Marker-Assisted Selection for Maize Streak Virus Resistance and Concomitant Conventional Selection for Downy Mildew Resistance in a Maize Population

By

Nothando Fowiza Mafu

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Plant Pathology

2013

Thesis abstract

Maize streak virus (MSV) disease, transmitted by leafhoppers (*Cicadulina mbila*, Naude), and maize downy mildew (DM) disease caused by *Peronosclerospora sorghi* (Weston and Uppal) Shaw, are major contributing factors to low maize yields in Africa. These two diseases threaten maize production in Mozambique, thus the importance of breeding Mozambican maize varieties that carry resistance to these diseases. Marker-assisted selection (MAS) was employed to pyramid MSV and DM disease resistant genes into a single genetic background through simultaneous selection. Firstly, it was essential to determine the genetic diversity of MSV disease resistance in 25 elite maize inbred lines to aid in the selection of suitable lines for the introgression of the *msv1* gene; and subsequently, to introduce the *msv1* resistance gene cluster from two inbred lines, CM505 and CML509, which were identified as the ideal parental lines for the introgression of MSV disease resistance into a locally adapted Mozambican inbred line LP23 that had DM background resistance.

Pyramiding the resistance genes by the use of simple sequence repeat (SSR) molecular markers to track the MSV gene cluster was investigated in 118 F_3 progeny derived from crosses of CML505 x LP23 and CML509 x LP23. High resolution melt (HRM) analysis using the markers umc2228 and bnlg1811 detected 29 MSV resistant lines. At the International Maize and Wheat Improvement Centre (CIMMYT) in Zimbabwe, MSV disease expression of the 118 F_3 progeny lines was assessed under artificial inoculation conditions with viruliferous leafhoppers and the effect of the MSV disease on plant height was measured. Thirty-seven family lines exhibited MSV and DM (DM incidence $\leq$50) disease resistance. Individual plants from a total of 41 progeny lines, that exhibited MSV disease severity ratings of 2.5 or less in both locations within each of the F_3 family lines, were selected based on the presence of the *msv1* gene based on SSR data, or field DM disease resistance, and were then advanced to the F_4 generation to be fixed for use to improve maize hybrids in Mozambique for MSV resistance. Simultaneous trials were run at Chokwe Research Station in Mozambique for MSV and DM disease assessment, under natural and artificial disease infestation, respectively. Thus the MSV and DM genes were effectively pyramided. Lines with both MSV and DM resistance

were advanced to the F_4 generation and will be fixed for use to improve maize hybrids in Mozambique for MSV and DM resistance, which will have positive implications on food security in Mozambique. This research discusses the results of combined selection with both artificial inoculation and the three selected SSR markers. It was concluded that a conventional maize breeder can successfully use molecular markers to improve selection intensity and maximise genetic gain.

Declaration

This MSc study was carried out at the African Centre for Crop Improvement (ACCI), in the School of Agricultural Sciences and Agribusiness, University of KwaZulu-Natal, Pietermaritzburg Campus, under the supervision of Prof. Mark Laing, Prof. John Derera and Dr Roobavathie Naidoo.

1. The research presented in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination any other university.
3. This thesis does not contain other scientists' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other scientists.
4. This thesis does not contain other scientists writing, unless specifically acknowledged a being sourced from other scientists. Where other written sources have been quoted, then:
    a. Their words have been re-written but the general information attributed to them has been referenced;
    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables taken from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references sections.

……..... …….....

Nothando F. Mafu (Candidate) Professor Mark Laing (Supervisor)

…… ……

Professor John Derera (Co-supervisor) Dr Roobavathie Naidoo (Co-supervisor)

Acknowledgements

This study was made possible by a grant from the Generation Challenge Programme (GCP). Thank you, GCP. I would like to express my sincere gratitude to the following: My supervisors Prof. Mark Laing, Prof. John Derera and Dr Roobavathie Naidoo for their supppport, guidance, enthusiasm and encouragement throughout the course of this study. With their enthusiasm, inspiration and support I grew to thoroughly enjoy and appreciate Crop Science and everything it entails. I am extremely indebted to Prof. Laing for having presented me with the opportunity to do an MSc. To Prof. John Derera, your confidence in my abilities has been humbling, thank you. To Dr Roobavathie Naidoo, thank you for the emotional support and for understanding, as only a woman can, of the emotional lows of trying to do your best and not always succeeding. I also want to thank Dr Jedidah Danson for her support and for the role she played in data collection and analysis for this study. I am extremely grateful to Mrs Amanda Madan, Mrs Lesley Brown, Mrs Rowelda Donnelly and the ACCI office for a great job done in organising my finances and the logistics for the trips that I had to make. I would like also to thank Dr Pedro Fato of IIAM-Mozambique, and Dr Cosmos Magorokosho and Mr Amini Mataka of CIMMYT-Zimbabwe for organising resources for the field and greenhouse screening for downy mildew and maize streak virus resistance trials in Mozambique and Zimbabwe, respectively. Part of this research was conducted at the Umbeluzi Research Station of Research Institute of Agriculture (IIAM) of Mozambique and at CIMMYT-Harare. I wish to acknowledge their support for hosting the research. Also, thank you to my best friend Tendekai Mahlanza for his undying support, for having taken time to read my work and going to great lengths to see to it that I had all the requirements I needed to finish this work and for his prayers. I would like to express my deepest appreciation to my family and friends for their undying support. To Mum, your love for me makes me want to be the best I can be. Mostly, thank you God Almighty for giving me the strength.

Table of contents

List of Tables

List of Figures

List of abbreviations

AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
CAPS	Cleaved amplified polymorphic sequence
CBI	Crop Breeding Institute
CIMMYT	International Maize and Wheat Improvement Centre
CIRAD	Agricultural Research Centre for International Development
CML	CIMMYT maize line
CSCE	Conformation sensitive capillary electrophoresis
CTAB	Cetyltrimethylammonium bromide
-dF/dT	Negative derivative
DM	Downy mildew
DMP	Days to mid-pollen shed
DMR	Downy mildew resistance
DMS	Days to mid-silking
DMSR	Downy mildew and maize streak resistant
dsDNA	Double stranded deoxyribonucleic acid
EDTA	Ethylene diaminetetraacetic acid
EMS	Ethyl methanesulfonate
F	Coefficient of inbreeding
FAO	Food and Agriculture Organisation of the United Nations
FTA	Flinders Technology Associates
GCP	Generation Challenge Programme
GGMV	Guinea grass mosaic virus

GLS	Grey leaf spot
GS	Genetic similarity
H_e	Expected heterozygosity
H_o	Observed heterozygosity
HR	Horizontal resistance
HRM	High resolution melt
IIAM	Agricultural Research Institute of Mozambique
IDT	Integrated DNA technologies
IITA	International Institute of Tropical Agriculture
ISSR	Inter-simple sequence repeats
KARI	Kenya Agricultural Research Institute
MALDI-TOF	Matrix-assisted laser desorption/ionization time of flight
MAS	Marker-assisted selection
MC	Major cluster
MCDV	Maize chlorotic dwarf virus
MCMV,	Maize chlorotic mottle stunt virus
MDMV	Maize dwarf mosaic virus
MStV	Maize stripe virus
MSV	Maize streak virus
MSVD	Maize streak virus disease
MYSV	Maize yellow stripe virus
N	Nitrogen
NaCl	Sodium chloride
NAM	Nested association mapping
NCLB	Northern corn leaf blight

NGS	Next-generation sequencing
NPK	Nitrogen: phosphorus: potassium fertiliser
NTSYS	Numerical taxonomy multivariate analysis system
PCR	Polymerase chain reaction
PIC	Polymorphic information content
PVP	Polyvinyl pyrrolidone
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RILs	Recombinant inbred lines
RNA	Ribonucleic acid
SDA	Sodium diethyldithiocarbamic acid
SDS	Sodium dodecyl sulphate
SE	Standard error
SNP	Single nucleotide polymorphism
ssDNA	Single stranded deoxyribonucleic acid
SSR	Simple sequence repeat
TE	Tris-EDTA
T_m	Melting temperature
UKZN	University of KwaZulu-Natal
UPGMA	Unweighted paired group method using arithmetic averages
USA	United States of America

INTRODUCTION TO THESIS

Importance of maize in Africa

Maize (*Zea mays* L.) is a staple food for over 100 million people in Africa (Magenya *et al*., 2009), making it an essential food crop for global food security. In most of Africa's rural economies, at least 85% of maize is used for human consumption, as compared with the developed world where most maize grain is used for animal feed and manufacturing industries (CIMMYT, 1990; Oluwafemi *et al*., 2008; Stevens, 2008). Maize is distributed worldwide and is the world's third highest produced cereal (Sharma and Misra, 2011). Despite this, the average yield per hectare of maize in Africa is the lowest in the world and consequently, food shortages are a perpetual problem in most Sub-Saharan countries (Magenya *et al*., 2008).

Low maize productivity in Africa is thus a major concern that requires urgent attention. According to FAOSTAT (2007), in the year 2007 Nigeria was the leading producer of maize on the African continent, followed by South Africa. This was mainly due to the large area of land dedicated to maize production in these countries (Table 1). Egypt, which was the third largest producer had the highest grain yield per hectare of land in Africa (81 163 hg ha^{-1}), followed by South Africa (28 759 hg ha^{-1}), Ethiopia (27 248 hg ha^{-1}), Malawi (20 400 hg ha^{-1}), Kenya (20 250 hg ha^{-1}), Cameroon (19 229 hg ha^{-1}) and Nigeria (16 595 hg ha^{-1}). The rest of the African countries, including Mozambique, had grain yield below 16 000 hg ha^{-1}. In Mozambique, maize was produced on 1 505 400 hg $ha^{-1}$of land, which is almost two times less land than South Africa. However, Mozambican production is almost five times less than the South African production of 7 338 738 tonnes at 1 579 400 tonnes. There is, therefore, a need to address factors affecting maize production in Mozambique because, as in most of Sub-Saharan Africa, maize production in Mozambique fails to meet the high demand in the country, despite the crop being grown in all the agro-ecological zones (Denic *et al*., 2001). A contributing factor to the low maize productivity in Mozambique is that 70%

of maize production is in the tropical lowland (≤ 800 meters above sea level (masl)) where downy mildew (DM) infection is prevalent (Fato, 2010).

Table 1: Maize production data from the top twenty maize producers in Africa.

Country	**Production (tonnes)**	**Area harvested**	**Yield hg ha^{-1}**
Nigeria	7 800 000	4 700 000	16 595
South Africa	7 338 738	2 551 800	28 759
Egypt	7 045 000	868 000	81 163
Ethiopia	4 000 000	1 468 000	27 248
Malawi	3 444 700	1 688 500	20 400
Tanzania	3 400 000	3 000 000	11 333
Kenya	3 240 000	1 600 000	20 250
Mozambique	1 579 400	1 505 400	10 491
Zambia	1 366 158	872 800	15 652
Uganda	1 262 000	844 000	14 952
DR Congo	1 155 000	1 480 000	7 804
Ghana	1 100 000	750 000	14 666
Zimbabwe	952 600	1 445 800	6 588
Cameroon	923 000	480 000	19 229
Benin	900 000	700 000	12 857
Angola	570 000	1 115 000	5 112
Togo	500 000	380 000	13 157
Chad	200 000	200 000	10 000
Rwanda	90 000	110 000	8 181
Sudan	60 000	80 000	7 500

Source: FAOSTAT, 2007.

Agriculture in Mozambique

Mozambique is located between latitudes 10° 27´ S and 27° 00´ S and longitudes 30° 12´ E and 40° 51´ E of the Greenwich, on the eastern seaboard of Southern Africa (Wulff and Torp, 2005). It is divided into three main agro-ecological zones: The

Southern (Inhambane, Gaza and Maputo), Central (Zambézia, Tete, Manica and Sofala) and the Northern (Niassa, Cabo Delgado and Nampula) (FAO and WFP Report, 2010). The important crops produced in the country include cassava, sorghum, beans, groundnuts, millet, rice and maize. Maize is the primary source of daily calories (44%) followed by cassava (36%) (SADC/FSTAU, 2003). Figure 1 indicates that maize followed closely by cassava is the highest produced crop in Mozambique. According to Wulff and Torp (2005) maize and cassava are the staple foods grown in Mozambique. Maize is produced in all three agro-ecological zones of Mozambique with the northern zone being the largest producer (FAO and WFP Report, 2010). Despite maize being a staple crop produced in all agro-ecological zones of the country, production still fails to meet domestic demand (Wulff and Torp, 2005). There is, therefore, a need for interventions to enhance production and close the gap between consumption and production.

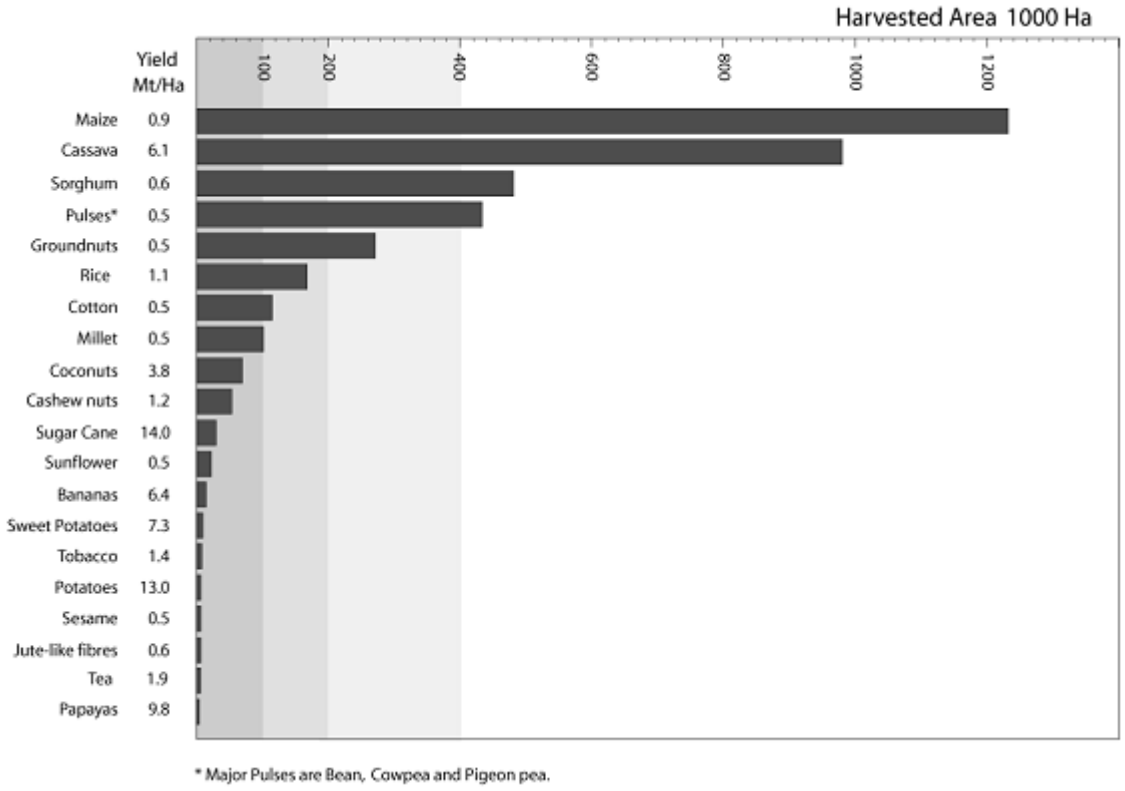

Figure 1: Yields of major crops harvest from 1 000 ha in Mozambique (2000-2004). (Source: FAOSTAT, 2004).

Agriculture in Mozambique is dominated by the rural populations which make up 70% of the country's population making it the most significant livelihood source (Wulff and Torp, 2005; FAO and WFP Report, 2010). Smallholders have limited access to capital thus typically practice cheap traditional farming methods like crop-rotation and early planting to avoid diseases and little or no purchased inputs and irrigation (FAO and WFP Report, 2010). While lack of access to modern farming techniques can prove to be a constraint on maize production, biotic factors like pests and diseases, further decrease yields. Furthermore, the high temperatures and humid tropical areas in Africa make some maize pests and diseases prevalent which negatively impacts maize yields (Fajemisin, 2003). Diseases especially prevalent in lowland areas and southern regions of Mozambique include DM, stem rot, ear rot, maize streak virus (MSV) and pests that include stem borers and grain weevils which cause massive losses in yields yearly (Denic *et al*., 2001; FAOSTAT, 2004; Fato *et al*., 2008).

Downy mildew (DM) disease in Mozambique

Downy mildew disease, caused by *Perenosclerospora sorghi* (Weston and Uppal) C.G. Shaw, infects both sorghum (*Sorghum bicolor* L.) and maize (Cardwell *et al*., 1997; Jeger *et al*., 1998; Bock *et al*., 2000) and has been attributed to 50-100% yield losses in Mozambique (IIAT, 1999). Denic *et al.* (2001) reported it as the most destructive foliar disease of maize in the lowland areas in central and southern Mozambique. In Mozambique, DM infection was reported on both sorghum and maize from regions with 400-800 mm rainfall per annum (Jeger *et al*., 1998). Conditions favourable for *P. sorghi* spore growth and the spread of DM are mild temperatures in a humid environment (Ahlawat, 2007). These conditions are prevalent in Mozambique, making the country prone to devastating losses when DM susceptible varieties are grown (IITA, 1999). For improved production of maize, use of resistant lines is the best control approach. Using resistant lines allows poor farmers to reduce the use of chemical pesticides or fertilisers which though popular because of their quick effective action cause great risks to the environment and human health (Adejumo, 2005).

Maize streak virus (MSV) disease in Mozambique

Numerous viral pathogens including MSV, maize chlorotic mottle stunt virus (MCMV), maize eyespot virus, guinea grass mosaic virus (potyvirus) and maize yellow stripe virus (MYSV) infect maize in Africa and reduce maize production (Thottappilly *et al*., 1993; Martin and Shepherd, 2009). The leafhopper-vectored MSV is considered as the most significant biological threat to food security in Sub-Saharan Africa (Thottappilly *et al*., 1993; Martin and Shepherd, 2009). It is the most prevalent viral disease, undermining the economic wellbeing of subsistence farmers throughout Africa (Bosque-Perez, 2000; Shepherd *et al*., 2007; Martin and Shepherd, 2009). Yield losses due to MSV disease range from a trace to virtually 100% when the virus attacks susceptible lines (Tefera *et al*., 2011). In Mozambique, MSV disease is prevalent in all maize production areas (Denic *et al*., 2001). The diagnosis and characterisation of MSV disease resistant maize populations is central to the breeding and selection of MSV resistant cultivars to control the disease.

Breeding for resistance and genetic diversity

Maize hybrid improvement for DM and MSV resistance is critical for Mozambique in order to support the large impoverished rural population. Characterisation of genetic diversity and similarities of maize inbred lines enables maximum efficiency in the determination of the best possible combination of parents for the development of new and improved inbred lines (Xia *et al.*, 2005; Legesse *et al*., 2007). Marker-assisted selection (MAS) is a biotechnology research tool adapted to enhance conventional breeding with accuracy and to accelerate variety development (Xu and Crouch, 2008). With regard to this, molecular markers, such as simple sequence repeats (SSRs) and inter-simple sequence repeats (ISSRs), are able to define genetic relationships of inbred lines at DNA level (Xia *et al*., 2005). The study aimed to contribute to the development of hybrids that have both DM and MSV disease resistance in Mozambique.

Global objective of the study

Breeding of MSV resistant maize is the best control measure for MSV disease in Mozambique. The MSV resistance gene cluster and flanking marker genes have been identified by the International Maize and Wheat Improvement Centre (CIMMYT) (CIMMYT, 2009). The goal of this study was, therefore, to introgress the MSV resistant gene cluster from CIMMYT MSV disease resistant lines CML505 and CML509 into the elite Mozambican maize line LP23, tracking the transfer of the gene cluster using MAS. Line LP23 has DM resistance and is adapted to the lowland areas in Mozambique but is susceptible to MSV disease.

Specific objectives

The specific objectives of the study were as follows:

1. To determine the genetic diversity in 25 elite maize inbred lines to aid in the selection of suitable lines for the introgression of the *msv1* gene (MSV disease resistant gene) enabling the production of the best possible MSV and DM disease resistant hybrids for Mozambique.
2. To evaluate the effectiveness of using MAS to transfer MSV resistance genes from CIMMYT donor lines (CML505 and CML509) into the selected Mozambican lines by evaluating the F_3 progeny.
3. To determine the effects of MSV disease on growth of the progeny, with emphasis on height of the infected maize plants.
4. To identify progeny lines that combined both MSV and DM disease resistance for potential use in developing MSV and DM resistant hybrids for Mozambique.

Research hypotheses

The following research hypotheses were tested in the thesis:

a) There is adequate genetic diversity among the elite lines in Mozambique which can be exploited in a breeding programme to generate MSV and DM disease resistant hybrids;

b) Molecular MAS can be effective in identifying lines that are resistant to MSV disease;

c) Resistance to MSV and DM diseases can be combined in a single inbred line and such lines can be obtained through simultaneous selection for MSV resistance and DM in one base population.

Thesis structure

The thesis structure is as follows:

Introduction to thesis

Chapter 1: Literature review

Chapter 2: Genetic diversity of maize germplasm lines and implications for breeding maize streak virus resistant hybrids.

Chapter 3: Detection of SSR markers linked to MSV disease resistance and high resolution melt (HRM) analysis of F_3 maize population samples stored on Whatman FTA^{TM} Elute cards.

Chapter 4: Phenotypic characterisation of progeny maize lines for maize streak virus (MSV) and downy mildew (DM) resistance.

Chapter 5: Genetic diversity among maize lines selected for downy mildew and maize streak virus resistance as determined by SNP markers.

Research overview

References

Adejumo, T.O. 2005. Crop protection strategies for major diseases of cocoa, coffee and cashews in Nigeria. African Journal of Biotechnology 4:143-150.

Ahlawat, Y.S. 2007. Plant Pathology Crop Diseases and their Management. Indian Agricultural Research Institute, New Delhi. pp. 98.

Bock, C.H., M.J. Jeger, K.F. Cardwell, L.K. Mughogho, and J. Sherington. 2000. Control of sorghum downy mildew in maize and sorghum in Africa. Tropical Science 40:47-57.

Bosque-Perez, N.A. 2000. Eight decades of maize streak virus research. Virus Research 71:107-121.

Cardwell, K.F., J.G. Kling, and C.H. Bock. 1997. Methods for screening maize against downy mildew *Peronosclerospora sorghi.* Plant Breeding 116:221-226.

CIMMYT. 1990. CIMMYT world maize 1989/90: Facts and Trends, Realizing the Potential of Maize in Sub-Saharan Africa. CIMMYT, Mexico. http://pdf.usaid.gov/pdf_docs/PNABT192.pdf (Accessed 15 Nov. 2010).

CIMMYT. 2009. CIMMYT Maize Inbred Lines (CML). CIMMYT International Maize and Wheat Improvement Center, Mexico. http://www.cimmyt.org/en/component/content/article/459-international-maize-improvement-network-imin/434-cimmyt-maize-inbred-lines-cml (Accessed 29 Oct. 2010).

