Cancer Biol Med 2019. doi: 10.20892/j.issn.2095-3941.2018.0184

ORIGINAL ARTICLE

Identification of portal vein tumor thrombus with an independent clonal origin in hepatocellular carcinoma via multi-omics data analysis

Shupeng Liu[1*], Zaixin Zhou[2*], Yin Jia[2], Jie Xue[3], Zhiyong Liu[2], Kai Cheng[1], Shuqun Cheng[3], Shanrong Liu[2]

[1]Clinical Research Center, Changhai Hospital, Second Military Medical University, Shanghai 200433, China; [2]Department of Laboratory Diagnostics, Changhai Hospital, Second Military Medical University, Shanghai 200433, China; [3]Department of Hepatic Surgery VI, Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai 200433, China

ABSTRACT

Objective: Multiple mechanisms underlying the development of portal vein tumor thrombus (PVTT) in hepatocellular carcinoma (HCC) have been reported recently. However, the origins of PVTT remain unknown. Increasing multi-omics data on PVTTs in HCCs have made it possible to investigate whether PVTTs originate from the corresponding primary tumors (Ts).

Methods: The clonal relationship between PVTTs and their corresponding primary Ts was investigated using datasets deposited in public databases. One DNA copy number variations dataset and three gene expression datasets were downloaded for the analyses. Clonality analysis was performed to investigate the clonal relationship between PVTTs and Ts from an individual patient. Differential gene expression analysis was applied to investigate the gene expression profiles of PVTTs and Ts.

Results: One out of 19 PVTTs had no clonal relationship with its corresponding T, whereas the others did. The PVTTs with independent clonal origin showed different gene expression and enrichment in biological processes from the primary Ts. Based on the unique gene expression profiles, a gene signature including 24 genes was used to identify pairs of PVTTs and primary Ts without any clonal relationship. Validation in three datasets showed that these types of pairs of PVTTs and Ts can be identified by the 24-gene signature.

Conclusions: Our findings show a direct evidence for PVTT origin and consolidate the heterogeneity of PVTTs observed in clinic. The results suggest that PVTT investigation at a molecular level is clinically necessary for diagnosis and treatment.

KEYWORDS Hepatocellular carcinoma; portal vein tumor thrombus; clonal origin; copy number variation; bioinformatics

Introduction

Portal vein tumor thrombus (PVTT) is one of the main hepatocellular carcinoma (HCC) complications, found in approximately 50% of HCC patients[1]. The prognosis of HCC patients with PVTT is poor, and the median overall survival (OS) time of patients with PVTT was primarily less than 1 year[2,3]. Although clinicians in Europe and America recommend sorafenib alone, and multidisciplinary therapy has been adapted in Asia[4,5], no worldwide consensus or guidelines have shown effectiveness for specific PVTT treatment, which might be due to a lack of knowledge about the nature of PVTT.

Although currently, hypoxia, non-coding RNAs, and cancer stem cells have been found to contribute to PVTT development[6-9], the clonal origins of PVTT still remain unknown. The identical gene expression profiles between PVTT and the corresponding primary HCCs demonstrated that PVTT could originate from HCC primary nodules by metastasis[10,11]. Contradictorily, in small HCC patients, some PVTTs are clinically defined, that are distant from the liver parenchyma tumor (PT) nodules, and some PVTTs are found without liver parenchyma tumor nodules[12,13]. Moreover, some PVTTs showed significantly different gene expression profiles compared with those of the paired primary tumors (Ts)[11]. These findings indicate that PVTTs display relatively high heterogeneity and may have different clonal origins from the paired primary Ts.

Recently, high-throughput technology, such as microarray and sequencing, has been widely used to investigate the

*These authors contributed equally to this work.
Correspondence to: Shanrong Liu
E-mail: liushanrong@hotmail.com
Received June 10, 2018; accepted October 30, 2018.
Available at www.cancerbiomed.org

mechanisms underlying the development of Ts, including HCC. Among them, 3 published studies involved the genomic or transcriptomic datasets of HCC as well as PVTT[9,11,14]. In the present study, previously published datasets (GSE77509, GSE69164, GSE77275, and GSE74656), and freshly obtained samples were used for analysis, and we found a type of PVTT that has a different clonal origin from that of the corresponding primary Ts in HCC. Additionally, this type of PVTT was found to express a specific set of genes, which varied from that of the corresponding Ts.

Materials and methods

Data resources

Four previously published datasets (Cytoscan HD array, GSE77275; RNA-seq, GSE77509 and GSE69164; and cDNA microarray, GSE74656) used in this study were downloaded from the NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/). GSE77275 included the copy number variation (CNV) profiles of 60 matched samples (PVTT/T/PT from 20 patients) assessed by Cytoscan HD array. GSE77509 included the gene expression profiles of these 60 samples assessed by RNA-seq. DNA and total RNA from the 60 samples were used for CNV and gene expression analysis, respectively, as previously reported[11].

The samples of patient 14 (P14), in these two datasets, were not included in the present study because they were reported to be contaminated[11]. GSE69164 included the gene expression of 33 matched samples from 11 patients assessed by RNA-seq[14]. GSE74656 included the gene expression data of 15 matched samples from 5 patients assessed by cDNA analysis.

Clonality testing

To investigate the clonal relationship between PVTTs and Ts from an individual patient, the CNV profiles of 19 pairs of PVTTs and Ts were extracted from GSE77275 using R package Affy2sv and loaded to R package Clonality for further analysis. The CNV of the chromosome region in each patient was estimated using the circular binary segmentation (CBS) analysis based on the normalized weighted Log_2 ratio. A significant CNV in chromosome regions was determined according to the threshold value (median absolute deviations = 1.25). The clonal relationship between PVTTs and Ts from an individual patient was investigated according to the likelihood ratio 2 (LR2)[15,16].

Differential gene expression analysis

The gene expression profiles of PVTTs and Ts from the datasets GSE77509, GSE69164, and GSE74656 were extracted using the appropriate methods[11]. EBSeq from NovelBrain BioCloud (https://cloud.novelbrain.com) was applied to analyze differentially expressed genes between PVTTs and Ts from individual patients in GSE77509. Genes with fold changes > 3 and FDR < 0.05 were identified as significantly differentially expressed.

Gene ontology and KEGG pathway enrichment analyses

The differentially expressed genes between PVTT13 and T13 were loaded to Metascape for Gene ontology (GO) and KEGG pathway enrichment analyses[17]. Protein-protein interaction (PPI) network analysis was also performed in Metascape, and the PPI network was exported using Cytoscape 3.4.

Gene signature enrichment analysis

Gene expression signature enrichment was evaluated using single-sample gene set enrichment analysis (ssGSEA) and nearest template prediction (NTP) analyses (Gene Pattern modules)[18]. Molecular Signature Database gene sets, (MSigDB, www.broadinstitute.org/msigdb) including hematopoietic stem cells and liver cancer stem cells, and previously published gene-expression signatures representing different cell types including bile duct cells (BDC), hepatocytes, hepatic progenitor cells (HPC), and hepatic stellate cells (HSC)[19] (**Supplementary Table S1**), were tested using ssGSEA. Previously, published HCC molecular classifications were also analyzed using NTP analysis[20,21].

Tumor tissues

Eight pairs of human PVTTs and Ts were obtained from HCC patients who underwent curative resection in the Eastern Hepatobiliary Surgery Hospital (Shanghai, China) after obtaining informed consent. All 8 pairs of PVTTs and Ts were used to perform real-time polymerase chain reaction (PCR) analysis. The study was approved by the Ethics Committee of Eastern Hepatobiliary Surgery Hospital.

Quantitative real-time PCR

Total RNA was isolated using TRIzol reagent (Invitrogen,

CA, USA), and reverse-transcription (RT) was performed using the PrimeScript RT Reagent Kit (Takara, Kusatsu, Japan), according to manufacturer's instructions. Real-time PCR was performed with SYBR Premix Ex Taq (Takara, Kusatsu, Japan) using the StepOnePlus Real-Time PCR System (Applied Biosystems, CA, USA). The primers used are listed in **Supplementary Table S2**. The gene expression levels were calculated relative to the expression of β-actin.

Cluster analysis

All pairs of PVTT and Ts were classified using one minus Pearson correlation hierarchical cluster analysis by the 24-gene signatures (Morpheus). The 24 genes are listed in **Supplementary Table S3**.

Statistical analysis

All *P* values and FDR values were obtained by the appropriate methods used in the different gene expression and enrichment analyses. Statistical analysis of the gene signature enrichment score from ssGSEA was not performed due to only one PVTT-T pair identified to have no clonal relationship.

Results

Analysis of the clonal relationship between PVTTs and corresponding primary Ts

To investigate the clonal relationship of PVTTs with corresponding Ts, DNA copy number variation (CNV) profiles of 19 paired PVTTs and Ts were analyzed using the R package Clonality[22]. The clonal-relatedness of the paired tissues was evaluated according to the LR2 value. Overall, 18 out of 19 PVTT-T pairs showed a significantly higher LR2 value ($P < 0.001$), whereas only one pair (patient 13, P13) showed a much lower LR2 value ($P > 0.05$) (**Table 1**). As a low LR2 value indicates no clonal relatedness, PVTT13 was identified to have an independent clonal origin from T13, whereas the other PVTTs had a clonal origin from the corresponding Ts. For further verification, the CNV of the chromosome regions and genes was investigated using the R package DNA copy[23]. The DNA CNV patterns of chromosome regions were estimated based on the threshold values (threshold = 0.3). Aberrant chromosome regions showed very different CNV patterns between PVTT13 and T13, whereas these differences were not observed in other pairs (**Figure 1A**). The region of 2q24.1-q31.1 was deleted in PVTT13 but amplified in T13. Regions of 5q13.2-q35.2 and 15q11.2-q21.1 were amplified in PVTT13 but deleted in T13. These three chromosome regions displayed similar CNV patterns between PVTTs and Ts from other patients (**Figure 1A**). CNV analysis of genes showed that 10 HCC driver genes have different copy numbers between the PVTTs and Ts from individual patients (**Figure 1B**). Less than 3 different mutant genes were observed in most PVTT-T pairs ($n = 15$), whereas 3 or more different mutant genes were observed in pairs from patient 6 (P6), patient 11 (P11), P13, and patient 22 (P22) (**Figure 1B**). Among these 4 PVTT-T pairs with more than 2 mutant genes with different copy numbers, only the pair from P13 had 2 mutant genes, with 2 and 3 different copies. Five copies of *ADCY2* and 3 copies of *CTNNB1* were observed in PVTT13, whereas 2 copies of *ADCY2* and 5 copies of *CTNNB1* were observed in T13 (**Figure 1B**). Additionally, Hoshida S1-S2-S3 classes of the PVTTs and Ts were evaluated by NTP analysis using their gene expression profiles[20]. All PVTTs and Ts were classified into one of the Hoshida S1-S2-S3 classes (Bonferroni $P = 0.019$, FDR < 0.01, **Figure 1C**). Among the 19 PVTT-T pairs, 6 pairs including PVTT13-T13 had their PVTTs and Ts in different subclasses

Table 1 Clonality analysis of PVTT and T ($n = 19$) according to copy number variations at the probe level

T	PVTT	LR2*	*P*
T3	PVTT3	48890042738	0
T6	PVTT6	3965105136	0
T7	PVTT7	1.27E+20	0
T8	PVTT8	1.21E+43	0
T10	PVTT10	7.68E+12	0
T11	PVTT11	43403254817	0
T12	PVTT12	4.67E+26	0
T13	PVTT13	0.005090782	0.201316
T15	PVTT15	1.49E+36	0
T16	PVTT16	2.73E+23	0
T17	PVTT17	1.68E+16	0
T18	PVTT18	3.61E+29	0
T19	PVTT19	1.64E+20	0
T20	PVTT20	1.75E+33	0
T21	PVTT21	6.16E+17	0
T22	PVTT22	9461268136	0
T24	PVTT24	1.07E+32	0
T25	PVTT25	2.10E+13	0
T26	PVTT26	3.81E+43	0

*LR2, likelihood ratio 2, quantifies the odds that the two tissues are clonal; T, primary tumor; PVTT, portal vein tumor thrombus.

(**Figure 1C**). PVTT13 was classified into the S3 class, which is associated with hepatocyte differentiation, and T13 was classified into the S1 class, which reflects an aberrant activation of the Wnt-signaling pathway (**Figure 1C**). Collectively, these data indicated that PVTT13 and T13 had different origins and that some PVTTs may not originate from the primary Ts.

