



Identifying germplasm for resistance to blackleg and Verticillium stripe diseases and mapping molecular markers for marker-assisted selection



**Mukhlesur Rahman
North Dakota State University, USA**

Impact of Blackleg (*Leptosphaeria maculans*) disease in ND

- **Yield loss = 5% to 20%, can reach over 50%.**
- **Economic loss = \$32 to \$128 million, could be \$322 million.**
- **New races can affect the Resistant cultivars.**
- **Can survive in infected stubbles and seeds for 3-5 years**

Disease Managements

- **Resistant cultivars:** the best practice, but new race is a big problem.
- **Crop rotation:** Crop rotation can help to reduce blackleg.
- **Fungicides:** Fungicide (Helix® Vibrance, Prosper® EverGol, Flutriafol, Azoxystrobin + Difenoconazole etc.) seed treatments and foliar spray can help manage blackleg, especially in fields at high risk. However, fungicide residues can contaminate water resources, harm aquatic ecosystems, and affect non-target organisms.

Impact of Verticillium Stripe disease in North Dakota

- **New disease to canola in North America, and 1st Reported in 2023 in the USA and 2014 in Canada.**
- **The symptoms are similar to those of blackleg disease, and both diseases may occur together at the same time.**
- **The infection starts from roots at seedling stage and moves to plant.**
- **The pathogen can survive in the soil over 10 years.**

Disease Managements

- **Resistant Cultivars** – when available.
- **Crop Rotation** – Rotate canola with non-host crops like cereals (wheat, barley).
- **Field Sanitation** – minimize soil movement, sanitize equipment.
- **Fungicides** – Currently, no effective fungicides are available.

Research Objectives

- 1. Germplasm identification:** Identification of germplasm resistant or tolerant to blackleg and verticillium stripe diseases in canola.
- 2. Markers Identification:** markers linked to genes associated with blackleg and verticillium stripe diseases.
- 3. Genomic Prediction:** Development of genomic prediction models for blackleg and verticillium stripe diseases.



Study of Blackleg Disease

1. Greenhouse Study
2. Field Study
3. GWAS Analysis
4. Genomic Prediction

Experiments

1. Greenhouse Experiments (cotyledon inoculation)

- Genotypes = 316 (RILs+ Germplasm)
- RCBD with 3 reps, 5 seedlings per rep
- GH Study repeated two times
- Phenotypic data
 - GH score at 7 DAI
 - GH score at 10 DAI
 - GH score at 14 DAI



Disease Score: 0 to 5 scale

Experiments

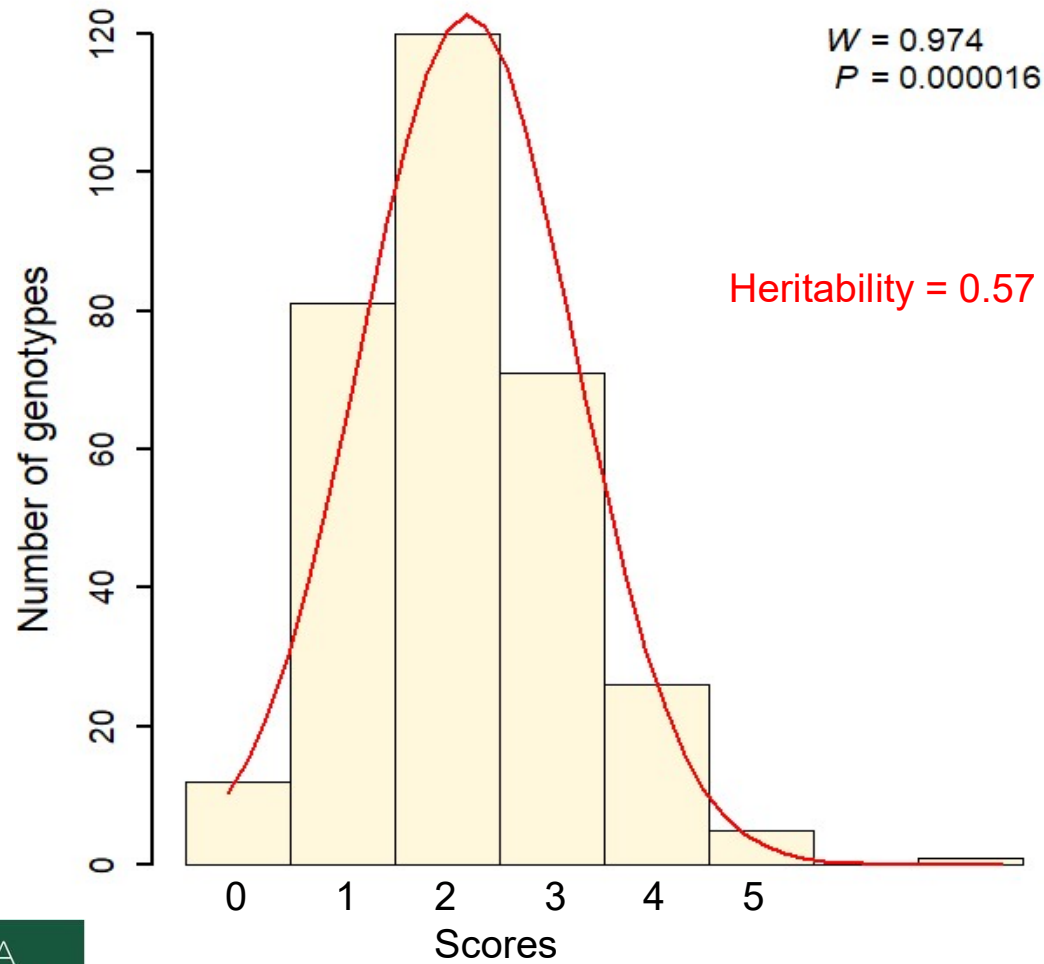
2. Field Experiments (artificial inoculation)