Denic, M.P., C. Chauque, M. Jose, D. Langa, Mariote, P. Fato, and W. Haag. 2001. Maize screening for multiple stress tolerance and agronomic traits. Eastern and Southern Africa Regional Maize Conference 7:88-91.

Fajemisin, J.M. 2003. Overview of maize viruses in sub-Saharan Africa. In: J.A. Hughes and B.O. Odu (eds.) Plant Virology in Sub-Saharan Africa, International Institute of Tropical Agriculture Conference. Ilesa, Nigeria. pp. 158-171.

FAO and WFP Report. 2010. Special report: Crop and Food Security Assessment Mission to Mozambique. Global Information and Early Warning Systems on Food and Agriculture. http://www.fao.org/docrep/012/ak350e/ak350e00.htm (Accessed 4 Jan. 2011).

FAOSTAT. 2004. Mozambique – Crops. http://faostat.fao.org/site/611/default.aspx. (Accessed 4 Aug. 2011).

FAOSTAT. 2007. http://faostat.fao.org/site/567/default.aspx (Accessed 17 Aug. 2010).

Fato, P. 2010. Investigation of heterotic patterns and genetic analysis of downy mildew resistance in Mozambican lowland maize (*Zea mays* L.) germplasm. African Crop Science Society Conference, Maputo, Mozambique. http://www.acss2011moz.uem.mz/doc/announcements/Volume%20of%20Abstracts%20&%20Programme.pdf (Accessed 16 Jul. 2011).

Fato P., P.S. Chauque, D. Cugala, and E. Carvalho. 2008. The status of development of maize resistance to field and storage pests in Mozambique - Consolidating

Experiences from IRMA I and II: Achievements, lessons and prospects. IRMA Project End-of-Phase II Conference, Nairobi, Kenya.

IITA. 1999. Stopping the scourge. IITA Annual Report 1999. http://www.iita.org/info/ar99/8-9.htm. (Accessed 5 Oct. 2010).

Jeger, M.J., E. Gilijamse, C.H. Bock, and H.D. Frinking. 1998. The epidemiology, variability and control of the downy mildews of pearl millet and sorghum, with particular reference to Africa. Plant Pathology 47:544-569.

Legesse, B.W., A.A. Myburg, K.V. Pixley, and A.M. Botha. 2007. Genetic diversity of African maize inbred lines revealed by SSR markers. Hereditas 144:10-17.

Magenya, O.E.V., J. Mueke, and C. Omwega. 2008. Significance and transmission of maize streak virus disease in Africa and options for management. African Journal of Biotechnology 7:4897-4910.

Magenya, O.E.V., J. Mueke, and C. Omwega. 2009. Association of maize streak virus disease and its vectors (*Homoptera: Cicadelidae*) with soil macronutrients and altitudes in Kenya. African Journal of Agricultural Research 4:1284-1290.

Martin, D.P., and D.N. Shepherd. 2009. The epidemiology, economic impact and control of maize streak disease. Food Security 1:305-315.

Oluwafemi, S., G. Thottappilly, and M.D. Alegbejo. 2008. Search for variability among mastreviruses showing symptoms of maize streak virus from different regions of Nigeria by polymerase chain reactions. Journal of Plant Protection Research 48:1-6.

SADC/FSTAU. 2003. SADC Brief No. 2, Harare, Zimbabwe.

Sharma, K., and B.S. Misra. 2011. Molecular approaches towards analyzing the viruses infecting maize (*Zea mays* L.). Journal of General and Molecular Virology 3:1-17.

Shepherd, D.N., T. Mangwende, D.P. Martin, M. Bezuidenhout, J.A. Thomson, and E.P. Rybicki. 2007. Inhibition of maize streak virus (MSV) replication by transient and transgenic expression of MSV replication-associated protein mutants. Journal of General Virology 88:325-336.

Stevens, R. 2008. Prospects for using marker-assisted breeding to improve maize production in Africa. Journal of the Science of Food and Agriculture 88:745-755.

Tefera, T., S. Mugo, Y. Beyene, H. Karaya, and R. Tende. 2011. Grain yield, stem borer and disease resistance of new maize hybrids in Kenya. African Journal of Biotechnology 10:4777-4783.

Thottappilly, G., N.A. Bosque-Perez, and I.I.W. Rossel. 1993. Viruses and virus diseases of maize in tropical Africa. Plant Pathology 42:494-509.

Wulff, E. and J. Torp. 2005. Seed sector country profile: Mozambique-Overview of Seed Supply Systems and Seed Health Issues. Danish Seed Health Centre for Developing Countries, Copenhagen, Denmark. http://www.dshc.life.ku.dk/Publications/~/media/migration%20folder/upload/shc/docs/pdf/country%20profile%20moz.pdf.ashx (Accessed 18 Nov. 2010)

Xia, X.C., J.C. Reif, A.E. Melchinger, M. Frisch, D.A. Hoisington, D. Beck, K. Pixley, and M.L. Warburton. 2005. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers: Subtropical, tropical mid-altitude, and highland maize inbred lines and their relationships with elite U.S. and European maize. Crop Science 45:2573-2582.

Xu, Y., and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: From publications to practice. Crop Science 48:391-407.

CHAPTER ONE

Literature review

1.1 Introduction

The following aspects are reviewed: 1) the importance of maize and the constraints associated with it in Mozambique, and in Africa; 2) the biology, symptoms, geographic locations and overall damage caused by maize streak virus (MSV) disease and downy mildew (DM) disease; 3) control methods for MSV and DM diseases, with emphasis on breeding for resistant maize genotypes as the best control measure; 4) comparing classical breeding and marker-assisted breeding; 5) reviewing the advantages and disadvantages of the different DNA markers utilised in marker-assisted selection (MAS), which include; restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNA (RAPDs), amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs); 6) discussing how high resolution melt (HRM) analysis can be used in this study for the identification of genotypes; and 7) discussing Flinders Technology Associates (FTA^{TM}) card technology as a tool for DNA sampling.

1.1.1 Maize and its importance in Africa

Maize, a member of the grass family *Poaceae* to which all major cereals belong, is parallel in importance in Africa to wheat in the Middle East and rice in Southeast Asia (Fajemisin, 2003). Approximately 15 million ha of maize are planted annually in Sub-Saharan Africa, mainly in rural areas, thus making it the mainstay of the continent's rural economies (Boomsma and Vyn, 2008). The International Food Policy Research Institute (2000) projected the annual maize demand in Sub-Saharan Africa to be 500 million tons by the year 2020, which will surpass the demand for both wheat and rice. In order to stabilise and increase global maize production for a rapidly growing world population, the development of maize varieties with enhanced disease tolerance,

amongst other factors, continues to be an important objective (Boomsma and Vyn, 2008).

1.1.2 Production constraints of maize in Africa

The rising demand for maize presents an urgent challenge for the developing countries of Africa. According to Denic *et al.* (2001) and the FAO and WFP Report (2010), in southern Mozambique the main constraints to maize production are: lack of adequate rain, the diseases MSV and DM, and pests such as borers and storage pests. Maize streak virus and downy mildew diseases are the major problems in the Manica and Sofala provinces. In the central and northern parts of Mozambique, however, the most important limitations, as in the rest of Africa, are low soil fertility and periodic droughts (Denic *et al*., 2001). Shortage of trained manpower, management expertise and poor cultivation practices such as inadequate intercropping, poor soil preparation, poor irrigation techniques, poor weeding and poor planting practices also hamper agricultural productivity (Odendo *et al*., 2001).

A survey carried out in the Siaya District of western Kenya revealed that while maize yield is between 0.5 - 0.7 t ha^{-1} in the area, in farm trials it can increase to 1.4-1.6 t ha^{-1} and this can be achieved by applying fertiliser and using improved maize varieties (Odendo *et al*., 2001). There is generally low adoption of technologies that improve productivity such as improved seed in Africa because these technologies come at much higher prices than can be afforded by the low income farmers (Odendo *et al*., 2001). Financial constraints limit access to credit facilities and agricultural extension services (Odendo *et al*., 2001; Fabiyi *et al*., 2007). The low adoption of productivity improving technologies is also associated with a lack of knowledge and understanding on the part of farmers, some of who still believe these technologies are inappropriate together with the application of fertilisers (Achieng *et al.*, 2001; Odendo *et al*., 2001). To increase the likelihood of adoption, variety preferences by farmers need to be taken into consideration when breeding for new improved varieties. By meeting the farmer's needs in terms of maize varieties and crop management packages provided, the large yield gap between farmers and researchers (Brink and Belay, 2006) can be closed.

1.1.3 Major biotic constraints

In addition to the use of unimproved production technologies and lack of access to modern farming techniques, there are also constraints like maize pathogen epidemics that further decrease maize yields. High temperatures and humid tropical conditions of some regions of Africa make some maize diseases and pests prevalent which take a severe toll on maize yield (Fajemisin, 2003).

About 32 viruses have been recorded to infect maize; however, only seven have been reported to occur in tropical Africa. These are MSV, maize stripe virus (MStV), maize chlorotic dwarf virus (MCDV), maize mottle virus, maize eyespot virus, maize dwarf mosaic virus (MDMV), and Guinea grass mosaic virus (GGMV) (Thottappilly *et al*., 1993). Of these seven viruses, the leafhopper-vectored MSV is considered the most important (Bosque-Perez, 2000**)**. The MSV disease is the most widely studied due to its high yield loss potential (Bosque-Perez, 2000; Shepherd *et al*., 2007).

1.2 Maize streak virus disease

Maize streak virus disease (MSVD) is the most widespread of all the maize viral diseases found in Africa (Harkins *et al*., 2009). Under favourable conditions, yield losses as a result of MSVD have been reported to be 100% (Sharma and Misra, 2011). These devastating low yields dangerously undermine the social development of the world's poorest people in Sub-Saharan Africa (Owor *et al*., 2007; Harkins *et al*., 2009). Martin and Shepherd (2009) stated that globally MSVD is considered to be the third most disastrous disease of maize after grey leaf spot (GLS) and northern corn leaf blight (NCLB) diseases. Economically MSVD is the most damaging disease of far greater consequence than both NCLB and GLS in Africa and its neighbouring islands where it is found (Thottappilly *et al*., 1993; Fakorede *et al*., 2003). Maize streak virus disease is therefore a major constraint on maize production and contributes to poverty throughout the tropical regions south of the Sahara (Martin and Shepherd, 2009).

1.2.1 Maize streak virus biology

MSV is a species of the genus *Mastrevirus* of the family *Geminiviridae* (Zhang *et al*., 2001). Previously characterised geminiviruses are known to have two circles of DNA while the MSV DNA virus is one DNA circle of 2 687 nucleotides (Mullineaux *et al*., 1984). It is obligately transmitted by leafhoppers of the genus *Cicadulina,* which feed on over 80 species of monocotyledonous plants belonging to the *Poaceae* family (Bosque-Perez, 2000; Harkins *et al*., 2009). These viruses have single-component single-stranded circular DNA genomes of 2.7 kb in size, and characteristic 'twinned' or geminate particles (Willment *et al*., 2001).

1.2.2 Maize streak virus disease geographic distribution

MSV is accepted as an endemic African virus that is confined to the African continent and its neighbouring islands (Bosque-Perez, 2000). It is a significant maize disease in the southern countries namely: Mozambique, Zimbabwe, South Africa, Kenya and Zaire and in the eastern and western countries of Africa (Fajemisin, 2003; Magenya *et al*., 2008).

1.2.3 Maize streak virus disease symptoms

Fuller (1901) stated that MSVD was first described in 1901 in Southern Africa and it has been affecting maize since the 1870s. The condition was first described as 'maize variegation' and was later renamed 'maize streak' by Storey in 1925 (Bosque-Perez, 2000). Symptoms of MSVD begin as spherical chlorotic spots which later coalesce into continuous longitudinal narrow streaks (Fig. 1.1).

Figure 1.1: Maize streak virus disease. Close-up of maize streak virus chlorotic streak symptoms and leafhoppers on a mature maize leaf. Photo by: Nothando F. Mafu, PMB, UKZN. Taken at CIMMYT Zimbabwe. Date: 16/12/10.

These narrow streaks are mainly along the veins of the leaf laminae and are distributed uniformly over the leaf surface (Magenya *et al*., 2008). The streaking pattern on the leaves corresponds to the existence of the virus whilst the density of streaking depends on varietal susceptibility. Chlorosis of the entire lamina can develop from the chlorotic streaking in highly sensitive varieties (Thottappilly *et al.*, 1993; Bosque-Perez, 2000; Fajemisin, 2003). Chlorosis is followed by the premature death of the plant, particularly if infection has occurred at an early stage of plant growth. The effects of MSVD on grain yield are most prominent when the infection takes place in young plants, and decreases with increased plant age (Thottappilly *et al.*, 1993; Bosque-Perez, 2000).

1.2.4 Damage caused by maize streak virus disease

Maize plants are vulnerable to the MSVD from emergence to tasseling and often infection at seedling stage results in no ear formation (Magenya *et al*., 2008). Infection at a later stage results in undersized and poorly filled ears (Fajemisin, 2003). If infection of a maize crop is in the first three weeks after planting, often this can result in 100% yield loss (Magenya *et al*., 2008). Infection at the 6-8 week stage after planting has little effect on the vigour of the plant (Fajemisin, 2003). In susceptible varieties yield reductions often exceed 70% depending on the stage of plant maturity when infection occurs (Magenya *et al*., 2008). Consequently effective control measures against MSV must be found to secure high yields.

1.2.5 Control of maize streak virus disease

1.2.5.1 Cultural practices

Agronomic practices such as crop rotation, incorporating barriers of bare ground between early and late planted maize fields, timely planting to avoid infestation, avoiding the planting of maize downwind from older cereal crops and the removal of infected plants, are used as cultural control measures to reduce leafhopper movement and the spread of MSVD between farms (Mawere *et al*., 2006; Martin and Shepherd, 2009).

Control of MSVD using cultural strategies is cheap and accessible to most African farmers. However, it is impossible to achieve complete MSVD control with these strategies given the inherent unpredictability of MSVD epidemiology (Adejumo, 2005; Martin and Shepherd, 2009). An example of the complexity involved is given by Martin and Shepherd (2009) of a farmer, who in the hope of avoiding the attack of MSVD-susceptible plants by migrating leaf-hoppers, changed planting dates. Without accurate long-term weather forecasts and a reasonably predictive MSVD epidemiological model, this proved a futile exercise.

1.2.5.2 Chemical practices

Carbamate insecticides such as carbofuran are used to control MSVD and have been shown to reduce its incidence in the field by killing leafhoppers. However, absolute protection against MSVD is also not achievable with insecticides, as they only provide partial control of MSVD (Bosque-Perez *et al*., 1998). Insecticidal control of leafhoppers generally requires repeated insecticide applications to control migrant leafhoppers (Magenya *et al*., 2008). Under severe leafhopper pressures, insecticides can be almost completely ineffective at controlling MSVD (Martin and Shepherd, 2009). Prices of imported chemicals and spraying equipment often limit access to this approach of controlling MSVD for small-scale farmers (Mawere *et al*., 2006). While insecticides have been used to control leafhopper vectors, Bosque-Perez (2000) stated that "resistance breeding is perceived as the most practical solution for the control of MSVD".

1.2.5.3 Use of maize streak virus disease-resistant maize genotypes

The use of resistant cultivars is probably the most economically viable approach to reducing losses that result from MSVD (Bosque-Perez, 2000; Mawere *et al*., 2006). Despite significant progress having been made in the breeding of MSVD resistant maize (Efron *et al*., 1989; Barrow, 1993; Welz *et al*., 1998; Barrow, 2003; Asea *et al*., 2008), in practice commercially available maize hybrids are at best moderately tolerant to MSV disease (Martin and Shepherd, 2009).

Collaborative efforts by several international and regional maize breeding programmes like the International Maize and Wheat Improvement Center (CIMMYT) and the International Institute of Tropical Agriculture (IITA) have produced a large collection of germplasm with improved MSV disease resistance (Welz *et al*., 1998; Bosque-Perez, 2000). Resistance to infection by MSVD has been identified in the CIMMYT inbred line CML202, which is adapted to the mid-altitude tropics (Welz *et al*., 1998; Asea *et al*., 2008). Other sources of resistance include C390 from the Agricultural Research Centre for International Development (CIRAD), IITA's Tzi3 and Tzi4, CIMMYT's

OSU231, PANNAR's A076 and Embu11 from the Kenya Agricultural Research Institute (KARI) (ISAAA, 1999). In the current study for the development of Mozambican MSV resistant hybrids, lines better adapted to the tropical lowland conditions other than the above mentioned lines were selected. The parental inbred lines for each study were chosen based on their diversity in disease resistance and adaptation to the environment.

1.3. Downy mildew disease

The production of maize in Sub-Saharan Africa is also threatened by DM disease, one of the most destructive maize diseases in the world (Frederiksen and Renfro, 1977; George *et al*., 2003). Downy mildew of maize is caused by at least ten different species of oomycete fungi that belong to the genera *Peronosclerospora*, *Sclerophthora* and *Sclerospora* (William and Grunwald, 2010). In sorghum and maize, the disease is caused by *Perenosclerospora sorghi* (Cardwell *et al*., 1997; Bock *et al*., 2000), a fungus that belongs to the order *Peronosporales* and the family *Peronosporaceae* (Jeger *et al*., 1998). Other downy mildew diseases that also affect maize are brown stripe downy mildew (*Sclerophthora rayssiae* R.G. Kenneth, Koltin and I. Wahl), crazy top downy mildew *(Sclerophthora macrospora* (Sacc.) Thirum., C.G. Shaw and Naras) and sugarcane downy mildew (*Peronosclerospora sacchari* (T.Miyake) Shirai and Hara) (George *et al*., 2003).

1.3.1 Downy mildew biology

The fungus *Perenosclerospora sorghi* reproduces asexually by means of conidia, and sexually via oospores (Jeger *et al*., 1998). Downy mildew disease is spread by soil infestations of sexually produced, thick-walled, long-lived oospores that enable the pathogen to survive dry periods and have been known to survive in the soil for up to ten years (Jeger *et al*., 1998; Adenle and Cardwell, 2000). Once infection is initiated in susceptible cultivars, the spread to neighbouring plants occurs via asexual conidia that are released from lower leaf surfaces following periods of high relative humidity (Bock

et al., 1998; Jeger *et al*., 1998). The conidia produced by *P. sorghi* on erect conidiophores which grow out through leaf stomata, are copiously produced in thin-walled structures that allow for the rapid polycyclic increase and spread of an epidemic within a season (Jeger *et al*., 1998; Adenle and Cardwell, 2000).

1.3.2 Geographic distribution of downy mildew disease

Downy mildew disease was first reported in India in 1907 and is now widely distributed in Asia and Africa (Frederiksen and Renfro, 1977; Bigirwa *et al*., 2000). In Asia, it is considered as one of the most destructive diseases of maize causing yield losses of up to 50%, thus making it a top priority biotic stress factor of maize (Pingali and Pandey, 2001; George *et al*., 2003). In Africa, DM disease outbreaks have been reported from Uganda, Mozambique and the Democratic Republic of the Congo (Ajala *et al*., 2003). In Nigeria the disease has been widespread because of continuous cultivation of maize throughout the year and is a serious constraint to maize production, especially in the forest zone (Kim *et al.*, 1994). In Mozambique, *P. sorghi* has been identified only on maize and not in sorghum (Plumb-Dhindsa and Mondjane, 1984).

1.3.3 Symptoms of downy mildew

Maize plants infected by *P. sorghi* at the seedling stage have a characteristic stunted appearance and may die prematurely approximately four weeks after infection (Jeger *et al*., 1998; Ajala *et al*., 2003). Leaves of older plants display the characteristic chlorotic streaking that begins at the base of the leaf with a clearly defined margin between diseased and healthy tissue (Fig. 1.2A). Infected plants have leaves that are narrower and more erect than healthy leaves (Craig and Frederiksen, 1983; Jeger *et al*., 1998; Ajala *et al*., 2003). Further, these plants may not seed, thus cobs fail to form and the tassel is replaced by a mass of twisted leaves exhibiting 'bushy' growth, hence the top is usually referred to as "crazy top" (Ajala *et al*., 2003).

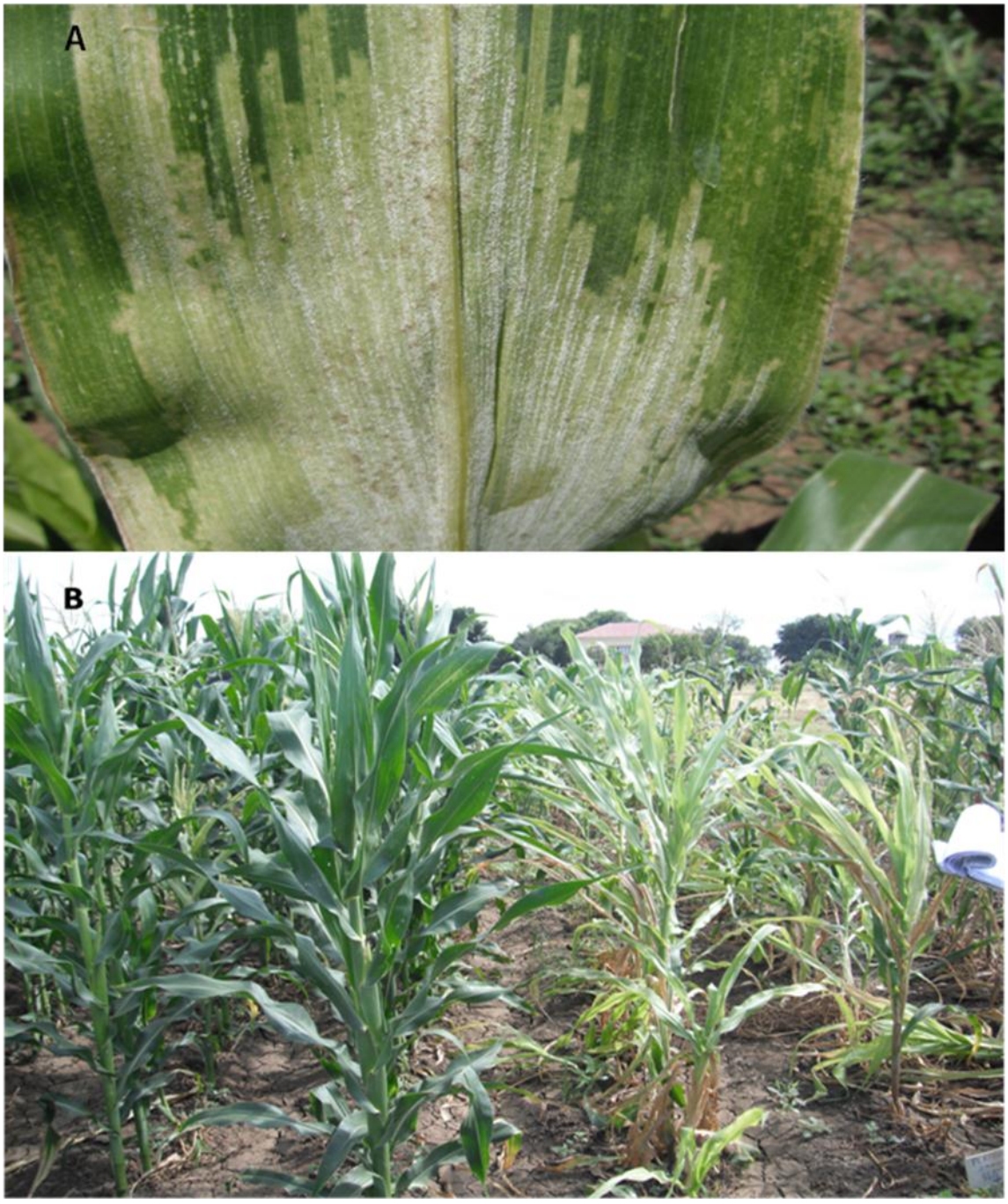

Figure 1.2: Downy mildew disease. **A**: Close-up of downy mildew disease symptoms of maize infected with *P. sorghi*, showing the characteristic white striping of leaves, which always includes the base; **B**: Comparison of downy mildew (DM) resistant and DM susceptible maize plants; on the right is maize with DM incidence approaching 100%, there is also stunting induced. DM resistant variety on the left is thriving. Photo by: Nothando F. Mafu, PMB, UKZN. Taken at IIAM Chokwe, Mozambique. Date: 13/03/11.