Analysis of aberrantly expressed genes in PVTT13

To investigate aberrantly expressed genes in PVTTs and Ts, EBSeq was used to analyze the differently expressed genes (DEGs) between PVTTs, Ts, and parenchyma tumor tissue (PT) from individual patients. Gene expression profiles of

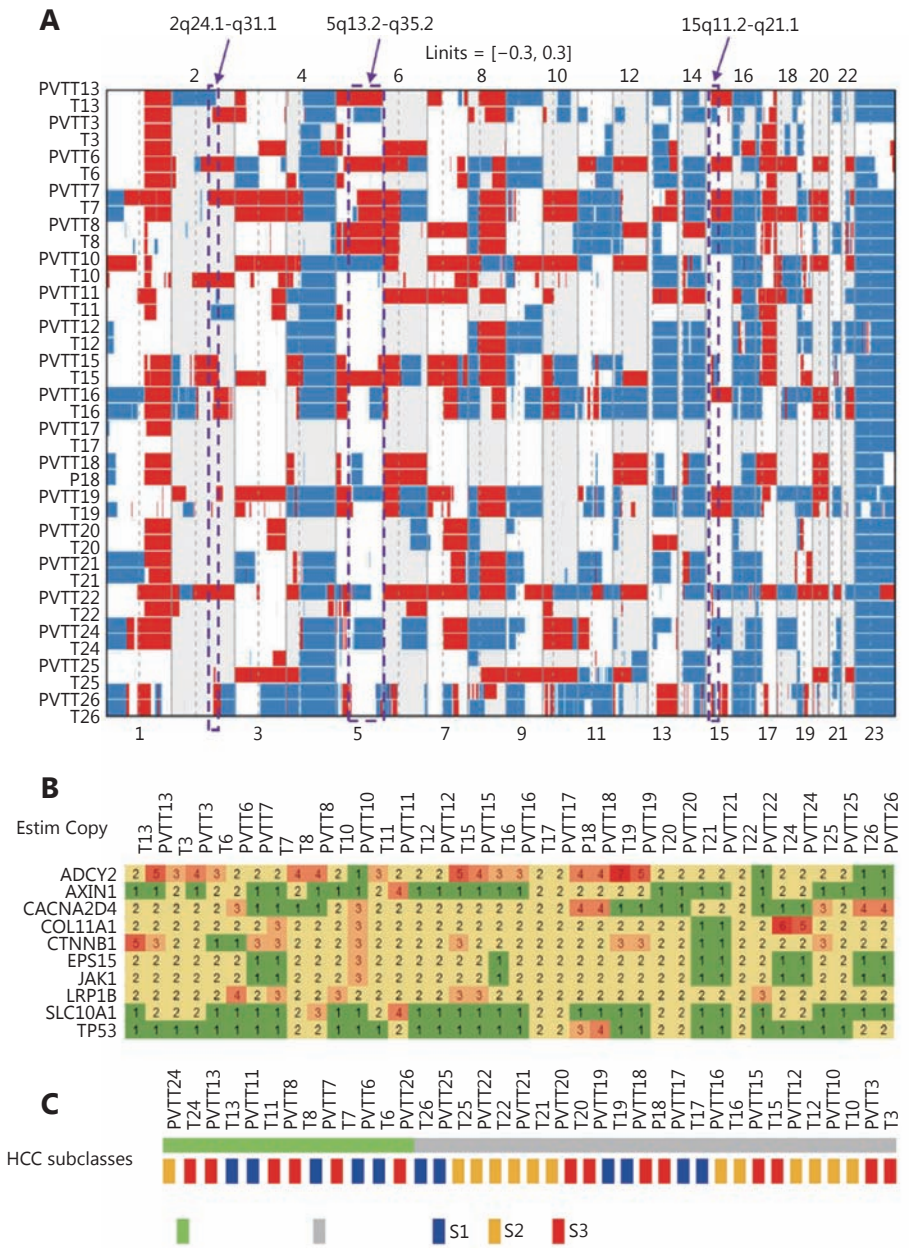

Figure 1 Analysis of the clonal relationship between PVTTs and corresponding primary Ts. (A) Copy number variations (CNVs) of chromosome regions in PVTTs and corresponding primary Ts. The specific CNV regions in PVTT13 and T13 were labeled by rectangles. The CNV threshold = 0.3; red, gain of CNV; blue, loss of CNV. (B) Copy numbers of HCC driver genes in PVTTs and Ts. (C) Molecular subtype of PVTTs and Ts according to Hoshida S1-S2-S3 classes.

these tissues (RNA-seq, GSE77509) were downloaded from the GEO database. Genes with fold change > 3 and FDR < 0.05 were identified as having significantly different expression. More than 2,000 DEGs were observed in 5 PVTT-T pairs including P13, whereas less than 1,000 DEGs were observed in the other pairs (**Figure 2A**). The Venn diagram in **Supplementary Figure S1A** shows that 512 and 3,257 DEGs were identified in PVTT13 and T13 compared with PT13, respectively. Among these 1,795 DEGs, 1,006 upregulated and 789 downregulated genes were identified as genes with specifically aberrant expression in PVTT13, and 1,540 DEGs with 595 upregulated and 945 downregulated genes were identified as genes with specifically aberrant expression in T13 (**Supplementary Figure S1B**). GO analysis showed that upregulated DEGs in PVTT13 were enriched in 15 biological processes (BPs) including the glutamate receptor signaling pathway, cell adhesion, and cell cycle (**Figure 2B**, **Supplementary Table S4**). Downregulated DEGs in PVTT13 were enriched in 16 BPs, including positive regulation of immune response, cytokine production, and peptidyl-tyrosine phosphorylation (**Figure 2B**, **Supplementary Table S4**). Upregulated DEGs in T13 were enriched in 15 BPs, including nuclear division, cell cycle checkpoints, and regulation of protein kinase activity (**Figure 2C**, **Supplementary Table S4**). Downregulated DEGs in T13 were enriched in 17 BPs, including inflammatory response, cell-cell adhesion, and negative regulation of cell proliferation (**Figure 2C**, **Supplementary Table S4**). KEGG enrichment analysis showed that upregulated DEGs in PVTT13 were enriched in pathways for nicotine addiction and type II diabetes mellitus, and The downregulated DEGs were enriched in signaling pathways, mainly PI3K-Akt, TGF-β, and p53 signaling (**Figure 2D**, **Supplementary Table S4**). The upregulated DEGs in T13 were enriched in signaling pathways, such as IL-17, p53 and cell cycle signaling, and the downregulated DEGs in T13 were enriched in signaling pathways including cGMP-PKG, Hippo and Rap1 (**Figure 2E**, **Supplementary Table S4**). The GO and KEGG analyses showed that DEGs in PVTT13 and T13 were enriched in different BPs and pathways. These data suggest that PVTT13 and T13 were different tumor types and that different mechanisms were involved in PVTT13 and T13 development.

Enrichment of hepatocyte gene signature in PVTT13

To further investigate the independent origin of PVTT13, enrichment of the gene signatures of the main cell types in liver tissues were analyzed using ssGSEA. The gene signatures of bile duct cells (BDCs), hepatocytes, hepatic progenitor cells (HPCs), and hepatic stellate cells (HSCs) were included due to their potential role in liver cancer[24].

The hepatocyte gene signature showed higher enrichment value in PVTT13 (7223.62) than that in T13 (5985.81), and the HPC gene signature showed lower enrichment value in PVTT13 (2608.81) than that in T13 (3212.07) (**Figure 3A**). Differences in the BDC and HSC gene signature enrichment values were not observed between PVTT13 and T13 (**Figure 3A**), whereas other PVTT-T pairs showed similar gene signature enrichment (**Figure 3A**). Further gene expression analysis showed a higher expression of hepatocyte-specific genes and lower expression of HSC-specific genes in PVTT13 than in T13 (**Figure 3B**). In addition, previous reports suggest that cancer stem cells (CSCs) and hematopoietic stem cells may be involved in PVTT development[8,12]. Our findings show that gene markers of these cells, such as *NANOG* and *C-kit*, were the nodes of the gene network or aberrantly expressed in PVTT13, further consolidating their potential roles in PVTT development (**Supplementary Figure S2**). Therefore, enrichment of gene signatures of hematopoietic stem cells and liver cancer stem cells were also investigated. A higher enrichment value of hematopoietic stem cell gene signatures was observed in PVTT13 (3114.07) than in T13 (1031.52), whereas a lower enrichment value of the $EpCAM^+$ liver cancer cell gene signature was observed in PVTT13 (–315.38) than in T13 (1401.61) (**Figure 3C**). Most PVTTs and Ts, except for PVTT6, PVTT11, PVTT19, T22, and pairs of P10, P16, and P25, had no enrichment of these signatures (**Figure 3C**). Gene expression analysis showed a higher expression of hematopoietic stem cell-specific genes, such as *CD34* and *KIT*, and a lower expression of $EpCAM^+$ liver cancer cell-specific genes, such as *EpCAM* and *CD44* in PVTT13 (**Figure 3D**). These data suggested that PVTT13 expressed high levels of hepatocyte- and hematopoietic stem cell-specific genes, and that these two cell types may be involved in PVTT13 development.

Gene signatures identifying PVTTs with no clonal relatedness with Ts

Having found that some PVTTs had no clonal relatedness with the corresponding Ts and that they have unique gene expression profiles, gene signatures identifying whether PVTTs have clonal relatedness with Ts may be very helpful for PVTT treatment. An integration analysis of CNV and gene expression profiles showed that 160 genes have differentially expressed copy number variations between PVTT13 and T13 (**Supplementary Figure S3**). The following PPI analysis found 24 genes to be the nodes in the network,

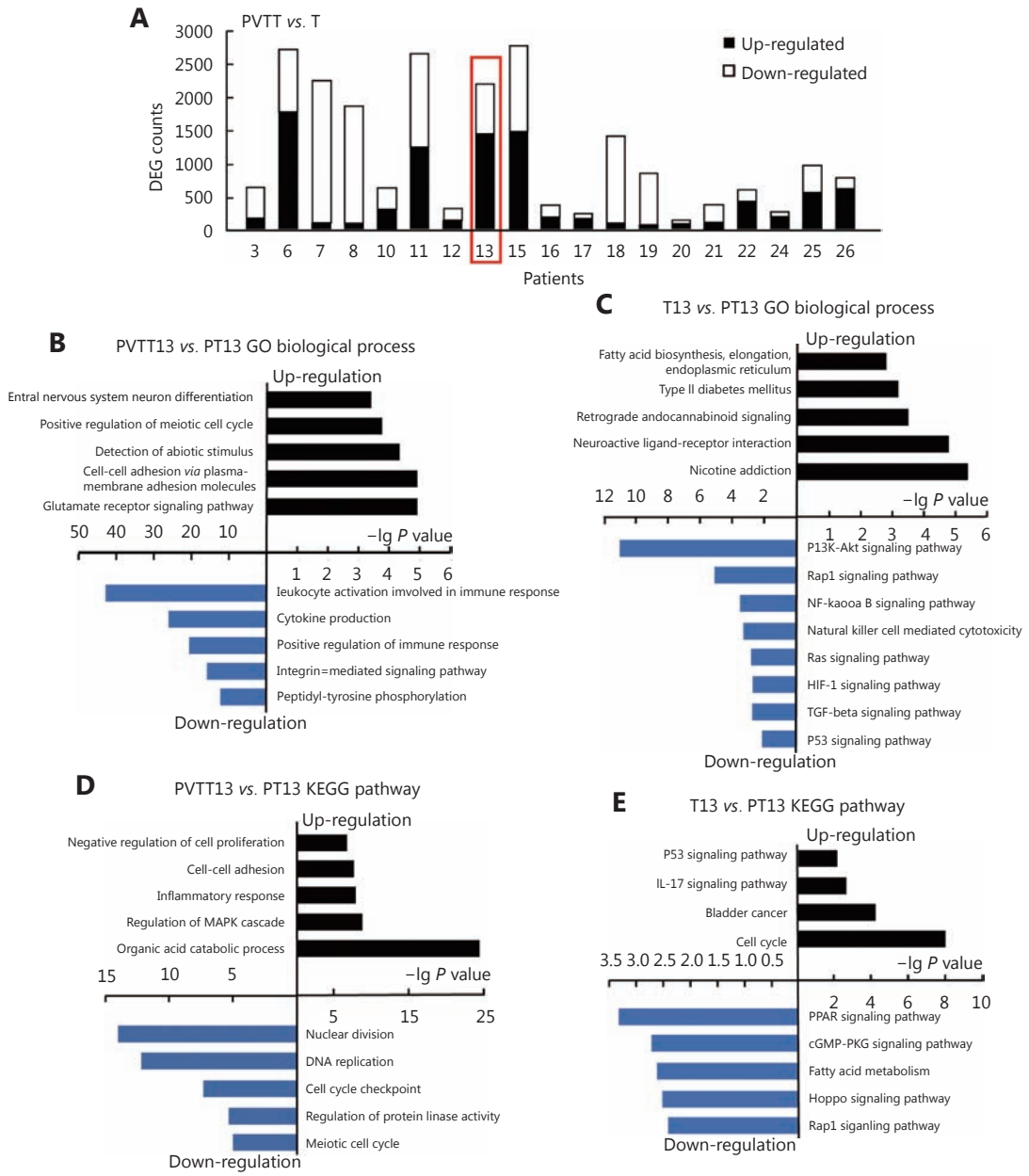

Figure 2 Analysis of aberrantly expressed genes in PVTT13 and T13. (A) Number of differently expressed genes between PVTTs and corresponding primary Ts. Top enriched biological processes in PVTT13 (B) and T13 (C) from GO enrichment analysis based on aberrantly expressed genes. Top enriched KEGG pathways in PVTT13 (D) and T13 (E) from KEGG pathway enrichment analysis based on aberrantly expressed genes.

