Doubled Haploid (DH) Technology in Tropical Maize Breeding

Vijay Chaikam

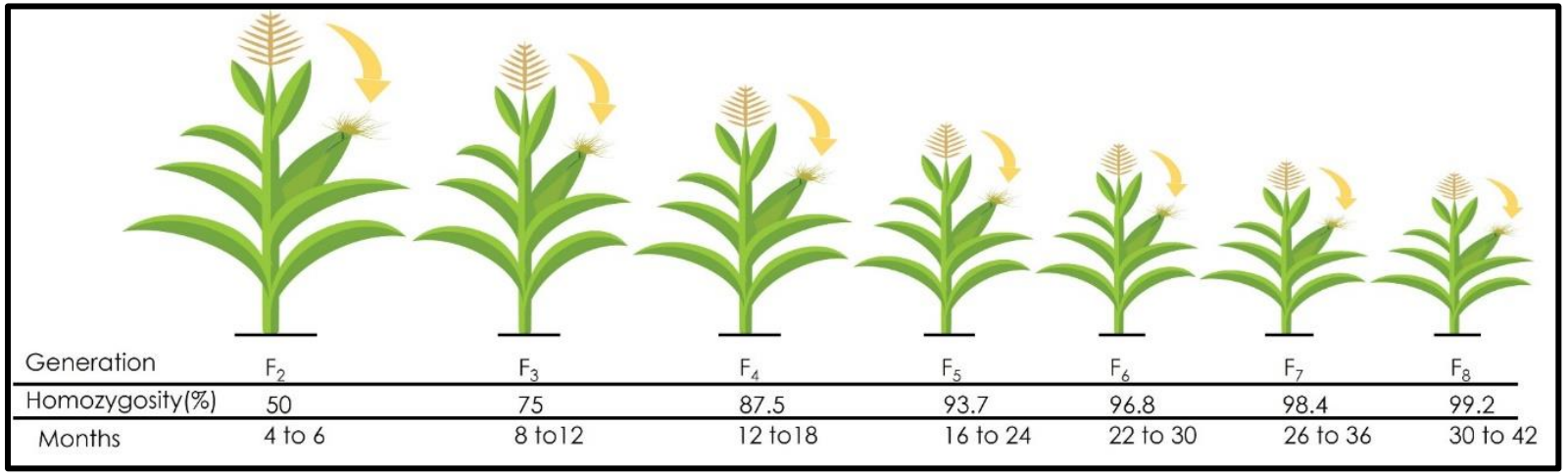
Crops to End Hunger Case Studies in Africa and Beyond: Supporting CGIAR Partners through Genotyping Services

April 11, 2024

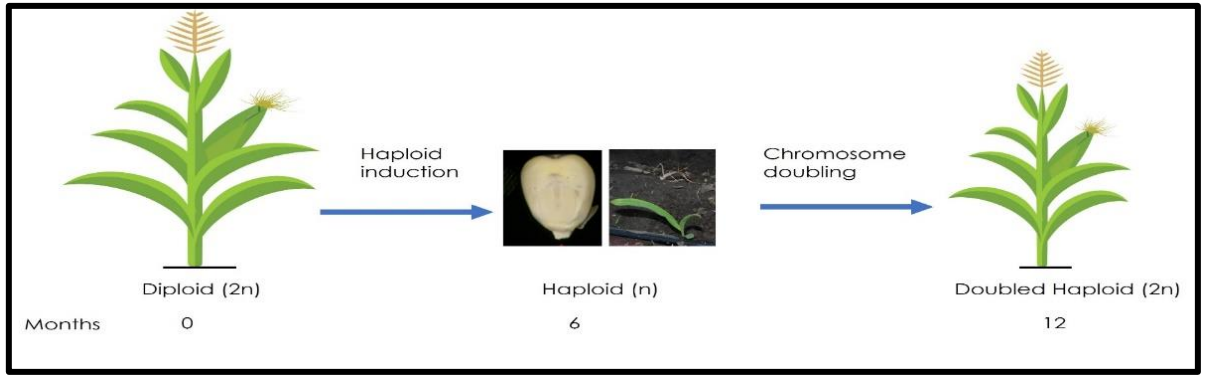
CIMMYT | CGIAR

Conventional Inbreeding vs. DH

B73 B73 X MO17 MO17



> 99% homozygous lines can be developed from F1 in **8 generations**



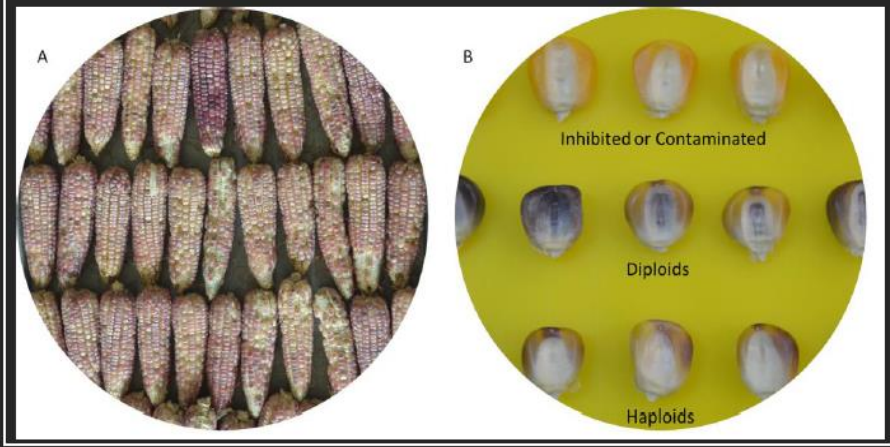
100% homozygous lines can be developed in just **two** generations

1% heterozygosity in maize = ~300 genes segregating

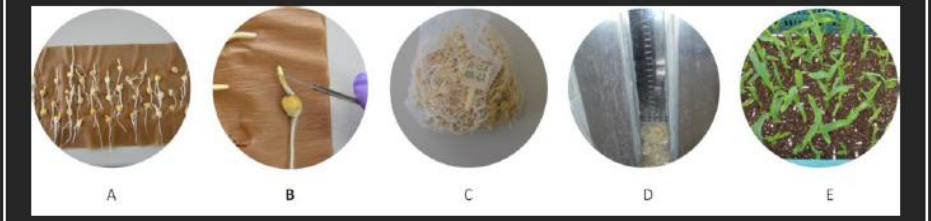
Maize DH line development process

Haploid induction

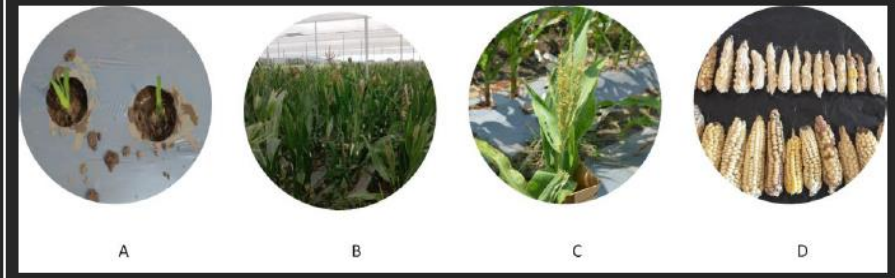
Haploid identification



Chromosomal doubling



Seed production from fertile doubled haploids



DH Line Production Pipelines in CIMMYT

Kiboko DH facility, Kenya
Established in 2012-13

Agua fria DH facility, Mexico
Established in 2011

Kunigal DH facility, India
Established in 2021

Year	CIMMYT		NARS partners in Africa & LatAm		Privite sector in Africa & LatAm		TOTAL	
	# of Populations	# of DH lines	# of Populations	# of DH lines	# of Populations	# of DH lines	# of Populations	# of DH lines
2017	409	18,723	8	1727	38	7944	455	28,394
2018	216	15,908	67	10,630	70	10,957	353	37,495
2019	366	13,287	58	9,145	35	5,410	459	27,842
2020	350	40,038	77	8,037	48	9,631	475	57,706
2021	164	21598	41	6182	36	2771	241	30551
2022	154	39916	47	6925	69	8406	270	55247
2023	396	53584	141	8978	53	5073	590	67635
Total	2055	203054	439	51624	349	50192	2843	304,870
		63%		17%		16%		

- At least 9 lines released as CMLs (notified and publicly available CIMMYT lines)
- Several DH-based CIMMYT hybrids are released in Sub-Saharan Africa and Mexico every year

CIMMYT | CGIAR

Improved Efficiencies in Maize DH Line Production

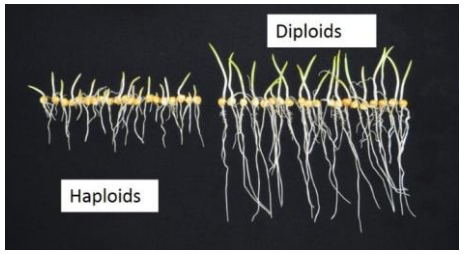
Development of tropicalized haploid inducer Lines(TAILs)

1st generation TAIL hybrid

2nd generation TAIL hybrid

Use of red root marker

Use of seedling traits

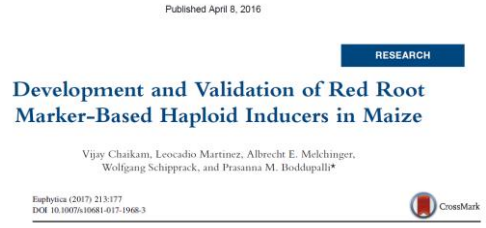


Theor Appl Genet (2015) 128:159–171
DOI 10.1007/s00122-014-2419-3

ORIGINAL PAPER

Analysis of effectiveness of *R1-nj* anthocyanin marker for in vivo haploid identification in maize and molecular markers for predicting the inhibition of *R1-nj* expression

Vijay Chaikam · Sudha K. Nair · Raman Babu · Leocadio Martinez · Jyothsna Tejomurtula · Prasanna M. Boddupalli



