

**Supplementary Table S1:** The hyperparameters learned through training and the description for different models using *Arabidopsis* Training Data

Model	Parameter tuned	Description
Logistic Regression	penalty: l2, C: 1, solver: saga, max iterations: 500	<b>Penalty</b> is a regularization parameter; <b>C</b> is the inverse of regularization strength; <b>solver</b> is the optimization algorithm; <b>max iterations</b> define the number of iterations the algorithm performs on data
Support Vector Machine	C: 0.5, kernel: linear, gamma: scale, degree: 4	<b>C</b> is the regularization parameter; <b>kernel</b> function is used to map data to higher dimensions; <b>gamma</b> controls the shape of the decision boundary; <b>degree</b> defines degree of the polynomial
Decision Tree	criterion: gini, min samples leaf: 5, min samples split: 5	<b>Criterion</b> is used to measure the quality of the split; <b>min samples leaf</b> is the minimal number of samples in a leaf. <b>min samples split</b> is the minimum number of samples that are required to split a node in tree.
K-Nearest Neighbors	metric: Manhattan, n neighbors: 6	<b>Metric</b> is used to calculate the distance between datapoints; <b>n neighbor</b> defines the number of neighbors considered for prediction
Random Forest	maximum depth: 20, min samples leaf: 5, min samples split: 2, estimators: 50	<b>Maximum depth</b> is the depth of each Decision Tree, <b>min samples leaf</b> is the minimal number of samples in a leaf and <b>min samples split</b> is the minimum number of samples required to split a node in tree, <b>estimators</b> are the number of Decision Trees constructed
Extremely Randomized Trees	maximum depth: 20, min samples leaf: 1, min samples split: 2, estimators: 50	<b>Maximum depth</b> is the depth of each Decision Tree, <b>min samples leaf</b> is the minimal number of samples in a leaf and <b>min samples split</b> is the minimum number of samples required to split a node in tree; <b>estimators</b> are the number of Decision Trees constructed
AdaBoost	estimators: 100, learning rate: 0.1	<b>Estimators</b> value is the number of Decision Trees constructed; <b>learning rate</b> controls the contribution of the weak classifier in final ensemble
Gradient Boosting	estimators: 100, learning rate: 0.1	<b>Estimators</b> value is the number of Decision Trees constructed; <b>learning rate</b> controls the contribution of the weak classifier in final ensemble
Bagging Classifier	estimators: 100, max samples: 10, max features: 1	Estimators value is the number of Decision Trees constructed; <b>max samples</b> define the maximum samples to use for each copy; <b>max features</b> define the maximum features to use for each copy of the model