Association mapping of Sclerotinia stalk rot resistance in domesticated sunflower plant introductions

Zahirul Talukder[1], Brent Hulke[2], Lili Qi[2], and Thomas Gulya[2]
[1]Department of Plant Sciences, NDSU
[2]USDA-ARS Northern Crop Science Laboratory
Fargo, North Dakota

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Background

- *Sclerotinia sclerotiorum* causes two serious diseases in Sunflower
  - Stalk rot and Head rot, genetics of resistance is different
  - Resistance is polygenic, no major resistant gene is known
  - **Resistance breeding relies on incorporating genetic factors from various partially-tolerant breeding lines**

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Quantitative disease resistance mapping

- Two approaches
  - Linkage analysis based QTL mapping, traditional tool
    - need to develop bi-parental mapping population by controlled crosses with known ancestry
    - few opportunities for recombination, resulting in low mapping resolution
    - QTLs that differ between two parental lines can only be detected in each mapping population

Quantitative disease resistance mapping

- Linkage disequilibrium (LD) based Association Mapping, new tool
  - no need to develop mapping population
  - exploits historical and evolutionary recombination events at the population level
  - mapping resolution is higher than bi-parental mapping population of the same size
  - suffer from risk of incurring false positives due to population structure and kinship among individuals

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Association Mapping

- Two approaches
  - Genome-Wide Association Mapping
    - whole genome may be scanned to identify markers that are associated with a particular phenotype
  - Candidate-Gene Association Mapping
    - alleles at a few selected functional candidate genes thought to be involved in controlling the trait of interest may be tested for association
- Candidate-Gene Association Mapping study is more hypothesis-driven than a Genome-Wide study

Candidate-Gene AM of Sclerotinia stalk rot resistance in sunflower

Methods

- AM population = 260 sunflower lines
  - 249 domesticated plant introductions (PIs)
  - 11 elite USDA inbred lines
  - 2 hybrid checks, susceptible (Car 270) & resistant (Croplan 305)
  - examined for stalk rot resistance in 2008 and 2009 in multi-location replicated trials
  - 52 genotypes with best resistant response, and 52 most susceptible genotypes were sequenced for selected candidate genes

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

- Six Candidate genes (*Arabidopsis thaliana* defense genes) :
  - *ABI1* (ABA Insensitive 1), and *ABI2* (ABA Insensitive 2) -involved in abscisic acid (ABA) signal transduction
  - *EIN2* (Ethylene Insensitive 2) - central regulator of ethylene signaling
  - *LACS2* (Long-chain Acyl-CoA Synthetase 2) - involved in cutin biosynthesis pathway
  - *DET3* (De-Etiolated 3) - involved in oxalic acid signaling, and
  - *COI1* (Coronatine Insensitive 1) jasmonate receptor
- Primer design:
  - Nucleotide sequences of the candidate genes were used to BLAST search against the NCBI EST database for sunflower EST sequences
  - Sunflower EST sequences with high score and e-value were then selected for each gene, and searched for contig assembly sequences in the Compositae Genome Project database (http://cgpdb.ucdavis.edu/)
  - The contig sequences were reverse BLAST against GenBank to confirm the gene identity
  - Multiple overlapping primer pairs were designed from contig sequences using the *Primer3* software

- DNA extraction, PCR amplification & sequencing:
  - DNA extracted from 260 individuals of AM population
  - PCR conditions optimized for each pair of primers
  - Cleaned PCR amplicons sent for sequencing to Genomics and Bioinformatics Research Unit at USDA-ARS, Stoneville, MS
- Sequence analysis and SNP survey
  - *DnaSP* v.5.1 software and web-based tool *SNiPlay* were used for sequence analysis and SNP survey within the gene sequences
- Population structure and Kinship analysis
  - SNP markers develop by NSA were used for both population structure and kinship analysis. Structure analysis was done with 750 randomly chosen SNP markers using *Structure* v.2.3.3 software and, kinship analysis was performed with 5244 SNPs using *SPAGeDi* v1.3a software
- Linkage disequilibrium and Association mapping
  - *Haploview* v.4.2 and *TASSEL* v.3.0 software were used for LD and association mapping analysis, respectively

