

## **Legends to Supplemental Figures.**

### **Supplementary Figure 1.**

Workflow of the procedure to reconstruct the gene models and test their correlation with outcome.

### **Supplementary Figure 2.**

Connectivity plot (left panels) and topological maps (right panels) describing the results of ARACNe analysis in the seven datasets.

The connectivity plot describes the degree probability distribution, i.e. the number of connections (degree  $k$ , x-axis) versus the number of genes having that number of connections ( $P(k)$ , y-axis). The degree probability distribution shows a power-law tail suggesting that the underlying structure of the network is scale-free, i.e. most nodes have only a small number of interactions whereas there are a few highly connected nodes (hubs). In all datasets, at low connectivity values, the degree distribution loses its linear progression probably as a consequence of the limited number of genes.

The topological maps (“neighbors degree distribution”, left panels) show the average connectivity of the neighbors of a node ( $\langle k_{nn} \rangle$ , y-axis) as compared to the node connectivity (degree  $k$ , x-axis). For most dataset but GSE2658 and GSE9782, the topological maps suggests an assortative behavior of the networks (i.e., the nodes tend to connect with nodes with a similar connectivity), thus partly implying a hierarchical structure. Instead, networks from GSE2658 and GSE9782 data show a partially disassortative behavior, most likely due to the large amount of nodes with a single connection. This difference may be explained considering that GSE2658 and GSE9782 signals have been quantified using MAS5.0 on probe sets instead of RMA on genes.

### **Supplementary Figure 3.**

Cytoscape representation of the *BLNK* network reconstructed from gene expression data of GSE13591 series. Primary and secondary neighbors are in red and yellow, respectively.