

Supplementary Data

Importance scores for iterative Random Forest (iRF) models

Importance scores are reported for each iRF model described in the manuscript.

Model 1: Alkene Minus (AM) vs Alkene Plus (AP) Classification

This classifier model was trained using all accessions included in the study. Model performance was evaluated using out-of-bag (OOB) estimation, yielding an OOB error rate of approximately 10%, indicating strong discrimination between AM and AP phenotypic classes.

Model 2: Z-9-Heptacosene Regression (AP individuals only)

This regression model was trained exclusively on AP individuals to evaluate predictors of quantitative variation in Z-9-heptacosene abundance.

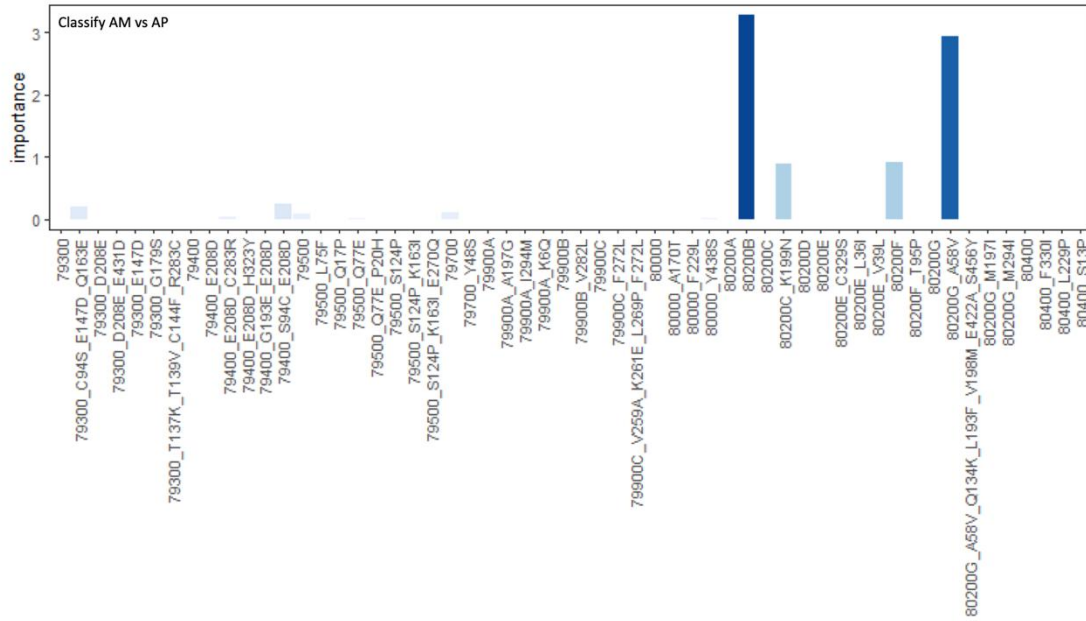
Model 3: Z-9-Pentacosene Regression (AP individuals only)

This regression model was trained exclusively on AP individuals to evaluate predictors of quantitative variation in Z-9-pentacosene abundance.

Importance scores reflect the relative contribution of predictors to model performance and are scaled according to the internal iRF feature importance metric.

Regression model results should be interpreted cautiously given the complexity of quantitative alkene abundance traits.

iRF Important scores using all AM and AP accessions



iRF Important scores using only AP accessions

