Intermediate Wheatgrass Breeding & Genomics at the University of Minnesota



Prabin Bajgain James A. Anderson

Kernza Conference July 1, 2019







Outline

- Overview of IWG breeding at UMN
- Kernza[®] variety release
- GWAS for yield traits
- Genomic selection with dominance, GxE, & haplotype blocks
- QTL mapping for *Fusarium* and bacterial leaf streak resistance

Timeline of Intermediate Wheatgrass Domestication



Intermediate Wheatgrass Breeding Goals

- Grain Yield
- Sustained Grain Yield
- Seed Size
- Shatter Resistance
- Free Threshing
- Plant Height
- Lodging Resistance
- Diseases (FHB, Ergot)
- End-use Quality





Genomic Selection

Training population





Genome wide markers

 $Y = \beta X$

Breeding population





 $\beta X = Y$

IWG Breeding at UMN with Genomic Selection



Breeding Progress: Trait Improvements



Cycle 2 (2015-2016) Cycle 3 (2017-2018)

10

0







Kernza Variety Release

Our first Kernza variety, MN1504, is approved for release in 2019!

- Produces an average of 500 lbs. of grain per acre (563 Kg ha⁻¹)
- Currently being increased on 100 acres (40 Ha) in MN well-head protection areas
- Variety release mechanism being decided



MN1504 Release

MN1504 is a good agronomic performer:

- Short plant height: 113 cm (3.7 ft)
- Minimal lodging, uniform maturity
- Moderately high seed threshability



Entry	Grain Yield	ткw	Seed Area	Height	Lodging	Threshability
	lbs/Ac	g	cm ²	cm	0-4	0-4
MN1505	600	5.70	6.60	114	2	2.0
MN1502	518	5.60	6.20	125	1	1.5
MN1504	502	6.00	6.50	113	0	1.5
MN1503	456	5.70	6.70	120	3	2.0
MN1501	423	8.50	8.30	125	3	3.5
TLI C5	417	6.50	6.70	127	2	1.5
TLI C4	395	6.40	6.90	118	3	2.0
TLI C3	369	6.00	6.50	118	3	3.0
Rush	251	5.30	5.90	118	4	3.5
No. of Envs	5	3	3	3	3	3
LSD (0.10)	155.30	0.40	0.31	5.40	0.70	0.42

Future UMN Kernza Variety Candidates



- Four MN16XX candidates planted in 2018 for small-plot seed increase in St. Paul; large-scale increase and state-wide trials in 2019
- Other candidates discontinued due to poor agronomic performance such as lodging, low yield, tall stands, ergot.

Yield Traits: QTL Mapping & Genomic Selection

Materials: Cycle 3 genets (560) of which ~450 used in GWAS Evaluated in 3 environments: 2017 and 2018 St. Paul, and 2018 Crookston

Traits:

- Grain yield
- Thousand Kernel Weight
- Seed length
- Seed width
- Spike weight
- Spike length
- No. of spikelets per spike

Marvin Seed Scanner



GWAS:

Genotyping by sequencing

Allele-calling using the latest v2.1 IWG reference genome Population structure determined using STRUCTURE Q-matrix used as covariate in GWAS



QTL for yield and seed traits





We identified **154 genomic regions** associated with yield and yield component traits in our cycle 3 IWG breeding population.

Applicability of QTL in Genomic Selection



Genetic mapping for FHB & BLS resistance

Materials & Methods:

A77-3: resistant to FHB (Fusarium head blight) & bacterial leaf streak (BLS) C20-7: susceptible

- Parents were selected based on field evaluations during 2011-2013
- Crossed to obtain 108 F_1 genets
- Syringe-inoculated using
 - FHB isolates collected from 2014 IWG plants
 - BLS culture grown from wheat-infecting pathogen
- Per plant: 5 spikes for FHB & 3 flag leaves for BLS
- Data recorded 21 dpi
- Joinmap for linkage map construction
- MapQTL for CIM



FHB & BLS phenotypes

Disease phenotypes observed in the IWG F_1 mapping population A77-3 x C20-7:

A: Fusarium head blight (FHB)B: Bacterial leaf streak (BLS)in 2016, 2017, and 2018.