suggesting that they may have crucial roles in PVTT13 development (**Supplementary Figure S3**). The ability of the 24-gene signature to identify PVTTs and Ts with no clonal relatedness was validated using three publicly available datasets with unsupervised clustering. In the dataset GSE77509, which includes 19 pairs of PVTT and Ts, the 24-gene signature classified PVTT13 and T13 into two different groups and classified other PVTTs and Ts from the same patient into the same group (**Figure 4A**). In the dataset GSE69164, including 11 pairs of PVTTs and Ts, PVTT9 and T9 were classified into two different groups, whereas other PVTTs and Ts from an individual patient were classified into the same group (**Figure 4B**). In the dataset GSE74656, which includes 5 pairs of PVTTs and Ts, all the PVTTs and Ts from an individual patient were classified into the same group (**Figure 4C**). In addition, the ability of the 24-gene signature

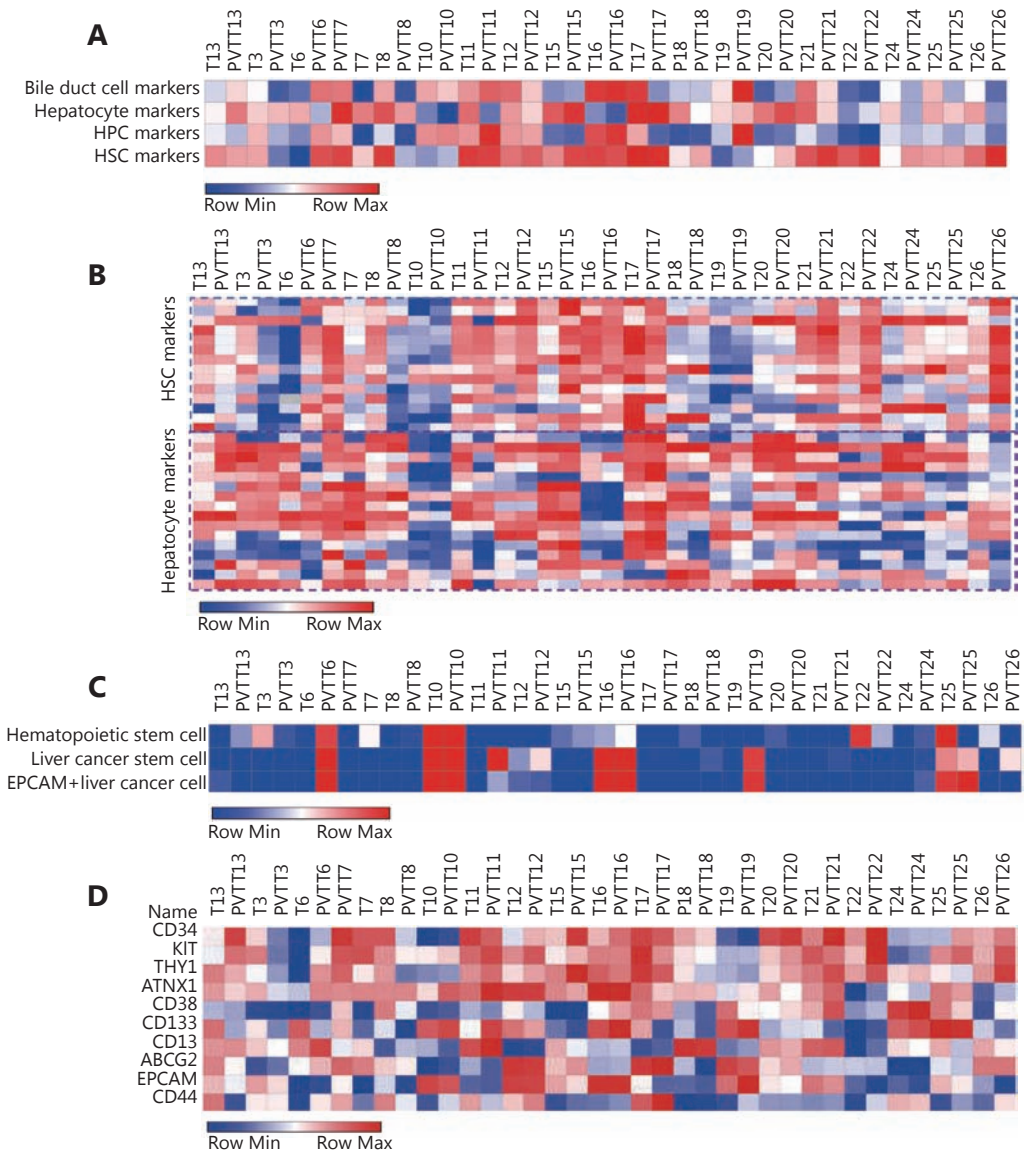

Figure 3 Enrichment analysis of gene signatures in PVTTs and Ts. (A) Enrichment of gene signatures of bile duct cells (BDC), hepatocytes, hepatic progenitor cells (HPC), and hepatic stellate cells (HSC) in PVTTs and Ts by ssGSEA based on their gene expression profiles. (B) Expression of the genes included in gene signatures of different liver cells in PVTTs and Ts from the dataset GSE77509. (C) Enrichment of gene signatures of hematopoietic stem cells and cancer stem cells in PVTTs and Ts by ssGSEA. (D) Expression of the genes included in gene signatures of stem cells in PVTTs and Ts from the dataset GSE77509.

to identify clonal relatedness was also validated using newly collected samples (n = 8). PVTT and T from patient 2 were, however, classified into two different groups (**Figure 4D**). All these data indicated that the 24-gene signature may be used to identify PVTTs and Ts with no clonal relatedness.

Discussion

In the present study, we found that PVTTs may have a different clonal origin than that of the corresponding HCC Ts. These types of PVTTs have different gene expression profiles from those of the corresponding Ts, whereas PVTTs originating from primary Ts have similar gene expression profiles to those of their origin Ts tissues. Previous studies reported the involvement of several factors, such as hypoxia and non-coding RNA in PVTT development, and similar gene expression profiles of PVTTs with Ts[7-10], suggesting that PVTT may be a special type of HCC intrahepatic metastasis.

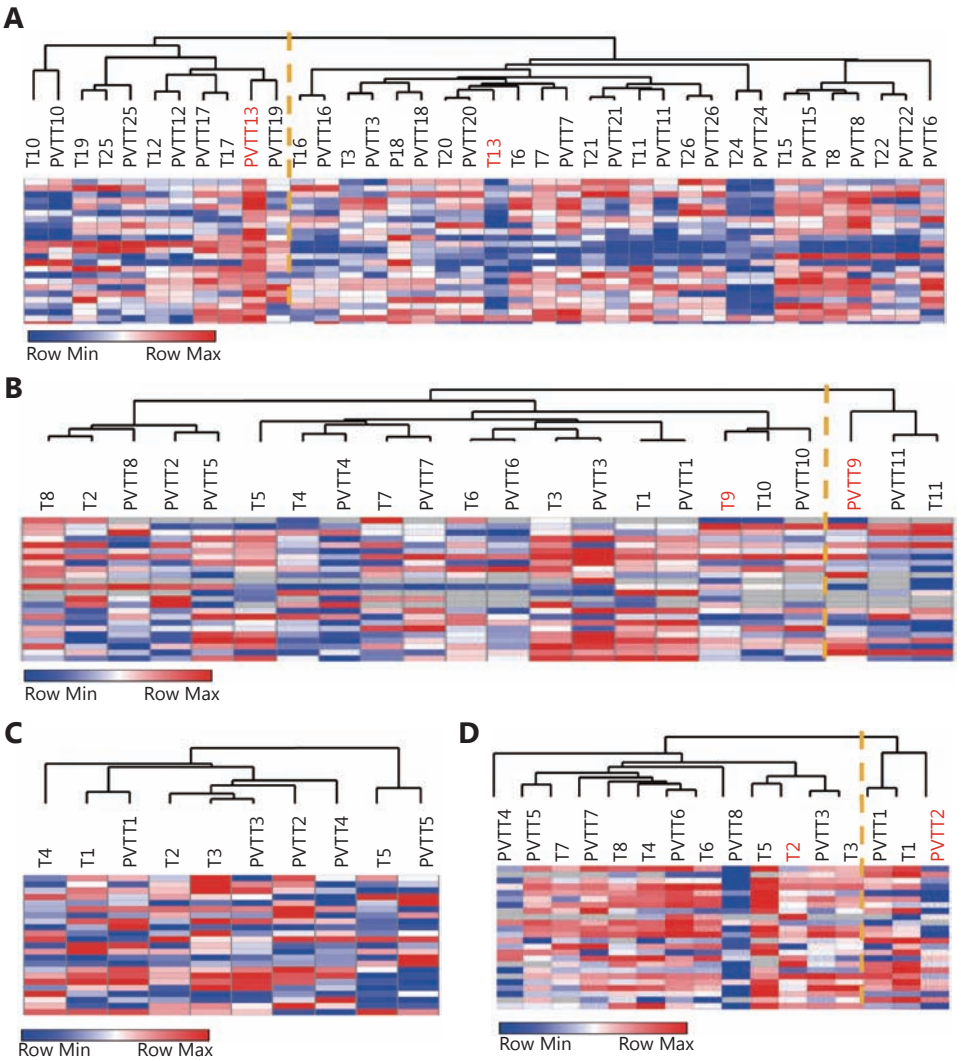

Figure 4 Validation of identification ability of 24-gene signature cluster analysis of pairs of PVTTs and Ts from GSE77509 (A), GSE69164 (B), GSE74656 (C), and newly collected samples (D) based on expression of 24 genes assessed by RNA-seq, cDNA microarray, or real time PCR. Orange line indicates the two subgroups classified based on the 24-gene signature.

Here, the clonality analysis, which was established to identify paired Ts with clonal origin, was applied to investigate the clonal relationship between PVTTs and Ts. Based on the clonality analysis and 24-gene signature identification, we found that most of the PVTTs (18 out of 19 in GSE77509, 10 out of 11 in GSE69164, and all 5 of GSE74656) possibly originate from the paired Ts (**Table 1**, **Figure 4**). Additionally, we also found another type of PVTTs, which received a very low LR2 value in the clonality analysis and were classified into a different group with the paired Ts in the cluster analysis (**Table 1**, **Figure 4**). These types of PVTTs were considered to represent tumor tissues having a different clonal origin than that of the paired Ts. These findings suggest the existence of two types of PVTTs according to their clonal origin, wherein one type has the same origin as the paired Ts and the other does not. Moreover, our findings can explain why PVTTs were observed in some patients without liver parenchyma tumor nodules[13]. The results can also explain why PVTTs from different patients showed very different gene expression profiles[11]. Of note is the observation that different CNV patterns were observed between PVTT13 and T13. These CNVs may be caused by gene translocation or genomic instability in the tumor cells. Genes transfer onto other chromosomes or extrachromosomal DNA (ecDNA) is amplified in tumor cells[25]. Genomic instability is a well-known characteristic of

tumor cells, inducing alterations in the genomes[26]. Hepatitis virus infection is another important factor that causes genomic instability in HCC[27]. Hence, the mechanism underlying the different CNV patterns between PVTTs and Ts needs more investigation.

Our study suggests that integrated analysis of the published datasets provides a better understanding of Ts, and that such analyses may lead to improved therapeutic strategies. With the development of high-throughput technology and multi-omics analyses, increasingly more omics data from Ts are uploaded into the public databases such as GEO and The Cancer Genome Atlas (TCGA)[28]. An integrated analysis of these datasets makes it possible to obtain a more comprehensive understanding of Ts at multidimensional levels. The comprehensive integrated analysis of HCC using six distinct data platforms determined subtypes of HCC associated with poorer prognosis, and the gene expression signatures correlated with poor survival and potential therapeutic targets for HCC[29]. Another multiplatform integrative analysis based on TCGA found that molecular signatures provide a higher accuracy for the clinical outcome prediction than that of the currently used tissue-of-origin-based classification in some cases[30]. By integrated analysis of CNV and gene expression profiles previously used to investigate aberrantly expressed lncRNAs[11], we analyzed the origin of PVTTs and developed a gene signature to identify the type of PVTTs. It has been reported that the "cell-of-origin"-based features mainly dominate the molecular taxonomy of tumor types[30]. Therefore, the gene signatures developed in our study may provide a platform for a high-accuracy identification of PVTTs with independent origins. However, follow-up studies and additional samples are necessary to validate the findings reported here.

Acknowledgements

This work was supported by grants of China National Funds for Distinguished Young Scientists (Grant No. 81425019), National Natural Science Foundation of China (Grant No. 81672899), the State Key Program of National Natural Science Foundation of China (Grant No. 81730076), Shanghai Science and Technology Committee Program (Grant No. 18XD1405300) and Specially-Appointed Professor Fund of Shanghai (Grant No. GZ2015009).

Conflict of interest statement

No potential conflicts of interest are disclosed.