- **Site: Historical Blackleg nursery (16 years continuous Blackleg inoculation)**
- **Genotypes = 194 (RILs + Germplasm)**
- **RCBD with 3 reps, 3-12 plants per rep**
- **Phenotypic data = Plant at Maturity**



Disease Score: 0 to 5 scale

Distribution of Scores – GH – 10 DPI



Genome-wide association mapping – GH – 10 DPI

GWAS analysis using GAPIT in R

Total genotypes used

- 316 RILs + Germplasm accessions

GH score at 7 DAI

GH score at 10 DAI

Total markers used

GH score at 14 DAI

- SNPs = 17,360

GAPIT Models (8) (R package GAPIT)

- Single-locus (5): GLM, MLM, CMLM, ECMLM, SUPER
- Multi-locus (3): MLMM, FarmCPU and BLINK

Model selection

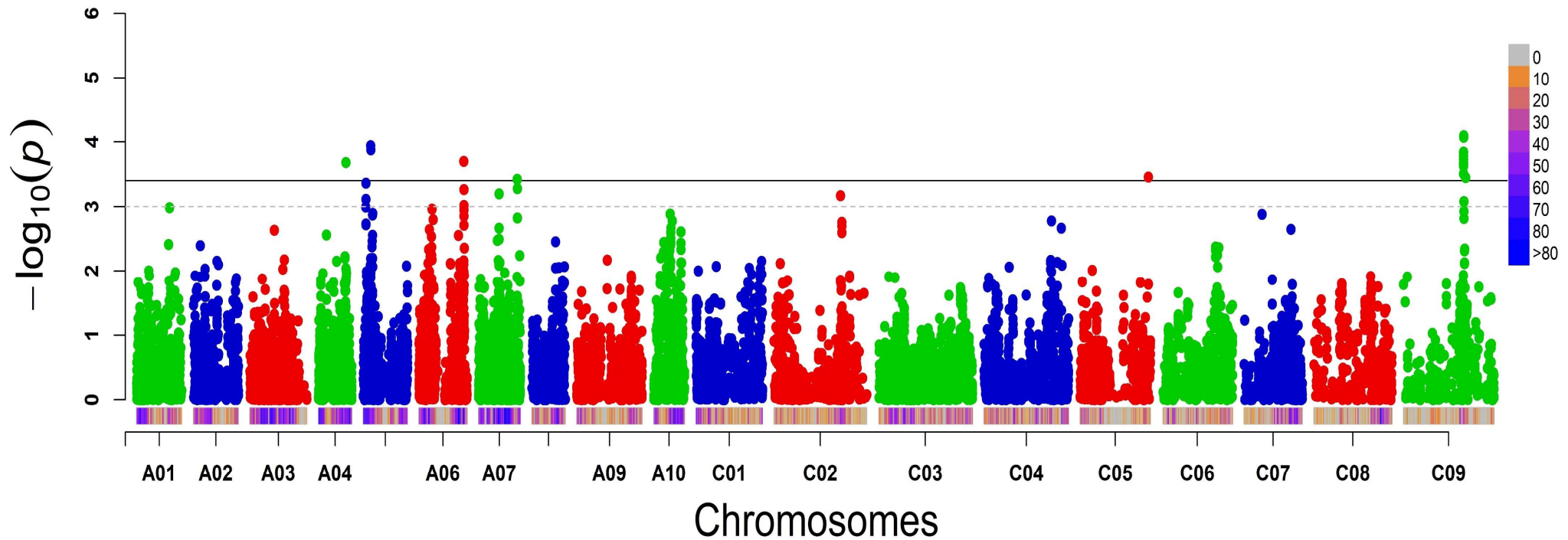
- Lowest mean of squared difference (MSD) values
- Quantile-quantile (QQ) plots

