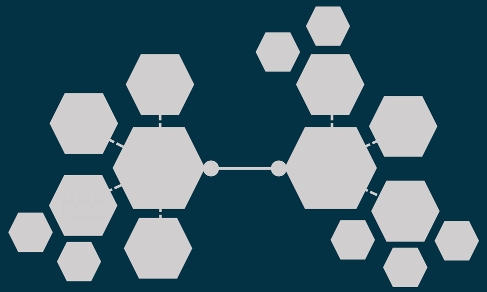


Chasing Genes and Making an Impact Where it Matters

Shantel A. Martinez, PhD

USDA-ARS Lincoln, NE

2019.11.20

ACKNOWLEDGMENTS

Cornell University

USDA ARS Wheat Health Unit
Steber, Garland Campbell Groups
Carter, Pumphrey Groups
Uauy Lab

Cornell Small Grains Group
Jannink & Sorrells Labs

PROJECT FUNDING

Education and Workforce Development (EWD)

Proposal #2017-07109

USDA

United States
Department of
Agriculture

National Institute
of Food and
Agriculture

USDA ars

WASHINGTON
GRAIN
COMMISSION

Eastern Washington Raised

Eastern Washington Raised

Dr. Camille M. Steber

BS Bioegineering

MS Crop Sci

PhD Mol Plant Sci

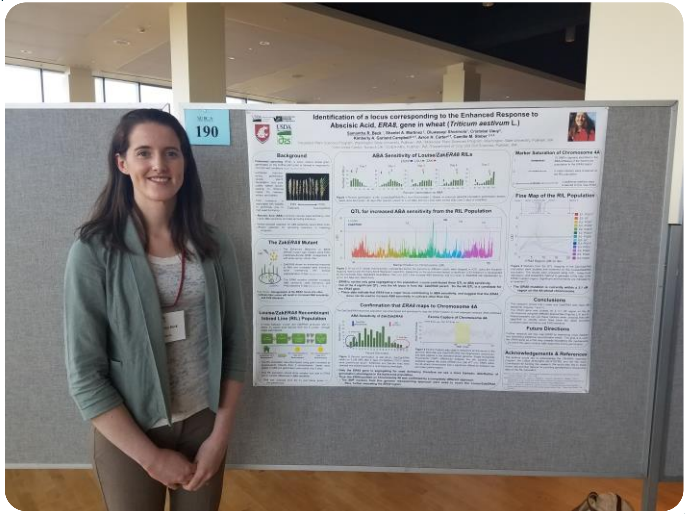
Dr. Kimberly A.
Garland Campbell

USDA NIFA EWD Fellowship

Dr. Mark E. Sorrells

CORNELL UNIVERSITY FOUNDED A.D. 1865

SERVICE & OUTREACH

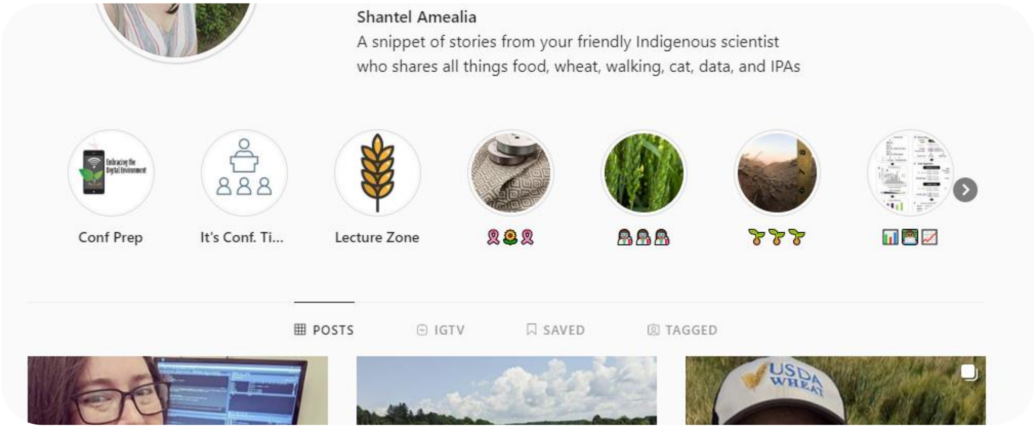


Mentoring women
in science (3 years)

LSAMP
Louis Stokes Alliance for Minority Participation

Volunteer, mentor (5 years)

Public Engagement
@s_amealia

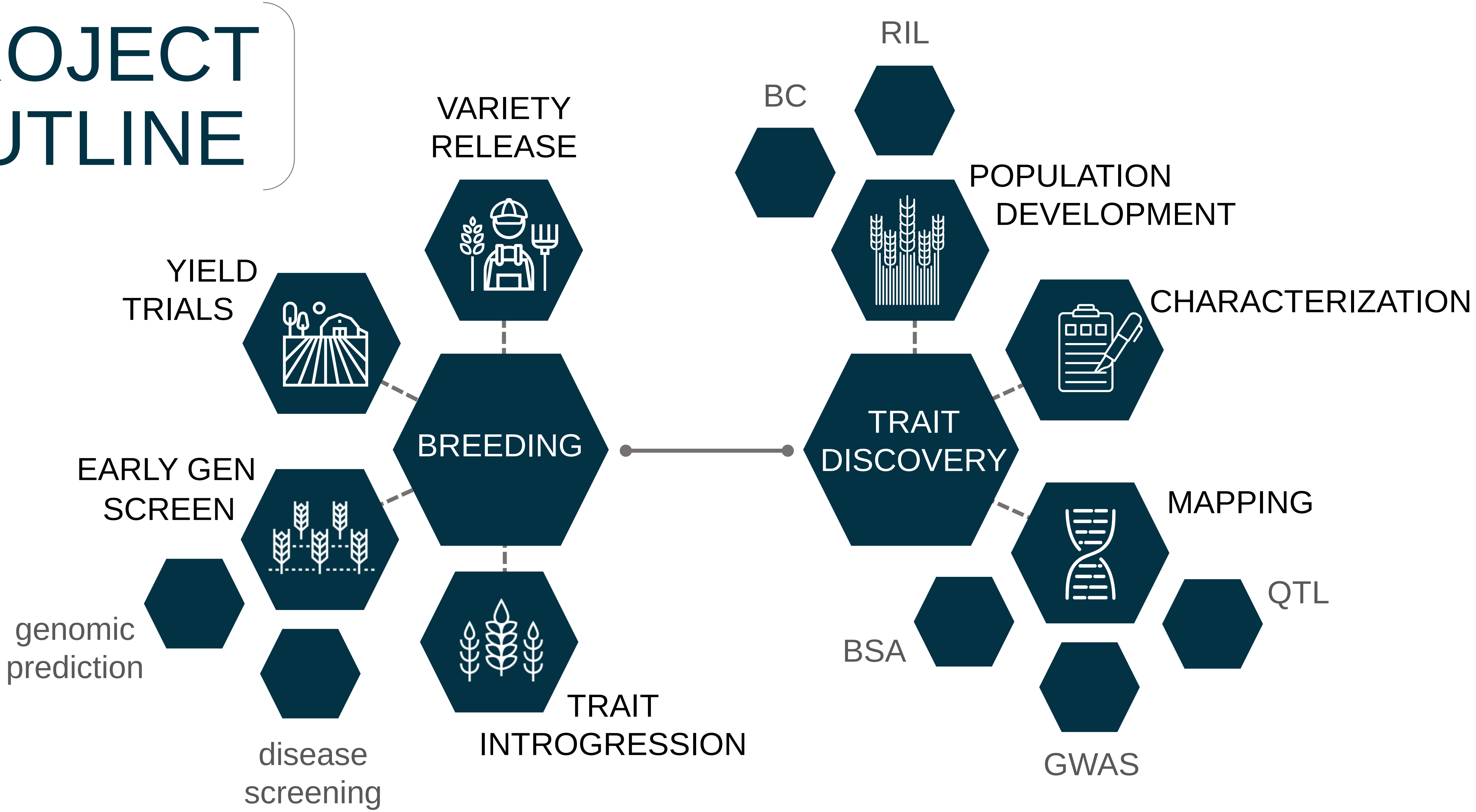


Shantel Amealia
A snippet of stories from your friendly Indigenous scientist
who shares all things food, wheat, walking, cat, data, and IPAs

UPWARD BOUND
PROGRAM
TRiO

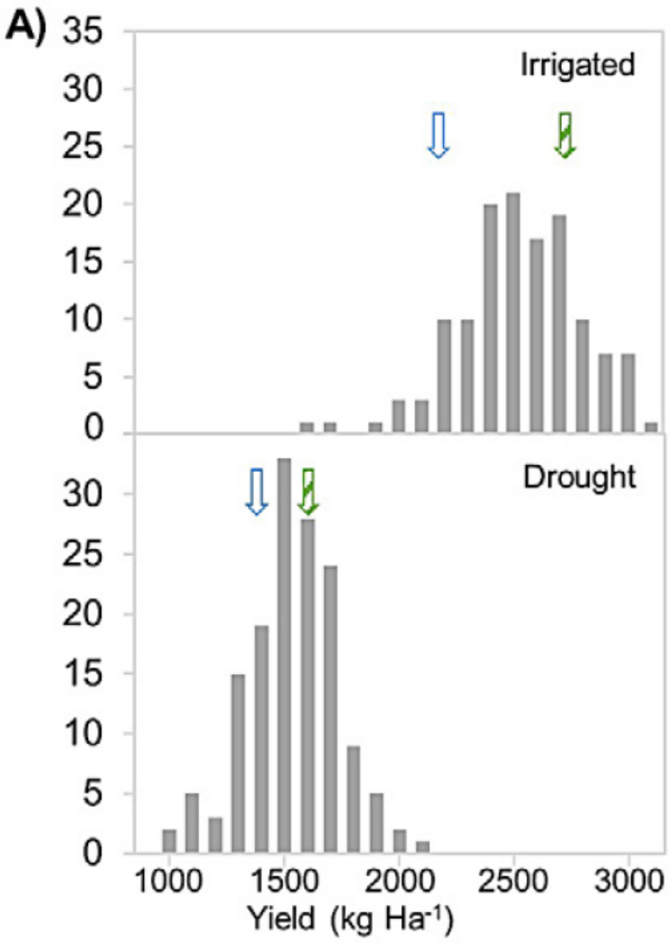
Recruiter, volunteer,
mentor (6 years)

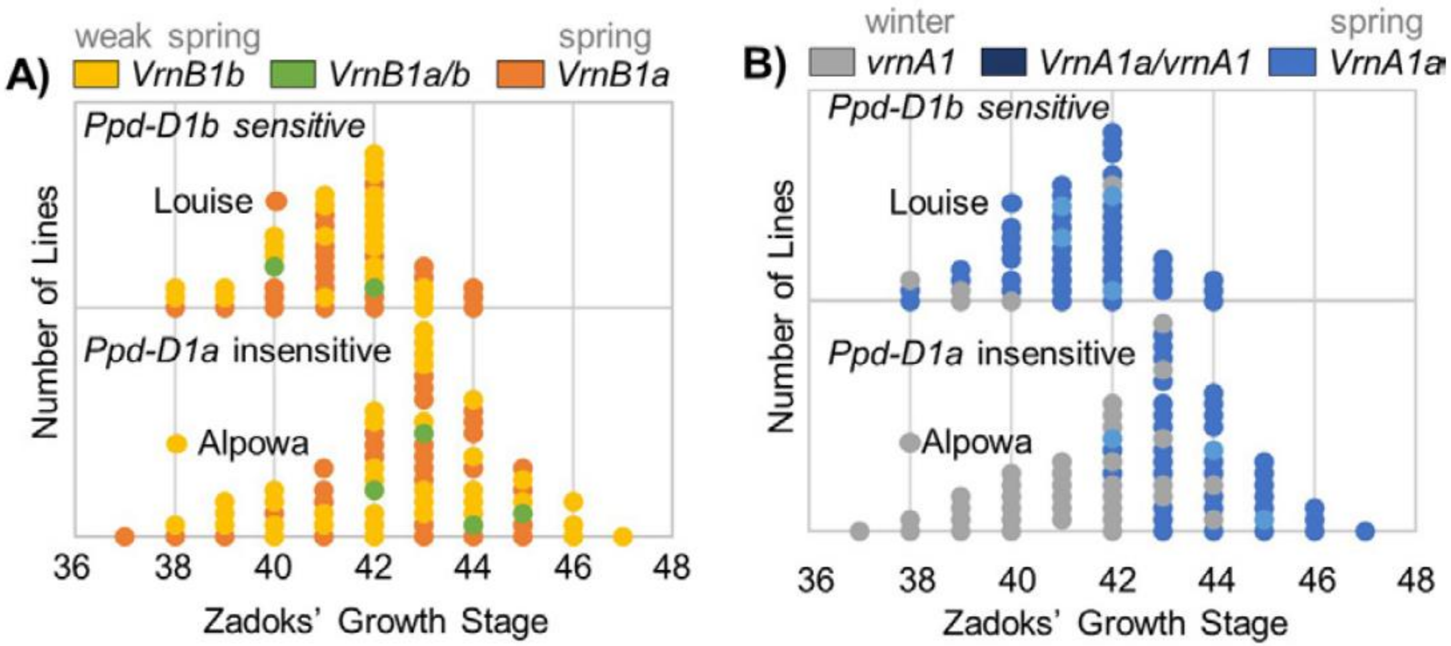
PROJECT OUTLINE



Mapping Population Release

Developed RIL population Louise x Alpowa for potential drought mapping studies





Martinez et al., 2018a

JOURNAL OF PLANT REGISTRATIONS

MAPPING POPULATION

Registration of the Louise/Alpowa Wheat Recombinant Inbred Line Mapping Population

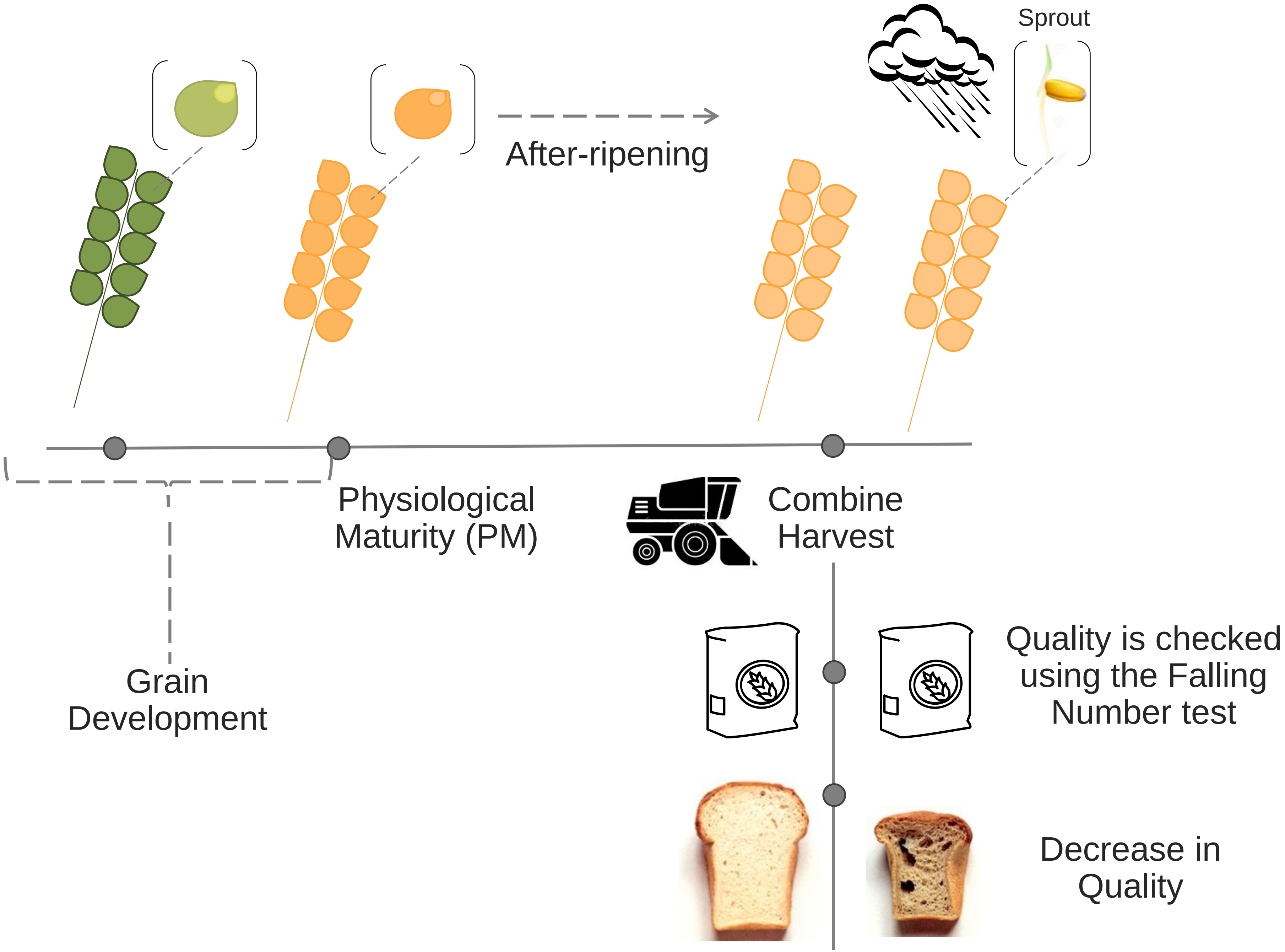
Shantel A. Martinez, Alison L. Thompson, Nuan Wen, Lesley Murphy, Karen A. Sanguinet, Camille M. Steber, and Kimberly A. Garland Campbell*

