**Supplementary Table 1.** **HCC data sets**

|  |  |  |  |
| --- | --- | --- | --- |
| **Accession** | **Brief introduction** | **GPL** | **Samples** |
| GSE54236 | Prospective gene expression analysis of human RNA samples from HCC in relation with growth rate and survival | GPL6480:Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version) | 78 primary HCC tumor samples and 78 adjacent cirrhotic samples |
| GSE14520 | Gene expression data of human HCC | GPL3921:[HT\_HG-U133A] Affymetrix HT Human Genome U133A Array | 225 primary HCC tumor samples and 220 adjacent samples |
| GSE25097 | Study of gene expression of human liver HCC | GPL10687:Rosetta/Merck Human RSTA Affymetrix 1.0 Microarray | 268 HCC tumor, 243 adjacent non-tumor, 40 cirrhotic and 6 healthy liver samples |
| GSE62232 | Large-scale gene expression profiling of 81 HCCs | GPL570:[HG-U133\_Plus\_2] Affymetrix Human Genome U133 Plus 2.0 Array | 81 liver tumors |
| GSE36376 | Gene expression profiles of both tumor and adjacent non-tumor liver tissue | GPL10558:Illumina HumanHT-12 V4.0 Expression Beadchip | 193 adjacent non-tumor liver tissue; 240 liver tumor |
| TCGA LIHC | Liver HCC from Xena | Illumina HiSeq\_HiSeqV2 | 370 primary HCC tumor samples and 50 normal solid tissues |
| TCGA pan-cancer | Pan-cancer | Illumina HiSeq\_RNASeqV2 | 11060 samples |

Abbreviation: GPL, GEO Platform.