

Genetic Mapping And Genomic Prediction For Northern Corn Leaf Blight (*Exserohilum Turcicum* (Pass.) Leonard And Suggs) Resistance

D.C. Balasundara

Department of Genetics and Plant Breeding, College of Agriculture, GVKV, University of Agricultural Sciences, Bangalore, Karnataka

H. C. Lohithaswa (✉ lohitaswa.chandappa@gmail.com)

Department of Genetics and Plant Breeding, College of Agriculture, V.C. Farm, Mandya, (University of Agricultural Sciences, Bangalore), 571 405, Karnataka, India

M. Rahul

Department of Genetics and Plant Breeding, College of Agriculture, V.C. Farm, Mandya, (University of Agricultural Sciences, Bangalore), 571 405, Karnataka, India

R. L. Ravikumar

Department of Biotechnology, College of Agriculture, GVKV, University of Agricultural Sciences, Bangalore, Karnataka

Anand Pandravada

Corteva Agriscience Pvt. Ltd., Kallinayakanahalli, Gauribidanur, Chikkaballapur, 561 213, Karnataka

Bhupendra S. Bhatia

Corteva Agriscience Pvt. Ltd., Kallinayakanahalli, Gauribidanur, Chikkaballapur, 561 213, Karnataka

Research Article

Keywords: *Exserohilum turcicum*, genomic prediction, heterosis, Linkage analysis

Posted Date: June 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-618501/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Northern corn leaf blight (NCLB) of maize caused by *Exserohilum turcicum* is a serious foliar disease. Resistance to NCLB is complexly inherited and the highly significant genotype x environment interaction effect makes selection of resistant genotypes difficult through conventional breeding methods. Hence an attempt was made to identify the genomic regions associated with NCLB resistance and perform genomic selection (GS) in two F_{2:3} populations derived from the crosses CM212 × MAI172 (Population-1) and CM202 × SKV50 (Population-2).

Results: Two populations, each comprising of 366 progenies, were phenotyped at three different locations in the disease screening nurseries. Linkage analysis using 297 polymorphic SNPs in Population-1 and 290 polymorphic SNPs in Population-2 revealed 10 linkage groups spanning 3623.88cM and 4261.92cM with an average distance of 12.40 cM and 14.9 cM, respectively. Location-wise and pooled data across locations indicated that QTL expression was population and environment specific. The genomic prediction accuracies of 0.83 and 0.79 were achieved for NCLB Population 1 and Population 2, respectively. The resistant progenies from both populations were advanced to derive inbred lines and crossed with four different testers in line x tester mating design to test for their combining ability. High overall general combining ability was exhibited by 21 inbred lines. Among crosses 48 % were assigned high overall specific combining ability status. Out of 136 single crosses, seven recorded significant positive standard heterosis over the best check for grain yield. The clustering pattern of inbred lines developed from the two populations revealed high molecular diversity.

Conclusions: In this study, comparatively better genomic prediction accuracies were achieved for NCLB and the worth of F₃ progenies with high genomic predictions was proved by advancing them to derive inbred lines and establishing their higher combining ability for yield and yield related traits.

Background

Maize is an important cereal crop and is cultivated widely throughout the world with far higher yield potential than any other cereals due to its C4 nature with a high rate of photosynthetic activity [1]. In many countries, it serves as an important staple food with major portion of maize produced goes for animal feed and many industrial applications. Higher genetic potential makes it an important component in tropical, subtropical, and temperate climates. Maize is the third most important crop in India after rice and wheat. There is a rapid increase in area under maize due to higher yields and easy cultivation. However, there is a wide gap between global productivity and that of India. This gap is due to the increased incidence of insect pests and diseases in the tropical Asian region. Northern corn leaf blight (NCLB), caused by the pathogen *Setosphaeria turcica* (anamorph *Exserohilum turcicum*), is a very serious disease causing more than 50% yield losses worldwide [2, 3]. NCLB is common in areas with high humidity coupled with moderate temperatures [4, 5]. NCLB can be effectively managed through the use of resistant hybrids [5]. Few genetic resources possess resistance to NCLB and high combining ability making it necessary to develop resistant inbreds under different genetic backgrounds.

Many studies indicated quantitative inheritance [5, 6–14] with resistance loci detected all over the genome [15]. With the development of comprehensive genetic maps, it became possible to dissect the genetic variation underlying quantitative variation, including disease resistance. As in the majority of Quantitative trait loci (QTLs) mapping experiments using biparentally derived populations, the QTLs were defined with relatively less precision with majority of them explaining less than a quarter of the variation and are often not detected in all environments. To capture the genetic variation generated by many small effect QTL Massman et al. [16] suggested employing genomic selection (GS), which is a more comprehensive marker selection strategy to advance the accuracy of prediction [17].

Genomic selection is also a form of marker-assisted selection wherein genetic markers covering the whole genome are used so that quantitative trait loci (QTL) are in linkage disequilibrium with at least one marker. This approach was made possible due to the discovery of large number of single nucleotide polymorphisms (SNP) by several genome sequencing efforts and the availability of potential new methods to efficiently genotype large number of SNPs [18]. In Genomic selection marker data of a genotyped and phenotyped training set are used for building a prediction model [19]. This model is then used to predict genotypic values of individuals for which only marker data are available. In genomic prediction all polymorphisms affecting a trait are modeled making it a highly useful approach in breeding for resistance to a complex trait like NCLB [20]. Keeping these points in view, the objective of this study was to predict the GEBVs of F₃ progenies for NCLB resistance, and to advance the top selected progenies from genomic predictions to derive inbreds with high combining ability.

Results

Meteorological conditions at Mandya, Hassan and Davanagere were ideal for the development of severe NCLB epidemics. The % disease severity during rainy season of 2014 did not follow the normal distribution, and hence data were transformed using arcsine transformation. Significant genetic differences among the F₃ progenies were revealed by the analysis of variance. Bartlett's chi-square test was non-significant which indicated the homogeneity of error mean sum of squares across three locations. Hence data on NCLB across locations were pooled. Pooled analysis of variance of F_{2:3} families indicated that variance due to genotypes and genotype x environment interaction was significant indicating that expression of disease incidence significantly varied among F₃ families and depended on testing environment (Table 1).

Table 1

Pooled Analysis of variance and estimates of mean, range of parents and $F_{2:3}$ progenies derived from crosses CM 212 x MAI 172 and CM 202 x SKV 50 for NCLB incidence during Kharif 2014 across three locations

Source	DF	Mean sum of Squares									
		CM 212 X MAI 172					CM 202 X SKV 50				
Replication	1	359.55*					4830.96**				
Genotype	367	448.40**					421.95**				
Location	2	14114.28**					18497.35**				
Genotype*Location	734	185.22**					232.39**				
Error	1103	70.32					71.78				
Estimates of mean, range of parents and 366 $F_{2:3}$ progenies											
	CM212 (Susceptible)	MAI172 (Resistant)	S v/s R (Pr < t)	F_3 Grand mean	F_3 Range	CM202 (Susceptible)	SKV50 (Resistant)	S v/s R (Pr < t)	F_3 Grand mean	F_3 Range	
Hassan	69.5(56.4)	29.5 (26.5)	0.0144	58.7(50.1)	7–99	84 (66.4)	38.5 (38.3)	0.00078	60.16 (51.0)	9–99	
Mandyā	71.0(57.4)	34 (37.3)	0.0034	50.7(45.3)	20– 81	70.5 (57.0)	26.5 (30.9)	0.00064	52.14 (46.1)	14– 88	
Davanagere	63.0 (52.5)	42.5(40.6)	0.0245	58.6(48.5)	20– 90	78.0 (62.0)	22.2 (28.2)	0.00117	61.8 (52.0)	20– 92	
Pooled	67.8 (55.5)	35.3 (34.8)	0.0238	56.0 (48.5)	7–99	77.5 (61.8)	29.06 (32.5)	0.00053	58.0 (49.7)	9–99	

*Significance at $p = 0.05$ ** Significance at $p = 0.01$

Estimation of mean, range, heritability and genetic advance

Parents differed significantly in their reaction to NCLB (Table 2). The parents MAI 172 and SKV 50 showed resistant reaction at all locations. The susceptible inbreds CM 212 and CM 202 recorded significantly higher disease incidence. The mean disease incidence of 57.92 %, 50.53 % and 57.93 % was observed at Hassan, Mandyā and Davanagere, respectively and it was 55.34 % when pooled across locations. Maximum range of disease incidence (7.00 to 99.00% for Population 1 and 9.00 to 99.00 % for Population 2) was recorded at Hassan followed by Davanagere (20.00 to 90.00% and 20.00 to 92.00 % for Population1 and 2, respectively) and Mandyā (20.00 to 81.00 % and 14.00 to 88.00 % for Population 1 and 2, respectively). The pooled NCLB incidence ranged from 7.00 to 99.00 % for the Population 1 and 9.00 to 99.00 for the Population 2. The frequency distribution pattern of F_3 families for NCLB was negatively skewed and platykurtic at all the three locations and across locations in population 1 and 2 (Table 2, Figs. 1 and 2).

Table 2 Test for normality, skewness, kurtosis and estimates of genetic components for NCLB disease incidence in $F_{2:3}$ populations derived from the crosses CM 212 x MAI 172 and CM 202 x SKV 50

Character	CM 212 x MAI 172				CM 202 x SKV 50			
	Hassan	Mandya	Davanagere	Across locations	Hassan	Mandya	Davanagere	Across locations
Skewness	-1.008	-0.428	-0.367	-0.213	-0.876	-0.355	-0.212	-0.278
Kurtosis	1.859	-0.294	0.499	0.223	1.636	1.240	-0.192	0.621
KS Test	0.111	0.077	0.118	0.048	0.117	0.077	0.113	0.041
Pr > D	<0.010	<0.010	<0.010	0.037	<0.010	<0.010	<0.010	0.146
Mean (%)	57.92	50.34	57.93	55.34	60.06	52.56	62.07	58.23
Range (%)	7-99	20-81	20-90	7-99	9-99	14 - 88	20 - 92	9-99
Genotypic Coefficient of Variation (GCV in %)	27.91	18.73	29.28	35.14	27.52	16.99	29.17	32.14
Phenotypic Coefficient of Variation (PCV in %)	31.41	27.57	29.66	42.89	30.83	27.32	29.61	41.45
Heritability (%)	78.92	46.14	97.49	67.12	79.67	38.70	97.06	60.11
Genetic Advance as <i>per cent</i> Mean (GAM in %)	51.07	22.77	59.57	56.66	50.60	21.78	59.20	51.33

KS test: Kolmogorov Smirnov goodness of fit

The estimates of phenotypic and genotypic coefficient of variation were high in both populations. High heritability and genetic advance which is a measure of genetic gain under selection were observed at all the locations and over locations (Table 2).

Construction of linkage map

Parental polymorphism survey using SNPs

The resistant (MAI 172 and SKV 50) and susceptible inbreds (CM 212 and CM 202) were genotyped using 768 SNP markers. The polymorphism percentage between the two parents was 38.67 % in Population 1 and 37.76 % in Population 2. The polymorphic SNPs were used for genotyping the two F_{2:3} populations (Population 1: CM 212 x MAI 172 and Population 2: CM 202 x SKV 50). Among these markers, only 297 SNP markers were polymorphic in population 1 and 290 SNP markers in population 2.

