**Figure S1: The relationship between positive signature and negative signature in the same node, related to Figure 1.**

The correlation between a positive signature's activities across the compendium and its corresponding negative signature's activities across the compendium has a close-to-normal distribution centered around 0.3. Only a small number of nodes have positive and negative signatures well correlated in their activities, indicating that the positive and negative signatures of the same node are most often not activated at the same time. Therefore, we considered the positive and negative sides of a node as two independent gene signatures rather than combining them.
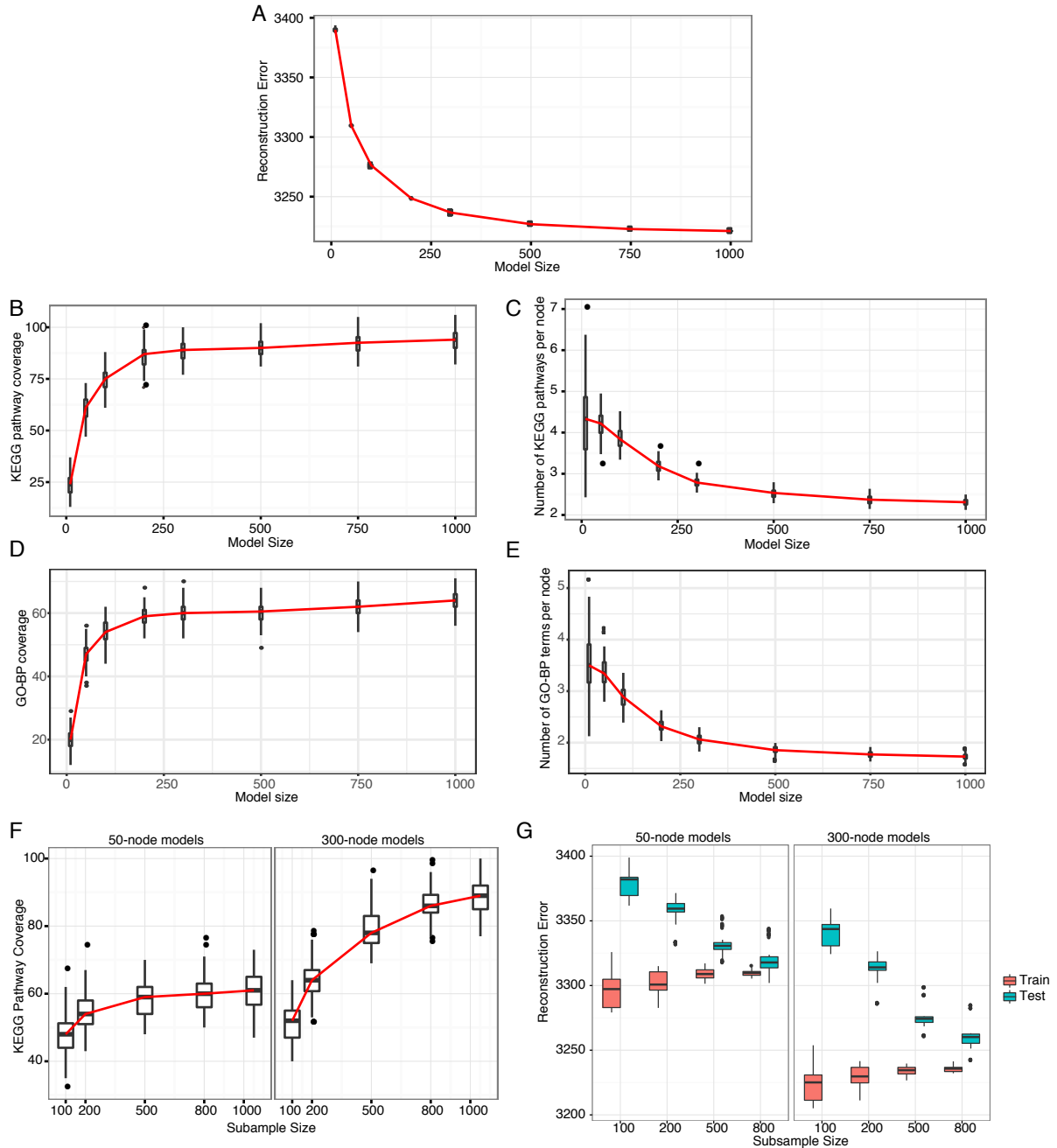
Figure S2: Knowledge- and data-driven heuristics for ADAGE, related to STAR Methods.

A: Data-driven model size heuristics on reconstruction error. As model size increases, the reconstruction error drops quickly at the beginning and levels out at size 300. The red line goes through the median reconstruction errors at each model size.

B: Knowledge-driven model size heuristics on KEGG pathway coverage. As model size increases, pathway coverage also increases at first and then levels out at size 300. The red line goes through the median value at each model size.