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Results

Summary table of sequence analysis for eight candidate gene sequences

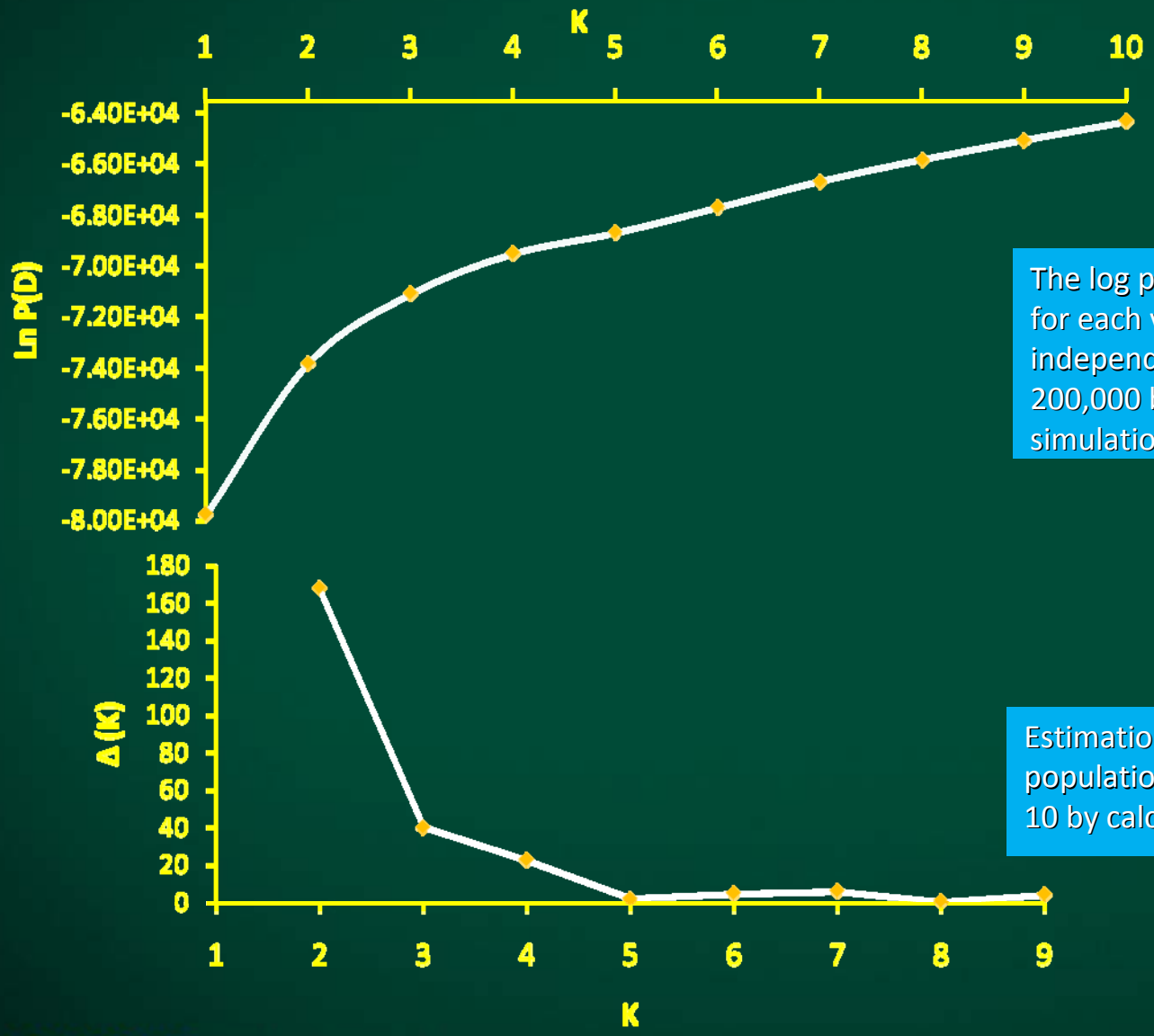
Sequence name	Length (bp)	LG	Map position (cM)	SNP	InDel	Coding region			Non-coding region		
						Length	SNP†	Freq.	Length	SNP	Freq.
ABI1N1	279	-	-	12	4	171	$1_{(0)}$	1/171	108	11	1/10
ABI2N2	189	10	44.9	5	0	189	$5_{(4)}$	1/38	0	-	-
COI24	668	14	50.9	23	1	668	$23_{(5)}$	1/29	0	-	-
COI57	660	14	7.0	12	0	660	$12_{(1)}$	1/55	0	-	-
DET1N1	808	16	68.3	61	24	192	$7_{(3)}$	1/27	616	54	1/11
EIN1	483	14	30.1	15	3	403	$13_{(2)}$	1/31	80	2	1/40
EIN2N1	239	-	-	3	0	239	$3_{(0)}$	1/80	0	-	-
LAC1N2	465	12	47.1	18	6	174	$2_{(0)}$	1/87	291	16	1/18
Total	3791			149	38	2696	$66_{(15)}$	1/41	1095	83	1/13