Abstract

A mapping population was developed from the cross of soft

THE INLAND northwest of the United States is a major wheat (*Triticum aestivum* L.)–producing region com-

Preharvest Sprouting



Sprout

After-ripening

Grain Development

Physiological Maturity (PM)

Combine Harvest

Quality is checked using the Falling Number test

Decrease in Quality

Wheat has quite a bit of genetic variation

Unfortunately, rain events do occur during harvest

England

Tobias Barber
@ekte_Toby

I guess this is what you'd call pre-harvest sprouting #Harvest18

Jim Thompson
@jimt_farmer

Think rain has stopped play #wheatharvest18 @AllpressF @LumleySean @coostiebarrey @chrisbettinson2

Kansas

Kyler Millershaski
@Shaski92

I'm always happy to have rain, but not the view I want during #WheatHarvest18 #kswx

3:36 PM - 22 Jun 2018

5 Retweets 37 Likes

Nebraska

Chris Cu11an
@ChrisCullan1

On the edge. #wheatharvest18

5:06 PM - 12 Jul 2018

1 Retweet 40 Likes

Canada

Fermes Chauvin Farms Ltd.🇨🇦
@MoeChauvin

#wheatharvest18 started here at fermeschauvinfarms.com in StoneyPoint. Decent yields for no rain ... #OntAg #AgMoreThanEver #goodineverygrain

Pacific Northwest

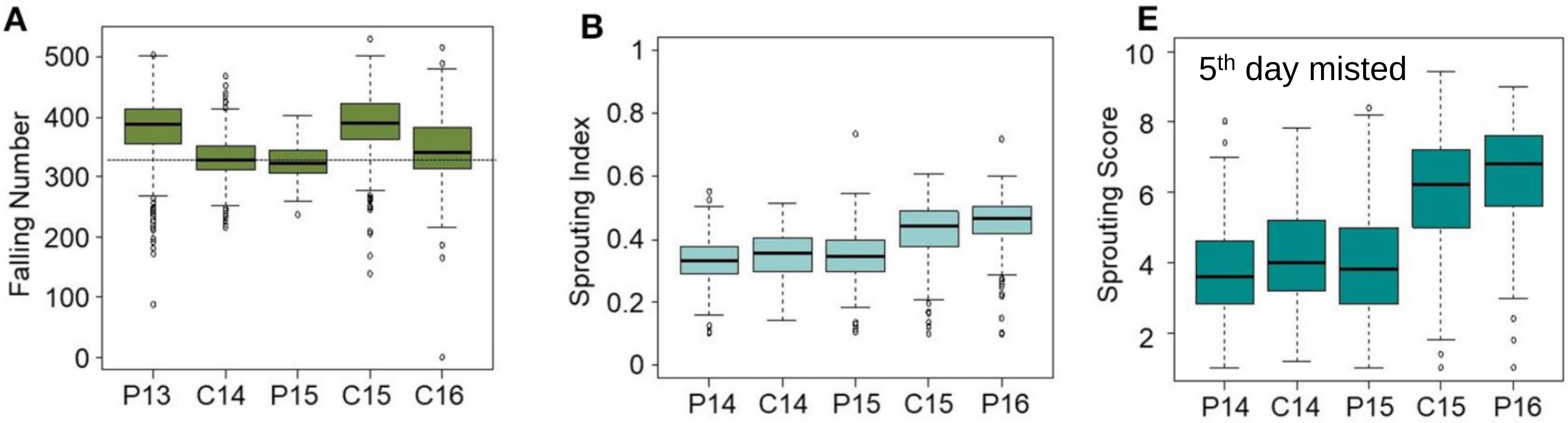
Twitter

Farms.com

“…in 2016 when losses [*due to low FN*] were staggering, in the tens of millions of dollars.” - Alex McGregor

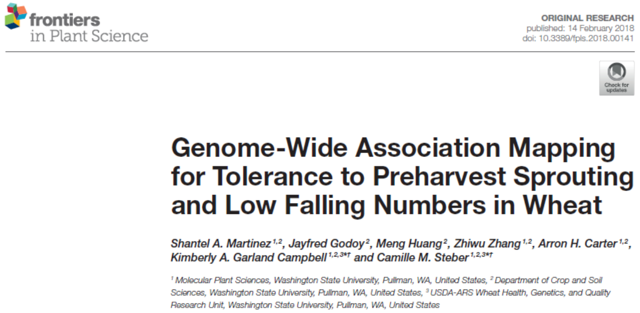
“…September 2016, as the dismal FN results came in, **growers, scientists, agencies, state legislators and Rep**. Cathy McMorris Rogers (R-Wash.) met …. make sure the FN challenge didn’t fall off the radar screen down the road as other issues came to the fore. It hasn’t.”

GWAS: PNW germplasm



5th day misted

Martinez et al., 2018b



frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 February 2018
doi: 10.3389/fpls.2018.00141

Check for updates

Genome-Wide Association Mapping for Tolerance to Preharvest Sprouting and Low Falling Numbers in Wheat

Shantel A. Martinez [1,2], Jayfred Godoy [2], Meng Huang [2], Zhiwu Zhang [1,2], Arron H. Carter [1,2], Kimberly A. Garland Campbell [1,2,3†] and Camille M. Steber [1,2,3*†]*

[1] Molecular Plant Sciences, Washington State University, Pullman, WA, United States, [2] Department of Crop and Soil Sciences, Washington State University, Pullman, WA, United States, [3] USDA-ARS Wheat Health, Genetics, and Quality Research Unit, Washington State University, Pullman, WA, United States

6 MTA were found in multi-env

19 / 34 *QPHS.wsu* loci and
4 / 9 *QFN.wsu* loci
co-localized with (54) previously published studies

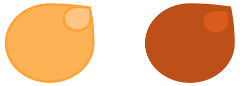
Martinez et al., 2018b

How does this study help?

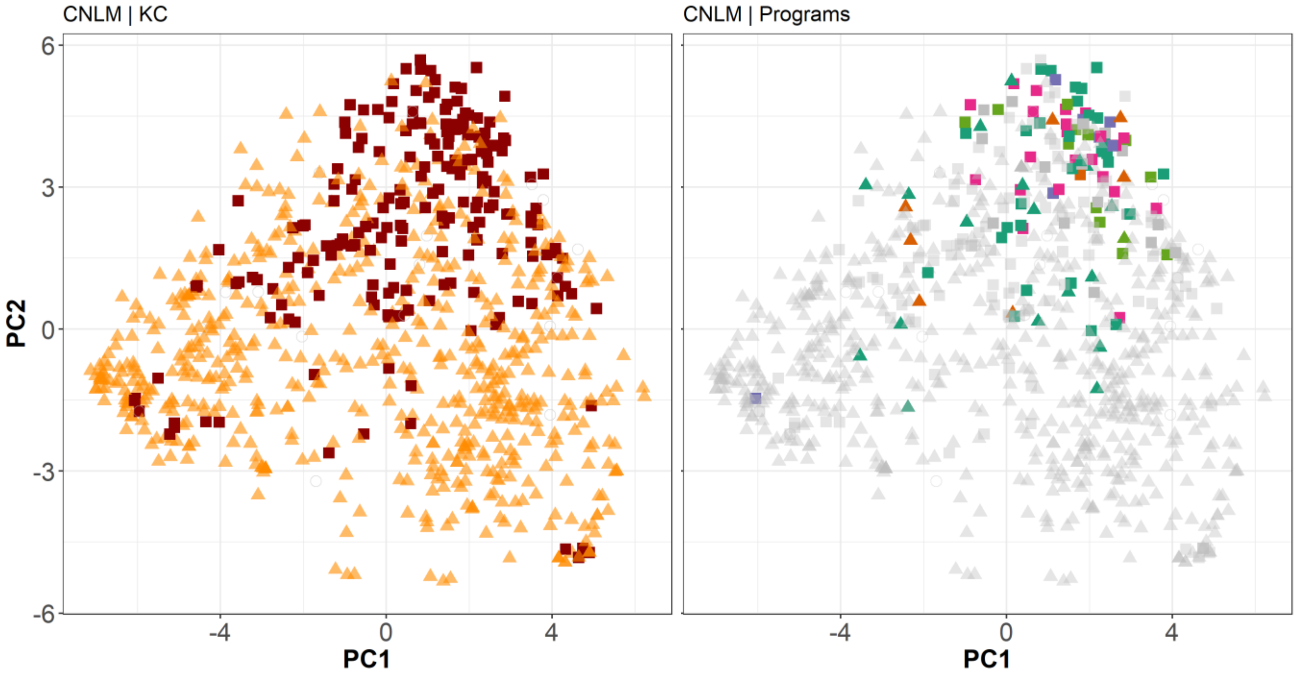
6 MTA made into KASP
markers for breeder selection

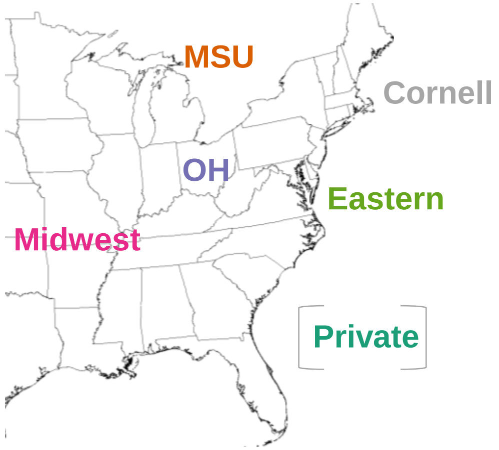
Identified germplasm that were
high in FN quality and/or PHS tolerant
versus germplasm that was too low in either

GWAS: Northeast germplasm



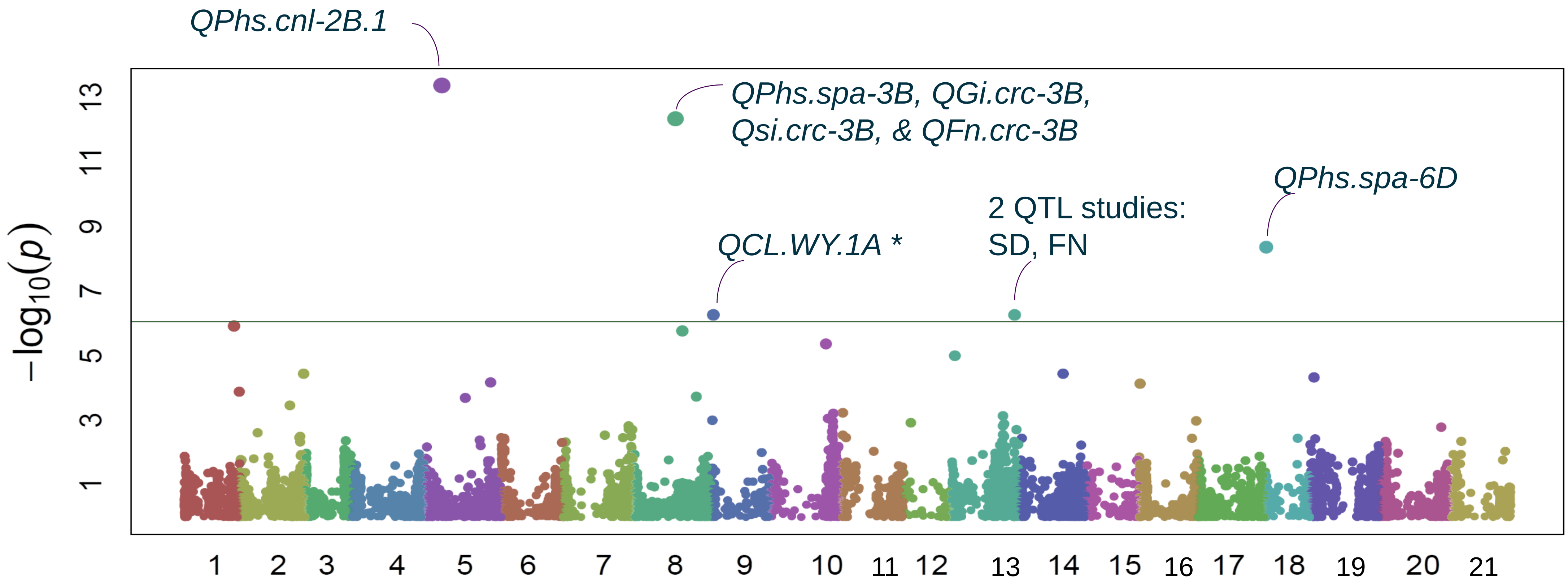
White Red





Martinez et al., unpublished

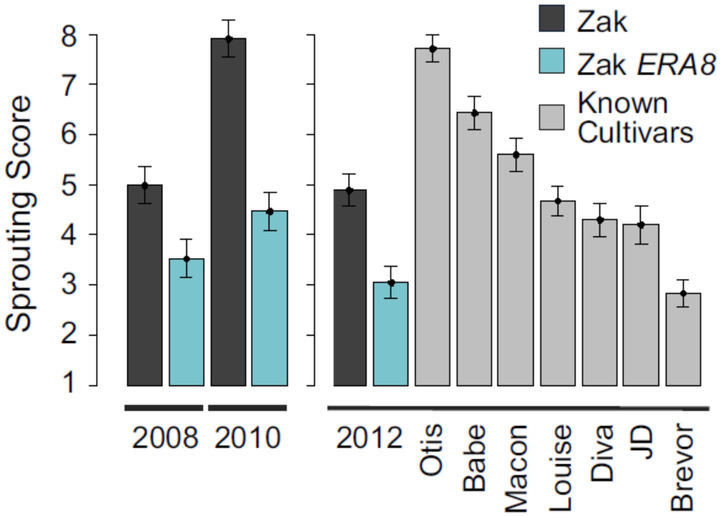
GWAS: Northeast germplasm

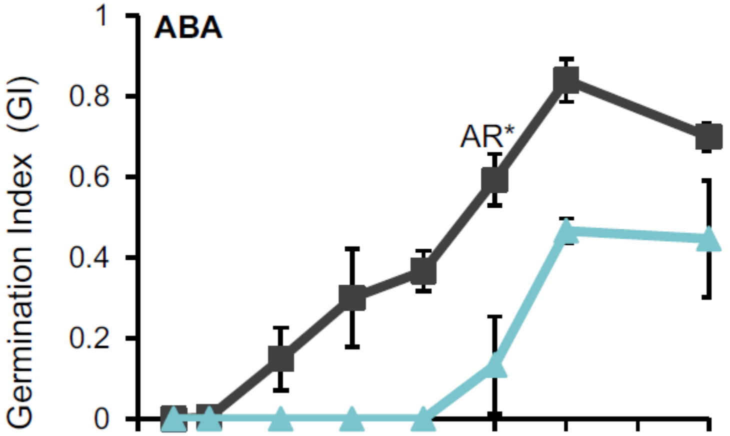


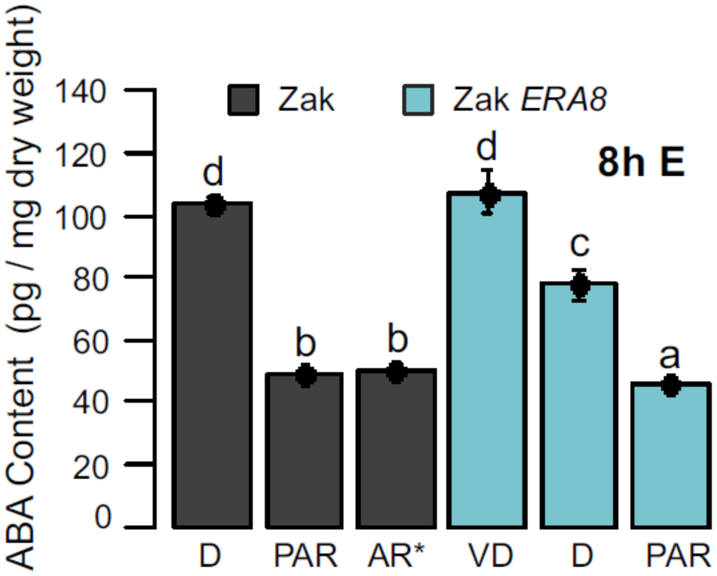
FarmCPU

Zanetti et al., 2000; Munkvold et al. 2009; Fofana et al. 2009; Zhang et al. 2013; Kumar et al. 2015; IWGSC, 2018; Martinez et al. 2018; Zuo et al., 2019

