Genomic Selection Using Maize ex-PVP Germlasm for the Prediction of N-use Traits
Adriano T. Mastrodomenico, Charles C. Hendrix, Patrick J. Brown, Alexander E. Lipka, and Frederick E. Below
Department of Crop Sciences, University of Illinois at Urbana-Champaign

Introduction:
Nitrogen use efficiency (NUE) in maize (Zea mays L.) is an important trait to help feed the world with minimal environmental impact.
Current breeding and biotechnological approaches have failed to develop a maize hybrid with optimal NUE.
Genomic selection for N-use traits may speed up the breeding cycle of research programs targeting improvements in maize NUE.

Objectives:
Identify genomic prediction accuracy for different N-use traits in maize using G-BLUP.
Evaluate the impact of training population sizes and training composition for effective application of genomic prediction of NUE traits in maize breeding programs.

Materials and Methods:
522 single-cross hybrids derived from 91 ex-PVP inbreds (33 SSS and 58 NSS) adapted to the U.S. Corn Belt (Figure 1) were tested in 10 environments during 2011 and 2016 (Figure 2).
On average, each SSS and NSS line was tested with 9 and 15 hybrids, respectively.

Figure 1. Principal component analysis (PCA) of 91 ex-PVP maize lines using 26,768 SNP’s. Colors represent the origin of the different lines.

Figure 2. Hybrids were evaluated at two N fertilizer rates (0 and 252 kg N ha⁻¹) in a RCB design: (a) number of environments planted across 3 locations in Illinois, (b) Infrared image detected from an UAV camera. Dark- and light-orange represent high and low N fertilizer treatments, respectively.

Genomic Selection Approach:
Data Analysis: Breeding value for each hybrid (GCA’s + SCA) was calculated for different phenotypic NUE traits (Figure 3) using REML.
G-BLUP: Untested hybrids were predicted using the genetic covariance matrix of untested and tested hybrids (C_T), the phenotypic covariance matrix of tested hybrids (C_YT), and the estimated breeding value. The genomic coefficient of each inbred was computed to the C_T and C_YT matrices.
Cross-Validation: Three cross-validation methods changing the training set (TS) composition were compared using 1,000 iterations:
1. T0: hybrids in which none of their parents were included in the TS or the validation set (VS).
2. T1: hybrids in which one of their parents were included in the TS and VS.
3. T2: hybrids in which both of their parents were included in the TS and VS.
In each cross validation method, a random subset of the T2 hybrids were used to predict either the T0, T1, or the remaining T2 hybrids.

Results:
Prediction accuracy increased 20% when the training composition changed from T0 to T1 and 17% when training composition changed from T1 to T2 hybrids when averaged across phenotypic traits and training sizes (Table 1). Lower H² under low N compared to high N was probably due to higher soil heterogeneity (e.g. NH₁₀⁻⁻⁰⁻⁻).

Yieldₐ₀⁻⁻ and HIₐ₀⁻⁻ exhibited the highest prediction accuracy under low N. Yieldₐ₀⁻⁻ and Proteinₐ₀⁻⁻ exhibited the highest accuracy under high N. NUE, Nₐ₀⁻⁻, and Nₐ₀⁻⁻ were positively correlated to Yieldₐ₀⁻⁻ (Table 2). GU and HI (low and high N) were positively correlated to Yieldₐ₀⁻⁻.

Table 1. Genomic prediction accuracy of T0, T1, and T2 hybrids (averaged across training set sizes). Training set was constituted from a constant number of SSS and NSS lines (16 and 30, respectively). Broad-sense heritabilities (H²) were estimated from ex-PVP maize hybrids grown across 10 environments at two levels of fertilizer N.