† number of SNP that led to changes in amino acid codons are shown in the parenthesis

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Estimation of number of sub-populations in the CG-AM population



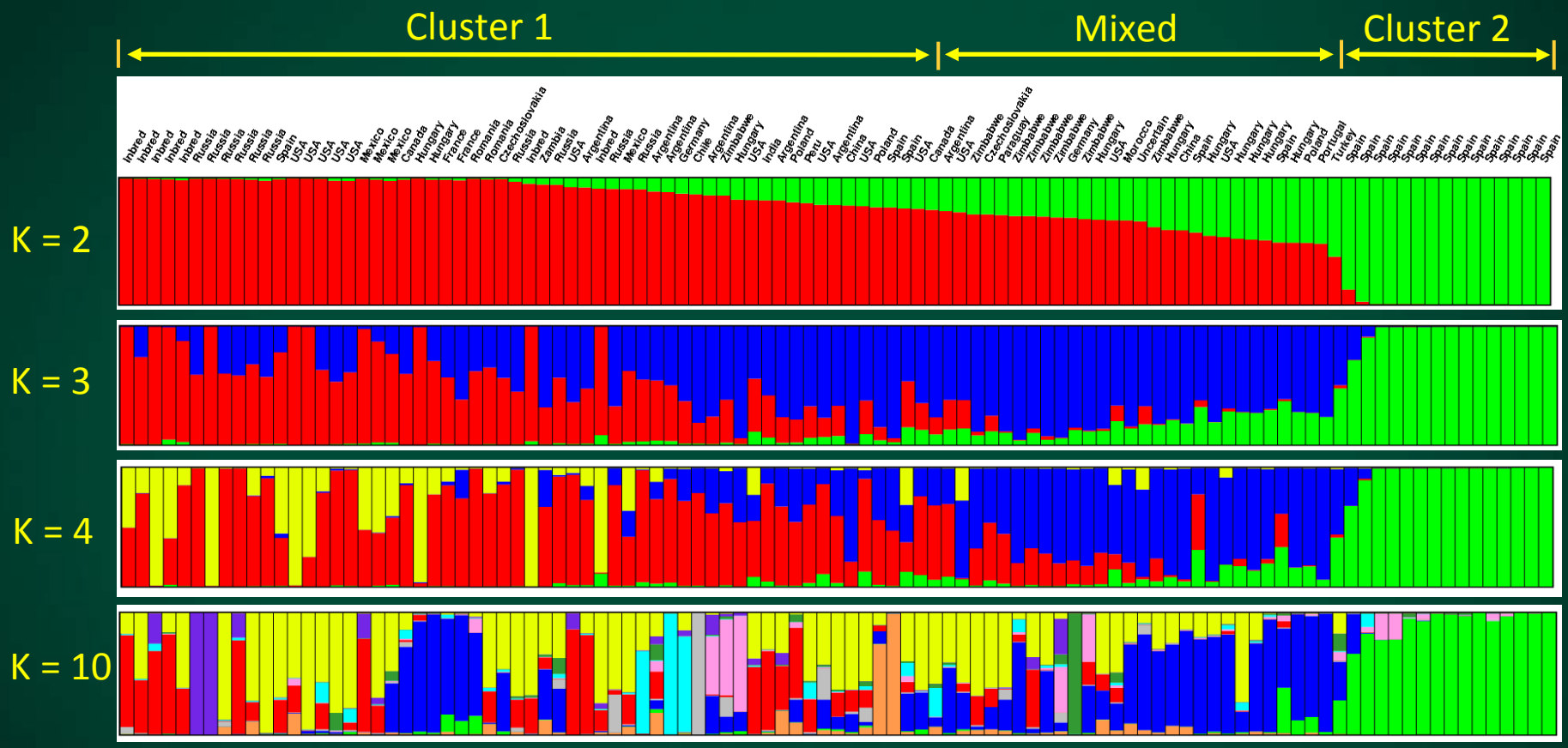
The log probability of data, Ln P(D) for each value of K averaged over 3 independent runs of structure with 200,000 burn-in steps and 200,000 simulation steps.

