Application of Advanced Technologies For Pulse Improvement: PulseBio
Basic Breeding Program

Create genetic diversity

1 year

Advance generations

F₂ - F₆

2 - 5 years

Selection of better adapted fixed lines

Disease testing

3 years

Yield trials

Commercialisation

3 years

Breeding cycle

c. 9 - 12 year
Trait Enrichment in Field Peas

- **F5 derived families**: Screen in all 4 assays to select for homozygotes.
- **Screen in field to select for better agronomic types**.
- **Enter best lines into yield trials**.

- **F2 population**: Boron tolerance (1 locus).
- **F5 population**: Bacterial Blight (3 loci).
- **F4 population**: Downy Mildew (2 loci).
- **F3 population**: Salinity tolerance (2 loci).
## Molecular Marker Implementation

<table>
<thead>
<tr>
<th>Method</th>
<th>% of F₅ Homozygous lines</th>
<th>No. lines for phenotypic testing</th>
<th>Early generation selection</th>
<th>F₅ phenotypic selection</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenotypic selection</td>
<td>16.8%</td>
<td>297</td>
<td>$2,543</td>
<td>$27,420</td>
<td>$29,964</td>
</tr>
<tr>
<td>Marker assisted selection</td>
<td>36.7%</td>
<td>50</td>
<td>$5,228</td>
<td>$5,978</td>
<td>$11,206</td>
</tr>
</tbody>
</table>

- Ability to screen for all selection traits early in the program
- More effective than phenotypic breeding
- Delivers same outcome for lower costs
- Limits on marker trait associations in use of parents
- Highly specific for simple traits
- Will not work on more complex traits
Genomic Selection - Background

- Traits of interest for breeding:
  - 1 gene
  - Specific disease resistance
  - Boron tolerance
  - Few genes
  - Salinity
  - Disease resistance for a variety of strains and pathogens
  - More genes
    - all genes
  - Yield
    - Grain weight
  - Drought tolerance
  - Flowering time
  - Grain quality

Prediction equation

Genomic Breeding Value = \( w_1X_1 + w_2X_2 + w_3X_3 + \ldots \)
Genomic Selection

Prediction equation

Genomic Breeding Value =

\[ w_1 x_1 + w_2 x_2 + \cdots + W_n x_n \]

• Agriculture Victoria invented science – 2001

• Only viable method for selecting traits where variation is contributed by large number of loci, e.g. yield

• Still works on all other traits, simple > complex

• Large benefit for traits that are difficult/expensive to measure, measured late in breeding cycle

• Accelerate genetic gain by reducing generation interval and increasing accuracy of selection

• Consensus that genomic selection is simplest and most robust method to use genomics for breeding
Genomic Selection

Figure 1: Genetic Trend of Young Holstein Bulls Released in Canada for Lifetime Profit Index (GLPI)

- Average Gain: 445 LPI points/year
- Average Gain: 210 LPI points/year
- Average Gain: 105 LPI points/year

Period of Semen Release in Canada
(Group 1=Jan-Jun, Group 2=Jul-Dec within Year)
Genomic Selection

Crop Species
- Modelled and implemented by Agriculture Victoria in:
  - Pasture grasses (ryegrass, phalaris)
  - Wheat
  - Canola
  - Potato
  - Lentil
  - Others: Maize, Rice, Soybean, Forestry, Lucerne
Genomic Selection - Overview

Reference Population
- Genotypes
- Phenotypes

Genotypes

Phenotypes

Prediction Equation
Genomic Breeding Value = \( w_1x_1 + w_2x_2 + w_3x_3 + \ldots \)

Selection Candidates
- Genotypes

Selected Parents
- Estimated breeding values
Crossing and seed multiplication
(5 years)

Multi-parent crossing and application of GS to identify elite parents
(1.5 years)

Generation advance and field trials to update GS prediction equation
(4 years)

Seed to breeding partner(s) for further evaluation

Selection for potential entry into next crossing cycle

Evaluation and generation advance of breeding material
(3 years)

Selection for potential entry into next crossing cycle

National variety trials
(1-3 years)
Genomic Selection in Lentil - Development & Implementation

- Reference transcriptome assembled (Australian cultivar - Cassab)
- 40 key historic and current parental lines genotyped
- A total of c. > 200,000 high-quality SNPs (from 30,000 contigs) identified from parental dataset - spread evenly across the genome
Genomic Selection – Historical Breeding Data

