Predictor traits from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat

Jessica Rutkoski

Co-authors: Jesse Poland, Suchismita Mondal, Enrique Autrique, Lorena González Pérez, José Crossa, Matthew Reynolds, and Ravi Singh
Wheat is the most extensively grown cereal crop
Breeding for grain yield and other key traits

- Many loci involved
- Gain occurs over time

Fischer and Edmeades, Crop Sci. 2009
Breeding pipeline

Wheat breeding pipeline

1. Crossing
2. Selected bulk via shuttle breeding
3. Advanced breeding lines
4. International nurseries
5. Advanced yield trial
6. Preliminary yield trial, optimal conditions
Shuttle breeding phase

Wheat breeding pipeline

Crossing → Selected bulk via shuttle breeding → Advanced breeding lines

- Selection for high-heritability traits: height, maturity, vigor, multiple disease resistance.
Line development

Wheat breeding pipeline

Crossing → Selected bulk via shuttle breeding → Advanced breeding lines

No longer bulk populations. Selection among lines (~30,000) for high-heritability traits.
Line development, details

1. Harvest individual F4 plants

2. Plant each F4:F5 in own small plot

3. Visual selection
Yield testing, year one

Wheat breeding pipeline

Selection among 7,000 lines for yield in one environment (Obregon)

International nurseries

Advanced yield trial

Preliminary yield trial, optimal conditions

Advanced breeding lines
Yield testing, year two

Wheat breeding pipeline

Selection among ~1,200 lines for yield under 6 management regimes (Obregon)

International nurseries

Advanced yield trial

Advanced breeding lines

Preliminary yield trial, optimal conditions
Dissemination and targeting of lines

Wheat breeding pipeline

~50 lines distributed for yield testing. Lines are targeted to different environment types.

International nurseries → Advanced yield trial → Preliminary yield trial, optimal conditions
Key factors affecting the rate of genetic gain

1. Accuracy of selection

2. Breeding cycle length

Vs.

Crossing
↓
Line development
↓
Selection of new parents
Exciting new technologies that could increase rates of genetic gain

Genomic selection (GS)-
Aim is to predict breeding values for traits of interest using genome-wide markers.

High-throughput phenotyping (HTP)-
Aim is to rapidly collect large a large volume of data on informative traits
Genomic selection, process

1. Training population is phenotyped & genotyped
2. Train prediction model
3. Predict genotyped selection candidates
Genomic prediction models ‘borrow’ information from relatives
Genomic relationships give better ‘resolution’ of relationships compared to pedigrees

Actual relationship between siblings

High-throughput phenotyping

- Uses multi-spectral, hyper-spectral, and thermal cameras mounted to drones and aircraft
- Reflectance data for hundreds of wavelengths
- Temperature
- Weekly data collection
High-throughput phenotyping, key traits measured

1. RNDVI
2. GNDVI
3. Canopy temperature-Water status

Biomass

1. Vegetative
2. Grain-filling

6 yield ‘predictor traits’
Goal is to predict yield prior to yield testing

Predict yield on ~30,000 lines. Improve selection for advancement and enable parent selection sooner.
Question: Can we use high-throughput phenotyping to help predict yield?
Integrate 6 predictor traits from high-throughput phenotypic data into genomic and pedigree prediction models and estimate prediction accuracy for yield
Materials and methods

• 557 lines after quality control pipeline

• 6 environments, 3 replicates each- HTP and yield data

• Genotyping-by-sequencing

• Prediction models:
  • Pedigree or genomic relationships
  • With or without HTP data
  • Corrected for days to heading and lodging
# Magnitude of genetic correlations, predictor traits & yield

<table>
<thead>
<tr>
<th></th>
<th>Across</th>
<th>Drought</th>
<th>Early Heat</th>
<th>Late Heat</th>
<th>Optimal</th>
<th>Severe Drought</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT_GF</td>
<td>0.52</td>
<td>0.49</td>
<td>0.44</td>
<td>0.73</td>
<td>0.44</td>
<td>0.42</td>
</tr>
<tr>
<td>CT_VEG</td>
<td>0.25</td>
<td>0.44</td>
<td>0.38</td>
<td>0.73</td>
<td>0.51</td>
<td>0.12</td>
</tr>
<tr>
<td>GNDVI_GF</td>
<td>0.32</td>
<td>0.19</td>
<td>0.34</td>
<td>0.8</td>
<td>0.36</td>
<td>0.2</td>
</tr>
<tr>
<td>GNDVI_VEG</td>
<td>0.08</td>
<td>0.06</td>
<td>0.11</td>
<td>0.07</td>
<td>0.49</td>
<td>0.48</td>
</tr>
<tr>
<td>RNDVI_GF</td>
<td>0.28</td>
<td>0.02</td>
<td>0.2</td>
<td>0.74</td>
<td>0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>RNDVI_VEG</td>
<td>0.37</td>
<td>0.28</td>
<td>0.38</td>
<td>0.76</td>
<td>0.41</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Magnitude of genetic correlations, predictor traits & yield

