**Supplementary Information:**

The pseudocode for random walk with restart

Random walk with restart algorithm of a gene-pathway pair

Input:
- V: genes of the pathway
- M: adjacency matrix of the pathway
- g: a given gene with CNA in the pathway, which is the seed node
- r: restart probability
- N: number of genes in the pathway

Begin algorithm:

01: for each $i \in V$, $d_{ii} = \sum m_{ij}$ ##Calculate the degree of genes in V of pathway

02: $W = D^{-1}M$ ##Construct the transfer matrix of the pathway

03: $P_0 = \{ p_1, p_2, ..., p_N \}$ ##Initiate the probabilities of genes in the pathway

04: if $V_i == g$

        $p_i = 1$

     else $p_i = 0$

05: $t = 0$ ##Start the random walk with restart from seed nodes

06: repeat

07:      if $t == 0$

08:          $P_{t+1} = (1-r) \times W \times P_t + r \times P_0$ ##Calculated the probabilities of genes at the (t+1)th step

09:      else

10:          $t = t+1$

11:          $P_{t+1} = (1-r) \times W \times P_t + r \times P_0$

12: until $\max(|P_{t+1} - P_t|) < 10^{-8}$ ##Determine steady-state of random walk with restart

end algorithm.

Output: the probability vector P_{t+1}

Supplementary Figure S1. The pseudocode of random walk with restart for identifying gene-pathway pairs.
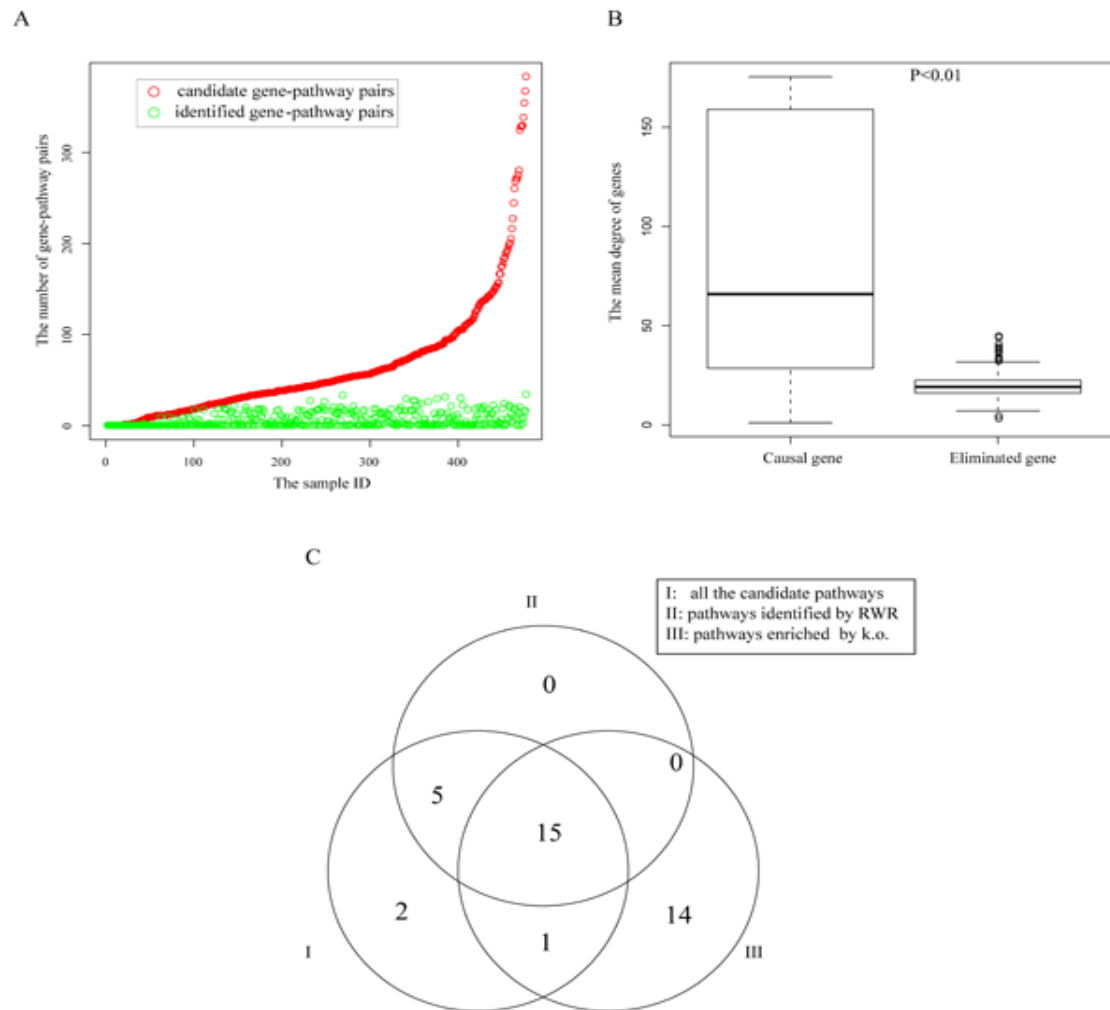
Supplementary Figure S2. (A) The number of candidate gene-pathway pairs (red) and identified gene-pathway pairs (green) across GBM samples. (B) The comparison of mean degree between identified genes and