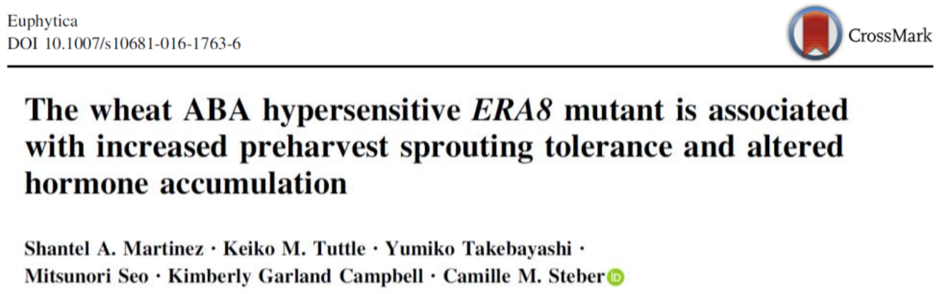
ERA8 Mutant Characterization







Martinez et al., 2016

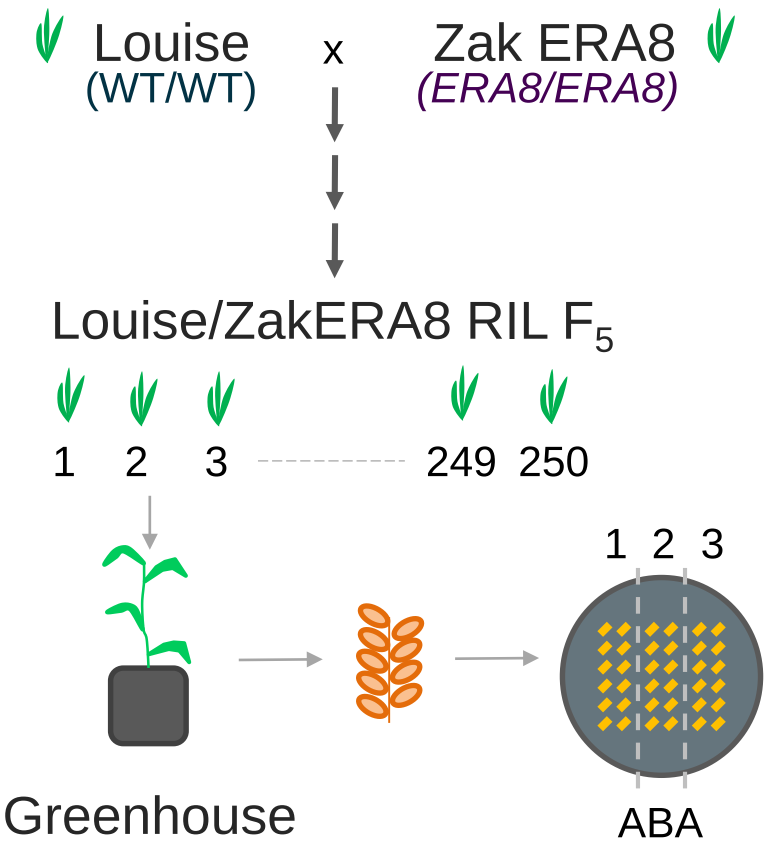

Euphytica
DOI 10.1007/s10681-016-1763-6

CrossMark

The wheat ABA hypersensitive *ERA8* mutant is associated with increased preharvest sprouting tolerance and altered hormone accumulation

Shantel A. Martinez · Keiko M. Tuttle · Yumiko Takebayashi · Mitsunori Seo · Kimberly Garland Campbell · Camille M. Steber

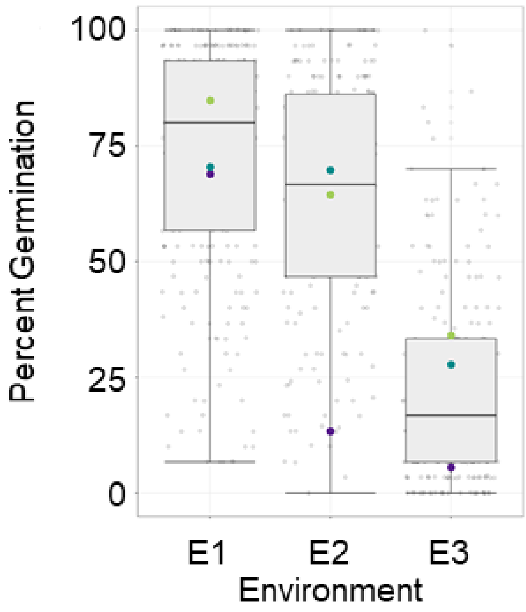
RIL Population Development


Louise (WT/WT) x Zak ERA8 *(ERA8/ERA8)*

Louise/ZakERA8 RIL F_5

1 2 3 ------------ 249 250

Greenhouse



Genetic cultivar differences in addition to the *ERA8* gene.

Martinez et al., 2020 (*accepted TAG*)


New Results

Comment on this paper

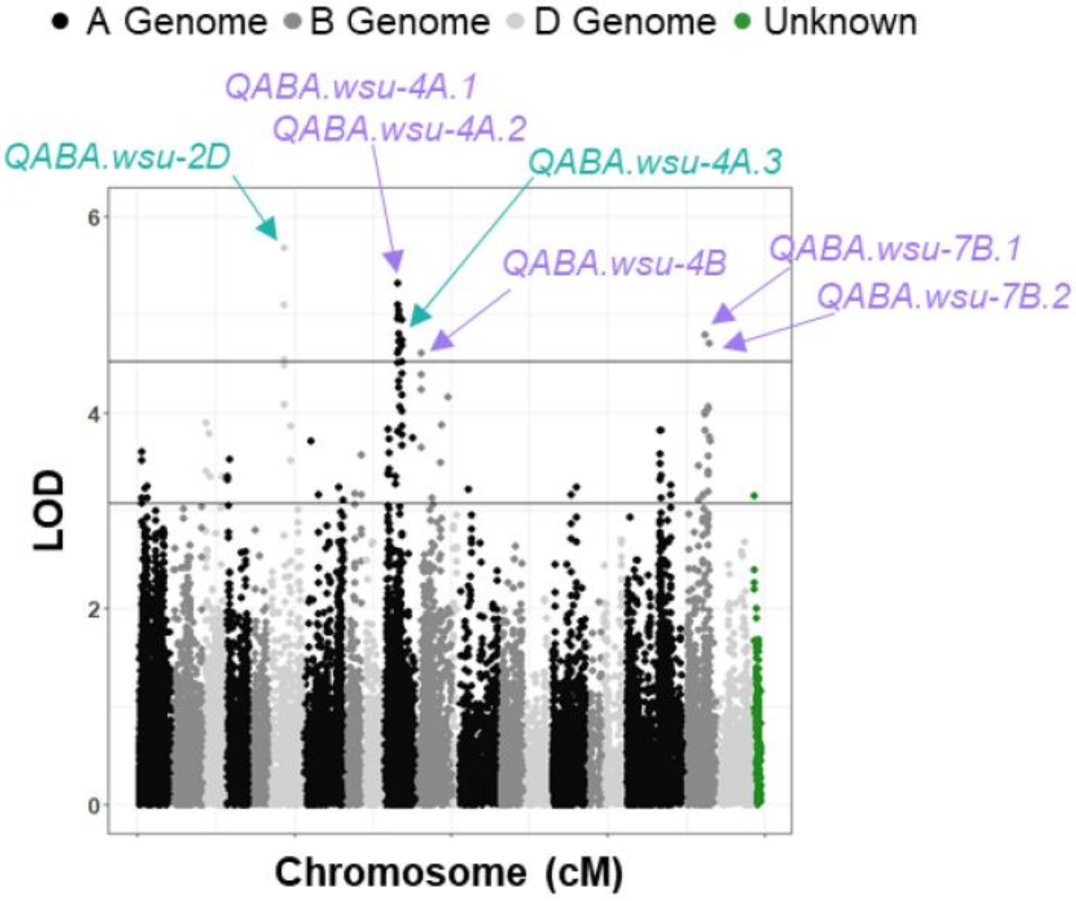
Exome sequencing of bulked segregants identified a novel *TaMKK3-A* allele linked to the wheat *ERA8* ABA-hypersensitive germination phenotype

Shantel A. Martinez, Oluwayesi Shorinola, Samantha Conselman, Deven See, Daniel Z. Skinner, Cristobal Uauy, Camille M. Steber

doi: https://doi.org/10.1101/784652

This article is a preprint and has not been certified by peer review [what does this mean?].

QTL Mapping



Multiple ABA sensitive QTL (increased dormancy) were found throughout the genome, contributed by either Louise or *ERA8*

Martinez et al., 2020 (*accepted TAG*)

Agronomic Characterization

Although technically ‘not significant’, still concerned about a slight dip in yield

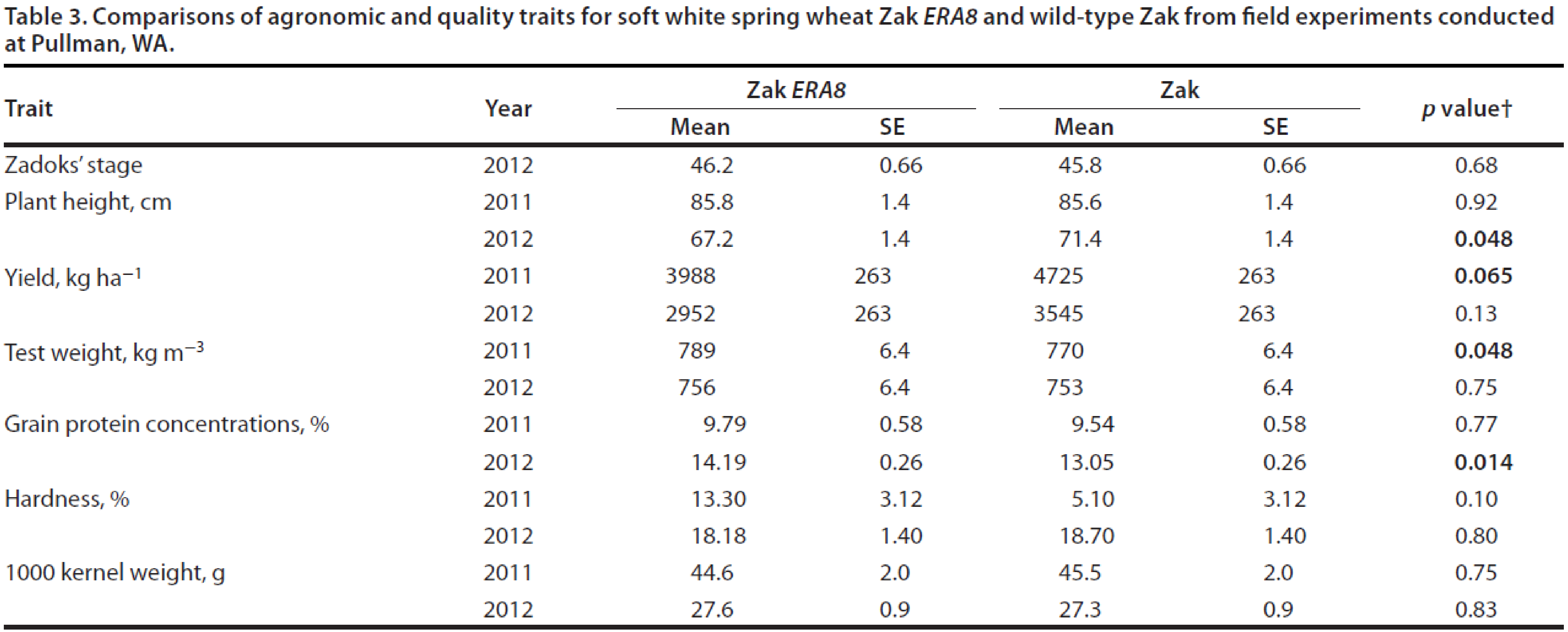


Table 3. Comparisons of agronomic and quality traits for soft white spring wheat Zak *ERA8* and wild-type Zak from field experiments conducted at Pullman, WA.

