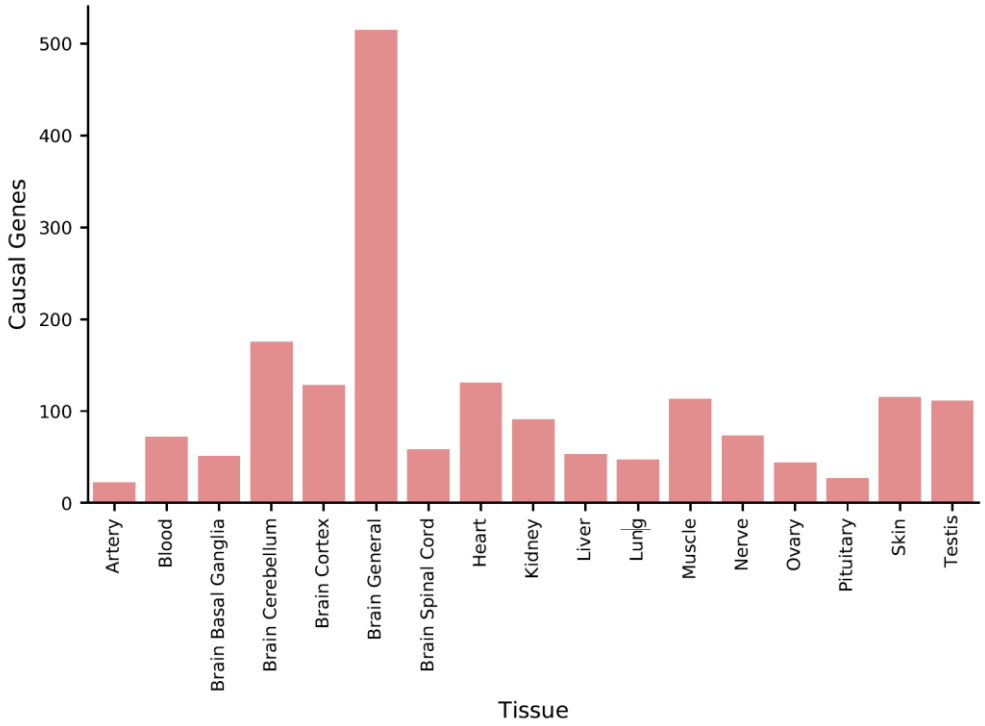
Supplementary Figures

Fig. S1. The number of pathogenic genes (A) and pathogenic variants (B) per tissue.

A.