eliminated genes in the HPRD protein interaction network. (C) For the case of AKT1, the overlap among all candidate pathways (I), pathways identified by random walk with restart (RWR) (II) and pathways enriched by differential expressed genes using AKT1 knock-down data (III).
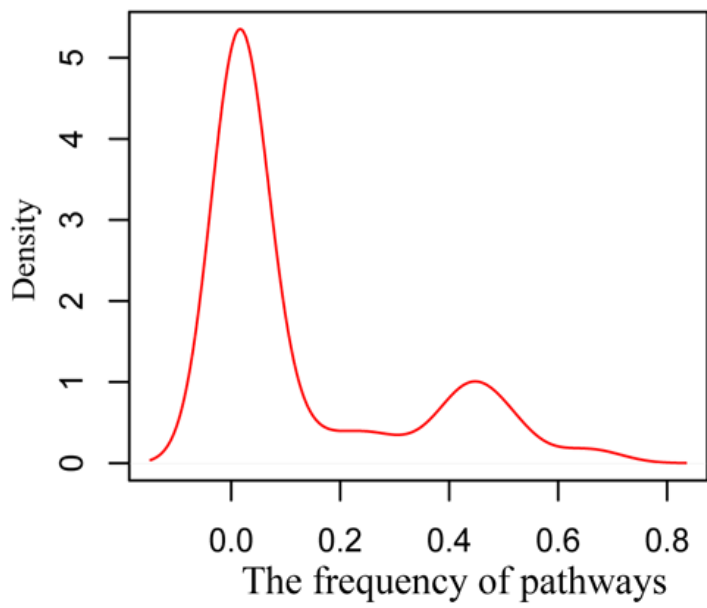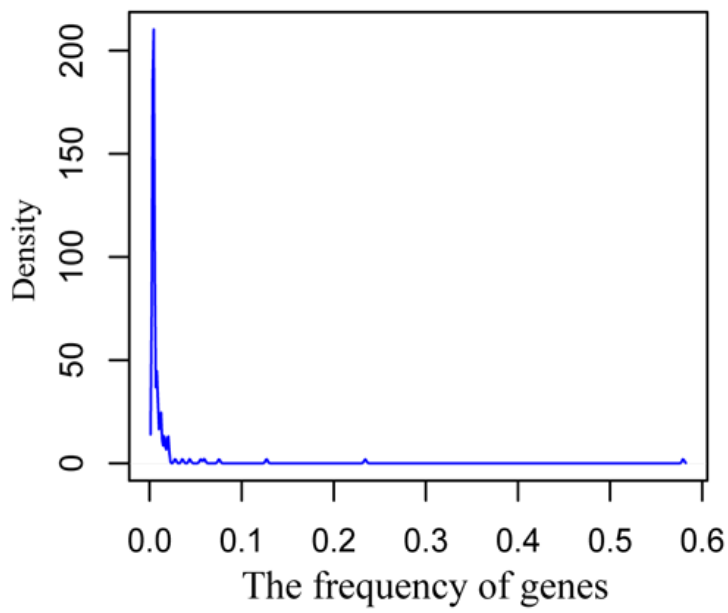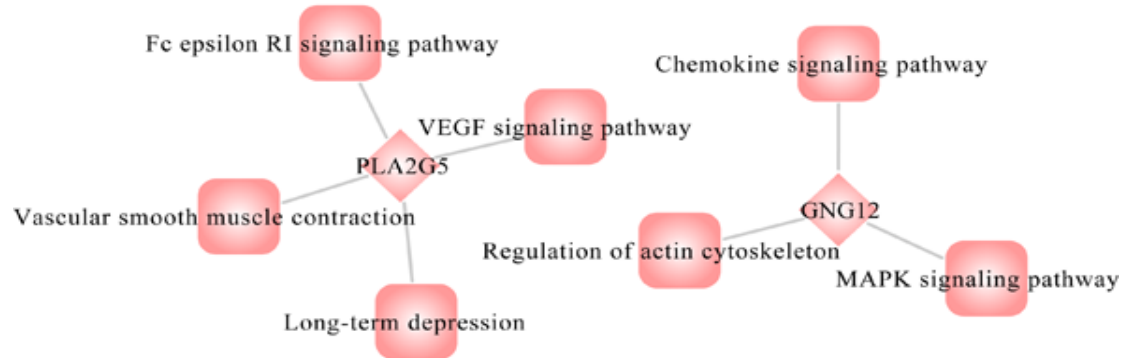
Supplementary Figure S3. The distribution of frequency of genes and pathways in gene-pathway pairs in GBM samples.
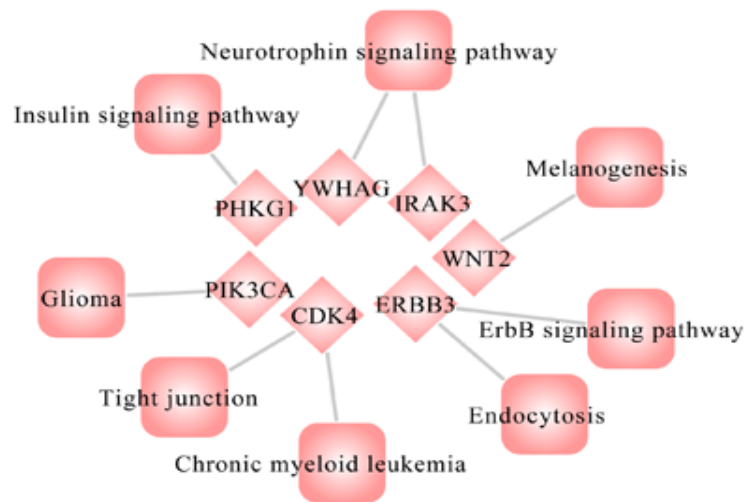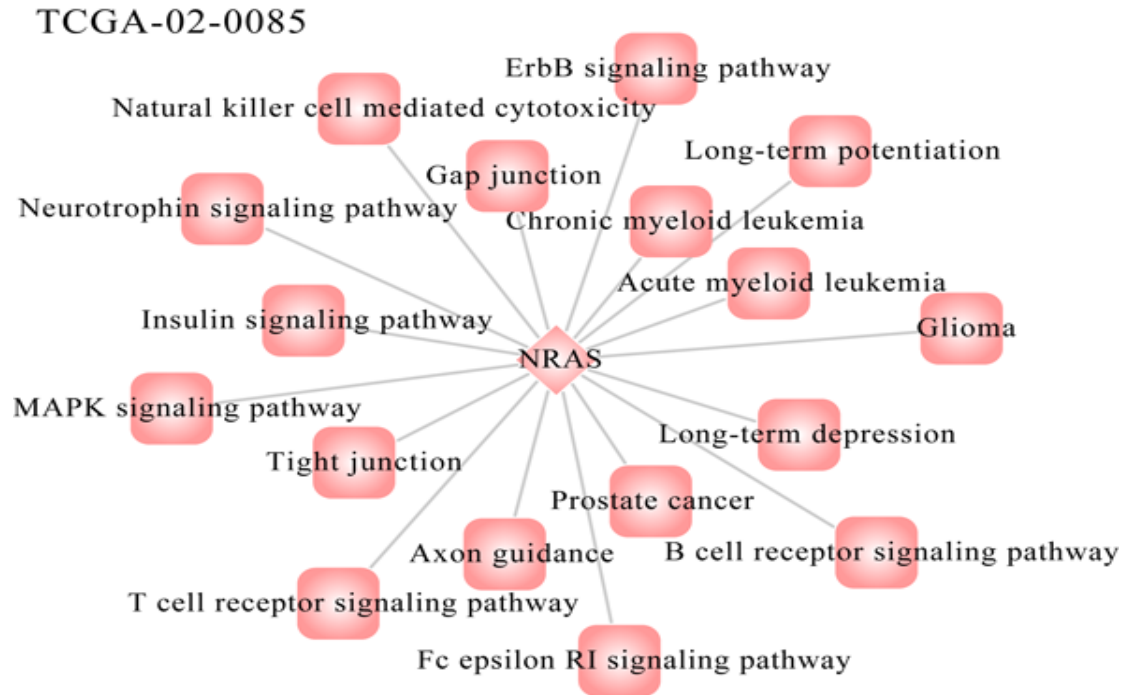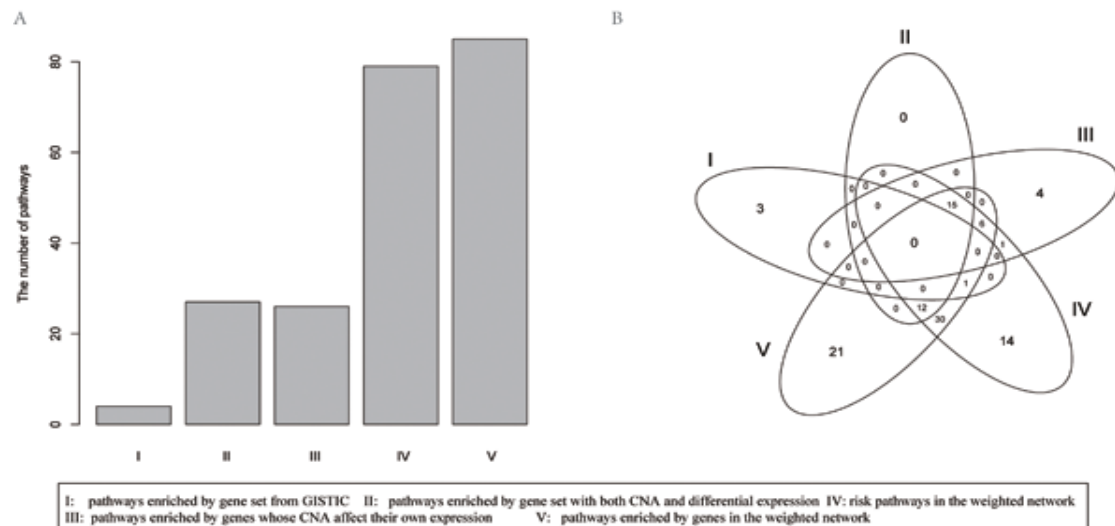
Supplementary Figure S4. The gene-pathway networks for two classical GBM individuals (TCGA-28-6450 and TCGA-06-6391).