References

1. Zhang ZM, Lai ECH, Zhang C, Yu HW, Liu Z, Wan BJ, et al. The strategies for treating primary hepatocellular carcinoma with portal vein tumor thrombus. Int J Surg. 2015; 20: 8-16.
2. Kokudo T, Hasegawa K, Matsuyama Y, Takayama T, Izumi N, Kadoya M, et al. Survival benefit of liver resection for hepatocellular carcinoma associated with portal vein invasion. J Hepatol. 2016; 65: 938-43.
3. Sakamoto K, Nagano H. Surgical treatment for advanced hepatocellular carcinoma with portal vein tumor thrombus. Hepatol Res. 2017; 47: 957-62.
4. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology. 2011; 53: 1020-2.
5. Cheng SQ, Chen MS, Cai JQ, National Research Cooperative Group for Diagnosis, Treatment of Hepatocellular Carcinoma with Tumor Thrombus. Chinese expert consensus on multidisciplinary diagnosis and treatment of hepatocellular carcinoma with portal vein tumor thrombus: 2016 edition. Oncotarget. 2017; 8: 8867-76.
6. Tang YF, Liu SP, Li N, Guo WX, Shi J, Yu HM, et al. 14-3-3ζ promotes hepatocellular carcinoma venous metastasis by modulating hypoxia-inducible factor-1α. Oncotarget. 2016; 7: 15854-67.
7. Ye LY, Chen W, Bai XL, Xu XY, Zhang Q, Xia XF, et al. Hypoxia-induced epithelial-to-mesenchymal transition in hepatocellular carcinoma induces an immunosuppressive tumor microenvironment to promote metastasis. Cancer Res. 2016; 76: 818-30.
8. Zhang J, Pan YF, Ding ZW, Yang GZ, Tan YX, Yang C, et al. Rmp promotes venous metastases of hepatocellular carcinoma through promoting il-6 transcription. Oncogene. 2015; 34: 1575-83.
9. Guo WX, Liu SP, Cheng YQ, Lu L, Shi J, Xu GX, et al. ICAM-1-related noncoding rna in cancer stem cells maintains ICAM-1 expression in hepatocellular carcinoma. Clin Cancer Res. 2015; 22: 2041-50.
10. Ye QH, Qin LX, Forgues M, He P, Kim JW, Peng AC, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nat Med. 2003; 9: 416-23.
11. Yang Y, Chen L, Gu J, Zhang HS, Yuan JP, Lian QY, et al. Recurrently deregulated lncRNAs in hepatocellular carcinoma. Nat Commun. 2017; 8: 14421.
12. Guo WX, Xue J, Shi J, Li N, Shao Y, Yu XY, et al. Proteomics analysis of distinct portal vein tumor thrombi in hepatocellular carcinoma patients. J Proteome Res. 2010; 9: 4170-5.
13. Saito M, Seo Y, Yano Y, Uehara K, Hara S, Momose K, et al. Portal venous tumor growth-type of hepatocellular carcinoma without liver parenchyma tumor nodules: a case report. Ann Hepatol. 2013; 12: 969-73.
14. Zhang H, Ye JY, Weng XL, Liu FT, He L, Zhou DZ, et al. Comparative transcriptome analysis reveals that the extracellular matrix receptor interaction contributes to the venous metastases of hepatocellular carcinoma. Cancer Genet. 2015; 208: 482-91.

15. Nemes S, Danielsson A, Parris TZ, Jonasson JM, Bülow E, Karlsson P, et al. A diagnostic algorithm to identify paired tumors with clonal origin. Genes Chromosomes Cancer. 2013; 52: 1007-16.
16. Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, Walker DA, et al. Tracking the clonal origin of lethal prostate cancer. J Clin Invest. 2013; 123: 4918-22.
17. Tripathi S, Pohl MO, Zhou YY, Rodriguez-Frandsen A, Wang GJ, Stein DA, et al. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe. 2015; 18: 723-35.
18. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. Genepattern 2.0. Nat Genet. 2006; 38: 500-1.
19. Ceulemans A, Verhulst S, Van Haele M, Govaere O, Ventura JJ, Van Grunsven LA, et al. Rna-sequencing-based comparative analysis of human hepatic progenitor cells and their niche from alcoholic steatohepatitis livers. Cell Death Dis. 2017; 8: e3164.
20. Hoshida Y, Nijman SMB, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res. 2009; 69: 7385-92.
21. Sia D, Jiao Y, Martinez-Quetglas I, Kuchuk O, Villacorta-Martin C, Castro De Moura M, et al. Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features. Gastroenterology. 2017; 153: 812-26.
22. Song FF, Li XC, Song FJ, Zhao YR, Li HX, Zheng H, et al. Comparative genomic analysis reveals bilateral breast cancers are genetically independent. Oncotarget. 2015; 6: 31820-9.
23. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007; 23: 657-63.
24. Sia D, Villanueva A, Friedman SL, Llovet JM. Liver cancer cell of origin, molecular class, and effects on patient prognosis. Gastroenterology. 2017; 152: 745-61.
25. Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature. 2017; 543: 122-5.
26. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144: 646-74.
27. Wang M, Xi D, Ning Q. Virus-induced hepatocellular carcinoma with special emphasis on HBV. Hepatol Int. 2017; 11: 171-80.
28. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol. 2015; 19: A68-77.
29. The Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell. 2017; 169: 1327-1341.e23.
30. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell. 2014; 158: 929-44.

Cite this article as: Liu S, Zhou Z, Jia Y, Xue J, Liu Z, Cheng K, et al. Identification of portal vein tumor thrombus with an independent clonal origin in hepatocellular carcinoma via multi-omics data analysis. Cancer Biol Med. 2019; 16: 147-56. doi: 10.20892/j.issn.2095-3941.2018.0184

Supplementary materials

Table S1 Gene set used in enrichment analysis

Molecular Signature Database gene sets	Description	Genes
HOSHIDA_LIVER_CANCER_SUBCLASS_S3	>Genes from 'subtype S3' signature of hepatocellular carcinoma (HCC): hepatocyte differentiation.	ABCB4, ABCC6, ABHD2, ABP1, ACAA2, ACADM, ACADS, ACADSB, ACADVL, ACO1, ACOX1, ACOX2, ACSL1, ACY1, ADH4, ADH6, ADK, AGL, AGXT, AKR1C1, ALAS1, ALDH1A1, ALDH1B1, ALDH2, ALDH3A2, ALDH4A1, ALDH6A1, ALDH7A1, ALDOB, ALPL, AMFR, AMT, ANXA6, APCS, APOA1, APOC2, APOC4, APOH, AQP7, ARG1, ARHGEF12, ARSA, ASCL1, ASGR1, ASGR2, ASL, ASS1, ATOX1, ATP5D, ATP5J, AZGP1, BAAT, BDH1, BHMT, BLOC1S1, BLVRB, BPHL, BTD, C1R, C1S, C4A, C4BPA, C8B, CA2, CAT, CBR1, CD14, CD302, CD81, CECR1, CES1, CFB, CFH, CGREF1, CNGA1, COL18A1, COX5B, CP, CPA3, CPA4, CPB2, CPS1, CRABP1, CRYAA, CRYM, CSTB, CTH, CTSO, CXCL2, CYB5A, CYFIP2, CYP21A2, CYP27A1, CYP2C9, CYP2J2, CYP3A7, DAO, DCAF8, DECR1, DNASE1L3, DPAGT1, DRG2, ECHS1, ECI1, EDNRB, EGFR, EHHADH, EMP2, EPAS1, EPHX1, ETS2, F11, F2, F5, FAH, FANCA, FGB, FGG, FH, FKBP2, FLT4, FMO4, FOXO1, FXR2, GCH1, GCHFR, GCKR, GGH, GHR, GJB1, GLYAT, GOT2, GPT, GPX2, GPX3, GSTA2, GSTO1, GSTZ1, HAAO, HADH, HGD, HMGCS2, HMOX2, HPD, HRG, HRSP12, HSD17B10, HSD17B4, ICAM3, IDH2, IDH3A, IFIT1, IGF1, IL13RA1, IL32, IL6R, IMPA1, INSR, IQGAP2, ISG15, ITIH1, ITIH3, ITIH4, ITPR2, IVD, KCNJ8, KLKB1, KMO, KNG1, LCAT, LONP1, LPIN1, LPIN2, MAOA, MAOB, MAPRE3, MGST2, MME, MSMO1, MT2A, MTHFD1, MTHFS, MUT, MYLK, MYO1E, NDUFV2, NFIB, NFIC, NFKBIA, NNMT, NRG1, PAH, PAPSS2, PCCA, PCCB, PCK1, PCK2, PDK4, PGM1, PGRMC1, PIK3R1, PKLR, PLA2G2A, PLCG2, PLG, PLGLB2, PNPLA4, POLD4, PON3, PPP2R1B, PROS1, PTGR1, PTS, QDPR, RARRES2, RBP5, RGN, RHOB, RNASE4, SBDS, SDC1, SDHB, SDS, SELENBP1, SEPP1, SERPINA3, SERPINA6, SERPINC1, SERPING1, SHB, SHMT1, SLC10A1, SLC16A2, SLC23A1, SLC23A2, SLC2A2, SLC35D1, SLC6A1, SLC6A12, SLC7A2, SLC9A3R2, SLCO2A1, SLPI, SMARCA2, SOAT1, SOD1, SOD2, SORL1, SPAM1, SPARCL1, SRD5A1, SREBF1, SULT2A1, TCEA2, TDO2, TGFBR3, TINAGL1, TJP2, TMBIM6, TMOD1, TOB1, TPMT, TST, UQCRB, VSIG2, ZNF160
HOSHIDA_LIVER_CANCER_SUBCLASS_S2	>Genes from 'subtype S2' signature of hepatocellular carcinoma (HCC): proliferation, MYC and AKT1 [GeneID=4609;207] activation.	ABCB10, ABCD3, ACP1, ADD3, AFP, AHCY, ARHGAP35, ARID3A, ATF2, ATM, ATP2B1, ATP2B2, ATP5F1, ATXN10, BCAM, BCLAF1, BRD3, BTG3, C5orf13, CASC3, CD46, CDK6, CHKA, CLK2, COL2A1, CPD, CSE1L, CSNK2A1, CSNK2A2, CTNNB1, CUL4A, CXADR, DDX1, DDX18, DEK, EIF4A2, EIF4B, ENPP1, EP300, ERBB3, FBL, FGFR3, FGFR4, FLNB, GBF1, GCN1L1, GLUD1, GNAI1, GPC3, GTF2I, GTF3C2, H1F0, HELZ, HMGCR, HNRNPA2B1, HNRNPC, HNRNPU, IDI1, IGF2, IGF2R, ITIH2, KLF3, LBR, MAPK6, MEST, NCOA4, NET1, NR2C1, NR5A2, NT5E, NUP153, PEG3, PHF3, PHKA2, PIGC, PLXNB1, PNN, POFUT1, PPARG, PPP2R1A, PRDX3, PTOV1, RAB4A, RBM39, RPL24, RPL27, RPL31, RPS19, RPS24, RPS25, RPS27, RPS5, RRP1B, SEPHS1, SLC6A2, SLC6A5, SMARCA1, SMARCC1, SNRPE, SNTB1, SREBF2, SSB, SUMO1, SUZ12, TARBP1, TBCE, TFIP11, TIA1, TIAL1, TM9SF4, TP53BP2, TPR, TRIM26, TTC3, UBE2K
HOSHIDA_LIVER_CANCER_SUBCLASS_S1	>Genes from 'subtype S1' signature of hepatocellular carcinoma (HCC): aberrant activation of the WNT signaling pathway.	ACP5, ACTA2, ADAM15, ADAM8, ADAM9, AEBP1, AIF1, AKT3, ALDOA, ALOX5AP, ANXA1, ANXA4, ANXA5, AP2S1, AQP1, ARF5, ARHGDIB, ARPC1B, ARPC2, ASAH1, ATP1B3, ATP6AP1, ATP6V0B, ATP6V1B2, ATP6V1F, BCL2A1, BLVRA, C1QB, C3AR1, CAPZA1, CBFB, CCL3, CCL5, CCND2, CCR7, CD151, CD37, CD3D, CD47, CD48, CD53, CD74, CD8A, CDC20, CDC25B, CDH11, CDK2AP1, CELF2, CHN1, COL11A1, COL15A1, COL1A2, COL3A1, COL4A1, COL4A2, COL5A2, COL6A1, COL6A2, CORO1A, CRABP2, CRIP2, CSRP3, CTGF, CTSC, CTSS, CXCL1, CXCR4, CYBA, CYBB, CYFIP1, CYP1B1, CYR61, DAB2, DCTN2, DDR1, DDR2, DDX11, DGKA, DGKZ, DNM2, DPYSL2, DUSP5, DUT, EFEMP1, EFNB1, F13A1, FBN1, FBRS, FCGBP, FCGR2A, FGL2, FHL3, FLNA, FUT4, FYB, GEM, GLIPR1, GNAI2, GNS, GPNMB, GRN, GSTP1, GUCY1A3, GYPC, HCLS1, HEXA, HIF1A, HK1, HLA-DMA, HLA-DOA, HLA-DPB1, HLA-DQA2, HLA-DQB1, ID3, IER3, IFI16, IFI30, IGFBP5, IGH@, IGKC, IGLL1, IKBKE, IL15RA, IL2RB, IL2RG, IL7R, IQGAP1, IRF1, ITGB2, ITPR3, KIAA0101, KLC1, KLF5, LAMB1, LAPTM5, LCP1, LDHB, LGALS1, LGALS3BP, LGALS9, LGMN, LHFPL2, LITAF, LMO4, LOX, LSP1, LTBP2, LTBP3, LTF, LUM, LYN, M6PR, MAP1B, ME1, MFAP1, MGP, MPHOSPH6, MSN, MTHFD2, MYCBP2, NBL1, NPC2, NSMAF, OAZ1, PAK1, PAM, PAPSS1, PEA15, PFN1, PGK1, PIM2, PKD2, PKMYT1, PKN1, PLAUR, PLD3, PLP2, PNP, POLD3, POSTN, PPIC, PPP1CB, PPP4C, PPP4R1, PRKD2, PRMT5, PROCR, PSMD2, PTPRC, PYGB, QSOX1, RAB31, RALB, RALGDS, RCC1, RHOA, RIN2, RIT1, RNASE1, RNASE6, RPA2, RSU1, S100A10, S100A11, S100A13, SLA, SLBP, SLC1A5, SLC2A1, SLC2A5, SLC39A6, SLC7A5, SMAD2, SMARCD1, SPAG8, SRGN, SRI, STK38, STX3, TAGLN, TAX1BP3, TCF4, TFF3, THY1, TIMP2, TMSB4X, TNFRSF1B, TP53BP1, TPM2, TRA@, TRAF3, TRAF5, TRIP10, TSPAN3, TUBA4A, VCAN, ZNF384