Genetic linkage mapping

Out of 768 SNP markers, 297 SNP markers in population 1 and 290 SNP markers in population 2 were segregating in a Mendelian fashion and linkage map was constructed using these markers employing ICIM software V3.1. Linkage groups were formed at threshold logarithm of odd (LOD) 2.5 and maximum recombination fraction of 0.30. A total of 297 SNP markers polymorphic for Population 1 were mapped on 10 linkage groups (LGs) spanning a total length of 3623.88 cM (Table 3). The linkage groups LG7 and LG9 consisted of least number of markers (16) and LG5 had maximum number of markers (45). The length of linkage group LG10 was minimum (210.12 cM) and it was maximum in LG5 (531.03 cM) with an average interval distance of 12.20 cM indicating comparatively high-density linkage map. The linkage map for the population 2 consisted of 290 polymorphic SNP markers which spanned 4261.92 cM. Minimum number of markers were mapped in the linkage group LG9 (15 markers) and it was maximum in LG1 (51 markers). The length of linkage groups ranged from 205.1cM (LG9) to 749.04 cM (LG1) with an average inter marker distance of 14.69 cM. The arrangement of markers was comparable with the consensus map constructed for maize. Linkage maps constructed on F_{2:3} mapping populations were used for identification and mapping of QTLs conferring resistance to NCLB disease.

Table 3
QTLs detected for NCLB at individual location and combined across locations using 366 F_{2:3} families (Threshold LOD Score = 2.50)

Location	Chromosome/ Bin	Flanking markers		Position (cM)	Maximum LOD Score	R ² (%)	Genetic effect		Gene action	Donor allele
		Left	Right				Additive	Dominance		
CM 212 x MAI 172										
Hassan	3.14	pza01791-2	pza00309-1	276	6.68	29.94	0.91	-0.65	PD	MAI172
	6.11	PZA00382-17	PHM4503-25	201	2.58	8.44	1.47	-1.44	D	MAI172
	7.03	PHM2691-31	PZA00132-17	47	3.53	30.14	0.78	-1.32	OD	MAI172
	8.04	PZA02955.3	PHM9695-8	63	2.71	8.84	0.10	0.70	OD	MAI172
	9.01	PZA01866.1	PZA00213-19	3	3.08	3.92	0.25	-0.11	PD	MAI172
	10.09	PZA00130.9	PZA00007.1	169	4.46	16.09	0.06	1.78	OD	MAI172
Mandya	3.04	PZA02090-1	PZA01765.1	73	2.68	5.34	0.23	-0.11	PD	MAI172
	4.10	PHM9635-30	PZA01187-1	193	2.67	3.24	-0.13	0.27	OD	CM212
	8.04	PZA02955.3	PHM9695-8	65	2.58	5.88	0.21	-0.13	PD	MAI172
Davanagere	4.09	PZA00125.2	PHM3112-5	163	3.06	3.66	0.11	-0.62	OD	MAI172
	6.03	PHM8327-18	PZA00540.3	48	2.52	6.32	-0.36	-0.01	A	CM212
	6.11	PZA01468.1	pza03577-1	214	3.79	4.53	0.13	0.59	OD	MAI172
Across locations	1.01	PHM2130-29	PZA00447.6	10	2.64	3.63	0.04	0.54	OD	MAI172
	6.03	PHM8327-18	PZA00540.3	48	4.86	11.08	-0.40	0.11	PD	CM212
	6.01	PZA01527.1	PZA03090-31	11	2.98	3.69	0.51	-0.15	PD	MAI172
CM 202 x SKV 50										
Hassan	5.12	PZA00472-2	PZA02862.10	240	4.30	14.45	0.23	-0.80	OD	SKV-50
	8.05	PHM4968-10	PHM7898-10	86	2.55	27.37	0.13	1.86	OD	SKV-50
	10.07	PHM11946-17	PHM4905-6	131	3.23	15.88	0.11	1.94	OD	SKV-50
Mandya	9.10	PHM14104-23	PHM2749-10	199	2.62	4.73	-0.20	0.01	A	CM202
Davanagere	4.07	PHM9635-30	PHM15427-11	138	3.84	4.08	0.10	-0.80	OD	SKV-50
	5.25	PHM1506-23	PHM3512-186	492	4.53	5.61	-0.31	0.23	PD	CM202
	10.03	PHM11226-13	PHM1576-25	51	3.25	3.52	0.23	0.17	PD	SKV-50
Across locations	3.12	PZA00827.1	PZA00817-2	232	2.86	6.55	-0.24	0.19	D	CM202

Nearest markers: Position of QTL peak as indicated by left and right flanking marker;

Position: Position of QTL peak as indicated by cumulative distance from the end of the short arm;

Max LOD: Likelihood of odds (LOD) scores = Likelihood ratios (LR)/4.6052. Critical thresholds of QTL were defined at LOD 2.5;

R²: Percentage of the phenotypic variance explained by genotype at Max LOD peak; Genetic effect: additive and dominance effect at QTL peak, Gene action was established by |d|/|a|: A (additive): 0 to 0.20, PD (partial dominance): 0.21 to 0.80, D (dominance): 0.81 to 1.20, and OD (over dominance) > 1.20;

Donor NCLBR allele detected by +/- of additive with reference to MAI 172. Positive values showed resistance alleles came from MAI172, negative values showed resistance alleles came from CM212.

Location	Chromosome/ Bin	Flanking markers		Position (cM)	Maximum LOD Score	R ² (%)	Genetic effect		Gene action	Donor allele
		Left	Right				Additive	Dominance		
8.15	PHM765-24	PHM3337-23	286	5.15	10.85	0.07	0.73	OD	SKV-50	
Nearest markers: Position of QTL peak as indicated by left and right flanking marker;										
Position: Position of QTL peak as indicated by cumulative distance from the end of the short arm;										
Max LOD: Likelihood of odds (LOD) scores = Likelihood ratios (LR)/4.6052. Critical thresholds of QTL were defined at LOD 2.5;										
R ² : Percentage of the phenotypic variance explained by genotype at Max LOD peak; Genetic effect: additive and dominance effect at QTL peak, Gene action was established by d / a : A (additive): 0 to 0.20, PD (partial dominance): 0.21 to 0.80, D (dominance): 0.81 to 1.20, and OD (over dominance) > 1.20;										
Donor NCLBR allele detected by +/- of additive with reference to MAI 172. Positive values showed resistance alleles came from MAI172, negative values showed resistance alleles came from CM212.										

QTL analysis

The QTLs were detected using the disease incidence data from three locations during rainy season of 2014 and pooled across locations.

In population 1, six QTL positions were identified for NCLB resistance at Hassan location (Table 3, Fig. 3). The QTLs on chromosome 7, 3, 10, 8 and 6 were major which explained 30.14, 22.94, 16.09, 8.84 and 8.44 % phenotypic variance whereas, QTL on chromosome 9 was minor with 3.92 % phenotypic variance. The favorable alleles for these QTLs were contributed by the resistant parent MAI 172. Three QTLs were detected on chromosome 3, 4 and 8 at Mandya which explained 5.34, 3.24, and 5.88 % phenotypic variation. Out of three QTLs found at Davanagere, the QTL located on chromosome 4 explained 3.66 % phenotypic variation and the other two QTLs were located on chromosome 6 which explained 6.32 and 4.53 % phenotypic variation, respectively. In combined QTL analysis three QTLs were identified with the QTL located on chromosome 1 explained 3.63 % phenotypic variation. The remaining QTL were located on chromosome 6 explaining 11.08 and 3.69 % phenotypic variation.

In population 2, three QTL positions were identified at Hassan (Table 3, Fig. 4) and the QTL located on chromosome 8 was the major with phenotypic variation of 27.37 % followed by the second major QTL located on chromosome 5 with 14.45 % phenotypic variation. The third major QTL located on chromosome 10 explained 15.58 % phenotypic variation. Only one QTL position was identified at Mandya on chromosome 9 which explained 4.73 % phenotypic variation. The favorable allele for this QTL was contributed by susceptible parent CM 202 which showed additive gene action. The First QTL located on chromosome 4 explained 4.08 % phenotypic variation with LOD of 3.84 at Davanagere. The second and third QTL located on chromosome 5 and 10 exhibited 5.61 and 3.52 % phenotypic variation. In combined QTL analysis, two QTL positions were identified for NCLB resistance with the QTL located on chromosome 3 explained 6.55 % phenotypic variation. The second QTL located on chromosome 8 explained 10.85 % phenotypic variation.

Genomic selection

Best linear unbiased predictors (BLUPs)

The genotypic data of 297 polymorphic SNPs of the Population 1 and 290 polymorphic markers of Population 2 were used to estimate the marker effects of each SNP using the RR-BLUP procedure. The marker effects estimated for Population 1 ranged from -0.0866 to 0.0857 and from -0.0101 to 0.0105 for Population 2 (Supplementary Table S1). The estimated effects were included in the model to predict the genomic estimated breeding values (GEBVs) of each individual (Supplementary Table S2).

Prediction accuracy and cross validation

The entire 366 F_{2:3} progenies from population 1 were cross validated 100 times in two different ratios of validation set and training set. The first case was with two-fold cross validation, where in each time validation set and training sets were randomly selected equally as 1:1. In this case the prediction accuracy of the GEBVs was 24% (Table 4). The second case was with five-fold cross validation, where the population was divided into five equal parts, where four sets were used as training set against one validation set. In this case the prediction accuracy of the GEBVs was 26%. The same kind of cross validation was done for Population 2 where GEBV prediction accuracies were 29% and 32% for the validation ratios of 1:1 and 1:5, respectively (Table 4, Figs. 5 and 6). The prediction accuracy measured in terms of correlation between the BLUPs and GEBVs was found positive (Table 5). The Population 1 had a correlation of 0.79 and for the Population 2 it was 0.83. However, when predictions were made using top 10 per cent selections correlation between GEBVs and their BLUPs was reduced to 0.35 in Population 1 and 0.50 in Population 2.

Table 4
Prediction accuracies of QTL mapping and Genomic selection for $F_{2:3}$ families derived from crosses CM 212 x MAI 172 and CM 202 x SKV 50

Location	CM 212 x MAI 172					CM 202 x SKV 50				
	Total Variance (%) explained by additive QTLs			GS Accuracy (%)		Total Variance (%) explained by additive QTLs			GS Accuracy (%)	
	PVE	GVE	Heritability	Two-fold cross validation	Five-fold cross validation	PVE	GVE	Heritability	Two-fold cross validation	Five-fold cross validation
Davanagere	6.56	10.25	0.97	-	-	18.95	37.16	0.98	-	-
Hassan	4.65	7.27	0.82	-	-	2.82	5.52	0.82	-	-
Mandya	3.23	5.05	0.45	-	-	2.87	5.63	0.47	-	-
Across location	6.34	9.90	0.64	24.0	26.0	4.66	9.13	0.51	29.0	32.0

PVE: Phenotypic Variance Explained; GVE: Genotypic Variance Explained

Table 5 Correlation between BLUEs and GEBVs estimated for NCLB incidence for all the individuals from Population 1 (CM 212 x MAI 172) and Population 2 (CM 202 x SKV 50) (**A**- All the individuals and **B**- 10 % individuals selected based on GEBVs)

A.