1.3.4 Damage caused by downy mildew disease

Downy mildew is one of the most destructive maize diseases in Nigeria and Mozambique (Pingali and Pandey, 2001; Ajala *et al*., 2003). In Mozambique, it causes crop losses in the lowland areas in the central and southern regions. The losses have resulted in some extensively grown varieties being withdrawn from the market during 2003-2006 due to their susceptibility to DM disease (Mariote, 2007). High relative humidity of about 90%, temperature ranges from 20-25°C and rainfall mostly favours disease development (Amusa and Iken, 2004). Yield losses as a result of DM infection have been reported to range from 10-100% (Fig.1.2B) (Gowda *et al*., 1987). Therefore, the disease requires effective control strategies to minimise economic damage.

1.3.5 Control methods of downy mildew disease

1.3.5.1 Cultural control

Cultural control methods can be used to manage DM disease in maize. These involve planting in well drained soils to reduce oospore growth, burying any infected crop debris to reduce inoculum sources and simultaneously cultivating maize with the alternative hosts (e.g. sorghum) of downy mildew disease (Bigirwa *et al*., 2000). The impact of diseases depends on the age of the plant; therefore adjusting planting times such that crops can escape high disease pressures can also be used as a means to control DM disease (Gilbert, 2002). For example, Frederiksen and Renfro (1977) reported that young plants outgrew DM disease when they were not infected at the seedling stage.

1.3.5.2 Chemical control

Chemical control involves the use of fungicides (George *et al*., 2003). While it has been effective, fungicides are not readily available in remote areas of Mozambique (Mariote, 2007). In addition, an emerging problem as a result of intensive use of fungicides is the build-up of chemical resistance in pathogens making some fungicides inefficient

(Perchepied *et al*., 2005). As with MSVD control, resistance breeding is perceived as the best solution for the control of DM disease (George *et al*., 2003; Perchepied *et al*., 2005). This strategy is advocated in the current study.

1.3.5.3 Use of downy mildew-resistant maize genotypes

In West Africa, breeding for downy mildew resistance (DMR) in maize started in the mid-1970s as a collaborative effort between the Nigerian National Maize Programme and the IITA. The effort resulted in the development of early and late maturing open-pollinated resistant varieties released for cultivation in the mid-1980s (Ajala *et al*., 2003). The use of resistant varieties is proving to be more cost-effective and an environmentally safe alternative for controlling DM diseases (George *et al*., 2003; Perchepied *et al*., 2005).

1.4. Breeding of disease resistance maize varieties

Breeding for disease resistance to MSV and DM in maize is an efficient control measure that is reliable and cost-effective. It is based on the identification and incorporation of major resistance genes into economically important varieties (Saxena and Hooker, 1968; Wisser *et al*., 2006). The two types of resistance that are recognized are qualitative and quantitative (Wisser *et al*., 2006). Their mechanisms are diverse in their specificity and durability and interactions with the virus in the host plant (Lecoq *et al*., 2004).

1.4.1 Qualitative resistance

Qualitative resistance is also referred to as vertical resistance in that it is either present or absent, and there are no intermediates (Robinson, 2006). This type of resistance is usually race-specific and because it is usually based on a single dominant gene, it confers a high level of resistance (Robinson, 2006; Wisser *et al*., 2006). It is easier to work with qualitative resistance in crop genetic studies and in plant breeding as the

genes follow a Mendelian pattern. However, this resistance is not stable because it is usually matched by virulent races of polygenic pathogens within 3-5 years. In contrast, quantitative resistance, also known as horizontal resistance, is often more useful in an agronomic context hence it is generally recommended for the small-scale and subsistence farmers (Wisser *et al*., 2006**)**.

1.4.2 Quantitative resistance

Quantitative resistance is usually assessed in the field and is considered to have a generally higher durability and broader specificity since it is controlled by multiple genes with small continuous phenotypic effects (Parlevliet, 1995). Quantitative resistance can occur at every level between a minimum and a maximum level (Parlevliet, 1995; Robinson, 2006; Wisser *et al*., 2006). Environmental and gene-for-gene interactions play important roles in the phenotypic expression of quantitative resistance (Young, 1996). Extensive field-testing is, therefore, required for assessment of quantitative resistance under multiple environments and also at different growth stages (Parlevliet, 1995; Robinson, 2006). Wisser *et al*. (2006) stated that "the majority of disease resistance deployed in elite maize varieties in the field is quantitative in nature" and because breeding for resistance can be a long and tedious method, it is most cost effective to breed for quantitative resistance, which is likely to provide long-term, durable protection (Lecoq *et al*., 2004).

1.4.3 Nature and mechanism of MSV disease resistance

Kyetere *et al*. (1999) and Mawere *et al*. (2006) reported a major quantitative trait locus (QTL) on the short arm of chromosome 1 (1S - bin1.04) and designated it *msv1*. Mawere *et al*. (2006) stated that "resistance in maize to MSV is controlled by a major gene, with two, three or 'few' modifying genes". The same locus was identified by Welz *et al.* (1998) in a population derived by crossing CML202, an MSVD resistant inbred, and Lo951, a susceptible inbred. Although most of the resistance was explained by the locus on chromosome 1, with the major MSV resistance gene being identified as

msv1 (Welz *et al*., 1998; Kyetere *et al*., 1999; Pernet *et al*., 1999; Asea *et al*., 2008), minor QTL effects have been detected at bins 3.06, 5.03 and 8.07 (Asea *et al*., 2008). Pernet *et al.* (1999) investigated QTL responsible for resistance to MSVD and showed that the resistance was quantitatively inherited. They detected at least five significant QTL on chromosomes 1, 2, 3, and 10 in resistant cultivar D211. MSV resistance is thus under the control of two genetic systems, one arising from a major gene on the short arm of chromosome 1 and the other conditioned by minor genes on chromosomes 2, 3 and 10, that confer quantitative resistance. Virus resistance is associated with one or two major resistance loci in most cases, which facilitates MAS, but resistance genes have been found to cluster in the maize genome (Redinbaugh *et al*., 2004). Stability of QTL across populations has been shown to be variable; however, this is not the case for the MSVD QTL (Pernet *et al*., 1999).

1.4.4 Nature and mechanism of downy mildew disease resistance

Agrama *et al.* (1999) identified two QTL on chromosome 1 and a third QTL on chromosome 9 that control the inheritance of resistance to DM disease. The genetics of DM resistance has been shown to be complex and polygenic in nature with additive effects predominantly contributing to the resistance (George *et al*., 2003).

1.5 Classical breeding

The fundamental basis of plant breeding is the selection of specific plant traits considered important by plant breeders. Classical breeding, also known as conventional breeding, involves crosses between selected parent plants that have desirable characteristics such as high yield or disease resistance (Visarada *et al*., 2009). Selection of superior plant traits involves visual assessment, thus the breeder's skills lie in selecting the best plants with desirable recombinant characteristics from the large segregating offspring populations (Visarada *et al*., 2009; Ulukan, 2011). Selection is postponed until later generations (F_5 or F_6) to enable alleles for traits of low heritability to be fixed, thus improving homozygosity of the progeny. The progeny are then

harvested in bulk and evaluated in replicated field trials. This process is expensive and laborious and takes about 5-10 years for elite lines to be developed (Collard and Mackill, 2008). A typical breeding programme can grow up to millions of individual plants, especially in the case of a large number of genes segregating, in order to identify specific gene combinations (Collard and Mackill, 2008).

The complexity of selection required in breeding programmes and the large size of the populations often required, point toward the need for new tools to assist breeders in plant selection (Xu and Crouch, 2008). It has been suggested that genetic engineering is best utilised when manipulating traits that depend on one or a few genes (Boopathi *et al*., 2011). Genetic engineering creates 'recombinant DNA' which is the result of direct human manipulation of an organism's genome involving the insertion of foreign DNA into that of another organism, which does not require the use of classical genetic methods (Gupta, 2008). However, Garcia-Olmedo (2002) stated that classical breeding is still the most effective approach when dealing with traits controlled by multiple genes distributed over the genome, especially with the aid of molecular markers. Using DNA markers via MAS has enormous potential to improve the efficiency and precision of conventional plant breeding (Collard *et al*., 2005).

1.6 Marker-assisted selection in plant breeding

Marker-assisted selection (MAS) in plant breeding refers to the use of molecular markers, usually DNA-based for the selection of plants with a region of DNA involved in the expression of a trait of interest (Collard *et al*., 2005; Stevens, 2008). Markers are tightly linked to agronomically important genes to assist in the selection of elite lines for the next generation crosses in crop improvement programmes, thus the marker is used to identify the gene (Semagn *et al*., 2006; Ibitoye and Akin-Idowu, 2010). Marker-assisted selection involves exploiting the presence or absence of a marker to facilitate phenotypic selection (Collard *et al*., 2005; Semagn *et al*., 2006). It is a more efficient and reliable approach than conventional plant breeding methodology as it is unaffected by environmental factors (Collard *et al*., 2005). This development has opened up a new

realm of possibilities in agriculture towards improvement of economically important crop varieties.

1.6.1 Marker-assisted breeding vs. conventional breeding

The advantage of MAS is that genotypes can be identified at the seedling stage, eliminating the time needed for plant maturation and reducing population sizes (Stevens, 2008). Conventional breeding methodology on the other hand, relies on phenotypic evaluation, which does not always accurately reveal the basic genomic information of the plant (Dreher *et al*., 2000; Collard *et al*., 2005). Environmental effects and genotype x environment interactions can significantly conceal the presence or absence of specific alleles, making it difficult for breeders to identify plants with the desired traits (Dreher *et al*., 2000). A solution is to use molecular markers that correspond to particular sequences of DNA in the plant genome. By informing the breeder of the plant's true identity and confirming the presence or absence of the desirable alleles through the use of markers, breeders cease to use time-consuming phenotypic evaluation methodology to determine whether or not alleles are present (Dreher *et al*., 2000). Therefore, MAS allows for a greater degree of selection precision whilst still greatly reducing the time required to achieve a particular breeding objective (Dreher *et al*., 2000; Collard *et al*., 2005; Lagat *et al*., 2008).

1.6.2 Application of markers for screening for disease resistance

In the case of disease resistance, marker-based selection is valuable for simplifying the pyramiding of several major resistance genes into one genetic background (Young, 1996). It is particularly useful in the screening for one resistance gene that interferes with the ability to screen for another, a common problem in disease resistance breeding (Young, 1996; Masojc, 2002). Efficient gene development and deployment can thus be accelerated through the use of marker-assisted breeding. Consequently, QTL from diverse donors can be rapidly introgressed into a desirable genetic background of commercial cultivars (Young, 1996; Collins *et al*., 2008). In most cases, virus resistance

is associated with one or two major resistance loci, which facilitates MAS (Redinbaugh *et al*., 2004). To make use of MAS, virus resistance must first be identified in maize germplasm and then mapped to specific regions of the maize genome. To aid in the identification of MSV or other virus resistance sources, identification and mapping of genes or QTL for virus resistance using markers must be available. This provides information on the number of genes or regions that must be transferred by breeding programmes (Redinbaugh *et al*., 2004).

A study was conducted by Asea *et al*. (2008) to determine the usefulness of molecular markers linked to consensus QTL controlling partial-resistance systems for NCLB, GLS and MSV in maize. The NCLB disease resistance QTL in chromosomal bins 3.06, 5.04 and 8.06; GLS QTL in bins 2.09 and 4.08; and a consensus MSV QTL in bin 1.04 were examined for selection in improving host resistance levels and pyramiding resistance loci of these diseases. Clustering of resistance genes is common in maize (McMullen and Simcox, 1995). Evaluations for each disease were done in a population of 410 $F_{2:3}$ lines derived from hybridisation between inbred line CML202 with known resistance to NCLB and MSV, and VP31, a breeding line with known resistance to GLS. The study concluded that markers linked to major resistance loci can facilitate pyramiding of resistance against multiple diseases during early generation selections. The major locus conferring resistance to MSV on chromosome 1 was significant ($P<0.05$) for resistance across seasons and phenotypic values indicated that QTL in bin 4.08 for GLS, bin 1.04 for MSV and bins 3.06 and 5.04 for NCLB significantly reduced disease severity.

1.6.3 Molecular genetic screening techniques

There are different marker systems used in the analysis of genetic diversity in plants in marker-assisted plant breeding programmes (Akkaya *et al*., 1992; Bolibok *et al*., 2005). These markers include RFLPs and PCR (polymerase chain reaction)-based molecular markers, such as RAPDs, SSRs, AFLPs and SNPs. They have an advantage as cultivar descriptors over morphological markers in that they are unaffected by environmental or physiological factors (Akter *et al*., 2008).

1.6.3.1 Restriction fragment length polymorphism (RFLPs)

Restriction fragment length polymorphism (RFLPs) was the first DNA marker system developed in the early 1980s (Clegg *et al*., 1999; Hoisington, 2001). Botstein *et al*. (1980) argued the possibility of constructing a complete human genome linkage map, where DNA polymorphisms were detected through restriction enzyme digestion. In maize RFLPs have been used extensively and successfully for polymorphism validation at DNA level among populations (Hai *et al*., 2000; Lubberstedt *et al*., 2000; Ignjatovic-Micic *et al*., 2003). The technique has been used for identification purposes in many crops like cowpea (Mignouna *et al*., 1998) and mungbean (Lakhanpaul *et al*., 2000). This technique however, proved time consuming and labour-intensive, thus with the development of PCR-based techniques, several markers emerged, namely RAPDs, AFLPs and SSRs (Welsch and McClelland, 1990; Williams *et al*., 1990; Lubberstedt *et al*., 2000; Hoisington, 2001; Ignjatovic-Micic *et al*., 2003).

1.6.3.2 Random amplified polymorphic DNA (RAPDs)

Random amplified polymorphic DNA analysis was developed independently by two different laboratories (Welsch and McClelland, 1990; Williams *et al*., 1990). A study by Ignjatovic-Micic *et al*. (2003) compared the effectiveness of the RFLP and RAPD techniques in the characterisation, identification and classification of 13 local maize populations from the Maize Research Institute 'Zemun Polje'. Characterisation using 20 and 30 RAPD and RFLP markers, respectively, revealed a high level of polymorphism among populations and the genetic distances calculated were highly similar leading to the conclusion that both methods can be successfully used for polymorphism validation.

The RAPD procedure has overcome the technical limitations of RFLPs, it is a simpler and faster method that has gained popularity due to the simplicity and decreased costs of the assay (Saliba-Colombani *et al*., 2000; Ignjatovic-Micic *et al*., 2003). Its assays

utilise short oligonucleotides of 4-10 bp of random sequences as primers to amplify relatively small DNA amplicons using low annealing temperatures by PCR amplification (Williams *et al*., 1990; Gonzalez-Chavire *et al.,* 2006).

Amplification products are separated by gel-electrophoresis and the exhibited polymorphisms are detected as the presence or absence of bands of particular size and are thus used as genetic markers (Bardakci, 2001; Gonzalez-Chavire *et al.,* 2006*)*. The presence of a RAPD band does not allow discrimination between heterozygous and homozygous states therefore; these are dominant markers (Williams *et al*., 1990; Collard *et al*., 2005). The main disadvantage with the method is that it is highly sensitive to small changes in laboratory conditions, and minor modifications of protocols, and therefore, there is a low reproducibility within and between laboratories (Ignjatovic-Micic *et al*., 2003; Agarwal *et al*., 2008; Ibitoye and Akin-Idowu, 2010).

1.6.3.3 Amplified fragment length polymorphisms (AFLPs)

Amplified fragment length polymorphisms technology developed by Vos *et al* (1995) overcame the limitation of reproducibility of RAPDs (Agarwal *et al*., 2008). The AFLP method is a DNA fingerprinting technique based on selective PCR amplification of restriction fragments from the genomic DNA of any origin or source without prior sequence knowledge and hence, has a relatively low start-up cost (Farooq and Azam, 2002; Gonzalez-Chavire *et al*., 2006; Vuylsteke *et al*., 2007; Agarwal *et al*., 2008). The technique can be used to distinctly differentiate closely related individuals at the sub-species level (Agarwal *et al*., 2008). Saliba-Colombani *et al*. (2000) stated that AFLP assays are more reliable than the RAPD technique. The AFLP markers have been employed to investigate (i) the genetic similarity among different accessions within maize inbreds (Lubberstedt *et al*., 2000; Heckenberger *et al*., 2003) and (ii) the relationship among four backcross generations and parents in cotton (Zhong *et al*., 2002).

The AFLP technique can be automated (Vuylsteke *et al*., 2007; Ibitoye and Akin-Idowu, 2010) and is highly multiplexed, which as stated by Vuylsteke *et al*. (2007),

"offers the potential to improve the efficiency and to increase the throughput of marker data production in organisms that lack the genomics platform necessary to allow the development of genotyping microarrays". In addition, AFLP markers have the advantage of being locus-specific but only at species level (Saliba-Colombani *et al*., 2000).

The disadvantages of AFLP markers include their dominant nature (Yuste-Lisbona *et al*., 2008; Ibitoye and Akin-Idowu, 2010), the fact that they are relatively costly, a high level of technical expertise is required to work with them, they are laborious to work with, and are not amenable for routine and quick screening (Yuste-Lisbona *et al*., 2008).

1.6.3.4 Simple sequence repeats (SSRs)

Simple sequence repeat (SSR) markers, also known as microsatellites (Molnar *et al*., 2003), are one of the most extensively used DNA marker types in the characterisation of germplasm collections in major cereal crops (He *et al*., 2003), e.g. maize (Warburton *et al*., 2001; Vigouroux *et al*., 2005; Kostova *et al*., 2006). Microsatellites consist of tandemly arranged bases that are spread throughout the genomes. The repeated sequences are often simple, consisting of two, three or four nucleotides (di, tri and tetra) that can be repeated 10-100 times (He *et al*., 2003; Santana *et al*., 2009). Different repeat numbers in SSRs can be treated as separate "alleles" and the site can be treated as a highly polymorphic site with multiple alleles for the detection of variations in populations (Akkaya *et al*., 1992).

Compared with other marker types, microsatellites are more advantageous because they are highly polymorphic, even among closely related cultivars, due to natural occurring mutations that can distinguish between closely related species (Brown *et al*., 1996). Microsatellites are highly abundant, simple to analyse, co-dominant, and are easily and economically assayed by PCR using primers specific to conserved regions flanking the repeat array (Yu *et al*., 2000). Another advantage is that they are accessible to other research laboratories via published primer sequences (Yu *et al*., 2000; He *et al*., 2003). Information gathered by these markers allows comparisons and information exchange

between different studies, especially for comparative genetic mapping (Grisi *et al*., 2007).

Danson *et al*. (2006) screened a total of 115 recombinant inbred lines for resistance to MSV disease using 52 SSR markers contained between bin 1.04 and 1.05 of maize chromosome 1. Of these, three microsatellite primers; bnlg1811, umc1917 and umc1144 targeting three loci of chromosome 1 were chosen on the basis of their polymorphism content. These markers were able to differentiate resistant from susceptible lines Markers for this current study were from Danson *et al*. (2006); Asea *et al*. (2008), the maize database (http://www.maizegdb.org) and from a study by Lagat *et al*., (2008) in which the QTL for resistance to MSVD in one population was mapped using SSRs. The study by Asea *et al*. (2008) determined the usefulness of the SSR molecular markers linked to consensus QTL that control partial-resistance systems. NCLB disease resistance QTL are located in chromosomal bins 3.06, 5.04 and 8.06; GLS QTL in bins 2.09 and 4.08; and a consensus MSV QTL in bin 1.04 for selection in improving host resistance. Multiple resistance was combined into the hybrid line derived from inbred line CML202 and VP31. It was concluded that a conventional maize breeder requires the use of molecular markers in order to improve selection intensity and maximise genetic gain (Collard *et al*., 2005).

1.6.3.5 Single nucleotide polymorphisms (SNPs)

Agarwal *et al*. (2008) stated that single nucleotide variations constitute the most abundant molecular markers and are widely distributed throughout genomes although their occurrence and distribution varies among species (Ryynanen *et al*., 2007; Yan *et al*., 2009). A considerable amount of screening effort is required for the development of SNPs and a large number of loci need to be assessed due to the low amount of information per marker, thus very high developmental costs can be incurred (Ryynanen *et al*., 2007; Ibitoye and Akin-Idowu, 2010). For instance, it has been estimated that for accurate parentage determination in natural populations, a considerable higher number of SNP markers would be required in order to have an equivalent discriminating power as with SSR marker loci that are multi-allelic (Ryynanen *et al*., 2007; Van Inghelandt *et al*., 2010). This characteristic of SSR markers thus makes them the markers of choice in

most plant genetics and breeding programmes (Varshney *et al*., 2007). However, with the recent development of several high throughput genotyping technologies that take advantage of the wealth of SNPs in eukaryotic genomes and the studies of nucleotide diversities using SNPs in wheat (Trick *et al*., 2012) and maize (Hamblin *et al*., 2007; Yan *et al*., 2009; Van Inghelandt *et al*., 2010), SNPs are fast becoming popular (Yan *et al*., 2009; Deulvot *et al*., 2010). The molecular marker advantages and disadvantages are further summarised below in Table 1.1.

Table 1.1: Advantages and disadvantages of different molecular marker types.