C: The relationship between model size and the number of KEGG pathways a node significantly associated with. Pathways associated with positive and negative signatures of a node were added together. When model is small, one node needs to account for multiple KEGG pathways.

As model size grows, more nodes become available and pathways also tend to spread into different nodes.

D: Knowledge-driven model size heuristic of GO-BP coverage.

E: The relationship between model size and the number of GO-BP terms with which a node is significantly associated.

F: Knowledge-driven sample size heuristics for pathway coverage. For 50-node models, pathway coverage increases with sample size and peaks at 500 samples. 300-node models cover more pathways than 50-node modes in general and maintain a slowly increasing trend of pathway coverage at the maximum sample size.

G: Data-driven sample size heuristics on reconstruction error. In both 50- and 300-node models, the reconstruction errors on the test set get closer to the reconstruction errors on the training set as sample size increases.
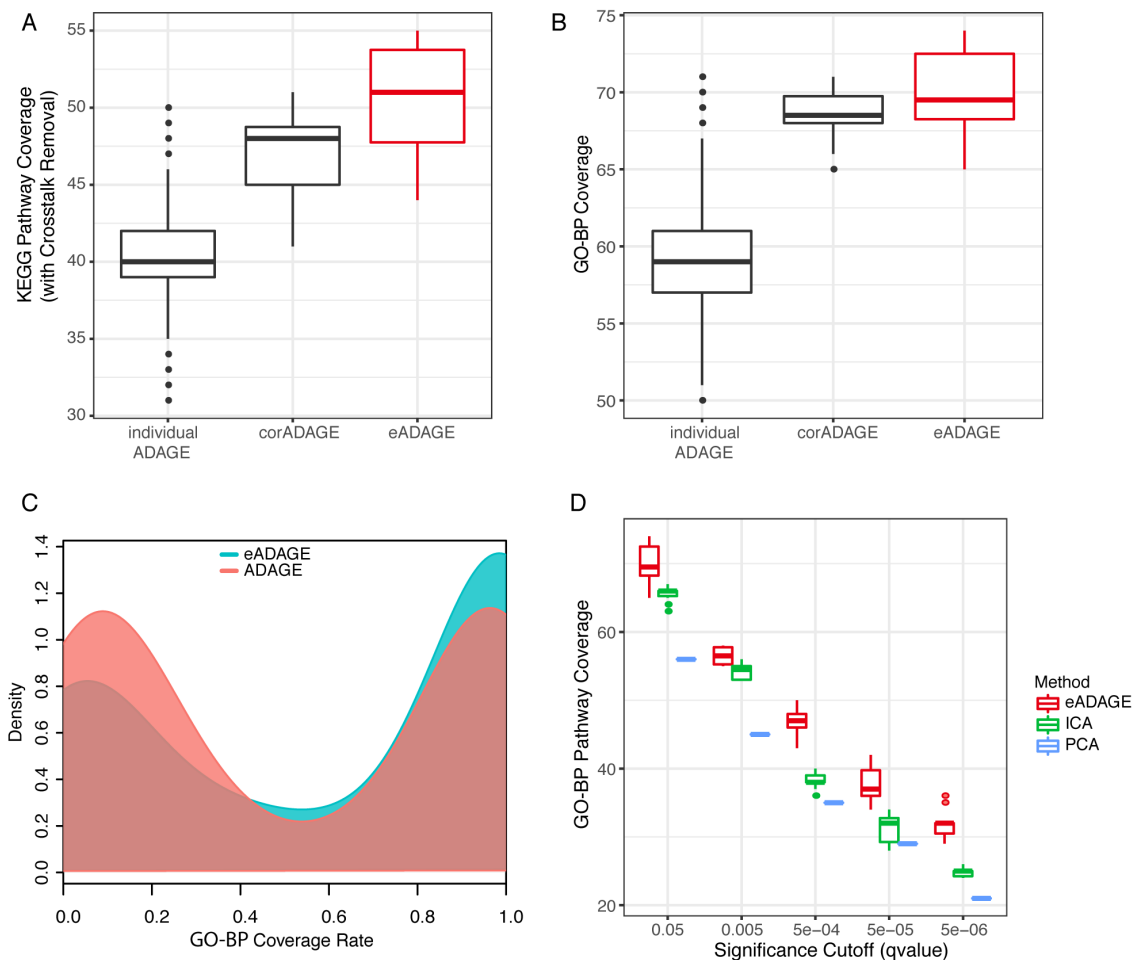
Figure S3: Additional performance evaluation of eADAGE, related to Figure 2.