Population 1		Population 2	
BLUEs	GEBVs	BLUEs	GEBVs
BLUEs	1	-	1
GEBVs	0.79	1	0.83

B.

Population 1		Population 2	
BLUEs	GEBV	BLUEs	GEBVs
BLUEs	1	-	1
GEBVs	0.35	1	0.50

Analysis of variance of combining ability of inbreds developed from NCLB resistant F_3 progenies

The analysis of variance for combining ability with respect to 12 quantitative traits indicated that the crosses exhibited high level of significance for all the traits (Table 6). The variance due to crosses was further divided into variance due to lines, testers and line \times testers. The variance due to lines was significant for ear height, kernel rows per cob and kernel per rows whereas, variance due to testers was significant for days to 50 per cent anthesis, days to 75 per cent dry husk, plant height, ear height, kernel rows per cob, test weight and plot yield. The line \times tester interaction variance was highly significant for all the traits. The GCA/SCA variance ratio was less than unity.

Table 6
Analysis of variance for combining ability in maize

Source of variation	DF	Days to 50% Anthesis	Days to 50% silking	Days to 75% husk	Plant height (cm)	Ear height (cm)	Cob length (cm)
Replication	1	0.83	6.62	2.76	12.22	6.89	1.40
Crosses	135	21.04 **	21.83 ***	28.11**	948.83**	489.28 **	7.59**
Line Effect	33	27.55	28.97	32.94	820.78	607.21 *	8.30
Tester Effect	3	7.31	7.51	181.68***	8949.92***	2906.91 ***	6.88
Line × Tester Effect	99	19.29 **	19.89 **	21.84**	749.06**	376.71 **	7.37 **
Error	173	3.47	3.59	0.81	22.43	9.97	1.03
Total	347	11.66	12.16	14.05	466.23	242.10	4.02
Source of variation	DF	Ear circumference (cm)	Kernel rows per cob	Kernels per row	Test weight (g)	Shelling %	Plot yield (Kg)
Replication	1.00	0.61	0.004	7.01	0.22	1.31	0.004
Crosses	135	2.20 **	4.44**	64.58 **	25.33**	35.94 **	0.50 **
Line Effect	33	2.35	7.29 **	107.64**	24.92	44.02	0.52
Tester Effect	3.00	5.19	11.01 *	10.95	164.59**	15.01	2.43 **
Line × Tester Effect	99	2.06 **	3.30 **	51.84 **	21.24**	33.88 **	0.44 **
Error	173	0.52	0.56	6.46	0.09	3.86	0.005
Total	347	1.29	2.31	30.87	12.72	18.83	0.24

* Significant at P = 0.05, ** Significant at P = 0.01

Identification of inbreds showing high general combining ability effects, crosses with high specific combining ability and heterosis

Out of 34 lines 17 showed significant *gca* effects for grain yield (Supplementary Table S3) and 21 lines (61%) were assigned high (H) overall *gca* status (Supplementary Table S4). The line MAI-E2-70 was the best overall general combiner followed by MAI-E9-183 and MAI-E9-39. The testers NAI137 and V351 were the best overall general combiners. High (H) overall *sca* status was exhibited by 65 out of 136 hybrids. The cross MAI-E₂-155 x MAI105 (L x L) was the best cross followed by MAI-E₂-179 x V351 (L x H) and MAI-E₂-72 x MAI105 (H x L) (Supplementary Table S5). Among 136 single crosses, 50 had high (H) overall heterotic status of which MAI-E₉-46 x MAI105 manifested best overall heterotic status followed by MAI-E₉-178 x MAI105 and MAI-E₉-46 x NAI137 (Supplementary Table S6).

The best single cross hybrids based on mean, *sca* effects and standard heterosis for grain yield were MAI-E2-72 x MAI105, MAI-E9-46 x NAI137, MAI-E2-70 x NAI137, MAI-E2-81 x V351, MAI-E9-220 x V351, MAI-E9-211 x NAI137, and CIL1218 x V351 compared to the best commercial check DKC9144 (Table 7).

Table 7
Best single cross hybrids based on mean, SCA effects and standard heterosis for grain yield

Hybrid	Mean	SCA effect	Standard Heterosis	Type of cross
CIL1218 x V351	3.165**	1.131 ***	22.75 **	L x H
MAI-E2-72 x MAI105	3.014 **	1.156 ***	16.89 **	H x L
MAI-E9-211 x NAI137	2.941 **	0.842 ***	14.08 **	H x H
MAI-E2-70 x NAI137	2.931 **	0.587 ***	13.67 **	H x H
MAI-E2-81 x V351	2.843 **	0.552 ***	10.26 **	H x H
MAI-E9-46 x NAI137	2.632 **	0.371 ***	2.06	H x H
MAI-E9-220 x V351	2.610**	0.450 ***	1.22	H x H

* Significant at p = 0.05, ** Significant at p = 0.01 and *** Significant at p = 0.001

Molecular diversity analysis in the resistant inbreds developed

The diversity was assessed in the 45 inbreds and most of the inbred lines could be discriminated with SSR markers. Out of 64 SSR markers tested 35 were found to be polymorphic for the genotypes studied (Supplementary Table S7). The PIC (polymorphism information content) value ranged from 0.11 (*bnilg1327*) to 0.85 (*umc1957*) with mean value of 0.45. Allelic frequency was in the range of 0.52 to 0.93. Markers *umc1139*, *mmc0111* and *umc2269* (0.93) showed high allelic frequency, followed by *umc1080* (0.90) and least was seen in case of *bnilg1449* (0.52). Gene diversity was in the range of 0.24 (*bnilg1942*) to 0.50 (*umc1139*, *bnilg1063*, *umc1085* and *bnilg1887*) with mean value 0.46.

Clustering of inbreds based on SSR marker data

The UPGMA based dendrogram was obtained from the data deduced from the DNA profiles of the samples analysed (Fig. 7). Forty-five maize inbreds were clustered by Jaccard's similarity coefficients which ranged from 0.17 to 0.81 per cent and at similarity coefficient of 0.41, four main groups were observed. The Cluster CI consisted of 21 inbred lines, followed by cluster CII (16), cluster CIII (6) and CIV was with least number of inbred lines (2).

Discussion

The parents differing for NCLB infection were used in the development of mapping populations which was reflected in significant differences among the progenies as revealed by the analysis of variance at each location and over three locations. Maximum number of progenies in both populations showed moderate resistance to NCLB. The disease pressure varied across locations and Hassan location had the highest disease incidence compared to Davanagere and Mandya. The genotype x environment interaction component was significant indicating the influence of environment on the expression of NCLB. The data from three locations were pooled as Bartlett's test proved the homogeneity of error mean sum of squares for NCLB data across locations.

Distribution of the F₃ progenies with respect to NCLB incidence

For a better understanding of the breeding material descriptive statistics is commonly employed. The nature of gene action [21] and number of genes responsible for the trait [22] are denoted by the coefficients of skewness and kurtosis, respectively. Frequency distribution of 366 F₃ progenies from the two crosses CM 212 x MAI 172 and CM 202 x SKV 50 revealed non-normal distribution. Negatively skewed distribution was observed at all locations and in pooled data. Distribution of disease expression was skewed towards susceptible parent CM 212 which denoted the dominance of susceptibility. However, the distribution was made near normal through arcsine transformation of the *per cent* disease incidence data. The near normal distribution of phenotypic data on F_{2:3} populations was reported by several workers [7, 9, 23, 24] with negative skewness for Ht1 gene [25]. The involvement of relatively large number of segregating genes with epistasis was indicated by platykurtic and skewed distribution of F_{2:3} populations in the inheritance of resistance to NCLB [26–29].

Genetic parameters in F_{2:3} mapping populations

Phenotypic co-efficient of variation (PCV) was higher than the genotypic co-efficient of variation (GCV) reflecting on the direct possibility of selecting resistant phenotypes for NCLB [30–32]. High heritability coupled with high genetic advance over mean also indicated the scope for selection of disease resistance genotypes in these populations [7, 10, 30, 33–35]. Mandya location had comparatively lower heritability and genetic advance and this could be due to low disease incidence.

QTL mapping for resistance to NCLB

Genetic linkage map construction

The availability of enormous genomic resources and high level of polymorphism in maize resulted in the development of useful genetic and physical maps [36,37]. The expected Mendelian segregation ratio of 1:2:1 was exhibited by 297 and 290 SNP markers which were then successfully used in the construction of linkage map in Population 1 and Population 2, respectively. The total map distance was 3623.88 cM with 297 markers in Population 1 (CM212 x MAI172) and 4261.92 cM with 290 markers in Population 2 (CM202 x SKV50). The average inter marker distance was 12.4 cM and 14.9 cM in Population 1 and Population 2, respectively. The marker density in our linkage map was 14.8 cM, which was comparatively less dense than the other linkage maps. But for mapping of QTL for important traits, marker density up to 20 cM can be used [38]. Therefore, our linkage map was conveniently used for detecting QTLs for NCLB. The markers used and their positions were confirmed with known positions on IBM2 Neighbor's consensus public linkage map (<http://www.maizegdb.org>).

QTL mapping for resistance to NCLB

The foremost requirement to perform QTL analyses is availability of mapping populations developed from the parents with contrasting reaction and it was clearly accomplished in the present study. Both populations showed near normal distribution pattern of F₃ progenies suggesting their reliability in the identification of QTLs for resistance to NCLB. Since substantial environmental variation was observed, each environment was separately analysed and pooled analysis was done to estimate the overall QTL effects. Most of the QTL were environment specific. The QTLs were identified on chromosomes 1, 3, 4, 6, 7, 8, 9 and 10 and the UMC reference map of maize [36] aid in the comparison of QTL positions across experiments and diverse genetic backgrounds [39, 40]. The major QTLs detected on chromosome 3, 7 and 10 were also reported earlier [8, 10, 23]. They also reported

minor QTLs on chromosome 3, 4, 6, 8 and 9 as in the present study. These NCLB resistant alleles present in resistant parents could be effectively used in MAS to support maize breeding programs [8, 9, 23, 35, 41, 42]. Since the QTL expression was population specific and fine mapping could be an option for clarity on this. The role and nature of QTL x environment interaction is also very crucial as we found most of the QTLs were environment specific and deserve more attention, critical analysis and deeper understanding of molecular basis for the benefit of disease resistance breeding.