Significant SNP selection

- p -value threshold by Li & Ji (2005) method

Genome-wide association mapping – GH – 10 DPI

Manhattan Plot (significant markers)



Genome-wide association mapping – GH – 10 DPI

SNP	Chr	Pos	P.value	MAF	Effect
Bn-scaff_15576_1-p375237	C09	39978725	0.00008	0.50	0.13
Bn-scaff_15576_1-p368057	C09	39968929	0.00008	0.50	0.13
Bn-A05-p4616151	A05	4897485	0.00011	0.33	0.13
Bn-A05-p4644984	A05	4933375	0.00013	0.34	0.12
Bn-scaff_15576_1-p335372	C09	39916387	0.00014	0.48	0.13
Bn-scaff_15576_1-p323211	C09	39907058	0.00015	0.45	0.13
Bn-scaff_15576_1-p307207	C09	39892931	0.00016	0.48	0.13

Chromosomes:

- A01 (1)
- A04 (1)
- A05 (5)
- A06 (3)
- A07 (3)
- C02 (1)
- C05 (1)
- C09 (13)

Greenhouse study: 28 Significant SNP's

Bn-scaff_15576_1-p265897	C09	39804986	0.00021	0.49	0.12
Bn-scaff_15576_1-p316507	C09	39900173	0.00023	0.48	0.12
Bn-scaff_15576_1-p329552	C09	39915031	0.00023	0.48	0.12
Bn-scaff_15576_1-p328628	C09	39914095	0.00024	0.47	0.12
Bn-scaff_15576_1-p337457	C09	39921128	0.00031	0.47	0.12
Bn-scaff_21634_1-p62623	C05	45137493	0.00035	0.10	-0.20
Bn-scaff_21311_1-p20912	C09	41245333	0.00035	0.42	-0.13
Bn-A07-p20580512	A07	25648921	0.00038	0.45	-0.14
Bn-A05-p1475369	A05	1665417	0.00044	0.34	0.12
Bn-A07-p20611621	A07	25690340	0.00053	0.41	-0.14
Bn-A06-p24267499	A06	29721348	0.00054	0.38	0.11
Bn-A07-p10409623	A07	13637355	0.00064	0.28	-0.12
Bn-scaff_17109_4-p159217	C02	44110056	0.00068	0.39	0.11
Bn-A05-p1499192	A05	1689501	0.00077	0.30	0.11
Bn-scaff_15576_1-p455734	C09	40082722	0.00084	0.30	0.12
Bn-scaff_15754_1-p496098	A06	29790879	0.00096	0.29	0.12
Bn-A05-p1462045	A05	1648082	0.00102	0.33	0.11
Bn-A01-p1462045	A01	21710005	0.00104	0.23	-0.14

Genome-wide association mapping – Greenhouse study

Data at 7 DAI = 3 sig markers

Data at 10 DAI = 28 sig markers

Data at 14 DAI = 17 sig markers

Genome-wide association mapping – Field – at maturity

GWAS analysis using GAPIT in R

Total genotypes used

- 194 RILs + Germplasm accessions

Total markers used

- SNPs = 16,415

GAPIT Models (8) (R package GAPIT)

- Single-locus (5): GLM, MLM, CMLM, ECMLM, SUPER
- Multi-locus (3): MLMM, FarmCPU and BLINK