Marker type	Advantages	Disadvantages	References
RFLPs	- no sequence information required - detects in related genomes - robust in usage - transferable across populations	- labour intensive and time consuming - large quantity of DNA needed - often very low level of polymorphism - fairly expensive	Tanksley *et al*., 1989; Collard *et al*., 2005; Ibitoye and Akin-Idowu, 2010.
RAPDs	- no sequence information required - can be automated - relatively low quality of DNA required - high polymorphism - inexpensive	- cannot be used across populations nor across species - highly sensitive to laboratory changes - often see multiple loci - low reproducibility	Williams *et al*., 1990; Collard *et al*., 2005; Ibitoye and Akin-Idowu, 2010.
AFLPs	- can be automated - no sequence information required - high levels of polymorphism generated	- marker clustering - technique is patented	Vos *et al*., 1995; Collard *et al*., 2005; Ibitoye and Akin-Idowu, 2010.
SSRs	- highly polymorphic - can be automated - small quantity of DNA required - robust and reliable - transferable between populations	- difficult interpretation because of stuttering - large amounts of time and labour required for development of primers	Powell *et al*., 1996; McCouch *et al*., 1997; Collard *et al*., 2005; Ibitoye and Akin-Idowu, 2010.
SNPs	- can be automated	-high development cost	Collard *et al*., 2005;

	- robust - suitable for high throughput	- require sequence information - technically challenging	Ibitoye and Akin-Idowu, 2010.

1.6.3.6 Analysis techniques

Simple sequence repeat polymorphisms and SNPs can be visualised by electrophoresis on polyacrylamide gels, which is sensitive for resolving differences of a single repeat length (Karp *et al*., 1996; Mader *et al*., 2008; Kinoshita *et al*., 2009). Agarose gels distinguish alleles differing by several repeats, thus visualisation of length polymorphisms in SSRs or SNPs by conventional gel electrophoresis is not only time consuming and labour intensive, but it is also insufficiently sensitive to detect single base pair differences and normally leads to problems in interpretation due to stutter bands (Karp *et al*., 1996; Taranenko *et al*., 1999; Mader *et al*., 2008; Studer *et al*., 2009). Sequencing the fragment, either manually or using an automated DNA sequencer will however resolve all the possible single base differences between samples (Karp *et al*., 1996; Meldrum, 2000).

The common methods used today for the assay of SSR and SNP analysis using multiplex PCR and high speed DNA sequencing by capillary electrophoresis, allow for the use of fluorescently labelled primers, which have to be optimised for multiplex PCR (Hayden *et al*., 2008; Mader *et al*., 2008; Yan *et al*., 2009). Analysis using multiplex PCR and capillary electrophoresis among others are post-PCR technologies that have proved to be expensive, time consuming to implement and require detailed knowledge on allelic sequence information (Mader *et al*., 2008; Studer *et al*., 2009). SNP genotyping by Sanger sequencing is limited by its low throughput and high cost per sample (Jenkins and Gibson, 2002). Application of the Illumina Golden Gate platform to SNP genotyping requires multiple preparation steps and is relatively expensive due to the initial high cost of probe production (Kinoshita *et al*., 2009).

An alternative simple, fast and cheaper method than electrophoresis analysis for genotyping both SSR and SNP markers is HRM analysis (Studer *et al*., 2009; Yu *et al*., 2011). Compared to other post-PCR technologies, HRM analysis has the advantages of

visualisation immediately following PCR amplification, lower costs and higher sensitivity than electrophoretic detection systems (Mader *et al*., 2008).

1.7 High resolution melt (HRM) analysis technology

Introduced in 2002 (Reed *et al*., 2007), HRM analysis has been widely adopted in clinical chemistry, human and plant sciences (White *et al*., 2007; Croxford *et al*., 2008; Steer *et al*., 2008). The technique was used in high throughput SNP discovery in tetraploid alfalfa (Han *et al*., 2012); sweet cherry (Marti *et al*., 2012) and high throughput SSR discovery in maize (Yu *et al*., 2011). High resolution melt analysis is used to characterise DNA samples by monitoring the melting behaviour of PCR amplicons as they transition from double stranded DNA (dsDNA) to single stranded DNA (ssDNA) with increasing temperature (Do and Dobrovic, 2009; Studer *et al*., 2009). HRM analysis is the only scanning technique that uses the closed-tube method where HRM analysis and PCR amplification are performed in the same tube (Gudrun and Wittwer, 2004). The technique has the advantages of reduced contamination and reduced processing time over other pre-sequencing scanning methods like sequencing based methods (Gudrun and Wittwer, 2004; Studer *et al*., 2009; Hondow *et al*., 2011). A closed-tube method means the subsequent data for analysis of genotypes and sequencing is immediately available with no further processing (Gudrun and Wittwer, 2004). The post-PCR analysis method scans entire amplicons and detects sequence variations using a saturating dsDNA binding dye. Saturating dyes commonly used include: LCGreen Plus, SYBR Green, ResoLight and SYTO9 (Stoep *et al*., 2009; Li *et al*., 2010). The amplicon is analysed by melting curves produced as temperatures increase and fluorescence decreases (Montgomery *et al*., 2007). With increasing temperatures, the amount of dsDNA decreases as shown in Fig. 1.3 and the intercalating dyes that bind specifically to dsDNA are released and thus have increased fluorescence (Ganopoulos *et al*., 2011). The change in fluorescence is caused by the release of the intercalating dye from a DNA duplex as it is denatured. The process is precisely monitored. Amplicons containing different sequences can be discriminated based on the melting transition of the PCR product and the resulting melt curve shape (Chateigner-Boutin and Small, 2007; Steer *et al*., 2008; Stoep *et al*., 2009).

High resolution melt analysis can be performed in approximately two minutes, thus analysis of PCR amplicons using this technique allows for rapid cost-effective genotyping, especially with large sample numbers, which makes it effective in clinical and epidemiological investigations (Gudrun and Wittwer, 2004; Steer *et al*., 2008).

1.8 FTA^{TM} technology

To sustain the large ever growing population of the world there is a need for breeders to better combat plant diseases and increase crop yields; thus cost effective tools for the molecular study of plant pathogens are of great importance (Ndunguru *et al*., 2005). Marker-assisted breeding involves the screening of large numbers of samples by PCR-based techniques, thus making the whole process of plant DNA sample collection, isolation and processing of the DNA for PCR-analysis labour intensive. This creates a

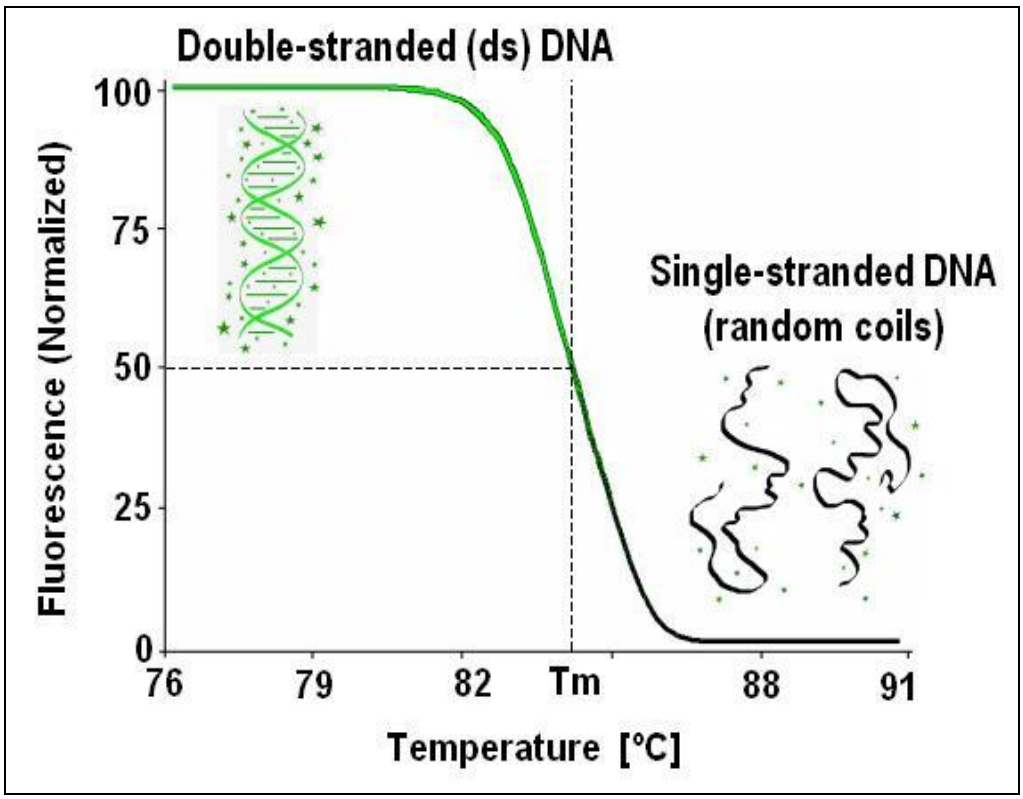

Figure 1.3: Fluorescent dye is intercalated into the strands of double-stranded DNA (dsDNA) and is released as temperatures are increased and the dsDNA structure dissociates into single-stranded DNA. (http://hrm.gene-quantification.info/).

bottleneck for analysing numerous samples (Drescher and Graner, 2002). Flinders Technology Associates (FTA™) card is a simple paper-based technology designed to reduce the steps of DNA collection, transportation, purification and storage (Mbogori *et al*., 2006). The collection of samples on filter papers started with the collection of blood samples in neonatal diagnostics (Guthrie and Susi, 1963), and later for PCR in medical and forensic applications (Carducci *et al*., 1992). This simple and feasible collection and storage method has been adopted for broad use in diagnostic screening, drug monitoring and genetic analysis, being particularly suitable for molecular epidemiologic studies in remote areas with tropical climates, where transport and storage conditions are difficult (Sultan *et al*., 2009). Conventional methods of plant DNA extraction are laborious and can be avoided with the use of FTA™ elute cards, thus making the process of purifying DNA ready for downstream applications more cost effective and less time consuming (Ndunguru *et al*., 2005; Mbogori *et al*., 2006). The FTA™ cards contain chemicals that lyse cellular material and bind and preserve DNA and RNA from degradation within the matrix of the paper. The technology prevents the degradation of genomic DNA at room temperature, making PCR amplification of viable DNA possible after long-term archiving (Ndunguru *et al*., 2005; Mbogori *et al*., 2006; Johanson *et al*., 2009).

1.8.1 Features and benefits of using FTA™ cards

An investigation was made between genomic DNA stored on FTA™ cards for over 14 years and that stored on non-FTA cards for over six months both at room temperature. The DNA stored on the FTA™ cards was successfully amplified by PCR whilst DNA stored on non-FTA did not amplify (Mbogori *et al*., 2006; McClure *et al*., 2009). This proved that FTA™ cards simplify the harvesting and storing of samples which can then be stored for long periods without the deterioration of the DNA. The FTA™ cards thus offer a compact room-temperature storage system that alleviates the need for valuable freezer space. Furthermore, transport of samples from long distant greenhouses or

experimental fields to the laboratory can be done without the need for freezers. This greatly reduces the labour involved in sample collection and management. In a study employing FTA™ cards in maize, Ndunguru *et al*. (2005) concluded that the working time required for sampling is shorter than with conventional DNA extraction methods based on liquid or frozen sample handling. The number of samples that can be collected within a given time and location is also significantly increased compared to conventional methods (Ndunguru *et al*., 2005). The costs involved are also greatly lowered as there is no need for freezers or specialised couriers for transportation as the cards are lightweight and thus economical (Mbogori *et al*., 2006).

The FTA™ elute cards eliminate the long sample processing times, the high cost of using purification kits to isolate DNA, and also the use of nucleic acid purification devices. The non-microbial growth and long-term room temperature storage of nucleic acid add further advantages to the FTA™ card method (Sultan *et al*., 2009). Use of FTA™ cards thus has the potential to reduce cost, sampling time and sample storage space (Ndunguru *et al*., 2005).

1.9 Genetic diversity and its importance for breeding

Molecular markers cannot give all the information needed in a plant breeding programme, therefore conventional breeding can never be rendered obsolete in crop improvement. However, MAS can complement conventional breeding to increase efficiency and optimise genetic gain in selection (Thormann *et al*., 1994; Danson *et al*., 2006; Karanja *et al*., 2009). Karanja *et al.* (2009) stated that compared to most crops "maize exhibits a wider range of morphological and molecular dynamism", thus the need to use both molecular markers and morphological markers in maize breeding programmes. Both molecular and morphological markers allow for extensive data estimation of genetic diversity and the levels of genetic variation in maize to identify elite inbred lines that can be crossed to create superior hybrids (Smith and Smith, 1989; Karanja *et al*., 2009). Maize breeding relies on the available genetic diversity which can be manipulated for maximum heterosis estimation in hybrid breeding programmes (Karanja *et al*., 2009). Heterosis, a state when offspring show superiority over their

parents, is also termed hybrid vigour as there is either increase of vigour, size or other reproductive factors (Duvick, 1999; Virmani *et al*., 2003). It is important that there is a constant search for a diverse genetic base for the development of superior commercial hybrids in plant breeding programmes. Some studies have shown that the pedigrees of some hybrids can be from 6-8 inbred lines (Karanja *et al*., 2009).

Summary

This review of the literature established that:

- Maize is of great importance in Sub-Saharan Africa and is the mainstay of most of the continent's rural economies.
- Maize streak virus and downy mildew diseases are widespread in Africa with devastating results on maize yields thus greatly undermining the social development of the world's poorest people.
- Several methods exist that can be used to control MSV and DM diseases, but the use of resistant cultivars was identified as the most efficient and economic, particularly for subsistence farmers in Mozambique.
- This review aimed to show that conventional breeding alone is not sufficient for effective plant breeding programmes that aim for highly adapted elite lines in a shorter space of time. The use of molecular marker technology can greatly assist by reducing generation times nearly by half. Markers are able to detect diversity at DNA sequence level, thereby inform the breeder of any desirable variation or genes.
- Molecular markers can be employed in the development of the best hybrid combinations, a process known as MAS. Different types of markers were evaluated and their advantages and disadvantages discussed.
- Simple and cost effective tools for the molecular study of pathogens are of great importance if breeders are to better combat plant diseases and increase crop yields to sustain the ever growing world populations. The technique known as HRM analysis is a post-PCR method that can be used for the identification of genotypes. The technique is less time consuming as compared to electrophoresis

analysis and has the added advantage of data storage and analysis capabilities via computer assimilation.

- Marker-assisted breeding involves the screening of large numbers of samples by PCR-based techniques, FTA^{TM} is a simple paper-based technology designed to reduce the steps of DNA collection, transportation, purification and storage. The technology thus makes the process of purifying DNA ready for downstream applications more cost effective and less time consuming.
- Maize breeding relies on the available genetic diversity, thus better hybrids of maize need to be developed by making use of information on the genetic relationships and diversity among elite materials which is of fundamental importance in hybrid crop improvement.

References

Achieng, J., D. Friesen, O.M. Odongo, and M. Odendo. 2001. Sustainability of fertiliser use on maize production in western Kenya through provision of credit. Eastern and Southern Africa Regional Maize Conference, Nairobi, Kenya 7:428-431.

Adejumo, T.O. 2005. Crop protection strategies for major diseases of cocoa, coffee and cashew in Nigeria. African Journal of Biotechnology 4:143-150.

Adenle, V.O., and K.F. Cardwell. 2000. Seed transmission of maize downy mildew (*Peronosclerospora sorghi*) in Nigeria. Plant Pathology 49:628-634.

Agarwal, M., N. Shrivastava, and H. Padh. 2008. Advance in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27:617-631.

Agrama, H.A., M.E. Moussa, M.E. Naser, M.A. Tarek, and A.H. Ibrahim. 1999. Mapping of QTL for downy mildew resistance in maize. Theoretical and Applied Genetics 99:519-523.

Ajala, S.O., J.G. Kling, S.K. Kim, and A.O. Obajimi. 2003. Improvement of maize populations for resistance to downy mildew. Plant Breeding 122:328-333.

Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131-1139.

Akter, J., M.S. Islam, A.A. Sajib, N. Ashraf, S. Haque, and H. Khan. 2008. Microsatellite markers for determining genetic identities and genetic diversity among jute cultivars. Australian Journal of Crop Science 1:97-107.

Amusa, N.A., and J.E. Iken. 2004. Control of downy mildew disease of maize in forest agroecologies of Southern Nigeria. pp. 11-14. *In* N. A. Amusa, and J. E. Iken, (eds.) Slovak and Czech Plant Protection Conference. Slovak Agricultural University of Nitra, Slovakia.

Asea, G., B.S. Vivek, G. Bigirwa, P.E. Lipps, and R.C. Pratt. 2008. Validation of consensus quantitative trait loci associated with resistance to multiple foliar pathogens of maize. Phytopathology 99:540-547.

Bardakci, F. 2001. Random Amplified Polymorphic DNA (RAPD marker). Turkish Journal of Biology 25:185-196.

Barrow, M.R. 1993. Increasing maize yield in Africa through the use of maize streak virus resistant hybrids. African Crop Science Journal 1:139-144.

Barrow, M.R. 2003.Development of maize hybrids resistant to maize streak virus. Crop Protection 11:267-271.

Bigirwa, G., E. Adipala, J.P. Esele, and K.F. Cardwell. 2000. Reaction of maize, sorghum and Johnson grass to *Peronosclerospora sorghi*. International Journal of Pest Management 46:1-6.

Bock, C.H., L.K. Mughogho, M.J. Jeger, E. Mtisi, and K.F. Cardwell. 1998. Production of conidia by *Peronosclerospora sorghi* on sorghum crops in Zimbabwe. Plant Pathology 47:243-251.

Bock, C.H., M.J. Jeger, K.F. Cardwell, L.K. Mughogho, and J. Sherington. 2000. Control of sorghum downy mildew in maize and sorghum in Africa. Tropical Science 40:47-57.

Bolibok, H., M. Rakocyz-Trojanowska, A. Hromada, and R. Pietrzykowski. 2005. Efficiency of different PCR-based marker systems in assessing genetic diversity among winter rye (*Secale cereale* L.) inbred lines. Euphytica 146:109-116.

Boomsma, C.R., and T.J. Vyn. 2008. Maize drought tolerance: Potential improvements through arbuscular mycorrhizal symbiosis. Field Crops Research 108:14-31.

Boopathi, N.M., B. Thiyagu, M. Urbi, A. Santhoshkumar, S. Gopikrishnan, G. Aravind, K. Swapnashri, and R. Ravikesavan. 2011. Marker-assisted breeding as next-

generation strategy for genetic improvement of productivity and quality: Can it be realized in cotton? International Journal of Plant Genomics 2011:1155-1171.

Bosque-Perez, N.A. 2000. Eight decades of maize streak virus research. Virus Research 71:107-121.

Bosque-Perez, N.A., S.O. Olojede, and I.W. Buddenhaggen. 1998. Effect of maize streak virus disease on the growth and yield of maize as influenced by varietal resistance levels and plant stage at the time of challenge. Euphytica 101:307-317.

Botstein, D., R.L. White, M. Skolnick, and R.W. Davis. 1980. Construction of a genetic linkage in man using restriction fragment length polymorphism. American Journal of Human Genetics 32:314-331.

Brink, M., and G. Belay. 2006. Plant Resources of Tropical Africa – Cereals and pulses. Backhuys Publishers, Wageningen. pp. 298.

Brown, S.M., M.S Hopkins, S.E. Mitchell, M.L. Senior, T.Y. Wang, R.R. F. Duncan, M. Gonzalez-Candelas, and S. Kresovich. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* (L.) Moench). Theoretical and Applied Genetics 93:190-198.

Carducci C., E. Ellul, I. Antonozzi, and A. Pontecorvi. 1992. DNA elution and amplification by polymerase chain reaction from dried blood spots. Biotechnology 13:735-737.

Cardwell, K.F., J.G. Kling, and C.H. Bock. 1997. Methods for screening maize against downy mildew *Peronosclerospora sorghi*. Plant Breeding 116:221-226.

Chateigner-Boutin, A.L., and I. Small. 2007. A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Research 35:1-8.

Clegg, M.T., M. Kobayashi, and J.Z. Lin. 1999. The use of molecular markers in the management and improvement of avocado (*Persea americana* Mill.). Revista Chapingo Serie Horticultura 5:227-231.

Collard, B.C.Y., and D.J. Mackill. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences 363:557-572.

Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer, and E.C.K. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196.

Collins, N.C., F. Tardieu, and R. Tuberosa. 2008. Quantitative trait loci and crop performance under abiotic stress: Where do we stand? Plant Physiology 147: 469-486.

Craig, J., and R.A. Frederiksen. 1983. Differential sporulation of pathotypes of *Peronosclerospora sorghi* on inoculated sorghum. Plant Disease 67:278-279.

Croxford, A.E., P.D.S. Caligari, and M.J. Wilkinson. 2008. High-resolution melt analysis to identify and map sequence-tagged site anchor points onto linkage maps: A white lupin (*Lupinus albus*) map as an exemplar. New Phytologist 180:594-607.

Danson J., M. Lagat, J. Ininda, and M. Kimani. 2006. Application of simple sequence repeats (SSRs) markers to study the resistance of locally adapted maize hybrids to damaging maize streak virus disease. African Journal of Biotechnology 5:1430-1434.

Denic, M.P., C. Chauque, M. Jose, D. Langa, D. Mariote, P. Fato, and W. Haag. 2001. Maize screening for multiple stress tolerance and agronomic traits. Eastern and Southern Africa Regional Maize Conference 7:88-91.

Deulvot, C., H. Charrel, A. Marty, F. Jacquin, C. Donnadieu, I. Lejeune-Hénaut, J. Burstin, and G. Aubert. Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. BMC Genomics 2010 11:468-478.

Do H., and A. Dobrovic. 2009. Limited copy number-high resolution melting (LCN-HRM) enables the detection and identification by sequencing of low level mutations in cancer biopsies. Molecular Cancer 8:82-92.

Dreher, K.M., M. Morris, J.M. Ribaut, S. Pandey, and G. Srinivason. 2000. Is marker-assisted selection cost-effective compared to conventional plant breeding methods? The International Consortium on Agricultural Fourth Annual Conference. Ravello, Italy. http://apps.cimmyt.org/Research/Economics/map/research_results/other_tech/marker_assisted.pdf (Accessed 5 Nov. 2011).

Drescher, A., and A. Graner. 2002. PCR-genotyping of barley seedlings using DNA samples from tissue prints. Plant Breeding 121:228-231.

Duvick D.N. 1999. Heterosis: Feeding people and protecting natural resources. pp. 19-29. *In* J.G. Coors and S. Pandey (eds.) The Genetics and Exploitation of Heterosis in Crops. CIMMYT, Mexico City.

Efron, Y., S.K. Kim, J.M. Fajemisin, J.H. Mareck, C.Y. Tang, Z.T. Dabrowski, H.W. Rossel, G. Thottappilly, and I.W. Buddenhagen. 1989. Breeding for resistance to maize streak virus: A multidisciplinary team approach. Plant Breeding 103:1-36.

Fabiyi, E.F., B.B. Danladi, K.E. Akande, and Y. Mahmood. 2007. Role of women in agricultural development and their constraints: A case study of Biliri local government area, Gombe State, Nigeria. Pakistan Journal of Nutrition 6:676-680.

Fajemisin, J.M. 2003. Overview of maize viruses in sub-Saharan Africa. pp. 158-171. *In* J.H. Hughes and J. Odu (eds.) Plant Virology in Sub-Saharan Africa. International Institute of Tropical Agriculture Conference, Ilesa, Nigeria.

