

Liver-enriched Genes are Associated with the Prognosis of Patients with Hepatocellular Carcinoma

Binghua Li^{1,*}, Tiancheng Xu^{1,*}, Chaohui Liu², Gang Meng², Yuchen Sun¹, Liyuan Qian¹,
Shaohe Wang¹, Jiwu Wei², Decai Yu¹, Yitao Ding¹

¹ Department of Hepatobiliary Surgery, the Affiliated Drum Tower Hospital, Medical
School of Nanjing University, Nanjing, 210093, China

² Jiangsu Key Laboratory of Molecular Medicine, Medical School of Nanjing University,
Nanjing, 210093, China

*These authors contributed equally to this study. Correspondence and requests for
materials should be addressed to D.C.Y. (email: dryudecai@qq.com) or Y.T.D. (email:
yitao_ding@163.com)

Figure S1. Expression pattern of liver-enriched genes. Heatmap showing the expression patterns of liver-enriched genes in non-tumor tissue and tumor tissue in (A) GSE14520 and (C) GSE54236. Red indicates higher expression and blue indicates lower expression. Quantitative analysis of the expression patterns of liver-enriched genes in non-tumor and tumor tissues in (B) GSE14520 and (D) GSE54236. (E) Heatmap showing the internal correlation of the 188 liver-enriched genes based on Spearman's correlation coefficient. The color in the heatmap represents the Spearman's correlation coefficient. Red indicates positive correlation, and blue indicates negative correlation.

Figure S2. Consensus clustering of liver-enriched genes in the TCGA and GEO datasets. Consensus clustering matrix of liver-enriched genes for $k=2$ to $k=5$ of (A) TCGA, (D) GSE14520, and (G) GSE54236. Consensus clustering CDF for $k=2$ to $k=6$ of (B) TCGA, (E) GSE14520, and (H) GSE54236. (C) Mean expression values for liver-enriched genes of the patients in the different groups in (C) TCGA, (F) GSE14520, and (I) GSE54236.

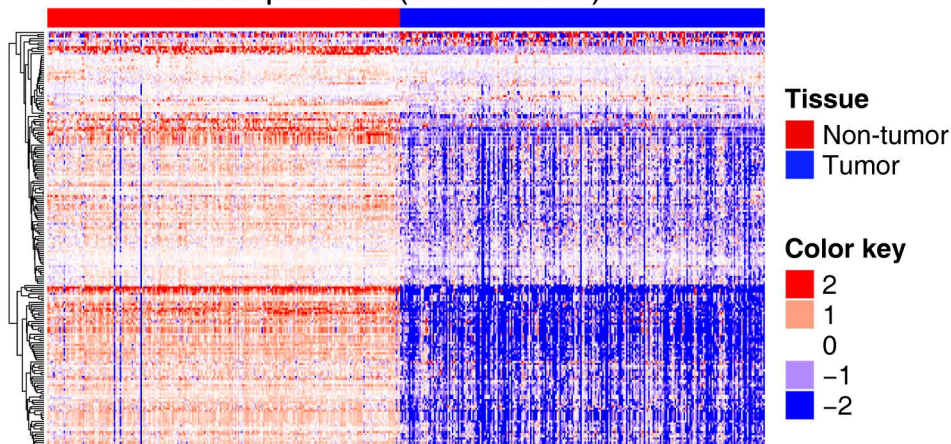
Figure S3. Time-dependent Brier score curves and ROC curves according to mean expression distance in the TCGA and GEO datasets. The time-dependent Brier score curve of (A) the TCGA LIHC cohort (B) the GSE14520 dataset, and (C) the GSE54236 dataset. The time-dependent ROC curve of (D) the TCGA LIHC cohort (E) the GSE14520 dataset, and (F) the GSE54236 dataset.

Table S1. The tissue-enriched gene list of 12 tissue types.

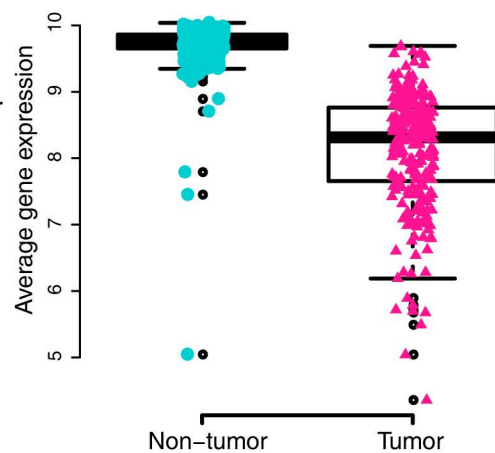
A

LIHC patients(GSE14520)

liver-enriched genes



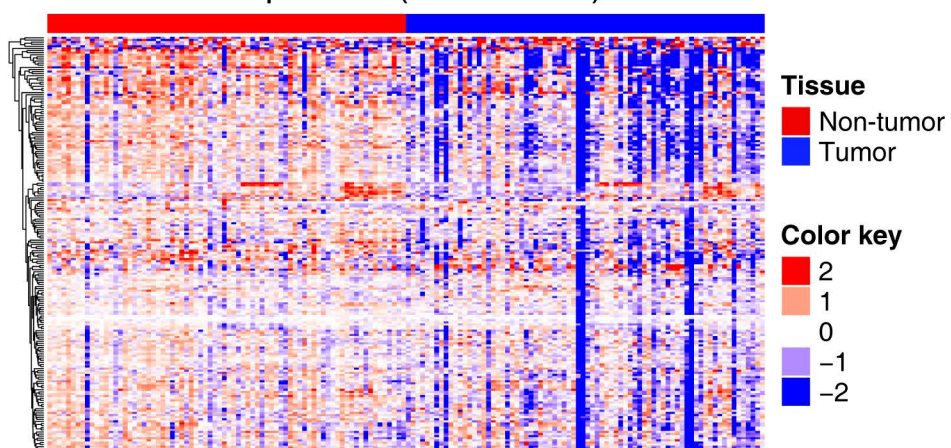
B

GSE14520
Kruskal-Wallis Test : $p < 0.0001$ 

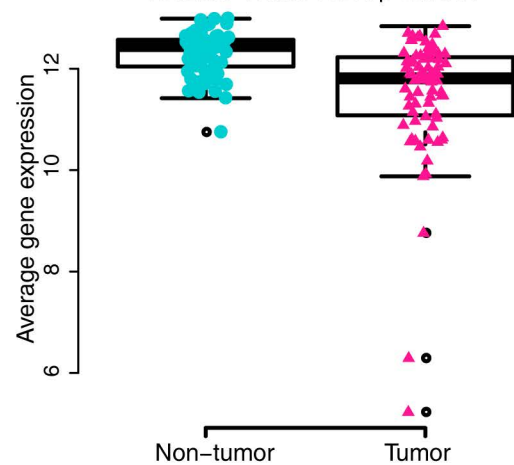
C

LIHC patients(GSE54236)

liver-enriched genes



D

GSE54236
Kruskal-Wallis Test : $p < 0.0001$ 

E