Published April 8, 2016

RESEARCH

Development and Validation of Red Root Marker-Based Haploid Inducers in Maize

Vijay Chaikam, Leocadio Martinez, Albrecht E. Melchinger, Wolfgang Schipprack, and Prasanna M. Boddupalli*

Euphytica (2017) 213:177
DOI 10.1007/s10681-017-1968-3

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 October 2018
doi: 10.3389/fpls.2018.01527

Marker-Assisted Breeding of Improved Maternal Haploid Inducers in Maize for the Tropical/Subtropical Regions

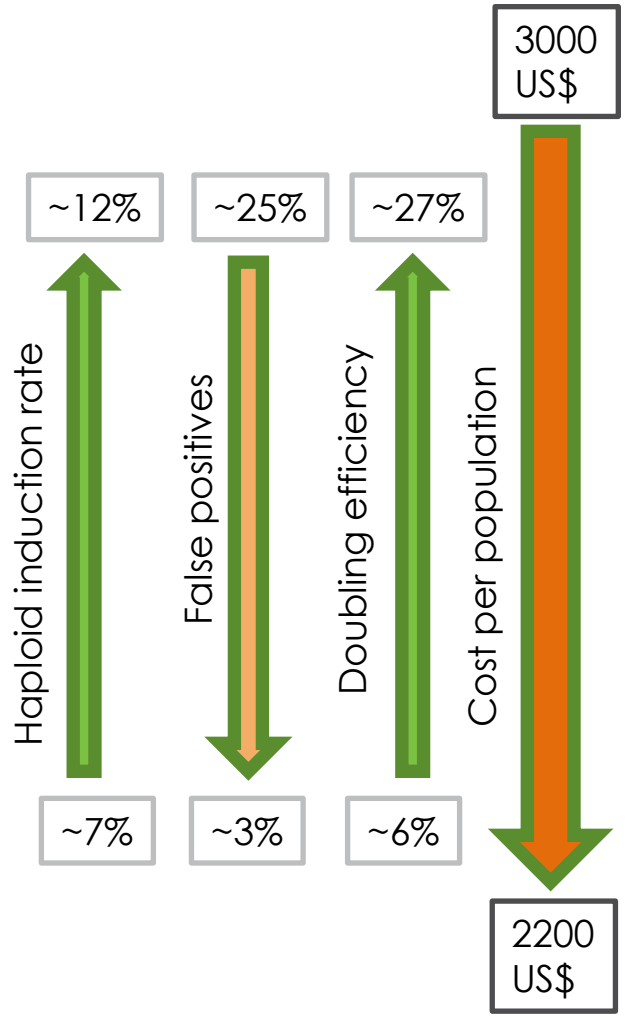
Vijay Chaikam[1], Sudha K. Nair[2], Leocadio Martinez[3], Luis Antonio Lopez[3], H. Friedrich Utz[4], Albrecht E. Melchinger[4] and Prasanna M. Boddupalli[1]*

plants

Article

Improving the Efficiency of Colchicine-Based Chromosomal Doubling of Maize Haploids

Vijay Chaikam [1], Manje Gowda [1], Leocadio Martinez [2], John Ochieng [1], Hamilton Amoshe Omar [1] and B.M. Prasanna [1,*]



Advantages of Using DH Lines in Maize Breeding

- DH offers quickest route to homozygosity-reduced time in inbreeding
- Increased efficiency of breeding operations
- Efficient use of breeder's time
- Increased phenotypic to genotypic correlations
- Greater amenability to molecular marker applications

Thank you for your interest!

CIMMYT

CGIAR

Genetic diversity and Identification of QTLs for low nitrogen tolerance in doubled haploid maize lines through mid-density genotyping

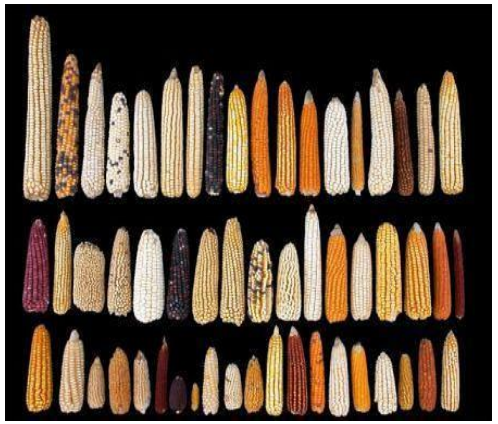
EHEMBA Georgina Lala

Outline

- Introduction
- Objectives
- Materials and methods
- Result and discussion
- conclusion

Introduction

Zea mays L.



Source: https://www.flickr.com/photos/croptrust/4534603184

- Maize is an important cereal crop,
- *Zea mays L*. (2n=2x=20) belong to *poaceae* family**,**
- First most important cereal in SSA,
- High genetic diversity with more than 32,000 genes.

Introduction

Nutritional Value per 100g of Maize

Energy	**360 kJ (86 kcal)**
Carbohydrates	18.7 g
Fat	1.35 g
Protein	3.27 g
Water	75.96 g
Zinc	0.46 mg
Phosphorus	89 mg
Potassium	270 mg
Vitamin C	6.8 mg
Iron	0.52 mg
Magnesium	37 mg

uses

Introduction

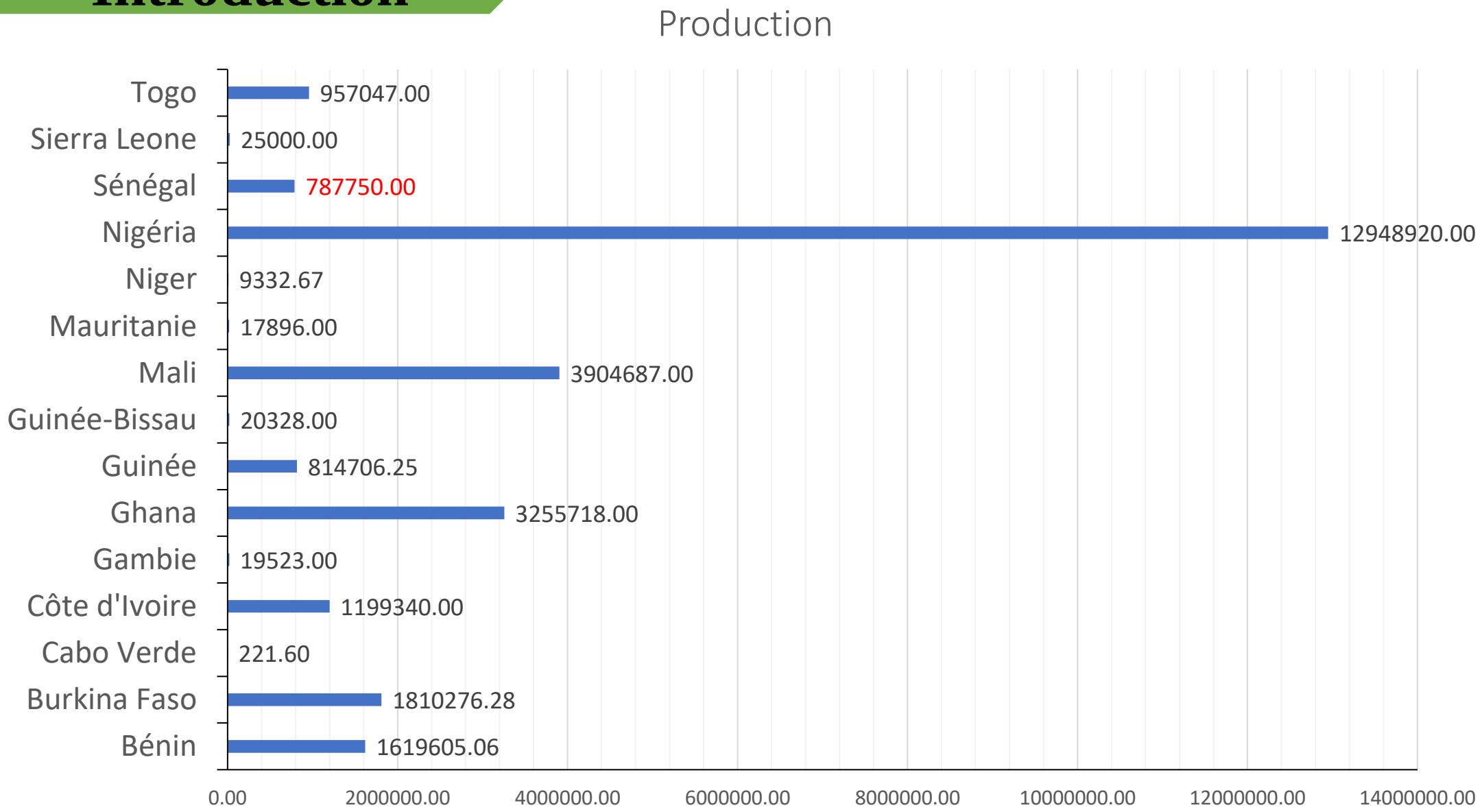
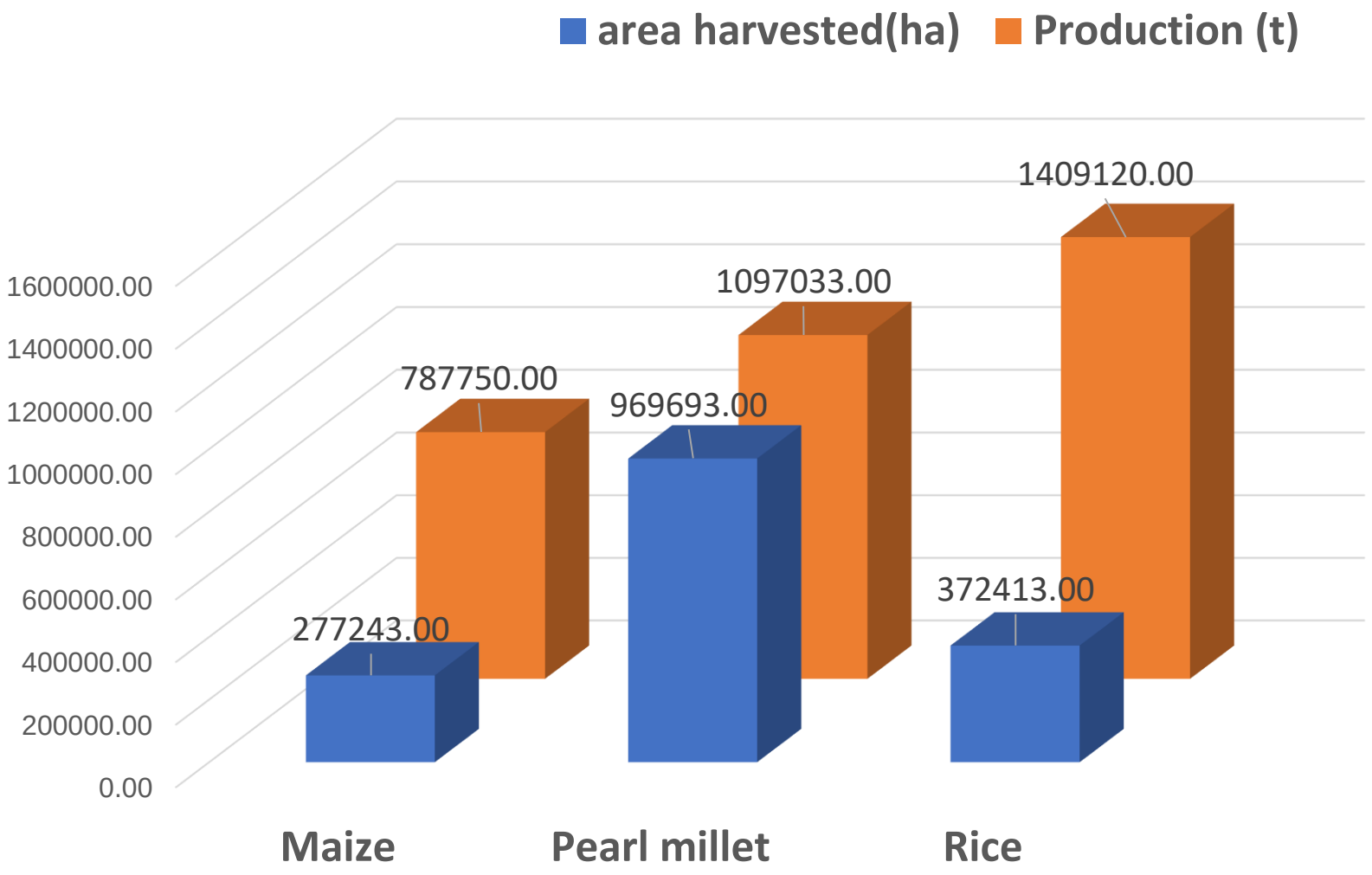