Estimation of the number of populations for K ranging from 1 to 10 by calculating delta K values.

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Bar plot of the STRUCTURE analysis

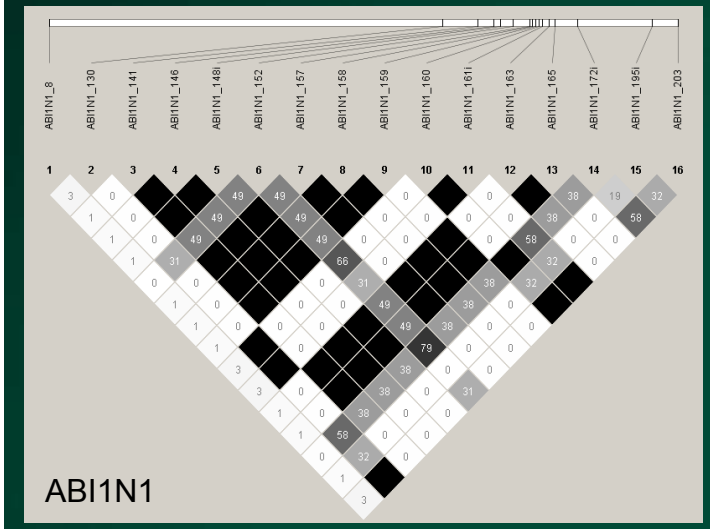


Each of the 104 genotypes is represented by a vertical bar, which is partitioned into K colored segments that represent the individual's estimated membership to the K clusters. The separation of cluster 1 and cluster 2 was done by the threshold for membership coefficient, Q ≥ 0.75.

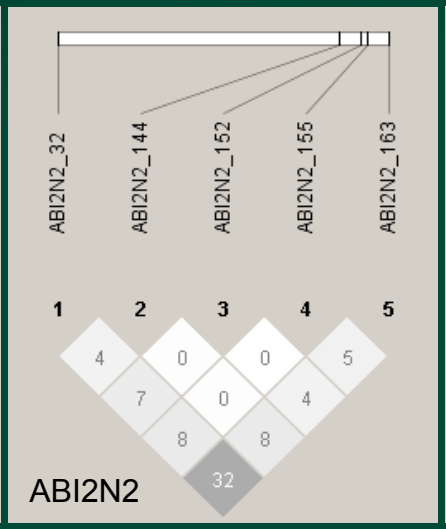
NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

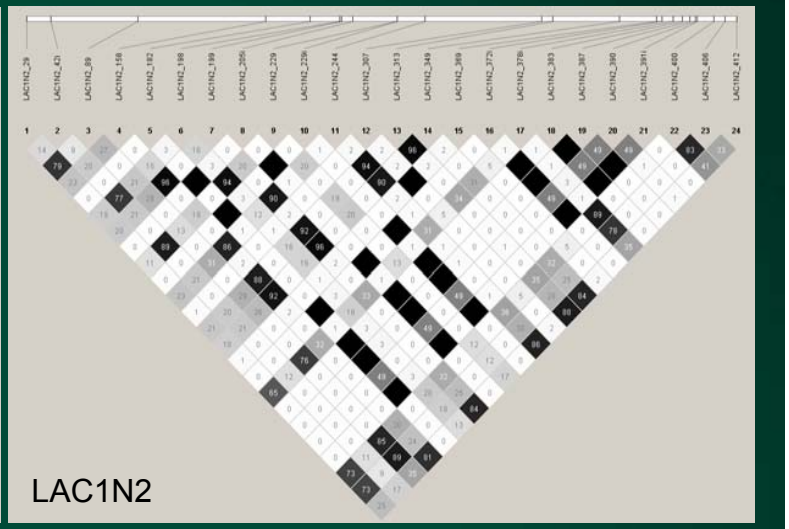
Patterns of linkage disequilibrium (LD) for eight candidate gene sequences