QTL for FHB and BLS resistance



15 QTL were associated with FHB resistance and 11 QTL with BLS.5 QTL were common between the two diseases.

Bajgain et al., in review

Disease reduction in 2-3 QTL models & Genomic Selection



Genomic Selection with Dominance effects

GBLUP (Genomic best linear unbiased prediction) vs. Bayesian models:

- Model with additive effects only (A variant)
- Model with additive and dominance effects (AD variant)



Genomic Selection with GxE interaction effects

- Models accounting for GxE effects (GBLUP only)
 - i. Single Kernel "MM" for main random genetic effect across all environments
 - ii. Multi-Kernels
 - a. "MDs" for main genetic effect in all environments and a single G×E effect for all environments
 - b. "MDe" for main genetic effect in all environments as well as a single $G \times E$ effect for each environment

Spike Length: H = 0.73Seed Width: H = 0.44TKW: H = 0.69Yield: H = 0.68Spike Length: H = 0.58Spike Weight: H = 0.43No. of Spikelets: H = 0.31Shatter Resistance: H = 0.75Threshability: H = 0.56

Genomic Selection with haplotype blocks

GBLUP (Genomic best linear unbiased prediction) vs. Bayesian models:

- Models with haplotype blocks
 - Blocks based on recombination breakpoints
 - Sliding window of 2, 5, 10 SNPs

Fixed-length haplotypes can improve genomic prediction accuracy in an admixed dairy cattle population

Melanie Hess 🖾 , Tom Druet , Andrew Hess and Dorian Garrick

 Genetics Selection Evolution
 2017
 49:54

 https://doi.org/10.1186/s12711-017-0329-y
 © The Author(s) 2017

 Received:
 18 October 2016
 Accepted: 26 June 2017
 Published: 3 July 2017

Assessing haplotype-based models for genomic evaluation in Holstein cattle

Z. Karimi,^a M. Sargolzaei,^{ab} J.A.B. Robinson,^a F.S. Schenkel^a

^aCentre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, ON N1G 2W1, Canada.

^bSemex Alliance, Guelph, ON N1H 6J2, Canada.

Corresponding author: Zahra Karimi (email: skarimi@uoquelph.ca).

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from <u>RightsLink</u>.

Published on the web 11 May 2018. Received January 23, 2018. Accepted May 2, 2018.

Haplotype-based genotyping-by-sequencing in oat genome research

Wubishet A. Bekele, Charlene P. Wight, Shiaoman Chao, Catherine J. Howarth, Nicholas A. Tinker 🔀

First published: 18 January 2018 | https://doi.org/10.1111/pbi.12888 | Cited by: 7

Performance of Single Nucleotide Polymorphisms versus Haplotypes for Genome-Wide Association Analysis in Barley

Aaron J. Lorenz 💿, Martha T. Hamblin 💿, Jean-Luc Jannink 🖂

Published: November 22, 2010 • https://doi.org/10.1371/journal.pone.0014079

Bajgain et al., in preparation

Additive vs. Additive + Dominance models



- No significant difference between additive only vs. additive + dominance models
 - In fact, additive only models were better in many cases
- GBLUP models were slightly better than Bayesian models, in general
- No single method or model gave the highest predictive ability for any particular trait
- No relationship between trait heritability and predictions

Better predictions when GxE effects considered



- Models incorporating GxE effects outperformed nearly all other models
- With GxE effects included in the models, predictive abilities improved by nearly two-folds for yield, spike length, spike weight, and threshability.

Haplotype Blocks increase trait prediction



- Overall, haplotype block length of 2 SNPs (sliding window method) was better for most traits
- Blocks constructed by considering recombination breakpoints (*Haploview*) were worse
- Compared to single SNP markers, the increase in predictively ability (~1-2% in UMN_C3) may not be worth the time it takes to convert genotypic data into haplotype blocks

Acknowledgments

University of Minnesota

Dept. of Agronomy and Plant Genetics James Anderson Kayla Altendorf

Brett Heim Jacob Jungers Donald Wyse Craig Sheaffer Ronald Faber

Dept. of Plant Pathology

Ruth Dill-Macky Rebecca Curland Carol Ishimaru

Dept. of Food Science & Nutrition

Pam Ismail	George Annor
Catrin Tvl	Tonva Schoenfuss

Green Lands Blue Waters, UMN Aaron Reser

The Land Institute

Lee DeHaan

Kansas State University

Jesse Poland Kevin Dorn

USDA-ARS, Logan, UT

Steve Larson

MN Crop Improvement Association

Funding partners:

Forever Green Initiative General Mills MN Department of Agriculture





Trait Improvements 2011-2018



Trait Distributions & Heritabilities



Trait

Boxplots of phenotypic data collected on UMN_C3 in St Paul in 2017 and 2018, and in Crookston, MN in 2018

IWG Linkage Disequilibrium & Recombination



Pairwise distance (cM)

Pairwise LD (r^2) and recombination frequencies in the IWG F₁ mapping population A77-3 x C20-7. Average genome-wide LD half-decay occurs within 7.5 cM; wheat is ~ 7 cM