Figure 1: Maize production (t) by West African countries

Introduction

3rd cereal in Senegal in term of area and production after the Rice and Pearl Millet.



Constraints of maize production

➢**Biotic constraints:**

✓ diseases

✓ pests

➢**Abiotic constraints:**

✓ Low soil nitrogen (10-50 % yield lost)

✓ Drought

✓ Heat

✓ Land degradation

Justification

- Low maize productivity in Senegal (2t/ha) with a high importation to meet food and feed demand (AAS, 2019).

- Lack of inbred lines and hybrids with low N tolerance in Senegal

- important to understand the genetic diversity among the newly developed DHLs to inform decision making in our breeding program

- SNPs through mid-density genotyping used to determine the genetic diversity among maize DH lines as well as other crops (Mayer et al., 2022; Semalaiyappan et al., 2023).

- A genome wide association study (GWAS) is needed to identify markers that can help to improve the population through MAS

Research Objectives

General objective: develop high yielding maize hybrids tolerant to drought and low nitrogen conditions

Specific objectives are to:

1. determine the genetic diversity and the population structure among the DH lines,
2. identify candidate genes associated with low N tolerance,

Materials and Methods

development of inbred lines through DH technology

- 5 maize populations:

Obatanpa, SUWAN1, Early Thai (from Senegal), 87036(A), TZE-W DT C4 STR (from IITA).

- 1000 kernels of each population cross with inducers lines at CIMMYT, Kenya.

Chaikam et al.(2019)

Materials and Methods (cont'd)

- **250 DHLs used in this study**
- **DNA extraction and genotyping**
- 10 days old seedlings for DNA extraction
- Intertek Australia
- DArTseq using the mid-density genotyping (3305 SNPs)

- **SNP Calling**
- Tassel software 5.0 was used to map DArTseq readings and detect the SNPs using the maize reference genome.
- MAF = 5% call rate of 80%
- A total of 2135 SNPs after filtering.

Materials and Methods (cont'd)

- **Data analysis**

- R statistical Computing Environment 4.2.1 (R Core Team, 2022) to compute MAF, PIC, GD, kinship matrix.

- The population structure among the DH lines was assessed based on admixture analysis in LEA package.

- The discriminant analysis of principal component (DAPC) in adegenet package (Jombart et al., 2010) used to further explore the genetic diversity.

Materials and Methods (cont'd)

Data Analysis (cont'd)

Candidate gene discovery

- Candidate genes identified via the maize genome data base (Maize GDB) genome browser tool (Schnable *et al.*,2009).

- B73 reference genome assembly version 3 was used

- Putative functional annotations of the candidate genes were retrieved via (NCBI) available in https://www.ncbi.nlm.nih.gov/

- Google used to complete the information given by the NCBI

Materials and Methods (cont'd)

Data Analysis

❑ Association Mapping analysis

- Genomic Association and Prediction Integrated Tool (GAPIT) version 3 integrated in R used to perform marker-trait association analysis.

- The Best Linear Unbiased Estimates (BLUE) computed using "*Lmer4*" in R.

- The Manhattan plot was created in R circular Manhattan plot (R/CMplot) using GAPIT output (Turner, 2018).

- A threshold of –log10 (p-value) = 3.3 was adopted to identify significant associations

Results and Discussion

Table Summary statistics

	MAF	GD	PIC
Minimum	0.05	0.003	0.1
Maximum	0.5	0.44	0.5
Mean	0.2	0.38	0.38

GD = genetic distance; PIC = polymorphism information content; MAF = minor allele frequency

Results and Discussion (cont'd)

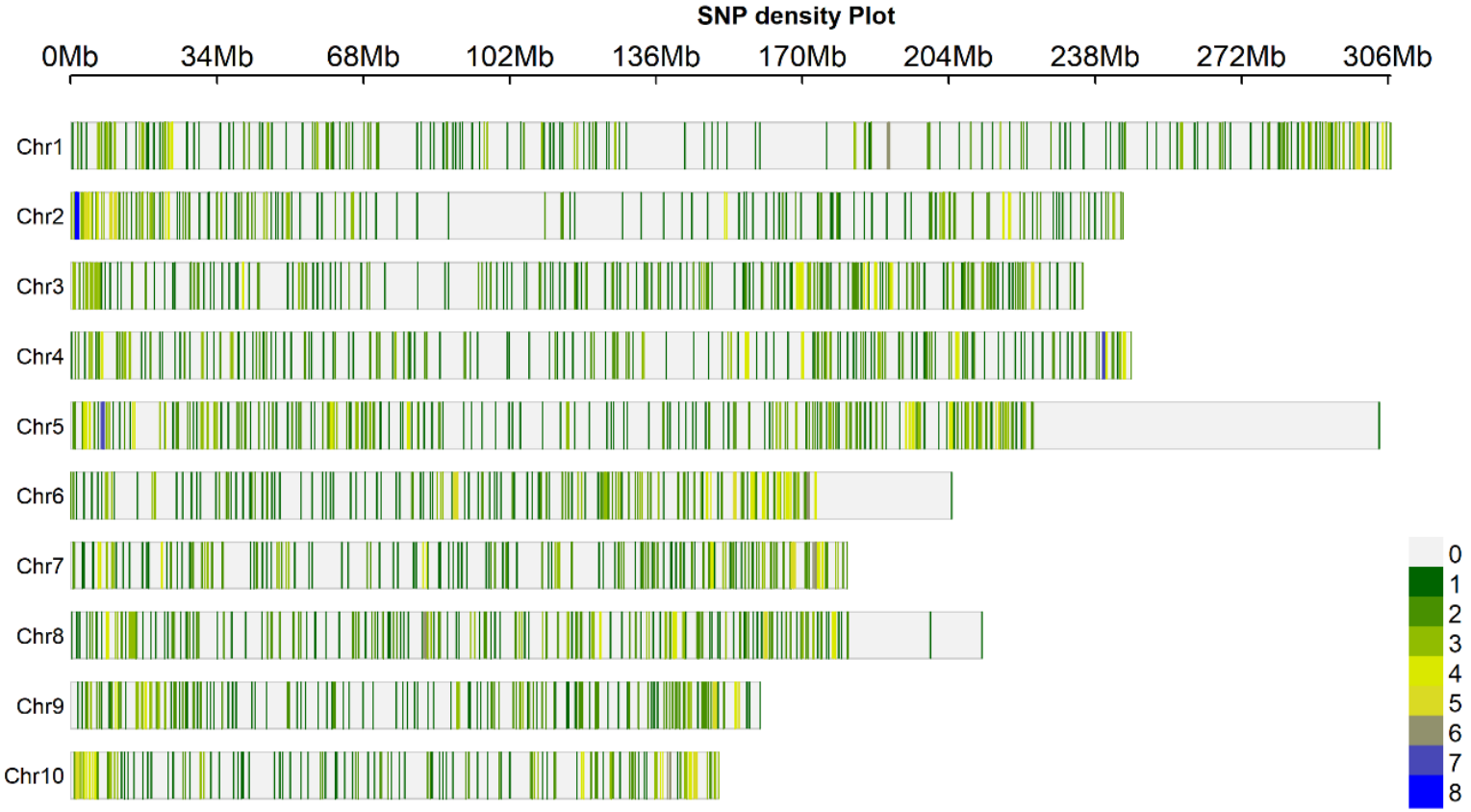


Figure 2: SNP density plot chromosome wise representing number of SNPs within 1Mb window size. The different colours depict SNP density, and the horizontal axis shows the chromosome lengths (mb).