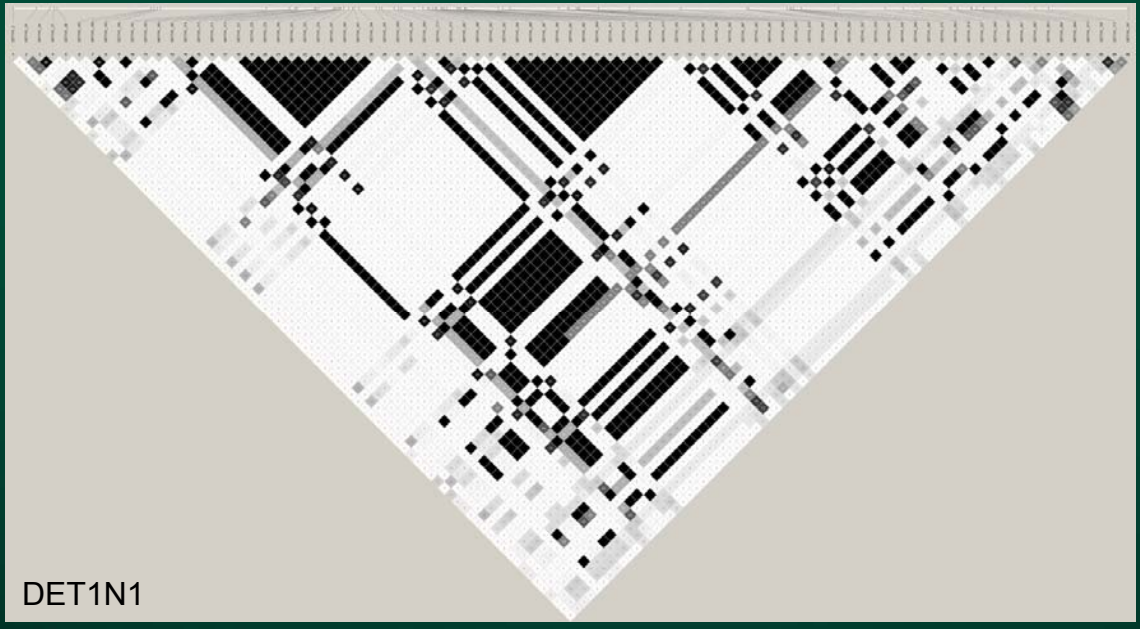


ABI1N1



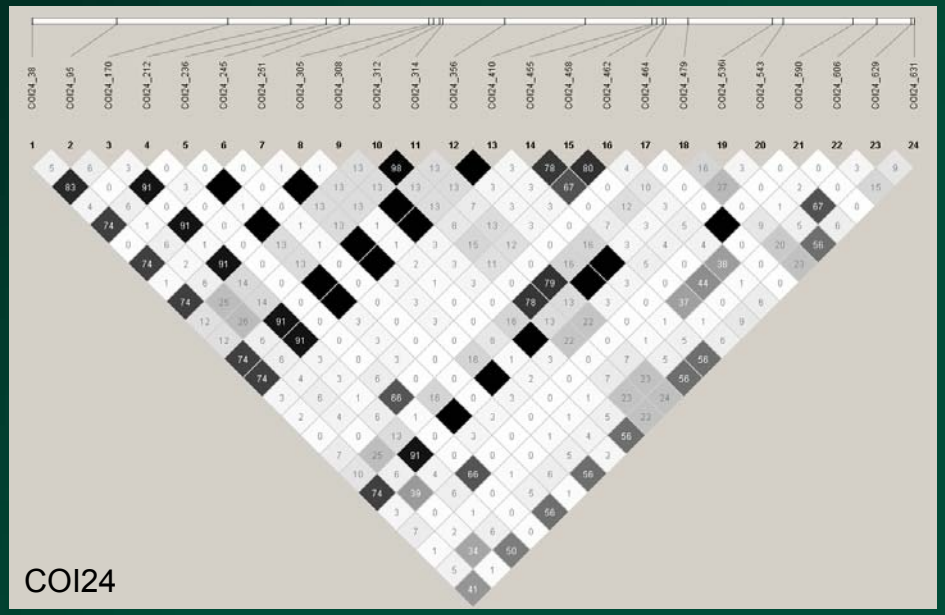
ABI2N2

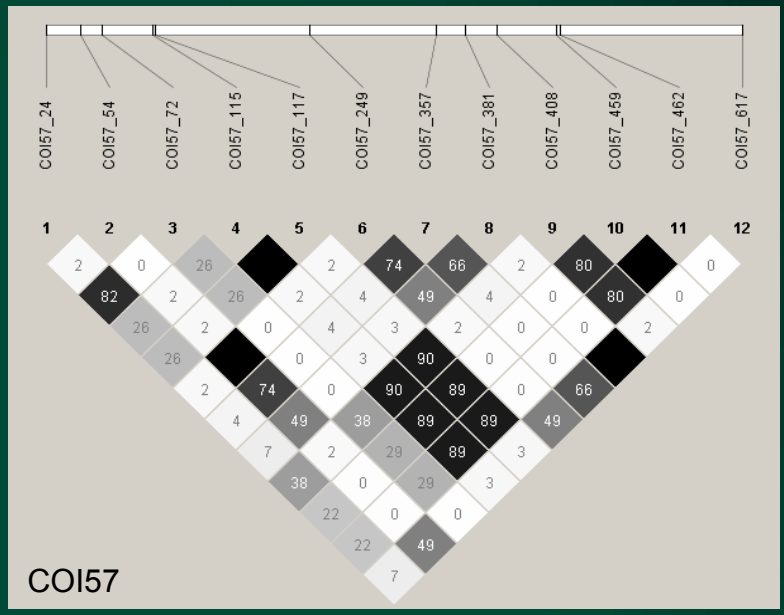

LAC1N2

