Table S1 Series Test of Cluster (STC) analysis for DEGs

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **GeneSymbol** | **SPOT** | **Profile** | **N** | **T** | **PVTT** |
| COL1A1 | ID\_20 | 4 | 0 | 142.09 | 2,754.76 |
| COL1A2 | ID\_21 | 4 | 0 | 3.62 | 188.62 |
| COL3A1 | ID\_22 | 4 | 0 | -38.66 | 134.78 |
| COL5A2 | ID\_85 | 7 | 0 | 49.19 | 99.38 |
| COL4A3BP | ID\_1050 | 7 | 0 | 22.06 | 65.97 |
| COL2A1 | ID\_1439;ID\_3688 | 1 | 0 | -342.59 | -388.64 |
| COL4A1 | ID\_1440 | 7 | 0 | 264.38 | 785.38 |
| COL6A1 | ID\_1441 | 2 | 0 | -865.87 | 62.11 |
| COL6A2 | ID\_1442;ID\_3781;ID\_3782 | 4 | 0 | -725.37 | 772.84 |
| COL9A2 | ID\_1443 | 6 | 0 | 442.21 | 351.32 |
| COL9A3 | ID\_1444 | 1 | 0 | -446.94 | -494.22 |
| COL15A1 | ID\_1445 | 7 | 0 | 29.19 | 69.78 |
| COL16A1 | ID\_1446 | 6 | 0 | 58.92 | 62.35 |
| COL6A3 | ID\_1914;ID\_3774;ID\_3775 | 4 | 0 | -140.03 | 403.45 |
| COLEC10 | ID\_2295 | 1 | 0 | -77.5 | -52.83 |
| COL11A2 | ID\_3806 | 1 | 0 | -68.41 | -86.89 |
| COLEC11 | ID\_4544 | 1 | 0 | -427.09 | -319.35 |

**SPOT**: Number of genes assigned in the STEM analysis. **Profile:** The gene is situated within the trend number. The follow-up columns: The median gene expression values for each group normalized by STEM.



**Supplementary Figure S1**

**A**. HE staining of paired PVTT samples (PVTT), primary tumors (T) and normal liver tissues (N). The representative images were shown. **B**. Immunohistochemistry staining of CD62P in paired PVTT, primary tumors (T) and normal liver tissues (N). **C**. Immunohistochemistry staining of Fibrinogen in paired PVTT, primary tumors (T) and normal liver tissues (N).



**Supplementary Figure S2**

**A**. UpSet diagram analysis of the differential gene expression of PVTT tissues, paired primary tumors (Ts), nontumorous tissues (Ns). **B**. Series Test of Cluster (STC) analysis for DEGs of collagen molecules in PVTT tissues, paired primary tumors (Ts), nontumorous tissues (Ns). Each box represents a gene expression trend model. The number on the top of the box indicates the trend number. The line segment during the experiment, from left to right, shows Ns, Ts and PVTT. The smaller P-value value on the left of the box indicates higher model credibility. Red and green colors indicate significant changes in trend. The screening criteria default to a p value of < 0.05. In this context, red signifies a significant upward trend, while green signifies a significant downward trend. **C**-**D**. Expression correlation of COL4A1 and TGFB1 (**C**) or KITLG (**D**) according to TCGA datasets. Linear regression analysis showed a positive correlation between mRNA levels of KITLG (R = 0.351, p < 0.0001) but a much weaker correlation between TGFB1 (R = 0.263, p < 0.0001) and COL4A1 from HCC samples in TCGA database.



**Supplementary Figure S3**

**A**. Immunohistochemistry staining of COL4A1 was performed in paired PVTT samples (PVTT), primary tumors (T) and normal liver tissues (N) from 6 patients. **B**. COL4A1 mRNA expression level in N, T, PVTT tissues according to GSE77509 (20 HCC with PVTT patients).



**Supplementary Figure S4**

**A**. Immunohistochemistry staining of KITLG was performed in paired PVTT samples (PVTT), primary tumors (T) and normal liver tissues (N) from 6 patients. **B.** KITLG mRNA expression level in N, T, PVTT tissues according to GSE77509 (20 HCC with PVTT patients).

**Supplementary Table S2**

DEGs between PVTT and non-tumorous tissues (P vs. N), paired primary tumors and non-tumorous tissues (T vs. N), and PVTT and paired primary tumors (P vs. T).