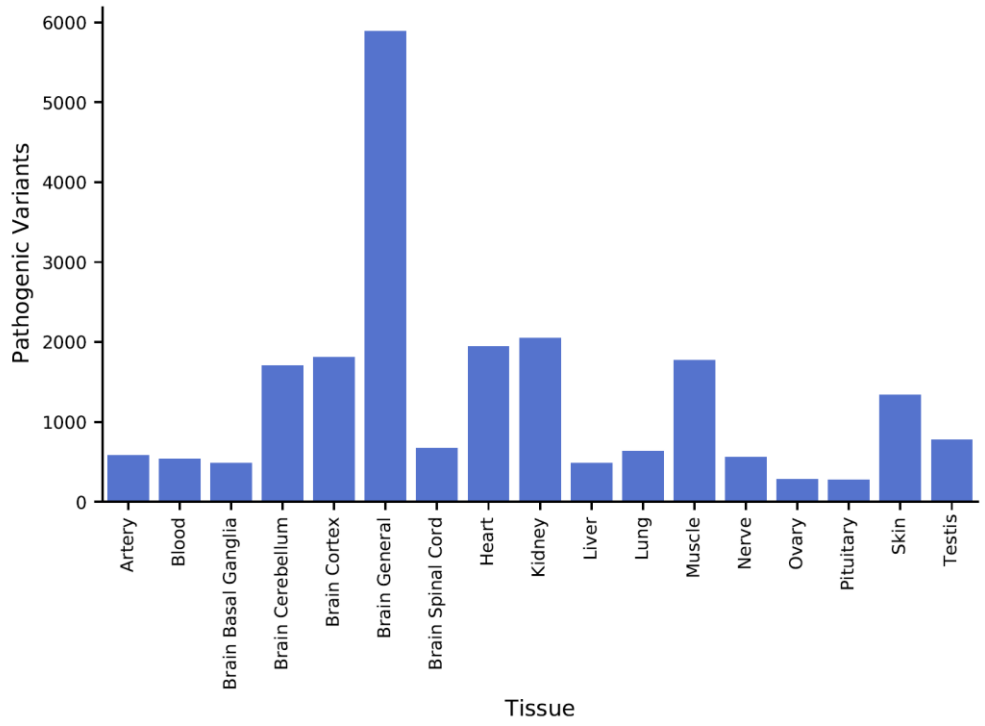
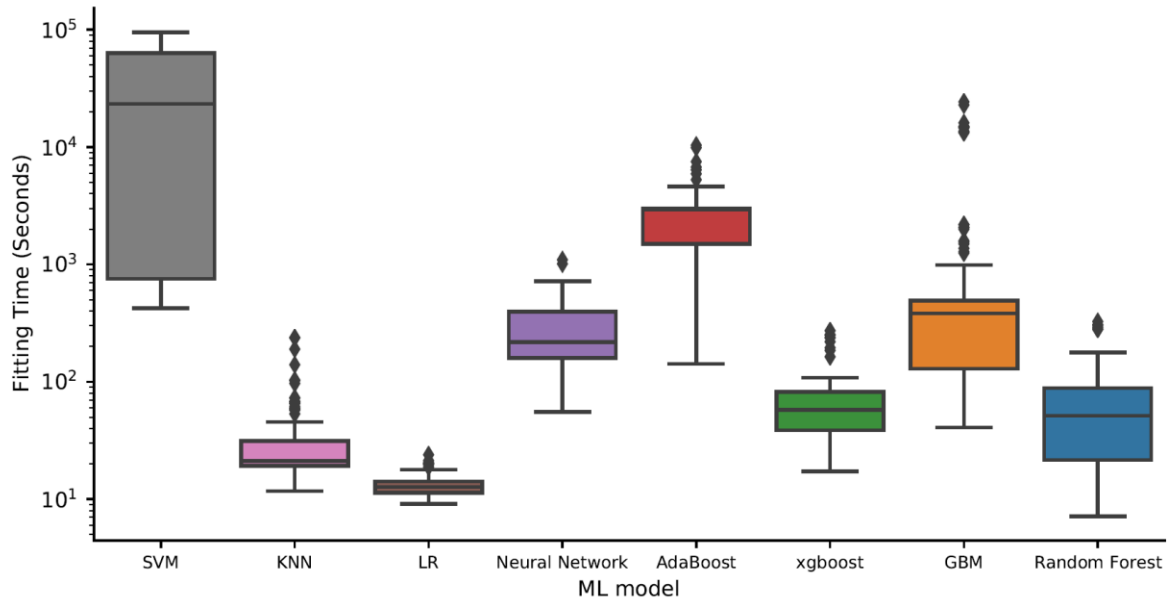


Fig. S2. Time measurements of eight machine-learning methods.

The training (A) and prediction (B) run times per method in 17 tissue models. Random forest training and prediction times were reasonable compared to other methods (median of 52 and 0.88 seconds, respectively).

SVM=support vector machine; KNN=K-nearest neighbors; LR=logistic regression; GBM=gradient boosting machine.

A.



B.