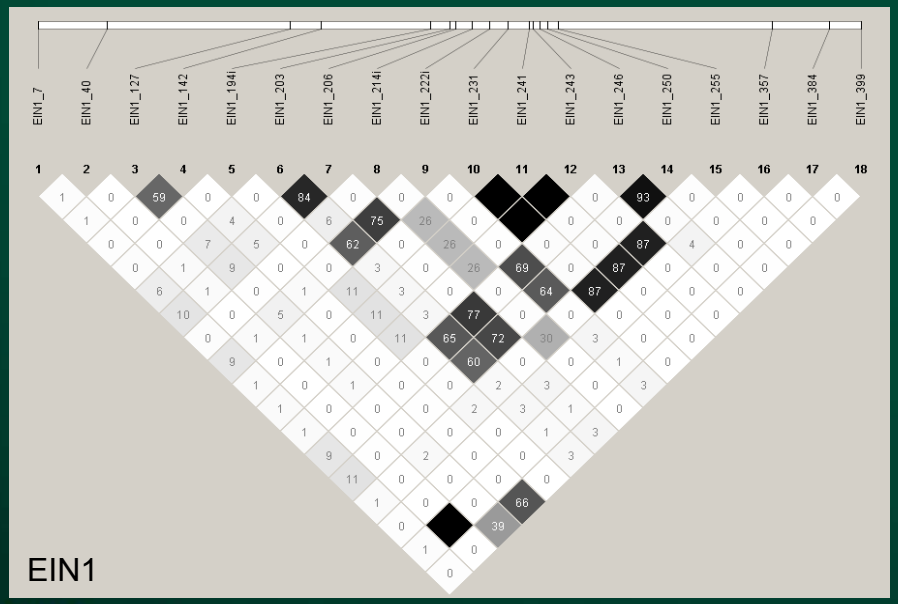

DET1N1

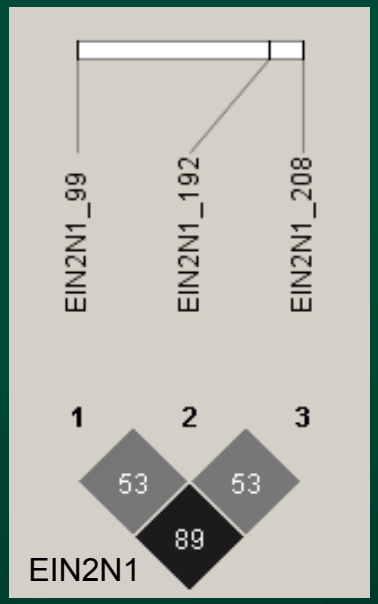
NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS





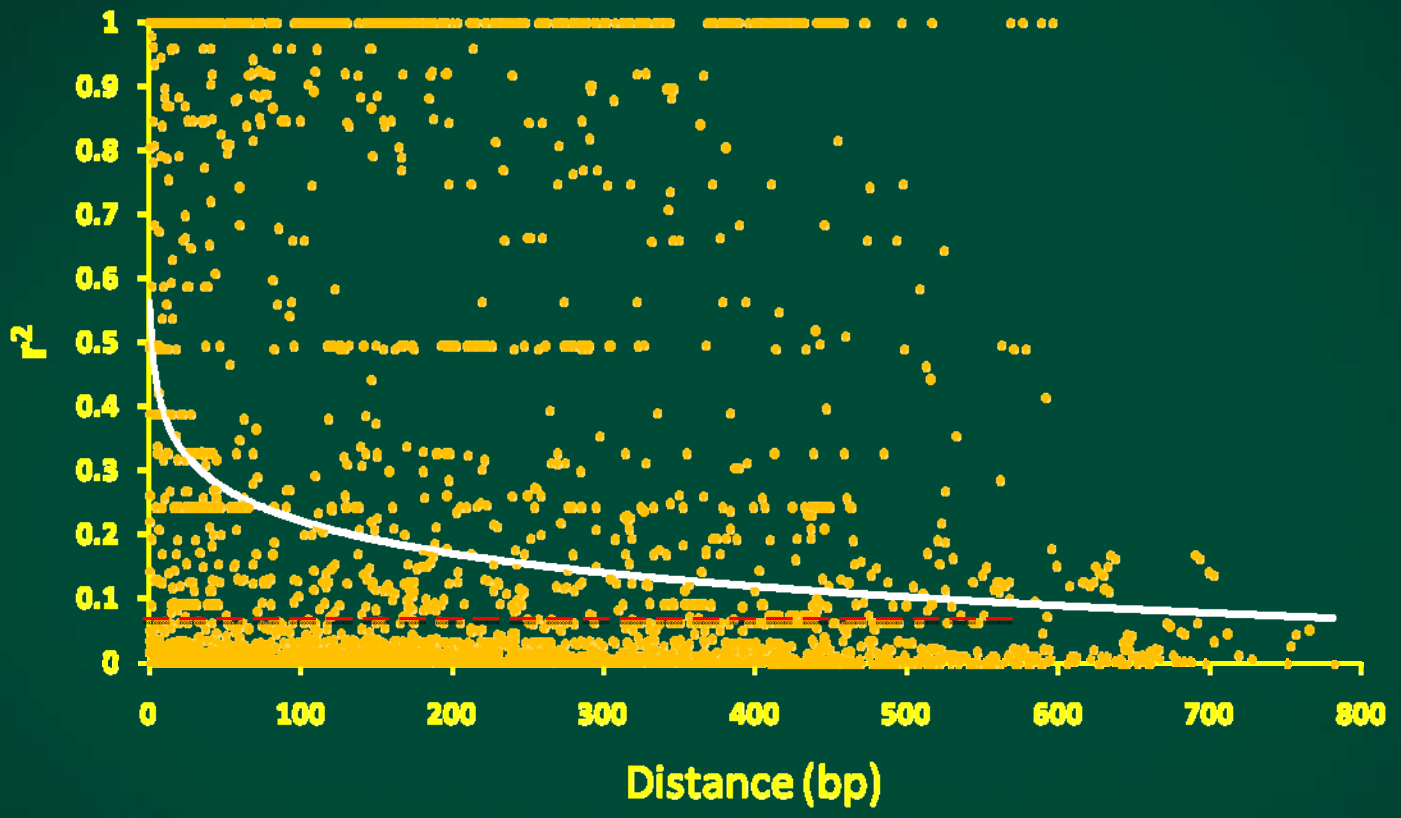




NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Linkage disequilibrium (LD) decay of eight candidate genes



LD decay of eight candidate genes was measured by plotting squired correlations of allele frequencies (r^2) against genetic distance (bp) between pairs of polymorphic sites. Inner fitted trend line is a nonlinear logarithmic regression curve.