TCGA-02-0085

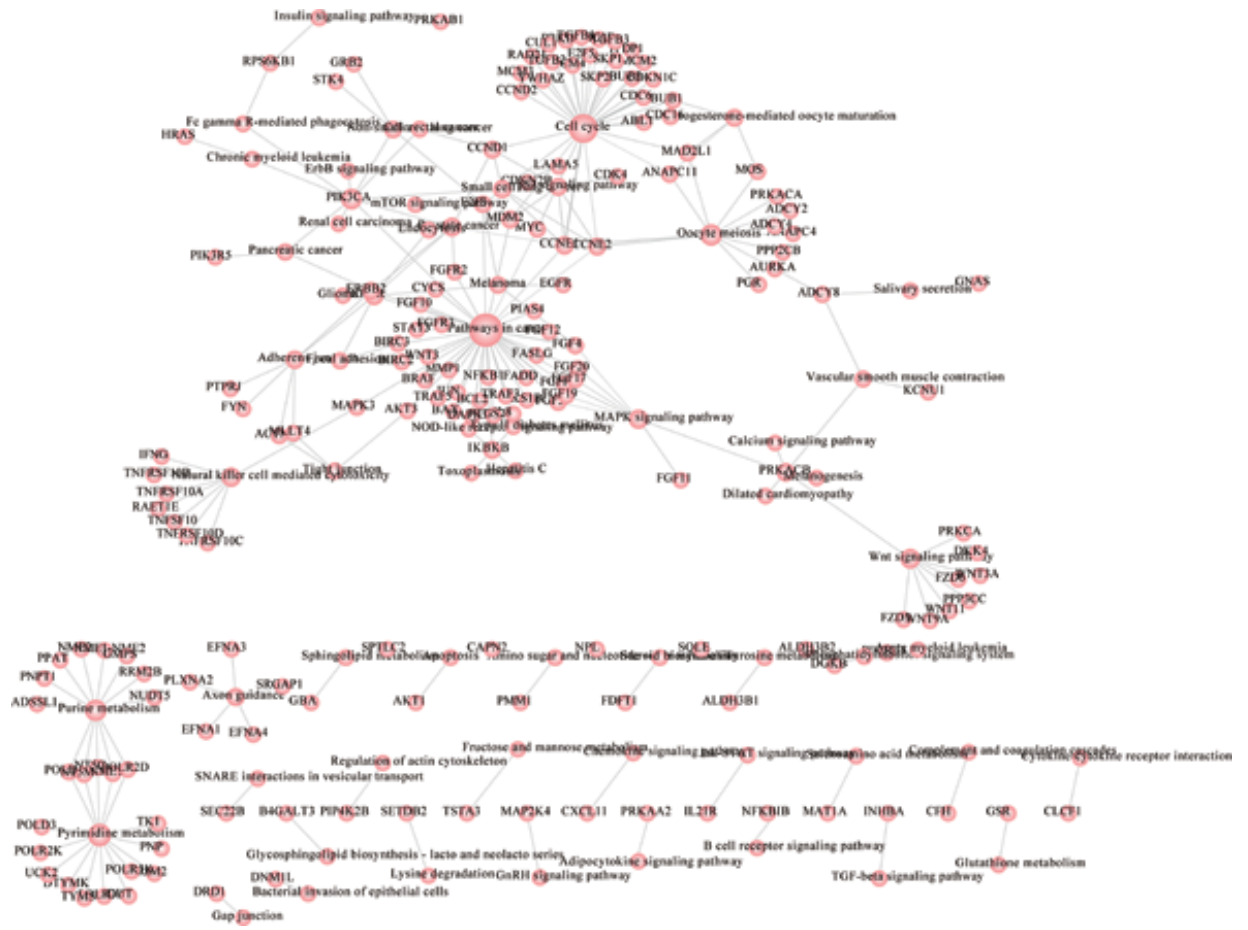**Supplementary Figure S5.** The NRAS-pathway network for a GBM individual (TCGA-02-0085).

