate congression	8	6.2E-09
response to DNA damage stimulus	35	8.3E-09
RNA metabolic process	61	1.9E-08
regulation of mitotic cell cycle	21	2.9E-08
cellular component biogenesis	63	3.5E-08
macromolecule metabolic process	227	4E-08
mitotic metaphase plate congression	7	5.5E-08
RNA processing	42	6.5E-08
rRNA processing	16	9E-08

rRNA metabolic process	16	1.6E-07
primary metabolic process	261	1.7E-07
chromosome organization	38	1.9E-07
metabolic process	280	5.3E-07
DNA-dependent DNA replication	12	0.000001
cellular response to stress	40	1.2E-06
cytokinesis	10	2.7E-06
cell cycle checkpoint	14	3.1E-06
regulation of mitosis	11	5.6E-06
regulation of nuclear division	11	5.6E-06
negative regulation of cell cycle process	8	6.3E-06
biosynthetic process	147	8.2E-06
cellular macromolecular complex subunit organization	28	0.000011
DNA packaging	15	0.000011
cellular response to stimulus	48	0.000015
DNA recombination	14	0.000015
spindle checkpoint	6	0.000016
cellular biosynthetic process	142	0.000018
regulation of cyclin-dependent protein kinase activity	10	0.000029
cellular macromolecular complex assembly	25	0.000034
meiosis	13	0.000036
M phase of meiotic cell cycle	13	0.000036
meiotic cell cycle	13	0.000044
establishment of mitotic spindle localization	5	0.000051
cytoskeleton organization	30	0.000055
pigment biosynthetic process	8	0.00014
spindle localization	5	0.00015
establishment of spindle localization	5	0.00015
gene expression	122	0.00019
macromolecular complex subunit organization	40	0.0002
regulation of protein ubiquitination	12	0.0002
negative regulation of mitotic metaphase/anaphase transition	5	0.00022
mitotic cell cycle spindle assembly checkpoint	5	0.00022
DNA strand elongation during DNA replication	4	0.00026
regulation of organelle organization	18	0.0003
cell proliferation	28	0.00032
negative regulation of nuclear division	5	0.00033
negative regulation of mitosis	5	0.00033
pigment metabolic process	8	0.00035
protein-DNA complex assembly	11	0.00039
regulation of mitotic metaphase/anaphase transition	6	0.00041
mitotic chromosome condensation	5	0.00046
macromolecular complex assembly	37	0.00046

DNA strand elongation	4	0.0005
double-strand break repair	9	0.00051
positive regulation of protein modification process	16	0.00053
regulation of ubiquitin-protein ligase activity	10	0.00053
nucleoside metabolic process	9	0.00063
regulation of DNA metabolic process	12	0.00063
regulation of ligase activity	10	0.0007
pyrimidine nucleoside metabolic process	6	0.00077
chromosome condensation	6	0.00077
macromolecule biosynthetic process	113	0.00079
tRNA metabolic process	12	0.00084
oxidoreduction coenzyme metabolic process	8	0.00087
DNA integrity checkpoint	8	0.00087
cellular macromolecule biosynthetic process	112	0.0009
positive regulation of protein ubiquitination	10	0.00092
interphase of mitotic cell cycle	11	0.0011
cellular protein metabolic process	96	0.0012
regulation of ubiquitin-protein ligase activity during mitotic cell cycle	9	0.0013
interphase	11	0.0013
pyrimidine ribonucleoside metabolic process	5	0.0014
regulation of cell proliferation	40	0.0014
positive regulation of cell cycle	8	0.0015
mitotic cell cycle checkpoint	7	0.0017
cellular component assembly	43	0.0023
cofactor metabolic process	15	0.0023
ribonucleoside metabolic process	7	0.0024
DNA damage checkpoint	7	0.003
RNA modification	7	0.003
negative regulation of organelle organization	9	0.0032
regulation of phosphate metabolic process	27	0.0032
regulation of phosphorus metabolic process	27	0.0032
nucleobase, nucleoside and nucleotide metabolic process	20	0.0033
positive regulation of cellular metabolic process	42	0.0034
negative regulation of cellular metabolic process	36	0.0035
regulation of protein modification process	19	0.0035
nucleosome assembly	9	0.0037
positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	8	0.0042
regulation of transferase activity	22	0.0043
positive regulation of metabolic process	43	0.0044
response to stress	70	0.0044
nitrogen compound biosynthetic process	20	0.0045
nucleobase metabolic process	5	0.0045

