

Supplementary Figures for

HepScope: CNN-based single-cell discrimination of malignant hepatocytes

Yunjin Go[†], Jestlin Tianthing Ng[†], Darren Wan-Teck Lim, Su Bin Lim

Correspondence to: sblim@ajou.ac.kr

This PDF file includes supplementary figures.

FIGURE S1. VALIDATION OF HEPSCOPE GENE SET IN STRNA-SEQ DATA	2
FIGURE S2. PROPORTION OF IMMUNE CELL'S SUBTYPE IN SAMPLES WITH LOW HEPSCOPE MODULE SCORES VS HIGH HEPSCOPE MODULE SCORES.	3
FIGURE S3. FOREST PLOT SUMMARIZING COX REGRESSION COEFFICIENTS OF HEPSCOPE GENE SET USED FOR RISK SCORE CALCULATION.	4
FIGURE S4. DESCRIPTION OF GEPLIVER SCRNA-SEQ DATA.	5
FIGURE S5. DESCRIPTION OF SCATLAS SCRNA-SEQ DATA.....	6
FIGURE S6. BATCH EFFECT CORRECTION OF BULK RNA-SEQ DATA.	7

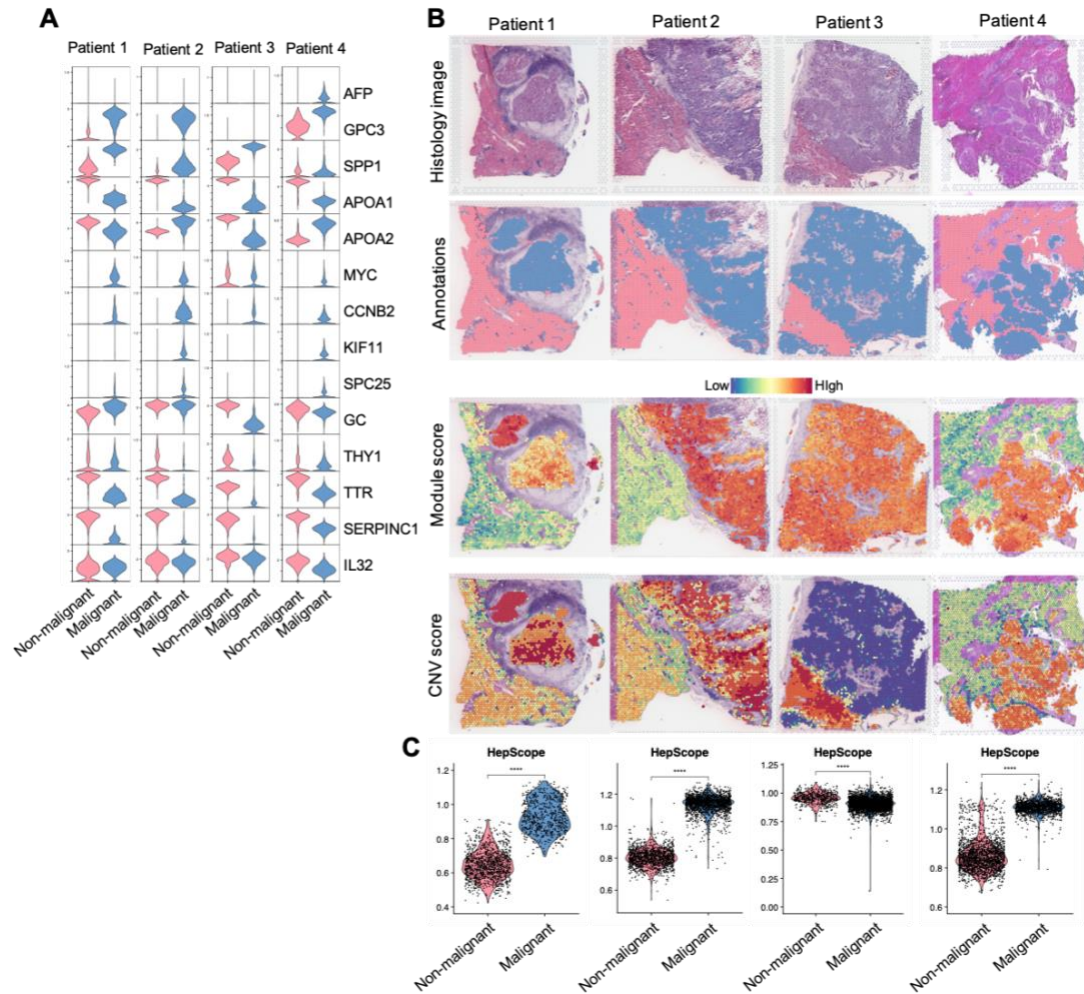