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Significant association between candidate gene SNP markers and Stalk rot resistance in sunflower AM population using four statistical models

Gene	Significant SNPs	Tag SNP	3 Environment mean				Davenport			
			GLM	GLM +Q	MLM +K	MLM +K+Q	GLM	GLM +Q	MLM +K	MLM +K+Q
COI1	COI57_459, COI57_462, COI57_115, COI57_117 & COI57_408	COI57_408	***	**	**	**	***	**	*	*
	COI57_24 and COI57_72	COI57_72	***	ns	ns	ns	***	ns	ns	ns
	COI24_236, COI24_251, COI24_308, COI24_356, COI24_410, COI24_536i, COI24_170 & COI24_38	COI24_251	***	ns	*	ns	***	ns	ns	ns
	COI24_631	COI24_631	***	ns	ns	ns	**	ns	ns	ns
	COI24_606	COI24_606	ns	ns	*	ns	ns	ns	ns	ns
	COI24_312 & COI24_314	COI24_312	**	ns	ns	ns	**	ns	ns	ns
ABI1 and ABI2	ABI2N2_32	ABI2N2_32	***	ns	ns	ns	**	ns	*	ns
	ABI2N2_163	ABI2N2_163	**	ns	ns	ns	ns	ns	ns	ns
	ABI2N2_155	ABI2N2_155	**	ns	ns	ns	ns	ns	ns	ns
	ABI1N1_141, ABI1N1_146, ABI1N1_148i, ABI1N1_157, ABI1N1_158, ABI1N1_158, ABI1N1_158, ABI1N1_159, ABI1N1_163, ABI1N1_165 & ABI1N1_172i	ABI1N1_146	ns	ns	ns	ns	*	ns	ns	ns
EIN2	EIN1_246, EIN1_250, EIN1_222i & EIN1_206	EIN1_250	**	**	**	ns	**	ns	ns	ns
	EIN1_127	EIN1_127	ns	ns	*	ns	ns	**	*	*
DET3	DET1N1_614i	DET1N1_614i	**	ns	ns	ns	**	ns	ns	ns
	DET1N1_535,DET1N1_537, DET1N1_217, DET1N1_109, DET1N1_661i,DET1N1_103,DET1N1_428i, DET1N1_181, DET1N1_540i,DET1N1_660, DET1N1_677i, DET1N1_538 & DET1N1_202	DET1N1_217	ns	ns	*	ns	ns	ns	ns	ns
	DET1N1_394i	DET1N1_394i	ns	ns	*	ns	ns	ns	ns	ns
	DET1N1_25	DET1N1_25	ns	ns	ns	ns	***	ns	*	ns
	DET1N1_34, DET1N1_37, DET1N1_107i & DET1N1_105i	DET1N1_37	ns	ns	ns	ns	***	ns	ns	ns

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Progress in Genome-Wide AM of stalk rot and head rot resistance

- Population structure and Kinship analysis of AM population is complete
- Mapping of NSA SNPs in the sunflower linkage map are in progress
- Phenotypic evaluation of AM population
  - Stalk rot is complete (2 yrs x 2 locations)
  - Head rot is in progress (1 yr x 2 locations is done)

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Future Plans

- Genotyping of the remaining lines of the AM population with significant CG-SNPs, and complete CG-AM work
- Complete GW-AM of Stalk rot resistance
- Continue working on association mapping of Head rot resistance and Phomopsis resistance

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS

Acknowledgement

- National Sclerotinia Initiative for funding
- National Sunflower Association for SNP marker development

- The authors thank following individuals for their technical assistance

Angelia Hogness
Leanne Matthiesen
Dana Weiskopf
Megan Ramsett
Nikolay Balbyshev
Arun Jani

NDSU NORTH DAKOTA STATE UNIVERSITY

USDA ARS