



**Figure S1. Related to Figure 2.**

(a) Cross-validated accuracy for models trained with varying feature subset sizes (2–14 fetures) using recursive feature elimination with  $5 \times 5$  RFECV. Error bars represent the standard deviation of accuracy across repeated folds. (b) Feature importance ranking of the RFECV-selected gene subset based on the MeanDecreaseGini metric from the random forest model. The six hub genes identified in this study are among the top-ranked features.