Figure S1. Validation of HepScope gene set in stRNA-seq data

(A) Violin plot showing expression of known HCC biomarkers between malignant hepatocytes and non-malignant hepatocytes in 4 Visium 10X stRNA-seq tissues obtained from HCC patients. **(B)** Histological image (1st row), Malignancy annotation of spots (2nd row), HepScope gene set module score (3rd row) and CNV score (4th row) in Visium 10X stRNA-seq data of HCC obtained from 4 patients. **(C)** Barplot of HepScope gene set module scores of malignant and non-malignant regions in the 4 tissues. Statistical comparison is performed with Wilcoxon ranked sum test with significance determined as $p < 0.05$.

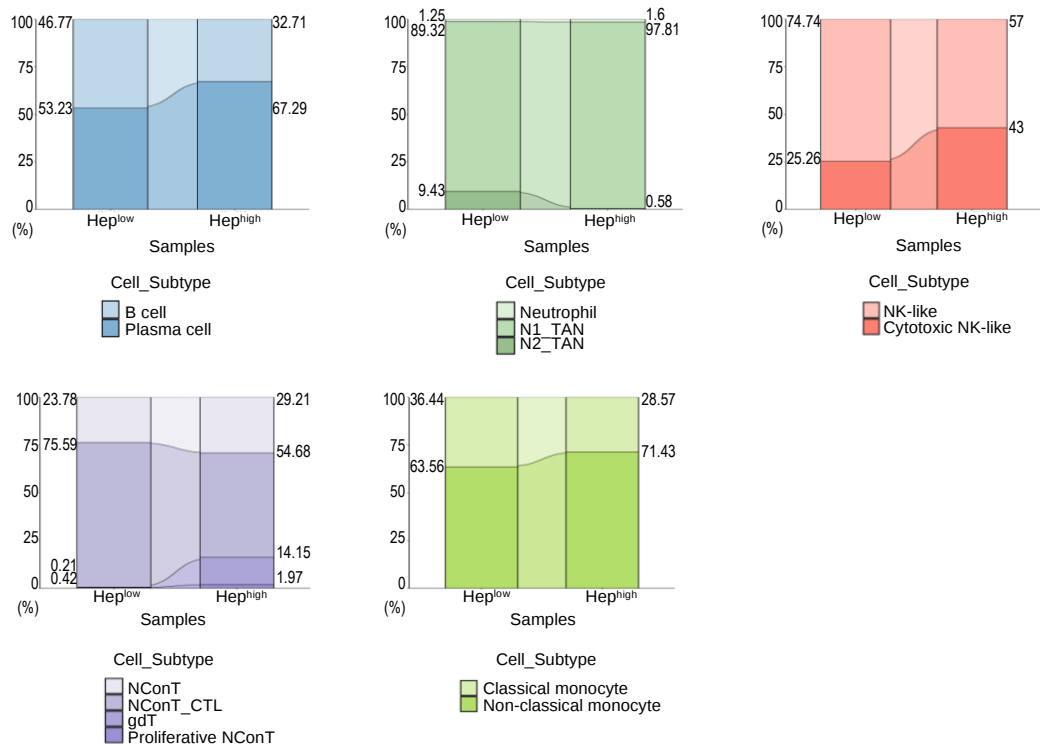


Figure S2. Proportion of immune cell's subtype in samples with low HepScore module scores vs high HepScore module scores.