Trait	Year	Zak *ERA8*		Zak		*p* value†
		Mean	SE	Mean	SE	
Zadoks' stage	2012	46.2	0.66	45.8	0.66	0.68
Plant height, cm	2011	85.8	1.4	85.6	1.4	0.92
	2012	67.2	1.4	71.4	1.4	**0.048**
Yield, kg ha^{-1}	2011	3988	263	4725	263	**0.065**
	2012	2952	263	3545	263	0.13
Test weight, kg m^{-3}	2011	789	6.4	770	6.4	**0.048**
	2012	756	6.4	753	6.4	0.75
Grain protein concentrations, %	2011	9.79	0.58	9.54	0.58	0.77
	2012	14.19	0.26	13.05	0.26	**0.014**
Hardness, %	2011	13.30	3.12	5.10	3.12	0.10
	2012	18.18	1.40	18.70	1.40	0.80
1000 kernel weight, g	2011	44.6	2.0	45.5	2.0	0.75
	2012	27.6	0.9	27.3	0.9	0.83

Martinez et al., 2014

Registration of Zak *ERA8* Soft White Spring Wheat Germplasm with Enhanced Response to ABA and Increased Seed Dormancy

Shantel A. Martinez, Elizabeth C. Schramm, Tracy J. Harris, Kimberlee K. Kidwell, Kimberly Garland-Campbell, and Camille M. Steber*

BC Population Development

To clean up background EMS mutations not related to our trait of interest

Zak ERA8 x Zak WT

↓

Zak ERA8 BC1 x Zak WT

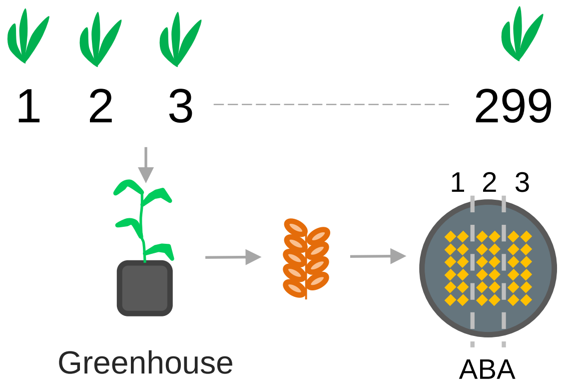
↓

Zak WT x Zak ERA8 BC2

(WT/WT) (ERA8/ERA8)

↓

Zak/Zak ERA8 BC3

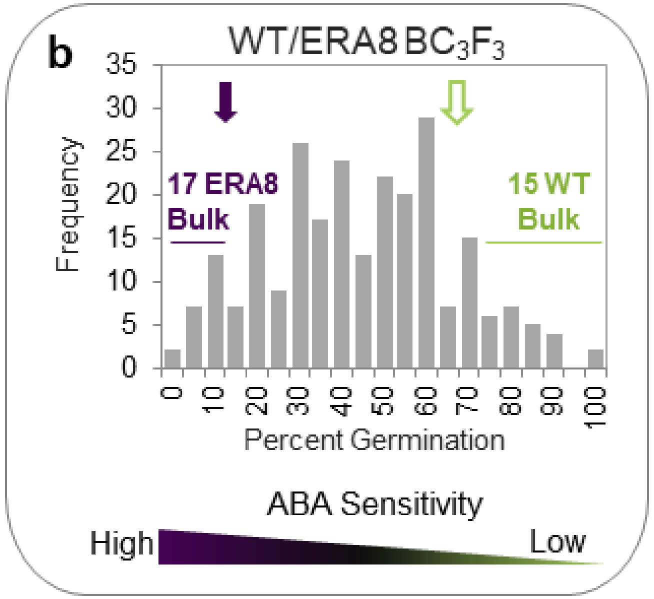


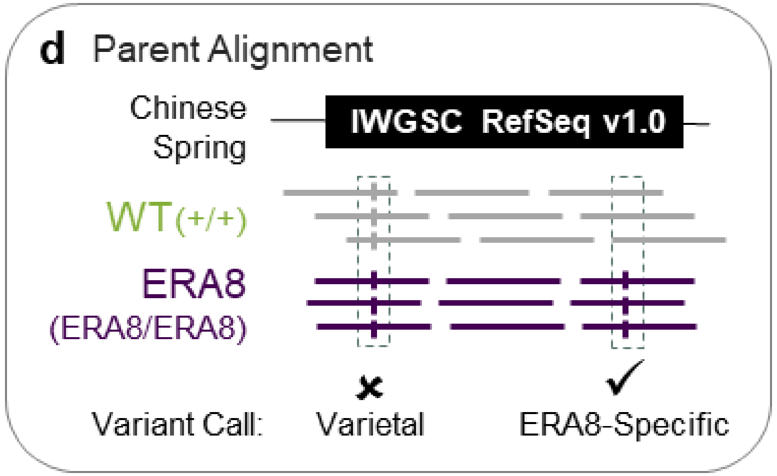
Only mutagen induced differences

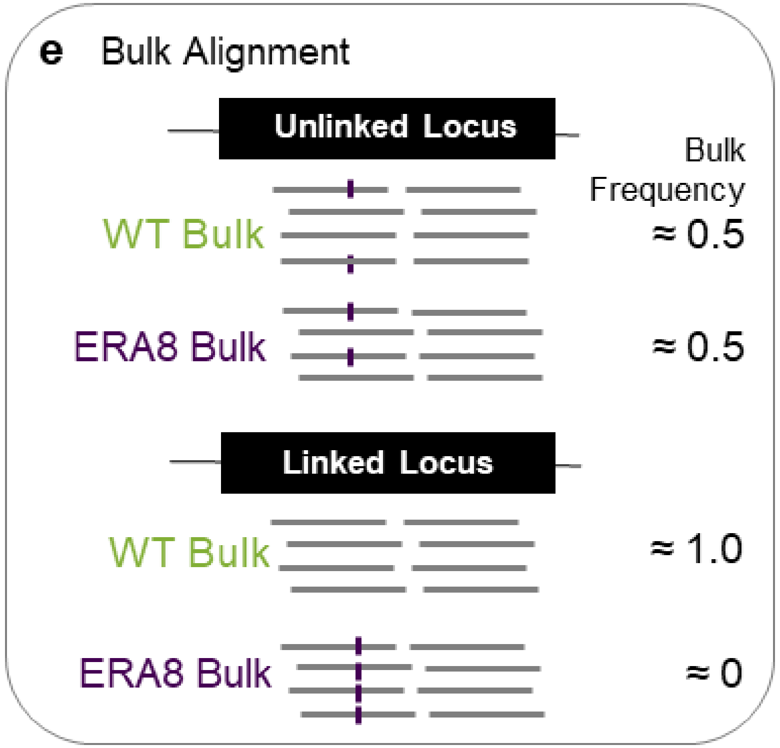
Martinez et al., 2020 (*accepted TAG*)

Bulked-Segregant Analysis

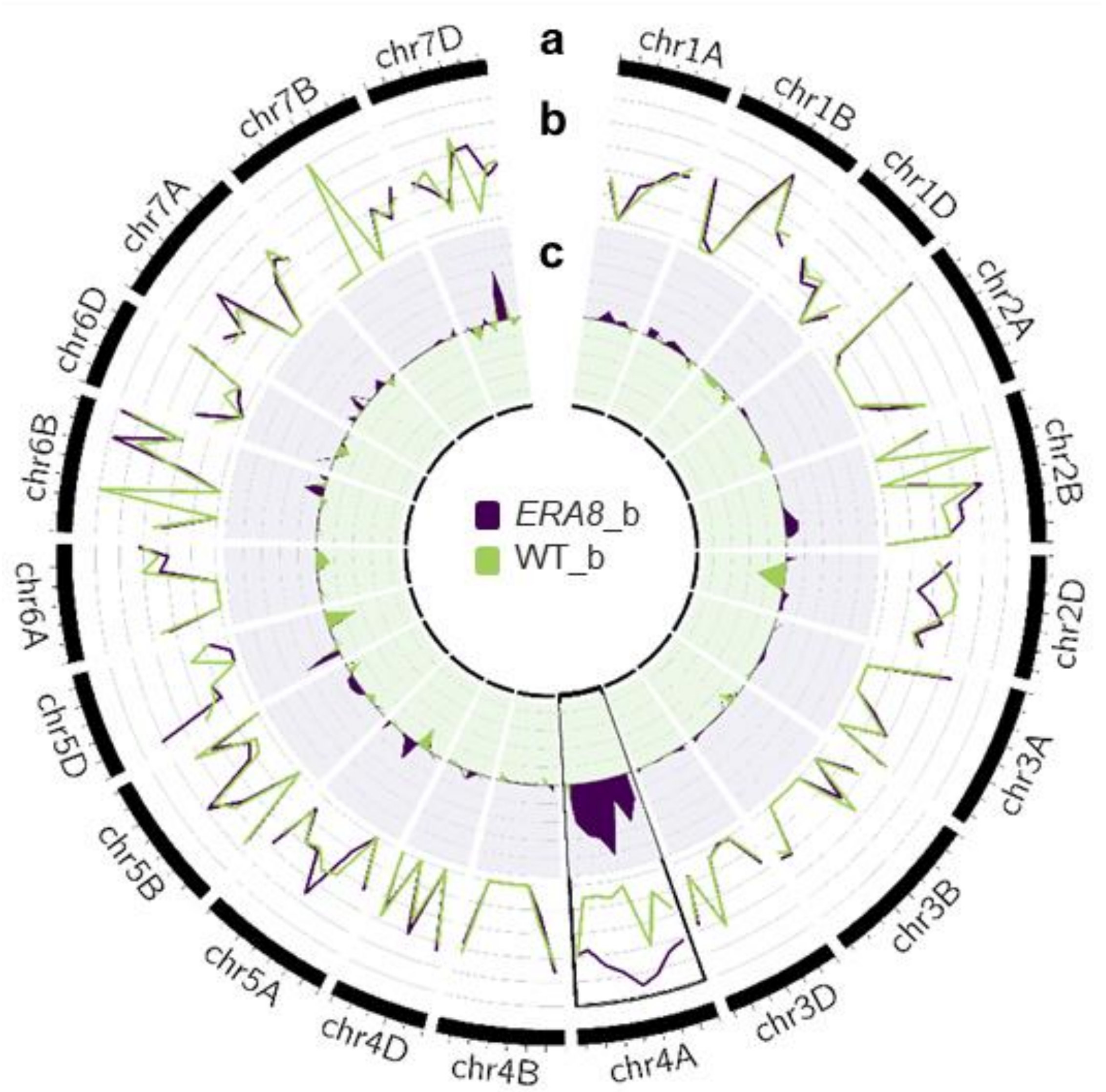
using exome-capture DNA sequencing

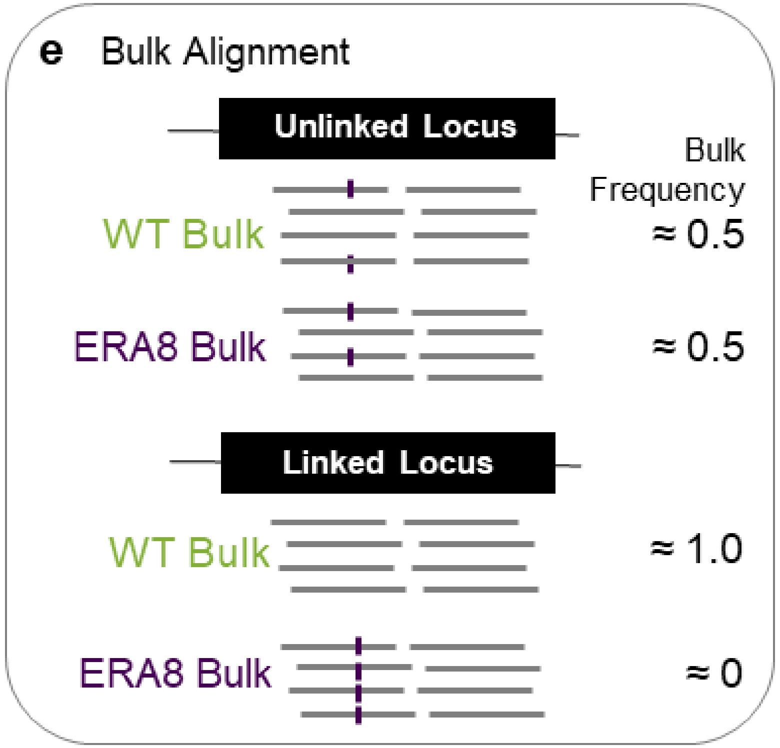






Martinez et al., 2020 (*accepted TAG*)





Linked locus on chromosome 4A

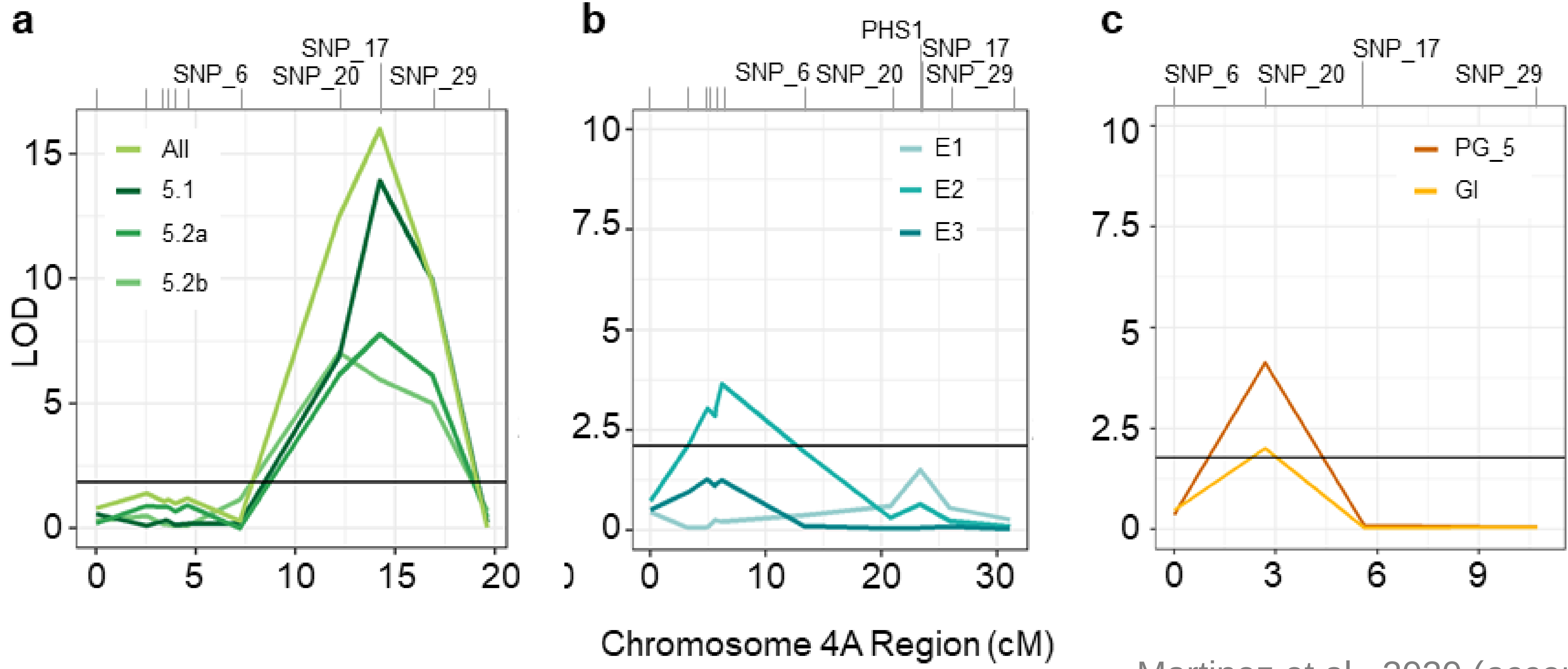
Martinez et al., 2020 (*accepted TAG*)

Fine Mapping in *ERA8*

Additional
Zak/ZakERA8 BC

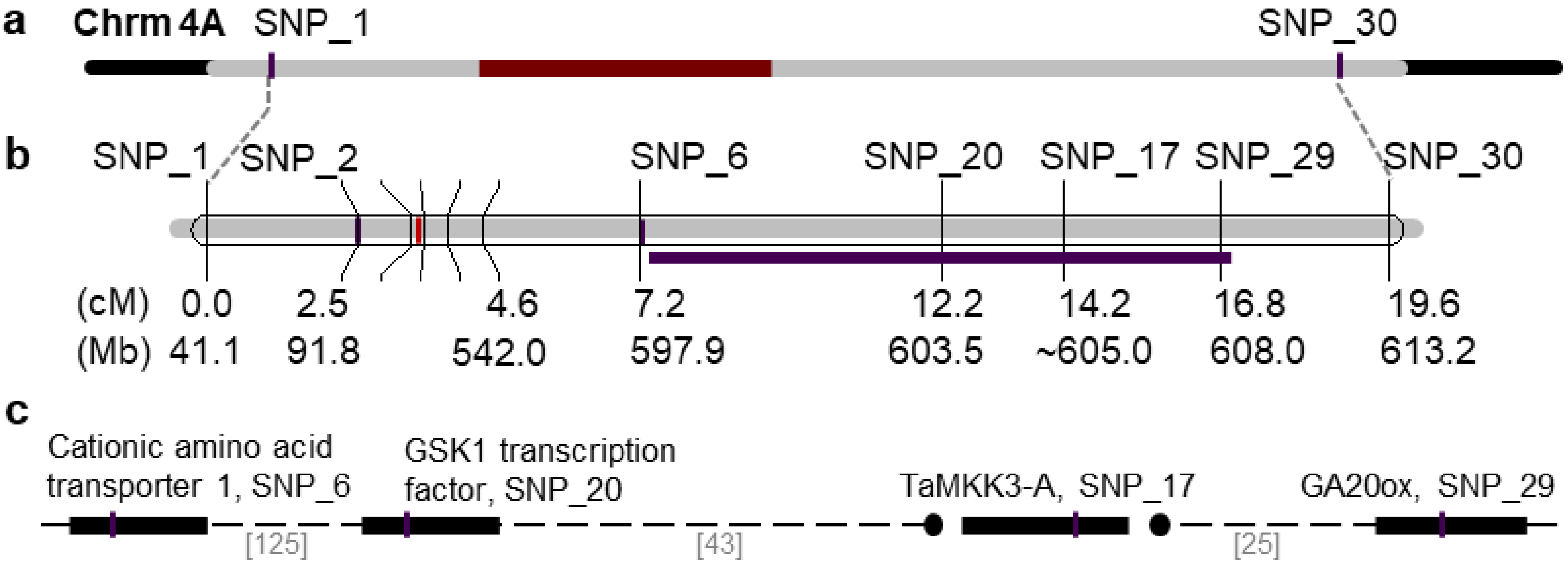
Louise/ZakERA8 RIL

Otis/ZakERA8 F2:F3



Martinez et al., 2020 (*accepted TAG*)