Fakorede, M.A.B., J.M. Fajemisin, J.L. Ladipo, S.O. Ajala, and S.K. Kim. 2003. Development and regional deployment of streak virus resistant maize germplasm. pp. 503-516. *In* J.H. Hughes and J. Odu (eds.) Plant Virology in Sub-Saharan Africa. International Institute of Tropical Agriculture Conference, Ilesa, Nigeria.

FAO and WFP Report. 2010. Special report - Crop and food security assessment mission to Mozambique. Global Information and Early Warning Systems on Food and Agriculture. http://www.fao.org/docrep/012/ak350e/ak350e00.htm (Accessed 4 Jan. 2011).

Farooq, S., and F. Azam. 2002. Molecular markers in plant breeding 1: Concepts and characterisation. Pakistan Journal of Biological Sciences 10:1135-1140.

Frederiksen, R.A., and B.L. Renfro. 1977. Global status of maize downy mildew. Annual Review of Phytopathology 15:249-275.

Fuller, C. 1901. Mealie variegation. First Report of the Government Entomologist 1899-1900. Davis and Sons, Pietermaritzburg. pp. 17-19.

Ganopoulos, I., A. Argiriou, and A. Tsaftaris. 2011. Microsatellite high resolution melting (SSR-HRM) analysis for authenticity testing of protected designation of origin (PDO) sweet cherry products. Food Control 22:532-541.

Garcia-Olmedo, F. 2002. Plant molecular breeding: Objectives, safety issues and prospects. European Review 10:409-416.

George, M.L.C., B.M. Prasanna, R.S. Rathore, T.A.S. Setty, F. Kasim, M. Azrai, S.K. Vasal, O. Balla, D. Hautea, A. Canama, E. Regalado, M. Vargas, M. Khairallah, D. Jeffers, and D. Hoisington. 2003. Identification of QTL conferring resistance to downy mildews of maize in Asia. Theoretical and Applied Genetics 107:544-551.

Gilbert, G.S. 2002. Evolutionary ecology of plant diseases in natural ecosystems. Annual Review of Phytopathology 40:13-43.

Gonzalez-Chavire, M.M., I. Torres-Pacheco and E. Villordo-Pineda. 2006. DNA markers. Advances in Agricultural and Food Biotechnology 2:99-134.

Gowda, K.T.P., S. Rajashekaraiah, B.J. Gowda, and B.S. Naidu. 1987. Mandya maize hybrid, a boon for the farmers in downy mildew disease endemic areas. Current Research 16:168-169.

Grisi, M.C.M., M.W. Blair, P. Gepts, C. Brondani, P.A.A. Pereira, and R.P.V. Brondani. 2007. Genetic mapping of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population BAT93 x Jalo EEP558. Genetics and Molecular Research 6:691-706.

Gudrun, H.R., and C.T. Wittwer. 2004. Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clinical Chemistry 50:1748-1758.

Gupta, P.K. 2008. Molecular biology and genetic engineering. Rastogi Publications, New Delhi, India. pp. 614.

Guthrie, R, and A. Susi. 1963. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Paediatrics 32:338-343.

Hai, L.X., F. Hua, Z.S. Huang, Y.L. Xing, and L.M. Shun. 2000. RFLP detection of genetic variation of maize inbred lines. Acta Botanica Sinica 11:1156-1161.

Hamblin, M.T., M.L. Warburton, and E.S. Buckler. 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. PLoS ONE 12:1-9.

Han, Y., D.M. Khu, and M.J. Monteros. 2011. High-resolution melting analysis for SNP genotyping and mapping in tetraploid alfalfa (*Medicago sativa* L.). Molecular Breeding 29:489-501.

Harkins, G.W., D.P. Martin, S. Duffy, A.L. Monjane, D.N. Shepherd, O.P. Windram, B.E. Owor, L. Donaldson, T.V. Antwerpen, R.A. Sayed, B. Flett, M. Ramusi, E.P. Rybicki, M. Peterschmitt, and A. Varsani. 2009. Dating the origin of the maize-adapted strain of maize streak virus, MSV-A. Journal of General Virology 90:3066-3074.

Hayden, M.J., T.M. Nguyen, A. Waterman, and K.J. Chalmers. 2008. Multiplex-ready PCR: A new method for multiplexed SSR and SNP genotyping. BMC Genomics 90:80-92.

He, G., R. Meng, M. Newman, G. Gao, R.N. Pittman, and C. Prakash. 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). Plant Biology 3:1-6.

Heckenberger, M., J.R. Van der Voort, A.E. Melchinger, and J. Peleman. 2003. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties: II. Genetic and technical sources of variation in AFLP data and comparison with SSR data. Molecular Breeding 12:97-106.

Hoisington, D. 2001. Application of biotechnology to maize improvement: Past, present and future prospects. Seventh Eastern and Southern African Regional Maize Conference, Mexico. pp. 7-11.

Hondow H.L., S.B. Fox, G. Mitchell, R.J. Scott, V. Beshay, S.Q. Wong, and A. Dobrovic. 2011. A high-throughput protocol for mutation scanning of the BRCA1 and BRCA2 genes. Cancer 11:265-276.

Ibitoye, D.O., and P.E. Akin-Idowu. 2010. Marker-assisted selection (MAS): A fast track to increase genetic gain in horticultural crop breeding. African Journal of Biotechnology 9:8889-8895.

Ignjatovic-Micic, D., T. Coric, D. Kovacevic, K. Markovic, and V. Lazic-Jancic. 2003. RFLP and RAPD analysis of maize (*Zea mays* L.) local populations for identification of variability and duplicate accessions. Maydica 48:153-159.

ISAAA. 1999. Advances in maize streak virus disease research in Eastern and Southern Africa. ISAAA Briefs 16:43.

Jeger, M.J., E. Gilijamse, C.H. Bock, and H.D. Frinking. 1998. The epidemiology, variability and control of the downy mildews of pearl millet and sorghum, with particular reference to Africa. Plant Pathology 47:544-569.

Jenkins S., and N. Gibson. 2002. High-throughput SNP genotyping. Comparative and Functional Genomics 3:57-66.

Johanson, H.C., V. Hyland, C. Wicking, and R.A. Sturm. 2009. DNA elution from buccal cells stored on Whatman FTA Classic Cards using a modified methanol fixation method. Biotechniques 46:309-311.

Karanja, J., N.O. Amugune, J. Ininda, J.N. Kimatu, and J.W. Danson. 2009. Microsatellite analysis of the correlation between molecular and morphological traits in assorted maize inbred lines. African Crop Science Journal 17:133-144.

Karp, A., O. Seberg, and M. Buiatti.1996. Molecular techniques in the assessment of botanical diversity. Annals of Botany 78:143-149.

Kim, S.K., J.G. Kling, J.E. Iken, K.F. Cardwell, A.O. Adenola, and V. Adenle. 1994. Breeding maize for downy mildew resistance eradication strategy for downy mildew. pp. 26-33. *In* K.F. Cardwell (ed.) Eradication Strategy for Downy Mildew. International Institute of Tropical Agriculture, Ibadan, Nigeria.

Kinoshita, E., E.K. Kikuta, and T. Koike. 2009. Phosphate-affinity polyacrylamide gel electrophoresis for SNP genotyping. Methods in Molecular Biology 578:183-192.

Kostova, A., E. Todorovska, N. Christov, V. Sevov, and A.I. Atanassov. 2006. Molecular characterisation of Bulgarian maize germplasm collection via SSR markers. Biotechnology and Biotechnological Equipment 20:29-36.

Kyetere, D.T., R. Ming, M.D. McMullen, R.C. Pratt, J. Brewbaker, and T. Musket. 1999. Genetic analysis of tolerance to maize streak virus in maize. Genome 42:20-26.

Lagat, M., J. Danson, M. Kimani, and A. Kuria. 2008. Quantitative trait loci for resistance to maize streak virus disease in maize genotypes used in hybrid development. African Journal of Biotechnology 7:2573-2577.

Lakhanpaul, S., S. Chadha, and K.V. Bhat. 2000. Random amplified polymorphic DNA (RAPD) analysis in Indian mungbean [*Vigna radiata* (L) Wilczek] cultivars. Genetica 109:227-234.

Lecoq, H., B. Moury, C. Desbiez, A. Palloix, and M. Pitrat. 2004. Durable virus resistance in plants through conventional approaches: A challenge. Virus Research 100:31-39.

Li, Y.D., Z. Chu, X. Liu, H. Jing, and Y. Liu. 2010. A cost-effective high-resolution melting approach using the EvaGreen dye for DNA polymorphism detection and genotyping in plants. Journal of Integrative Plant Biology 52:1036-1042.

Lubberstedt, T., A.E. Melchinger, C. Duble, M. Vuylsteke, and M. Kuiper. 2000. Relationships among early European maize inbreds: IV. Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD and pedigree data. Crop Science 40:783-791.

Mader, E., B. Lukas, and J. Novak. 2008. A strategy to setup codominant microsatellite analysis for high-resolution-melting-curve analysis (HRM). Genetics 9:69.

Magenya, O.E.V., J. Mueke, and C. Omwega. 2008. Significance and transmission of maize streak virus disease in Africa and options for management. African Journal of Biotechnology 7:4897-4910.

Maize genetics and genomics database. http://www.maizegdb.org/ (Accessed 6 Jul. 2010).

Mariote, D. 2007. Response to selection for downy mildew (*Peronosclerospora sorghi*) and maize streak virus resistance in three quality protein maize populations in Mozambique. PhD Thesis, University of KwaZulu-Natal. Pietermaritzburg.

Marti, A.F., B. Athanson, T. Koepke, C.F. Forcada, A. Dhingra, and N. Oraguzie. 2012. Genetic diversity and relatedness of sweet cherry (*Prunus avium* L.) cultivars based on single nucleotide polymorphic markers. Plant Science 3:116-128.

Martin, D.P., and D.N. Shepherd. 2009. The epidemiology, economic impact and control of maize streak disease. Food Security 1:305-315.

Masojc, P. 2002. The application of molecular markers in the process of selection. Cellular and Molecular Biology Letters 7:499-509.

Mawere, S., V. Vincent, J. De Meyer, and K.V. Pixley. 2006. Resistance of four inbred maize lines to inoculation with 20 isolates of maize streak virus from Zimbabwe. Plant Disease 90:1485-1489.

Mbogori, M.N., M. Kimani, A. Kuria, M. Lagat, and J.W. Danson. 2006. Optimization of FTA technology for large scale plant DNA isolation for use in marker-assisted selection. African Journal of Biotechnology 5:693-696.

McClure, M.C., S.D. McKay, R.D. Schnabel, and J.F. Taylor. 2009. Assessment of DNA extracted from FTA cards for use on the IIIuminai Select Bead Chip. BMC Research Notes 2:107-110.

McCouch, S.R., X. Chen, O. Panaud, S. Temnykh, Y. Xu, Y. Cho, N. Huang, T. Ishii, and M.W. Blair. 1997. Microsatellite marker development, mapping and applications in rice. Genetics and Breeding 35:89-99.

McMullen, M.D., and K.D. Simcox. 1995. Genomic organisation of disease and insect resistance in maize. Molecular Plant-Microbe Interactions 8:811-815.

Meldrum, D. 2000. Automation for genomics, Part Two: sequences, microarrays and future trends. Genome Research 10:1288-1308.

Mignouna, H.D., N. NQ, J. Ikea, and G. Thottappilly. 1998. Genetic diversity in cowpea as revealed by random amplified polymorphic DNA. Journal of Genetics and Breeding 53:151-159.

Molnar, S., J.S. Rai, M. Charette, and E.R. Cober. 2003. Simple sequence repeat (SSR) markers linked to E1, E3, E4 and E7 maturity genes in soybean. Genome 46:1024-1036.

Montgomery, J., C.T. Wittwer, R. Palais, and L. Zhou. 2007. Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2:59-66.

Mullineaux, P.M., J. Danson, B.A.M. Morris-Krsinich, M.I. Boulton, and J.W. Davies. 1984. The nucleotide sequence of maize streak virus DNA. The EMBO Journal 3:3063-3068.

Ndunguru, J., N.J. Taylor, J. Yadar, H. Aly, J.P. Legg, T. Areling, G. Thompson, and C.M. Fauquet. 2005. Application of FTA technology for sampling, recovery and

molecular characterisation of viral pathogens and virus-derived transgenes from plant tissues. Virology Journal 2:45.

Odendo, M., H. De Groote, and O.M. Odongo. 2001. Assessment of farmers' preferences and constraints to maize production in the moist mid-altitude zone of Western Kenya. Fifth International Conference of the African Crop Science Society, Lagos, Nigeria. http://www.syngentafoundation.org/temp/kenya_irma_maize_production.pdf (Accessed 16 Feb. 2011).

Owor, B.E., D.P. Martin, D.N. Shepherd, R. Edema, A.L. Monjane, E.P. Rybicki, J.A. Thomson, and A. Varsani. 2007. Genetic analysis of maize streak virus isolates from Uganda reveals widespread distribution of a recombinant variant. Journal of General Virology 88:3154-3165.

Parlevliet, J.E. 1995. Genetic and breeding aspects of durable resistance of crops to pathogens. African Crop Science Journal 3:1-13.

Perchepied, L., M. Bardin, C. Dogimont, and M. Pitrat. 2005. Relationship between loci conferring downy mildew and powdery mildew resistance in melon assessed by quantitative trait loci mapping. Phytopathology 95:556-566.

Pernet, A., D. Hoisington, J. Franco, M. Isnard, D. Jewell, C. Jiang, J.L. Marchand, B. Reynaud, J.C. Claszmann and D. Gonzalez de Leon. 1999. Genetic mapping of maize streak virus resistance from the mascarene source. Resistance in line D211 and stability against different virus clones. Theoretical and Applied Genetics 99:524-539.

Pingali, P.L., and S. Pandey. 2001. Meeting world maize needs: Technological opportunities and priorities for the public sector. CIMMYT, Mexico. http://apps.cimmyt.org/Research/Economics/map/facts_trends/maizeft9900/pdfs/maizeft9900.pdf. (Accessed 25 Nov. 2010).

Plumb-Dhindsa, P., and A.M. Mondjane. 1984. Index of plant diseases and associated organisms of Mozambique. Tropical Pest Management 30:407-429.

Powell, W.G., G. Machray, and J. Provan.1996. Polymorphism revealed by simple sequence repeats. Trends in Plant Science 1:215-222.

Redinbaugh, M.G., M.W. Jones, and R.E. Gingery. 2004. The genetics of virus resistance in maize (*Zea mays* L.). Maydica 49:183-190.

Reed, G.H., J.O. Kent, and C.T. Wittwer. 2007. High resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics 8:597-608.

Robinson, R.A. 2006. Return to resistance: Breeding crops to reduce pesticide dependence. Sharebooks Publishing, Winnipeg. pp. 148.

Ryynanen, H.J., A. Tonteri, A. Vasemagi, and C.R. Primmer. 2007. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). Journal of Heredity 98:692-704.

Saliba-Colombani, V., M. Causse, L. Gervais, and J. Philouze. 2000. Efficiency of RFLP, RAPD and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome 43:29-41.

Santana, Q.C., M.P.A. Coetzee, E.T. Steenkamp, O.X. Mlonyeni, G.N.A. Hammond, M.J. Wingfield, and B.D. Wingfield. 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. Biotechniques 46:217-223.

Saxena, K.M.S., and A.L. Hooker. 1968. On the structure of a gene for disease resistance in maize. Genetics 61:1300-1305.

Semagn K., N. Bjornstad, and M.N. Ndjiondjop. 2006. Progress and prospects of marker assisted back-crossing as a tool in crop breeding programmes. African Journal of Biotechnology 5:2588-2603.

Sharma K., and M.S. Misra. 2011. Molecular approaches towards analyzing the viruses infecting maize (*Zea mays* L.). Journal of General and Molecular Virology 3:1-17.

Shepherd, D.N., T. Mangwende, D.P. Martin, M. Bezuidenhout, J.A. Thomson, and E.P. Rybicki. 2007. Inhibition of maize streak virus (MSV) replication by transient and transgenic expression of MSV replication-associated protein mutants. Journal of General Virology 88:325-336.

Smith J.S.C., and O.S. Smith. 1989. Comparison of heterosis among hybrids as a measure of relatedness with that to be expected on the basis of pedigree. Maize Genetics 63:86-87.

Steer, P.A., N.C. Kirkpatrick, D. O'Rourke, and A.H. Noormohammadi. 2008. Classification of fowl adenovirus serotypes by use of high-resolution melting-

curve analysis of the hexon gene region. Journal of Clinical Microbiology 47:311-321.

Stevens, R. 2008. Prospects for using marker-assisted breeding to improve maize production in Africa. Journal of the Science of Food Agriculture 88:745-755.

Stoep, N., C. Paridon, T. Janssens, P. Krenkova, A. Stambergova, M. Macek, G. Matthijs, and E. Bakker. 2009. Diagnostic guidelines for high-resolution melting curve (HRM) analysis: An interlaboratory validation of *brca1* mutation scanning using the 96-well LightScanner^{TM}. Human Mutation 30:899-909.

Studer, B., L.B. Jensen, A. Fill, and T. Asp. 2009. ‘‘Blind’’ mapping of genetic DNA sequence polymorphisms in *Lolium perenne* L. by high resolution melting curve analysis. Molecular Breeding 24:191-199.

Sultan, D.M., M.M. Khalil, A.S. Abdouh, W.F. Poleh, and A.A.M. Al Muthanna. 2009. Imported malaria in United Arab Emirates: Evaluation of a new DNA extraction technique using nested PCR. Korean Journal of Parasitology 47:227-233.

Tanksley, S.D., N.D. Young, A.H. Paterson, and M. Bonierbale. 1989. RFLP mapping in plant breeding: New tools for an old science. Biotechnology 7:257-264.

Taranenko, N.I., N.T. Potter, S.L. Allman, V.V. Golovlev and C.H. Chen. 1999. Detection of trinucleotide expansion in neurodegenerative disease by matrix-assisted laser desorption: Ionization time-of-flight mass spectrometry. Genetic Analysis: Biomolecular Engineering 15:25-31.

Thormann, C.E., M.E. Ferreira, L.E.A. Camargo, J.G. Tiranga, and T.C. Osborn.1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theoretical and Applied Genetics 88:973-980.

Thottappilly, G., N.A. Bosque-Perez, and I.I.W. Rossel. 1993. Viruses and virus diseases of maize in tropical Africa. Plant Pathology 42:494-509.

Trick, M., N.M. Adamski, S.G. Mugford, C.C. Jiang, M. Febrer and C. Uauy. 2012. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biology 12:14-31.

Ulukan, H. 2011. The use of plant genetic resources and biodiversity in classical plant breeding. Soil and Plant Science 61:97-104.

Van Inghelandt, D., A.E. Melchinger, C. Lebreton, and B. Stich. 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theoretical and Applied Genetics 120:1289-1299.

Varshney, R.K., K. Chabane, P.S. Hendre, R.K. Agarwal, and A. Graner. 2007. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barley. Plant Science 173:638-649.

Vigouroux, Y., S. Mitchell, Y. Matsuoka, M. Hamblin, S. Kresovidi, J.S.C. Smith, J. Jaqueth, O.S. Smith, and J. Doebley. 2005. An analysis of genetic diversity across the maize genome using microsatellites. Genetics 169:1617-1630.

Virmani, S.S., Z.X. Sun, T.M. Mou, A.A. Jauhar, and C.X. Mao. 2003. Two-Line Hybrid Rice Breeding Manual. International Rice Research Institute, Los Baños, Philippines. pp. 88.

Visarada, K.B.R.S., K. Meena, C. Aruna, S. Srujana, N. Saikishore, and N. Seetharama. 2009. Transgenic breeding: Perspectives and prospects. Crop Science 49:1555-1563.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hoernes, A. Frijters, J. Pot, J.D. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Vuylsteke, M., J.D. Peleman, and M.J.T. Van Eijk. 2007. AFLP technology for DNA fingerprinting. Nature Protocols 2:1387-1398.

Warburton, M., X. Xianchun, S. Ambriz, L., Diaz, E. Villordo, and D. Hoisington. 2001. Use of molecular markers in maize diversity studies at CIMMYT. Eastern and Southern Africa Regional Maize Conference 7:130-133.

Welsh, J., and M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18:7213-7218.

Welz, H.G., A. Schechert, A. Pernet, K.V. Pixley, and H.H. Geiger. 1998. A gene for resistance to the maize streak virus in the African CIMMYT maize inbred line CML202. Molecular Breeding 4:147-154.

White, H.E., V.J. Hall, and N.C.P. Cross. 2007. Methylation-sensitive high-resolution melting-curve analysis of the SNRPN gene as a diagnostic screen for Prader-Willi and Angelman syndromes. Clinical Chemistry 53:1960-1962.

William, F., and N.J. Grunwald. 2010. Introduction to Oomycetes. http:www.apsnet.org/edcenter/intropp/PathogenGroups/Pages/IntroOomycetes.aspx (Accessed 17 May 2011).

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535.

Willment, J.A., D.P. Martin, and E.P. Rybicki. 2001. Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. Journal of Virological Methods 93:75-87.

Wisser, R.J., P.J. Kurti-Balint, and R.J. Nelson. 2006. The genetic architecture of disease resistance in maize: a synthesis of published studies. Phytopathology 96:120-129.

Xu, Y., and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: From publications to practice. Crop Science 48:391-407.

Yan J., X. Yang, T. Shah, H. Sa'nchez-Villeda, J. Li, M. Warburton, Y. Zhou, J.H. Crouch and Y. Xu. 2009. High-throughput SNP genotyping with the Golden Gate assay in maize. Molecular Breeding 25:441-451.

Young, N.D. 1996. QTL mapping and quantitative disease resistance in plants. Annual Reviews of Phytopathology 34:497-501.

Yu, K., S.J. Park, V. Poysa, and P. Gepts. 2000. Markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). The Journal of Heredity 91:429-434.

Yu, R.H., X.H. Shan, S. Wang, X.H. Li, Y. Jiang, H. Tan, and Y.D. Li. 2011. A screening method for detecting simple sequence repeat (SSR) polymorphism of *Zea mays* using high-resolution melting curve analysis. African Journal of Biotechnology 10:16443-16447.

Yuste-Lisbona, F.J., C. Capel, J. Capel, R. Lozano, M.L. Gomez-Guillamon, and A.I. Lopez-Sese. 2008. Conversion of an AFLP fragment into one dCAPS marker linked to powdery mildew resistance in melon1. *In* M. Pitrat (ed.) EUCARPIA Meeting on Genetics and Breeding of *Cucurbitaceae.* INRA, Avignon.

Zhang, W., N.H. Olson, T.S. Baker, L. Faulkner, M. Agbandje-McKenna, M.I. Boulton, J.W. Davies, and R. McKenna. 2001. Structure of the maize streak virus geminate particle. Virology 279:471-477.

Zhong, M., J.C. McCarthy, J.N. Jenkins, and S. Saha. 2002. Assessment of day-neutral backcross populations of cotton using AFLP markers. Journal of Cotton Science 6:97-103.