Estimation of Genomic Estimated Breeding values (GEBVs) of the selection candidates for the NCLB resistance and prediction of GS accuracies

All methods employed in genomic selection efficiently utilize genome wide markers to predict any trait with accuracy enabling selection on that prediction alone. These breath-taking developments enabled potential acceleration of the breeding cycle and also increase in selection intensity [43, 44]. It is possible to reduce four-year breeding cycle that include three years of field testing, to only four months for growing and crossing a plant. It also helps in the evaluation of thousands of selection candidates without ever taking them out to the field. However, field trials are still a very much part of breeding program using GS, and selection is not based on the phenotypes. Phenotypes are just used to train a prediction model [45] (Heffner et al. 2009). The statistical methods used by GS are relatively new to the plant-breeding community. But it is already clear that GS has the ability to redirect resources and activities in plant breeding programs to accelerate crop improvement. The disease scores recorded at all the three locations were pooled and used in the estimation of the BLUEs. The estimated BLUEs ranged from 1.96 to 6.42 for the Population 1 across locations and 1.66 to 6.68 for the Population 2. For GS, same set of $F_{2:3}$ families used in QTL analysis were used. The genotypic data obtained from the polymorphic SNPs were used in marker effect estimation by RR-BLUP. The marker effects estimated for Population 1 were ranging from -0.0866 to 0.0857 and from 0.0101 to 0.0105 for Population 2. These marker effects were included in the model to predict the genomic estimated breeding values (GEBV) of each selection candidate. The entire 366 $F_{2:3}$ populations from Population 1 were cross validated 100 times in two different ratios of validation set: training set. The first case was with 1:1, where each time validation set and training sets were randomly selected as 1:1. In this case the prediction accuracy of the GEBVs was 24 %. The second case was with 1:5, where the population was divided into five equal parts, where four sets were used as training set against one validation set. In this case the prediction accuracy of the GEBVs was 29%. The same kind of cross validation was done for Population 2 where GEBV prediction accuracies were 29% and 32% for the validation ratios of 1:1 and 1:5, respectively. An attempt was made to select the individuals purely based on the GEBVs wherein top 10% of the individuals with high GEBVs were selected which resulted in significant reduction in the correlations. Population 1 had 0.35 and Population 2 had 0.50, thereby indicating the importance of the accuracy levels in the estimation of the GEBVs. The prediction accuracies in the present investigation were considered to be comparatively low when individuals with high GEBVs were selected as compared to the earlier reported prediction accuracies of GS in corn and other crops for different traits. This variation across populations and studies could be due to quantitative nature of the trait, characteristics of the populations, environmental effects, use of a smaller number of polymorphic markers and small training / population size [46, 47]. Higher prediction accuracies of 0.706 (dent) and 0.690 (flint) in estimating the GEBVs were reported based on the training set size for NCLB resistance [48]. Average prediction accuracies based on genomic data were also high for a complex trait like grain yield (0.72–0.74) [49], grain moisture (0.90) [50, 51], Goss's wilt (0.69) [52], Gray leaf spot (0.84) [53], Fusarium ear rot (0.46) and fumonisin contamination (0.67) [54], Fusarium ear rot (ranging from 0.34 to 0.4) [55], and maize chlorotic mottle virus (0.32, 0.78, 0.47 and 0.21) [56] in maize. Higher prediction accuracies of 0.72 and 0.80 were reported for different agronomic traits in sugar beet [57]. However, prediction accuracies in the present study were comparatively better considering biparental nature of the mapping populations. This result agrees with previous studies which also employed biparental plant populations [58–63].

Evaluation of combining ability of inbreds derived from F_3 progenies

Highly significant line × tester interaction variance indicated that both additive and non-additive gene effects were important in the inheritance of most of the traits studied and also the presence of heterosis [64–68]. The GCA/SCA variance ratio indicated the predominance of non-additive genetic variation [65, 69–71]. The inbred lines with high overall general combining ability status have to be used in hybrid breeding. The seven crosses which showed higher heterosis over the best commercial check DKC9144 have to be further evaluated in large scale trials before commercialization.

Higher variation in the inbred lines was evident from high range of PIC values and the mean PIC value was comparable with the previous findings [72, 73]. The inbred lines derived from the two crosses E2 (CM212 × MAI172), E9 (CM202 × SKV50) were distributed in all the four clusters indicating the presence of molecular diversity among the inbred lines developed.

Conclusion

In this study, comparatively better genomic prediction accuracies were achieved for NCLB and the worth of F_3 progenies with high genomic predictions was proved by advancing them to derive inbred lines and establishing their higher combining ability for yield and yield related traits.

Methods

Plant material

The plant material required for the present study was developed at the Zonal Agricultural Station (ZARS), V.C. Farm, Mandya, Karnataka, India, using two resistant (MAI172 and SKV50) and two susceptible inbreds (CM212 and CM202). Two F_1 s were produced (CM212 x MAI172 and CM202 x

SKV50) during summer 2013. They were grown during rainy season of 2013 and selfed. The F₂ populations consisting of 366 plants each were planted during winter 2013-14 and selfed.

Phenotyping of F_{2:3} mapping populations

The 366 F_{2:3} families developed in the crosses CM212 × MAI172 and CM202×SKV50 were evaluated along with their parents employing Randomized Complete Block Design with two replications during rainy season of 2014 in the National Disease Screening Nurseries for NCLB at ZARS, V.C. Farm, Mandya, Pioneer Hi-Bred disease screening nurseries at Davanagere and Hassan in the Karnataka state, India. The F_{2:3} populations were planted in a 4-meter single row plots with a spacing of 60 cm between rows and 20 cm between plants. The susceptible checks for NCLB, CM202 and Pioneer Hi-bred Seeds commercial hybrid P3522 were planted after every 10th row to assess the disease pressure and as spreader rows. The artificial inoculation procedure was followed for creating disease epiphytotic condition.

Creation of artificial epiphytotic condition

For uniform disease development, artificial inoculation was carried out [74]. The diseased leaves were obtained from the field, and washed three times with sterile distilled water. These leaf tissues were cultured on potato dextrose agar medium to obtain the pathogen inoculum. The sorghum seeds soaked overnight were transferred to sterilized conical flasks and the pathogen inoculum was added. These flasks were shaken once every two days and after one-week equal amount of fresh sorghum seeds were added. The infected sorghum seeds were ground to fine powder, and 1 to 1.5 gram of the ground inoculum was applied to the leaf whorl. Light spray of water was given to create humidity and the inoculum was sprayed 20 days after sowing between 3.00 to 6.00 PM. Another inoculum spray was made after one week.

Disease scoring methodology

The northern corn blight severity was recorded at flowering (60 days after planting) and at dough stage (80 days after sowing) using a standard scale ranging from 1 (Susceptible) to 9 (Resistant) (www.pioneer.com). The disease severity was assessed based on lesion spot development in the middle to upper part leaves.

Score	% Leaf loss	Remarks
9	0–5	Resistant
8	6–10	
7	11–24	
6	25–44	Moderately Resistant
5	45–55	
4	56–66	
3	67–77	Susceptible
2	78–88	
1	89 & above	

The disease scores were converted into *per cent* disease severity using the following formula [75] for data analysis.

Sum of individual disease ratings

$$\text{Per cent Disease Severity} = \frac{\text{Sum of individual disease ratings}}{\text{Total number of plants observed} \times \text{Maximum grade}}$$

Statistical analysis

Since the data did not follow normal distribution, the arcsine transformation was done [76] with the expectation that the means and variances become independent and normally distributed. Using PROC GLM procedure of SAS package version 9.3, the analysis of variance was conducted on transformed phenotypic data. Bartlett's test was used to test for homogeneity between data obtained from three environments before combining data [77]. The analysis of variance was performed considering seasons, replicates and F_{2:3} families as random in the statistical model. Transformed entry means were used to carry out the combined analyses of variance and covariance across seasons [78]. Estimates of variance components σ^2_g (genotypic variance), σ^2_{ge} (G × E) interaction variance), and σ^2_e (error variance) of F_{2:3} families were computed [79](Searle, 1971). The heritability (h^2) was calculated using the following formula [80].

$$\frac{\sigma^2 g}{h^2} = \frac{\sigma^2 g}{Nr} + \frac{\sigma^2 \epsilon}{n}$$

Where, r is number of replications and n is number of test environments.

Genotyping of F_{2:3} mapping population

We employed 768 public plex SNPs covering maize whole genome for genotyping of parents (www.panzea.com). The polymorphic SNPs between parents were then used to genotype 366 F_{2:3} progenies in both crosses. Leaf samples were pooled from eight random plants in each F_{2:3} family, and lyophilized in 96 well plates. Samples were loaded to the Illumina Bead Xpress Vera Code Reader for genotyping following Illumina protocols (<http://www.illumina.com/>).

Scoring of SNP marker generated bands

The markers were scored on F_{2:3} families with A = homozygous maternal genotype, B = homozygous paternal genotype, H = heterozygote genotype and - = Missing samples.

Linkage map construction and QTL detection by IciMapping v3.1

Using IciMapping v3.1 software, linkage map was constructed with polymorphic SNPs [81]. With logarithm of odds (LOD) score of 2.5 linkage groups were identified and the Kosambi mapping function was used to convert recombination frequency into centimorgans [82]. The QTLs were detected employing inclusive composite-interval mapping (CIM) and the results were confirmed by the IciMapping v3.1 software with ICIM-ADD mapping method [81](Wang et al. 2011). The stepwise regression was used to estimate phenotypic variance (R²) explained by each QTL and additive effects at LOD peaks. The gene action was determined by the ratio of the absolute value of the estimated dominance effect divided by the absolute value of the estimated additive effect |d|/|a| [83]. The value between 0–0.20 was considered additive, 0.21–0.80 as partial dominance, 0.81–1.20 as dominance and > 1.20 as overdominance.

Genomic Selection (GS)

The estimation of Best Linear Unbiased Estimates (BLUEs) and Best Linear Unbiased Predictors (BLUPs) are the integral part of genomic selections and were estimated by rrBLUP implemented in R Package [84](Endelman, 2011). We estimated environment wise adjusted phenotypic values of each F3 progenies initially (in step 1) using ASReml version 3.0. Further adjusted phenotypic values in each environment were used for genomic prediction. We estimated combined BLUEs across environment using adjusted phenotypic values. Then BLUEs were used as phenotypes to fit whole genome prediction model.

Ridge regression-Best linear unbiased prediction (RR-BLUP)

By using ridge regression best linear unbiased prediction (RR-BLUP), the marker effects were estimated [19, 85].

GS Prediction Accuracies

Evaluating GEBV accuracy through Cross Validation (CV)

The GS studies on empirical data use cross validation (CV). In CV the data were split into a training set and validation set. We employed two-fold and a five-fold CV in this study. The data set was randomly divided into five sets in five-fold CV. The four sets were combined to form the training set and the remaining set was denoted as the validation set. In two-fold, the observations were divided into training and validation sets equally. In any case, a model was developed using the training data and GEBVs were calculated using genotypes of the validation set. The sampling of training and validation sets was repeated 100 times. The marker effects estimated were used to predict the genomic breeding values in each family across locations. The data sets were divided into training set (TS) to estimate marker effects, and a validation test set (VS), in which the predictive ability (Pearson correlation rMP) between observed BLUEs and the predicted genotypic values was estimated as a measure of prediction accuracy. Correlations were estimated either as accuracy of prediction rp = rMP or as standardized accuracy of prediction. The accuracy of GS was expressed as,

$$rGS = rMP / h$$

Where, h is the square root of heritability [86, 87]

Development and evaluation of inbred lines for combining ability

We have advanced 103 F_{2:3} progenies from Population-1 and 93 from Population-2 with resistance to NCLB to F₈ stage with selection for productivity traits to derive inbred lines. Forty-five inbred lines were selected at F₈ stage and only 34 among them produced successful crosses with all four testers (CM500, MAI105, NAI137 and V351) in line x tester method. The resulting 136 hybrids, 34 inbred lines and four testers along with four commercial check hybrids namely MAH-14-5 (UAS, Bangalore), DKC9141, DKC 9144 (Monsanto India Pvt. Ltd.,) and P3501 (Pioneer Hi-Bred Seeds) notified by the Central Variety Release Committee (CVRC) for South Zone in India, were evaluated under the Completely Randomised Block Design [88] with two replications during rainy season of 2018. Each genotype was sown in a single row of 3m length with 20cm between the plants and 60 cm between the rows. The data recorded on 12 morphological characters were analysed for combining ability and heterosis. Five plants in each genotype were tagged randomly in both replications and same were used to record observations. Mean of five plants was used for the combining ability analysis [89]. Combining ability analysis helps in identifying superior lines to be used in breeding programs or to identify promising cross combinations for the development of varieties [90]. General combining ability (GCA) describes the breeding value of a parent and is generally associated with additive genetic effects, while specific combining ability (SCA) is the relative performance of a cross that is associated with non-additive gene action, predominantly contributed by dominance, epistasis, or genotype × environment interaction effects [91, 92]. Therefore, both gca and sca effects are important in the selection or development of breeding populations [93]. Heterosis over the best commercial check (standard heterosis) was computed [94, 95].