Fine Mapping in *ERA8*



Martinez et al., 2020 (*accepted TAG*)

Trait Introgression of *ERA8*

JOURNAL OF PLANT REGISTRATIONS
GERMPLASM

Martinez et al., 2014

WSU Winter Wheat

5 crosses
female/ (BC$_1$ERA8/male)
(~200)
20 crosses
(BC$_1$ERA8/male) / male
(~200)

USDA Club Wheat

24 crosses
Jasper / BC$_3$ERA8 (~800)
Bruehl / BC$_3$ERA8 (~800)
Kaseberg / BC$_3$ERA8
Brevor / BC$_3$ERA8
X010263-3C / BC$_3$ERA8
ARS010719-4L / BC$_3$ERA8

WSU Spring Wheat

3 crosses
Otis / BC$_2$ERA8 (529)
Louise / ERA8
x5 (698)
x12 (286)
BC$_1$ERA8 / Diva

RAGT Seed, UK

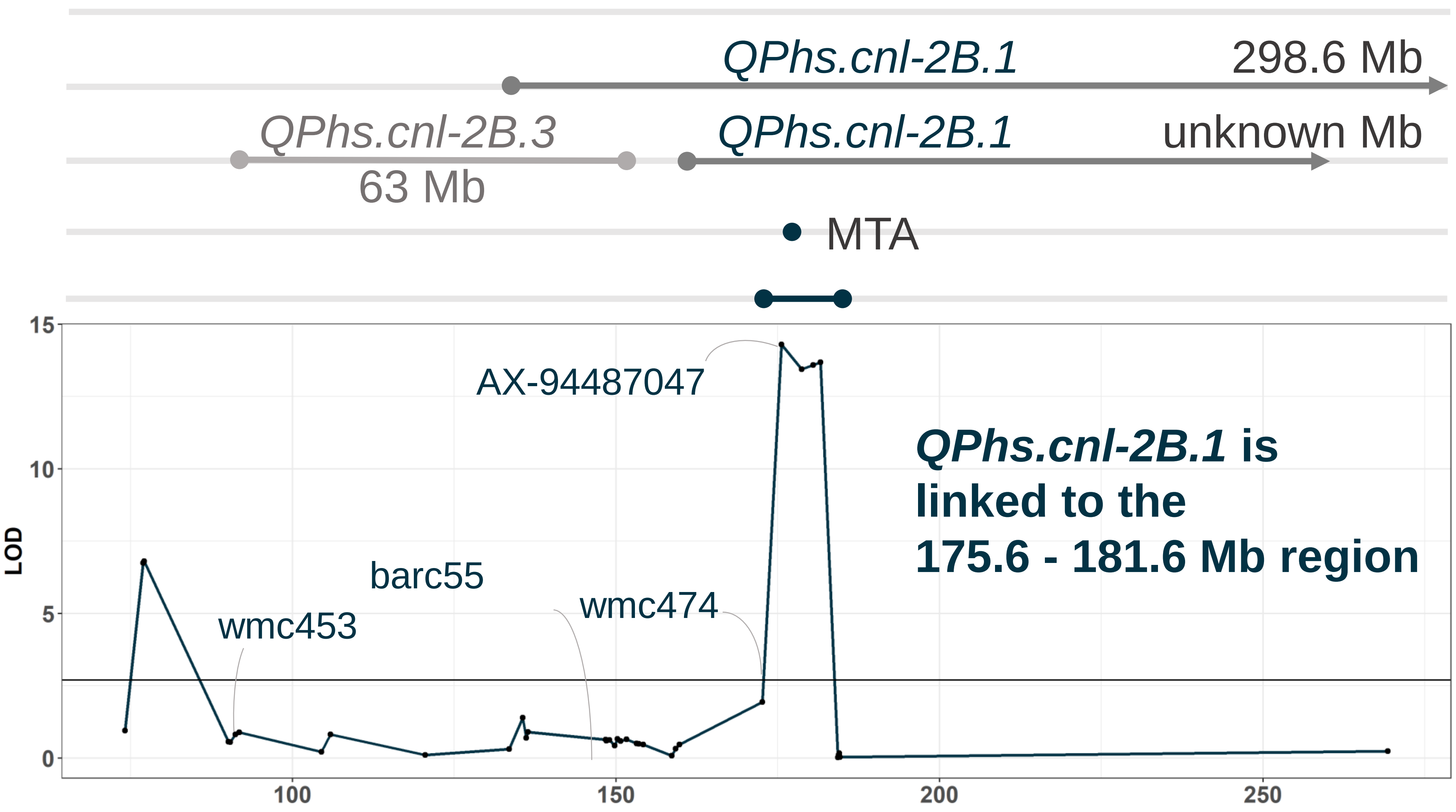
In collaboration with JIC

f Genome-specific KASP primers

Allele 1	FAM-CCTCTGCTATTTGCTTTAATCTCTc
Allele 2	HEX-CCTCTGCTATTTGCTTTAATCTCTt
Common	GGACTTGGCAGCATATGTCA

Martinez et al., 2020

Fine Mapping *QPhs.cnl-2B.1*



Chrm 2B

Munkvold et al. 2009

Somyong et al. 2014

GWAS

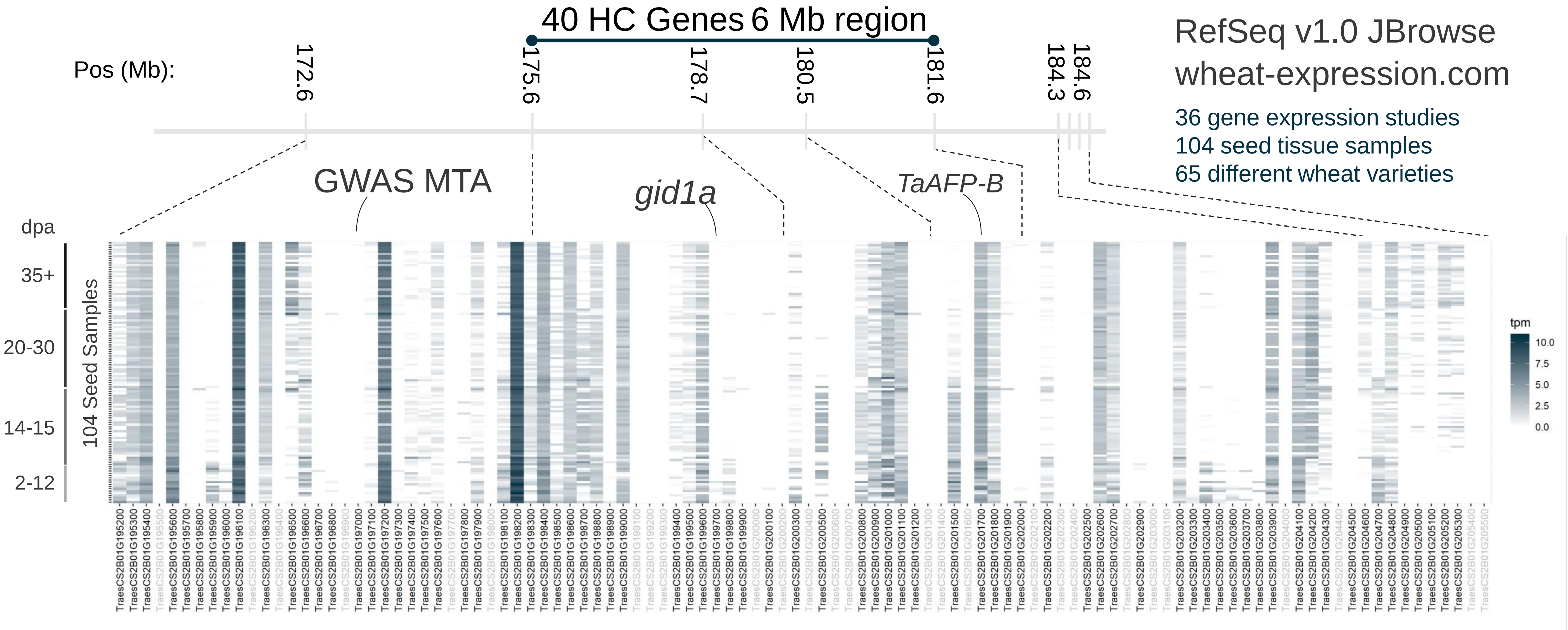
QTL Analysis

***QPhs.cnl-2B.1* is linked to the 175.6 - 181.6 Mb region**

430 BC1F5:8
PHS BLUP Score;
7 environments
SSR, Axiom, KASP markers
`cim` of qtl package in R

Martinez et al., unpublished

Fine Mapping *QPhs.cnl-2B.1*



RefSeq v1.0 JBrowse
wheat-expression.com

36 gene expression studies
104 seed tissue samples
65 different wheat varieties

(Borrill et al., 2016; IWGSC, 2018 Ramírez-González et al., 2019)

Trait Introgression *QPhs.cnl-2B.1*

Club wheat

USDA ARS

WA winter

NY winter

CORNELL UNIVERSITY
FOUNDED A.D. 1865

158 lines
elite nursery
past 10 years

Japanese Germplasm

NARO
National Agriculture and Food Research Organization

‘Tamaizumi’

Can we integrate genomic prediction in the breeding program to make more accurate decisions?

Genomic Prediction

fixed effects

random effects

Observations/Phenotype

u: GEBV

$$\mathbf{u} \sim N(0, K\sigma_g^2)$$

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{WZu} + \varepsilon$$

u is the vector of breeding values with estimated relationship matrix K

design matrix

e: experimental error

Models can calculate K differently:

Ridge Regression

Gaussian Kernel

design matrix

$$y_{obs} \sim \text{Loc} + \text{Yr} + \text{HarvDate} + (1|\text{variety})$$

Endelman 2011

Martinez et al., unpublished

Genomic Prediction

Prediction Accuracy (PA)

cor (y_{obs} , $GEBVs_{model}$)

cor (y_{obs_test} , $GEBVs_{test}$)

1.0 Perfect Prediction

0.0 Inaccurate

Prediction Model

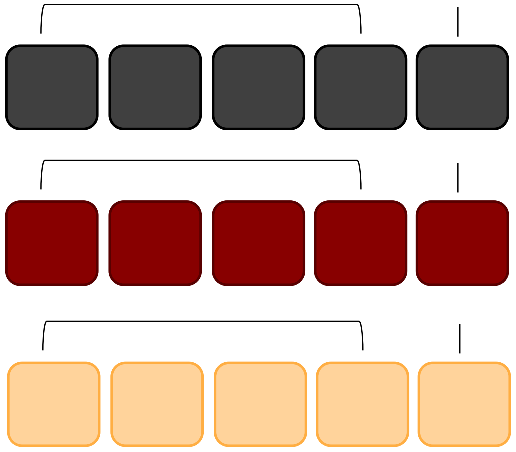
Gaussian kernel

Ridge regression

Output:
GEBVs $_{train+test}$

Five-fold CV

Phenotype



All Env	total n
Both	1287
Red	369
White	904

Genomic Prediction

Prediction Accuracy (PA)

Prediction Model

$\text{cor}\ (\ y_{obs}\ ,\ GEBVs_{model})$

$\mathbf{cor}\ (\ \mathbf{y}_{\mathbf{obs\_test}}\ ,\ \mathbf{GEBVs}_{\mathbf{test}})$

Gaussian kernel

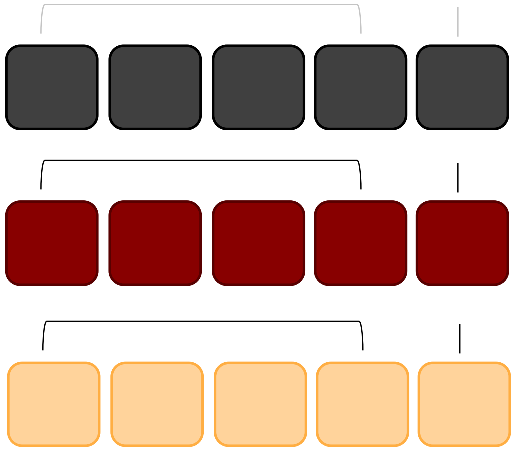
Ridge regression

Output:
$GEBVs_{\ train+test}$

Phenotype

80% Train

20% Test



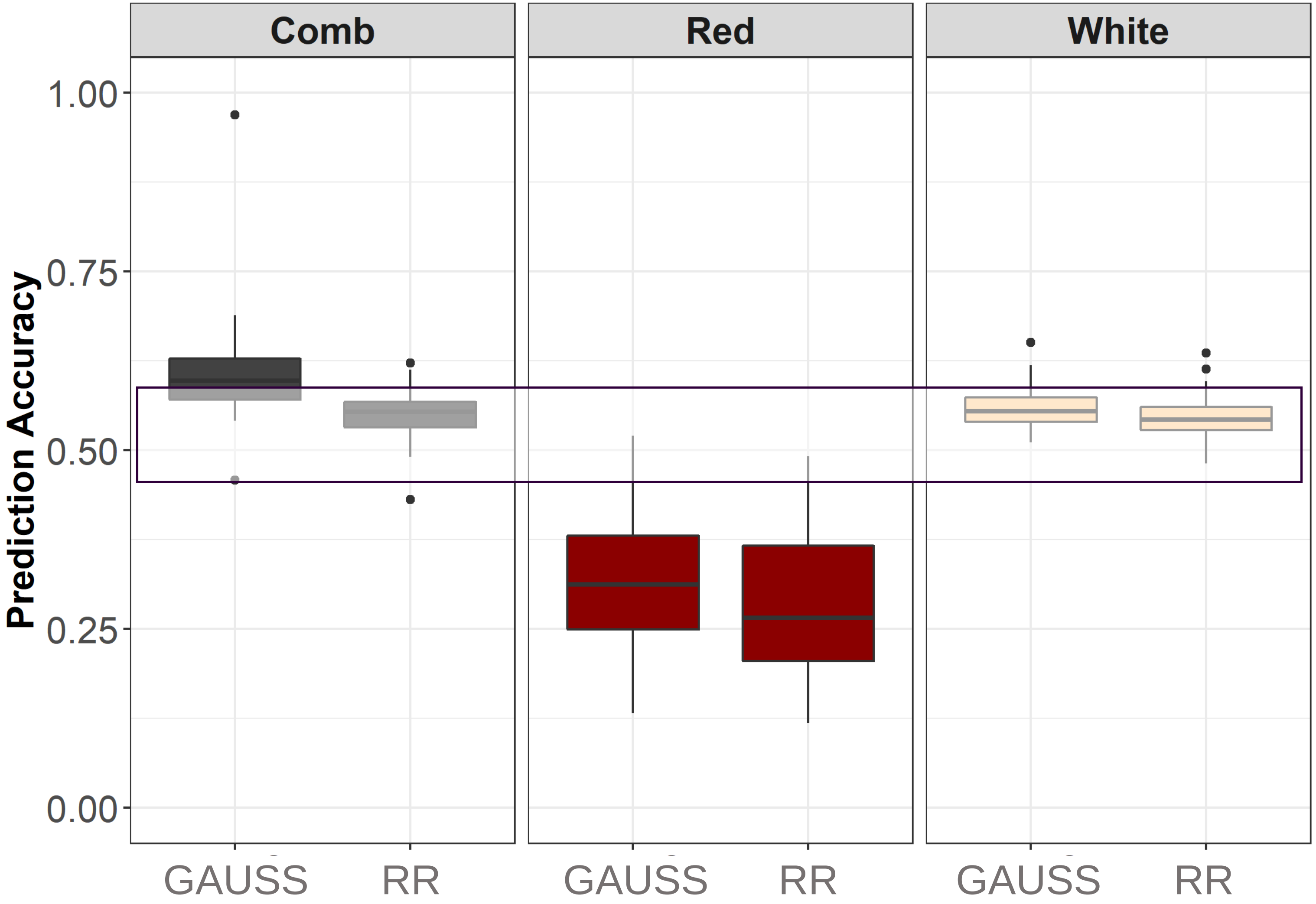
All Env	total n
Both	1287
Red	369
White	904

Both $y_{obs\_both} \sim \text{Loc} + \text{Yr} + \text{HarvDate} + \text{KColor} + (1|\text{variety})$

Red $y_{obs} \sim \text{Loc} + \text{Yr} + \text{HarvDate} + (1|\text{variety})$

White $y_{obs} \sim \text{Loc} + \text{Yr} + \text{HarvDate} + (1|\text{variety})$

A comparison of Ridge Regression and Gaussian kernel models across seed coat colors



Previous PHS GP studies	**KC**
Heffner et al. 2011 **PA** 0.52 – 0.53	White
Heslot et al. 2013 **PA** 0.47 – 0.57	White
Moore et al. 2017 **PA** 0.49 – 0.59	White & Red

GAUSS: Gaussian kernel model
RR: Ridge Regression model

rrBLUP

What does an accuracy of 0.6 even mean?