Results and Discussion (cont'd)

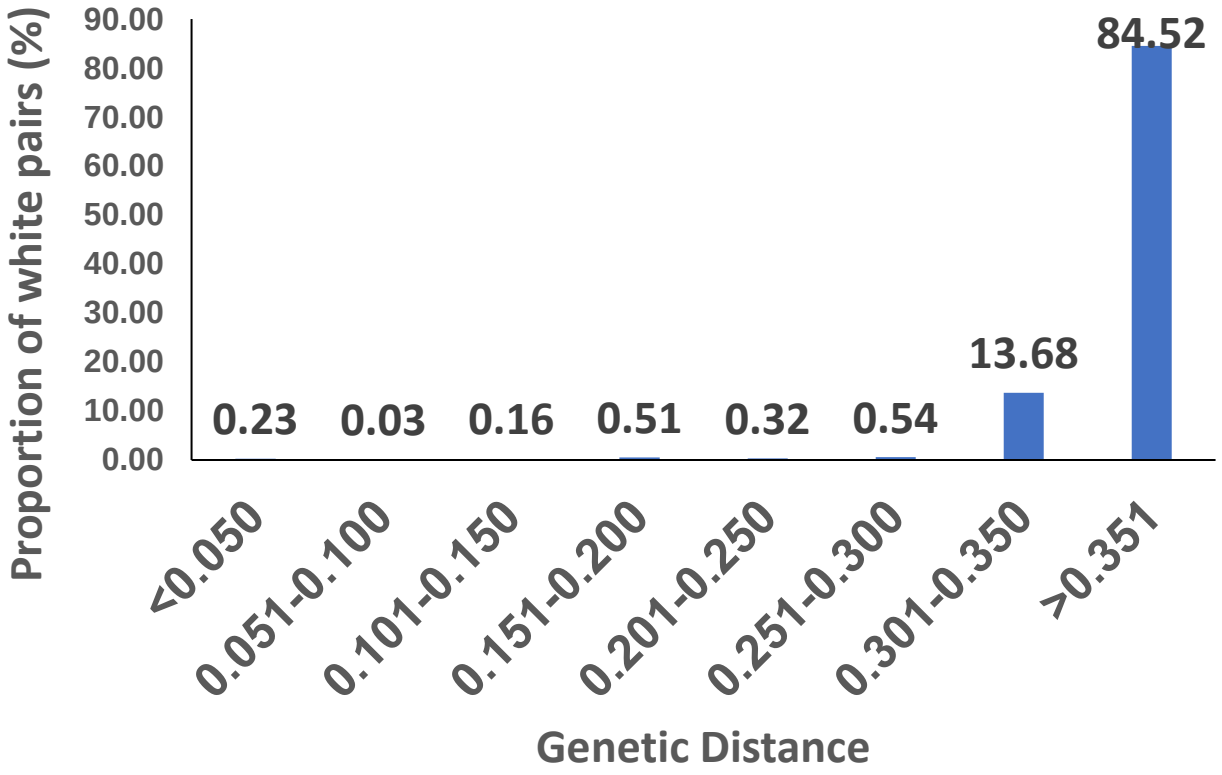


Figure 3. 1: Frequency distribution of the pairwise genetic distance using 2135 SNPs markers of 161 double-haploid white maize lines

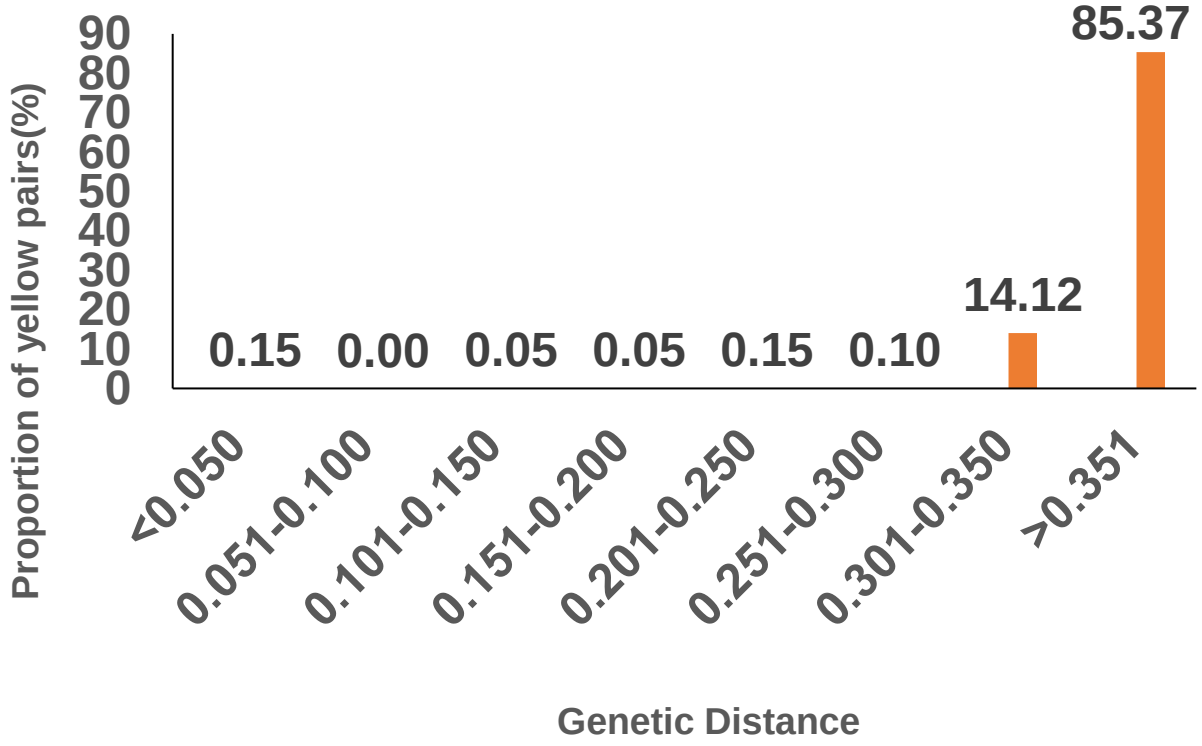


Figure 3. 2: Frequency distribution of the pairwise genetic distance using 2135 SNPs markers of 89 doubled haploid yellow maize lines.

Results and Discussion (cont'd)