Overall general combining ability (gca) status of parents, specific combining ability (sca) and Heterotic status of crosses

The *gca* status of parents and *sca* or heterosis of crosses were estimated and expressed in the desirable direction. The overall status of a parent or cross with respect to *gca* and *sca* or heterosis over a number of characters was assessed [96, 97].

Assessment of molecular diversity in the inbred lines using SSR markers

Genomic DNA from 45 inbreds developed was extracted from the leaves of three-week-old maize seedlings using modified CTAB procedure [98]. The genotyping of the inbred lines was carried out with 64 SSR markers. For each primer, the amplified DNA fragments were scored as 1 or 0 based on whether the DNA fragment is present or absent. The binary data thus obtained were used to estimate pair wise similarity coefficients [99]. The NTSYS software was employed to subject the resulted similarity matrix to hierarchical cluster analysis using the Unweighted Pair Group Method with Arithmetic average (UPGMA) algorithm [100].

Declarations

Acknowledgements

The authors are grateful to Dr. Manje Gowda, Maize Molecular Breeder, CIMMYT, Kenya, for his help in the analysis of data.

Author contributions

HCL designed the research. DCB, MR, HCL and BSB performed experiments and analyzed the data. AP helped in genotyping of mapping populations. HCL wrote the manuscript with inputs from DCB, MR and RLR. All authors approved the final version of manuscript.

Funding

This work was funded by M/s Pioneer Hi-Bred Seeds Pvt. Ltd., India.

Data availability statement

All datasets generated for this study are included in the Supplementary Material.

Ethics approval and consent to participate

Relevant permissions were obtained from an appropriate authority, and the studies comply with local and national regulations.

Consent for publication

Not applicable.

Competing interest

The authors declare no competing interests.

Author details

¹Department of Genetics and Plant Breeding, College of Agriculture, GVK, University of Agricultural Sciences, Bangalore, Karnataka 571 405, India. ²Department of Genetics and Plant Breeding, College of Agriculture, V.C. Farm, Mandya (University of Agricultural Sciences, Bangalore), Karnataka 571

References

1. Xu Y, Skinner DJ, Wu H, Palacios-Rojas N, Araus JL, Yan J, Gao S, Warburton ML, Crouch JH. Advances in Maize Genomics and Their Value for Enhancing Genetic Gains from Breeding. *International J Plant Genomics*. 2009; doi:10.1155/2009/957602
2. Raymundo AD, Hooker AC. Measuring relationship between northern leaf blight of maize and yield losses. *Plant Dis*. 1981; 65: 325–327.
3. Perkins JM, Pederson WL. Disease treatment and yield loss associated with northern leaf blight of corn. *Plant Dis*. 1987; 71: 940–943.
4. Adipala E, Takan JP, Ogena-Latigo MW. Effect of planting density of maize on the progress and spread of northern leaf blight from *Exserohilum turcicum* infested residue source. *European J Plant Pathol*. 1995; 101: 25–33.
5. Dingerdissen AL, Geiger HH, Lee M, Schechert A, Welz HG. Interval mapping of genes for quantitative resistance of maize to *Setosphaeria turcica*, cause of northern leaf blight, in a tropical environment. *Mol Breed*. 1996; 2: 143–156.
6. Brewster VA, Carson ML, Wicks ZW. Mapping components of partial resistance to northern leaf blight of maize using reciprocal translocations. *Phytopathol*. 1992; 82, 225–229.
7. Schechert AL, Welz HG, Geiger HH. QTL for resistance to *Setosphaeria turcica* in tropical African maize. *Crop Sci*. 1999; 39:514–523.
8. Welz HG, Geiger HH. Genes for resistance to northern corn leaf blight in diverse maize populations. *Plant Breed*. 2000, 119: 1–14.
9. Welz HG, Xia XC, Bassetti P, Melchinger AE, Lubberstedt T. QTLs for resistance to *Setosphaeria turcica* in an early maturing dent × flint maize population. *Theor Appl Genet*. 1999; 99:649–655.
10. Balint-Kurti PJ, Yang J, Esbroeck GV, Jung J, Smith ME. Use of a maize advanced intercross line for mapping of QTL for Northern leaf blight resistance and multiple disease resistance. *Crop Sci*. 2010; 50:458–466.
11. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci*. 2011; 108:6893–6898.
12. Inghelandt VD, Melchinger AE, Martinant JP, Stich B. Genome-wide association mapping of flowering time and northern corn leaf blight (*Setosphaeria turcica*) resistance in a vast commercial maize germplasm set. *Plant Biol*. 2012; 12:56–70.
13. Ranganatha H M, Lohithaswa HC, Pandravada AS. Understanding the genetic architecture of resistance to northern corn leaf blight and southern corn rust in maize (*Zea mays* L.). *Indian J Genet*. 2017; 77(3): 357–363, DOI: 10.5958/0975-6906.2017.00048.7
14. Ranganatha HM, Lohithaswa HC, Pandravada, A. Mapping and Validation of Major Quantitative Trait Loci for resistance to Northern Corn Leaf Blight along with the determination of the relationship between resistances to multiple foliar pathogens of maize (*Zea mays* L.). *Front Genet*. 2021 Jan 29;11:548407. doi: 10.3389/fgene.2020.548407. PMID: 33584784; PMCID: PMC7878677.
15. Wisser RJ, Balint-Kurti PJ, Nelson RJ. The genetic architecture of disease resistance in maize: A synthesis of published studies. *Phytopathol*. 2006; 96:120–129.
16. Massman, J, Cooper B, Horsley R, Neate S, Dill-Maeky R, Chao S, Dong Y, Schwarz P, Muehlbauer GJ, Smith KP. Genome-wide association mapping of Fusarium head blight resistance in contemporary barley breeding germplasm. *Mol Breed*. 2010; 27:439–454. doi:10.1007/s11032-010-9442-0
17. Desta ZA, Ortiz R. Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci*. 2014; 19 (9): 592–601. https://doi.org/10.1016/j.tplants.2014.05.006
18. Bhat JA, Ali S, Salgotra RK, Mir ZA, Dutta S, Jadon V, Tyagi A, Mushtaq M, Jain N, Singh PK, Singh GP, Prabhu KV. Genomic Selection in the Era of Next Generation Sequencing for Complex Traits in plant Breeding. *Front. Genet*. 2016; 7:221. doi: 10.3389/fgene.2016.00221
19. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 2001; 157:1819–1829.
20. Galiano-Carneiro AL, Miedaner T. Genetics of Resistance and Pathogenicity in the Maize/*Setosphaeria turcica* Pathosystem and Implications for Breeding, *Frontiers Plant Sci*. 2017, 8: 1490. DOI = 10.3389/fpls.2017.01490.
21. Fisher RA, Immer FR, Tendin D. The genetical interpretation of statistic of the third degree in the study of quantitative inheritance. *Genetics*. 1932; 17:107–124.
22. Robson DS. Application of K4 statistics to genetic variance component analysis. *Biometrics*. 1956; 12: 433–444.
23. Brown AF, Juvik JA, Pataky JK. Quantitative trait loci in sweet corn associated with partial resistance to Stewart's Wilt, northern corn leaf blight, and common rust, *Phytopathol*. 2001;91: 293–300.
24. Asea G, Vivek BS, Bigirwa G, Lipps PE, Pratt RC. Validation of consensus quantitative trait loci associated with resistance to multiple foliar pathogens of maize. *Phytopathol*. 2009; 99: 540–547.
25. Campana A, Pataky JK. Frequency of the *Ht1* gene in populations of sweet corn selected for resistance to *Exserohilum turcicum* race 1. *The American Phytopathol Soc*. 2005; 95(1):85–91.

