

Supporting Information Legends

Figure S1. The expression patterns identified from seven lung cancer datasets. (A) GSE4573. **(B)** GSE37745-SCC. **(C)** TCGA. **(D)** GSE32863. **(E)** GSE26939. **(F)** GSE11969. **(G)** GSE37745-AD. For each cancer dataset, we identified several distinct expression patterns across tumor progression stages. Among these results, up- and down-regulated patterns (color labeled) were defined and further analyzed in our study.

(TIF)

Figure S2. The distribution of overall slopes in cancer progression patterns. (A) The frequency of the absolute values of all overall slopes. **(B)** The count of up- and down-regulated patterns when threshold s equal to 0.05, 0.1 and 0.15.

(TIF)

Figure S3. The function associations between lung development and lung cancer progression patterns. (A) We identified significant GO terms enriched by genes with certain expression pattern in development and cancer, and further compared the function associations using hypergeometric test. The cell color showed the P-value performance after hypergeometric test. **(B)** The representative GO terms which were involved in more than two dataset pairs were shown.

(TIF)

Figure S4. Gene interaction networks in (A) devD-canU association. **(B)** devU-canD association. Colors denoted different lung cancer datasets and bias lines denoted macaque development. Known lung cancer genes and drug targets were labeled using underline and yellow background.

(TIF)

Figure S5. The interaction relationship between our risk signature and other previously published signatures. Using protein-protein interaction network from

HPRD (<http://www.hprd.org/>), we viewed the interaction relationship among all these lung cancer survival signatures. We identified direct protein-protein interactions (blue edges), and indirect interactions via at most one shared protein (green edges). In total, 119 interactions linked the 10 genes of 13-gene risk signature (pink nodes) to genes of other 9 studies. To simplify the graph, remaining edges among other signatures were faded.

(TIF)

Figure S6. The development-cancer pattern associations when slope cutoff equal to (A) (C) 0.1. (B) (D) 0.15.

(TIF)

Table S1. The detailed information of genes with inverse expression patterns in two development-cancer associations.

(XLS)

Table S2. Significant sub-pathway results from KEGG enriched in the two development-cancer associations.

(XLS)

Table S3. The associations between the development-cancer enriched pathways and lung cancer genes (or drug target genes).

(XLS)

Table S4. The detailed information of survival genes in Cell Cycle sub-pathway.

(XLS)

Table S5. The detailed information of the 13-gene risk signature.

(XLS)