Continued

Continued

Molecular Signature Database gene sets	Description	Genes
JAATINEN_ HEMATOPOIETIC_ STEM_CELL_UP	>Genes up-regulated in CD133+ [GeneID=8842] cells (hematopoietic stem cells, HSC) compared to the CD133- cells.	ABCC1, ACACA, ACSF2, ADAM28, ADAT2, AKR1A1, AKR1C3, AKT3, ALDH1A1, ALG10B, ANGPT1, ANKRD28, ANKRD6, AREG, AREGB, ARHGAP22, ARMCX1, ARMCX2, ASPH, ATP2C1, ATP6V0A2, ATP9A, B4GALT6, BAALC, BCAT1, BCL11A, BEND4, BEX2, BIVM, BSPRY, BTBD3, C10orf58, C11orf54, C11orf95, C12orf24, C19orf77, C1QTNF4, C3orf64, C5orf13, C5orf35, C9orf93, CALN1, CBX2, CBX5, CCDC6, CCNB1IP1, CD109, CD34, CDCA7, CDK2AP1, CDK4, CDK6, CEP170, CHRDL1, CHST13, CMAHP, CNKSR3, COL24A1, COL5A1, CPA3, CPSF3, CPT1A, CREBZF, CRHBP, CRIM1, CRISPLD1, CTHRC1, CYTL1, CYYR1, DAPK1, DDAH1, DEPTOR, DKC1, DNMT3B, DOCK7, DPPA4, DPY19L2, DPY19L3, DPY19L4, DPYSL3, DSG2, DST, DTL, EBPL, EFHA2, EMP1, ERG, ERLIN2, ERMP1, ETV6, F2R, F2RL1, FABP5, FAIM, FAM115A, FAM171B, FAM175A, FAM69B, FAM92A1, FAM98B, FANCL, FGD5, FHL1, FLT3, FLVCR1, FRMD4B, FRMD6, GALNT7, GATA2, GATM, GCSH, GNA15, GNAI1, GOLIM4, GPR125, GPR126, GUCY1A3, H2AFY, HADH, HDGFRP3, HHEX, HLF, HMGA2, HMGCR, HOXA10, HOXA3, HOXA5, HOXA9, HOXB3, HPGDS, HSH2D, HSPB1, HSPD1, IGFBP7, IGLL1, IL18, IPO11, IPO7, IQGAP2, ISYNA1, ITGA9, ITM2C, ITPRIPL2, JUP, KCTD15, KCTD3, KDELC1, KDM5B, KHDRBS3, KIAA0125, KIAA0368, KIAA1211, KIT, LAPTM4B, LIMCH1, LOC100287017, LOC100506844, LOC400464, LPIN1, LRCH2, MAP7, MAP9, MAST4, MBLAC2, MCM5, ME3, MEG3, MEGF6, MEIS1, MEST, MIR155HG, MLLT3, MMP28, MPL, MRPS27, MSI2, MSRB3, MUM1, MYB, MYCT1, MYO5C, MZB1, NAP1L3, NASP, NDN, NEGR1, NKAIN2, NKX2-3, NME1, NPR3, NRIP1, NT5DC2, OBSL1, PAICS, PAIP1, PAM, PDE1A, PDGFC, PDZD2, PGBD1, PHACTR1, PHF16, PKD2, PLA2G12A, PLA2G4A, PLAGL1, PLCB1, PLCB4, PLEKHA5, PLS3, PM20D2, POGZ, PON2, PPM1H, PRDM16, PRDX4, PREX2, PRMT5, PROM1, PSMB5, PSMG1, PTPLA, PXDN, QSER1, RAB34, RAVER2, RBM10, RBPMS, RCN1, RDX, RHOBTB1, RPS15A, RUNX2, RUVBL2, SAMD13, SCD, SCHIP1, SCN3A, SCN9A, SCRN1, SEPP1, SERPING1, SH3BP4, SH3RF1, SHANK3, SLC16A14, SLC25A27, SLC27A2, SLC39A10, SLC39A8, SLITRK4, SMAD1, SMARCA1, SMARCA2, SMYD3, SOCS2, SOCS6, SPG20, SPIN4, SPINK2, SPOCK3, SRD5A3, SRSF3, SSBP2, STK3, STMN1, SV2A, SYPL1, TAF1D, TANC1, TBC1D24, TCEAL4, TFEC, TFPI, TGFBRAP1, TMEM163, TMEM200A, TMEM38B, TMEM44, TMEM5, TNFRSF21, TNS3, TRIM24, TRIM73, TRIP6, TRO, TSPYL5, TUG1, TUSC1, TXNRD3, U BR5, UBTD2, UHRF1, UMODL1, UNG, VAV3, VWDE, WASF1, WBP5, WDR17, WDR49, WDR91, ZBED3, ZBTB8A, ZC3HAV1L, ZMAT1, ZNF117, ZNF165, ZNF302, ZNF512B, ZNF521, ZNF618, ZNF709, ZNF711, ZNFX1-AS1, ZNRF1
YAMASHITA_LIVER_ CANCER_STEM_ CELL_UP	>Genes up-regulated in hepatocellular carcinoma (HCC) cells with hepatic stem cell properties.	ARHGEF15, ARPC5L, ASAP1, BACE2, BAK1, CCNB2, CDKN2A, CLDN1, CLDN2, CLDN4, CSNK1E, CYCS, DDR1, DKK1, DUSP1, ELF3, ESR1, FGFR3, FZD5, GADD45A, GRB7, GYS2, HNF4A, IGF1, ITGB5, LINGO1, MAP2K2, MAP3K5, MAPK13, MERTK, MMP11, MYLK, PDGFRB, PIK3C2G, PIK3C3, PLCB1, PRKCD, RARG, ROBO2, SEMA3C, SEMA4G, SMAD5, SOX9, TCF7L2, UBD, WEE1, YWHAZ
HPC marker	>HPC	KRT19, KRT8, CLDN4, KRT17, TACSTD2, EPCAM, ITGB1, SOX9, DDR1, PROM1, ID4, TPBG, ABCC4, ALDH1A2, HNF1B
Hepatocyte marker	>hepatocyte	CYP3A4, CYP2E1, ALB, TTR, APOC1, B2M, CYP2C9, CYP2C19, UGT2B7, ABCB1, ABCC2, ABCG2, CYP1A1, CYP1A2, HNF4A, HMGCS2, SLC2A2
HSC marker	>HSC	VIM, PDGFRB, APOB, PDGFD, ACTA2, ACTG2, COL1A1, COL1A2, COL3A1, COL4A1, LOX, LOXL1, LOXL2, NTM
Bile duct cell markers	>bile duct cell	KRT7, KRT19, SOX9, CFTR, ASBT, SCTR, EPCAM, AQP1, TGR5, GGT1, JAG1

Table S2 The primer sequence used in Real-time PCR

Gene symbol		Sequence （5′ → 3′）
RNF14	Forward Primer	ACTTTGTGCAGGTTGACCTAC
	Reverse Primer	CGCTTGCAGGTATTCATTTCGTA
GABRA1	Forward Primer	AGCCGTCATTACAAGATGAACTT
	Reverse Primer	TGGTCTCAGGCGATTGTCATAA
NDST1	Forward Primer	CTGCCTGTTCAGCGTTTTCAT
	Reverse Primer	CGAGTAGAGGCTCTCCACAAA
FBXW11	Forward Primer	GGAACATCATCTGTGATCGTCTC
	Reverse Primer	TGGTAAAGCGGTAATAAAGTCCC
TJP1	Forward Primer	CAACATACAGTGACGCTTCACA
	Reverse Primer	CACTATTGACGTTTCCCCACTC
HSPA9	Forward Primer	CTTGTTTCAAGGCGGGATTATGC
	Reverse Primer	GCAGGAGTTGGTAGTACCCAAA
CLINT1	Forward Primer	CGAGAGGCAACGAACGATGAT
	Reverse Primer	CCTTATGAGGTAAGCTAGGAGCA
THBS4	Forward Primer	TGCTGCCAGTCCTGACAGA
	Reverse Primer	GTTTAAGCGTCCCATCACAGTA
GABRG2	Forward Primer	ACTTCGGCCTGATATAGGAGTG
	Reverse Primer	ACGTCTGTCATACCACGTTTG
FGF1	Forward Primer	GACAGGAGCGACCAGCACATTC
	Reverse Primer	TCCAGCCTTTCCAGGAACAAACATT
APC	Forward Primer	AAAATGTCCCTCCGTTCTTATGG
	Reverse Primer	CTGAAGTTGAGCGTAATACCAGT
SAR1B	Forward Primer	TACAGTGGTTTCAGCAGTGTG
	Reverse Primer	AGTGGGATGTAATGTTGGGACA
GRIA1	Forward Primer	TGGGTTTTATGAACGTAGGACTG
	Reverse Primer	AAAGCTCGGCGTAATGAAGCA
FGF18	Forward Primer	ACTTGCCTGTGTTTACACTTCC
	Reverse Primer	GACCTGGATGTGTTTCCCACT
ACTC1	Forward Primer	TCCCATCGAGCATGGTATCAT
	Reverse Primer	GGTACGGCCAGAAGCATACA
F2RL1	Forward Primer	CAGTGGCACCATCCAAGGAA
	Reverse Primer	CAGGGCCATGCCGTTACTT
GABRB3	Forward Primer	GATAAAAGGCTCGCCTATTCTGG
	Reverse Primer	GATCATGCGGTTTTTCACTGTC
CHST14	Forward Primer	CGCCACATCCTCGTAAGTGAC
	Reverse Primer	CCCGCTTCCAGTTAGAGCA
ADRB2	Forward Primer	TTGCTGGCACCCAATAGAAGC
	Reverse Primer	CAGACGCTCGAACTTGGCA
OTUD7A	Forward Primer	GTGTTGGGCAGCACTTCTACA
	Reverse Primer	CGTGGACCGAACAAAGTCTG
ADRA1B	Forward Primer	TGGGGCGGATCTTCTGTGA
	Reverse Primer	GTGACCAGCGTGGGATACTG
SYNPO	Forward Primer	ATGGAGGGGTACTCAGAGGAG
	Reverse Primer	CTCTCGGTTTTGGGACAGGTG
GABRG3	Forward Primer	AACCGACCGTAATTGACGTTG
	Reverse Primer	CTGTCCAGGTCTGAGCAAAAA
OCLN	Forward Primer	ACAAGCGGTTTTATCCAGAGTC
	Reverse Primer	GTCATCCACAGGCGAAGTTAAT
β-actin	Forward Primer	CATGTACGTTGCTATCCAGGC
	Reverse Primer	CTCCTTAATGTCACGCACGAT

Table S3 The 24 genes set identifying pairs of PVTTs and Ts with no clonal relationship

Gene symbol	Gene names
RNF14	Ring finger protein 14
GABRA1	Gamma-aminobutyric acid type A receptor alpha1 subunit
NDST1	N-deacetylase and N-sulfotransferase 1
FBXW11	F-box and WD repeat domain containing 11
TJP1	Tight junction protein 1
HSPA9	Heat shock protein family A (Hsp70) member 9
CLINT1	Clathrin interactor 1
THBS4	Thrombospondin 4
GABRG2	Gamma-aminobutyric acid type A receptor gamma2 subunit
FGF1	Fibroblast growth factor 1
APC	APC, WNT signaling pathway regulator
SAR1B	Secretion associated Ras related GTPase 1B
GRIA1	Glutamate ionotropic receptor AMPA type subunit 1
FGF18	Fibroblast growth factor 18
ACTC1	Actin, alpha, cardiac muscle 1
F2RL1	F2R like trypsin receptor 1
GABRB3	Gamma-aminobutyric acid type A receptor beta 3 subunit
CHST14	Carbohydrate sulfotransferase 14
ADRB2	Adrenoceptor beta 2
OTUD7A	OUT deubiquitinase 7A
ADRA1B	Adrenoceptor alpha 1B
SYNPO	Synaptopodin
GABRG3	Gamma-aminobutyric acid type A receptor gamma 3 subunit
OCLN	Occludin

Table S4 KEGG pathways and GO biological processes enriched in PVTT13 and T13. Only KEGG pathways were displayed. The full Table S4 can be found in the online version of this article