26. Pooni HS, Jinks JL, Cornish MA. The causes and consequences of non-normality in predicting the properties of recombinant inbred lines. *Heredity*. 1977; 38:329–338.
27. Kimbeng CA, Bingham ET. Population improvement in lucerne (*Medicago sativa* L.): components of inbreeding depression are different in original and improved populations. *Australian J Exp Agri*. 1998; 38: 831–836.
28. Roy D. Plant Breeding – The Analysis and exploitation of variability, Narosa Publishing House. New Delhi, India, 2000; pp. 198.
29. Choo TM, Reinbergs E. Analysis of skewness and kurtosis for detecting gene interaction in a double haploid population. *Crop Sci*. 1982; 22: 231–235.
30. Hakiza JJ, Lipps PE, Martin SS, Pratt RC. Heritability and number of genes controlling partial resistance to *Exserohilum turicum* in maize inbred H99. *Maydica*. 2004; 49:173–182.
31. Vieira RA, Scapim CA, Moterle LM, Tessmann DJ, Conrado TV, Amaral AT. Diallel analysis of leaf disease resistance in inbred Brazilian popcorn cultivars. *Genet. Mol Res*. 2009; 8: 1427–1436.
32. Vieira RA, Scapim CA, Tessmanne DS, Hata FT. Diallel analysis of yield, popping expansion, and southern rust resistance in popcorn lines. *Revista Ciência Agronômica*. 2011; 42(3):774–780.
33. Freymark PJ, Lee M, Woodman WL, Martinson CA. Quantitative and qualitative trait loci affecting host-plant response to *Exserohilum turicum* in maize (*Zea mays* L.). *Theor Appl Genet*. 1993; 87:537–544.
34. Freymark PJ, Lee M, Woodman WL, Martinson CA. Molecular-marker facilitated investigation of host-plant response to *Exserohilum turicum* in maize (*Zea mays* L.): Components of resistance. *Theor Appl Genet*. 1994; 88: 305–313.
35. Zwonitzer JC, Coles ND, Krakowsky MD, Arellano C, Holland JB, McMullen MD, Pratt RC, Balintkurti PJ. Mapping resistance quantitative trait loci for three foliar diseases in a maize recombinant inbred line population-Evidence for multiple disease resistance?. *Phytopathol*. 2010; 100: 72–79.
36. Davis GL, McMullen MD, Baysdorf C, Musket T, Grant D, Staebell M, Xua G, Polacco M, Kosterd L, Melia-Hancock S. A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. *Genetics*. 1999; 152:1137–1172.
37. Sharopova N, McMullen MD, Schultz L, Schroeder S, Sanchez-Villeda H, Gardiner J, Bergstrom D, Houchins K. Development and mapping of SSR markers for maize. *Plant Mol. Biol*. 2002; 48(5–6): 463–481.
38. Doerge WR. Mapping and analysis of quantitative trait loci in experimental populations. *Genetics*. 2002; 3: 43–52.
39. Lin YR, Keith FS, Paterson AH. Comparative Analysis of QTLs affecting plant height and maturity across the poaceae, in reference to an interspecific sorghum population. *Genetics*. 1995; 141:391–411.
40. Tuberosa R, Salvi S, Sanguineti MC, Landi P, Maccaferri M, Conti S. Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. *Ann Bot*. 2002; 89: 941–963.
41. Chung CL, Jamann T, Longfellow J, Nelson R. Characterization and fine-mapping of a resistance locus for northern leaf blight in maize bin 8.06. *Theor Appl Genet*. 2010; 121:205–227.
42. Chung CL, Poland P, Kump K, Benson J, Longfellow J, Walsh E, Balint-Kurti PJ, Nelson R. Targeted discovery of quantitative trait loci for resistance to northern leaf blight and other diseases of maize. *Theor Appl Genet*. 2011; 123: 307–326
43. Schaeffer LR (2006) Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet* 123: 218–223.
44. Heffner EL, Lorenz AJ, Jannink JL, Sorrells ME. Plant breeding with genomic selection: Gain per unit time and cost. *Crop Sci*. 2010; 50:1681–1690.
45. Heffner EL, Sorrells ME, Jannink JL. Genomic selection for crop improvement. *Crop Sci*. 2009; 49:1–11.
46. Crossa J, Pérez P, Hickey J, Burguen J, Ornella L, Ceron-Rojas J, Zhang X, Dreisigacker S, Babu R, Li Y, Bonnett D and Mathews K. Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity*. 2014; 112: 48–60. <https://doi.org/10.1038/hdy.2013.16>
47. Jiang S, Cheng Q, Yan J, Fu R, Wang X. Genome optimization for improvement of maize breeding. *Theoret Appl Genet*. 2019; <https://doi.org/10.1007/s00122-019-03493-z>
48. Technow F, Burger A, Melchinger AE. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3-Genes Genome Genet*. 2013; 3:197–203.
49. Albrecht T, Wimmer V, Auinger HJ, Erbe M, Knaak C, Ouzunova M, Simianer H, Schön CC. Genome-based prediction of testcross values in maize. *Theor. Appl. Genet*. 2011; 123: 339–350. doi:10.1007/s00122-011-1587-7
50. Zhao Y, Gowda M, Liu W, Würschum T, Maurer HP, Longin FH, Ranc N, Reif JC Accuracy of genomic selection in European maize elite breeding populations. *Theor. Appl. Genet*. 2012a; 124: 769–776.
51. Zhao Y, Gowda M, Longin CFH, Würschum T, Ranc N, Reif JC. Impact of selective genotyping in the training population on accuracy and bias of genomic selection. *Theor Appl Genet*. 2012b; 125:707–713.
52. Cooper JS, Rice BR, Shenstone EM, Lipka AE, Jamann TM. Genome-wide analysis and prediction of resistance to Goss's wilt in maize. *Plant Genome*. 2019; 12:180045. doi:10.3835/ plantgenome2018.06.0045

53. Kibe M, Nair SK, Das B, Bright JM, Makumbi D, Kinyua J, Beyene Y, Olsen MS, Boddupalli P, Gowda M. Genetic Dissection of Resistance to Gray Leaf Spot by Combining Genome-wide Association, Linkage Mapping and Genomic Prediction in Tropical Maize Germplasm. *Frontiers Plant Sci.* 2020; 11:1602.
54. Holland JB, Marino TP, Manching HC, Wisser RJ. Genomic prediction for resistance to Fusarium ear rot and fumonisin contamination in maize. *Crop Sci.* 2020; 60(4):1863–1875.
55. Kuki MC, Pinto RJB, Bertagna FAB, Tessmann DJ, Teixeira do Amaral A, Scapim CA, Holland JB. Association mapping and genomic prediction for ear rot disease caused by *Fusarium verticillioides* in a tropical maize germplasm. *Crop Sci.* 2020; 60(6): 2867–2881.
56. Suresh LM, Beyene Y, Olsen MS, Makumbi D, Oliver K, Das B, Bright JM, Mugo, S, Crossa J, Tarekegne A, Prasanna, BM. Genetic architecture of maize chlorotic mottle virus and maize lethal necrosis through GWAS, linkage analysis and genomic prediction in tropical maize germplasm. *Theor Applied Genet.* 2019; 132(8): 2381–2399.
57. Wurschum T Reif JC, Kraft T, Janssen G, and Zhao Y. Genomic selection in sugar beet breeding populations. *BMC Genet.* 2013; 14: 85.
58. Bernardo, R. Genomewide selection for sapid introgression of exotic germplasm in maize. *Crop Sci.* 2009; 49(2), 419–425.
59. Lorenzana R, Bernardo R. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theor Appl Genet.* 2009; 120:151–161.
60. Heffner EL, Jannink JL, Iwata H, Souza E, Sorrells M. Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 2011; 51:2597–2606.
61. Riedelsheimer C, Endelman JB, Stange M, Sorrells ME, Jannink JL, Melchinger AE. Genomic predictability of interconnected biparental maize populations maize populations. *Genetics.* 2013; 194: 493–503. doi:10.1534/genetics.113.150227/-DC1.
62. Zhang X, Pérez-Rodríguez P, Semagn K, Beyene Y, Babu R, López-Cruz MA, San Vicente F, Olsen M, Buckler E, Jannink JL, Prasanna BM, Crossa J. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity.* 2015; 114:291–299. doi:10.1038/hdy
63. Beyene Y, Semagn K, Mugo S, Tarekegne A, Babu R, Meisel B. Genetic gain in grain yield through genomic selection in eight bi-parental maize populations under drought stress. *Crop Sci.* 2015; 55:154–163. doi:10.2135/cropsci2014.07.0460
64. Joshi VN, Dubey RB, Marker S. Combining ability for polygenic traits in early maturity hybrids of maize (*Zea mays L.*), Indian J Genet. 2002; 62: 312–315.
65. Kambegowda R, Kage U, Lohithaswa HC, Shekara BG, Shobha D. Combining ability studies in maize (*Zea mays L.*). *Mol Plant Breed.* 2013; 4(14): 116–127.
66. Kumar GP, Reddy VN, Kumar SS, Rao PV. Combining ability studies in newly developed inbred lines in maize (*Zea mays L.*). *Int J Plant Animal Environ Sci.* 2014; 4(4): 229–234.
67. Kumar R, Mandal SS, Mishra AK, Smriti, Rahul S, Kumar P. Heterosis and combining ability for yield and its contributing traits of kharif maize (*Zea mays L.*). *The Bioscan.* 2015; 10(4): 2049–2056.
68. Seyoum A, Wegary D, Alamerew S. Combining ability of elite highland maize (*Zea mays L.*) inbred lines at Jimma dedo, South West Ethiopia. *Adv Crop Sc. Tec.* 2016; 4(2): 102–113.
69. Amiruzzaman M, Islam MA, Hassa L, Kadir M, Rohman MM. Heterosis and combining ability in a diallele among elite inbred lines of maize (*Zea mays L.*). *Emirates J Food Agric.* 2013; 25(2): 132–137.
70. Aminu D, Mohammed SG, Kabir BG. Estimates of combining ability and heterosis for yield and yield traits in maize population (*Zea mays L.*) under drought conditions in the Northern Guinea and Sudan Savanna zones of Borno State, Nigeria. *Int J Agric Innovations Res.* 2014; 2(5): 824–830.
71. Ruswandi D, Supriatna J, Makkulawu AT, Waluyo B, Marta H, Suryadi E, Ruswandi S. Determination of combining ability and heterosis of grain yield components for maize mutants based on linextester analysis. *Asian J Crop Sci.* 2015; 7(1): 19–33.
72. Patto MV, Satovic Z, Pego S, Fevereiro P. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. *Euphytica.* 2004; 137(1): 63–72.
73. Nepolean T, Ishwar S, Firoz H, Neha P, Gupta HS. Molecular characterization and assessment of genetic diversity of inbred lines showing variability for drought tolerance in maize. *J Pl Biochem Biotech.* 2013; 22(1):71–79.
74. Shekhar M, Kumar S. Technical bulletin; Inoculation methods and disease rating scales for maize diseases. Directorate of Maize Research, New Delhi, 2012.
75. Wheeler BEJ. An Introduction to Plant Diseases. John Wiley and Sons Ltd., London, United Kingdom, 1969; pp. 301.
76. Little TM, Hills FJ. Agricultural Experimentation: Design and Analysis. New York: John Wiley and Sons, 1978
77. Gomez KA, Gomez AA. *Statistical Procedure for Agricultural Research.* New York: John Wiley and Sons, 1984
78. Bohn M, Khairallah MM, Gonzalez-de-Leon D, Hoisington D, Utz HF, Deutsch JA, et al. QTL mapping in tropical maize: I. Genomic regions affecting leaf feeding resistance to sugarcane borer and other traits. *Crop Sci.* 1996; 36, 1352–1361. doi: 10.2135/cropsci1996.0011183x003600050045x
79. Searle SR. Linear Models. New York: John Wiley & Sons, 1971

80. Hallauer AR, Miranda JB. Quantitative genetics in maize breeding. Ames, IA: Iowa State University Press, 1981
81. Wang J, Li H, Zhang L, Li C, Meng L. Users' Manual of QTL IciMapping v3.1. Webpage: <http://www.isbreeding.net>. 2011
82. Kosambi DD. The estimation of map distances from recombination values. Ann Eugen. 1943; 12: 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x
83. Stuber CW, Edwards MD, Wendel JF. Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci. 1987; 27: 639–648. doi: 10.2135/cropsci1987.0011183x002700040006x
84. Endelman JB. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome. 2011; 4:250–255. doi: 10.3835/plantgenome2011.08.0024
85. Whittaker JC, Thompson R, Denham MC. Marker-assisted selection using ridge regression. Genet Res. 2000; 75: 249–252.
86. Lande R, Thompson R. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics. 1990; 124: 743–756.
87. Dekkers J. Prediction of response to marker-assisted and genomic selection using selection index theory. J Anim Breed Genet. 2007; 124(6):331–341.
88. Yates F. A new method of arranging variety trials involving a large number of varieties. J Agric Sci. 1936; 26: 424–455.
89. Kempthorne O. An introduction to genetic statistics John Wiley and Sons, New York, 1957
90. Acquaah G. Principles of Plant Genetics and Breeding. Oxford: Wiley-Blackwell, 2007
91. Rojas BA, Sprague GF. A comparison of variance components in corn yield trials: III. General and specific combining ability and their interaction with locations and years. Agron J. 1952. 44, 462–466. doi: 10.2134/agronj1952.00021962004400090002x
92. Falconer DS, Mackay TFC. Introduction to Quantitative Genetics, 4th Edn. Harlow: Pearson Prentice Hall. 1996.
93. Viana JMS, Matta FP. Analysis of general and specific combining abilities of popcorn populations, including selfed parents. Genet Mol Biol. 2003; 26: 465–471. doi: 10.1590/S1415-47572003000400010
94. Turner JK. A study of heterosis in upland cotton and combining ability and inbreeding effects. Agron J. 1953; 45: 487–490.
95. Hayes, HK, Immer FR, Smith DC. Methods of Plant Breeding. Mc. Grow Hill Book Co., Inc., New York. 1955; pp.165–180
96. Arunachalam V, Bandopadhyay A. Are multiple cross- multiple pollen hybrids an answer for productive populations in *Brassica campestris* variety brown sarson? Theor Appl Genet. 1979; 54: 203–20.
97. Mohan Rao A, Lakshminarayudu G, Kulkarni RS, Ramesh S, Lalitha Reddy SS. Prediction of heterosis based on genetic divergence of parents through regression analysis in sunflower. Helia. 2004; 27(41): 51–58.
98. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA. 1984; 81: 8014–8018.
99. Jaccard P. Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat. 1908; 44, 223–270.
100. Rohlf FJ. NTSYS-pc. Numerical taxonomy and multivariate analysis system. V.2.1. Exeter Software, Setauket, New York, USA, 2000