Model selection

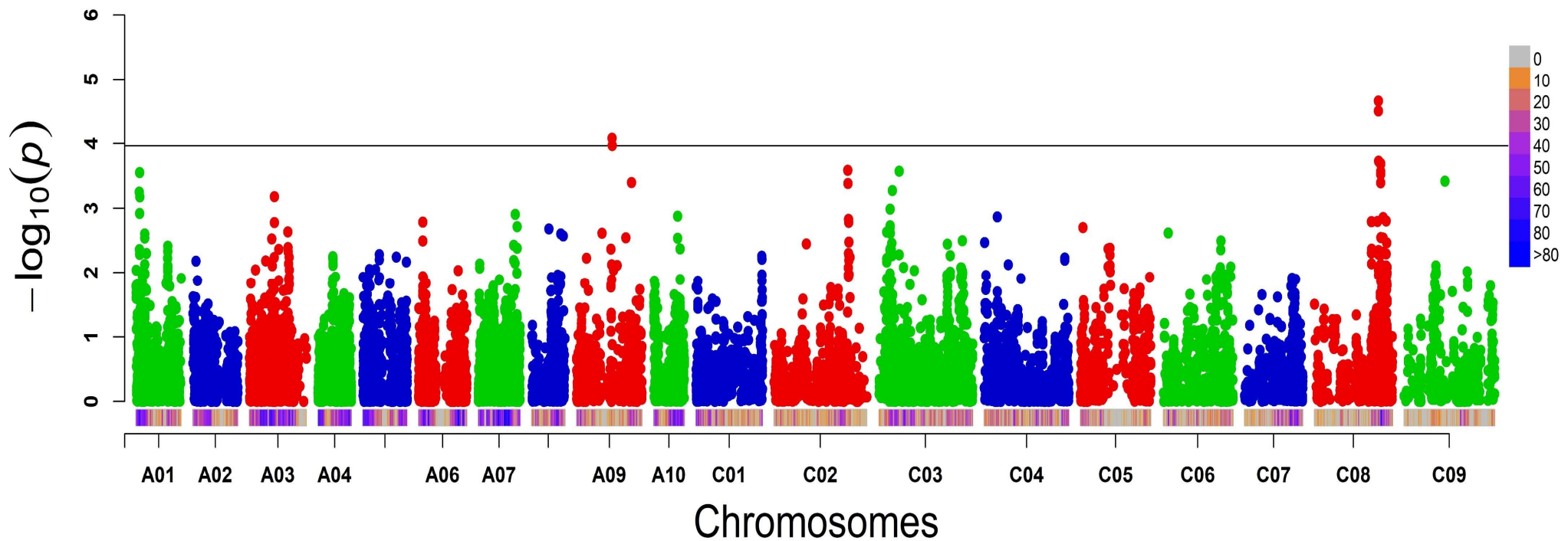
- Lowest mean of squared difference (MSD) values
- Quantile-quantile (QQ) plots

Significant SNP selection

- p -value threshold by Li & Ji (2005) method

Genome-wide association mapping – Field – at maturity

Manhattan Plot (significant markers)



Genome-wide association mapping – **Field** – at maturity

SNP	Chr	Pos	P.value	MAF	Effect
Bn-scaff_16197_1-p3082648	C08	42511399	0.00002	0.31	-0.38
Bn-scaff_16197_1-p3089528	C08	42508303	0.00003	0.32	-0.38
Bn-A09-p18258259	A09	23464377	0.00008	0.07	0.60
Bn-A07-p5737991	A09	23498332	0.00011	0.09	0.53
Bn-A07-p5738022	A09	23498364	0.00014	0.09	0.53

Chromosomes:

A01 (3)

A03 (1)

A09 (4)

C02 (2)

C03 (2)

C08 (9)

C09 (1)