Up-regulated pathways in PVTT13

Pathway ID	Description	%	type	Count	P Value	genes
hsa05033	Nicotine addiction	1	up	10	3.98E-06	CACNA1B, GABRA1, GABRA4, GABRG2, GABRG3, GABRR1, GRIA1, GRIA4, GRIN2B, SLC17A6
hsa04080	Neuroactive ligand-receptor interaction	2.8	up	28	1.58E-05	ADRA2C, AVPR1B, BDKRB1, CHRNB3, CRHR2, GABRA1, GABRA4, GABRG2, GABRG3, GABRR1, GLP1R, GLRA1, MCHR1, GRIA1, ……
hsa04723	Retrograde endocannabinoid signaling	1.3	up	13	0.00031623	ADCY2, CACNA1B, GABRA1, GABRA4, GABRG2, GABRG3, GABRR1, GRIA1, GRIA4, KCNJ9, RIMS1, PLCB1, SLC17A6
hsa04930	Type II diabetes mellitus	0.81	up	8	0.00063096	CACNA1B, CACNA1E, IKBKB, KCNJ11, PKLR, ABCC8, ADIPOQ, MAFA
hsa04727	GABAergic synapse	1.1	up	11	0.001	ADCY2, CACNA1B, GABRA1, GABRA4, GABRG2, GABRG3, GABRR1, GAD2, HAP1, GABBR2, SLC12A5
hsa_M00415	Fatty acid biosynthesis, elongation, endoplasmic reticulum	0.41	up	4	0.00158489	HACD3, ELOVL2, ELOVL7, HACD2
hsa05032	Morphine addiction	1.1	up	11	0.00158489	ADCY2, CACNA1B, GABRA1, GABRA4, GABRG2, GABRG3, GABRR1, KCNJ9, OPRM1, PDE1C, GABBR2
hsa04724	Glutamatergic synapse	1.2	up	12	0.00316228	ADCY2, GRIA1, GRIA4, GRIN2B, GRM3, GRM4, SLC1A6, HOMER1, PLCB1, SLC17A6, PLA2G4E, PLA2G4D
hsa00564	Glycerophospholipid metabolism	1	up	10	0.00630957	CDS1, PLA2G1B, DGKI, LPGAT1, LYPLA1, AGPAT4, GPAM, PLA2G2F, PLA2G4E, PLA2G4D
hsa04911	Insulin secretion	0.91	up	9	1.00E-02	ADCY2, GCG, GLP1R, KCNJ11, KCNN1, ABCC8, RIMS2, PLCB1, KCNU1

Dwon-regulated pathways in PVTT13

Pathway ID	Description	%	type	Count	P Value	genes
hsa04380	Osteoclast differentiation	4	down	31	1.00E-18	ACP5, BTK, CSF1R, CYBB, FCGR1A, FCGR3A, FCGR3B, IL1R1, ITGB3, JUNB, LCP2, NCF4, PIK3CD, SPI1, SYK, TGFB1, FOSL1, SOCS1, ……
hsa04151	PI3K-Akt signaling pathway	5.1	down	40	1.00E-11	CCND3, COL4A3, COL6A1, COL6A2, COL6A3, COL9A2, CSF1R, CSF3, CSF3R, LPAR1, FGF2, FGF7, FGF9, NR4A1, TNC, IGF1R, IL2RG, IL6, ……
hsa05144	Malaria	1.9	down	15	1.00E-11	CR1, CSF3, GYPC, HBA1, HBA2, HBB, IL6, IL18, ITGB2, TGFB1, TGFB3, THBS1, THBS2
hsa04062	Chemokine signaling pathway	3.5	down	27	1.00E-10	ADCY7, CXCR5, CCR1, CCR5, CCR6, DOCK2, PTK2B, FGR, CXCL2, HCK, CXCR1, CXCR2, JAK2, PIK3CD, PLCB2, PRKCB, RAC2, CCL11, VAV1, ……
hsa04512	ECM-receptor interaction	2.3	down	18	1.58E-10	CD44, COL4A3, COL6A1, COL6A2, COL6A3, COL9A2, TNC, ITGA4, ITGB3, ITGB4, ITGB7, ITGB8, RELN, THBS1, THBS2, ITGA8, GP6, LAMA1
hsa05152	Tuberculosis	3.3	down	26	2.51E-10	CD14, PLK3, CR1, CTSS, FCER1G, FCGR1A, FCGR3A, FCGR3B, IL6, IL10RA, IL18, ITGB2, JAK2, LSP1, SYK, TGFB1, TGFB3, TLR1, TLR2, TLR4, ……

Continued

Continued

Pathway ID	Description	%	type	Count	P Value	genes
hsa04510	Focal adhesion	3.3	down	26	2.51E-09	CCND3, COL4A3, COL6A1, COL6A2, COL6A3, COL9A2, FLNA, TNC, IGF1R, ITGA4, ITGB3, ITGB4, ITGB7, ITGB8, PIK3CD, PRKCB, RELN, ...
hsa04145	Phagosome	2.8	down	22	7.94E-09	CD14, CTSS, CYBB, FCGR1A, FCGR3A, FCGR3B, ITGB2, ITGB3, MPO, MSR1, NCF4, THBS1, THBS2, TLR2, TLR4, TUBA4A, MRC2, CORO1A, ...
hsa05140	Leishmaniasis	1.9	down	15	1.26E-08	CR1, FCGR1A, FCGR3A, FCGR3B, ITGA4, ITGB2, JAK2, NCF4, PRKCB, PTGS2, TGFB1, TGFB3, TLR2, TLR4, NCF1
hsa04611	Platelet activation	2.3	down	18	1.26E-07	ADCY7, BTK, FCER1G, ITGB3, ITPR3, LCP2, P2RX1, PIK3CD, PIK3CG, PLA2G4A, PLCB2, PTGS1, SYK, TBXAS1, AKT3, PIK3R5, GP6, FERMT3
hsa04666	Fc gamma R-mediated phagocytosis	1.9	down	15	3.16E-07	DOCK2, FCGR1A, FCGR3A, HCK, INPP5D, PIK3CD, PLA2G4A, PRKCB, RAC2, SYK, VAV1, WAS, AKT3, LAT, NCF1
hsa04060	Cytokine-cytokine receptor interaction	3.5	down	27	3.98E-07	CXCR5, BMPR1B, TNFSF8, CCR1, CCR5, CCR6, CSF1R, CSF2RB, CSF3, CSF3R, CXCL2, IL1R1, IL2RG, IL6, CXCR1, CXCR2, IL10RA, IL18, ...
hsa04640	Hematopoietic cell lineage	1.9	down	15	6.31E-07	CD4, CD14, CD22, CD37, CD44, CR1, CSF1R, CSF3, CSF3R, FCGR1A, IL1R1, IL6, ITGA4, ITGB3, MME
hsa04621	NOD-like receptor signaling pathway	2.6	down	20	0.000001	CYBB, CXCL2, IFI16, IL6, IL18, ITPR3, MEFV, PLCB2, TLR4, PSTPIP1, RIPK3, LRP1, TRPV2, NLRC4, CARD9, ANTXR1, CARD6, NLRP12, ...
hsa04664	Fc epsilon RI signaling pathway	1.5	down	12	2.00E-06	BTK, FCER1G, INPP5D, LCP2, PIK3CD, PLA2G4A, MAP2K3, RAC2, SYK, VAV1, AKT3, LAT
hsa04662	B cell receptor signaling pathway	1.5	down	12	3.16E-06	BTK, CD22, CD72, INPP5D, PIK3CD, PRKCB, RAC2, SYK, VAV1, AKT3, LILRB3, DAPP1
hsa05200	Pathways in cancer	4.1	down	32	3.16E-06	ADCY7, COL4A3, CSF1R, CSF3R, DAPK1, LPAR1, FGF2, FGF7, FGF9, FZD2, GSTP1, IGF1R, IL6, MMP2, PIK3CD, PLCB2, PRKCB, PTCH1, PTGS2, ...
hsa04670	Leukocyte transendothelial migration	1.9	down	15	5.01E-06	CYBB, PTK2B, ITGA4, ITGB2, MMP2, NCF4, PIK3CD, PRKCB, RAC2, TXK, VAV1, CXCR4, CLDN2, MYL9, NCF1
hsa04933	AGE-RAGE signaling pathway in diabetic complications	1.8	down	14	5.01E-06	COL4A3, CYBB, IL6, JAK2, MMP2, PIM1, PIK3CD, PLCB2, PRKCB, STAT5A, TGFB1, TGFB3, THBD, AKT3
hsa04015	Rap1 signaling pathway	2.7	down	21	7.94E-06	ADCY7, CSF1R, LPAR1, FGF2, FGF7, FGF9, FPR1, FYB, IGF1R, ITGB2, ITGB3, LCP2, PIK3CD, PLCB2, PRKCB, MAP2K3, RAC2, THBS1, LPAR2, ...
hsa04072	Phospholipase D signaling pathway	2.2	down	17	7.94E-06	ADCY7, DGKA, LPAR1, PTK2B, FCER1G, CXCR1, CXCR2, PIK3CD, PIK3CG, PLA2G4A, PLCB2, SYK, LPAR2, AKT3, AGPAT2, PIK3R5, CYTH4
hsa04974	Protein digestion and absorption	1.7	down	13	7.94E-06	COL4A3, COL5A1, COL6A1, COL6A2, COL6A3, COL9A2, COL10A1, KCNN4, MME, PRCP, SLC8A1, SLC7A7, SLC7A8
hsa05146	Amoebiasis	1.7	down	13	1.58E-05	CD14, COL4A3, IL1R1, IL6, ITGB2, PIK3CD, PLCB2, PRKCB, TGFB1, TGFB3, TLR2, TLR4, LAMA1
hsa05410	Hypertrophic cardiomyopathy (HCM)	1.5	down	12	2.00E-05	ACE, IL6, ITGA4, ITGB3, ITGB4, ITGB7, ITGB8, SLC8A1, TGFB1, TGFB3, ITGA8, PRKAG2

Continued

Continued

Pathway ID	Description	%	type	Count	P Value	genes
hsa05418	Fluid shear stress and atherosclerosis	2	down	16	2.00E-05	BMPR1B, CYBB, GSTM2, GSTP1, IL1R1, ITGB3, MMP2, PIK3CD, RAC2, THBD, TP53, AKT3, KLF2, TRPV4, NCF1, GSTT2B
hsa05202	Transcriptional misregulation in cancer	2.3	down	18	3.16E-05	RUNX2, CD14, CD86, CSF1R, DEFA3, DUSP6, FCGR1A, FUT8, GRIA3, ID2, IGF1R, IL6, ITGB7, MPO, SPI1, TP53, NR4A3, PROM1
hsa05205	Proteoglycans in cancer	2.4	down	19	5.01E-05	CD44, FGF2, FLNA, FZD2, HCLS1, IGF1R, ITGB3, ITPR3, MMP2, PIK3CD, PRKCB, PTCH1, TGFB1, THBS1, TLR2, TLR4, TP53, AKT3, HPSE
hsa04630	Jak-STAT signaling pathway	2	down	16	6.31E-05	CCND3, CSF2RB, CSF3, CSF3R, IL2RG, IL6, IL10RA, JAK2, LIF, PIM1, PIK3CD, STAT5A, STAT6, SOCS1, SOCS3, AKT3
hsa04810	Regulation of actin cytoskeleton	2.4	down	19	7.94E-05	CD14, FGF2, FGF7, FGF9, NCKAP1L, ITGA4, ITGB2, ITGB3, ITGB4, ITGB7, ITGB8, PIK3CD, RAC2, VAV1, WAS, ITGA8, MYL9, BAIAP2, FGD3
hsa05134	Legionellosis	1.2	down	9	7.94E-05	CD14, CR1, CXCL2, IL6, IL18, ITGB2, TLR2, TLR4, NLRC4
hsa05150	Staphylococcus aureus infection	1.2	down	9	7.94E-05	C1S, FCGR1A, FCGR3A, FCGR3B, FPR1, FPR2, CFH, ITGB2, PTAFR
hsa04514	Cell adhesion molecules (CAMs)	1.9	down	15	0.0001	CD4, CD22, CD86, ICAM3, ITGA4, ITGB2, ITGB7, ITGB8, SIGLEC1, SPN, ITGA8, CLDN2, NLGN2, LRRC4, NTNG2
hsa05145	Toxoplasmosis	1.7	down	13	0.0001	CCR5, IL10RA, JAK2, PIK3CG, MAP2K3, TGFB1, TGFB3, TLR2, TLR4, SOCS1, AKT3, PIK3R5, LAMA1
hsa05143	African trypanosomiasis	0.9	down	7	0.00012589	HBA1, HBA2, HBB, IL6, IL18, PLCB2, PRKCB
hsa04010	MAPK signaling pathway	2.6	down	20	0.00031623	CD14, DUSP6, FGF2, FGF7, FGF9, FLNA, NR4A1, IL1R1, GADD45B, PLA2G4A, PRKCB, MAP2K3, PTPN7, RAC2, TGFB1, TGFB3, TP53, ……
hsa04064	NF-kappa B signaling pathway	1.4	down	11	0.00031623	BTK, CD14, CXCL2, IL1R1, GADD45B, PRKCB, PTGS2, SYK, TLR4, ……
hsa04750	Inflammatory mediator regulation of TRP channels	1.4	down	11	0.00039811	ADCY7, HTR2B, IL1R1, ITPR3, PIK3CD, PLA2G4A, PLCB2, PRKCB, MAP2K3, TRPV2, TRPV4
hsa04650	Natural killer cell mediated cytotoxicity	1.7	down	13	0.00050119	PTK2B, FCER1G, FCGR3A, FCGR3B, ITGB2, LCP2, PIK3CD, PRKCB, RAC2, SYK, VAV1, NCR1, LAT
hsa04371	Apelin signaling pathway	1.7	down	13	0.00063096	ADCY7, APLNR, CTGF, ITPR3, PIK3CG, PLCB2, SLC8A1, AKT3, KLF2, PIK3R5, PRKAG2, GNG2, GNB4
hsa05323	Rheumatoid arthritis	1.3	down	10	0.00079433	ACP5, CD86, IL6, IL18, ITGB2, TGFB1, TGFB3, TLR2, TLR4, TNFSF13B
hsa05414	Dilated cardiomyopathy	1.3	down	10	0.00079433	ADCY7, ITGA4, ITGB3, ITGB4, ITGB7, ITGB8, SLC8A1, TGFB1, TGFB3, ITGA8
hsa04672	Intestinal immune network for IgA production	0.9	down	7	0.001	CD86, IL6, ITGA4, ITGB7, TGFB1, CXCR4, TNFSF13B
hsa05161	Hepatitis B	1.7	down	13	0.001	EGR3, PTK2B, IL6, PIK3CD, PRKCB, STAT5A, STAT6, TGFB1, TGFB3, TLR2, TLR4, TP53, AKT3
hsa04725	Cholinergic synapse	1.4	down	11	0.00125893	ADCY7, ITPR3, JAK2, PIK3CD, PIK3CG, PLCB2, PRKCB, AKT3, PIK3R5, GNG2, GNB4