Figures

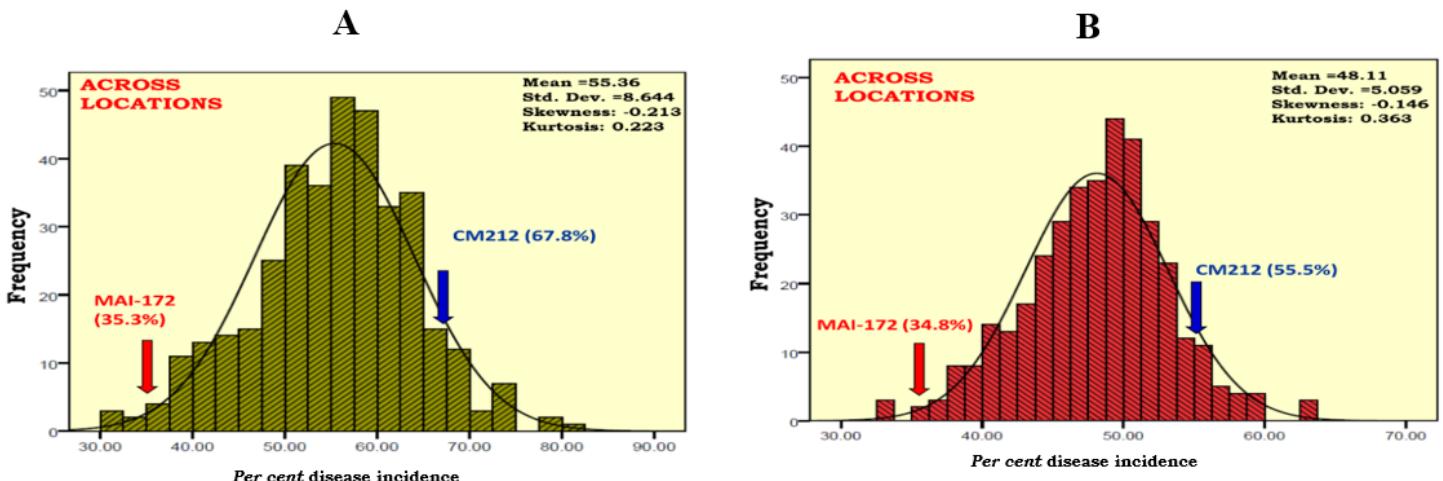


Figure 1

Frequency distribution of per cent NCLB disease incidence in F2:3 population derived from the cross CM 212 x MAI 172 across three locations (A=original; B= Arcsine transformed)

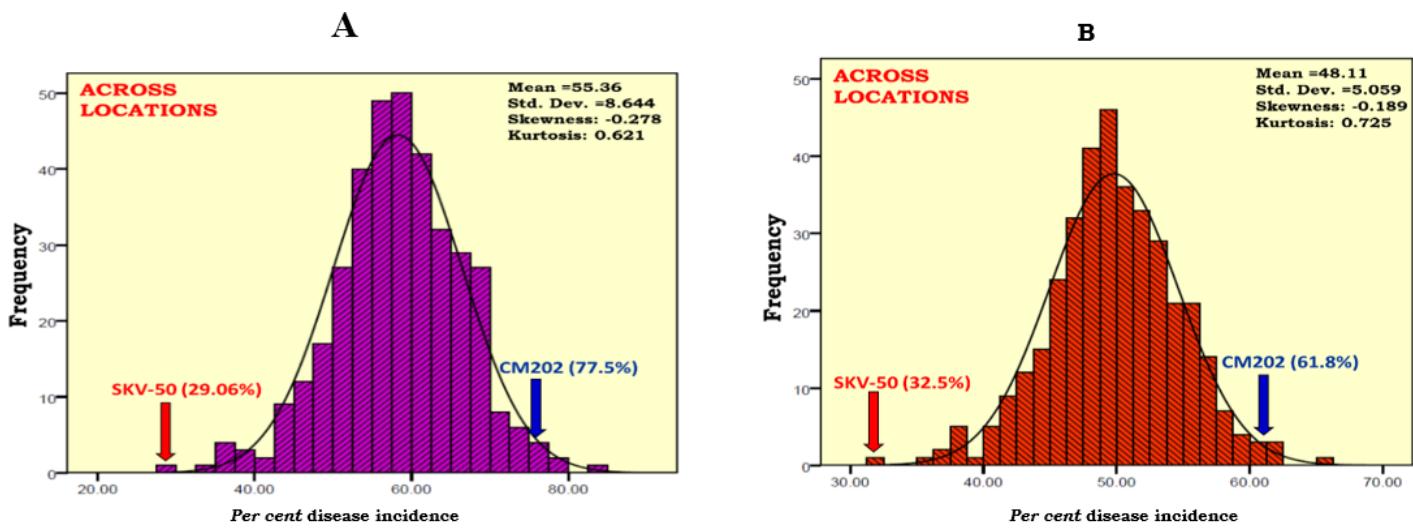


Figure 2

Frequency distribution of per cent NCLB disease incidence in F2:3 population derived from the cross CM 202 x SKV 50 across three locations (A=original; B= Arcsine transformed)

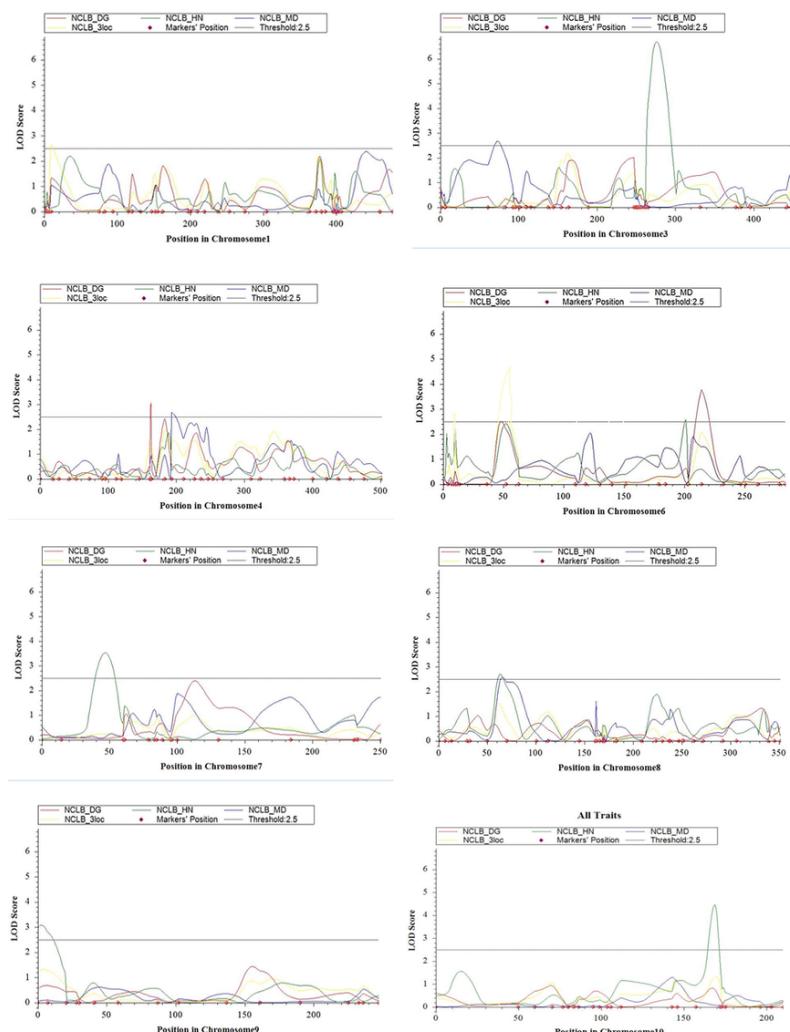


Figure 3

LOD peak for NCLB resistant QTLs identified on chromosome 1, 3,4, 6,7,8,9 and 10 in F2:3 population derived from the cross CM 212 x MAI 172 during rainy season of 2014 at locations and in pooled analysis HN: Hassan; DG: Davanagere; MD: Mandya

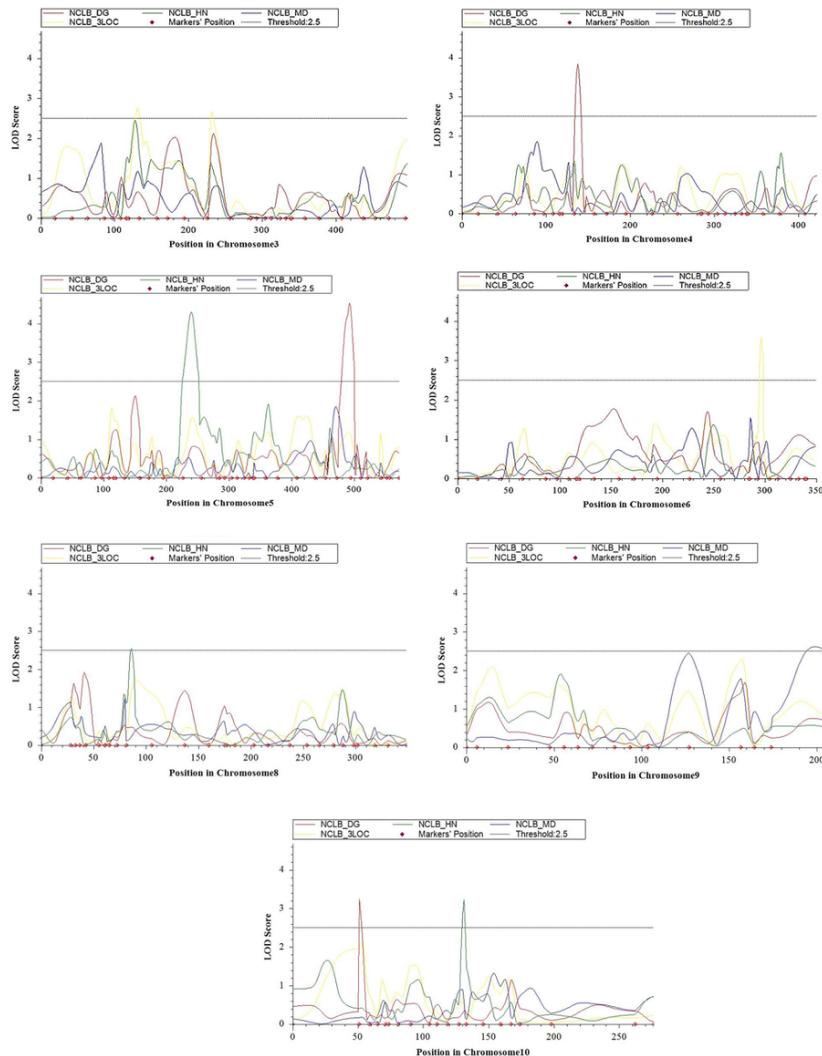


Figure 4

LOD peak for NCLB resistant QTLs identified on chromosome 3,4,5,6,8,9 and 10 in F2:3 population derived from the cross CM 202 x SKV 50 during rainy season of 2014 at three locations and in pooled analysis HN: Hassan; DG: Davanagere; MD: Mandya