CHAPTER TWO

Genetic diversity of maize germplasm lines and implications for breeding maize streak virus resistant hybrids

Abstract

Evaluation of genetic distances between maize inbred lines can be used to identify lines that can maximise heterosis in hybrids. This study was conducted to determine genetic distances among 25 maize inbred lines to enable predictions of the best combinations for developing maize streak virus resistant hybrids. The inbred lines were genotyped using 19 simple sequence repeat (SSR) markers which are known to be associated with maize streak virus resistance in maize. All amplification products were in the range of 68-290 bp. In total, 94 SSR alleles were detected, with a mean of 4.95 alleles per locus. The average polymorphic information content (PIC) value was 0.56. Gene diversity (He) values ranged from 0.00-0.853 with an average of 0.594, while heterozygosity (Ho) values attained an average of 0.039, ranging from 0.00-0.095. The 25 lines were grouped into three major clusters and five sub-clusters, from which potential breeding lines could be sampled. These results confirm the diversity found among the maize lines used in this study. In general, the grouping of the inbred lines by SSR markers is consistent with pedigree information. The genetic distance data obtained using SSR markers was useful in identifying lines that could be used to design new hybrids and new breeding populations. Overall the findings form the basis for organising the germplasm lines in the breeding programme.

Keywords: Hybrid vigour, maize, molecular breeding, simple sequence repeats (SSRs).

2.1 Introduction

Maize is the chief support of most of Africa's rural economies (Oluwafemi *et al*., 2008; Stevens, 2008), including Mozambique, in which it is the staple food and is grown in all of its agro-ecological zones (Denic *et al*., 2001). Despite this, the average yield of maize

in Africa is the lowest in the world and, as a result, fails to meet the high demands in most Sub-Saharan countries (Denic *et al*., 2001; Magenya *et al*., 2008). Numerous viral pathogens affect maize productivity, with maize streak virus (MSV) disease being considered as the most significant biological threat to food security in Africa (Bosque-Perez, 2000; Shepherd *et al*., 2007; Martin and Shepherd, 2009). This is a major concern and can only be addressed by solving production constraints to improve crop yields.

Virus resistance is usually associated with one or two major resistance loci, which facilitate marker-assisted selection (MAS), but resistance genes have been found to cluster in the maize genome (Redinbaugh *et al*., 2004). Genomic regions associated with resistance to the MSV disease have been identified in several studies using different populations in diverse environments and these studies have revealed that resistance is quantitatively inherited with a varying number of genes involved (Pernet *et al*., 1999; Welz *et al*., 1998; Mawere *et al*., 2006). Mawere *et al*. (2006) reported that MSV resistance is expressed by a major gene and two or three modifying genes. Pernet *et al*. (1999) also proposed that MSV resistance was controlled by two genetic systems, one from a major gene on the short arm of chromosome 1 and the other conditioned by minor genes on chromosomes 2, 3 and 10, that confer quantitative resistance. Minor quantitative trait loci (QTL) effects have been detected at bins 3.06, 5.03 and 8.07 (Asea *et al*., 2008). The major QTL, designated *msv1*, was identified on the short arm of chromosome 1 (1S – bin1.04) (Welz *et al*., 1998; Kyetere *et al*., 1999; Mawere *et al*., 2006). The stability of QTL across populations has been shown to be variable; however, this is not the case for maize streak virus disease (MSVD) QTL (Pernet *et al*., 1999).

A study conducted by Danson *et al*. (2006) used three DNA markers: bnlg1811, umc1917 and umc1144, which are contained between bin 1.04 and 1.05 of maize chromosome 1 to screen 115 recombinant inbred lines (RILs) for resistance to MSV disease. These markers were able to differentiate resistant from susceptible lines**.** A study by Asea *et al*. (2008) further examined a consensus MSV QTL in bin 1.04 as a potential target for selection in improving host resistance. Maize streak field evaluations and subsequent selections were conducted in Zimbabwe in a population of 410 $F_{2:3}$ lines

derived from hybridisation between inbred line CML202 with known resistance to MSV and the susceptible line VP31. It was concluded that the major locus conferring resistance to MSV on chromosome 1 was significant ($P<0.05$) for resistance across seasons and explained 23% of phenotypic variations in the $F_{2:3}$ generation. Markers used for this current study were developed by Danson *et al*. (2006); Asea *et al*. (2008), the maize database (http://www.maizegdb.org) and from a study by Lagat *et al*. (2008) in which the QTL for resistance to MSVD in one resistant source MAL13 crossed to one elite line, MAL9, were mapped using SSRs. Conventional maize breeders may benefit from the use of molecular markers in order to improve selection intensity and maximise genetic gain (Collard *et al*., 2005).

The adoption of hybrids in maize production has resulted in increased yields across the world (Warburton *et al*., 2002). Maize breeding relies on the available genetic diversity (Karanja *et al*., 2009). Improved hybrids of maize are developed by making use of information on the genetic relationships and diversity among elite materials (Dias *et al*., 2003; Diniz *et al*., 2005). Evaluating genetic diversity among the elite lines aids in the estimation of genetic variation and thus the degree of heterosis to be expected among segregating progeny for pure-line cultivar development (Biswas *et al*., 2008; Salem *et al*., 2008; Karanja *et al*., 2009).

Heterosis, also known as hybrid vigour, is a phenomenon in which the offspring show superiority over their parents either in yield, vigour, increased size, rate of growth or other reproductive factors (Duvick, 1999; Virmani *et al.,* 2003). The term coined by Shull (1952) can be used for the expression of adaptive traits like increased resistance to disease and drought tolerance, with the hybrid of choice exceeding the best parent in superiority. However, superiority is lost with every successive generation of self-fertilisation, thus maximum heterosis is expressed in the F_1 generation (Meyer *et al*., 2004). The manifestation of heterosis depends on genetic divergence of the two parental varieties (Hallauer and Miranda, 1988). Morphological, pedigree, physiological, biochemical and molecular data can be used to identify elite inbred lines to be crossed for a superior hybrid (Smith and Smith, 1989). However, molecular markers can detect variation at the DNA sequence level (Diniz *et al*., 2005) and genetic distances (GD) are

used to group similar germplasm as the first step in identifying potentially useful heterotic patterns (Melchinger, 1999).

Simple sequence repeat (SSR) markers, also known as microsatellites (He *et al*., 2003; Molnar *et al*., 2003), have been extensively used to characterise germplasm collections in major cereal crops including wheat (Salem *et al*., 2008; Ijaz and Khan, 2009) and maize (Taramino and Tingey, 1996; Smith *et al*., 1997; Li *et al*., 2002; Danson *et al*., 2006; Aguiar *et al*., 2008; Cholastova *et al*., 2011). Different repeat numbers in SSRs can be treated as separate "alleles" and the site can be treated as highly polymorphic with multiple alleles for the detection of variation in populations (Akkaya *et al*., 1992). Microsatellites are highly abundant, simple to analyse, co-dominant, economical and are easily assayed using PCR with primers specific to conserved regions flanking the repeat array (Yu *et al*., 2000). Compared with other marker types, SSRs are advantageous due to their abundance in plant genomes and large number of alleles per locus making them highly polymorphic even among closely related cultivars due to naturally occurring mutations, and thus they can distinguish between closely related species (Brown *et al*., 1996; Weising *et al*., 2005), providing greater power of discrimination. Hence, they are useful for assigning heterotic groups for maize lines (Enoki *et al*., 2002; Li *et al*., 2002; Xia-Su *et al*., 2004).

The objectives of the study were, therefore, to determine the genetic diversity among the 25 maize inbred lines using 19 SSR markers which are known to be associated with MSV disease resistance in maize. The information will be used in the selection of the most appropriate parents out of the potential MSV resistance donors for the introgression of the MSV resistant gene, *msv1* into the different Mozambican lines that are adapted to the lowland environment but are susceptible to MSV. The information would be crucial in devising future hybrid breeding programmes that will emphasise MSV resistance, in Mozambique.

2.2 Materials and methods

2.2.1 Plant material

A total of 25 maize inbred lines were used in this study. All maize varieties, along with their pedigrees, MSV and DM disease ratings are listed in Table 2.1. Some of the maize inbred lines were from CIMMYT, Zimbabwe (CML505, CML509, P13, P14, P15, P16, P17, P18) and PA1 from the Crop Breeding Institute (CBI) in Zimbabwe. The CIMMYT lines are potential donors for MSV resistance (CIMMYT, 2009). Line PA1 has high yielding potential but is susceptible to MSV and adapted to mid-altitude conditions, hence was included as a standard check. Although the P lines (P13, P14, P15, P16, P17 and P18) are MSV resistant, they are susceptible to DM disease and not adapted to the tropical conditions of the lowlands of coastal Mozambique. Lines CML505 and CML509 have baseline resistance to DM and are adapted to the tropical lowland conditions. The LP group of lines and E designated inbred lines developed by the maize research programme in Mozambique are susceptible to the MSV disease (Denic, 2005; Fato, 2010). These lines were therefore suitable recipients of the MSV resistance gene.

2.2.2 DNA extraction

The 25 maize inbred lines were grown in a tunnel at the University of KwaZulu-Natal (UKZN), Pietermaritzburg during 2009. At four weeks from planting, leaf tissue samples (five plants per line) were harvested and total DNA was extracted from each of the genotypes using a CTAB (cetyltrimethylammonium bromide) method. The tissue was macerated in centrifuge tubes with three 5 mm stainless steel beads and shaken in a TissueLyser bead beaker for 2 min. Immediately 500 µl of pre-warmed (65°C) (CTAB) extraction buffer [100 mM Tris-HCl (pH8.0); 700 mM sodium chloride (NaCl); 10mM ethylene diamine tetraacetic acid (EDTA); 2% CTAB; 2% polyvinyl pyrrolidone (PVP-40); 5 mM ascorbic acid and 4 mM sodium diethyldithiocarbamic acid (SDA)] and vortexed for a few seconds. The samples were incubated at 65°C for 60 min followed by cooling on ice for 15 min. Subsequently 200 µl chloroform/isoamyl alcohol (24:1) was

added under the hood and phases were separated by centrifugation for 10 min at 12 000 rpm at room temperature. Then 200 µl of the supernatant was transferred into new sterile tubes, which contained 200 µl cold isopropanol. The precipitated DNA was then collected by centrifugation for 15 min with 12 000 rpm. The supernatant was drained off immediately and then pellets washed with 100 µl 70% cold ethanol. The pellet was dissolved in 100 µl of TE buffer (10 mM Tris and 1 mM EDTA, pH 8.0) and stored at -20^{o}C.

Table 2.1: Features of the 25 maize inbred lines used in the study

Inbred lines	Pedigree / Source population	MSV status	DM Status
CML505	[92SEW1-2/[DMRESR-W]EarlySel-#I-2-4-B/CML386]-B-11-1-B-2-#-BB	R	MR
CML509	[92SEW1-2/[DMRESR-W]EarlySel-#L-2-1-B/CML386]-B-22-1-B-4-#-1-BB	R	MR
P13	HA07145-13-B	R	S
P14	HA07145-14-B	R	S
P15	HA07145-15-B	R	S
P16	HA07145-16-B	R	S
P17	HA07145-19-B	R	S
P18	HA07145-20-B	R	S
PA1	Salisbury white	S	S
LP19	Matuba-6-2-1-1-1-1-X-B-10-2-4-B	S	HR
LP37D	Pop44-1-1-1-4-6-6-X-B-12-2-1-B	S	S
E80	ZM621-56-1-6-1-3-4-B	S	S
LP21	Unknown	S	R
LP37F	Pop44-1-1-1-4-6-6-X-B-12-1-1-B	S	R
E75	ZM621-56-1-2-1-1-3-B	S	S
E77	ZM621-56-1-2-1-1-5-B	S	S
E72	ZM621-30-1-4-1-3-1-B	S	S
E27	ZM621-96-1-2-1-1-3-B	S	S
E71	ZM621-30-1-4-1-1-2-B	S	S
LP23	Tzi- 4	S	HR
E7	(P501SR0/P502SRO) F5-26-1-1-2-2-1-B	S	S
E24	Unknown	S	S
E66	ZM621-19-1-1-1-1-1-B	S	S
E46	ZM521-21-2-1-3-1-1-B	S	S
E47	ZM521-21-2-1-3-1-2-B	S	S

S=Susceptible, MR=Mild resistance, R=Resistant, HR=High resistance.

TZI maize populations are IITA populations improved for downy mildew resistance and some of them for both DM and MSV resistance.

2.2.3 Simple sequence repeats primer selection

Nineteen SSR gene-specific markers for *msv1* which confers resistance to the MSV disease were used to genotype the lines (Table 2.2). The primers which were identified in previous studies (Danson *et al*., 2006; Asea *et al*., 2008; Lagat *et al*., 2008) and from the MaizeGDB database (www.maizegbd.org) were chosen based on bin locations to detect genetic diversity at that region of the genome. Markers linked to the *msv1* gene have mostly been identified on bin 1.04, thus more primers on this bin were selected.

Table 2.2: Primer sequences of the 19 SSR markers used in this study.

Primer name	Forward primer (5´- 3´)	Reverse primer (3´- 5´)	Bin
bnlg490	GCCCTAGCTTGCTAATTAACTAACA	ACTGTAAGGGCAGTGGACCTATA	4.00
bnlg105	GACCGCCCGGGACTGTAAGT	AGGAAAGAAGGTGACGCGCTTTTC	5.00
umc1122	CACAACTCCATCAGAGGACAGAGA	CTGCTACGACATACGCAAGGC	1.06
phi227562	TGATAAAGCTCAGCCACAAGG	ATCTCGGCTACGGCCAGA	1.11
umc1917	ATTTCCACTTCACCAGCCTTTTC	GGAAAGAAGAGCCGCTTGGT	1.04
umc1811	AGATAGCCGCCGAGACCAAG	ACTCACTCGACGGACTTCTCGAC	1.06
umc1144	ATGGCCCACTCATCATATCTCTGT	TCTGTTGATTAGCAGCGGATAAAA	1.04
umc1551	CACCGGAACACCTTCTTACAGTTT	CGAAACCTTCTCGTGATGAGC	2.09
umc2077	CTGGTTCGGATGCAAGTAGTCAG	AAACTCACTGAACATGATCCTGGC	2.09
umc1086	CATGAAAGTTTTCCTGTGCAGATT	GGGCAACTTTAGAGGTCGATTTATT	4.08
umc1559	CTTGCTAGAGTCGGTGAACAACAA	AACCAAGCTCCTTAATGAGGTCAC	4.08
umc1644	CCATAAACTGTTCCTTTGGCACAC	CTTTCACGTGTTAAGGGAGACACC	3.06
umc2169	ACTACTCCTCGGATAGCCACG	GACGAGTAGAGGCTCTGGGAC	3.06
phi330507	GTAAAGTACGATGCGCCTCCC	CGGGGTAGAGGAGAGTTGTG	5.04
umc1221	GCAACAGCAACTGGCAACAG	AAACAGGCACAAAGCATGGATAG	5.04
umc1724	GTCTCAAGTGAAACAACCACGCTT	CCACATGAGATGAGATTGCCATT	8.06
umc0181	CTAATCACCAACCACCAACAC	AGTCCGTCCTCTGTCCTCGTC	8.06
umc1169	TAGCCAACAGTCCAACATTTTTCA	CAGGCTAGAATAACATCCCGAAGA	1.04
bnlg2086	CGGAACCTGCTGCAGTTAAT	GAGATGCAGGAATGGGAAAA	1.04

2.2.4 PCR amplification and detection conditions

The PCR reactions were carried out in a final volume of 12 µl, containing 1 × PCR buffer), 3 mM $MgCl_2$, 2.5 mM dNTPs, 10 mM forward and reverse primers, 0.6 U of *Taq* DNA polymerase and approximately 40 ng of DNA template with the final volume made up with double-distilled sterile water. The PCR reaction was carried out in a touchdown fashion with a first denaturation at 94°C for 2 min, followed by 8 cycles of (1) denaturation at 94°C for 30 s; (2) annealing at 60°C for 30 s; and (3) extension at 72°C for 45 s, with the annealing temperature being reduced by 1°C per cycle. This procedure was followed by 25 cycles of (1) denaturation at 94°C for 30 s; (2) annealing at 52°C for 30 s; and (3) extension at 72°C for 45 s, and a final extension at 72°C for 5 min. The target sequences were amplified by using forward primers fluorescently labelled with either VIC (Green), FAM (Blue), PET (Red) or NED (Yellow). The PCR products were denatured by heating at 95°C for 5 min and separated by capillary gel electrophoresis using an ABI3130 automatic DNA sequencer (Applied Biosystems, Johannesburg, SA). The allele sizes of amplified PCR fragments were identified on the basis of size in comparison with DNA molecular weight markers. The size of the amplified fragments was determined by the software programme GeneMapper 4.1 (Applied Biosystems, Johannesburg, South Africa).

2.2.5 Data analysis

PowerMarker v3.25 (Liu and Muse, 2005) was used to determine PIC, gene diversity and heterozygosity values for each SSR marker used in the study. The expected heterozygosity (He) and observed heterozygosity (Ho) were used to evaluate the genetic diversity within the set of cultivars. Expected heterozygosity, i.e. the probability that two alleles from the same locus would be different when chosen at random, was calculated for each SSR locus according to Nei (1973):

$He = 1 - \sum (pi)^2$.

Observed heterozygosity was calculated by dividing the number of heterozygous individuals by the number of individuals scored. Polymorphic information content (PIC) for the SSR markers in the sample DNA was calculated as:

PIC = $1- \Sigma p_i^2$ where p_i is the frequency of the i^{th} allele in a locus for individual p.

For co-dominant markers like SSRs and RFLPs, data can be scored as allele frequencies and as binary traits (1= allele presence; 0 = allele absence) (CIMMYT, 2002). The allele frequency data from PowerMarker v3.25 was used to export the data in binary format for analysis with NTSYS-PC (Numerical Taxonomy and Multivariate Analysis for Personal Computers) v2.1. NTSYS-PC v2.1 only accepts binary data coding. The 0/1 matrix was used to calculate genetic similarity (GS) based on the DICE coefficient. A dendrogram was constructed using the UPGMA (unweighted paired group method using arithmetic averages) method as implemented in NTSYS-PC v2.1 to infer genetic relationships and phylogeny. The MXComp function was used to determine the R value for the dendrogram (Rohlf, 1998).

2.3 Results

2.3.1 Genetic diversity

A total of 94 alleles were amplified among 25 maize inbred lines and the numbers of alleles scored for SSR loci ranged from one to nine. The average number of alleles was 4.95. The maximum number of alleles (nine) was detected at the bnlg105 locus. The PIC value of the SSR markers, which is a measure of allele diversity at a locus, ranged from 0.000 to 0.838 with an average of 0.556 (Table 2.3). Nine SSR loci (bnlg490, bnlg105, umc1122, umc1811, umc1086, umc1551, umc2169, phi330507 and umc1221) exhibited PIC values higher than 0.6 indicating that they were potentially informative in detecting differences among the inbred lines.

Di- and tri-nucleotides repeat markers were the most abundant. The expected heterozygosity (He) values were in the range of 0.00 (phi227562, umc1917) to 0.853 (bnlg105) with a mean gene diversity value of 0.594. The observed heterozygosity (Ho)

values ranged from 0.00 to 0.318 with an average of 0.0387. The most polymorphic SSRs based on PIC value (≥0.75), genetic diversity (≥0.75) and alleles (≥6) were bnlg105, umc1551 and umc1086.

2.3.2 Analysis of maize genotype associations

The estimates of similarity coefficients among the 25 inbred lines ranged from 15% to 94%. The highest similarity index of 0.94 was obtained between P14 and P13, followed by E77 and E75 (0.89) and E47 and E46 (0.86). The lowest similarity value of 0.15 was between E75 and LP21 and also between E77 and LP21 (Table 2.4). The results on similarity were also substantiated by the dendrogram (Figure 2.1).

Table 2.3: Details of polymorphisms and genetic analysis of 19 maize microsatellite markers across the 25 maize inbred lines.

Marker	Repeat types	Allele no.	Size range	PIC	Gene diversity (He)	Heterozygosity (Ho)
bnlg490	-	6	120-140	0.744	0.773	0.000
bnlg105	Di	9	90-120	0.838	0.853	0.000
umc1122	Tri	5	166-191	0.618	0.675	0.000
phi227562	Tri	1	68-88.6	0.000	0.000	0.000
umc1917	Tri	1	170-220	0.000	0.000	0.000
umc1811	Di	7	166-220	0.638	0.671	0.095
umc1144	Di	5	160-180	0.551	0.597	0.280
umc1559	Tri	3	150-160	0.589	0.664	0.042
umc1551	Tetra	6	160-185	0.775	0.803	0.000

umc2077	Tri	5	170-184	0.553	0.597	0.000
umc1086	Di	7	100-120	0.764	0.792	0.000
umc1644	Tri	6	170-188	0.535	0.560	0.000
umc2169	Quad	4	90-110	0.643	0.691	0.000
phi330507	Hepta	5	148-165	0.721	0.760	0.000
umc1221	Di	5	85-117	0.682	0.726	0.000
umc1724	Tetra	5	140-158	0.538	0.613	0.000
umc0181	-	6	90-138	0.571	0.632	0.000
umc1169	Tri	3	160-165	0.468	0.531	0.000
bnlg2086	Di	5	250-290	0.338	0.354	0.318
Mean		**4.95**		**0.556**	**0.594**	**0.039**

The majority of bifurcations in the dendrogram occurred at genetic distances above 0.30. All the 25 maize genotypes were grouped into three major clusters: 1, 2 and 3 at a 40% similarity. Major Cluster 1 (MC1) was further divided into 3 sub-clusters at a 44% similarity coefficient. The sub-clusters derived from MC1 were as follows: Sub-cluster 1.1 consisted of Mozambican inbred lines LP19, LP37D, LP21 and E80; Sub-cluster 1.2 grouped CIMMYT lines P13, P14, P17, P15, P16 and P18 and some Mozambican lines LP23, E27 and E71; and Sub-cluster 1.3 which had the Mozambican line LP37F. Major Cluster 2 (MC2) was divided into 2 sub-clusters at a 48% similarity with the CIMMYT lines CML505 and CML509 being grouped together in Sub-cluster 2.1 at a 54% similarity. The LP23 line was in a different cluster from lines CML505 and CML509, which reflects the wide genetic distances among the lines. The similarity value between LP23 and CML509 was 0.28 and 0.41 between LP23 with CML505.

Lines P13 and P14 of MC1, Sub-cluster 1.2 were the most related with a 94% similarity value. The lowest similarity value of 0.15 was between the lines E75 and LP21 and between E77 and LP21 followed by E27 and E24, E77 and LP37F and also between E72 and LP21 (0.21).

Most of the Mozambican E lines: E66, E7, E46, E47 and E24 were grouped into Sub-cluster 2.2 of MC2 with the CBI line PA1 at the 46% similarity. Major Cluster 3 (MC3) consisted of the lines E72, E75 and E77 at 64% similarity. All the CIMMYT inbred P lines and the Mozambican inbred LP lines clustered in MC1. The R value for the dendrogram was 0.85.

Table 2.4: Similarity matrix for the 25 maize inbred lines based on 19 SSR markers.