I: pathways enriched by gene set from GISTIC II: pathways enriched by gene set with both CNA and differential expression IV: risk pathways in the weighted network
III: pathways enriched by genes whose CNA affect their own expression V: pathways enriched by genes in the weighted network
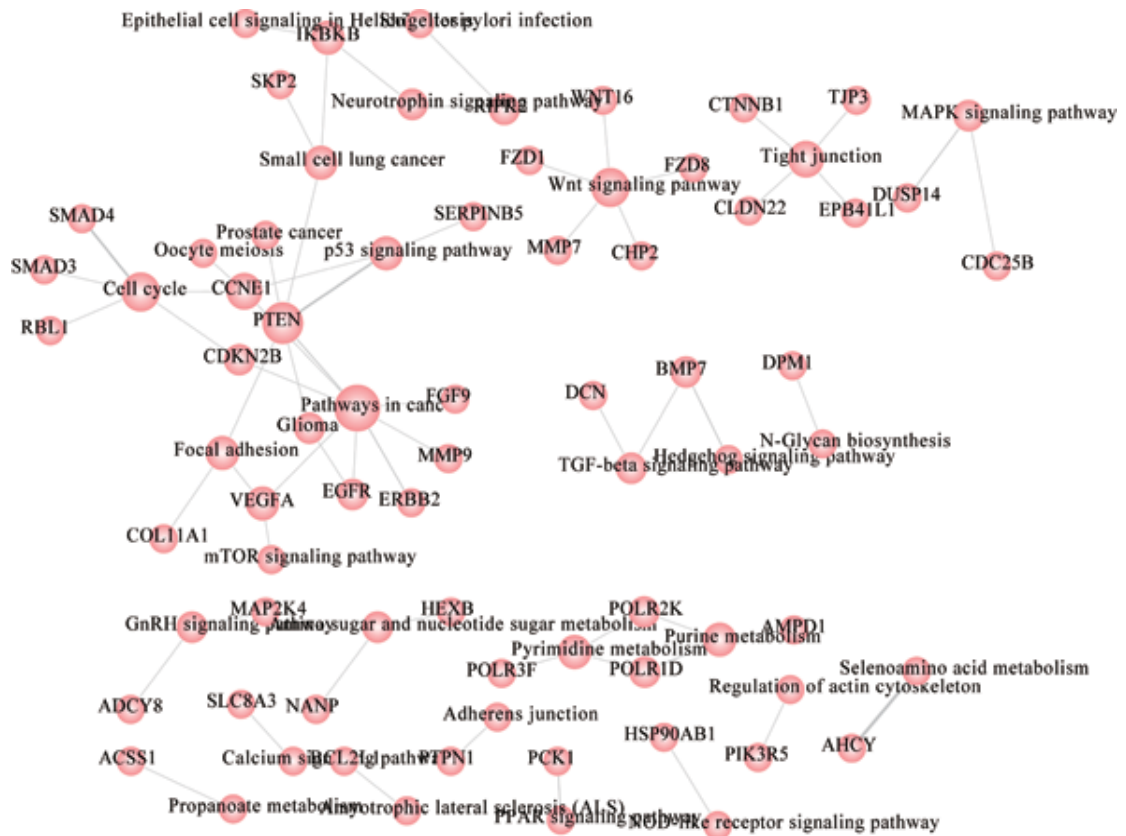
Supplementary Figure S6. (A) The bar plot of the numbers of pathways identified from different