Table S1. Nonlinear regression estimates for modeling the accuracy of phenotyped lines:

\[ r^2_i = \frac{n_1(1-c)}{n_1 + b} + c, \]

where \( n_1 \) is the training population size.

<table>
<thead>
<tr>
<th>Population</th>
<th># Plots</th>
<th>( b )</th>
<th>( c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>1</td>
<td>747</td>
<td>0.12</td>
</tr>
<tr>
<td>Barley</td>
<td>2</td>
<td>624</td>
<td>0.23</td>
</tr>
<tr>
<td>Barley</td>
<td>3</td>
<td>681</td>
<td>0.32</td>
</tr>
<tr>
<td>Maize</td>
<td>1</td>
<td>742</td>
<td>0.12</td>
</tr>
<tr>
<td>Maize</td>
<td>2</td>
<td>528</td>
<td>0.25</td>
</tr>
<tr>
<td>Maize</td>
<td>3</td>
<td>502</td>
<td>0.40</td>
</tr>
</tbody>
</table>
Figure S1. $R_{\text{max}}$ as a function of $n_2$ (# unphenotyped entries) and $k_2$ (# selected from unphenotyped set) at a budget of 150 YPU, L+G cost of 0.5 YPU, two plots per entry, and 10 selected entries, using accuracies from the maize case study. The figure illustrates that $R_{\text{max}}$ is unimodal with respect to $k_2$. Although not visible in the figure, for some parameters $R_{\text{max}}$ is not unimodal with respect to $n_2$. 
Figure S2. The effect of resource allocation on barley prediction accuracy. See Fig. 3 caption for more details.
Figure S3. Optimal PYT designs for the maize case study, at a budget of 200 YPU. See Fig. 5 caption for more details.
Figure S4. Genetic gain with GS compared to PS for the maize case study, at a budget of 200 YPU and a line cost of 0.5 YPU. See Fig. 6 caption for more details.
Figure S5. $R_{\text{max}}$ vs. budget for the maize case study (# selected = 5, L cost = 0.5 YPU). The figure illustrates that the genetic gain with GS increases relative to the gain with phenotypic selection (PS) as the budget increases.
Selection index with genomic predictions

Consider a selection index \( I = \sum_i b_i z_i \) with predictors \( z_i \) and weights \( b_i \), and a measure of genetic merit \( H = \sum_j a_j g_j \) that is a weighted average of the genetic values \( (g_j) \) for several traits. The optimal index weights are chosen to maximize the correlation between the selection index and genetic merit:

\[
b^* = \operatorname{argmax} \rho(I, H) = \operatorname{argmax} \frac{\sigma\left(\sum_i b_i z_i, \sum_j a_j g_j\right)}{\sigma\left(\sum_i b_i z_i\right)} = \operatorname{argmax} \frac{b'Sa}{(b'Tb)^{1/2}} \tag{S1}
\]

where \( S_{ij} = \sigma(z_i, g_j) \) and \( T_{ij} = \sigma(z_i, z_j) \). By differentiating the final expression in Eq. S1 with respect to \( b \), one can derive the following first-order condition for optimality:

\[
(b'Tb)Sa = (b'Sa)Tb \tag{S2}
\]

and the reader can verify by substitution that the solution to Eq. S2 is

\[
b^* = T^{-1}Sa \tag{S3}
\]

When the predictors are phenotypes for the different traits, then \( T = P \) is the phenotypic covariance matrix, \( S = G \) is the genetic covariance matrix, and Eq. S3 is the well-known Smith-Hazel formula (Hazel, 1943; Walsh and Lynch, 2008).

Now consider an index for a single trait but using two predictors: \( z_1 = y \) is the phenotype and \( z_2 = \text{BLP}(g) = \hat{g} \) is the best linear predictor (BLP) calculated from \( y \) (and potentially the phenotypes of other individuals in the population). The genetic merit for this single-trait example is \( H = g \). From the statistical properties of best linear prediction (Searle et al., 1992) one can write \( \sigma(y, \hat{g}) = \sigma(y, g) = \sigma^2(g) \) and \( \sigma(g, \hat{g}) = \sigma^2(\hat{g}) \). This leads to the following expressions for \( T, T^{-1}, \) and \( S \):
\[
\begin{align*}
T & = \begin{bmatrix} \sigma^2(y) & \sigma^2(g) \\ \sigma^2(g) & \sigma^2(\hat{g}) \end{bmatrix}, \\
T^{-1} & = \Delta^{-1} \begin{bmatrix} \sigma^2(\hat{g}) & -\sigma^2(g) \\ -\sigma^2(g) & \sigma^2(y) \end{bmatrix}, \\
S & = \begin{bmatrix} \sigma^2(g) \\ \sigma^2(\hat{g}) \end{bmatrix}
\end{align*}
\]

where \( \Delta = \sigma^2(\hat{g})\sigma^2(y) - \sigma^4(g) \) is the determinant of \( T \). The final result \( b^* = T^{-1}S = [0 \ 1]' \) proves the optimal index weight for the phenotype is zero.

References