<table>
<thead>
<tr>
<th>Trait</th>
<th>T0</th>
<th>T1</th>
<th>T2</th>
<th>Range</th>
<th>H²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yieldₐ₀⁻⁻</td>
<td>0.36</td>
<td>0.56</td>
<td>0.75</td>
<td>0.31 / 0.96</td>
<td>0.30</td>
</tr>
<tr>
<td>Yieldₐ₀⁻⁻</td>
<td>0.28</td>
<td>0.53</td>
<td>0.73</td>
<td>0.22 / 0.93</td>
<td>0.60</td>
</tr>
<tr>
<td>Nₐ₀⁻⁻</td>
<td>0.11</td>
<td>0.31</td>
<td>0.43</td>
<td>0.07 / 0.66</td>
<td>0.59</td>
</tr>
<tr>
<td>Nₐ₀⁻⁻</td>
<td>0.16</td>
<td>0.34</td>
<td>0.45</td>
<td>0.12 / 0.67</td>
<td>0.27</td>
</tr>
<tr>
<td>HIₐ₀⁻⁻</td>
<td>-0.16</td>
<td>0.17</td>
<td>0.36</td>
<td>-0.17 / 0.45</td>
<td>0.11</td>
</tr>
<tr>
<td>GU</td>
<td>0.19</td>
<td>0.44</td>
<td>0.61</td>
<td>0.08 / 0.98</td>
<td>0.58</td>
</tr>
<tr>
<td>Proteinₐ₀⁻⁻</td>
<td>0.16</td>
<td>0.37</td>
<td>0.69</td>
<td>0.15 / 0.90</td>
<td>0.71</td>
</tr>
<tr>
<td>Proteinₐ₀⁻⁻</td>
<td>0.20</td>
<td>0.36</td>
<td>0.52</td>
<td>0.18 / 0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>HIₐ₀⁻⁻</td>
<td>0.44</td>
<td>0.61</td>
<td>0.73</td>
<td>0.32 / 0.98</td>
<td>0.63</td>
</tr>
<tr>
<td>HIₐ₀⁻⁻</td>
<td>0.24</td>
<td>0.36</td>
<td>0.45</td>
<td>0.19 / 0.65</td>
<td>0.71</td>
</tr>
<tr>
<td>NH₁₀⁻⁻</td>
<td>-0.30</td>
<td>-0.19</td>
<td>-0.18</td>
<td>-0.23 / -0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>NH₁₀⁻⁻</td>
<td>0.14</td>
<td>0.33</td>
<td>0.44</td>
<td>0.10 / 0.73</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 2. Phenotypic correlations between yield and N-use efficiency (NUE), N-uptake efficiency (Nₐ₀⁻⁻), genetic utilization efficiency (GU), genetic utilization (GU), and grain protein concentration, harvest index (HI), and nitrogen harvest index (NHI) at both low and high N.

<table>
<thead>
<tr>
<th>NUE</th>
<th>Nₐ₀⁻⁻</th>
<th>Nₐ₀⁻⁻</th>
<th>GU</th>
<th>Protein</th>
<th>HI</th>
<th>NHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yieldₐ₀⁻⁻</td>
<td>-0.38</td>
<td>NS</td>
<td>NS</td>
<td>0.67</td>
<td>-0.39</td>
<td>0.64</td>
</tr>
<tr>
<td>Yieldₐ₀⁻⁻</td>
<td>0.74</td>
<td>0.63</td>
<td>0.46</td>
<td>NS</td>
<td>-0.38</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

*Pearson coefficients indicate significance at P ≤ 0.001, NS indicate non-significant correlations.

Conclusions:
Changes in prediction accuracy as a result of training population and composition was dependent on the N use trait.
Both GU and HIₐ₀⁻⁻ exhibit high prediction accuracy and correlation to Yieldₐ₀⁻⁻. These traits may be integrated into maize breeding programs targeting improved performance under low N conditions (Figure 4).
Similarly, NUE and Nₐ₀⁻⁻ are desirable traits in breeding programs developing hybrids for agricultural systems using high N fertilizer inputs.

Figure 3. N-use traits evaluated for genomic selection: N-Use Efficiency (NUE), N-utilization efficiency (Nₐ₀⁻⁻), N-Uptake Efficiency (Nₐ₀⁻⁻), and Genetic Utilization (GU). Harvest Index (HI), N-harvest index (NHI), and grain protein concentration were measured at low and high N conditions. Plant biomass (BM) and N concentration were measured at physiological maturity.

Figure 4. Desirable selection traits for NUE maize breeding programs.

Future Research:
Integrating the genotype by environment effect into genomic selection models may improve the prediction accuracy of low heritability traits (e.g. Nₐ₀⁻⁻).

References: