**Table S1 –** *p*-value for differences in prediction accuracy between tested procedures and *GBLUP*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Trait | Pop. | Liberal test (Naïve t-test) | | | | |
| **a) Target included in CS** | | | **b) Target omitted from CS** | |
| ***GBLUP-Target*** | ***MPM-Mixture*** | ***MPM-Matérn*** | ***MPM-Mixture*** | ***MPM-Matérn*** |
| PH | WS4U-C2 | 0.019 | 0.491 | 0.114 | 0.002 | <0.001 |
| Liberty-C2 | 0.390 | 0.139 | 0.534 | 0.063 | <0.001 |
| U4X-N | <0.001 | 0.315 | 0.157 | 0.020 | 0.678 |
| L4X-NE | <0.001 | <0.001 | 0.034 | 0.838 | <0.001 |
| HD | WS4U-C2 | 0.969 | 0.083 | 0.079 | 0.313 | 0.005 |
| Liberty-C2 | 0.898 | 0.069 | 0.508 | 0.148 | 0.065 |
| U4X-N | 0.758 | <0.001 | <0.001 | <0.001 | 0.002 |
| L4X-NE | <0.001 | <0.001 | 0.065 | <0.001 | <0.001 |
| St | WS4U-C2 | 0.025 | 0.541 | 0.568 | 0.006 | 0.008 |
| Liberty-C2 | 0.004 | 0.028 | 0.098 | 0.815 | 0.616 |
| U4X-N | 0.459 | 0.621 | 0.764 | 0.803 | <0.001 |
| L4X-NE | <0.001 | 0.113 | 0.139 | <0.001 | <0.001 |
| Trait | **Pop.** | **Conservative test (Corrected t-test of Nadeau and Bengio 2003)** | | | | |
| **a) Target included in CS** | | | **b) Target omitted from CS** | |
| ***GBLUP-Target*** | ***MPM-Mixture*** | ***MPM-Matérn*** | ***MPM-Mixture*** | ***MPM-Matérn*** |
| PH | WS4U-C2 | 0.309 | 0.778 | 0.507 | 0.147 | 0.050 |
| Liberty-C2 | 0.723 | 0.535 | 0.798 | 0.431 | 0.126 |
| U4X-N | 0.100 | 0.678 | 0.555 | 0.314 | 0.865 |
| L4X-NE | 0.018 | 0.067 | 0.363 | 0.934 | <0.001 |
| HD | WS4U-C2 | 0.987 | 0.464 | 0.458 | 0.677 | 0.212 |
| Liberty-C2 | 0.958 | 0.441 | 0.786 | 0.545 | 0.434 |
| U4X-N | 0.900 | 0.050 | 0.113 | 0.023 | 0.168 |
| L4X-NE | 0.004 | 0.017 | 0.433 | 0.016 | <0.001 |
| St | WS4U-C2 | 0.331 | 0.802 | 0.815 | 0.226 | 0.244 |
| Liberty-C2 | 0.203 | 0.343 | 0.486 | 0.924 | 0.837 |
| U4X-N | 0.761 | 0.840 | 0.902 | 0.919 | 0.094 |
| L4X-NE | 0.083 | 0.505 | 0.536 | 0.059 | 0.017 |

Validation scheme: (a) whole-sample calibration, including the target population in the calibration set; (b) cross-population calibration, omitting all individuals from the target population in the calibration set. Trait: plant height (PH), heading date (HD) or standability (St). Population: Population used as target for prediction. Prediction accuracies are averaged over 20 cross-validation replicates. Models differ by prediction procedure, under the same mean structure (*Intercept*: intercept-only model). *GBLUP*: Single-population model; *GBLUP-Target*: GBLUP model where the CS includes only the individuals from the same population as the TS; *MPM*: Multi-population model with among-population correlations based on admixture coefficients (*MPM-Mixture*) or PC distances (*MPM-Matérn*).



**Figure S1 –** Error from 10-fold cross-validation error, as implemented in ADMIXTURE (Alexander et al. 2011), by the number of inferred population clusters (*K*); *K*=7 was chosen because the error reached a plateau at that value.



**Figure S2 –** Histograms of genomic relationships in the sample

(a) Marginal genomic relationships, scaled as in VanRaden (2008): off-diagonal elements of , with (: estimated allele frequency at marker *l*). (b) Genomic relationships conditional on population structure, as captured by PCs, also scaled as in VanRaden (2008): off-diagonal elements of , with ( consists of the first PCs of ).

**Figure S3 –** Histograms of node degree (number of neighbors by individual) based on the graphs of conditional relationships inferred by the graphical LASSO. Each graph represents the relationships as depicted from , where is the regularized matrix of conditional relationships as obtained by restricted maximum likelihood for each trait (the absence of edge between two individuals is indicated by a zero *ij*-element in ). The parameter λ represents the degree of regularization on relationships, fitted by maximum restricted likelihood for each trait (Appendix 1).

**Figure S4 –** Relationship between improvement in prediction accuracy and improvement in fit in *MPM-Matérn*

Validation scheme: (a) whole-sample calibration, where the target population is included in the calibration set; (b) cross-population calibration, where all individuals from the target population are omitted in the calibration set. Improvement in accuracy: difference in prediction compared to *GBLUP*. ΔBIC: difference in the Bayesian information criterion compared to *GBLUP* (the lower ΔBIC, the stronger the improvement in fit); where *L0* and *L1* are the restricted maximum likelihoods of *GBLUP* and *MPM-Matérn*, respectively, and characterizes the penalty on model complexity due to two additional parameters in *MPM-Matérn* compared to *GBLUP*, with the number of individuals in the CS.



**Figure S5 –** Significance of marker-by-panel interactions. Each plot represents the restricted likelihood, on a log-scale, as a function of the correlation between panels. The correlation between panels is the parameter γ such that ; and , where is the *n* x *n* matrix of ones and **Z** is the *n* x 2 design matrix attributing individuals to panels (BP or AP). *p*-values were obtained from a one-degree-of-freedom *χ2*-ditribution based on where *L0* and *L­1* are the restricted maximum likelihoods of *GBLUP* and the aforementioned model, respectively.



**Figure S6 –** Significance of difference in genomic variance by panel. Each plot represents the restricted likelihood, on a log-scale, as a function of the relative variance of AP, i.e., the ratio of genomic variance in AP over the sum of variances in AP and BP. This relative variance of AP is the parameter γ such that ; and , where **Z** is the *n* x 2 design matrix attributing individuals to AP and BP, respectively. *p*-values were obtained from a one-degree-of-freedom *χ2*-ditribution based on where *L0* and *L­1* are the restricted maximum likelihoods of *GBLUP* and the aforementioned model, respectively.