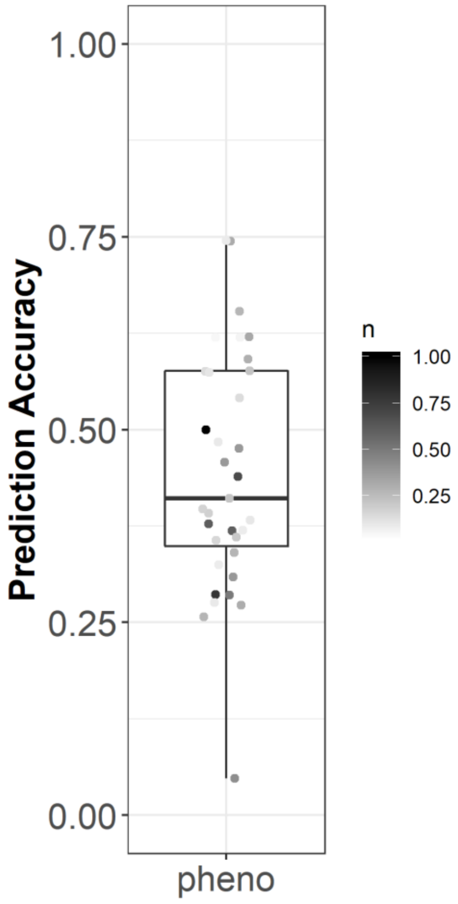
PA_{pheno}

Accuracy of an phenotypic estimate (without genetic data)

PA = cor(Obs_EnvA, Obs_EnvB)

cor(Obs2008, Obs2009)
.
.
.
cor(Obs2015, Obs2017)

Average PA
0.45

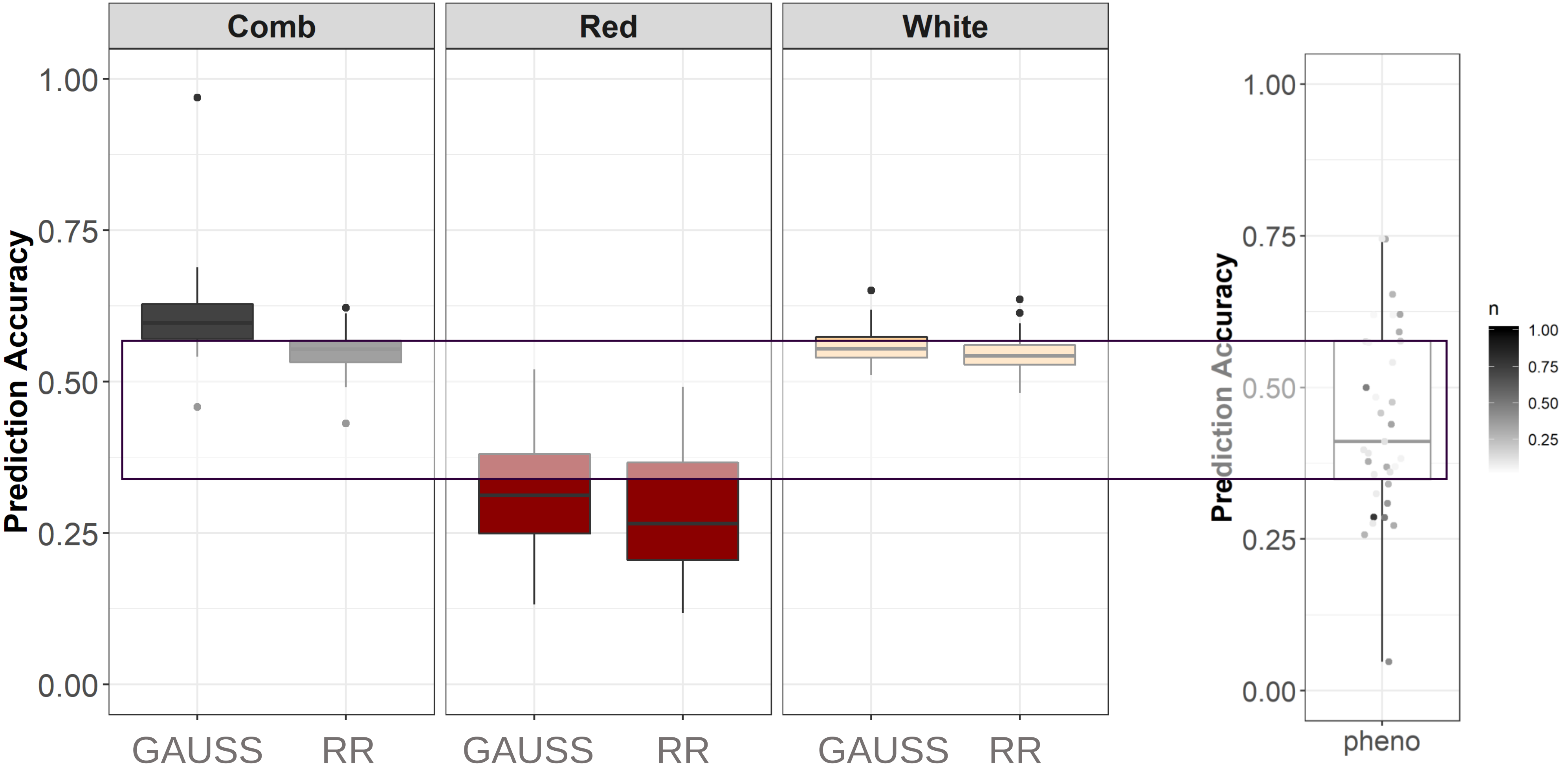


rrBLUP

The correlation (PA) from one environment to another environment is comparable, to genomic prediction accuracies.

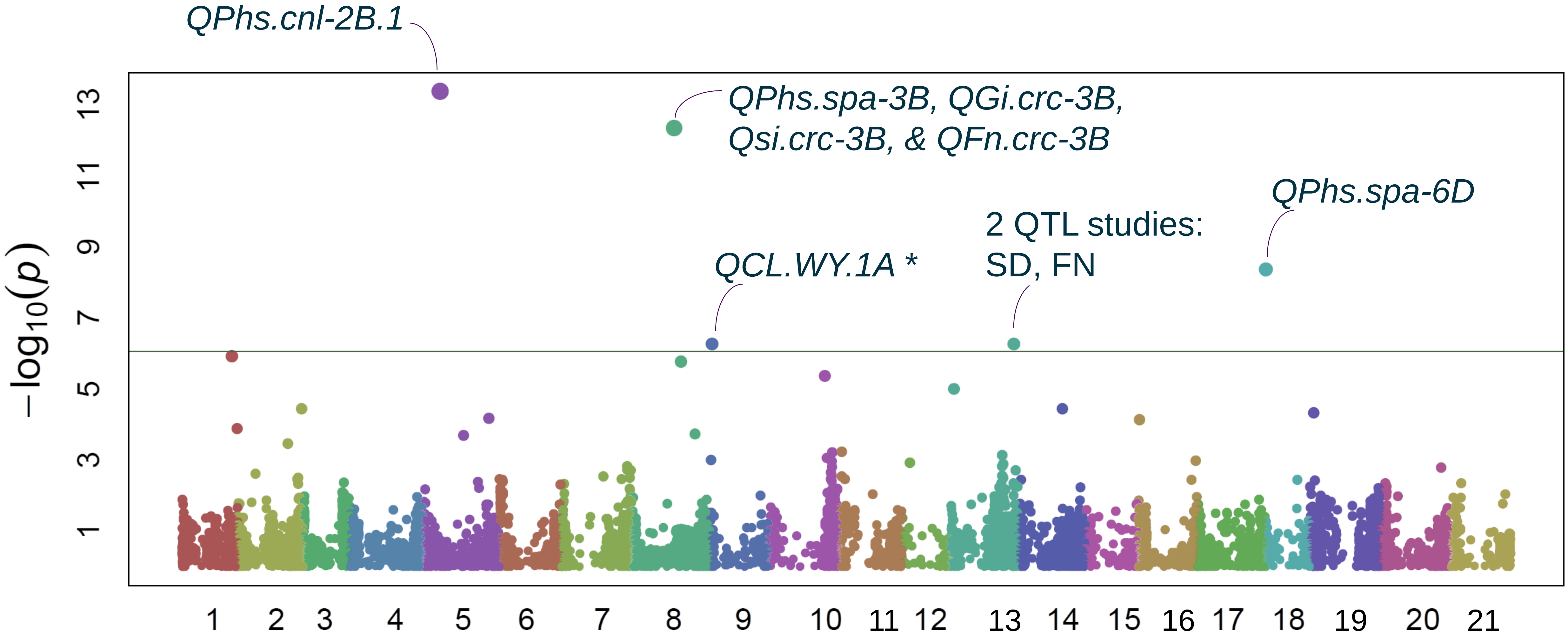
PA_{pred}

PA_{pheno}



rrBLUP

White kernel GWAS found multiple significant loci



FarmCPU

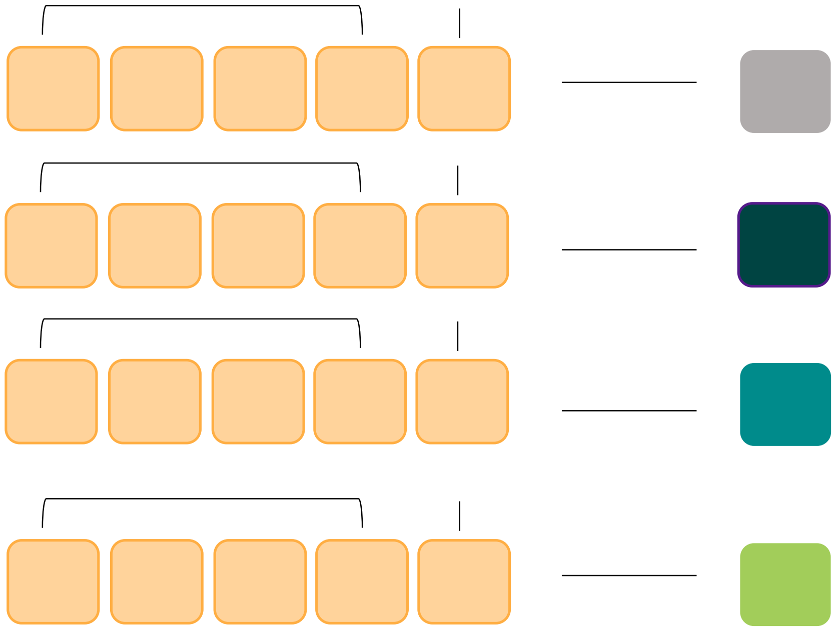
Zanetti et al., 2000; Munkvold et al. 2009; Fofana et al. 2009; Zhang et al. 2013;
Kumar et al. 2015; Martinez et al. 2018; Zuo et al., 2019

Will associated QTL improve prediction?

Observed phenotype mixed model

y ~ Loc+Yr+HarvD+(1|variety)

Five-fold CV



Prediction mixed model

y ~ Loc+Yr+HarvD+(1|variety) **(base model)**

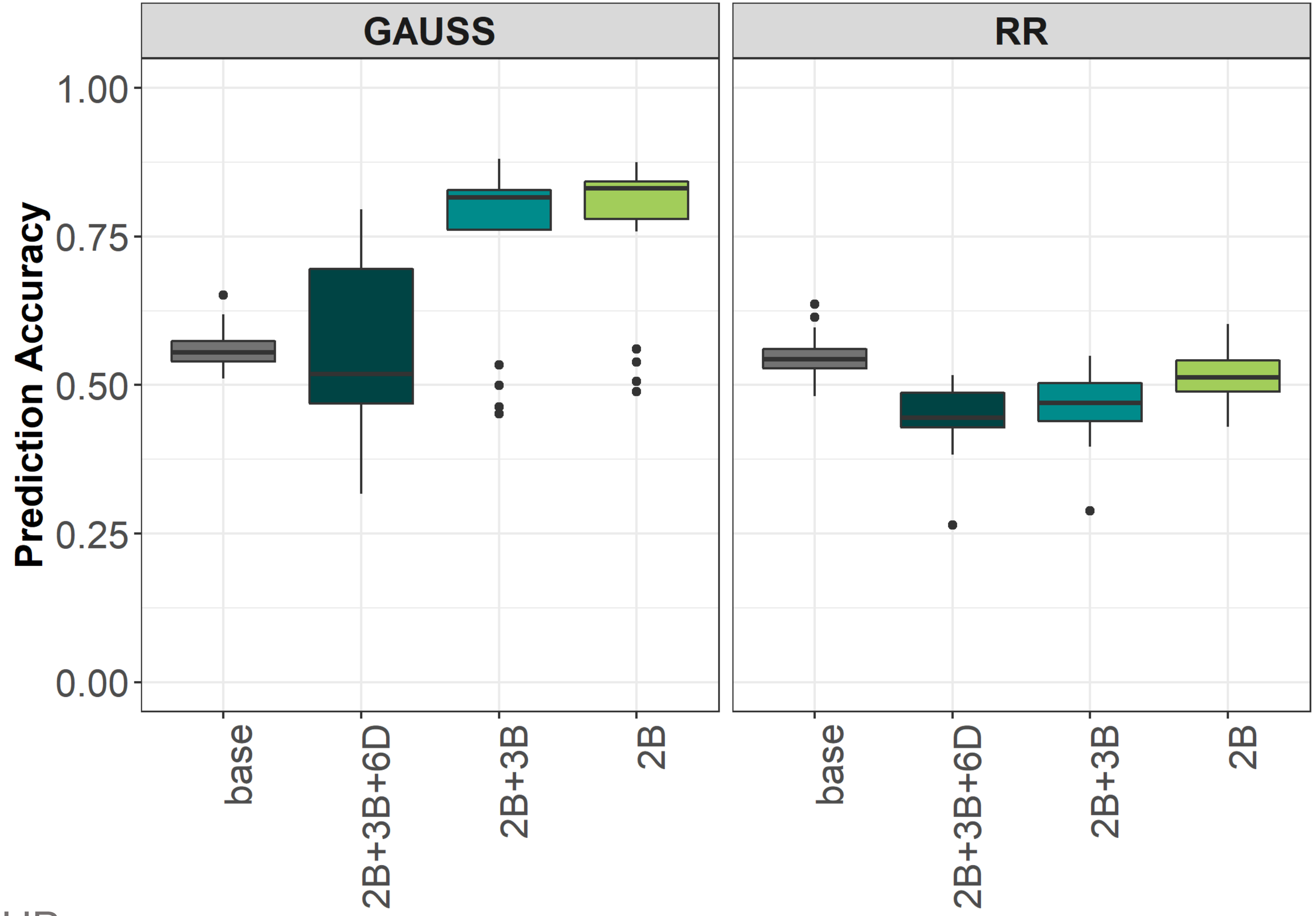
y ~ Loc+Yr+HarvD+**2B+3B+6D** +(1|variety)

y ~ Loc+Yr+HarvD+**2B+3B** +(1|variety)

y ~ Loc+Yr+HarvD+**2B** +(1|variety)

$$PA = cor\,(\,y_{obs\_test}\,,\,GEBVs_{test})$$

Adding significant QTL as fixed effects seems to improve accuracy for Gaussian kernel model



