**Table S1 Effect of training set size (*N*P) on prediction ability (*r*MP) when performing cross-validation across the entire genomic selection dataset (EGSD method). Marker density (*N*M) fixed at all SNPs.**

|  |  |  |  |
| --- | --- | --- | --- |
| ***N*P** | **Yield *r*MP** | **Protein *r*MP** | **Oil *r*MP** |
| 50 | 0.055 | 0.627 | 0.540 |
| 100 | 0.097 | 0.697 | 0.609 |
| 150 | 0.137 | 0.749 | 0.639 |
| 200 | 0.177 | 0.764 | 0.664 |
| 250 | 0.189 | 0.796 | 0.684 |
| 300 | 0.242 | 0.799 | 0.699 |
| 350 | 0.236 | 0.806 | 0.710 |
| 400 | 0.255 | - | - |
| LSD (0.05)*a* | 0.038 | 0.019 | 0.023 |

*a*Fisher’s least significant difference threshold at α = 0.05

**Table S2 Effect of marker density (*N*M) on prediction ability (*r*MP) when performing cross-validation across the entire genomic selection dataset (EGSD method). Training set size (NP) fixed at 350 for protein and oil and 400 for yield.**

|  |  |  |  |
| --- | --- | --- | --- |
| ***N*M** | **Yield *r*MP** | **Protein *r*MP** | **Oil *r*MP** |
| All SNPs | 0.255 | 0.806 | 0.710 |
| Tag SNPs | 0.266 | 0.789 | 0.694 |
| Half tag SNPs | 0.243 | 0.784 | 0.692 |
| 4th tag SNPs | 0.283 | 0.751 | 0.685 |
| 8th tag SNPs | 0.298 | 0.743 | 0.644 |
| LSD (0.05)*a* | 0.035 | 0.016 | 0.019 |

*a*Fisher’s least significant difference threshold at α = 0.05

**Table S3 Effect of training set size (*N*P) on prediction ability (*r*MP) when performing cross-validation to predict individual bi-parental families (Pop1-4) using the within population (WP) method versus the across population (AP) method. Prediction ability was averaged across Pop1-4. Marker density (*N*M) fixed at all SNPs.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **GS method** | ***N*P** | **Yield *r*MP** | **Protein *r*MP** | **Oil *r*MP** |
| WP | 50 | 0.125 | 0.603 | 0.522 |
| AP | 50 | 0.041 | 0.336 | 0.209 |
| AP | 100 | 0.059 | 0.439 | 0.244 |
| AP | 150 | 0.082 | 0.472 | 0.259 |
| AP | 200 | 0.100 | 0.503 | 0.271 |
| AP | 250 | 0.083 | 0.530 | 0.270 |
| AP | 300 | 0.108 | 0.552 | 0.296 |
| AP | 350 | 0.119 | - | - |
| LSD (0.05)*a* | 0.031 | 0.028 | 0.032 |

*a*Fisher’s least significant difference threshold at α = 0.05

**Table S4 Effect of training set size (*N*P) on prediction ability (*r*MP) when performing cross-validation to predict individual bi-parental families (Pop1-4) using the within population (WP) method versus the across population (AP) method. Prediction ability was displayed for each individual validation population. Marker density (*N*M) fixed at all SNPs.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **GS method** | ***N*P** | **Yield *r*MP** |  | **Protein *r*MP** |  | **Oil *r*MP** |
| **Pop1** | **Pop2** | **Pop3** | **Pop4** |  | **Pop1** | **Pop2** | **Pop3** | **Pop4** |  | **Pop1** | **Pop2** | **Pop3** | **Pop4** |
| WP | 50 | 0.036 | 0.209 | 0.245 | 0.011 |  | 0.644 | 0.727 | 0.607 | 0.434 |  | 0.644 | 0.361 | 0.629 | 0.455 |
| AP | 50 | 0.020 | -0.009 | 0.065 | 0.086 |  | 0.281 | 0.447 | 0.435 | 0.180 |  | 0.116 | 0.148 | 0.366 | 0.207 |
| AP | 100 | 0.105 | 0.015 | 0.087 | 0.029 |  | 0.428 | 0.478 | 0.551 | 0.298 |  | 0.113 | 0.198 | 0.393 | 0.270 |
| AP | 150 | 0.081 | 0.045 | 0.068 | 0.135 |  | 0.433 | 0.496 | 0.584 | 0.374 |  | 0.105 | 0.215 | 0.441 | 0.274 |
| AP | 200 | 0.090 | 0.058 | 0.092 | 0.162 |  | 0.510 | 0.507 | 0.612 | 0.383 |  | 0.116 | 0.217 | 0.432 | 0.321 |
| AP | 250 | 0.117 | 0.004 | 0.094 | 0.118 |  | 0.553 | 0.545 | 0.627 | 0.397 |  | 0.060 | 0.255 | 0.480 | 0.285 |
| AP | 300 | 0.112 | 0.060 | 0.107 | 0.155 |  | 0.567 | 0.552 | 0.635 | 0.453 |  | 0.109 | 0.253 | 0.459 | 0.363 |
| AP | 350 | 0.117 | 0.101 | 0.080 | 0.180 |  | - | - | - | - |  | - | - | - | - |
| LSD (0.05)*a* | 0.058 | 0.063 | 0.062 | 0.059 |  | 0.051 | 0.053 | 0.040 | 0.054 |  | 0.059 | 0.055 | 0.048 | 0.056 |

*a*Fisher’s least significant difference threshold at α = 0.05