A: Pathway coverage comparison between individual ADAGE and ensemble ADAGE after correcting pathway crosstalk effects. eADAGE models (n=10) covers significantly more pathways than both corADAGE (n=10) and ADAGE (n=1000).
B: GO term coverage comparison between individual ADAGE and ensemble ADAGE. Both corADAGE (n=10) and eADAGE (n=10) covers significantly more GO terms than individual ADAGE (n=1000).
C: The distribution of GO-BP term coverage rate in ADAGE models (n=1000) and eADAGE models (n=10). eADAGE shows a higher density in distribution on the high coverage side.
D: Comparison among PCA, ICA, and eADAGE in GO-BP term coverage at different significance levels.
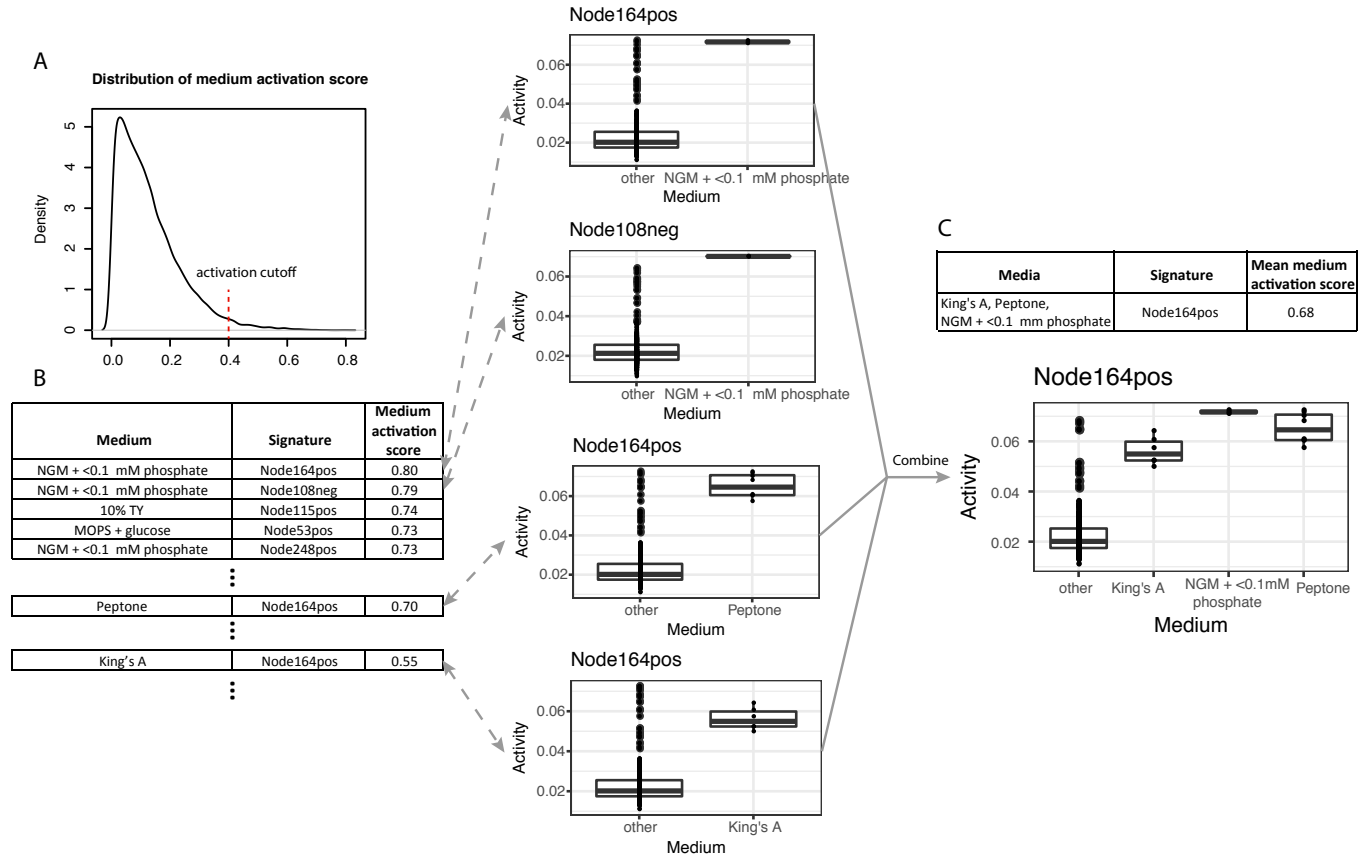
Figure S4: Medium activation score, related to Figure 3 and STAR Methods.

A: The distribution of medium activation score. The activation cutoff is set to 0.4.

B: Examples of signature activation in one medium. A high activation score indicates that the mean activity of samples in that medium is highly different from the mean activity of the rest samples.

C: An example of a signature (Node164pos) activation in multiple media. A mean activation score is calculated when a signature is active in multiple media.