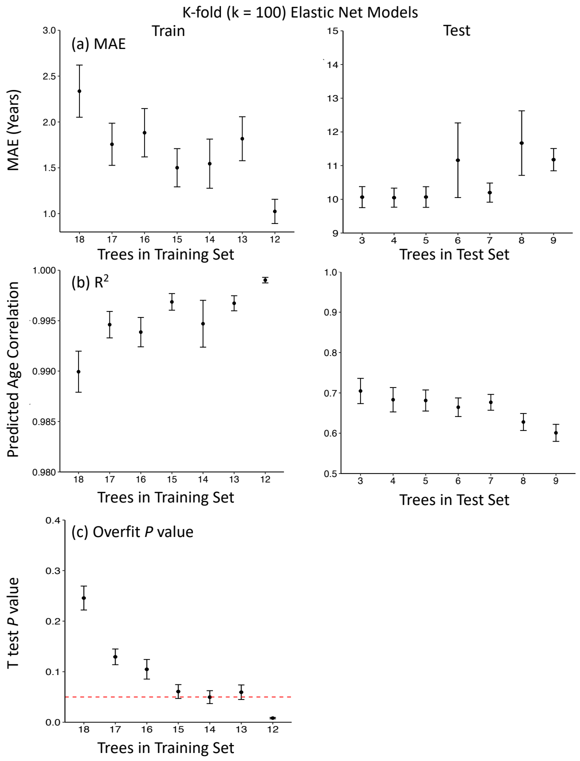
MOLECULAR ECOLOGY
RESOURCES

Supplemental Information for:

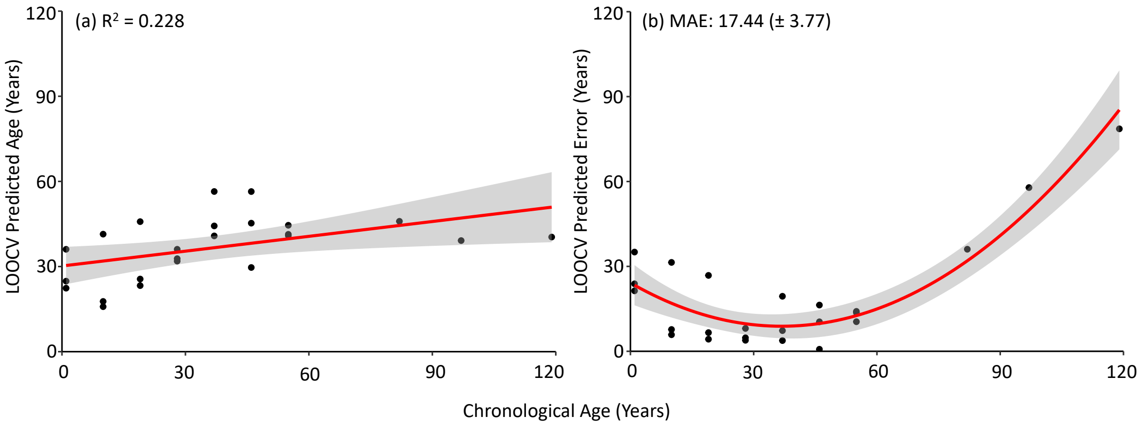
Development of DNA methylation-based epigenetic age predictors in loblolly pine (*Pinus taeda*)

Steven T. Gardner, Emily M. Bertucci, Randall Sutton, Andy Horcher, Doug Aubrey, Benjamin B. Parrott

Table of Contents:



Supplemental Figure 1. Performance metrics of elastic net models based on numbers of trees used in model training and test sets, using K-fold (k = 100) validations. (A) Following k=100 validations, as the number of trees in training sets decreased and test set numbers increased, MAE values decreased for training set trees (*P* = 0.0023) and increased for test set trees (*P* = 0.0397). (B) Correlations of chronological age estimates for trees in training sets increased as the number of trees used in each training set decreased (P = 0.009). Conversely, correlations of chronological age estimates for trees in test sets decreased as the number of trees in test sets increased (*P* = 0.0014). (C) As the number of trees used in training sets decreased, *P* values of t tests used to assess over fitness of each model to training set trees decreased (*P* < 2e-16), and approached significance (indicated by the dashed red line).



Supplemental Figure 2. Accuracy and precision of age predictions using elastic net models to initially predict chronological age for n = 24 *P. taeda* trees (ages 1-119 years) using a leave-one-out-cross-validation (LOOCV) approach. (A) There was poor correlation between predicted age and chronological age ($R^2 = 0.23$). (B) Precision of predicted ages compared to chronological ages of trees was assessed as mean absolute error (MAE) in years. MAE increased with chronological age, with the three oldest individuals (ages 82, 97, and 119 years) having the highest predicted errors.