ribonucleoprotein complex assembly	8	0.0046
chromatin assembly	9	0.0046
positive regulation of cellular protein metabolic process	16	0.0046
negative regulation of macromolecule metabolic process	36	0.0047
phosphoinositide-mediated signaling	9	0.0049
positive regulation of ubiquitin-protein ligase activity	8	0.0049
regulation of catalytic activity	40	0.005
cellular localization	43	0.0051
translation	20	0.0054
quinone cofactor metabolic process	4	0.0061
positive regulation of ligase activity	8	0.0062
regulation of molecular function	44	0.0065
coenzyme metabolic process	12	0.0065
negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	27	0.0066
negative regulation of protein ubiquitination	8	0.0067
positive regulation of protein metabolic process	16	0.0068
nucleosome organization	9	0.0068
negative regulation of metabolic process	37	0.0069
negative regulation of cellular process	68	0.0069
positive regulation of cell proliferation	23	0.0071
negative regulation of biological process	73	0.0073
DNA unwinding during replication	4	0.0077
nucleobase biosynthetic process	4	0.0077
negative regulation of nitrogen compound metabolic process	27	0.0077
tRNA processing	8	0.0077
meiosis I	6	0.0081
cofactor biosynthetic process	9	0.0087
secondary metabolic process	8	0.0095
establishment of localization in cell	39	0.0095
positive regulation of macromolecule metabolic process	39	0.01

Table S4 Functional gene annotation analysis of significantly regulated genes in stage II alone

S4-1 GO terms BP categories of 706 up-regulated genes.

Go Term-BP ALL	Count	P-Value
cellular ketone metabolic process	41	6.8E-09
oxoacid metabolic process	40	1.2E-08
carboxylic acid metabolic process	40	1.2E-08
organic acid metabolic process	40	1.5E-08
alcohol metabolic process	33	6.7E-08
glycolysis	11	2.5E-07
cellular amino acid metabolic process	22	3.8E-07
monosaccharide metabolic process	21	1.2E-06
cellular amino acid and derivative metabolic process	27	1.4E-06
amine metabolic process	29	1.6E-06
glucose catabolic process	11	1.9E-06
glucose metabolic process	17	0.000002
tRNA aminoacylation	10	2.1E-06
tRNA aminoacylation for protein translation	10	2.1E-06
amino acid activation	10	2.1E-06
hexose metabolic process	19	2.3E-06
cellular amine metabolic process	24	4.2E-06
hexose catabolic process	11	9.8E-06
cellular carbohydrate catabolic process	12	0.000011
monosaccharide catabolic process	11	0.000013
alcohol catabolic process	11	0.000041
carbohydrate catabolic process	12	0.00011
cellular carbohydrate metabolic process	24	0.00017
tRNA metabolic process	12	0.00023
carbohydrate metabolic process	29	0.00026
cellular amino acid biosynthetic process	8	0.00031
generation of precursor metabolites and energy	20	0.00049
response to chemical stimulus	53	0.00062
response to hypoxia	12	0.00068
L-serine metabolic process	4	0.00086
response to oxygen levels	12	0.001
amine biosynthetic process	9	0.0011
response to organic substance	33	0.0019
response to endogenous stimulus	22	0.0019
cellular carbohydrate biosynthetic process	8	0.0019
regulation of myeloid cell differentiation	8	0.0021
oxidation reduction	30	0.0022
serine family amino acid biosynthetic process	4	0.0024
metabolic process	223	0.0039

amine catabolic process	8	0.0039
response to antibiotic	5	0.0041
serine family amino acid metabolic process	5	0.0041
sterol metabolic process	9	0.0045
response to stress	61	0.0053
steroid metabolic process	13	0.0062
carbohydrate biosynthetic process	9	0.0063
ncRNA metabolic process	14	0.0066
negative regulation of myeloid cell differentiation	5	0.0069
hexose biosynthetic process	5	0.0078
primary metabolic process	202	0.0084
regulation of cell differentiation	23	0.0086
glutamate biosynthetic process	3	0.0092
cholesterol metabolic process	8	0.0095
nitrogen compound biosynthetic process	17	0.01

S4-2 GO terms BP categories of 615 up-regulated genes.

