

## Supplementary materials

**Supplementary Table 1.** Multicollinearity assessment and key correlations VIF analysis for all 17 phenotypic predictors and correlation coefficients for traits showing  $r > 0.40$ . All VIF values were below 3.2, indicating acceptable multicollinearity levels

| Trait_Pair | Correlation_Coefficient | P_Value   |
|------------|-------------------------|-----------|
| SRL-RDW    | 0.67                    | 0.001     |
| TW-GYPP    | 0.58                    | 0.003     |
| RDW-RFW    | 0.72                    | 0         |
| MRL-SRL    | -0.45                   | 0.012     |
| PH-NTP     | 0.34                    | 0.045     |
| DTF-DM     | 0.78                    | 0         |
| HI-GYPP    | 0.43                    | 0.018     |
| Trait      | VIF_Value               | Tolerance |
| MRL        | 2.34                    | 0.427     |
| RV         | 1.89                    | 0.529     |
| RFW        | 3.12                    | 0.32      |
| RDW        | 2.67                    | 0.374     |
| RSR        | 1.45                    | 0.69      |
| NTP        | 1.78                    | 0.562     |
| PH         | 2.91                    | 0.344     |
| ESPP       | 1.56                    | 0.641     |
| DM         | 1.23                    | 0.813     |
| TW         | 1.87                    | 0.535     |
| DTF        | 2.45                    | 0.408     |
| SL         | 2.03                    | 0.493     |
| PL         | 1.67                    | 0.599     |
| NPPP       | 2.34                    | 0.427     |
| SRL        | 2.89                    | 0.346     |
| HI         | 1.34                    | 0.746     |
| ML         | 1.92                    | 0.521     |

**Supplementary Table 2.** Bootstrap confidence intervals 95 % confidence intervals for significant regression coefficients based on 1000 bootstrap resamples, confirming coefficient stability and statistical robustness

| Trait | Estimate | Bootstrap_Mean | Bootstrap_SE | CI_Lower_95 | CI_Upper_95 | Bias  | Bootstrap_Iterations |
|-------|----------|----------------|--------------|-------------|-------------|-------|----------------------|
| SRL   | 35.091   | 35.124         | 3.892        | 28.142      | 42.04       | 0.033 | 1000                 |
| RDW   | 0.108    | 0.109          | 0.021        | 0.071       | 0.146       | 0.001 | 1000                 |
| TW    | 0.172    | 0.174          | 0.041        | 0.095       | 0.249       | 0.002 | 1000                 |

**Supplementary Table 3.** Cross-validation performance comparison 10-fold cross-validation results comparing MLR, LASSO and random forest models, including overfitting assessment and generalization capacity

| <b>Model</b>               | <b>CV_R_Squared</b> | <b>CV_RMSE</b> | <b>CV_MAE</b> | <b>Training_R_Squared</b> | <b>Training_RMSE</b> | <b>CV_Folds</b> | <b>Overfitting_Index</b> |
|----------------------------|---------------------|----------------|---------------|---------------------------|----------------------|-----------------|--------------------------|
| Multiple Linear Regression | 0.891               | 0.315          | 0.245         | 0.986                     | 0.296                | 10              | 0.095                    |
| LASSO Regression           | 0.874               | 0.342          | 0.267         | 0.923                     | 0.312                | 10              | 0.049                    |
| Random Forest              | 0.729               | 1.032          | 0.823         | 0.967                     | 0.458                | 10              | 0.238                    |