- 5 years of multi environment plot trials accessed
- ~ 200 varieties per year
- Grain yield (Under low-moderate genetic control)
- Grain weight (Under high genetic control)
- Phenotypic data analysed using advanced statistical models (4x extended factor analysis with spatial effects fitted and variety as a random effect)
- Seed samples retained from all sown varieties
- Genotyped with GBS transcriptome method
- 26,000 high quality segregating SNP loci used for genomic prediction
Pedigree Relationships Across Years

- Forward prediction will provide the best indication of GS accuracy
- Using previous years stage 2 trials as the training population, the performance of subsequent year will be predicted and compared to the observed phenotypes
- High level of relatedness (pedigree) across years
- Gradual transition in main parent usage across years

![Lentil Parent Usage Across Years](image-url)
Genomic Prediction - Grain Yield

- Within year clusters of trials were connected across years with seasonal rainfall data.

Rainfall:
- VIC, SA, VIC, SA, VIC, VIC + Kingsford (SA)
Genomic Prediction - Grain Yield

Moderate prediction accuracy observed across years

- Complex trait with large environmental interaction
- Development of 3 reference populations for different environments
- Additional environmental factors to be explored
- 2015/2016 genotype and phenotype data will help expand reference population size
- Ability to predict performance of new varieties across all environments
Forward Prediction Accuracy for Grain Yield

- Able to predict with moderate accuracy across multiple environments lentil performance.
- Accuracy will improve with additional years and additional environmental profiles.
- Although rainfall was similar the seasonal pattern was opposite.
  - 2011 was a dry winter.
  - 2012 was a dry spring.
Correlation between Genomic Prediction and Phenotypic Data

A

Grain Yield - 2014 Low Rainfall

Genomically Predicted Yield (t/ha)

Phenotypically Observed Yield (t/ha)

B

Grain Weight - 2014 Very Low Rainfall

Predicted Grain Weight

Observed Grain Weight
Genomic Prediction – Effects

- Yield t/ha
- Grain Weight
- Boron Tolerance
Optimum Haploid Value

Selection on Optimal Haploid Value Increases Genetic Gain and Preserves More Genetic Diversity Relative to Genomic Selection

Hans D. D’Cruz,1,3 Matthew J. Hayden,2* German C. Spangenberg,* and Ben J. Hayes1,2*

1AgRlife, Center for AgBioResource, Department of Economic Development, Jobs, Transport and Resources (DEIJTR), Bundaberg, Victoria 2003, Australia, and 2School of Applied Systems Biology, La Trobe University, Bundaberg, Victoria 2006, Australia

ORCID: iD: 0000-0002-3819-3949 (H.D.D.)

ABSTRACT: Doubled haploids are routinely created and phenotypically selected in plant breeding programs to accelerate the breeding cycle. Genomic selection, which makes use of both phenotypes and genotypes, has been shown to further improve genetic gain through prediction of performance before or without phenotypic characterization of novel germplasm. Additional opportunities exist to combine genomic prediction methods with the creation of doubled haploids. Here we propose an extension to genomic selection, optimal haploid value (OHV) selection, which predicts the best doubled haploid that can be produced from a segregating plant. This method focuses selection on the haplotype and optimizes the breeding program toward its end goal of generating an elite fixed line. We rigorously tested OHV selection breeding programs, using computer simulation, and show that it results in up to 5x standard deviations more genetic gain than genomic selection. At the same time, OHV selection preserved a substantially greater amount of genetic diversity in the population than genomic selection, which is important to achieve long-term genetic gain in breeding populations.

KEYWORDS: genomic selection; haploids; genetic gain; genetic diversity; doubled haploid; shared data resources; GerHard
Optimum Haploid Value

Optimum Haplotype Value

\[ \text{Haplotype Value} = \sum_{j=1}^{l} x_j \beta_j \]

\[ \text{OHV} = 2 \sum_{o=1}^{n} \max(HV_o) \]

\[ \text{GEBV} = +2.5 \]
Multi-parent Crossing & Family Advancement

Parents and multi-parent crossing design selected based on predicted DHV of derived F₁ individuals and subsequent F₁ lines.

Selection of F₂ individuals based on simulated family segregation and resulting genetic potential.

Rapid generation advancement of F₂ individuals to Fₙ.

Selection of elite derived lines based on GEBV.