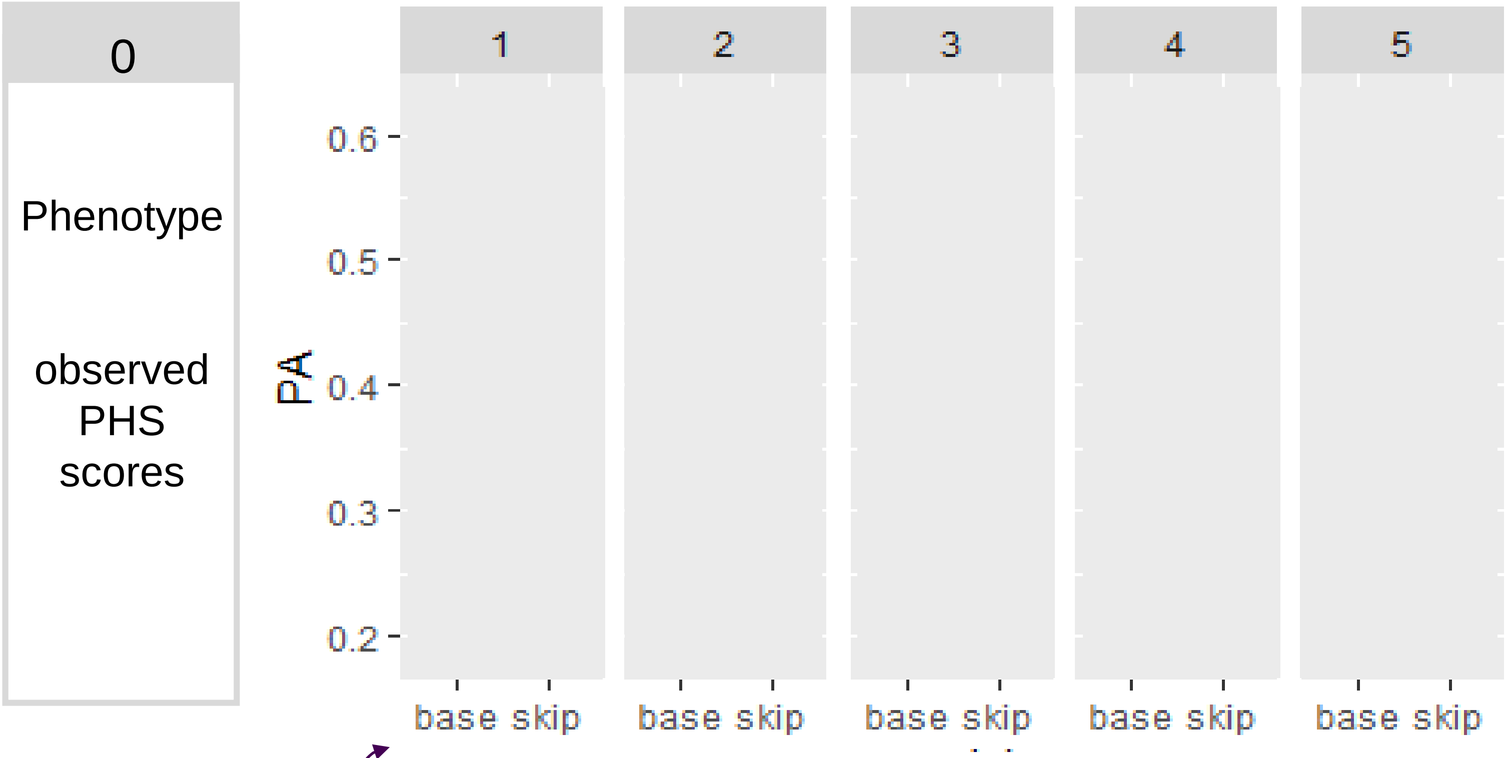
GAUSS additive + non-additive
RR only additive

GAUSS vs RR
The real question is, what do you want to use it for?
Want additive RR for selection
Want +non-add for prediction

Need to investigate if adding significant QTL markers as a fixed effect is overfitting the model?

rrBLUP

How many years between phenotyping can a breeding program skip without losing substantial PA?



Phentoype mixed model

GEBVs$_{yr}$ & y$_{obs}$

y ~ Loc+Yr+HarvD+(1|variety)

Prediction Accuracy

base: if phenotype every year, and predict the following year

PA = cor($y_{obs\_yr0+--+yr(N-1)}$, GEBVs$_{yrN}$)

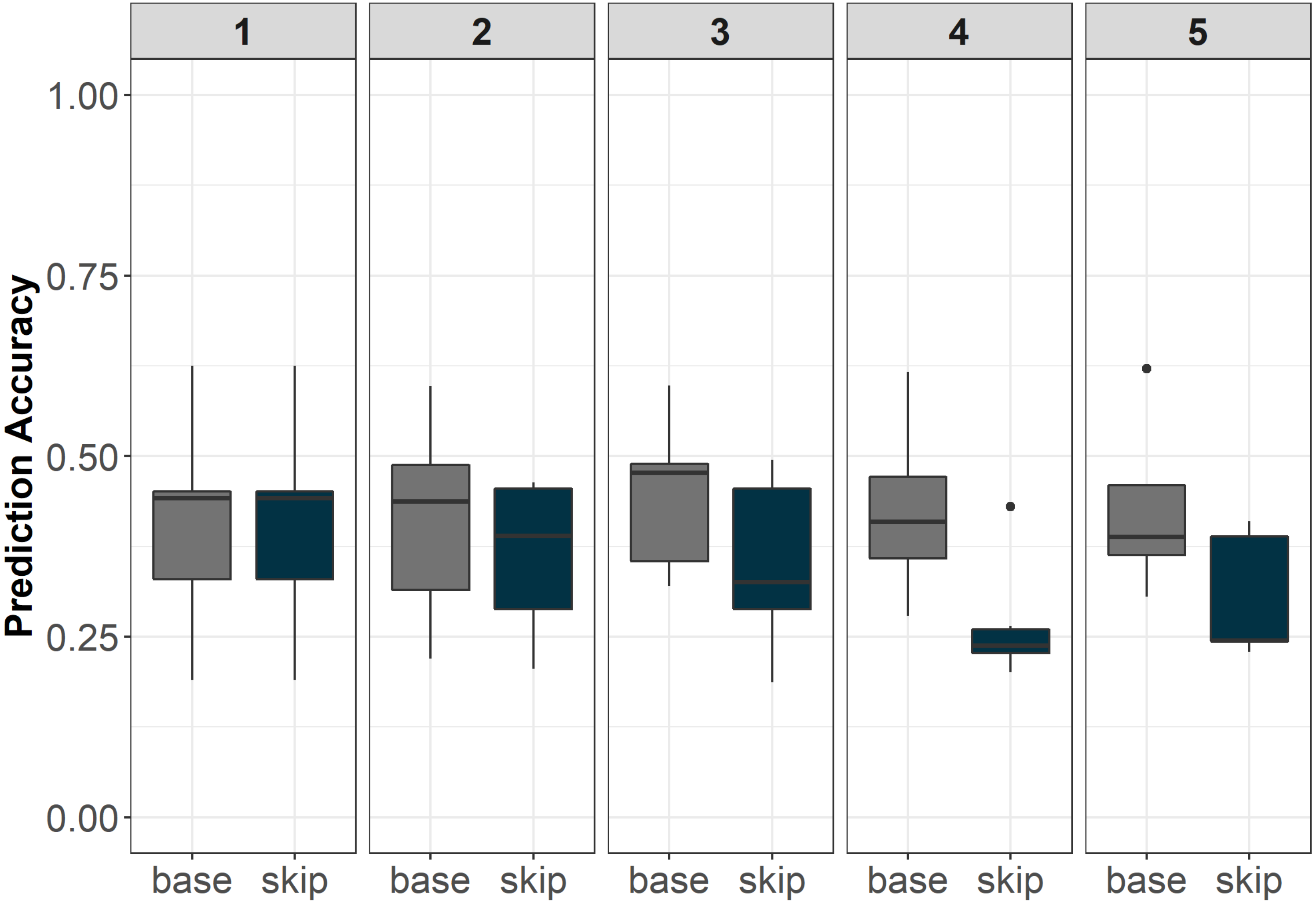
skip: 1, 2, 3, 4, 5 years between phenotyping

PA = cor(y_{obs_yr0} , GEBVs$_{yrN}$)

rrBLUP

Ridge Regression
White KC only

How many years between phenotyping can a breeding program skip without losing substantial PA?



It may be possible to phenotype for PHS tolerance every couple of years, without effecting accuracy

Something to consider:
- Increased genetic diversity year to year will reduce ability to accurately calculated GEBVs
- Year to year GEBV calculations tend to have lower PA than GEBV calculated from hundreds of individuals over multiple environments

rrBLUP

Ridge Regression
White KC only

Why use genomic prediction for this trait?

Phenotyping occurs at
the end of the growth cycle

No “mid-cycle” selections before harvest

These phenotypic hurdles are not exclusive to PHS screening

Very labor intensive

Precise sampling

Disease Screening

Traits dependent on inoculation, severity, etc can benefit from genotypic predictions of the phenotype

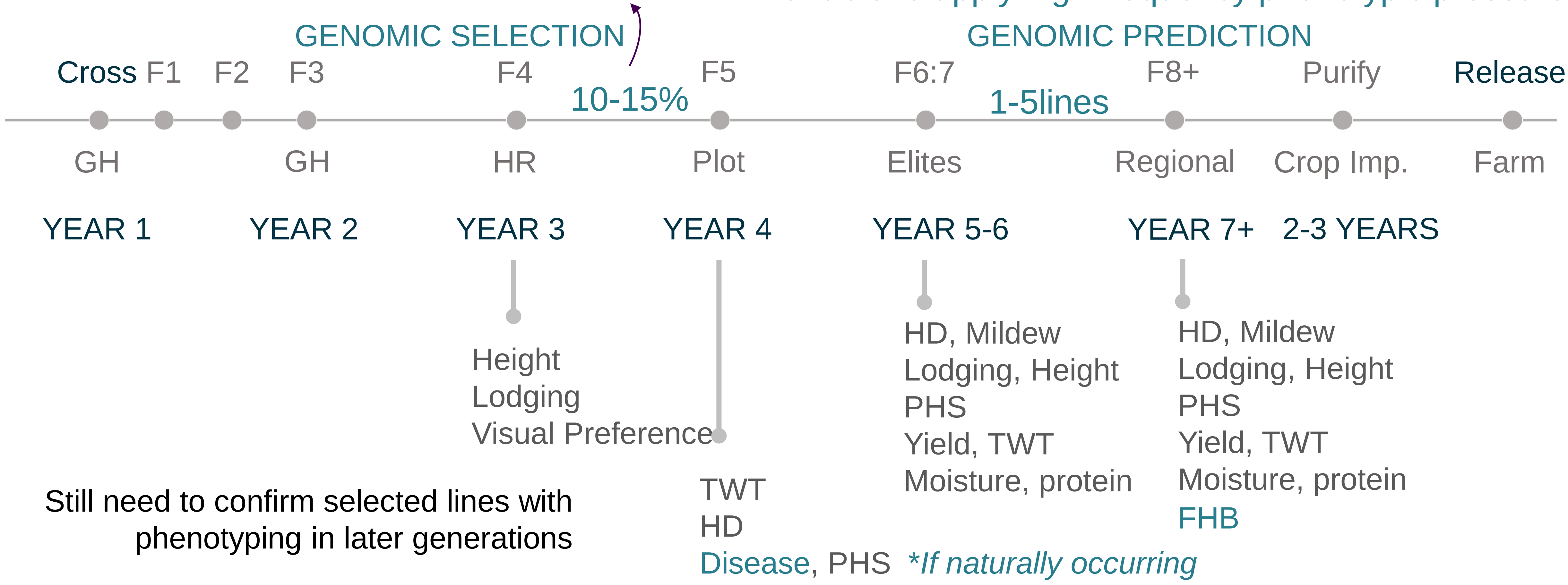
FHB (2 field seasons)

Stripe Rust (7 field seasons)

Breeding with Genomic Prediction

GEBVs can help guide selection; less likely to throw out PHS Tolerant lines when applying selection pressure

Assist in predicting PHS tolerance for variety selection, if unable to apply high frequency phenotypic pressure



GENOMIC SELECTION

GENOMIC PREDICTION

Cross F1 F2 F3 F4 F5 F6:7 F8+ Purify Release

10-15%

1-5lines

GH GH HR Plot Elites Regional Crop Imp. Farm

YEAR 1 YEAR 2 YEAR 3 YEAR 4 YEAR 5-6 YEAR 7+ 2-3 YEARS

Height
Lodging
Visual Preference

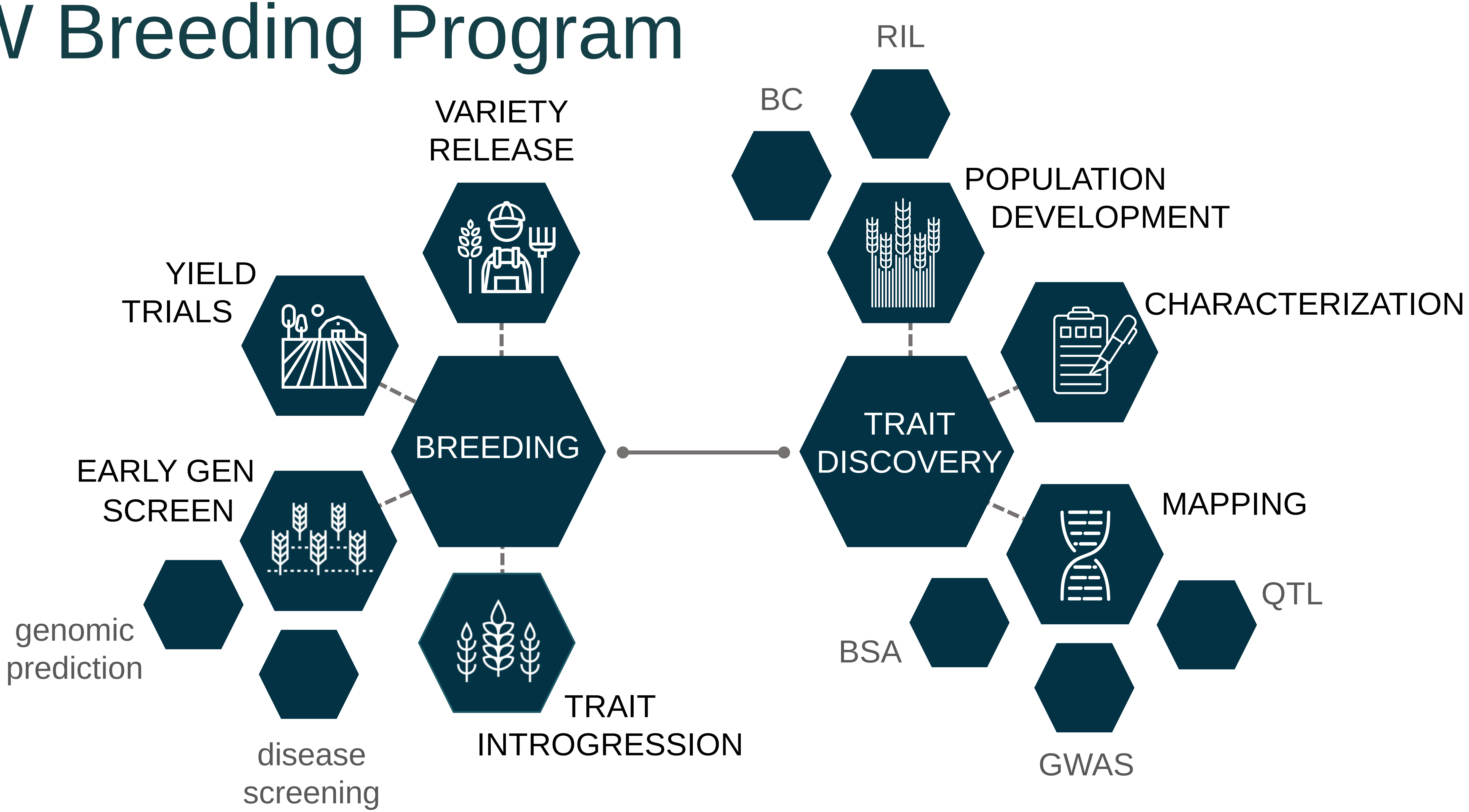
TWT
HD
Disease, PHS *If naturally occurring

