**Supplementary Figure 1. Validation procedure.** A) For internal validation, 10-fold cross-validation is used. In each of the ten folds, the classifier, including the gene selection phase, is rebuilt from scratch. A single estimate of the molecular risk score is made for each of the 105 samples, based on the predictons from the left-out samples. This entire procedure is repeated 100 times. The final reported risk score for each biopsy (i.e. those in Fig. 2) is the average risk score over all 100 iterations. B) For external validation, the classifier is built from the full set of 105 late biopsies, then used to predict either early biopsies (validation set 1) or late biopsies from a different centre (validation set 2). C) Classifier built as in A, but using all 173 patients (early + late) from the U of Alberta and U of Illinois.
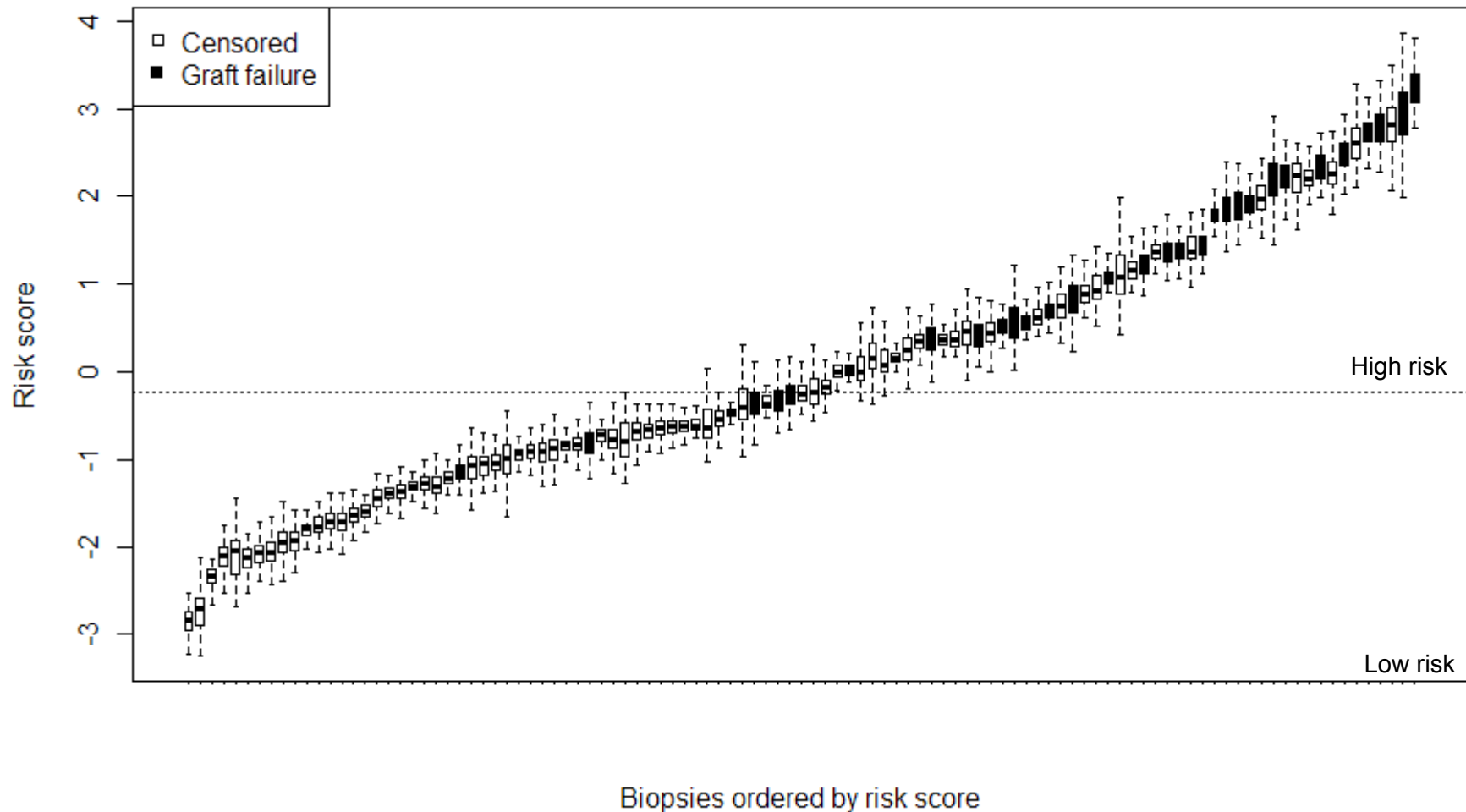
Supplementary Figure 2. Boxplots of molecular risk scores in individual biopsies over all 100 iterations of the classifier. See Figure 2 for details. Each biopsy's risk score is estimated 100 times in separate iterations of a 10-fold cross-validation procedure. The boxes represent the interquartile range of the 100 estimates. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from each box.