Continued

Continued

Pathway ID	Description	%	type	Count	P Value	genes
hsa05321	Inflammatory bowel disease (IBD)	1	down	8	0.00125893	IL2RG, IL6, IL18, STAT6, TGFB1, TGFB3, TLR2, TLR4
hsa04014	Ras signaling pathway	2.2	down	17	0.00158489	CSF1R, FGF2, FGF7, FGF9, IGF1R, PIK3CD, PLA2G4A, PRKCB, RAC2, AKT3, RASA3, LAT, GNG2, GNB4, RASAL3, RASGRP4, RASA4B
hsa04068	FoxO signaling pathway	1.5	down	12	0.00158489	PLK3, IGF1R, IL6, GADD45B, PIK3CD, TGFB1, TGFB3, S1PR4, AKT3, KLF2, GADD45G, PRKAG2
hsa05162	Measles	1.5	down	12	0.00158489	CCND3, IL2RG, IL6, JAK2, PIK3CD, STAT5A, TLR2, TLR4, TP53, AKT3, RAB9B, TLR7
hsa04066	HIF-1 signaling pathway	1.3	down	10	0.00199526	CYBB, HK3, IGF1R, IL6, PFKFB3, PIK3CD, PRKCB, TIMP1, TLR4, AKT3
hsa04350	TGF-beta signaling pathway	1.2	down	9	0.00199526	BMP6, BMPR1B, ID2, INHBA, LTBP1, SMAD9, TGFB1, TGFB3, THBS1
hsa04360	Axon guidance	1.8	down	14	0.00199526	BMPR1B, DPYSL2, PIK3CD, PLXNB3, PTCH1, RAC2, RGS3, CXCR4, SEMA7A, SEMA4D, SEMA3C, LRRC4, SEMA4A, NTNG2
hsa04080	Neuroactive ligand-receptor interaction	2.4	down	19	0.00251189	ADORA3, APLNR, CHRNE, LPAR1, F2RL2, FPR1, FPR2, GPR35, GRIA3, HRH2, HTR2B, MC1R, NTSR1, OXTR, P2RX1, PTAFR, PTGDR, S1PR4, ...
hsa04620	Toll-like receptor signaling pathway	1.3	down	10	0.00251189	CD14, CD86, IL6, PIK3CD, MAP2K3, TLR1, TLR2, TLR4, AKT3, TLR7
hsa04020	Calcium signaling pathway	1.8	down	14	0.00316228	ADCY7, ATP2A3, PTK2B, HRH2, HTR2B, ITPR3, NTSR1, OXTR, P2RX1, PLCB2, PRKCB, PTAFR, SLC8A1, ORAI2
hsa04668	TNF signaling pathway	1.3	down	10	0.00316228	CXCL2, IL6, JUNB, LIF, PIK3CD, MAP2K3, PTGS2, SOCS3, AKT3, RIPK3
hsa04726	Serotonergic synapse	1.3	down	10	0.00398107	HTR2B, ITPR3, KCND2, PLA2G4A, PLCB2, PRKCB, PTGS1, PTGS2, GNG2, GNB4
hsa04610	Complement and coagulation cascades	1	down	8	0.00501187	C1S, CR1, F2RL2, CFH, ITGB2, THBD, PROCR, VSIG4
hsa04071	Sphingolipid signaling pathway	1.3	down	10	0.00630957	ADORA3, FCER1G, PIK3CD, PLCB2, PRKCB, RAC2, TP53, S1PR4, AKT3, SGPP2
hsa04144	Endocytosis	2.2	down	17	0.00630957	CCR5, CSF1R, IGF1R, IL2RG, CXCR1, CXCR2, TGFB1, TGFB3, WAS, WIPF1, CXCR4, RAB31, SPG20, PSD4, CYTH4, VPS37B, RAB11FIP1
hsa04550	Signaling pathways regulating pluripotency of stem cells	1.4	down	11	0.00630957	BMPR1B, FGF2, FZD2, ID2, IGF1R, INHBA, JAK2, LIF, SMAD9, PIK3CD, AKT3
hsa04723	Retrograde endocannabinoid signaling	1.2	down	9	0.00630957	ADCY7, GRIA3, ITPR3, PLCB2, PRKCB, PTGS2, GNG2, GNB4, SLC17A8
hsa04913	Ovarian steroidogenesis	0.77	down	6	0.00630957	ADCY7, BMP6, CYP1B1, IGF1R, PLA2G4A, PTGS2
hsa05142	Chagas disease (American trypanosomiasis)	1.2	down	9	0.00630957	ACE, IL6, PIK3CD, PLCB2, TGFB1, TGFB3, TLR2, TLR4, AKT3
hsa04115	p53 signaling pathway	0.9	down	7	0.00794328	CCND3, CD82, GADD45B, PMAIP1, THBS1, TP53, GADD45G
hsa04540	Gap junction	1	down	8	0.00794328	ADCY7, LPAR1, HTR2B, ITPR3, PLCB2, PRKCB, TUBA4A, TUBB6
hsa05132	Salmonella infection	1	down	8	0.00794328	CD14, FLNA, CXCL2, IL6, IL18, TLR4, WAS, NLRC4

Continued

Continued

Pathway ID	Description	%	type	Count	P Value	genes
hsa05218	Melanoma	0.9	down	7	0.00794328	FGF2, FGF7, FGF9, IGF1R, PIK3CD, TP53, AKT3
hsa04659	Th17 cell differentiation	1.2	down	9	0.01	CD4, IL1R1, IL2RG, IL6, JAK2, STAT5A, STAT6, TGFB1, LAT
hsa05130	Pathogenic Escherichia coli infection	0.77	down	6	0.01	CD14, HCLS1, TLR4, TUBA4A, WAS, TUBB6
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	0.9	down	7	0.01	ITGA4, ITGB3, ITGB4, ITGB7, ITGB8, SLC8A1, ITGA8

Up-regulated pathways in T13

Pathway ID	Description	%	Type	Count	*P*	Genes
hsa04110	Cell cycle	2.9	up	17	1.26E-08	BUB1, CCNA2, CDK1, CDK4, CDKN2A, MAD2L1, MCM3, MCM7, ORC1, PLK1, SKP2, PKMYT1, ESPL1, DBF4, CHEK2, ORC6, ANAPC11
hsa00480	Glutathione metabolism	1.4	up	8	5.01E-05	GCLM, GPX2, GSTA2, GSTA4, IDH1, PGD, RRM2, SRM
hsa03460	Fanconi anemia pathway	1.4	up	8	6.31E-05	BLM, BRCA1, FANCA, FANCD2, FANCB, RAD51C, FANCI, BRIP1
hsa05219	Bladder cancer	1.2	up	7	6.31E-05	CDK4, CDKN2A, EGF, CXCL8, MMP1, MMP9, SRC
hsa03440	Homologous recombination	1	up	6	0.00050119	BARD1, BLM, BRCA1, RAD51C, RAD54L, BRIP1
hsa00410	beta-Alanine metabolism	0.85	up	5	0.001	ALDH3A1, ALDH1B1, AOC2, SRM, SMOX
hsa01230	Biosynthesis of amino acids	1.2	up	7	0.00251189	BCAT1, IDH1, PKM, TKT, TKTL1, PHGDH, PSAT1
hsa04657	IL-17 signaling pathway	1.4	up	8	0.00251189	CXCL3, CXCL8, LCN2, MMP1, MMP9, MMP13, MAPK13, CCL7
hsa04914	Progesterone-mediated oocyte maturation	1.4	up	8	0.00251189	BUB1, CCNA2, CDK1, MAD2L1, PLK1, MAPK13, PKMYT1, ANAPC11
hsa_M00004	Pentose phosphate pathway (pentose phosphate cycle)	0.51	up	3	0.00316228	PGD, TKT, TKTL1
hsa00561	Glycerolipid metabolism	1	up	6	0.00316228	ALDH1B1, GLA, LIPF, PLPP2, GPAT3, MOGAT1
hsa_M00050	Guanine ribonucleotide biosynthesis IMP ≥ GDP, GTP	0.51	up	3	0.00398107	IMPDH1, IMPDH2, PKM
hsa00600	Sphingolipid metabolism	0.85	up	5	0.00630957	GBA, GLA, PLPP2, SPHK1, ASAH2
hsa00603	Glycosphingolipid biosynthesis - globo and isoglobo series	0.51	up	3	0.00630957	FUT2, GLA, B3GALNT1
hsa04115	p53 signaling pathway	1	up	6	0.00794328	CDK1, CDK4, CDKN2A, RRM2, CHEK2, GTSE1
hsa04978	Mineral absorption	0.85	up	5	0.01	CLCN2, FTH1, SLC26A6, TRPM6, STEAP2

Down-regulated pathways in T13

pathway	Description	%	Type	Count	*P* Value	Genes
hsa04610	Complement and coagulation cascades	2.1	down	20	1.00E-11	A2M, C3, C8B, C8G, CPB2, F2, F9, F12, F13A1, FGA, FGG, CFI, KNG1, SERPINA1, PLG, SERPINF2, PROC, VTN, VWF, MASP2
hsa00380	Tryptophan metabolism	1.1	down	10	2.51E-06	ACAT1, ALDH2, CAT, CYP1A1, GCDH, TDO2, KMO, AADAT, OGDHL, ACMSD
hsa00071	Fatty acid degradation	1.1	down	10	6.31E-06	ACADL, ACADM, ACADSB, ACADVL, ACAT1, ADH6, ALDH2, CPT2, GCDH, ACAA2
hsa00630	Glyoxylate and dicarboxylate metabolism	0.85	down	8	0.00001	ACAT1, AGXT, AMT, CAT, HAO2, HAO1, MCEE, HOGA1
hsa00280	Valine, leucine and isoleucine degradation	1.1	down	10	1.58E-05	ABAT, ACADM, ACADSB, ACAT1, ALDH2, BCKDHB, IVD, ACAA2, AGXT2, MCEE
hsa04514	Cell adhesion molecules (CAMs)	1.8	down	17	6.31E-05	CD80, CD40, CDH1, CDH3, CDH5, CDH15, CLDN3, ICAM2, CLDN5, VCAM1, CLDN8, NRXN3, NRXN1, NLGN1, CLDN14, JAM2, OCLN
hsa00650	Butanoate metabolism	0.75	down	7	7.94E-05	ABAT, ACAT1, BDH1, ALDH5A1, ACSM5, ACSM2A, ACSM2B