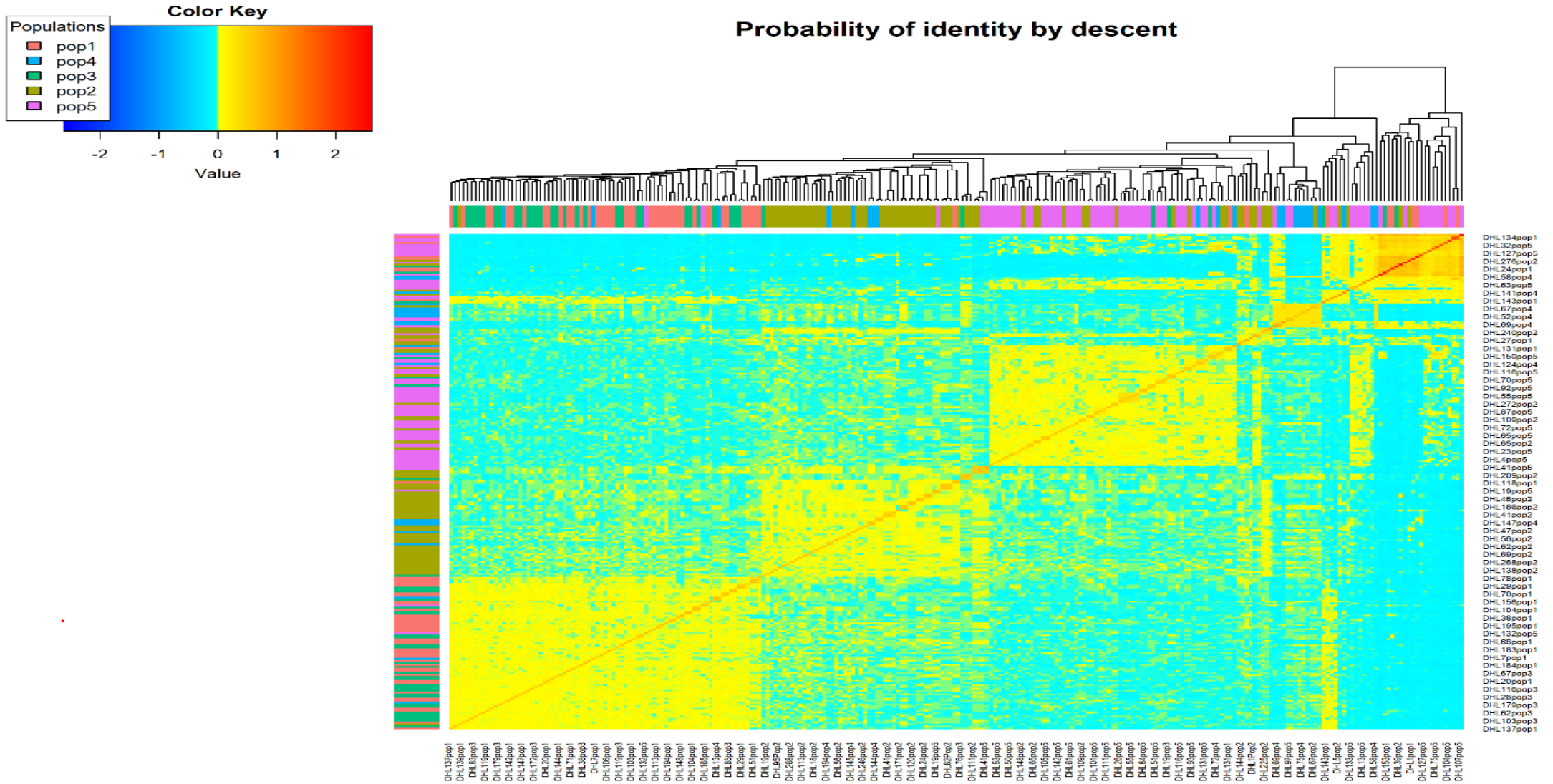
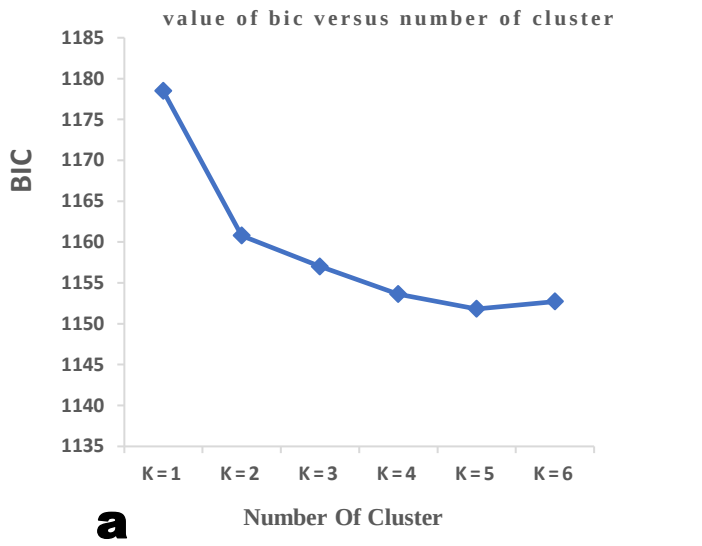
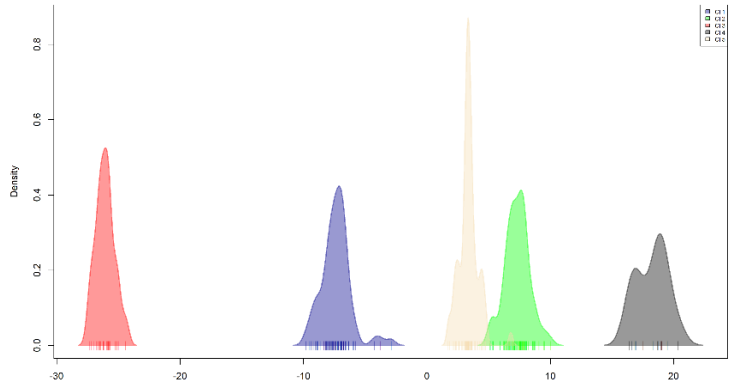


Figure 4: The heat map of SNPs among 250 DH lines based on genetic distances.

Results and Discussion (cont'd)



a



b

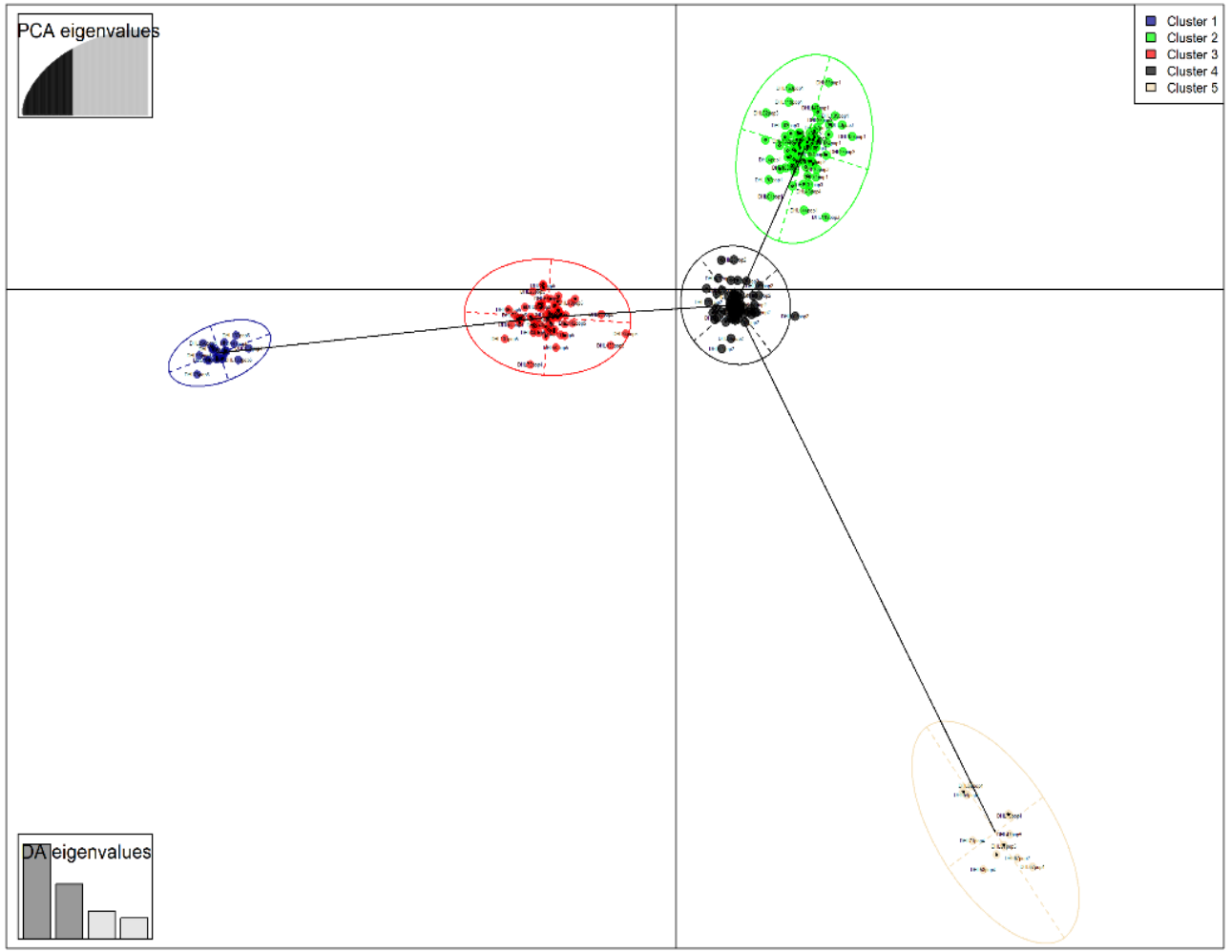


Figure 5: Genetic clustering of 250 DHLs by discriminant analysis of principal components (DAPC); a= bayesien information content; b= discriminant

Results and Discussion (cont'd)

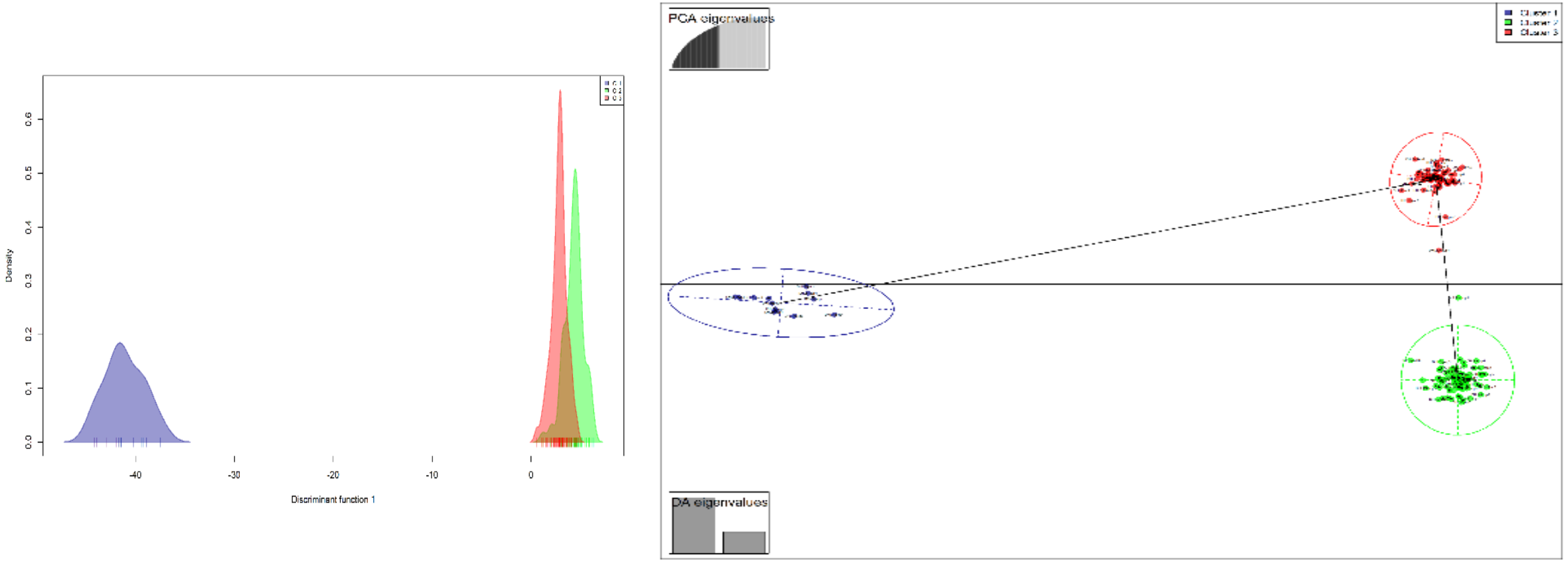


Figure 6: Genetic clustering of the white DHLs by discriminant analysis of principal components (DAPC)