methods. (B) The overlap of pathways identified from different
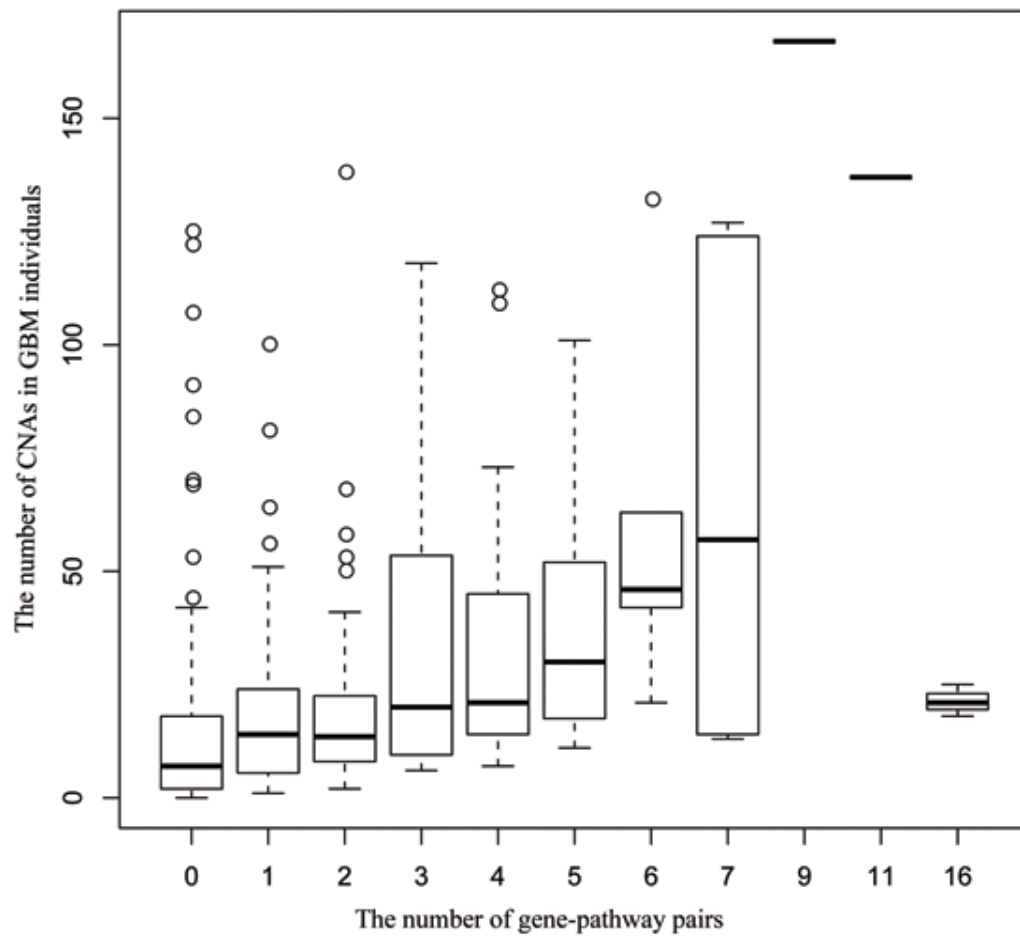
methods.

Supplementary Figure S7. The weighted gene-pathway network of breast cancer.

Supplementary Figure S8. The weighted gene-pathway network of colorectal adenocarcinoma.

Supplementary Figure 9. The numbers of CNAs in the GBM samples corresponding

to with different number of identified gene-pathway pairs.