<table>
<thead>
<tr>
<th></th>
<th>Across</th>
<th>Drought</th>
<th>Early Heat</th>
<th>Late Heat</th>
<th>Optimal</th>
<th>Severe Drought</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT_GF</td>
<td>0.52</td>
<td>0.49</td>
<td>0.44</td>
<td>0.73</td>
<td>0.44</td>
<td>0.42</td>
</tr>
<tr>
<td>CT_VEG</td>
<td>0.25</td>
<td>0.44</td>
<td>0.38</td>
<td>0.73</td>
<td>0.51</td>
<td>0.12</td>
</tr>
<tr>
<td>GNDVI_GF</td>
<td>0.32</td>
<td>0.19</td>
<td>0.34</td>
<td>0.8</td>
<td>0.36</td>
<td>0.2</td>
</tr>
<tr>
<td>GNDVI_VEG</td>
<td>0.08</td>
<td>0.06</td>
<td>0.11</td>
<td>0.07</td>
<td>0.49</td>
<td>0.48</td>
</tr>
<tr>
<td>RNDVI_GF</td>
<td>0.28</td>
<td>0.02</td>
<td>0.2</td>
<td>0.74</td>
<td>0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>RNDVI_VEG</td>
<td>0.37</td>
<td>0.28</td>
<td>0.38</td>
<td>0.76</td>
<td>0.41</td>
<td>0.04</td>
</tr>
</tbody>
</table>
# Heritabilities

<table>
<thead>
<tr>
<th></th>
<th>Across</th>
<th>Optimal</th>
<th>Drought</th>
<th>Severe Drought</th>
<th>Late Heat</th>
<th>Early Heat</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-GF</td>
<td>0.71</td>
<td>0.34</td>
<td>0.67</td>
<td>0.61</td>
<td>0.41</td>
<td>0.37</td>
</tr>
<tr>
<td>CT-VEG</td>
<td>0.86</td>
<td>0.66</td>
<td>0.65</td>
<td>0.4</td>
<td>0.38</td>
<td>0.46</td>
</tr>
<tr>
<td>GNDVI-GF</td>
<td>0.65</td>
<td>0.68</td>
<td>0.7</td>
<td>0.63</td>
<td>0.54</td>
<td>0.49</td>
</tr>
<tr>
<td>GNDVI-VEG</td>
<td>0.61</td>
<td>0.64</td>
<td>0.58</td>
<td>0.55</td>
<td>0.31</td>
<td>0.64</td>
</tr>
<tr>
<td>GY</td>
<td>0.56</td>
<td>0.39</td>
<td>0.45</td>
<td>0.51</td>
<td>0.67</td>
<td>0.4</td>
</tr>
<tr>
<td>RNDVI-GF</td>
<td>0.62</td>
<td>0.66</td>
<td>0.73</td>
<td>0.55</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>RNDVI-VEG</td>
<td>0.54</td>
<td>0.47</td>
<td>0.64</td>
<td>0.48</td>
<td>0.59</td>
<td>0.59</td>
</tr>
</tbody>
</table>
## Heritabilities

<table>
<thead>
<tr>
<th></th>
<th>Across</th>
<th>Optimal</th>
<th>Drought</th>
<th>Severe Drought</th>
<th>Late Heat</th>
<th>Early Heat</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-GF</td>
<td>0.71</td>
<td>0.34</td>
<td>0.67</td>
<td>0.61</td>
<td>0.41</td>
<td>0.37</td>
</tr>
<tr>
<td>CT-VEG</td>
<td>0.86</td>
<td>0.66</td>
<td>0.65</td>
<td>0.4</td>
<td>0.38</td>
<td>0.46</td>
</tr>
<tr>
<td>GNDVI-GF</td>
<td>0.65</td>
<td>0.68</td>
<td>0.7</td>
<td>0.63</td>
<td>0.54</td>
<td>0.49</td>
</tr>
<tr>
<td>GNDVI-VEG</td>
<td>0.61</td>
<td>0.64</td>
<td>0.58</td>
<td>0.55</td>
<td>0.31</td>
<td>0.64</td>
</tr>
<tr>
<td>GY</td>
<td>0.56</td>
<td>0.39</td>
<td>0.45</td>
<td>0.51</td>
<td>0.67</td>
<td>0.4</td>
</tr>
<tr>
<td>RNDVI-GF</td>
<td>0.62</td>
<td>0.66</td>
<td>0.73</td>
<td>0.55</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>RNDVI-VEG</td>
<td>0.54</td>
<td>0.47</td>
<td>0.64</td>
<td>0.48</td>
<td>0.59</td>
<td>0.59</td>
</tr>
</tbody>
</table>
## Predictor trait and yield data used in prediction models

### Univariate prediction

<table>
<thead>
<tr>
<th></th>
<th>CT-GF</th>
<th>CT-VEG</th>
<th>GNDVI-GF</th>
<th>GNDVI-VEG</th>
<th>RNDVI-GF</th>
<th>RNDVI-VEG</th>
<th>GY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

**Legend:**
- Phenotypic data present
- Phenotypic data absent

### Multivariate prediction

<table>
<thead>
<tr>
<th></th>
<th>CT-GF</th>
<th>CT-VEG</th>
<th>GNDVI-GF</th>
<th>GNDVI-VEG</th>
<th>RNDVI-GF</th>
<th>RNDVI-VEG</th>
<th>GY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

**Legend:**
- Phenotypic data present
- Phenotypic data absent
HTP data improves prediction accuracies within environment
HTP data improves prediction accuracies within environment

![Graph showing prediction accuracies across different environments.](image)

- Low predictor trait correlations & moderate heritabilities
- Low predictor trait heritabilities & moderate correlations
HTP data improves prediction accuracies across environment
HTP + pedigree may be better than ordinary genomic selection
Next step, validate in small plots and improve accuracy
Acknowledgements

Field Crew
Thank you for your interest!

Questions?