Results and Discussion (cont'd)

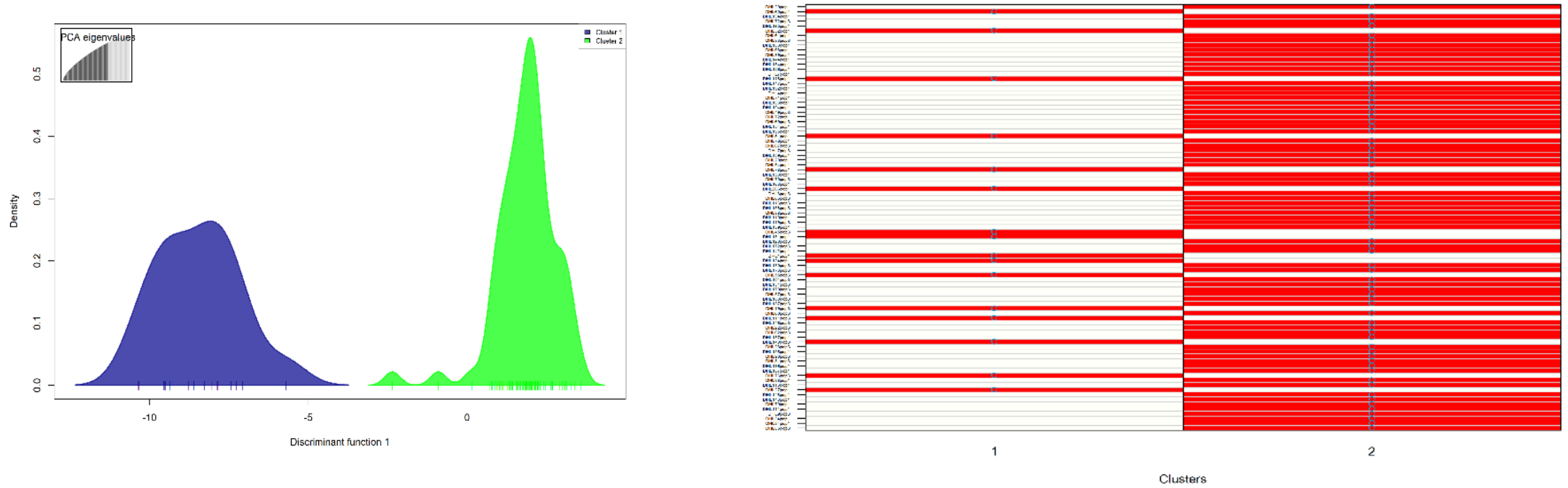


Figure 7: Genetic clustering of the white DHLs by discriminant analysis of principal components (DAPC)

Results and Discussion (cont'd)

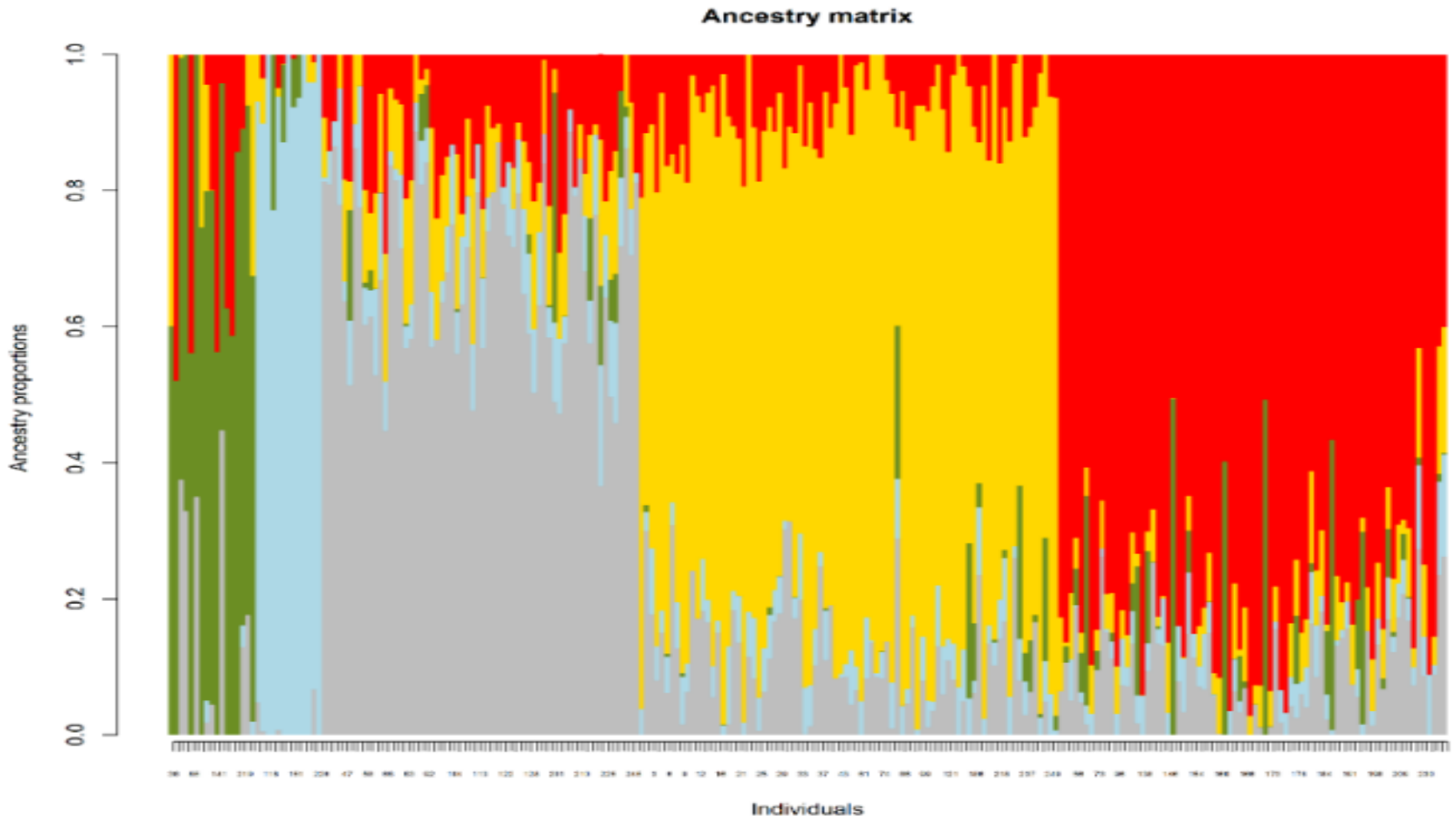


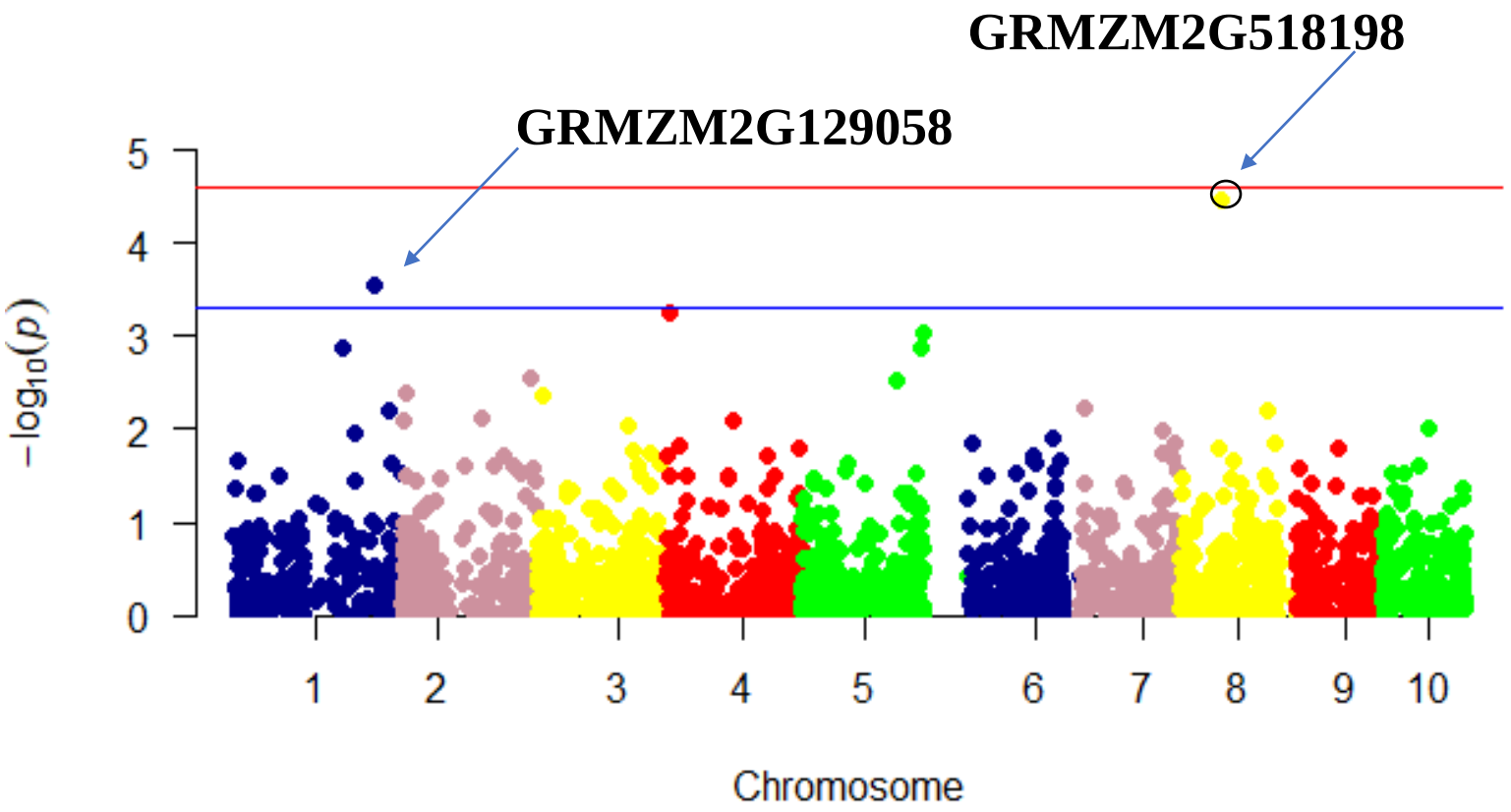
Figure 8: Population structure of 250 doubled haploid lines based on 2135 polymorphic SNP (admixture)

