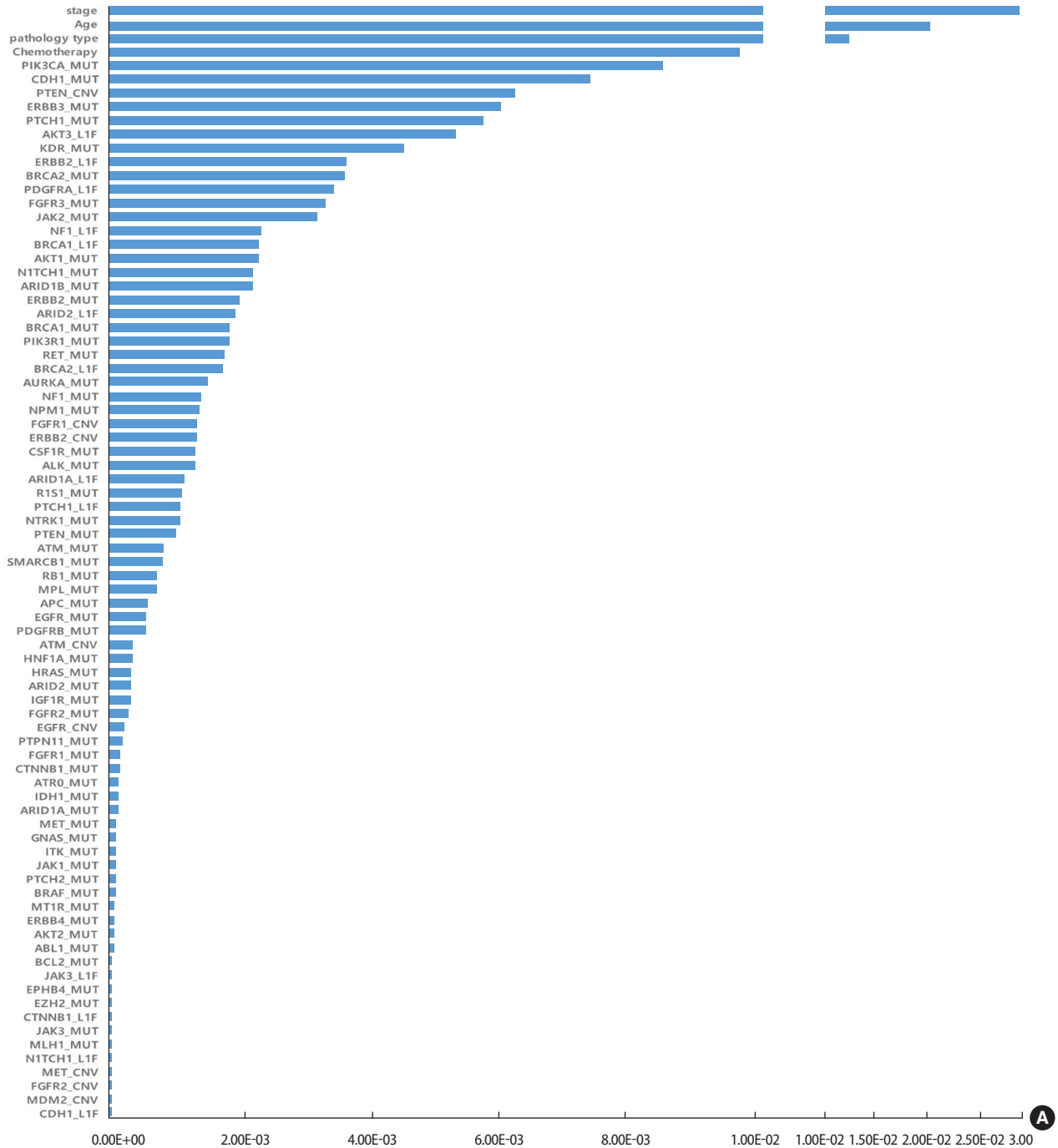
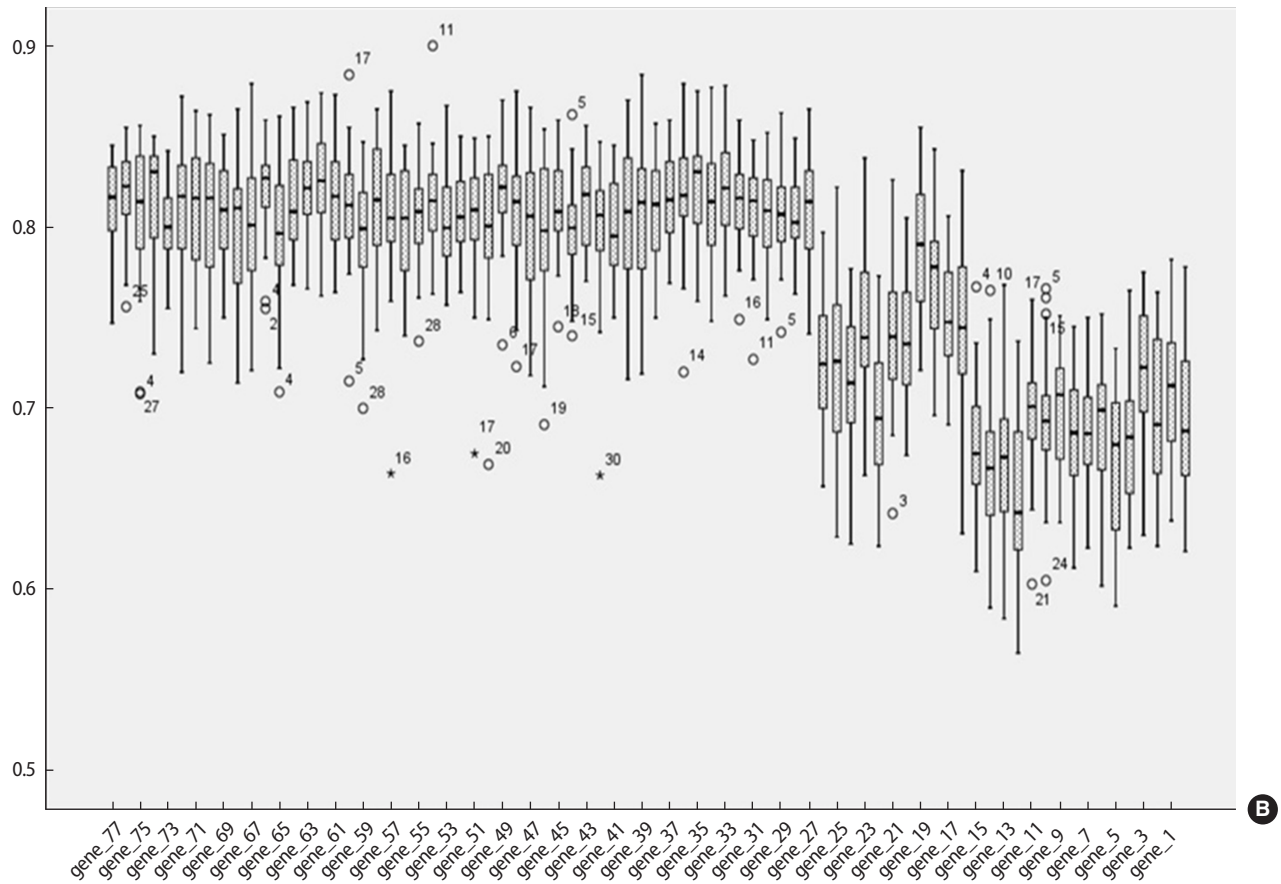


(*Continued to the next page*)



Supplementary Figure 1. (A) List of genes in order of importance score estimated by gradient boosting: A total of 243 features (alterations, including loss-of-function, mutation, and copy number variation in 81 genes) were used as input features. The target value was alive or dead at the 5-year mark. The genes were sorted in descending order of importance. (B) Selection for the optimal number of genes: Bootstrap resampling (n = 100) was performed by sequentially reducing the less important genetic features, where the training sets (85%) and their corresponding test sets (15%) were resampled 100 times. The performance of each model was evaluated using the AUC values. The AUC was sequentially reduced by reducing the number of features. The AUC values of the 27 important genes (0.81; 95%CI: 0.76–0.85) were not significantly different from those with a higher number of genes. However, the AUC values of 26 important genes (0.724; 95% CI: 0.80–0.66) were significantly lower than those of the 27 important genes ($p = 0.000$).

https://doi.org/10.14449/jbd.2023.11.2.39