Go Term-BP ALL	Count	P-Value
nervous system development	51	4.7E-08
multicellular organismal development	94	1.1E-06
anatomical structure development	85	1.8E-06
system development	80	1.9E-06
developmental process	99	3.3E-06
cell-cell signaling	31	5.9E-06
neuron differentiation	24	0.000036
anatomical structure morphogenesis	46	0.000048
multicellular organismal process	120	0.000049
neuron development	20	0.000073
regulation of cell communication	41	0.000078
cell communication	34	0.000088
cell development	29	0.00014
generation of neurons	26	0.00021
cellular developmental process	57	0.00021
cell differentiation	55	0.00025
regulation of signal transduction	35	0.00026
neurogenesis	27	0.00027
transmission of nerve impulse	19	0.00033
axon guidance	10	0.00035
cell morphogenesis	19	0.0004
synaptic transmission	17	0.00044
cellular component morphogenesis	20	0.00055
cell recognition	7	0.00087
organ development	55	0.0011

response to external stimulus	34	0.0011
cell morphogenesis involved in differentiation	14	0.0016
cell projection morphogenesis	14	0.0016
camera-type eye morphogenesis	6	0.002
cell part morphogenesis	14	0.0024
negative regulation of cellular component organization	10	0.0026
sensory organ development	13	0.0027
eye morphogenesis	7	0.0028
response to inorganic substance	12	0.0034
regulation of biological quality	46	0.0038
cell projection organization	17	0.0039
cell morphogenesis involved in neuron differentiation	12	0.0039
neuron projection morphogenesis	12	0.0045
response to metal ion	9	0.005
response to calcium ion	6	0.0052
regulation of cellular component organization	19	0.0064
axonogenesis	11	0.0065
neuron projection development	13	0.0066
reproductive process	27	0.0075
behavior	19	0.0081
reproduction	27	0.0083
neuron recognition	4	0.0086
neuromuscular process	6	0.0087
response to abiotic stimulus	16	0.009

Table S5 Functional gene annotation analysis of significantly regulated genes shared in two stages during the hHPCs differentiation

GO Term-BP ALL (95 genes, I, II up)	**Count**	**P-Value**
negative regulation of defense response	6	0.0000016
response to insulin stimulus	7	0.00002
negative regulation of inflammatory response	5	0.000025
response to chemical stimulus	20	0.000062
skeletal system development	10	0.00007
negative regulation of response to external stimulus	5	0.00017
regulation of multicellular organismal process	16	0.00019
positive regulation of biological process	25	0.00021
response to peptide hormone stimulus	7	0.00022
negative regulation of response to stimulus	6	0.00024
GO Term-BP ALL (73 genes, I, II down)	**Count**	**P-Value**
blood coagulation	6	0.000022
coagulation	6	0.000022
hemostasis	6	0.000028
wound healing	7	0.000038
regulation of body fluid levels	6	0.0001
DNA replication	6	0.00041
cell cycle	10	0.0009
positive regulation of leukocyte migration	3	0.0011
response to wounding	8	0.0017
regulation of leukocyte migration	3	0.002
GO Term-BP ALL (177 genes, I up, II down)	**Count**	**P-Value**
cell surface receptor linked signal transduction	31	0.000049
enzyme linked receptor protein signaling pathway	12	0.000066
organ development	29	0.0001
system development	34	0.00025
regulation of multicellular organismal process	19	0.00028
anatomical structure development	35	0.00051
developmental process	40	0.0009
transmembrane receptor protein serine/threonine kinase signaling pathway	6	0.0012
multicellular organismal development	37	0.0012
female pregnancy	6	0.0016
GO Term-BP ALL (169 genes, I down, II up)	**Count**	**P-Value**
amino acid transport	7	0.000035
alcohol metabolic process	13	0.000037
carboxylic acid transport	8	0.000068
organic acid transport	8	0.000071
amine transport	7	0.00016
nitrogen compound metabolic process	40	0.0027

cellular nitrogen compound metabolic process	39	0.003
carbohydrate metabolic process	11	0.0034
monosaccharide metabolic process	7	0.0043
cellular carbohydrate metabolic process	9	0.0051

