

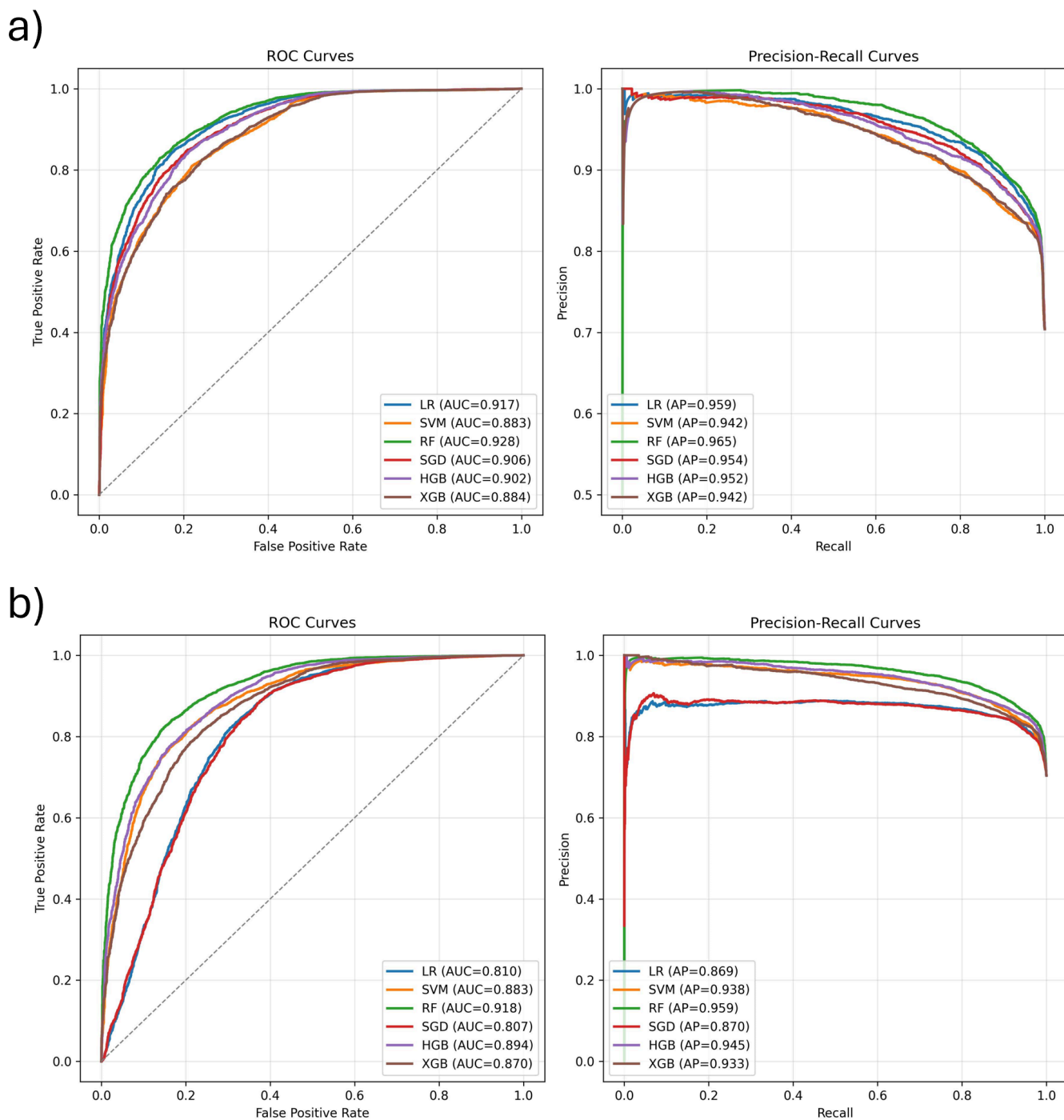


Figure S1: Receiver operating characteristic (ROC) curves (left) and precision-recall (PR) curves (right) comparing the performance of six classification models: logistic regression (LR), support vector machine (SVM), random forest (RF), stochastic gradient descent (SGD), histogram-based gradient boosting (HGB), and XGBoost (XGB) when 2.048 ECFP4 (a) and 15 theoretical features (b) have been used for training.