Cycle back as potential parents.

Infield phenomic evaluation.

Supply to commercial breeding sector for generation advancement and selection.

Cycle back as potential parents.

Infield phenomic evaluation.

Supply to commercial breeding sector for evaluation and release.
Crossing and seed multiplication (5 years)

Evaluation and generation advance of breeding material (3 years)

National variety trials (1-3 years)

Multi-parent crossing and application of GS to identify elite parents (1.5 years)

Generation advance and field trials to update GS prediction equation (4 years)

Selection for potential entry into next crossing cycle

Seed to breeding partner(s) for further evaluation
• State-of-the-art glasshouse for high-throughput phenotyping of crop plants
• Conveyor system to transport plants to the imaging stations
• Computer controlled plant tracking, imaging and image analysis
• Enables measurement of traits previously not practical
• Unparalleled tracking of plant response to treatments

Plant Phenomics Victoria, Horsham
Plant Growth
Plant Phenomics – In Field

NDVI data can be extracted to a 1 cm resolution and geo-referenced for automated data processing.
Grain Phenomics – In Laboratory

Roundness Index

\[ RI = 1 - \frac{1}{m} \sum_{i=1}^{m} \left( 1 - r_i \right) \]

Circularity

\[ C = \frac{4A}{PD_{equi}} \]

Sphericity

\[ S = RI \times p \]
PulseBio – Vision of Success in 2021

**BETTER SEEDS**

- **2X**
  - RATE OF GENETIC GAIN FOR YIELD
  - Through breeding cycle reduced to half

- **+15%**
  - VALUE OF CROP
  - Through new high value quality and end-use traits

- **+20%**
  - STABLY EXPANDED CROP
  - Through foundation, yield stability and adaptation traits

**SAFER SEEDS**

- **DE-RISK**
  - Through biosecurity proofing of seeds and biologicals

- **+15%**
  - YIELD INCREASE
  - Through tailored biologicals

Economic Development Jobs, Transport and Resources
Conclusions

• Modern genomics applied to lentil breeding
• Rapid improvements possible to all traits
• Methods applicable to all pulses
• Multi environment prediction of performance possible
• Novel crossing designs to maximise genetic gain
• Automated plant phenomics under development
• Expanded traits to be integrated into breeding/prebreeding
• Delivering better options for farming systems faster
Acknowledgments
Pulse Bio 1 - Outline

- **P1 X P2**
  - F1
  - F1i
    - Genotype - GEBVs
  - F2-Pop
    - F3-Pop multiplication
      - F4-Pop Multiplication & Field Trial
        - Updated Equation
          - Transfer to Pulse Breeding
          - Genotype for hand-over
PulseBio 1 - Workflow

- **2016**
  - P1 X P2 → F1
  - P3 X P4 → F1
  - F1 → F2-Pop

- **2017**
  - P1 X P2 → F1
  - P3 X P4 → F1
  - F1 → F2-Pop
  - Genotype - GEBVs

- **2018**
  - P1 X P2 → F1
  - P3 X P4 → F1
  - F1 → F2-Pop

  - Transfer to Pulse Breeding

- **2019**
  - P1 X P2 → F1
  - P3 X P4 → F1
  - F1 → F2-Pop
  - Genotype - GEBVs

  - In field multiplication

- **2020**
  - P1 X P2 → F1
  - P3 X P4 → F1
  - F1 → F2-Pop

  - In field multiplication

  - Updated Equation
  - Transfer to Pulse Breeding
  - Genotype for hand-over

  - F4-Pop Multiplication & Field Trial

  - Updated Equation
  - Transfer to Pulse Breeding
  - Genotype for hand-over

  - F3-Pop multiplication

  - F4-Pop Multiplication & Field Trial
PulseBio Projects

• PulseBio 1 – Genomics assisted pre-breeding in pulse crops (2 x rate of genetic gain for yield)
• PulseBio 2 – Pulse quality (15% increase in value of the crop)
• PulseBio 3 – Yield stability in pulse crops (20% expansion in area for pulse cropping)
• PulseBio 4 – De-risking (Biosecurity proofing of seeds and biologicals)
• PulseBio 5 – Tailored biologicals (15% yield increase)
Genomic Prediction – Simple Traits

Accuracy of 0.92 or 0.98 when considered as a binary trait

*\( k \)-fold validation was performed for this single trial