Table S6-1 The differentially expressed genes in stage 1 involved in cell morphogenesis

SEQ_ID	Gene_Name	Description	P-value	Fold Change	Regulation
NM_003565	ULK1	unc-51-like kinase 1 (C. elegans)	0.001112004	2.1749628	up
NM_001025930	TTLL3	tubulin tyrosine ligase-like family, member 3	0.022438414	2.9282434	up
NM_006288	THY1	Thy-1 cell surface antigen	0.002236558	2.1170537	up
NM_003243	TGFBR3	transforming growth factor, beta receptor III (betaglycan, 300kDa)	0.004257658	3.6808603	up
NM_003239	TGFB3	transforming growth factor, beta 3	7.27E-07	3.9477339	up
NM_005631	SMO	smoothened homolog (Drosophila)	0.026544534	3.0802407	up
BC098106	SLITRK5	SLIT and NTRK-like family, member 5	0.039970133	2.7017598	up
NM_003966	SEMA5A	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	8.97E-04	6.3031693	up
NM_002961	S100A4	S100 calcium binding protein A4 (calcium protein, calvasculin, metastasin, murine placental homolog)	0.013041517	2.7580175	up
BC007925	RXRA	retinoid X receptor, alpha	0.041113526	2.2438416	up
AF112221	RUFY3	RUN and FYVE domain containing 3	0.04312	2.0639415	up
BC008623	ROBO3	roundabout, axon guidance receptor, homolog 3 (Drosophila)	0.042731248	2.2812417	up
NM_020663	RHOJ	ras homolog gene family, member J	0.001351383	2.8807347	up
NM_002821	PTK7	PTK7 protein tyrosine kinase 7	0.00578	2.1254432	up
BC022016	PRKCI	protein kinase C, iota	0.029417548	2.3961835	up
NM_012401	PLXNB2	plexin B2	0.006030819	2.8451014	up
NM_002673	PLXNB1	plexin B1	0.006926047	3.457886	up
BC071566	PARD3	par-3 partitioning defective 3 homolog (C. elegans)	0.007274893	2.5061336	up
BC042830	OFD1	oral-facial-digital syndrome 1	0.037263248	2.4516008	up
NM_005964	MYH10	myosin, heavy polypeptide 10, non-muscle	3.82E-04	5.0942516	up
NM_001025100	MBP	myelin basic protein	0.00518265	4.2023287	up
AB209512	LIMA1	LIM domain and actin binding 1	0.011375751	2.2395325	up
NM_002291	LAMB1	laminin, beta 1	0.00272	2.1057777	up
NM_000213	ITGB4	integrin, beta 4	0.021688038	2.39776	up
NM_181501	ITGA1	integrin, alpha 1	0.00338	2.1475258	up
NM_000165	GJA1	gap junction protein, alpha 1, 43kDa (connexin 43)	0.00032	4.0409064	up
NM_182828	GDF7	growth differentiation factor 7	0.015491038	2.1976216	up
NM_002037	FYN	FYN oncogene related to SRC, FGR, YES	0.014496534	2.4656248	up
BC019895	FBLIM1	filamin binding LIM protein 1	0.012971541	2.389736	up
BC052979	EFNB1	ephrin-B1	0.04150975	2.6089635	up
NM_000109	DMD	dystrophin (muscular dystrophy, Duchenne and Becker types)	0.00826136	2.106343	up
NM_015404	DFNB31	deafness, autosomal recessive 31	0.040571257	2.02652	up