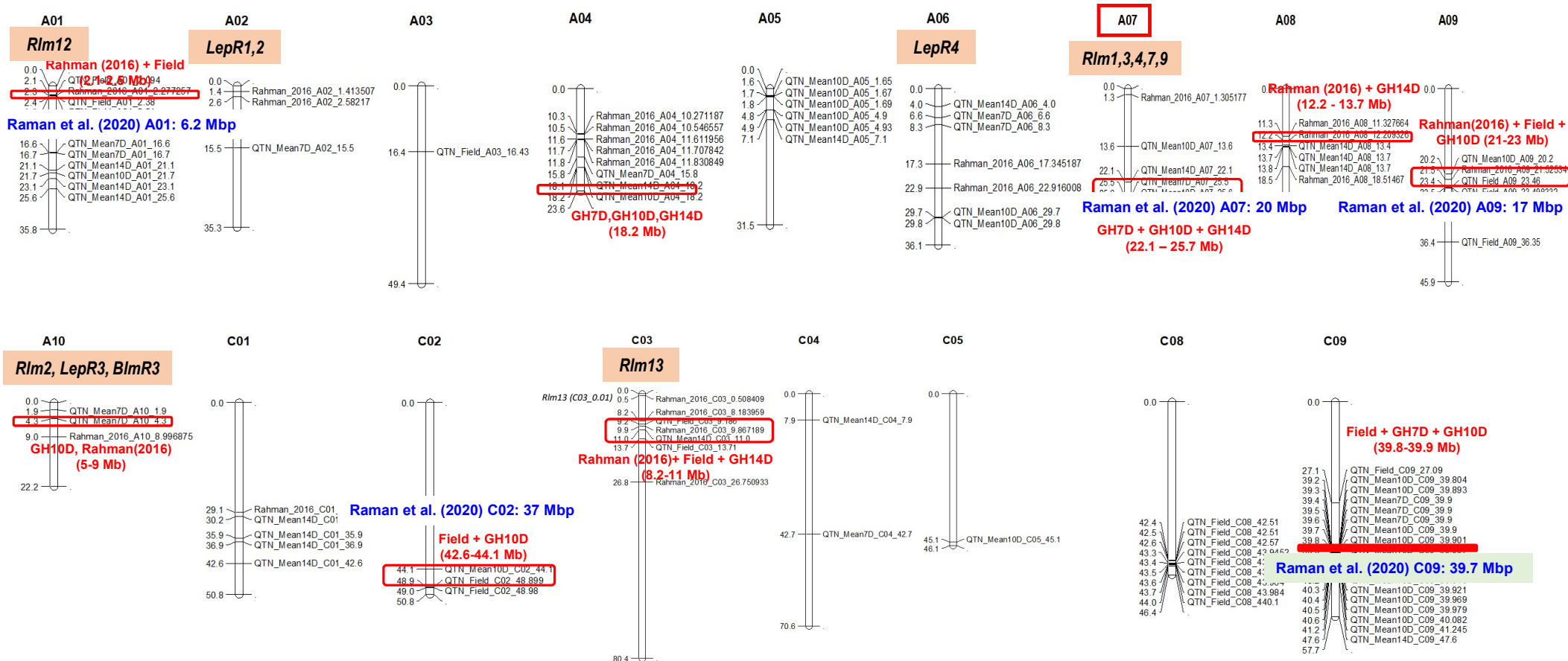
Field study: 22 Significant SNP's

Bn-scaff_16197_1-p2605576	C08	43984498	0.00020	0.22	0.36
Bn-scaff_16197_1-p2626407	C08	43963779	0.00021	0.22	0.36
Bn-scaff_16139_1-p643296	C02	48899814	0.00026	0.39	-0.31
Bn-scaff_21312_1-p1332032	C03	13711606	0.00027	0.30	-0.37
Bn-scaff_16197_1-p2636039	C08	43954285	0.00027	0.22	0.34
Bn-A01-p2329932	A01	2381239	0.00028	0.22	0.38
Bn-scaff_16197_1-p2581072	C08	44005117	0.00030	0.26	0.34
Bn-A09-p17657662	C09	27087500	0.00038	0.07	0.51
Bn-scaff_16197_1-p2582618	A09	36351489	0.00040	0.28	0.31
Bn-scaff_16197_1-p2645092	C08	43945228	0.00040	0.22	0.34
Bn-scaff_16139_1-p721705	C02	48983108	0.00042	0.40	-0.29
Bn-scaff_22728_1-p1393497	C03	9186207	0.00053	0.43	0.27
Bn-A01-p2042209	A01	2099441	0.00056	0.29	0.31
Bn-A03-p15906810	A03	16427867	0.00066	0.22	0.33
Bn-A01-p2453758	A01	2509071	0.00067	0.22	0.35

Significant markers – Greenhouse + FIELD + Rahman et al. (2016)

Markers used (92): Greenhouse (3+28+17), FIELD (22), Rahman et al. (2016) (22)

7 DAI = 3 sig markers; 10 DAI = 28 sig markers; 14 DAI = 17 sig markers





**Training Population
(~200 lines)**



Repeated Phenotyping



TTCGTACGGAT
TTCGTACTGAA
TTCGTACGGAT
TTCGTACTGAA
TTCGTACTGAA
TTCGTACTGAA

**Genotyping
~200 lines**



**Identify Genomic
Prediction Model**

- RR-BLUP
- BayesA
- BayesB
- BayesC
- LASSO
- GBLUP
- ...
- SVM
- RKHS
- RF
- RBFNN
- ...

Genomic Prediction



**Breeding or unknown
Population (~1000)**



No Phenotyping

TTCGTACGGAT
TTCGTACTGAA
TTCGTACTGAA
TTCGTACTGAA

**Genotyping
~1000 lines**



**Best lines (eg. 100)
selection based on
the best Pred model**



Genomic Prediction (GH study) – Blackleg disease in Canola

Method: BWGS pipeline in R using CCAST

Data (same phenotypic data at 10 DPI)

- Rapeseed/canola germplasm accessions = 316
- Design = RCBD with 3 technical reps, 5 biological reps (i.e. 5 seedlings/ technical rep)
- Environments = Repeated GH study (2024-2025)

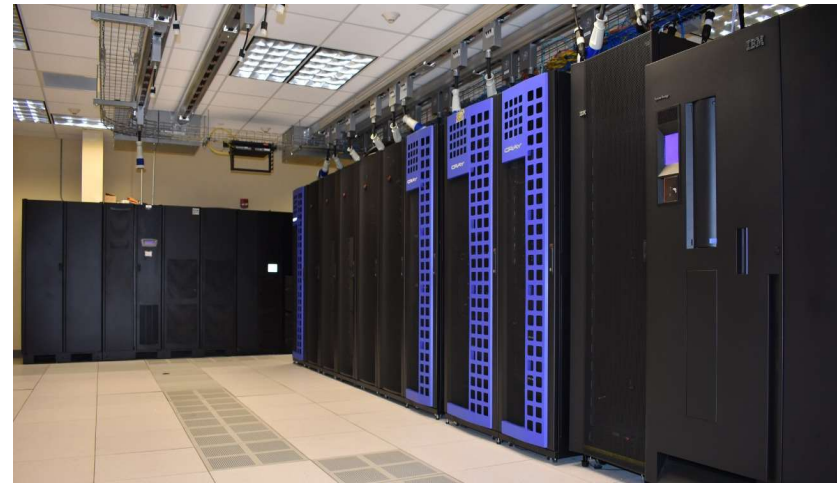
Markers used: 16,423 SNP

Models (14)