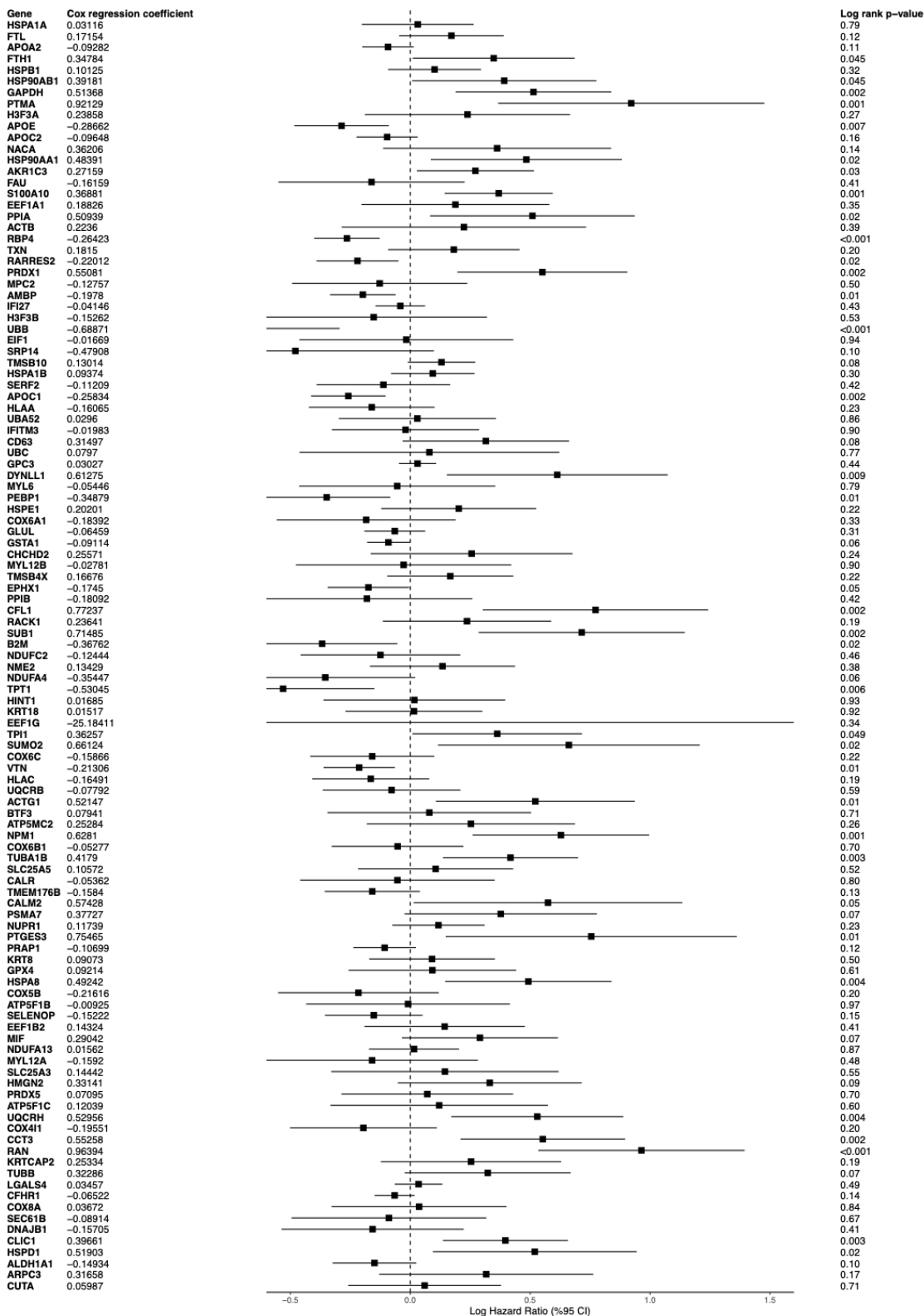


Figure S3. Forest plot summarizing cox regression coefficients of HepScope gene set used for risk score calculation.