HD, Mildew
Lodging, Height
PHS
Yield, TWT
Moisture, protein

HD, Mildew
Lodging, Height
PHS
Yield, TWT
Moisture, protein
FHB

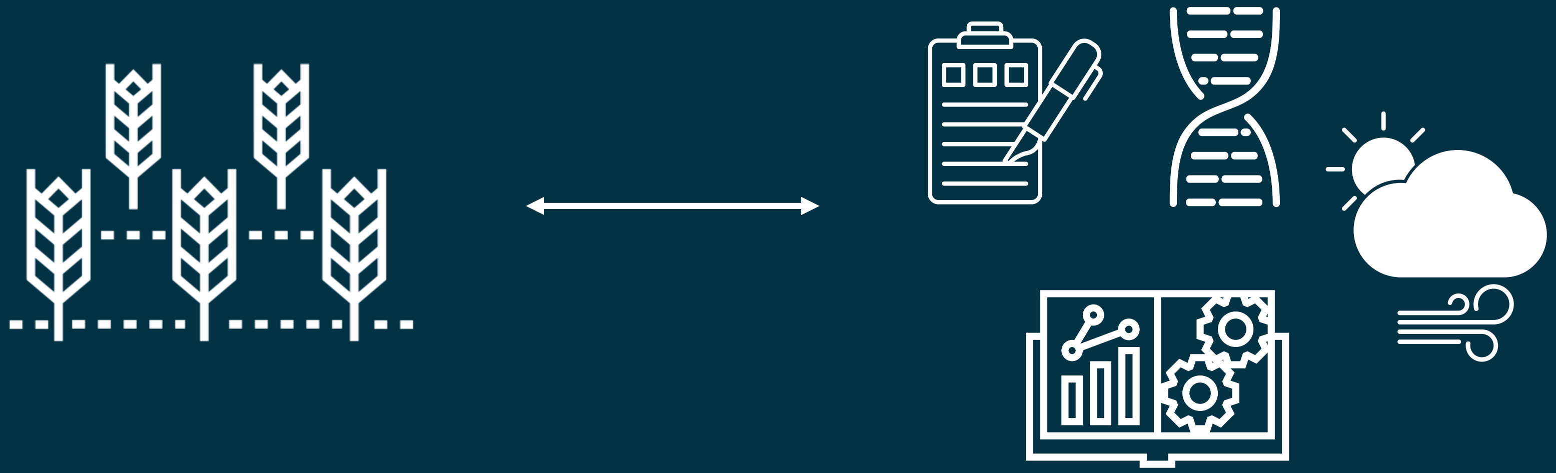
Still need to confirm selected lines with phenotyping in later generations

HRW Breeding Program



As your next generation plant breeder:

I aim to fuse wheat breeding with predictive analytics



Thank you for your attention

LinkedIn

shantel-a-martinez

Website

shantel-martinez.github.io

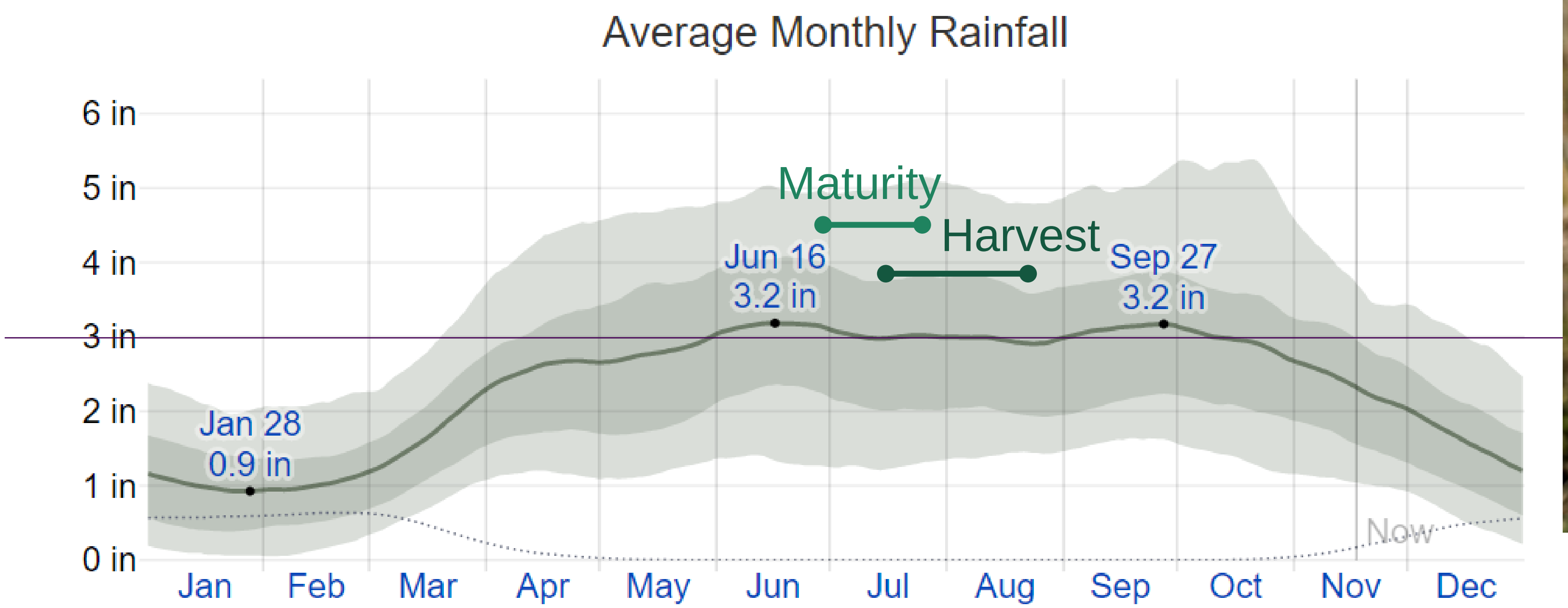
Email

shantel.a.martinez@gmail.com
sam594@cornell.edu

Twitter

@s_amealia

How can my PHS skills be used?

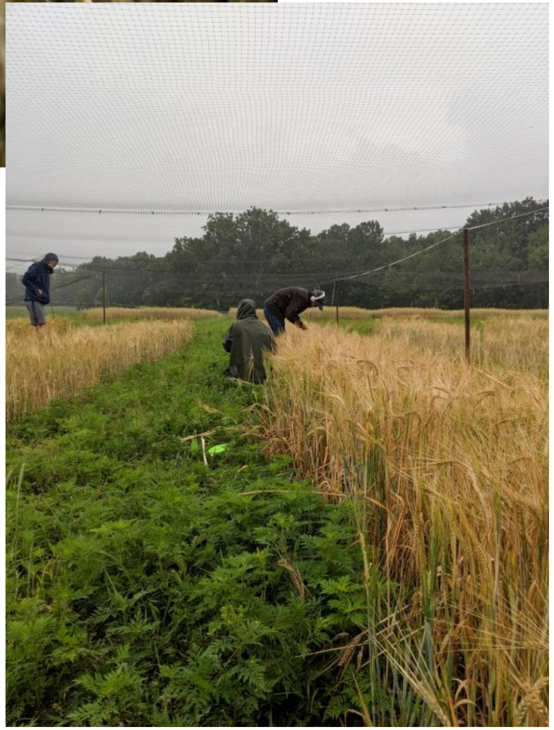


Ithaca, NY

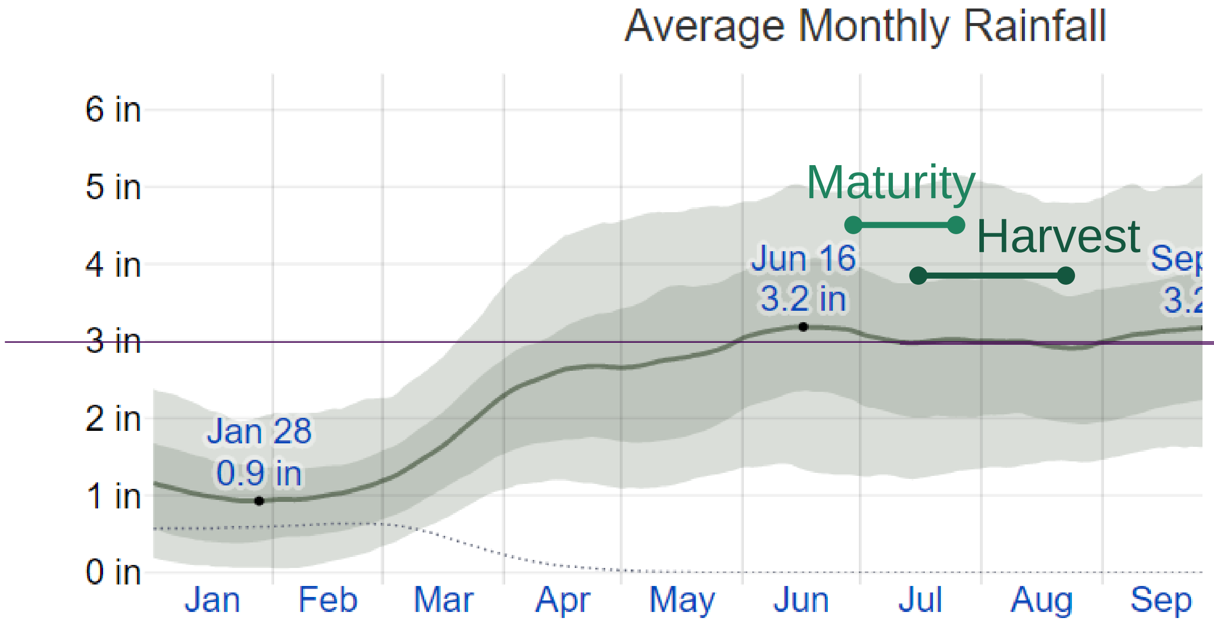
weatherspark.com

1.5 days
rain 2018

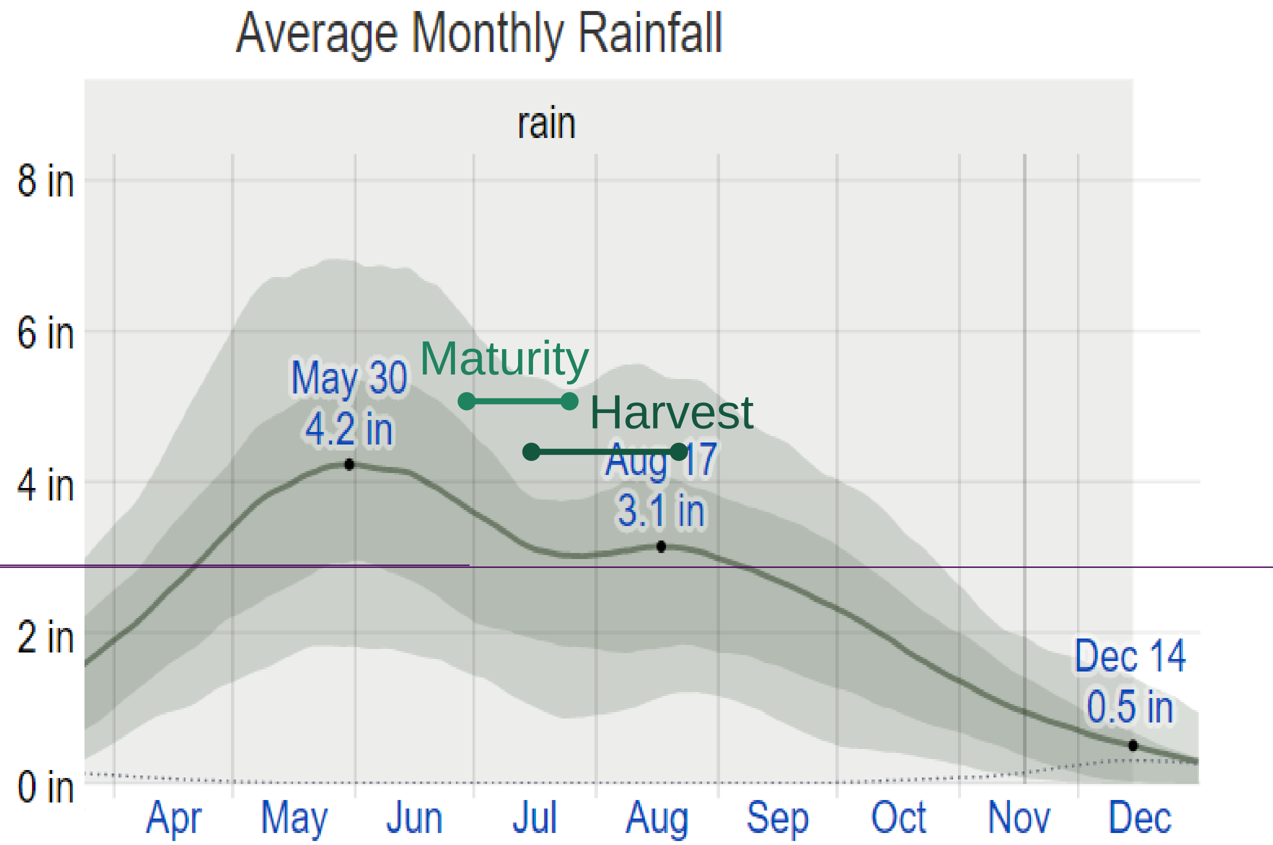
2019



My experience can be useful if a program starts growing malting barley in NE



Ithaca, NY



Lincoln, NE

weatherspark.com

Mixed Model

fixed effects

random effects

Observations/Phenotype

g: BLUP

y_{obs}

$$y = \mu + X\beta_{loc} + X\beta_{yr} + X\beta_{harv} + Zg + \varepsilon$$

population mean

design matrix

loc: location

yr: year

harv: sampling date

design matrix

e: experimental error

$$y_{obs} \sim \text{Loc} + \text{Yr} + \text{HarvDate} + (1|\text{variety})$$