Continued

Continued

Pathway ID	Description	%	Type	Count	*P*	Genes
hsa04976	Bile secretion	1.2	down	11	0.0001	ADCY8, AQP1, ATP1A2, ATP1A3, FXYD2, ABCB4, SLC10A1, SLCO1A2, SLCO1B1, ABCG8, SLC51A
hsa05204	Chemical carcinogenesis	1.3	down	12	0.0001	ADH6, CYP1A1, CYP2A13, CYP2C18, CYP3A5, GSTM5, SULT1A2, SULT1A1, UGT2B15, UGT2B28, UGT1A4, CYP3A43
hsa04918	Thyroid hormone synthesis	1.2	down	11	0.00015849	ADCY8, ALB, ASGR1, ASGR2, ATP1A2, ATP1A3, FXYD2, GPX3, PLCB4, TTR, IYD
hsa00830	Retinol metabolism	1.1	down	10	0.00025119	ADH6, CYP1A1, CYP2C18, CYP3A5, RDH5, UGT2B15, RDH16, HSD17B6, UGT2B28, UGT1A4
hsa00140	Steroid hormone biosynthesis	0.96	down	9	0.00039811	CYP1A1, CYP3A5, SRD5A1, AKR1D1, UGT2B15, HSD17B6, CYP7B1, UGT2B28, UGT1A4
hsa03320	PPAR signaling pathway	1.1	down	10	0.00050119	ACADL, ACADM, APOA1, APOC3, CPT2, FABP1, RXRG, SCP2, SORBS1, PLIN5
hsa00330	Arginine and proline metabolism	0.85	down	8	0.00079433	ALDH2, ARG1, DAO, NOS1, P4HA1, PRODH, CARNS1, HOGA1
hsa_M00087	beta-Oxidation	0.43	down	4	0.001	ACADL, ACADM, ACADVL, ACAA2
hsa00640	Propanoate metabolism	0.64	down	6	0.00158489	ABAT, ACADM, ACAT1, BCKDHB, MLYCD, MCEE
hsa00982	Drug metabolism - cytochrome P450	0.96	down	9	0.00158489	ADH6, CYP2D6, CYP3A5, FMO3, FMO4, GSTM5, UGT2B15, UGT2B28, UGT1A4
hsa04973	Carbohydrate digestion and absorption	0.75	down	7	0.00158489	AMY2B, ATP1A2, ATP1A3, FXYD2, G6PC, SLC37A4, SLC2A2
hsa00340	Histidine metabolism	0.53	down	5	0.00199526	ALDH2, HAL, FTCD, CARNS1, AMDHD1
hsa04022	cGMP-PKG signaling pathway	1.6	down	15	0.00199526	ADCY8, ADORA1, ADRA1B, ADRB2, AGTR1, ATP1A2, ATP1A3, FXYD2, CNGA1, EDNRB, GUCY1A2, KCNJ8, NPR1, PLCB4, IRS2
hsa04261	Adrenergic signaling in cardiomyocytes	1.5	down	14	0.00199526	ACTC1, ADCY8, ADRA1B, ADRB2, AGTR1, ATP1A2, ATP1A3, FXYD2, CACNB2, CAMK2B, PLCB4, PPP1R1A, PPP2R2B, SCN4B
hsa04916	Melanogenesis	1.2	down	11	0.00199526	ADCY8, CAMK2B, EDNRB, GNAO1, KIT, PLCB4, WNT1, WNT7A, WNT8B, WNT9A, FZD8
hsa00250	Alanine, aspartate and glutamate metabolism	0.64	down	6	0.00251189	ABAT, AGXT, GPT, ALDH5A1, AGXT2, ADSSL1
hsa00980	Metabolism of xenobiotics by cytochrome P450	0.96	down	9	0.00251189	ADH6, CYP1A1, CYP2A13, CYP2D6, CYP3A5, GSTM5, UGT2B15, UGT2B28, UGT1A4
hsa01212	Fatty acid metabolism	0.75	down	7	0.00251189	ACADL, ACADM, ACADSB, ACADVL, ACAT1, CPT2, ACAA2
hsa04060	Cytokine-cytokine receptor interaction	2.2	down	21	0.00251189	AMHR2, FAS, CD40, EPO, INHBB, INHBC, KIT, CCL8, CCL24, XCL1, CX3CL1, TNFSF11, TNFRSF25, TNFSF10, TNFSF18, IL20RA, ...
hsa05200	Pathways in cancer	3	down	28	0.00251189	ADCY8, AGTR1, FAS, AR, CCND1, RUNX1T1, CDH1, EDNRB, ERBB2, FGF1, FGF10, FGFR2, GNG11, ITGA3, KIT, NTRK1, PLCB4, PTGER3, ...
hsa04390	Hippo signaling pathway	1.5	down	14	0.00316228	CCND1, CDH1, FGF1, PPP2R2B, PRKCZ, WNT1, WNT7A, WNT8B, WNT9A, FZD8, NKD1, WTIP, GDF7, AMOT
hsa04530	Tight junction	1.6	down	15	0.00316228	CCND1, CLDN3, ERBB2, MYH4, PPP2R2B, PRKCZ, TIAM1, CLDN5, CLDN8, CLDN14, AMOTL2, JAM2, CGNL1, AMOT, OCLN

Continued

Continued

Up-regulated pathways in T13						
Pathway ID	Description	%	Type	Count	*P*	Genes
hsa04970	Salivary secretion	1.1	down	10	0.00316228	ADCY8, ADRA1B, ADRB2, ATP1A2, ATP1A3, FXYD2, GUCY1A2, NOS1, PLCB4, SLC12A2
hsa00053	Ascorbate and aldarate metabolism	0.53	down	5	0.00398107	ALDH2, UGT2B15, RGN, UGT2B28, UGT1A4
hsa04015	Rap1 signaling pathway	1.8	down	17	0.00398107	ADCY8, CDH1, EFNA2, EFNA3, EPHA2, FGF1, FGF10, FGFR2, GNAO1, GRIN2A, KIT, PLCB4, PRKCZ, TIAM1, SKAP1, MAGI2, PDGFD
hsa04080	Neuroactive ligand-receptor interaction	2.2	down	21	0.00398107	ADORA1, ADRA1B, ADRB2, AGTR1, CALCR, DRD1, EDNRB, F2, F2RL1, GABRB3, GABRE, NPBWR1, GRIN2A, GRM7, HTR2A, PLG, PTGER3, ……
hsa04360	Axon guidance	1.6	down	15	0.00398107	CAMK2B, EFNA2, EFNA3, EPHA2, EPHA4, EPHA7, PRKCZ, SLIT3, SEMA5A, RND1, PAK5, NTN4, GDF7, SEMA3D, EPHA6
hsa04146	Peroxisome	0.96	down	9	0.00501187	AGXT, CAT, DAO, ECH1, SCP2, MLYCD, HAO2, HAO1, NUDT7
hsa04713	Circadian entrainment	1.1	down	10	0.00501187	ADCY8, CAMK2B, GNAO1, GNG11, GRIN2A, GUCY1A2, NOS1, PLCB4, RYR1, PER3
hsa00410	beta-Alanine metabolism	0.53	down	5	0.00630957	ABAT, ACADM, ALDH2, MLYCD, CARNS1
hsa01200	Carbon metabolism	1.2	down	11	0.00630957	ACADM, ACAT1, AGXT, AMT, CAT, GPT, RGN, HAO2, HAO1, OGDHL, MCEE
hsa04020	Calcium signaling pathway	1.6	down	15	0.00630957	ADCY8, ADRA1B, ADRB2, AGTR1, CAMK2B, DRD1, EDNRB, ERBB2, ERBB4, GRIN2A, HTR2A, NOS1, PLCB4, PTGER3, RYR1
hsa_M00032	Lysine degradation, lysine acetoacetyl-CoA	0.32	down	3	0.00794328	GCDH, AADAT, OGDHL
hsa02010	ABC transporters	0.64	down	6	0.00794328	ABCA3, ABCC6, ABCB4, ABCC9, ABCA8, ABCG8
hsa05205	Proteoglycans in cancer	1.7	down	16	0.00794328	ANK3, FAS, CCND1, CAMK2B, ERBB2, ERBB4, IGF2, TIAM1, TIMP3, VTN, WNT1, WNT7A, WNT8B, WNT9A, FZD8, CBLC

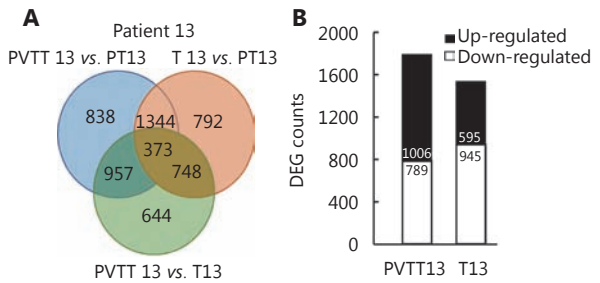

Figure S1 Analysis of aberrantly expressed genes in PVTT13 and T13.(A) Venn diagram of differently expressed genes between PVTT13, T13 and PT13. Genes with fold change>3 and FDR < 0.05 were identified as differently expressed between tissues.(B) Number of genes aberrantly expressed specifically in PVTT13 and T13. Up-regulation and down-regulation of genes was estimated relatively to the expression in PT13.

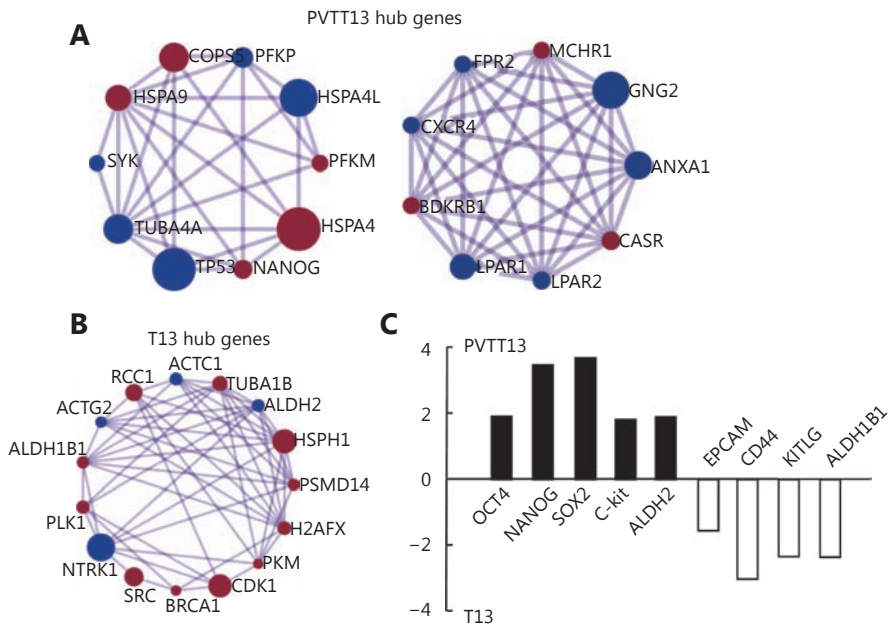

Figure S2 Hub genes and expression of cancer stem cell genes in PVTT13 and T13. (A) Hub genes (degree > 40) in PVTT13 found by PPI analysis using DEGs between PVTT13 and PT13. (B) Hub genes (degree > 50) in T13 found by PPI analysis using DEGs between T13 and PT13. Genes with degree higher than 50 were displayed. PPI analysis was performed using metascape database, and the result was viewed using Cytoscape 3.4. (C) Expression of cancer stem cell markers in PVTT13 relative with T13. FDR < 0.05; FC, fold change.

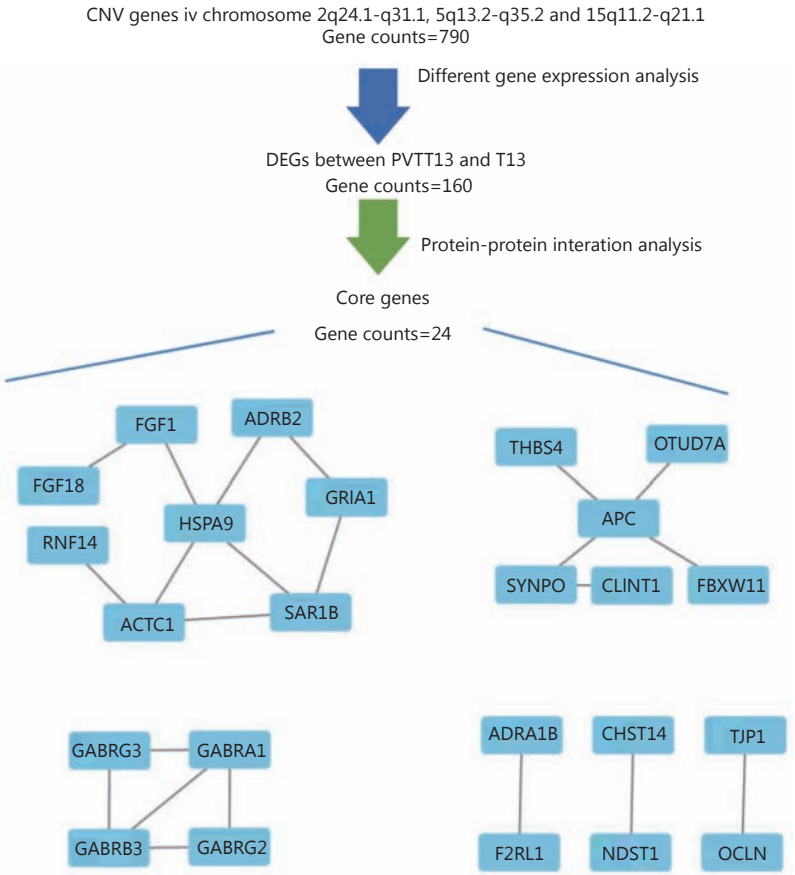

Figure S3 Analysis of differently expressed genes between PVTT13 and T13. 790 genes were found located in Chromosome regions with CNV. 160 out of these genes were found differently expressed between PVTT13 and T13. 24 DEGs were found to be node genes in the gene network established by PPI analysis.