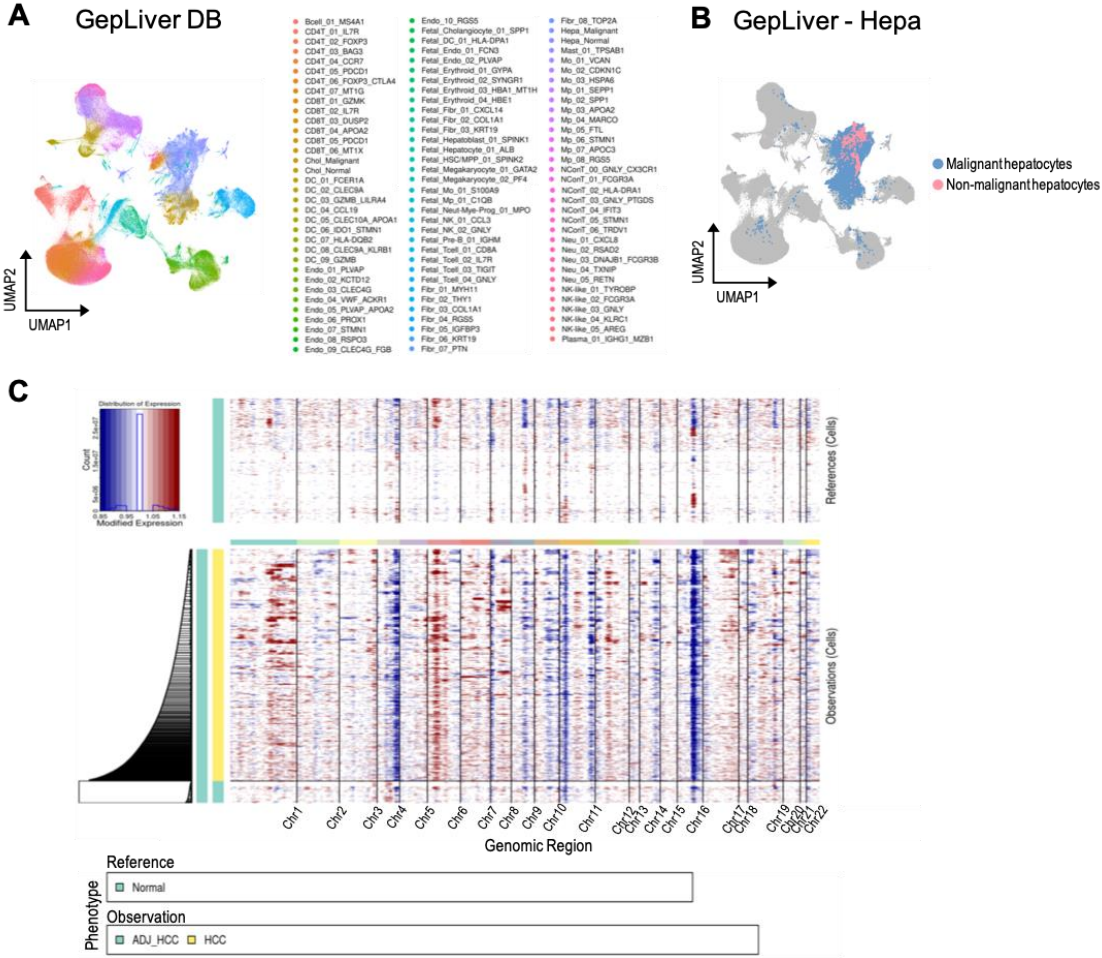
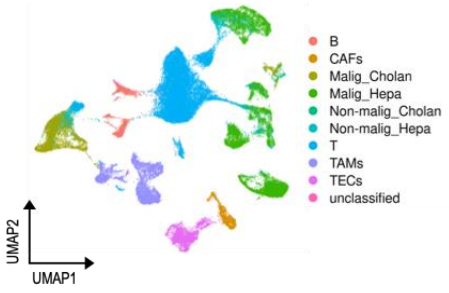
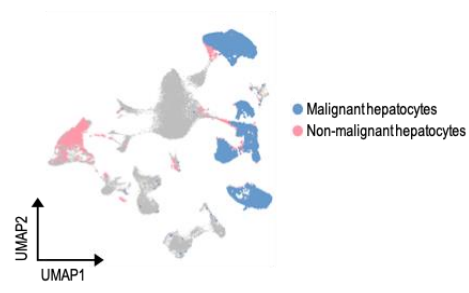
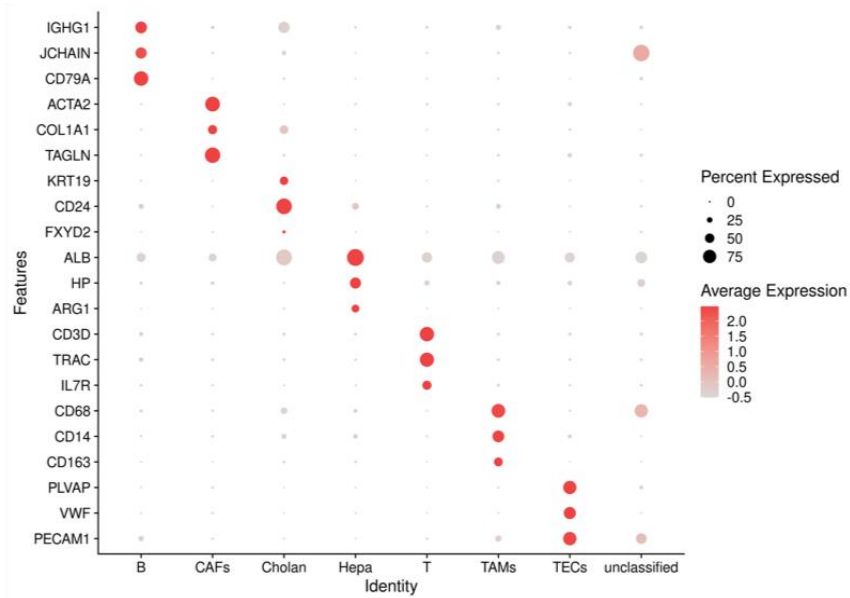


