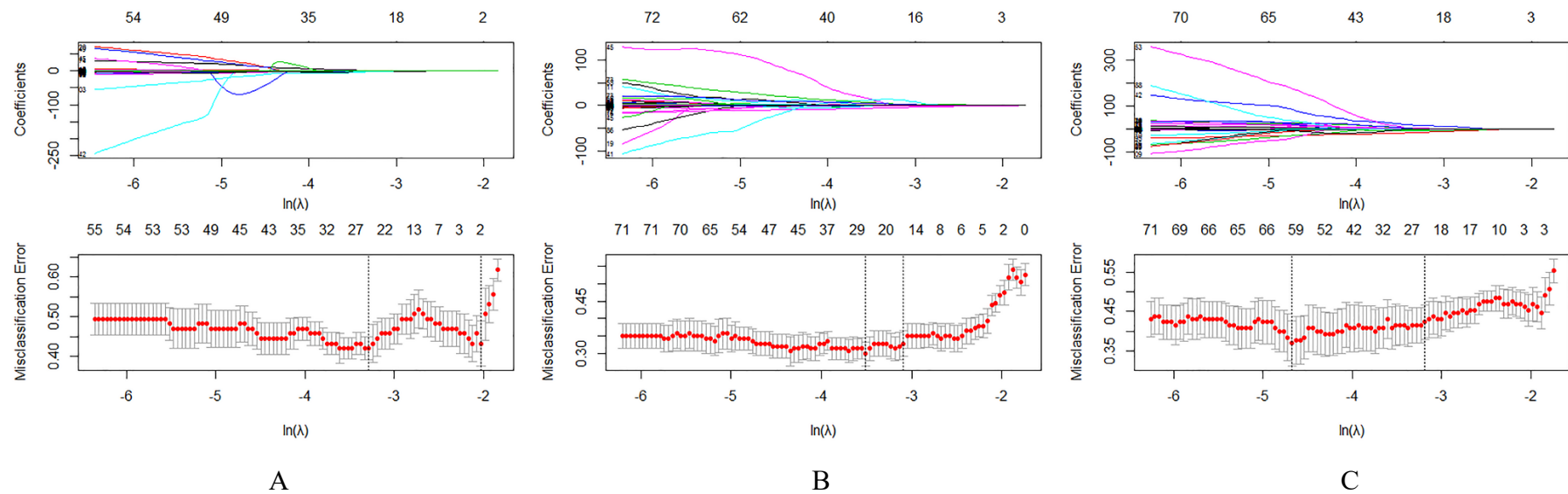**Supplementary Materials**

Supplementary Figure 1.

Fig.1 The epidermal growth factor receptor (EGFR) mutation status the least absolute shrinkage and selection operator (LASSO) binary logistic regression model for the feature selection

As shown, the vertical dotted line was drawn at the value selected by the 10-fold cross-validation, where the 23 optimal coefficients were obtained. (b) The relationship between the misclassification error and the parameter (ln(λ)) was visually shown. In order to avoid overfitting the model, the number of features was as few as possible. When the value ln(λ) increased to -3.19, the misclassification error reached the peak again with the appropriate number of features according to the 10-fold cross-validation.

Supplementary Figure 2.

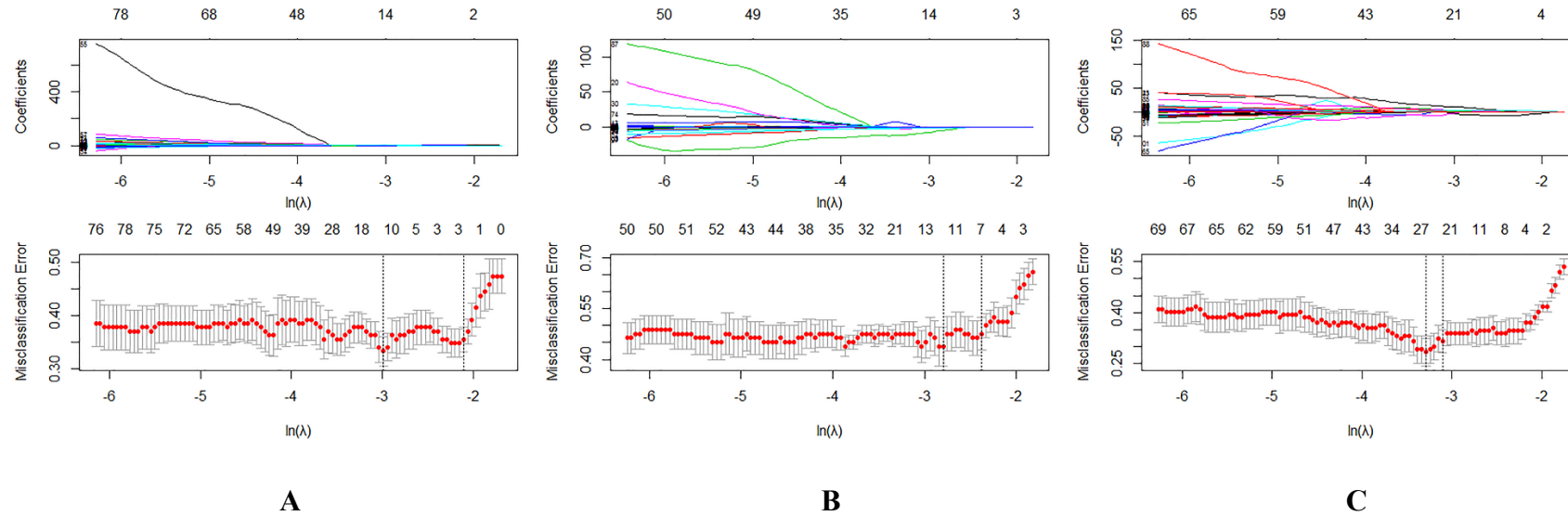|		
A	**B**	**C**

Fig.2 screening of a sensitive population treated with tyrosine kinase inhibitors (TKIs) the least absolute shrinkage and selection operator (LASSO) binary logistic regression model for the feature selection

As shown, the vertical dotted line was drawn at the value selected by the 10-fold cross-validation, where the 22 optimal coefficients were obtained. (b) The relationship between the misclassification error and the parameter ($\ln(\lambda)$) was visually shown. In order to avoid overfitting the model, the number of features was as few as possible. When the value $\ln(\lambda)$ increased to -3.1038, the misclassification error reached the peak again with the appropriate number of features according to the 10-fold cross-validation.