BC003064	DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	0.009875146	2.4454923	up
BC010514	CLU	clusterin	0.009535163	2.5454612	up
BC019270	CDC42EP3	CDC42 effector protein (Rho GTPase binding) 3	0.004088029	2.123438	up
BC109065	BBS1	Bardet-Biedl syndrome 1	0.006018481	2.1862755	up
NM_000041	APOE	apolipoprotein E	2.18E-04	13.660188	up
NM_001164	APBB1	amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)	0.026795521	2.7240481	up
NM_018153	ANTXR1	anthrax toxin receptor 1	0.004252684	2.323502	up
NM_001124	ADM	adrenomedullin	0.003690694	2.307919	up
NM_001025366	VEGF	vascular endothelial growth factor	0.004783762	3.4873211	down
NM_004181	UCHL1	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	0.008277879	2.0350597	down
NM_153712	TTL	tubulin tyrosine ligase	0.017611364	2.2168527	down
BC010866	TRAPPC4	trafficking protein particle complex 4	0.03279	2.1353054	down
NM_015915	SPG3A	spastic paraplegia 3A (autosomal dominant)	0.006669874	2.005465	down
NM_005904	SMAD7	SMAD, mothers against DPP homolog 7 (Drosophila)	0.010041158	4.098476	down
NM_005902	SMAD3	SMAD, mothers against DPP homolog 3 (Drosophila)	1.92E-04	2.097961	down
BC010704	SH2B	SH2-B homolog	0.00529	2.226669	down
NM_002872	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	0.02610	2.396448	down
BC040105	NOX4	NADPH oxidase 4	0.021824026	2.087286	down
NM_006101	KNTC2	kinetochore associated 2	0.033449423	3.0935338	down
NM_052899	KIAA1893	G protein-regulated inducer of neurite outgrowth 1	0.00792	2.0695422	down
NM_000600	IL6	interleukin 6 (interferon, beta 2)	0.00588	2.5422394	down
NM_000194	HPRT1	hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	0.03721	2.2548966	down
NM_005524	HES1	hairy and enhancer of split 1, (Drosophila)	0.002381294	4.978615	down
NM_005270	GLI2	GLI-Kruppel family member GLI2	5.91E-05	2.9623184	down
BC035625	EGR2	early growth response 2 (Krox-20 homolog, Drosophila)	0.002568145	6.537261	down
NM_006094	DLC1	deleted in liver cancer 1	0.017850777	2.1001132	down
NM_001809	CENPA	centromere protein A, 17kDa	0.004895133	3.4486568	down
NM_001794	CDH4	cadherin 4, type 1, R-cadherin (retinal)	0.012968633	2.1490993	down
NM_044472	CDC42	cell division cycle 42 (GTP binding protein, 25kDa)	4.67E-04	2.9767895	down
BC004372	CD44	CD44 molecule (Indian blood group)	0.01603	3.7392714	down
NM_017785	CCDC99	coiled-coil domain containing 99	1.09E-04	5.442056	down

Table S6-2 The differentially expressed genes in stage II involved in cell morphogenesis