- GBLUP, RRBLUP, EGBLUP, LASSO, EN, BRR, BL, BA, BB, BC, RKHS, MKRKHS, SVM, RF

Predictive ability: The correlation between the observed phenotype and genomic estimated breeding value (GEBV) of validation set.

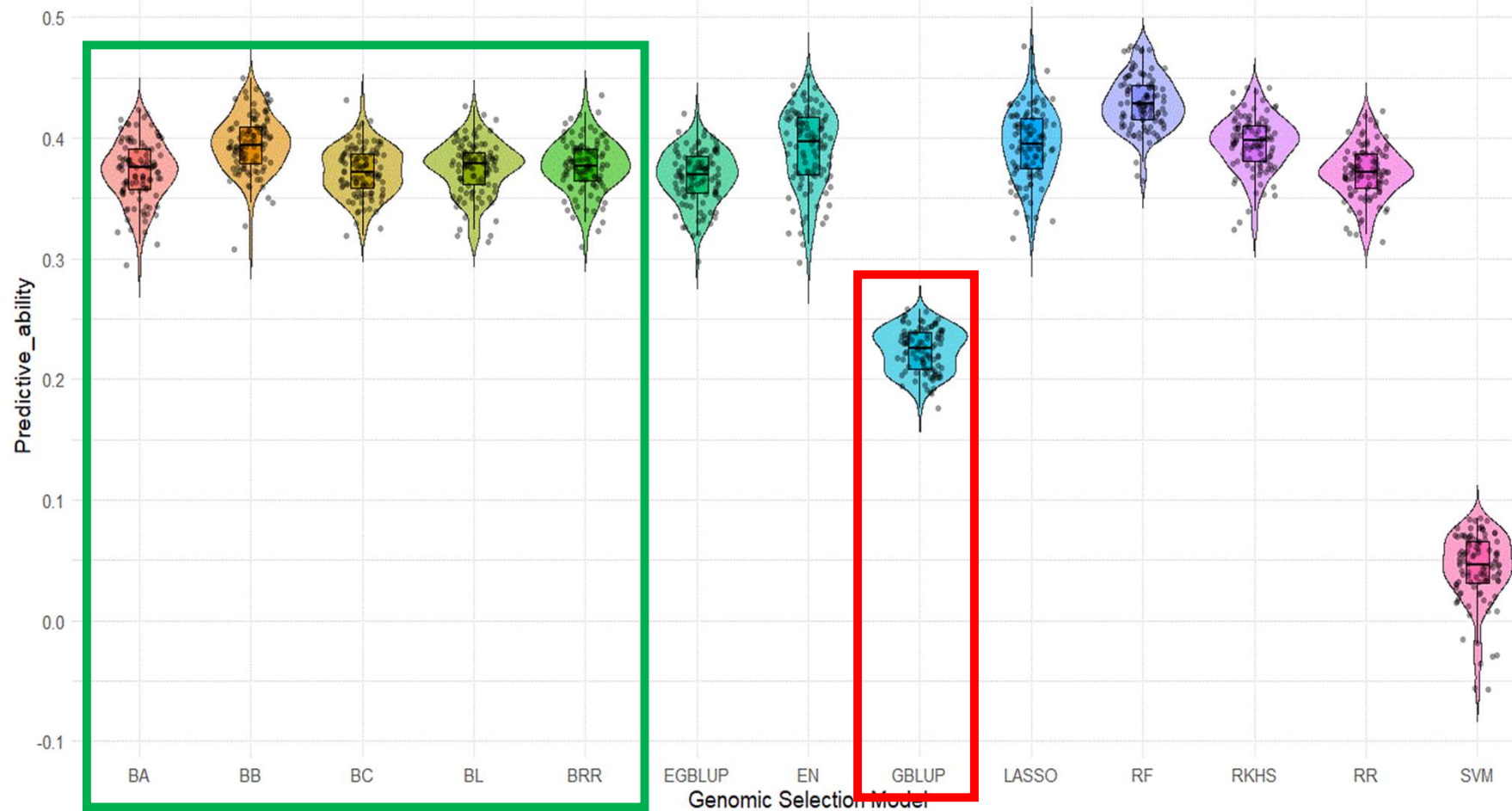
Model efficiency: Five fold cross-validation schemes (80% training, 20% validation), 100 iteration.



