Supplementary Figure Legends

Fig S1. The distribution of correlation coefficients between gene copy numbers and expression levels. (*A*) For the 9,981 distinct genes, the distribution of the correlation coefficients of Pearson's correlation tests between gene copy numbers and the corresponding expression levels were plotted. Mean correlation coefficient (mean = 0.1416, red line) was positively shifted against zero (blue line). The significance of mean change correlation coefficient was determined by one-sample *T*-test ($P < 2.2 \times 10^{-16}$). (*B*) Pair-wise correlation matrix between gene copy numbers and expression levels of all genes in individual patient are indicated with different color scale. The highest scores in the diagonal line in the correlation matrix indicate the best correlation of each sample copy numbers to the corresponding sample's gene expression levels. (*C,D*) The correlation between copy numbers and gene expression levels were shown in two corCNA genes, *PORL2K* (C) and *C1orf43* (*D*). The DNA copy numbers of these genes in 52 HCC samples (including 15 samples used in CGH analysis) were estimated by qPCR, while the gene expression values were obtained from the corresponding gene expression profiles.

Fig.S 2. Cross-validation of the prognostic groups defined by 50 corCNA genes and 30 corCNA genes in an independent SNU dataset. (*A,B*) Kaplan-Meir plot analyses for the recurrence free survival are performed on the prognostic groups stratified by the predictors trained by the expressions of 50 corCNA (*A*) or 30corCNA (*B*) genes in 139 HCC profiles. 5 out of 6 predictors for both 50 and 30 corCNA genes could stratify the prognostic groups of recurrence-free survival in SNU dataset ($P < 0.05$, log-rank test). Class prediction was performed by six class prediction algorithms (*i.e.*, CCP; Compound Covariate Predictor, LDA; Linear Discriminant Analysis, NC; Nearest Centroid, k-NN; k-Nearest Neighbor, SVM; Support Vector Machine).

Fig.S 3. siRNA-mediated knockdown of target genes . (*A,B*) Each of siRNAs (15 nM) targeting the indicated genes are transfected to HepG2 (*A*) or HuH-7 (*B*) cells for 48 hrs and the mRNA expression levels of the target genes are measured by real time qRT-PCR assay. Each bar indicates the relative expression levels of the target genes in siRNA-treated cells compared to that of non-targeting siRNA transfected control (NT-CTL) (** $P < 0.01$, *** $P < 0.001$). (*C*) For each of 5 effective target genes, individual two siRNAs were tested to validate the target silencing effect on HepG2 cell viability. After 96 hr transfection with the indicated siRNAs (15 nM), cell viability was assessed by MTT assay. Significance of the % cell viability compared to NT-CTL was determined by two-sample *T*-test (** $P < 0.01$, *** $P < 0.001$). Error bars indicate the mean ± standard deviation in three replicates.

Fig. S4. Functional enrichment of the *in trans* correlated genes. The significance of functional enrichment of the *in trans* correlated genes for each of the 50 corCNA genes are calculated based on cumulative hypergeometric test for Gene Ontology (GO) biological processes. In order to obtain representative and significantly enriched terms, at least three genes are considered for each term in our calculation. The significance of *P*-values are categorized into <0.05, <0.01, and <0.001, and hierarchical clustering was performed on the GO terms which are significant ($P < 0.05$) at least in two signatures. Commonly enriched GO terms are indicated (yellow box). N.S: not significant.