SEQ_ID	Gene_Name	Description	P-value	Fold Change	Regulation
NM_001025366	VEGF	vascular endothelial growth factor	0.000642692	8.081642	up
NM_170744	UNC5B	unc-5 homolog B (C. elegans)	6.33E-04	2.5263813	up
NM_006288	THY1	Thy-1 cell surface antigen	0.000275988	2.7471893	up
BC015749	STXBP1	syntaxin binding protein 1	0.049957436	2.793498	up
BC038850	STK4	serine/threonine kinase 4	0.041938737	2.686412	up
NM_015915	SPG3A	spastic paraplegia 3A (autosomal dominant)	0.001163982	2.4745915	up
NM_004172	SLC1A3	solute carrier family 1 (glial high affinity glutamate transporter), member 3	0.00047	4.652989	up
NM_147150	PALM2-AKAP2	PALM2-AKAP2 protein	0.04941	3.9649818	up
NM_004756	NUMBL	numb homolog (Drosophila)-like	0.021706646	2.7153475	up
NM_006162	NFATC1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1	0.00353	2.3109372	up
BC041374	LRRC4C	leucine rich repeat containing 4C	0.020886647	3.514237	up
NM_000425	L1CAM	L1 cell adhesion molecule	3.78E-04	5.399565	up
NM_003709	KLF7	Kruppel-like factor 7 (ubiquitous)	0.02407098	2.8363883	up
BC045542	KIF3A	kinesin family member 3A	0.00316	2.0065503	up
NM_004972	JAK2	Janus kinase 2 (a protein tyrosine kinase)	0.017939717	2.075862	up
NM_002185	IL7R	interleukin 7 receptor	1.65E-04	1.33E+01	up
NM_001530	HIF1A	hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	6.80E-04	3.1373606	up
NM_001024215	FBLIM1	filamin binding LIM protein 1	0.021484122	2.2699285	up
NM_004429	EFNB1	ephrin-B1	0.014348468	2.496405	up
AK125894	DKFZP586P0123	hypothetical protein	0.049991205	2.6673484	up
NM_004385	CSPG2	chondroitin sulfate proteoglycan 2 (versican)	0.00501118	2.2604275	up
NM_031361	COL4A3BP	collagen, type IV, alpha 3 (Goodpasture antigen) binding protein	0.001295015	2.3176868	up
BC112940	CLASP1	cytoplasmic linker associated protein 1	0.008049205	2.657452	up
NM_006449	CDC42EP3	CDC42 effector protein (Rho GTPase binding) 3	0.011114754	2.0954468	up
NM_001706	BCL6	B-cell CLL/lymphoma 6 (zinc finger protein 51)	0.006807243	3.7061498	up
NM_032146	ARL6	ADP-ribosylation factor-like 6	0.001387085	2.695869	up
NM_000041	APOE	apolipoprotein E	6.83E-05	7.895473	up
NM_001124	ADM	adrenomedullin	0.002432253	2.5417335	up
NM_033131	WNT3A	wingless-type MMTV integration site family, member 3A	0.012537284	2.8236485	down
NM_020335	VANGL2	vang-like 2 (van gogh, Drosophila)	0.034692552	3.3394382	down
NM_001001894	TTC3	tetratricopeptide repeat domain 3	0.001122245	2.321824	down
NM_001018008	TPM1	tropomyosin 1 (alpha)	0.00738	2.4083729	down
NM_003239	TGFB3	transforming growth factor, beta 3	3.21E-04	4.1152616	down
NM_012445	SPON2	spondin 2, extracellular matrix protein	0.005216891	2.478143	down
NM_006080	SEMA3A	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	3.66E-04	4.18516	down

NM_002872	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	0.048959512	2.160753	down
NM_002855	PVRL1	poliovirus receptor-related 1 (herpesvirus entry mediator C; nectin)	0.0068376	3.0776873	down
NM_002840	PTPRF	protein tyrosine phosphatase, receptor type, F	0.003784943	2.068883	down
BC029822	PDGFB	platelet-derived growth factor beta polypeptide (simian sarcoma viral (v-sis) oncogene homolog)	0.00266	2.0785835	down
NM_004852	ONECUT2	one cut domain, family member 2	0.019351047	2.632819	down
NM_004822	NTN1	netrin 1	0.012867027	3.741368	down
NM_004796	NRXN3	neurexin 3	0.0012135	5.0871463	down
NM_003872	NRP2	neuropilin 2	0.001059144	3.3427832	down
NM_016931	NOX4	NADPH oxidase 4	0.007502841	3.3562307	down
NM_002506	NGFB	nerve growth factor, beta polypeptide	0.009710856	7.3024135	down
BC033900	NDE1	nudE nuclear distribution gene E homolog 1 (A. nidulans)	9.87E-04	2.0401835	down
NM_005964	MYH10	myosin, heavy polypeptide 10, non-muscle	5.74E-04	2.5268657	down
NM_001025100	MBP	myelin basic protein	0.004283231	5.5663433	down
NM_016357	LIMA1	LIM domain and actin binding 1	0.001033113	2.0585601	down
NM_001005731	ITGB4	integrin, beta 4	0.00221	3.0821695	down
NM_000600	IL6	interleukin 6 (interferon, beta 2)	0.04903364	3.8792894	down
NM_001010931	HGF	hepatocyte growth factor (hepapoietin A; scatter factor)	4.06E-04	2.5206928	down
NM_182828	GDF7	growth differentiation factor 7	0.001337855	2.7772722	down
NM_003644	GAS7	growth arrest-specific 7	0.024752568	2.0029397	down
NM_002045	GAP43	growth associated protein 43	0.02072	2.3222516	down
NM_018351	FGD6	FYVE, RhoGEF and PH domain containing 6	7.59E-04	2.358395	down
NM_001986	ETV4	ets variant gene 4 (E1A enhancer binding protein, E1AF)	0.003483942	5.3934455	down
NM_004443	EPHB3	EPH receptor B3	0.019268857	3.8572438	down
BC094761	EGFR	epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian)	0.001494542	2.8196142	down
NM_004946	DOCK2	dedicator of cytokinesis 2	0.022490676	2.5767684	down
AK056190	DFNB31	deafness, autosomal recessive 31	0.00035	3.107523	down
NM_030582	COL18A1	collagen, type XVIII, alpha 1	0.003695049	6.7027907	down
BC075040	CHRNB2	cholinergic receptor, nicotinic, beta 2 (neuronal)	0.014531849	3.1385067	down
NM_006366	CAP2	CAP, adenylate cyclase-associated protein, 2 (yeast)	0.048545163	2.6019742	down
NM_145200	CABP4	calcium binding protein 4	0.036602996	2.7718177	down
NM_001709	BDNF	brain-derived neurotrophic factor	0.002600991	5.089972	down
NM_001649	APXL	apical protein-like (Xenopus laevis)	0.026382333	3.9293947	down
AK025429	ANTXR1	anthrax toxin receptor 1	0.002507914	2.5154662	down