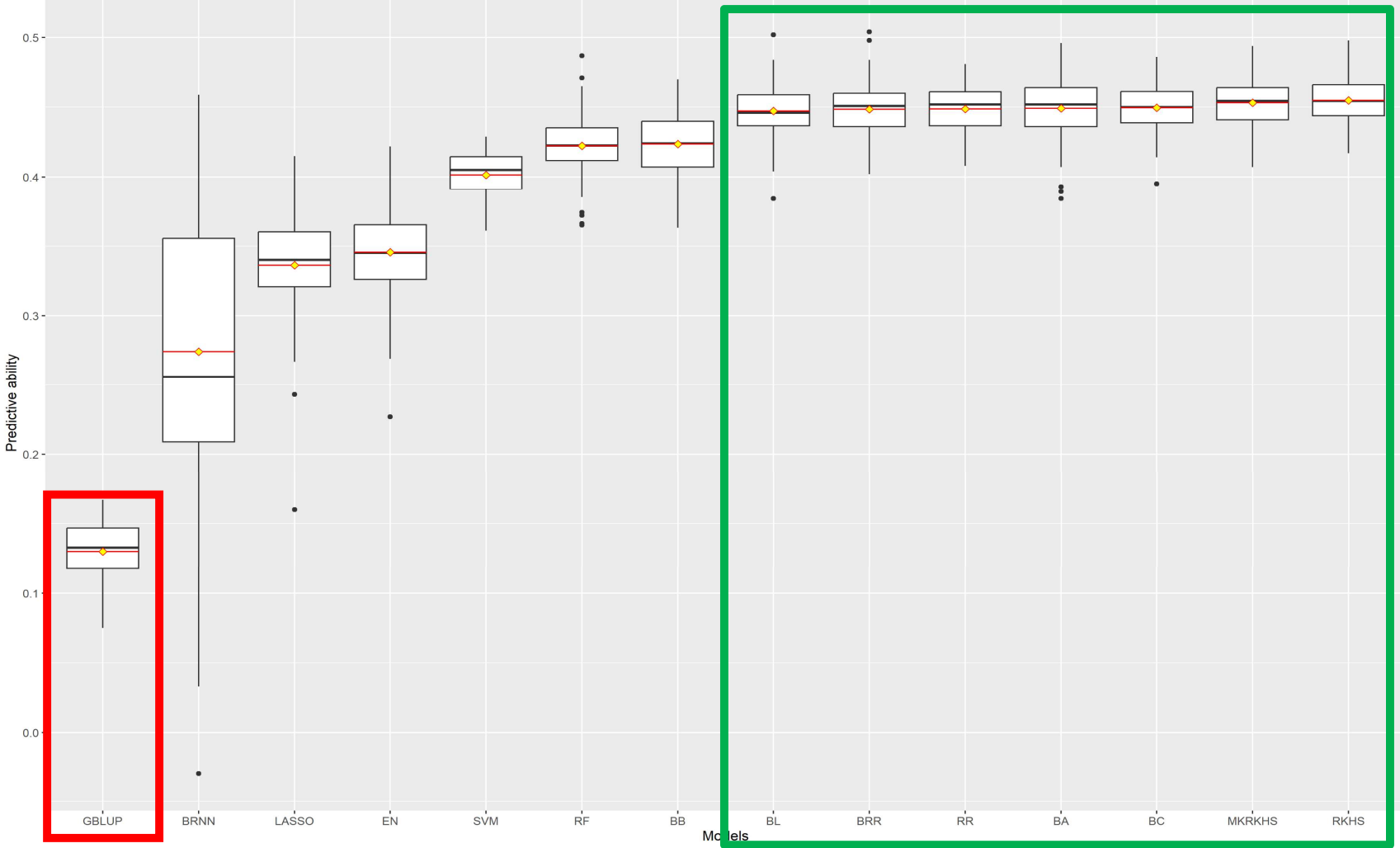
CCAST Platform

CCAST (Center for Computationally Assisted Science and Technology) is the largest academic supercomputing facility in the state of North Dakota, with more than 12,000 CPU cores and 70 GPUs