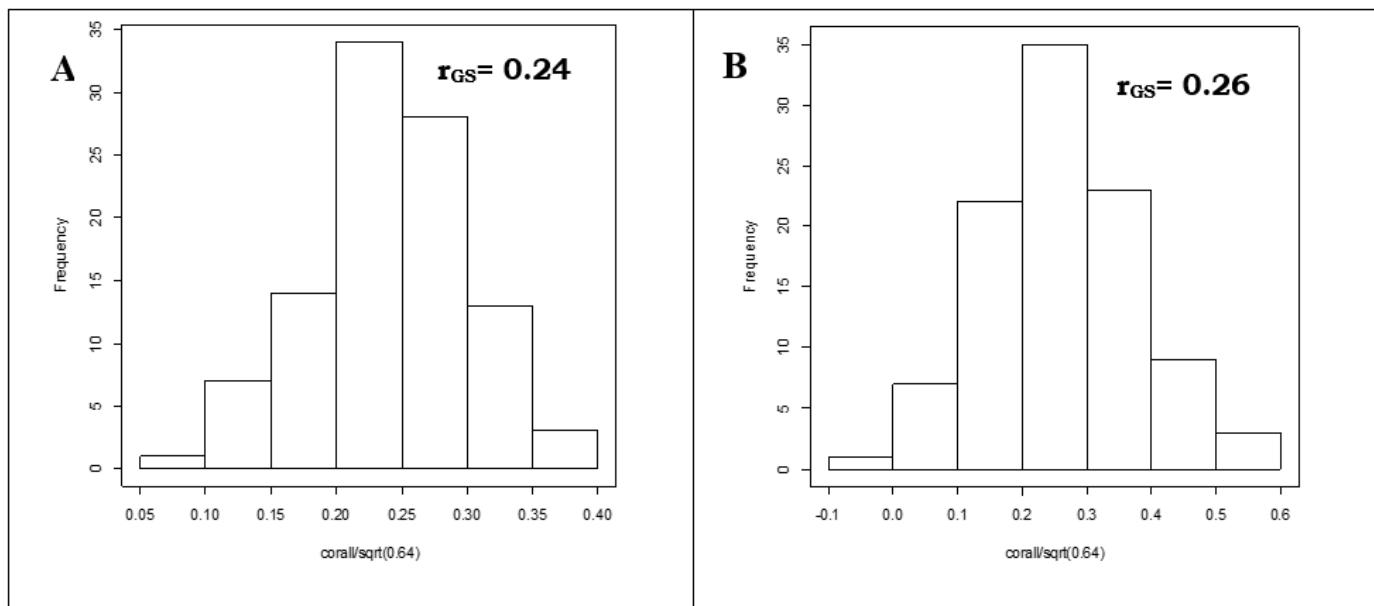


Figure 5

Distribution of accuracy of genomic prediction (rGS) in F2:3 population of the cross CM 212 x MAI 172 (A-1:1 ratio, 100 CVs, B-1:5 ratio, 100 CVs)

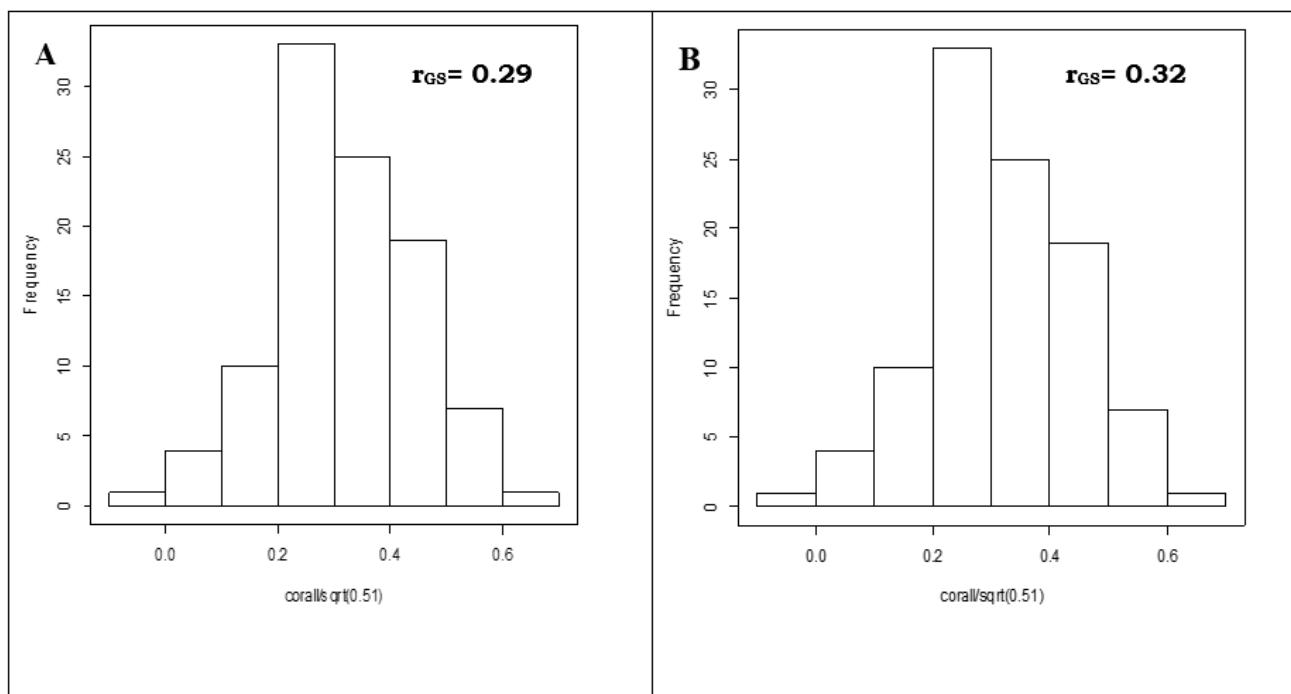


Figure 6

Distribution of accuracy of genomic prediction (rGS) in F2:3 population of the cross CM 202 x SKV 50 (A-1:1 ratio, 100 CVs, B-1:5 ratio, 100 CVs)

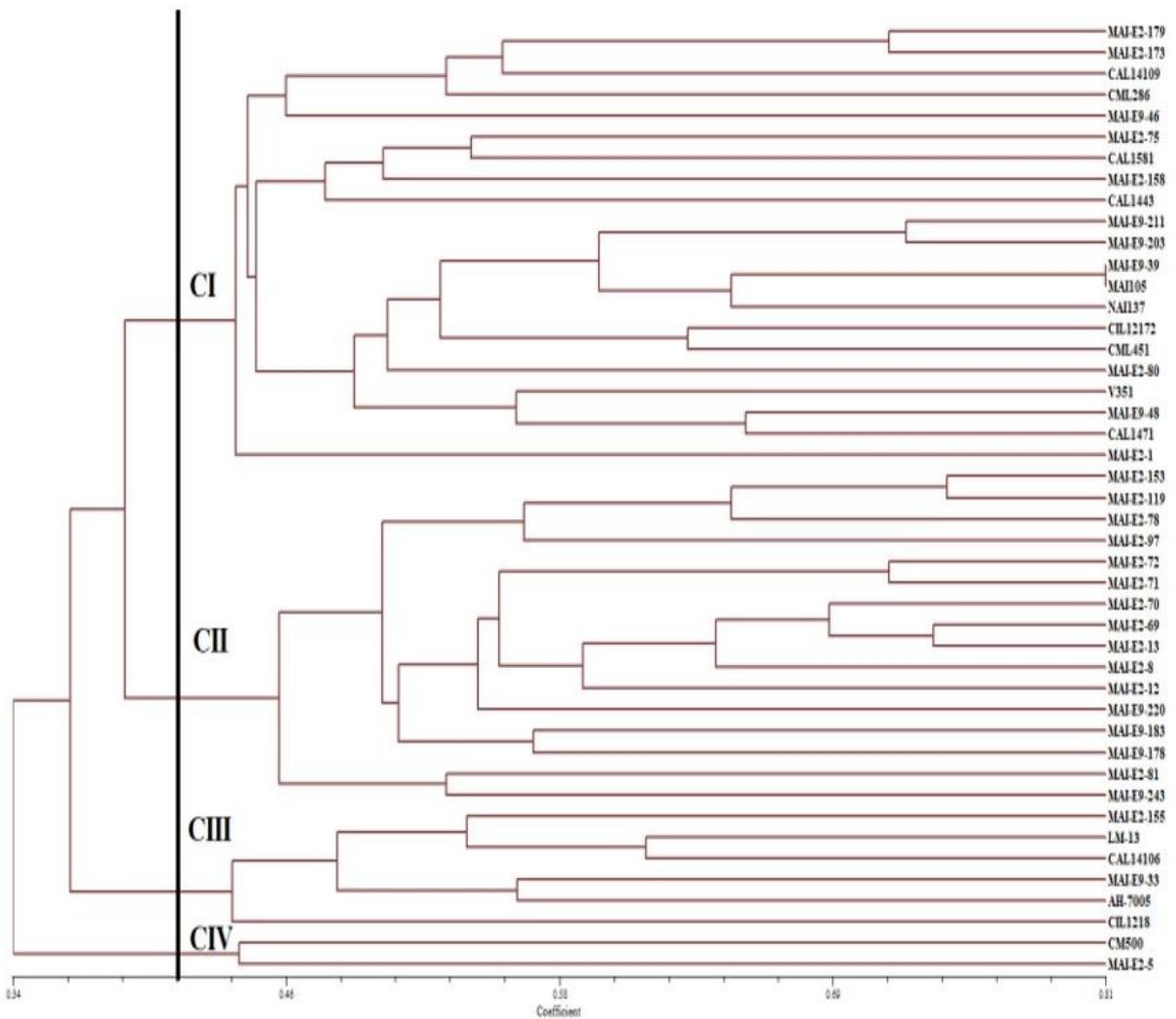


Figure 7

Dendrogram of 45 inbred lines along with designated clusters at similarity coefficient of 0.42

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suppl.TableS1.docx](#)
- [Suppl.TableS2.docx](#)
- [Suppl.TableS3.doc](#)
- [Suppl.TableS4.docx](#)
- [Suppl.TableS5.docx](#)
- [Suppl.TableS6.docx](#)
- [Suppl.TableS7.docx](#)