Table S6-3 List of genes involved in cell morphogenesis shared by two stages

SEQ ID	Gene Name	Description	Stage I			Stage II		
			P-value	Fold Change	Regulation	P-value	Fold Change	Regulation
NM_001124	ADM	adrenomedullin	0.00369069	2.307919	up	0.00243225	2.5417335	up
NM_018153	ANTXR1	anthrax toxin receptor 1	0.00425268	2.323502	up	0.00250791	2.5154662	down
NM_000041	APOE	apolipoprotein E	2.18E-04	13.660188	up	6.83E-05	7.895473	up
NM_006449	CDC42EP3	CDC42 effector protein (Rho GTPase binding) 3	0.00408803	2.123438	up	0.01111475	2.0954468	up
NM_015404	DFNB31	deafness, autosomal recessive 31	0.04057126	2.02652	up	0.00035	3.107523	down
NM_004429	EFNB1	ephrin-B1	0.04150975	2.6089635	up	0.01434847	2.496405	up
NM_001024215	FBLIM1	filamin binding LIM protein 1	0.01297154	2.389736	up	0.02148412	2.2699285	up
NM_182828	GDF7	growth differentiation factor 7	0.01549104	2.1976216	up	0.00133786	2.7772722	down
NM_000600	IL6	interleukin 6 (interferon, beta 2)	0.00588	2.5422394	down	0.04903364	3.8792894	down
NM_000213	ITGB4	integrin, beta 4	0.02168804	2.39776	up	0.00221	3.0821695	down
NM_016357	LIMA1	LIM domain and actin binding 1	0.01137575	2.2395325	up	0.00103311	2.0585601	down
NM_001025100	MBP	myelin basic protein	0.00518265	4.2023287	up	0.00428323	5.5663433	down
NM_005964	MYH10	myosin, heavy polypeptide 10, non-muscle	3.82E-04	5.0942516	up	5.74E-04	2.5268657	down
NM_016931	NOX4	NADPH oxidase 4	0.02182403	2.087286	down	0.00750284	3.3562307	down
NM_002872	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	0.02610	2.396448	down	0.04895951	2.160753	down
NM_015915	SPG3A	spastic paraplegia 3A (autosomal dominant)	0.00666987	2.005465	down	0.00116398	2.4745915	up
NM_003239	TGFB3	transforming growth factor, beta 3	7.27E-07	3.9477339	up	3.21E-04	4.1152616	down
NM_006288	THY1	Thy-1 cell surface antigen	0.00223656	2.1170537	up	0.00027599	2.7471893	up
NM_001025366	VEGF	vascular endothelial growth factor	0.00478376	3.4873211	down	0.00064269	8.081642	up