	LP19	LP21	LP37D	LP37F	CML505	CML509	PA1	P13	P14	P15	P16	P17	P18	LP23	E7	E24	E27	E46	E47	E66	E71	E72	E75	E77	E80
LP19	1.00																								
LP21	0.55	1.00																							
LP37D	0.63	0.47	1.00																						
LP37F	0.40	0.40	0.47	1.00																					
CML505	0.37	0.37	0.33	0.37	1.00																				
CML509	0.36	0.31	0.38	0.41	0.54	1.00																			
PA1	0.46	0.36	0.38	0.41	0.43	0.42	1.00																		
P13	0.56	0.50	0.47	0.39	0.44	0.39	0.32	1.00																	
P14	0.56	0.50	0.47	0.39	0.50	0.45	0.39	0.94	1.00																
P15	0.39	0.26	0.41	0.53	0.58	0.53	0.60	0.52	0.58	1.00															
P16	0.33	0.47	0.43	0.41	0.40	0.34	0.28	0.53	0.60	0.55	1.00														
P17	0.42	0.55	0.39	0.44	0.45	0.31	0.50	0.65	0.58	0.53	0.48	1.00													
P18	0.31	0.31	0.33	0.45	0.53	0.45	0.39	0.47	0.53	0.55	0.64	0.45	1.00												
LP23	0.48	0.41	0.43	0.34	0.41	0.28	0.48	0.62	0.69	0.57	0.57	0.64	0.50	1.00											
E7	0.31	0.31	0.27	0.39	0.53	0.52	0.65	0.33	0.40	0.55	0.43	0.52	0.60	0.43	1.00										
E24	0.26	0.21	0.32	0.36	0.38	0.47	0.53	0.32	0.39	0.47	0.34	0.38	0.45	0.34	0.45	1.00									
E27	0.46	0.41	0.38	0.46	0.27	0.26	0.47	0.58	0.52	0.47	0.41	0.56	0.45	0.48	0.52	0.21	1.00								
E46	0.31	0.26	0.32	0.26	0.27	0.42	0.47	0.39	0.45	0.40	0.28	0.38	0.32	0.41	0.45	0.53	0.32	1.00							
E47	0.42	0.26	0.44	0.32	0.26	0.43	0.54	0.38	0.44	0.45	0.27	0.32	0.27	0.40	0.47	0.43	0.38	0.86	1.00						
E66	0.39	0.39	0.41	0.44	0.44	0.40	0.69	0.45	0.52	0.60	0.41	0.53	0.48	0.62	0.55	0.46	0.46	0.51	0.56	1.00					
E71	0.45	0.25	0.37	0.30	0.42	0.41	0.36	0.48	0.55	0.63	0.58	0.35	0.55	0.58	0.42	0.41	0.51	0.41	0.42	0.44	1.00				
E72	0.37	0.26	0.28	0.37	0.54	0.32	0.32	0.45	0.52	0.53	0.48	0.39	0.53	0.48	0.47	0.32	0.38	0.22	0.27	0.51	0.47	1.00			
E75	0.41	0.15	0.27	0.26	0.43	0.37	0.37	0.26	0.32	0.40	0.28	0.25	0.45	0.33	0.45	0.32	0.37	0.26	0.38	0.34	0.41	0.65	1.00		
E77	0.36	0.15	0.22	0.21	0.43	0.26	0.32	0.26	0.32	0.33	0.28	0.25	0.45	0.34	0.39	0.37	0.32	0.26	0.27	0.34	0.46	0.65	0.89	1.00	
E80	0.58	0.37	0.61	0.37	0.33	0.27	0.38	0.40	0.47	0.41	0.50	0.30	0.47	0.55	0.40	0.27	0.32	0.27	0.39	0.35	0.42	0.39	0.43	0.38	1.00

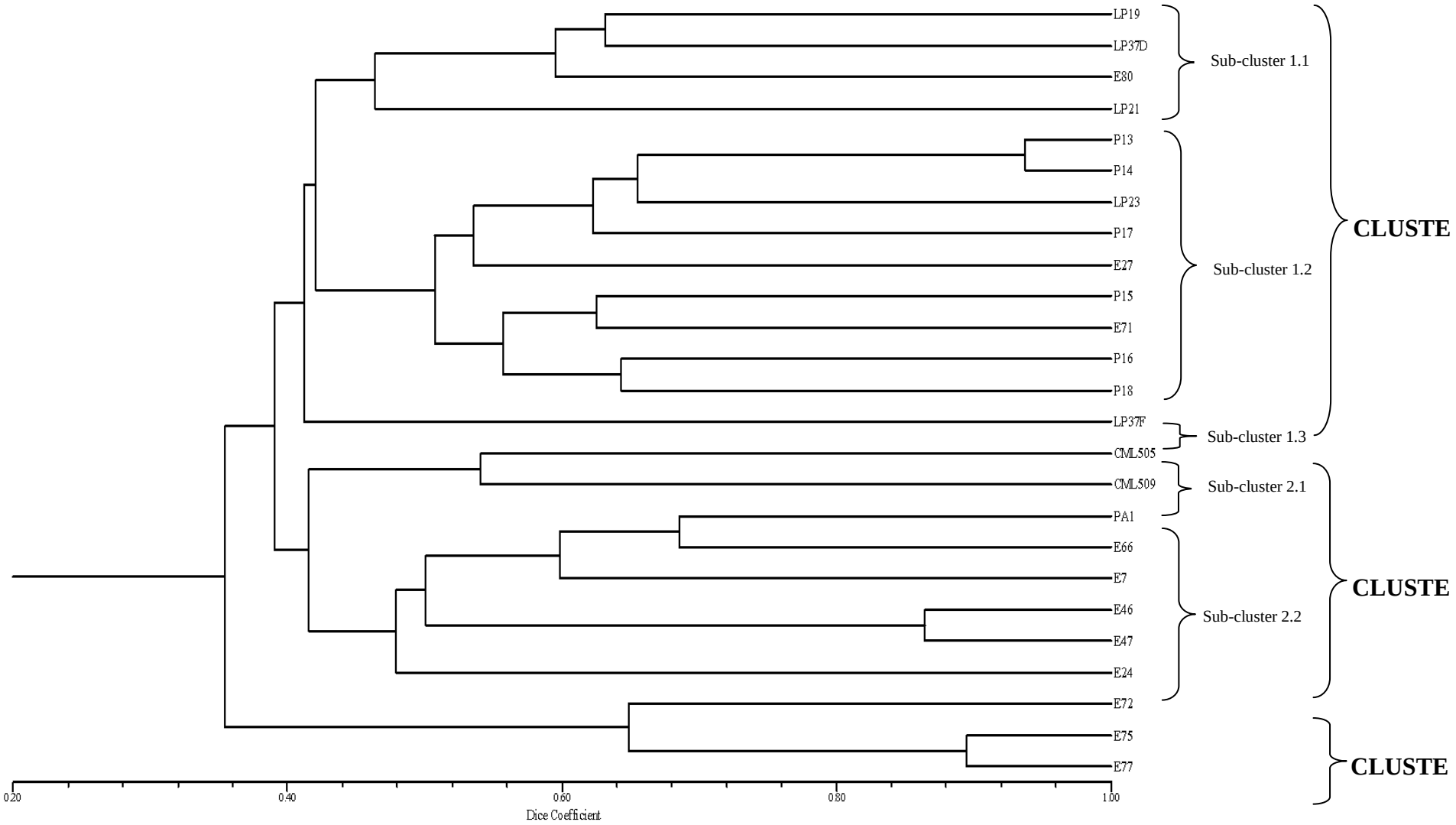

Figure 2.1: UPGMA dendrogram of 25 maize inbred lines based on the dice coefficient calculated using 19 SSR markers.

2.4 Discussion

2.4.1 Analysis of genetic diversity using simple sequence repeats

Genetic diversity within a population can be expressed as the number of alleles detected at a single locus (Hoxha *et al.,* 2004). The average number of alleles obtained per primer in the present study was 4.95, which is comparable to the 5.2 alleles which were reported for 50 SSR markers in 85 tropical maize lines (Laborda *et al*., 2005). However, the average number of alleles observed is smaller than the 7.4 alleles which were reported for 79 SSR markers in 15 lines (Xia-Su *et al*. 2004). The total number of alleles in diversity studies is usually proportional to the sample size (Xia-Su *et al*., 2004), which can partly explain the observed differences between the studies. A study by Liu *et al*. (2003) on 260 inbred lines across tropical, sub-tropical and temperate maize inbred lines characterised with 94 SSRs detected a total of 2 039 alleles with a mean of 21.7 alleles per locus. Higher genetic diversity in that study can be attributed to the more diverse and larger number of inbreds (Vigouroux *et al*., 2005). Sharopova *et al*. (2002) developed 1 051 maize SSR primers and concluded that polymorphism increased significantly with the increase in the number of repeat units. The inclusion of more dinucleotide repeat SSRs, which tend to be more polymorphic than longer repeat motifs, also increases polymorphism levels (Masi *et al*., 2003; Vigouroux *et al*., 2005).

Three of the primers used, bnlg105, umc1551 and umc1086, exhibited high polymorphism. These can be effectively used in future molecular breeding programmes for the identification of highly diverse maize lines in hybrid development focusing on MSV disease resistance. The parameters used for the determination of polymorphic SSR loci were also previously used by Babu *et al*. (2012) and Yao *et al*. (2007) to screen inbred lines in India and China, respectively.

The majority of bifurcations in the dendrogram occurred at genetic distances above 0.30, which was also observed in a study by Weising *et al*. (2005) on indigenous Swedish mandate cultivars. This indicates that the 25 lines tested in the present study were not closely related to one another. Xia *et al*. (2005) used SSR markers to assess genetic diversity in CIMMYT maize lines and the study did not reveal clear clustering amongst the lines assessed. However, in the current study, the genetic similarity index using SSR markers identified three major clusters, MC1, MC2 and MC3. All the CIMMYT P lines were grouped into MC1, which is

consistent with pedigree information. These lines were developed from a common genetic background for MSV resistance. The CIMMYT maize lines, CML505 and CML509, did not cluster in MC1 with the rest of the CML lines, which is also in agreement with pedigree information. These observations are supported by the study by Yuan *et al*. (2002) in which all the CIMMYT lines clustered together according to pedigree. According to Warburton *et al*. (2002), CIMMYT inbred lines are usually drawn from a mixture of populations that contain a broad genetic base. The CML505 and CML509 lines were derived from the same population with adaptation to the tropical lowland and resistance to downy mildew disease. In agreement with pedigree information lines E46 and E47 were placed in the same cluster 2, because both lines were derived from ZM521. However, it was observed that the E-group of lines which were derived from the same population "ZM621" (except E7) were allocated to different clusters with three lines in MC3, two in MC2 and three in MC1, which is not congruent with pedigree information. This can be explained by the fact that the base population ZM621 is broad-based. It was developed as a synthetic hybrid population between heterotic group A and group B lines involving more than eight lines, and has gone through several cycles of recombination at CIMMYT in Harare.

The clustering of the Mozambican inbred LP lines also partly reflects the pedigree information and origin of the lines. The lines LP37D and LP37F are sister lines derived from the same base population Pop44 from CIMMYT, but they are placed in different sub-clusters since LP37D has dent grain texture while the LP37F has flint grain. Lines LP23 and LP19 were developed under lowland tropical conditions in Mozambique, and Nigeria, respectively, and were derived from different populations; hence they were grouped in different sub-clusters within MC1. The actual pedigree data for the line LP21 could not be established, so the observation that it was classified in the same sub-cluster as LP19 suggests that they have a similar genetic background or have similar allele frequencies. These lines may be placed in the same heterotic group in the programme in Mozambique. The importance of the origin of the lines is also reflected by the classification of PA1 in sub-cluster 2.2 together with line E46, E47 and E66, among others, which were derived from the mid-altitude populations ZM521 and ZM621 in Zimbabwe. The genetic distance matrices derived from the SSRs were highly correlated (r = 0.85), indicating that the SSRs have distinguishable power to detect polymorphism and are appropriate for genetic diversity analysis among maize inbred lines.

2.4.2 Implications for breeding new hybrids

Studies report that for the production of hybrids with better yield performance, it is best to use lines with larger genetic distances as parents as these increase genetic variation (Yuan *et al*., 2002; Shahnejat-Bushehri *et al*., 2005; Biswas *et al*., 2008). Small genetic distances reveal that lines are closely related thus should not be crossed with each other if hybrid vigour is to be maximised. Among the 25 inbred lines, E75 and LP21 as well as E77 and LP21 were the most distantly related parental lines with the lowest similarity values of 0.15, indicating that the two lines are divergent and contain different allele frequencies which can be exploited to make hybrids. The lines E24 and LP19 with a low similarity value of 0.26 would also be recommended for use in making hybrids.

The lowland adapted and MSV resistant donor lines (CML505 and CML509) had low similarity values, ranging between 27 and 54%, with the set of lines from Mozambique (P and E series), indicating the existence of large genetic distances and, therefore, are divergent. Hence, it would be expected that productive hybrids could be obtained by making crosses between the CMLs and the Mozambican lines because they have complimentary gene frequency that might result in high heterosis. The fact that LP21 and LP23 were placed in a different cluster with the MSV donor lines CML505 and CML509 suggests that productive hybrids can be produced by combining the DM resistant lines with the MSV resistant lines, respectively. The lines CML505 and CML509 were previously characterised, and displayed MSV resistance with ratings of 1.5 and 2, respectively (CIMMYT, 2009); whereas Denic (2005) characterised the LP inbred lines, including LP23 and LP21, and reported that they are highly resistant to DM. Most importantly, the results indicate that LP19 and LP21 are in the same sub-cluster while LP23, which originates in Nigeria, is in a different sub-cluster, which suggests that a three-way cross hybrid among these lines (LP19/LP21/LP23) should be productive. This analysis has been confirmed by the maize breeders in Mozambique from practical experience with these lines ([1]Fato, 2011; pers. comm.).

[1]Dr Pedro Fato, Chokwe Research Station, Institute of Agriculture of Mozambique (IIAM).

2.4.3 Implications for developing new inbred lines

Additionally, classification of MSV resistance donor lines in two different clusters MC1 (P13 through to P18) and MC2 (CML505 and CML509) indicates that they can be regarded as different heterotic groups. This has implications for managing the breeding programme in Mozambique. In order to fix inbred lines over shorter periods, it is crucial to develop breeding populations by crossing lines with similar allele frequencies. Lines CML505 and CML509 are the potential donors for MSV resistance to the Mozambican lines in the cluster 2 (E66, E7, E46, E47 and E24). On the other hand, the MSV potential donor lines P13, P14, P15, P16, P17 and P18 could be used to improve MSV resistance in the other Mozambican lines in the cluster 1 (E80, E27, E71, LP19, LP37D, LP37F and LP23) because they are likely to have a similar gene frequency. Unfortunately there are no MSV resistance donor lines which were fitted in cluster 3; hence another set of potential donor lines for use to improve resistance of these lines (E72, E75 and E77) to MSV should be found.

2.5 Conclusions

The following conclusions were drawn:

1) There is genetic variation among the 25 inbred lines as they were assigned to three major clusters, and in general, the genotypic analysis confirmed the pedigree data. This indicates that breeding new MSV resistant hybrids from this set of germplasm would be viable.
2) Existence of genetic distance between the lines indicates that productive hybrids can be designed such that lines from different groups can be crossed so that they can complement each other in hybrids.
3) The study also identified the potential MSV resistance donors for use to improve the lines in Mozambique, which also forms the basis for devising a breeding strategy which is appropriate for the breeding programme in Mozambique.

References

Aguiar, C.G., I. Schuster, A.T. Amaral, Juniour, C.A. Scapim, and E.S.N. Viera. 2008. Heterotic groups in tropical maize germplasm by testcrosses and simple sequence repeat markers. Genetics and Molecular Research 4:1233-1244.

Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1132-1139.

Asea, G., B.S. Vivek, G. Bigirwa, P.E. Lipps, and R.C. Pratt. 2008. Validation of consensus quantitative trait loci associated with resistance to multiple foliar pathogens of maize. Phytopathology 99:540-547.

Babu, B.K., P. Pooja, J.C. Bhatt, and P.K. Agrawal. 2012. Characterisation of Indian and exotic quality protein maize (QPM) and normal maize (*Zea mays* L.) inbreds using simple sequence repeat (SSR) markers. African Journal of Biotechnology 11:9691-9700.

Biswas, M.K., M.A.A. Mondal, M. Hossain, and R. Islam. 2008. Utilization of genetic diversity and its association with heterosis for progeny selection in potato breeding programs. Science 6:882-887.

Bosque-Perez, N.A. 2000. Eight decades of maize streak virus research. Virus Research 71:107-121.

Brown, S.M., M.S. Hopkins, S.E. Mitchell, M.L Senior, T.Y. Wang, R.R. Duncan, F. Gonzalez-Candelas, and S. Kresovich. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* L. Moench). Theoretical and Applied Genetics 93:190-198.

Cholastova, T., M. Soldanova, and R. Pokorny. 2011. Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) marker efficacy for maize hybrid identification. African Journal of Biotechnology 10:4794-4801.

CIMMYT. 2002. CIMMYT: Data analysis in the CIMMYT applied biotechnology centre: for fingerprinting and genetic diversity studies. http://www.cimmyt.cgiar.org/ABC/Protocols/manualABC.html (Accessed 9 Nov. 2012).

CIMMYT. 2009. CIMMYT: International Maize and Wheat Improvement Center. CIMMYT Maize Inbred Lines (CML). http://www.cimmyt.org/en/component/content/article/459-international-maize-

improvement-network-imin/434-cimmyt-maize-inbred-lines-cml (Accessed 9 Aug. 2011).

Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer, and E.C.K. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169-196.

Danson, J., M. Lagat, J. Ininda, and M. Kimani. 2006. Application of simple sequence repeats (SSRs) markers to study the resistance of locally adapted maize hybrids to damaging maize streak virus disease. African Journal of Biotechnology 5:1430-1434.

Denic, M. 2005. Maize breeding program: Breeding activities and national trials. National Agriculture Research Institute Annual Progress Report. Department of Agriculture and Farming Systems. Maputo, Mozambique.

Denic, M.P., C. Chauque, M. Jose, D. Langa, D. Mariote, P. Fato, and W. Haag. 2001. Maize screening for multiple stress tolerance and agronomic traits. Proceedings of the Eastern and Southern Africa Regional Maize Conference 7:88-91.

Dias, L.A.S., J.M. Cruz, C.D. Barros, and T.M.F. Salamao. 2003. Genetic distance and its association with heterosis in cacao. Brazilian Archives of Biology and Technology 46:339-347.

Dice, L.R. 1945. Measures of the amount of ecologic association between species. Ecology 26:297-302.

Diniz, L.E.C., C.F. Ruas, V.P. Carvalho, F.M. Torres, E.A. Ruas, M.O. Santos, T. Sera, and P.M. Ruas. 2005. Genetic diversity among forty coffee varieties assessed by RAPD markers associated with restriction digestion. Brazilian Archives of Biology and Technology 48:511-521.

Duvick, D.N. 1999. Heterosis: Feeding people and protecting natural resources. pp. 19-29. *In* J.G. Coors and S. Pandey (eds.) The Genetics and Exploitation of Heterosis in Crops. An International Symposium. CIMMYT, Mexico City, Mexico.

Enoki, H., H. Sato, and K. Koinuma. 2002. SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan. Theoretical and Applied Genetics 104:1270-1278.

Fato, P. 2010. Investigation of heterotic patterns and genetic analysis of downy mildew resistance in Mozambican lowland maize (*Zea mays* L.) germplasm. African Crop Science Society Conference, Maputo, Mozambique. http://www.acss2011moz.uem.mz/doc/announcements/Volume%20of%20Abstracts%20&%20Programme.pdf (Accessed 16 Jul. 2011).

Hallauer, A.R., and J.B. Miranda. 1988. Quantitative Genetics in Maize Breeding. Iowa State University Press. Ames, USA.

He, G., R. Meng, M. Newman, G. Gao, R.N. Pittman, and C. Prakash. 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biology 3:1-6.

Hoxha, S., M.R. Shariflou, and P. Sharp. 2004. Evaluation of genetic diversity in Albanian maize using SRR markers. Maydica 49:97-103.

Ijaz, S., and Khan I.A. 2009.Molecular characterisation of wheat germplasm using microsatellite markers. Genetics and Molecular Research 8:809-815.

Karanja, J., N.O. Amugune, J. Ininda, J.N. Kimatu, and J.W. Danson. 2009. Microsatellite analysis of the correlation between molecular and morphological traits in assorted maize inbred lines. African Crop Science Journal 17:133-144.

Kyetere, D.T., R. Ming, M.D. McMullen, R.C. Pratt, J. Brewbaker, and T. Musket. 1999. Genetic analysis of tolerance to maize streak virus in maize. Genome 42:20-26.

Laborda, P.R., K.M. Oliveira, A.A.F. Garcia, M.E.A.G.Z. Paterniani, and A.P. de Souza. 2005. Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers? Theoretical and Applied Genetics 111:1288-1299.

Lagat, M., J. Danson, M. Kimani, and A. Kuria. 2008. Quantitative loci for resistance to maize streak virus disease in maize genotypes used in hybrid development. African Journal of Biotechnology 7:2573-2577.

Li, Y., J. Du, T. Wang, Y. Shi, Y. Song, and J. Jia. 2002. Genetic diversity and relationships among Chinese maize inbred lines revealed by SSR markers. Maydica 47:93-101.

Liu, K., and S. Muse. 2005. PowerMarker: an integrated analysis environment for genetic marker analysis. Genetics and Population Analysis 21:2128-2129.

Liu, K., M. Goodman, S. Muse, J.S. Smith, E. Buckler, and J. Doebley. 2003. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. Genetics 165:2117-2128.

Magenya, O.E.V., J. Mueke, and C. Omwega. 2008. Significance and transmission of maize streak virus disease in Africa and options for management. African Journal of Biotechnology 7:4897-4910.

Maize genetics and genomics database. http://www.maizegdb.org/ (Accessed 21 Nov. 2011).

Martin, D.P., and D.N. Shepherd. 2009. The epidemiology, economic impact and control of maize streak disease. Food Security 1:305-315.

Masi, P., P.L. Spagnoletti-Zeuli, and P. Donini. 2003. Development and analysis of multiplex microsatellite markers sets in common bean (*Phaseolus vulgaris* L.) Molecular Breeding 11:303-313.

Mawere, S., V. Vincent, J. De Meyer, and K.V. Pixley. 2006. Resistance of four inbred maize lines to inoculation with 20 isolates of maize streak virus from Zimbabwe. Plant Disease 90:1485-1489.

Melchinger, A.E. 1999. Genetic diversity and heterosis. pp. 99-118. *In* J.G. Coors and S. Pandey (ed.) The Genetics and Exploitation of Heterosis in Crops. American Society of Agronomy/Crop Science Society of America Inc., Madison, USA.

Meyer, R.C., T. Otto, M. Becher, and T. Altmann. 2004. Heterosis of biomass production in Arabidopsis establishment during early development. Plant Physiology 134:1813-1823.

Molnar, S., J.S. Rai, M. Charette, and E.R. Cober. 2003. Simple sequence repeat (SSR) markers linked to E1, E3, E4 and E7 maturity genes in soybean. Genome 46:1024-1036.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.