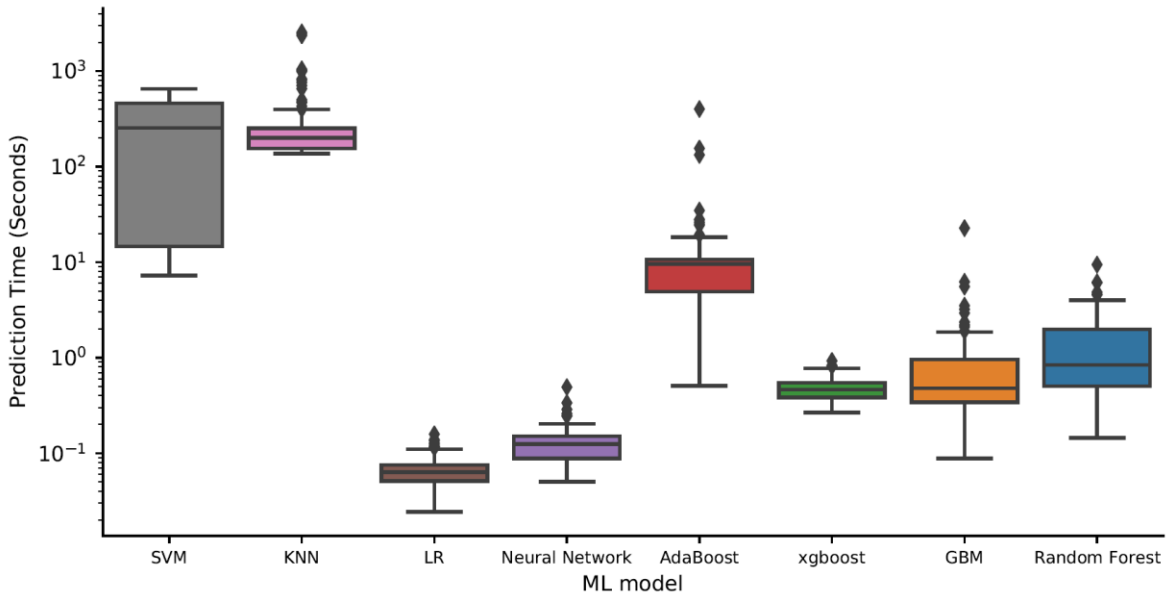


Fig. S3. The contribution of different types of tissue-specific features to each tissue model.

Color scale and dot size reflect the normalized importance values.

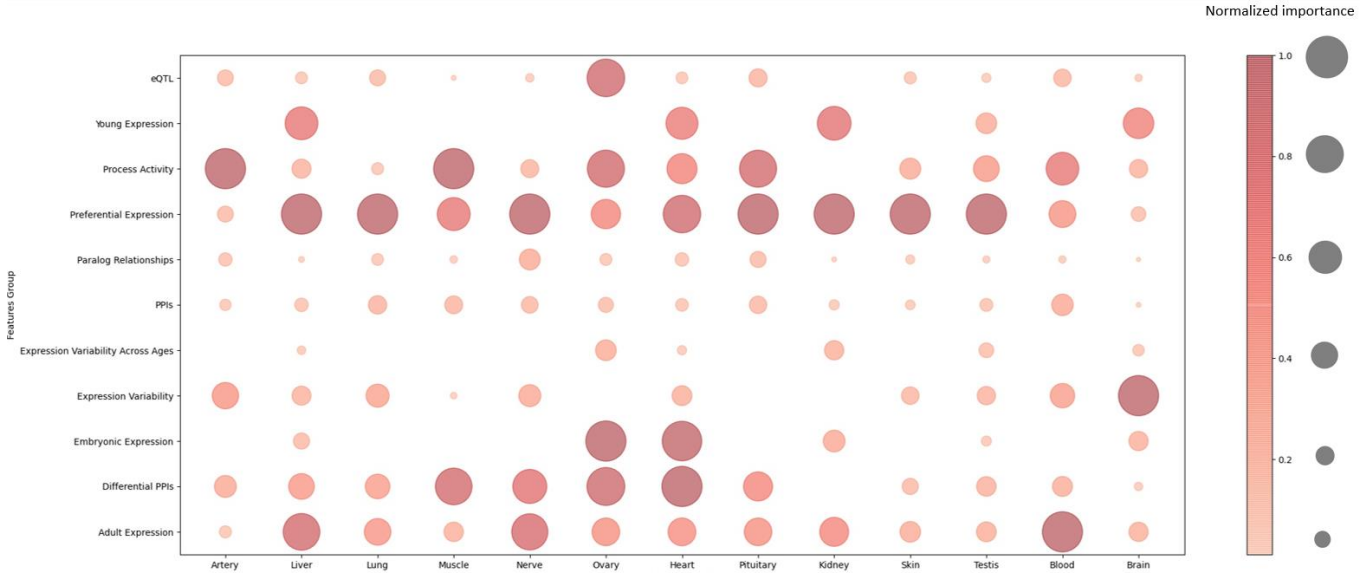


Fig. S4. The Spearman correlation between the auPRC of TRACEvar tissue models and the number of pathogenic variants per tissue.