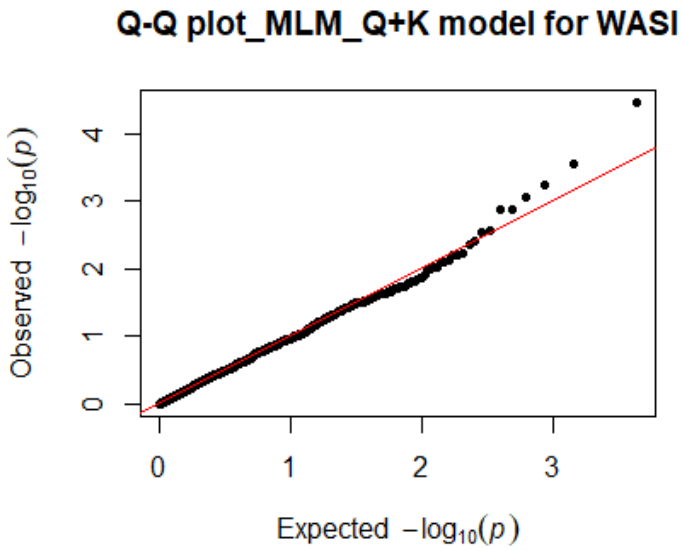
Results and Discussion (cont')

Table 3.3 Association table

Trait	SNP-name	Chr	Position	P-values	R^2	MAF	Minor Allele	Minor Allele effect	Putative Candidate gene	Predicted function of candidate gene
KYield	S1_187589268	1	189779623	9.30E-05	0.08	0.16	C/T	-239.1	GRMZM2G129058	G-type lectin S-receptor-like serine/threonine-protein kinase At2g19130
KASI	PZB01013_1	2	232254011	0.00027	0.09	0.48	A/C	-2.06	GRMZM2G372068	hydroquinone glucosyltransferase
KPASP	PHM2290_12	3	126130837	0.00019	0.15	0.45	A/G	-0.82	GRMZM2G151717	transcription initiation factor TFIID subunit 8
KSTYG	S3_14863959	3	14145145	0.00041	0.1	0.25	A/C	0.02	GRMZM2G352891	RNA binding (RRM/RBD/RNP motifs) family protein
WASI	S8_77392607	8	77392607	6.85E-05	0.08	0.42	G/A	-0.51	GRMZM2G518198	S-type anion channel SLAH3
WASI	PHM16605_19	1	255275067	0.00017	0.07	0.48	A/G	-0.82	GRMZM2G121293	subtilisin-like protease
WSTYG	C7_77617557_v3.22	7	80234401	0.00014	0.08	0.21	C/A	0.34	GRMZM2G031761	FLU (protein fluorescent in blue light chloroplastic)
WSTYG	RS10_83607968	10	83607968	0.00053	0.06	0.3	A/C	0.28	GRMZM2G021277	Tyrosine decarboxylase 1
KEARHARV	C8_141484739_v3.22	8	146131199	0.000173	0.17	0.19	G/A	- 1.31	GRMZM2G044343	putative C3HC zinc finger-like family protein
WEARHARV	PHM3352_19	3	221304526	0.0001729	0.09	0.4	A/T	-0.69	GRMZM2G121878	carbonic anhydrase
WEARHARV	10K06049	6	94664506	0.000427	0.09	0.34	C/T	0.61	GRMZM2G039383	nonspecific lipid-transfer protein
WEARHARV	PZA03612_1	8	133840884	0.000519	0.09	0.16	A/G	-0.63	GRMZM2G330475	transcription factor MYB2
Total R^2					0.98					

Results and Discussion (cont')





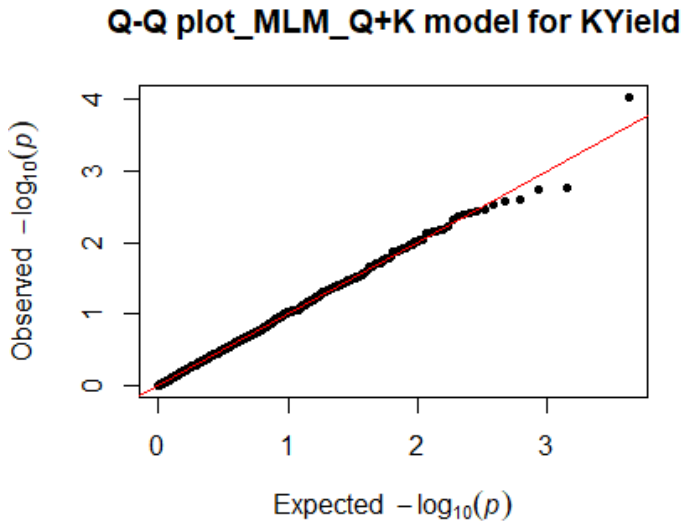
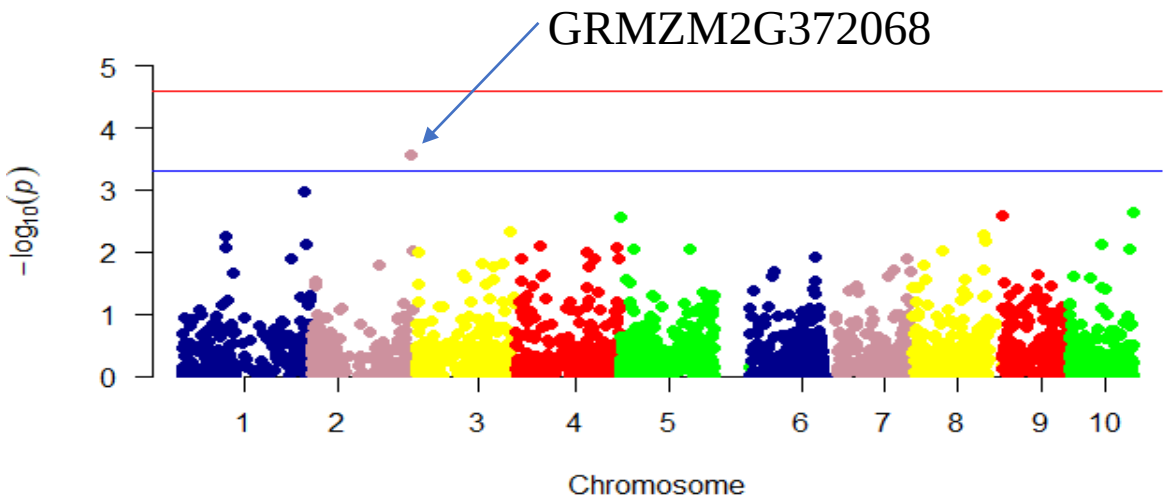
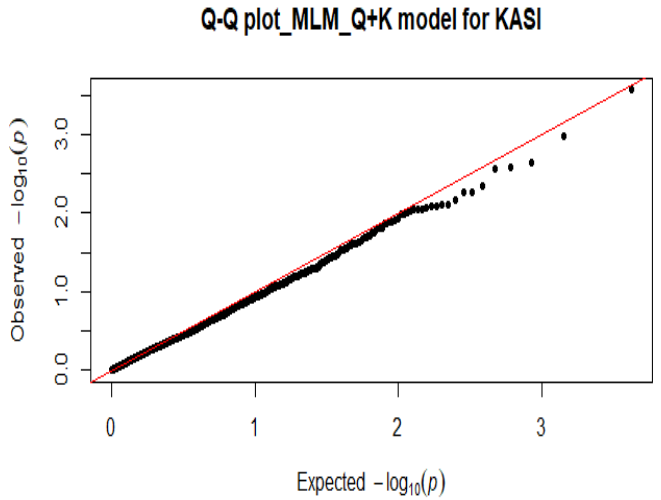
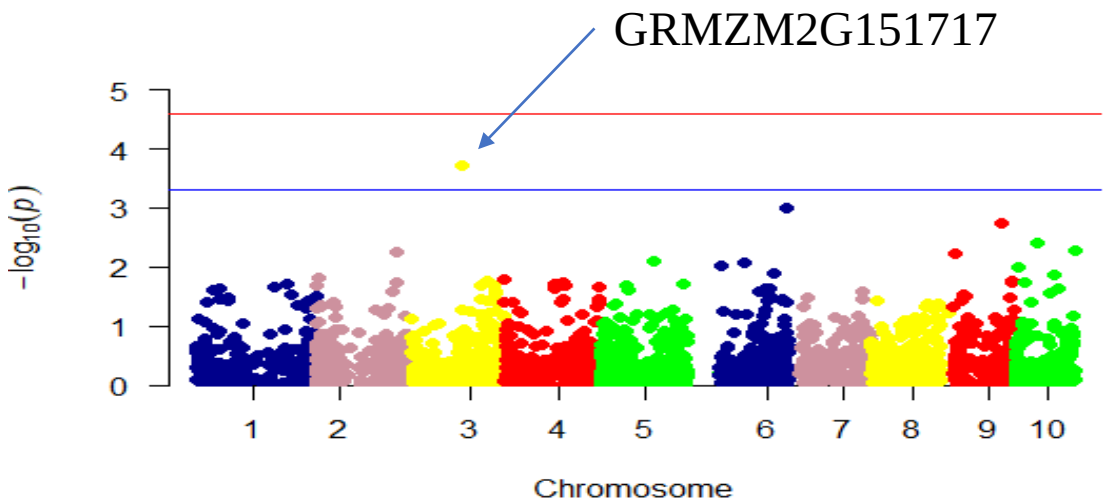