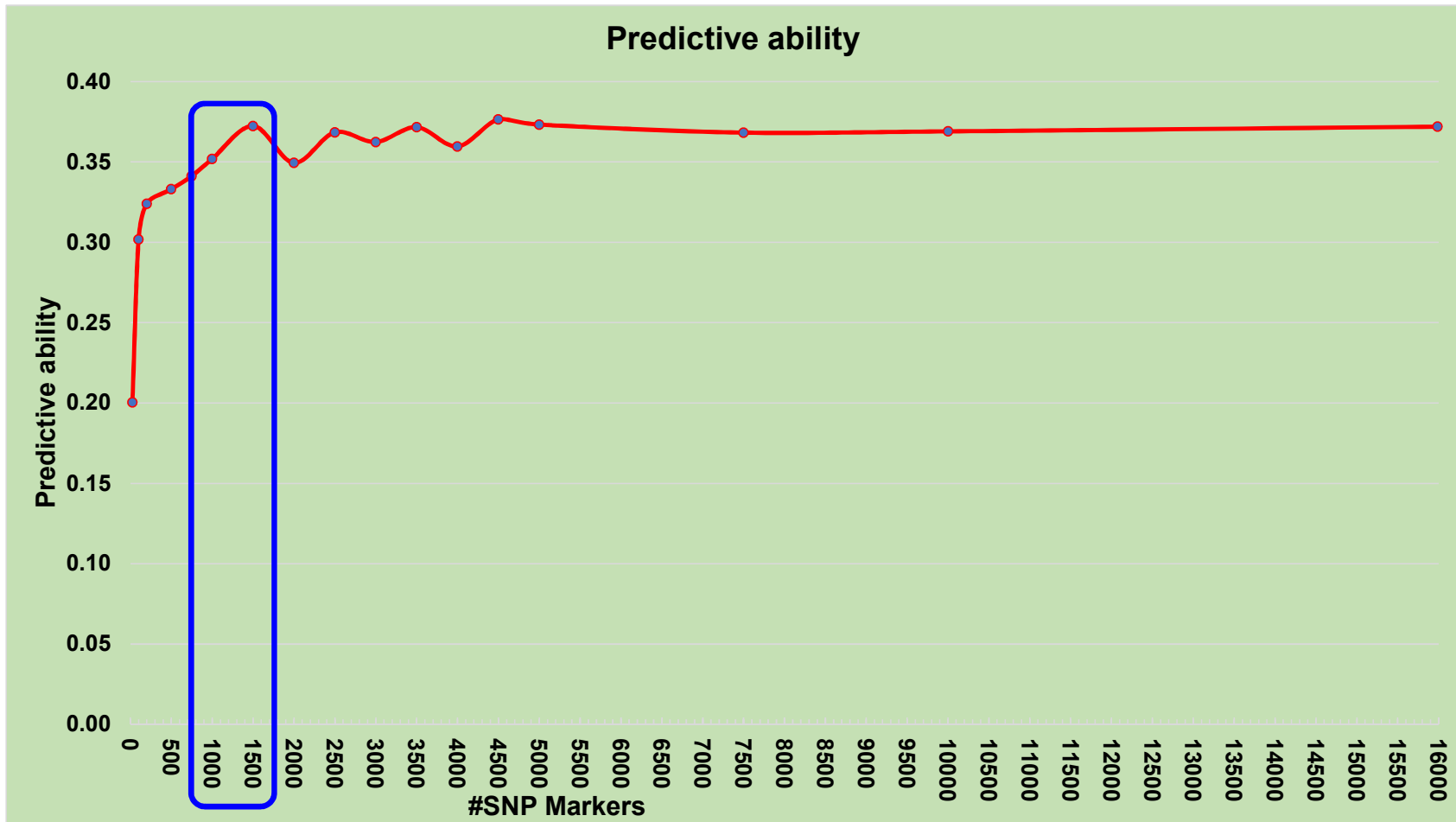
Genomic Prediction based on All Markers – Greenhouse study



Genomic Prediction based on All Markers – Field study



Markers required for genomic prediction analysis



Hoque et al. (2024) Scientific Report
Hossain et al. (2025) Euphytica

Verticillium Stripe disease screening

Screening methods

(i) Petiole inoculation

(ii) Root-dip inoculation (Cui et al., 2023) [Stephen Strelkov's Lab]

(iii) Grain inoculation (Hwang et al., 2014)

In Progress

Acknowledgements (Breeding team)



Kreg Kercher
Research Specialists

Current Graduate Students



Zahangir Alam
(Ph.D. from 2021)



Mohammad Jony
(Ph.D. from 2022)



Md. Mehedi Hafiz
(Ph.D. from 2023)



Shahab Ahm
(Ph.D. from 2021)



Fahad Hasan
(M.S. from 2023)



Jannatun Naher
(M.S. from 2023)



Ashok Mandal
(Ph.D. from 2022)



Aizaz Ali
(Ph.D. from 2023)



Farhan Tanvir
(M.S. from 2023)

Role: Academic Supervisor

Role: Academic Advisor

Acknowledgements (Funding)



**National
Sclerotinia
Initiative**



**AFRI Foundational and Applied
Science Program**



North Central Region of Canola Research

State Board of Agricultural Research and Education
SBARE
Who we are and what we do



Thank you



NDSU NORTH DAKOTA
STATE UNIVERSITY