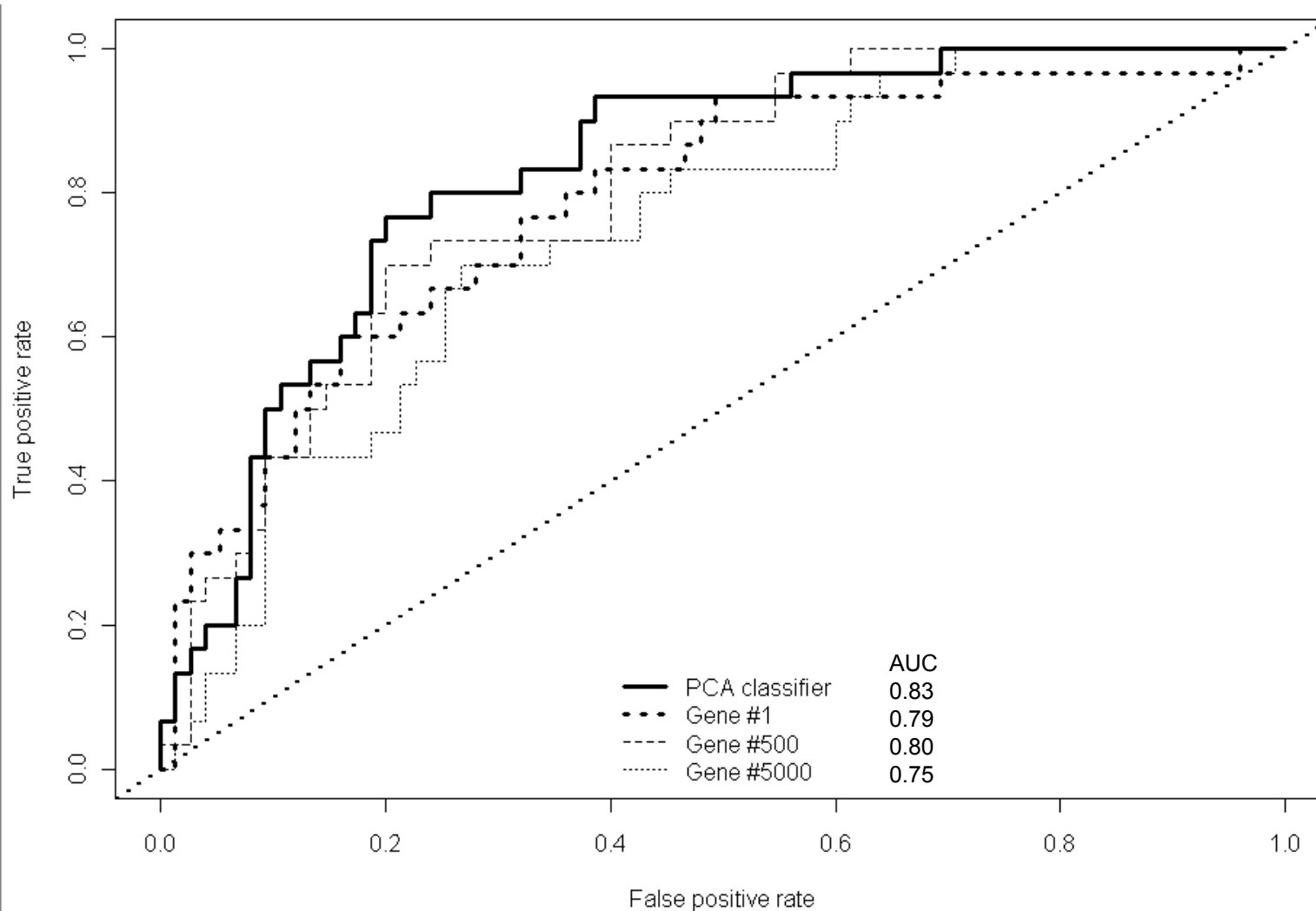
Supplementary Figure 3. ROC curves for single gene classifiers. The main classifier result based on the supervised PCA classifier is compared to those from single-gene classifiers. Ranking of genes in each iteration of the training sets was based on the p-value from a Cox regression.