Application of Genomic Selection in commercial corn breeding and crop improvement

Klaus Koehler  Dow AgroSciences

Abstract:
In a breeding program, variety testing presents a significant part of the Cost. This cost is expected to increase. At the same time the cost of genotyping is being reduced annually while the amount of information (data points per sample) is being dramatically increased.

As tools are being developed to predict the performance of a genotype and its correlation to its field performance is more and more improved, genomic prediction is becoming a more and more desirable component of a routine breeding program. It will have a significant change on plant breeding.

The fist generations of a breeding cycle typically are characterized by low heritability which results in low heritability of selection in early breeding stages. The question is how much precision is needed in genomic selections to become a viable alternative to actual testing.

The presentation will show results from a genomic prediction experiment in a DAS cooperation project with Purdue. While the overall accuracy of prediction of untested lines is still relatively low for yield, the correlation between two years of testing for yield is even lower in the study. Genomic prediction shows higher correlation with yield in 2012 than actual yield trials in 2011.

This indicates that genomic prediction has the potential to better predict the performance of a genotype in a future year than previous year yield trials.
APPLICATION OF GENOMIC SELECTION IN COMMERCIAL CORN BREEDING AND CROP IMPROVEMENT

Klaus Koehler  Dow AgroSciences
Dow AgroSciences

- Global Headquarters in Indianapolis, Indiana
- Annual Sales: >$ 7 billion in 2013
- Achieved milestone of $1 in seed sales in 2011
- Fastest growing seed company
- Employs 7000+ employees globally
- Serving markets in more than 150 nations
- Backing and leverage of a $60 B+ parent company

Solutions for the Growing World
Global Operations

Rapid worldwide expansion.
Facilities in more than 40 countries and products sold in more than 130 countries.
More than 7,000 employees worldwide.
Genomics and Corn breeding

• In corn breeding yield testing represents half of breeding budget: in search for yield, stability and related traits that reward the farmer

• Field testing will stay expensive (sophisticated equipment, land values, commodity pricing is up)

• Genotyping is becoming cheaper (currently less than 1/3 of yield test per candidate line)

• Genomic pre-selection will change plant breeding:
  > Increase response to selection (=better mean of population to select)
  > Change allocation of program for first years field testing
  > We must as much as possible integrate into our routine breeding cycle

• Can Genome-based selection provide better advancements?
The reality of breeding ....

... and how we can improve

- Understand the important genetics for inheritance of major selection traits
  - Yield and yield related (quantitative traits – many controlling loci – small effects)
  - Typical breeding programs have **low heritability** (repeatability) of the **most important phenotypic trait = yield**, especially in first year(s) of testing: many entries - few replications = lower accuracy of selection

Multi-stage Select for:
- Hybrid yield
- Inbred per-se

- GS= genomic selection:
  How much precision do we need?

- High selection intensities in most plant breeding stages
  \( \alpha = 0.05 \) to 0.1
Breeding cycle

MAB pre-select

Pre DH

Post DH

Create Variability
Recombine

DH Nursery

1st YEAR

2nd YEAR

3rd Y. and above

Recycling

? years
Germplasm improvement- Hybrid Development

\[ \frac{G}{Y} = \left( h^2 i \sigma \right) / Y \]

**Trait Y**

**Hybrid Level**

**Selected Fraction**

**Selection Response**

**Population Level**

**Time**

C1  C2  C3  C4

---

Dow AgroSciences

DOW RESTRICTED
Project with J. Morales and T. Rocheford; Purdue University unpublished
QTL for Yield in F_{2:3} Population of LH51xPHG35

Figure 2-3. LOD profile (top) and additive effects (below) for yield in DAS (red), LH119 (blue), and PHG39 (green) testcrosses. Chromosomes are separated by vertical black lines. LOD threshold equal to p=0.05 is represented by a horizontal blue line.

Red = DAS Tester
Blue = LH119
Green = PHG39

358 lines
407 polymorphic SNP markers from 1536 total
Figure 2-4. LOD profile (top) and additive effects (below) for moisture in DAS (red), LH119 (blue), and PHG39 (green) testcrosses. Chromosomes are separated by vertical black lines. LOD threshold equal to $p=0.05$ is represented by a horizontal blue line.
Figure 2-6. LOD profile (top) and additive effects (below) for traits with significant QTL on chromosome 1 in LH119 testcross. LOD threshold equal to p=0.05 is represented by a horizontal black line.
Training population tested in 7 environments using unreplicated design with replicated checks. Spatially corrected BLUPs were used to calculate predictions. Allelic effects of 284 markers were calculated.
Figure 4-1. Cross-validation of Genomic Selection model. 358 $F_{2:3}$ families testcrossed to LH119 were evaluated for yield in 2010 and 2011 in 7 environments and BLUPs were calculated across environments (observed yield). BLUPs were used to estimate alleleic effects which were then used to estimate a genome estimated breeding value (GEBV) for the training population.
GEBV and 2012 yield data on 180 untested F2:3 families were testcrossed to LH119. Arrows indicate families that were selected based on GEBVs and induced.
Comparison of 2010/2011 yield data and 2012 yield data in lines selected from the training population. 13 top, 13 bottom, and 76 random F2;3 families from the training population were grown along side the untested (predicted) families and evaluated for yield. The correlation between 2010/2011 performance and 2012 performance is very similar to the correlation between predicted and observed performance in untested families.
Comparison of GEBV and 2012 yield data in lines selected from the training population. 13 top, 13 bottom, and 76 random F_{2:3} families from the training population were grown along side the untested (predicted) families and evaluated for yield.
• Genomic performance prediction can be better than phenotypic performance prediction based on previous year performance

• It might be easier to predict the lower performance distribution of our breeding populations through GS than the high

• Genomic selection in early breeding stages can improve efficiency of breeding programs as it might have a better potential to predict future performance, i.e. estimate true(er) breeding value than the phenotype for a given future year
Thank You