Figure S4. Description of GepLiver scRNA-seq data.

(A) Umap showing total cell types in GepLiver dataset. **(B)** Umap showing hepatocytes in GepLiver dataset. **(C)** Result obtained from inferCNV after running with 5,000 downsampled HCC and ADJ_HCC hepatocytes referred by hepatocytes from normal sample.

A scAtlas(HCC)**B** scAtlas(HCC) - Hepa**C****Figure S5. Description of scAtlas scRNA-seq.**

(A) Umap showing total cell types in scAtlas dataset. **(B)** Umap showing hepatocytes in scAtlas dataset. **(C)** Dot plot showing annotation of cell type by using canonical markers for each cell type.

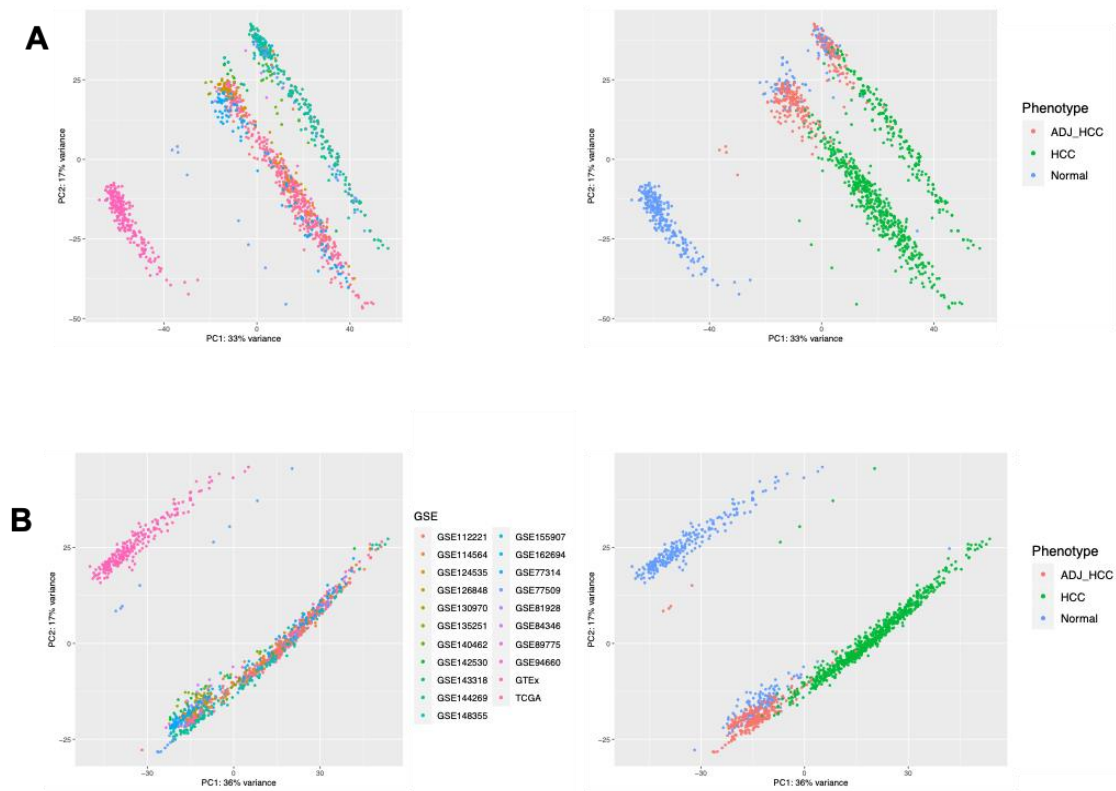


Figure S6. Batch effect correction of bulk RNA-seq data.

(A) PCA plots of bulk RNA-seq samples split between datasets before batch effect correction.

(B) after batch correction.