Figure 9: Manhattan plot for anthesis silking interval and yield

Results and Discussion (cont')







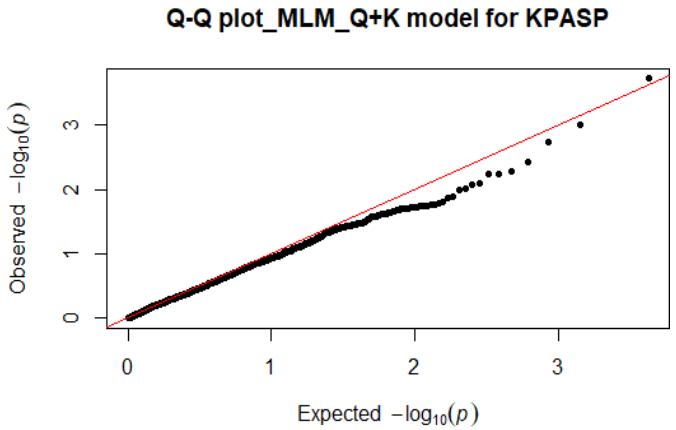


Figure 10: Manhattan plot for anthesis silking interval and plant aspect

Results and Discussion (cont')

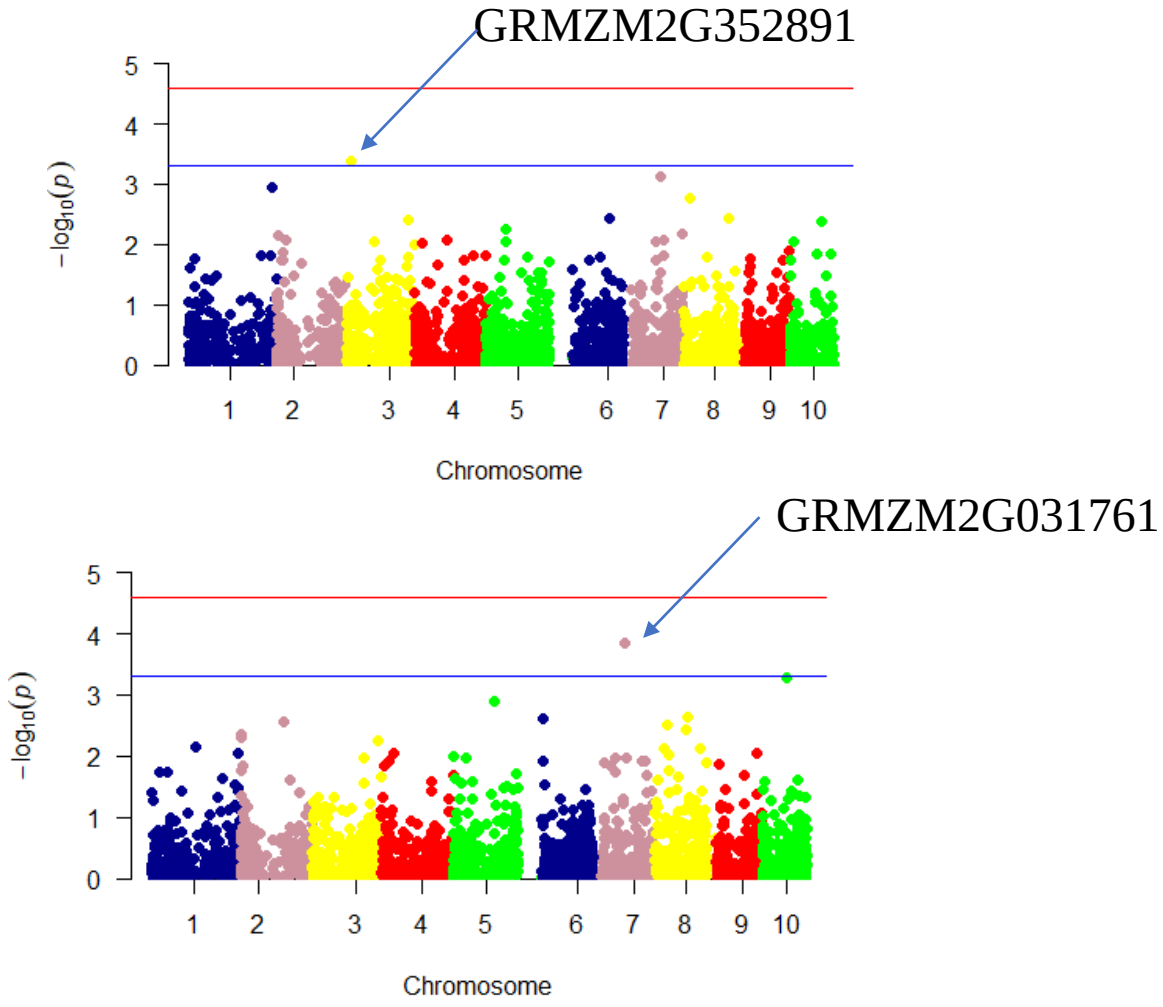
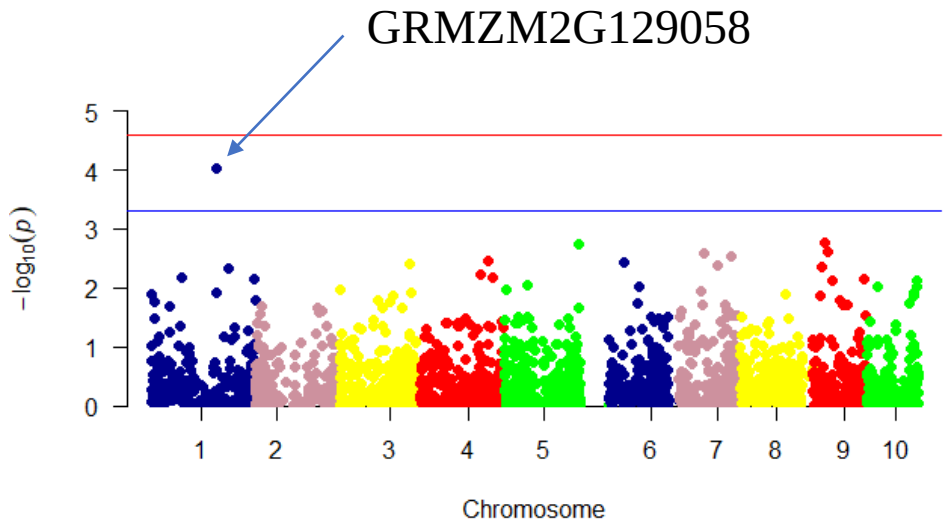


Figure 11: Manhattan plot for stay green

Results and Discussion (cont')



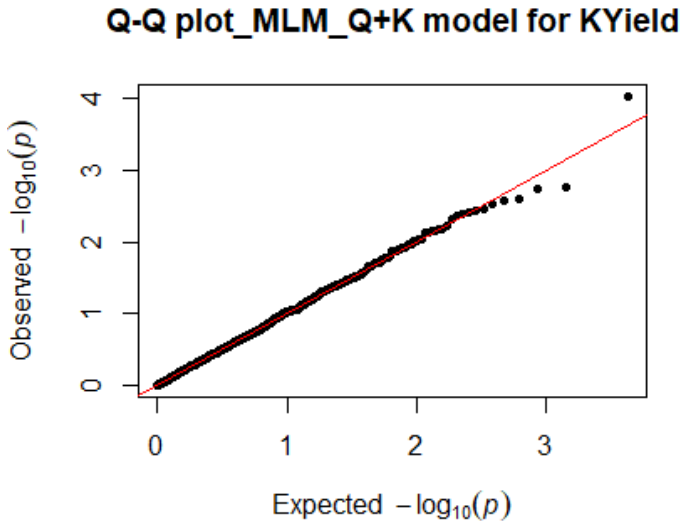


Figure 12: Manhattan plot for yield

Conclusion

- High genetic diversity and population structure were observed within the DH lines.
- Five sub-groups and five clusters were obtained in this study by both the admixture and the DAPC.
- Twelve SNP markers were found associated with the stay green, yield, anthesis silking interval, plant aspect.
- The candidate genes functional annotations revealed that most of them were involved in photosynthesis probably to help the plants to stay green or produce enough chlorophyll under low nitrogen.
- Genes were involved in biotic and abiotic stress tolerances.

Acknowledgements

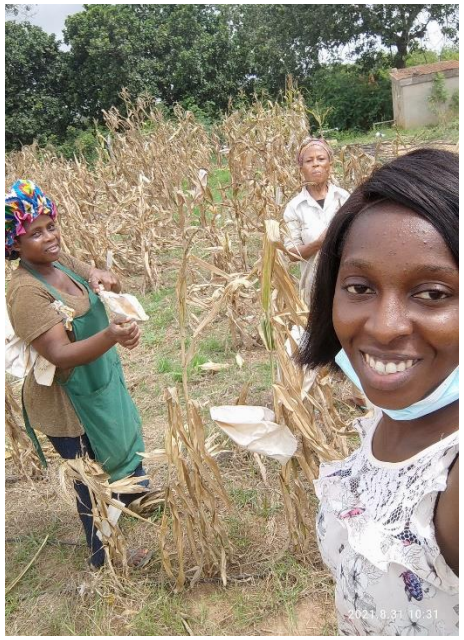
Institut Sénégalais de Recherches Agricoles

CIMMYT

WACCI







THANK YOU