Oluwafemi, S., G. Thottappilly, and M.D. Alegbejo. 2008. Search for variability among mastreviruses showing symptoms of maize streak virus from different regions of Nigeria by polymerase chain reactions. Journal of Plant Protection Research 48:1-6.

Pernet, A., D. Hoisington, J. Franco, M. Isnard, D. Jewell, C. Jiang, J.L. Marchand, B. Reynaud, J.C. Claszmann, and D. Gonzalez de Leon. 1999. Genetic mapping of maize streak virus resistance from the mascarene source. Resistance in line D211 and stability against different virus clones. Theoretical and Applied Genetics 99:524-539.

Redinbaugh, M.G., M.W. Jones, and R.E. Gingery. 2004. The Genetics of virus resistance in maize (*Zea mays* L.). Maydica 49:183-190.

Rohlf, F. J. 1998. NTSYSpc numerical taxonomy and multivariate analysis system version 2.0. Setauket, NY, USA: Exeter Software. pp. 31.

Salem, K.F.M., A.M. El-Zanaty, and R.M. Esmail. 2008. Assessing wheat (*Triticum aestivum* L.) genetic diversity using morphological characters and microsatellite markers. World Journal of Agricultural Sciences 5:538-544.

Shahnejat-Bushehri, A.A., S. Torabi, M. Omidi, and M.R. Ghannadha. 2005. Comparison of genetic and morphological distance with heterosis with RAPD markers in hybrids of Barley. International Journal of Agriculture and Biology 7:592-595.

Sharopova, N., M.D. McMullen, L. Schultz, S. Schroeder, H. Sanchez-Villeda, J. Gardiner, D. Bergstrom, K. Houchins, S. Melia-Hancock, and T. Musket T. 2002. Development and mapping of SSR markers for maize. Plant Molecular Biology 48:463-81.

Shepherd, D.N., T. Mangwende, D.P. Martin, M. Bezuidenhout, J.A. Thomson, and E.P. Rybicki. 2007. Inhibition of maize streak virus (MSV) replication by transient and transgenic expression of MSV replication-associated protein mutants. Journal of General Virology 88:325-336.

Shull, G.H. 1952. Beginnings of the heterosis concept. pp.14-48. *In* Gowen, J.W. (ed). Heterosis. Iowa State Univ. Press, Ames, IA.

Smith, J.S.C., and O.S. Smith. 1989. Comparison of heterosis among hybrids as a measure of relatedness with that to be expected on the basis of pedigree. Maize Genetics Cooperative Newsletter 63:86-87.

Smith, J.S.C., E.C.L. Chin, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovich, and J. Ziegle. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. Theoretical and Applied Genetics 95:163-173.

Stevens, R. 2008. Prospects for using marker-assisted breeding to improve maize production in Africa. Journal of the Science of Food and Agriculture 88:745-755.

Taramino, G., and S. Tingey. 1996. SSR for germplasm analysis and mapping in maize. Genome 39:277-287.

Vigouroux, Y., S. Mitchell, Y. Matsuoka, M. Hamblin, S. Kresovidi, J.S.C. Smith, J. Jaqueth, O.S. Smith, and J. Doebley. 2005. An analysis of genetic diversity across the maize genome using microsatellites. Genetics 169:1617-1630.

Virmani, S.S., Z.X. Sun, T.M. Mou, A.A. Jauhar, and C.X. Mao. 2003. Two-line Hybrid Rice Breeding Manual. International Rice Research Institute, Los Baños, the Philippines. pp. 88.

Warburton, M.L., X. Xianchun, S. Ambriz, L. Diaz, S.K. Vasal, and F. Gonzalez. 2002. Genetic relationships among the CIMMYT maize lines developed for Asia and from other CIMMYT breeding programs. Proceedings of the 8th Asian Regional Maize Workshop, Bangkok, Thailand. http://apps.cimmyt.org/english/docs/proceedings/armw/armw.pdf (Accessed 9 Mar. 2011).

Weising, K., H. Nybom, K. Wolff, and G. Kahl. 2005. DNA Fingerprinting in Plants: Principles, Methods and Applications. Taylor and Francis/CRC Press, London, UK. pp. 444.

Welz, H.G., A. Schechert, A. Pernet, K.V. Pixley, and H.H. Geiger. 1998. A gene for resistance to the maize streak virus in the African CIMMYT maize inbred line CML202. Molecular Breeding 4:147-154.

Xia-Su, X., J. Liu, and G. Liu. 2004. The use of SSRs for predicting the hybrid yield and yield heterosis in 115 key inbred lines of Chinese maize. Hereditas 141:207-215.

Xia, X.C., J.C. Reif, A.E. Melchinger, M. Frisch, D.A. Hoisington, D. Beck, K. Pixley, and M.L. Warburton. 2005. Genetic diversity among CIMMYT maize inbred lines investigated with SSR markers. Crop Science 45:2573-2582.

Yao, Q., K. Yang, G. Pan, and T. Rong. 2007. Genetic diversity of maize (*Zea mays* L.) landraces from Southwest China based on SSR data. Journal of Genetics and Genomics 34:851-860.

Yu, K., S.J. Park, V. Poysa, and P. Gepts. 2000. Markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). Journal of Heredity 91:429-434.

Yuan, L., S. Zhang, M.L. Warburton, X. Li, J. Fu, and M. Li. 2002. Assessment of genetic similarities among maize inbred lines using SSR markers. Proceedings of the Asian Regional Maize Workshop 8:50-54.

CHAPTER THREE

Detection of SSR markers linked to MSV disease resistance and high resolution melt (HRM) analysis of F_3 maize population samples stored on Whatman FTA™ elute cards

Abstract

Diagnosis and characterisation of maize populations resistant to the maize streak virus is essential for selection and improvement of maize streak virus (MSV) resistant cultivars. The objectives of the current investigation were to establish the reliability of FTA™ cards for the sampling of genotypes and develop a polymerase chain reaction-high resolution melt (PCR-HRM) analysis assay, utilising simple sequence repeat (SSR) molecular markers to detect MSV disease resistance in maize lines. This entailed the transfer of the MSV resistance gene cluster from CML505 and CML509 (MSV resistant donor lines) into a Mozambican adapted line LP23 (MSV susceptible). A total of 118 F_3 family lines derived from two F_2 populations (CML505 x LP23 and CML509 x LP23) were genotyped using SSR markers and HRM analysis. The melting profiles were characterised by one peak at a melting temperature (T_m) of 82.47°C for CML505 and two peaks at T_m 79.97°C and 82.20°C for LP23, with the SSR marker bnlg1811. The parental lines CML509 and LP23 assayed with the SSR marker umc2228 were characterised by single peaks at T_m 83.75°C and 85.06°C, respectively. Twenty-nine maize lines were classified as MSV disease resistant. Simple sequence repeats and HRM analyses of maize genotypes successfully differentiated between the parental lines and detected the *msv1* gene, which is responsible for conferring resistance to the maize streak virus disease. This investigation established DNA sampling using FTA™ cards followed by post-PCR analysis using HRM curve analysis as a feasible approach for the rapid screening of large numbers of maize DNA samples and for rapid *msv1* gene detection.

Keywords: Maize streak virus, resistance, FTA™ elute cards, high-resolution melt (HRM) analysis, simple sequence repeat (SSR) markers.

3.1 Introduction

Maize streak virus disease can cause devastating yield losses of up to 100% when the virus infects susceptible plant lines (Sharma and Misra, 2011; Tefera *et al*., 2011). Such negative impact of the disease makes it one of the major biotic constraints on maize production, with a significant biological threat to food security in Sub-Saharan Africa (Thottappilly *et al*., 1993; Martin and Shepherd, 2009). This undermines the well-being of small-scale farmers throughout Africa (Bosque-Perez, 2000; Owor *et al*., 2007; Shepherd *et al*., 2007; Martin and Shepherd, 2009). Breeding for MSV disease resistant maize lines has been suggested to be the most cost-effective measure for reducing yield losses due to the disease (Saxena and Hooker, 1968; Wisser *et al*., 2006).

Plant breeding relies on genetic variation and makes use of selection to improve characteristic traits that are of consumer interest (Asins, 2002). Characters like quality, yield and disease resistance reaction that show continuous variation are difficult to detect and transfer through conventional plant breeding (Asins, 2002; Farooq and Azam, 2002). These characters are controlled by multiple loci known as quantitative trait loci (QTL) and environmental factors, and thus cannot be measured by individually recognisable loci under normal conditions of measurement (Farooq and Azam, 2002). However, considerable progress in predicting response to selection has been made since 1980, before which the genetics of such traits were studied using statistical techniques (e.g. means, variances and heritabilities), rather than in terms of individual gene effects (Luo, 1998). The location of loci affecting quantitative traits is detected by the joint analysis of segregation of marker genotypes and phenotypic values of individual lines (Luo, 1998). Marker-assisted selection (MAS) is based on genetic information retrieved through the application of molecular markers. It can be employed to speed up the development of cultivars with enhanced traits and also to unveil masked beneficial alleles (Asins, 2002).

The process of sample collection and processing by conventional means can be laborious (Ndunguru *et al*., 2005) as collected plant tissue needs to be preserved under appropriate conditions in order to maintain the integrity of the nucleic acids to be analysed (Cortes *et al*., 2009). Furthermore, failure to obtain reliable high quality nucleic acids from infected plant material in the field is usually the result of late processing and/or storage of samples before

they spoil, especially when there are large sample numbers and when working in fields distant to laboratory facilities. Studies are, therefore, constrained by resources required for preservation and transportation of collected samples in order to maintain their integrity (Ndunguru *et al*., 2005).

Flinders Technology Associates have developed a card based technology called FTA^{TM} elute that provides a simple and robust sample collection and storage method, which is suitable for molecular epidemiologic studies in remote areas with tropical climates, where transport and storage conditions are difficult (Moscoso *et al*., 2005; Sultan *et al*., 2009). The laborious isolation and purification steps that come with conventionally used methods for plant DNA extraction can be avoided with the use of FTA^{TM} elute cards that provide an extraction-free means of preparing DNA templates (Orlandi and Lampel, 2000). The paper-based FTA^{TM} card technology is designed to reduce the steps of DNA collection, transportation, purification and storage. The technology thus makes the process of purifying DNA for downstream applications more cost effective and less time-consuming (Orlandi and Lampel, 2000; Ndunguru *et al*., 2005).

Due to the increased demand for rapid genetic analysis and high-throughput diagnosis of pathogens, there has been a growing focus on evaluation methods that rely on the melting characteristics of amplicons (Reed *et al*., 2007). This applies particularly to PCR-based melting-curve and high resolution melting-curve analyses, which substantially decrease the time required for testing of samples (Pangasa *et al*., 2009). Introduced in 2002 (Reed *et al*., 2007), HRM analysis is an automated analytical molecular technique that measures dissociation of double stranded DNA to single stranded DNA with increasing temperature of a PCR product amplified in the presence of a saturating fluorescent dye (Studer *et al*., 2009; Talmi-Frank *et al*., 2010). Amplicons containing different sequences can be discriminated based on the melting transition of the PCR product and the resulting melt curve shape (Chateigner-Boutin and Small, 2007; Steer *et al*., 2008; Stoep *et al*., 2009), thus HRM analysis can be used for genotyping and species determination (Lin *et al*., 2008). High resolution melt analysis can be performed in a Rotor-Gene 6000 real-time rotary analyser, which performs both PCR amplification and HRM analysis in the same tube (Corbett Research, 2006). High resolution melt analysis is a simple method for determining sequence variation and can achieve high mutation detection rates for small (usually 100-250 bp) amplicons (Reed *et al*., 2007; Pangasa *et al*., 2009; Hondow *et al*., 2011).

A screening method to detect SSR polymorphism of the F_2 generation in different maize inbred lines using HRM curve analysis was developed by Yu *et al*. (2011). The study concluded that the HRM-SSR system can substitute gel electrophoresis for analysis of PCR products. The HRM technique was also applied for mapping of single nucleotide polymorphism (SNP) markers linked to a covered smut resistance gene in barley (Lehmensiek *et al*., 2008). This was the first report on the application of HRM for SNP detection and rapid scoring of known cleaved amplified polymorphic sequence (CAPS) markers in plants. A study by Hofinger *et al*. (2009) demonstrated that the HRM analysis of cDNA-derived PCR amplicons is a rapid, simple and cost-effective method for identification of novel *eIF4E* gene alleles in barley. Gady *et al*. (2009) adapted two techniques used in human genetic diagnostics, Conformation Sensitive Capillary Electrophoresis (CSCE) and HRM, for the characterisation of a large tomato ethyl methanesulfonate (EMS) mutated population. The results demonstrated that CSCE and HRM are fast, affordable and sensitive techniques for mutation detection in DNA pools and therefore allow the rapid identification of new allelic variants in a mutant population.

In the present study, the main objective was to develop a PCR-HRM protocol for the screening of the *msv1* gene to assign genotypes of 118 maize samples. This was enabled by the ability of HRM to simultaneously detect and genotype DNA polymorphisms (Montgomery *et al*., 2007). The focus was to establish whether the *msv1* resistance gene had been introgressed from the MSV disease resistant CML505 and CML509 lines into the MSV disease susceptible LP23 Mozambican maize line. High resolution melt analysis, which is less time consuming compared to electrophoresis analysis (Steer *et al*., 2008), was used for microsatellite marker post-PCR analysis. It has the added advantage of data storage and analysis capabilities via computer assimilation. Another objective was to establish the reliability of FTA^{TM} cards for sampling and retrieval of DNA samples collected from maize plants.

3.2 Materials and methods

3.2.1 Germplasm and generations

Three inbred parental maize lines were used in this study. Resistance to MSV disease has been identified in inbred lines CML505 and CML509 from the International Maize and

Wheat Improvement Center (CIMMYT) (CIMMYT, 2009). These lines were used as the MSV donor parental lines. LP23 is an inbred line which was derived from TZ14 at the International Institute of Tropical Agriculture (IITA) in Nigeria and is considered to be a reliable source of downy mildew (DM) resistance, which is also an important disease of maize in Mozambique (Denic, 2005; Fato, 2010). The LP23 line is used as a potential hybrid parent in Mozambique. However, LP23 is susceptible to the MSV disease, thus requires resistance to the MSV disease to be incorporated. The initial F_1 hybrids between CML505 and LP23 and between CML509 and LP23 were self-pollinated to develop two segregating F_2 populations. Superior single plants interms of agronomic traits like plant height and lodging were further advanced to the F_3 generation at Makhathini Research Station (27° 38´ 15; 32° 10´ E) in the KwaZulu-Natal (KZN) Province of South Africa resulting in 118 $F_{2:3}$ families (Table 3.1).

Table 3.1: Pedigree names and the 118 $F_{2:3}$ family lines used in the study.

Name	Pedigree	Origin	No. of plants	F_2 : F_3 Ears
GCPMOZ7	CML505/LP23-F_2B	09MAK28-7	49	28
GCPMOZ12	CML509/LP23-F_2B	09MAK28-12	188	90

Previous screen house and field evaluations carried out at CIMMYT, Zimbabwe, identified the maize donor lines CML505 and CML509 to have MSV resistance with disease ratings 1.5 and 2, respectively, on a scale of 1-5 (CIMMYT, 2009), with a rating of 1 being highly resistant and 5 being highly susceptible. The lines were characterised as having molecular markers associated with the resistance property at bin 1.04 of chromosome 1 (CIMMYT, 2009). The donor lines are also known to have high levels of resistance to DM, thus were the preferred donors for breeding for MSV resistance in Mozambican lines due to the prevalence of DM in coastal Mozambique (Denic, 2005; Fato, 2010). The common parent LP23 is the candidate Mozambican line for improvement because it is adapted to the lowland environment (≤ 500 m altitude) and is the potential parent of productive hybrids. However, hybrid vigour in these lines is compromised by the susceptibility of the LP23 line to the MSV disease.

3.2.2 Tissue sampling and DNA collection

The F_3 seed harvested from Makhathini Research Station were planted out at the University of KwaZulu-Natal tunnels in Pietermaritzburg. The 118 $F_{2:3}$ progeny lines were planted out together with parental lines (CML505, CML509 and LP23). Three seeds were planted per pot and four pots per line, and were thinned to one plant per pot. DNA was isolated from the single plants from each of the 121 lines. Equal amounts of DNA of each plant were then bulked for each line during analysis, because the breeding aimed at emphasising selection between the families. Selection of the best individuals within a family was delayed to the F_5 under field conditions so that individuals which combine MSV resistance with good agronomic traits would be advanced in the breeding programme. The data represented in Tables 3.3 and 3.4 is that of bulked DNA.

FTA^{TM} elute cards (Merck, Johannesburg, South Africa) were used for the collection of DNA. Leaf samples were collected at three weeks after emergence. A pair of pliers with a rounded end was used to press the leaf sample onto the FTA^{TM} elute paper until both sides of the FTA^{TM} were soaked with leaf sap. Gloves were worn and the pliers' end was wrapped with parafilm and wiped with 70% ethanol after each sampling to prevent sample to sample contamination. The FTA^{TM} elute cards were left to dry at room temperature for 2-5 hours and stored at room temperature until required. To establish the reliability of FTA^{TM} cards for the sampling of genotypes, a comparison test of DNA extracted from FTA^{TM} elute cards with crudely extracted DNA was set-up.

3.2.3 Comparison of DNA extracted from FTA^{TM} elute cards with crudely extracted DNA

Three comparison tests were carried out with FTA^{TM} eluted DNA and crude DNA in order to determine whether FTA^{TM} eluted DNA could generate distinct melting curves compared to those generated by DNA eluted directly from the plant without the means of a paper matrix. The DNA was purified from FTA^{TM} cards in three steps. Firstly, a 2 mm disk was punched out of the FTA^{TM} matrix impregnated with plant material. The disk was placed into a 1.5 ml

microcentrifuge tube after which it is washed with 500 µl of sterile H_2O by pulse vortexing three times for a total of 5 s. The disc was then transferred to a new 0.5 ml microcentrifuge tube containing 30 µl sterile H_2O ensuring that the disc was completely submerged. The tube was transferred to a 95°C heat block for 30 min. The sample was removed from the block and pulse vortexed for at least 5 s. The sample was centrifuged for 30 s, to separate the matrix from the eluent. The FTA™ elute matrix disc was removed from the eluent that contained the purified DNA, using a sterile pipette tip and discarded. The eluted DNA was stored at -20°C until required for PCR amplification.

Crude genomic DNA was extracted using a modified Edwards *et al*. (1991) protocol. Leaf tissue was collected using the lid of a sterile microcentrifuge tube to punch out a disc of material into the tube. The extraction buffer (400 µl) [(200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% sodium dodecyl sulphate (SDS)] was then added to the punch, and macerated, after which the sample was vortexed for five seconds. Samples were then heated at 65°C for 10 minutes in a heating block. The extracts were centrifuged at 13 000 rpm for 2 min and 350 µl of the supernatant was transferred to a clean microcentrifuge tube. The supernatant was mixed gently with 350 µl cold isopropanol and left at -20°C for 30 min. Following centrifugation at 13 000 rpm for 5 min, the pellet was vacuum dried and dissolved in 20 µl 1 X Tris-EDTA (TE) buffer. DNA was stored at 4°C.

The melt profiles of the parental lines generated from DNA eluted from FTA™ cards were then compared to the conventionally extracted DNA to determine the efficiency and reliability of FTA™ cards for PCR and HRM analysis.

3.2.4 DNA samples

The unknown 118 F_3 progeny were genotyped for the *msv1* gene by using the known reference genotypes; parental lines CML505 and CML509 which were used to diagnose the presence of the *msv1* gene, and LP23 in which the gene was absent. The aim of this study was to develop a high throughput HRM protocol in order to detect the homozygous *msv1* gene, therefore a heterozygous reference was not included though both homozygous and heterozygous sequence variants can be reliably differentiated using HRM analysis (Gundry *et al*., 2003; Wittwer *et al*., 2003). A negative control that had all the reagents excluding the

DNA was included in every PCR run to ensure that the PCR reagents were free of contamination.

3.2.5 Simple sequence repeats (SSR) analysis

For amplification reactions, 10 SSR markers were used (Table 3.2). Some of these markers were included in the screening of the possible parental lines but most were taken from Danson *et al.* (2006) and the maize database (www.maizegdb.org). These were markers previously proven to provide greater complementarity and reproducibility, as well as the presence of polymorphism for the MSV gene. The primers for these SSR markers were manufactured by Integrated DNA Technologies (IDT) (Whitehead Scientific, Cape Town, South Africa). The generated melt profiles showed that two of the ten markers used, primers bnlg1811 and umc2228 gave the most notable differences in melt profiles of the parental lines thus enabling the identification of MSV marker disease resistant progeny. The use of the rest of the primers was discontinued as some failed to amplify the *msv1* region for these particular lines resulting in indistinguishable or no peaks.

Table 3.2: List of potential 10 SSR markers for *msv1* introgression (Chr 1.04-0.5) [Danson *et al*., 2006; Maize database (www.maizegdb.org)].

Name	Forward primer (5´- 3´)	Reverse primer (5´- 3´)
bnlg1811	GTAGTAGGAACGGGCGATGA	ACACAAGCCGACCAAAAAAC
bnlg2086	CGGAACCTGCTGCAGTTAAT	GAGATGCAGGAATGGGAAAA
umc2390	GAAATGGCAGGGAAACTTGTTTAT	AAGAGGCAAGCAAGTGTACAGTGA
umc1676	AGTCGTACGATGACGGAGGC	GCACCACCGACTGATCAAGA
umc1243	AACTGCAGAGTCGCCTGATCC	AAGCAGACTATGCTATGCTACGCC
umc1144	ATGGCCCACTCATCATATCTCTGT	TGTGTTGATTAGCAGCGGATAAAA
bnlg1884	TTCGGATGCATGTGTAACGT	CGGAAGTCCCATCTGTTTGT
bnlg2295	CGGAGGAGTGGTTCTTGAAA	GGTTAGTGAAAGGGTTGCCA
umc2228	ACCATACCTCTCTGAACATGAGCC	GTGAGGTGAAAATGAAGCTGGAAC
umc1917	ACTTCCACTTCACCAGCCTTTTC	GGAAAGAAGAGCCGCTTGGT

3.2.6 PCR and HRM conditions

PCR amplifications were performed in 20 µl reaction volumes consisting of approximately 20 ng of genomic DNA template, 10 µl of 2 x Quantace SensiMix for the PCR reaction components and 200 nm of forward and reverse primers. A negative control (PCR mix without DNA template) was included in each set of PCR reactions to ensure non-contamination of PCR reagents. The primers were optimised in terms of the number of FTA^{TM} discs per elute to be used for each PCR reaction and range of melting temperatures in HRM analysis. The cycling profile was as follows: hold at 95°C for 10 min; 40 cycles of 95°C for 5 s, 60°C for 10 s, 72°C for 10 s and finally hold at 72°C for 3 min. A pre-melt hold at 90°C for 30 s was allowed prior to HRM analysis of 81-88°C with a 1°C increment per step. The melt profile generated for each amplicon following real time PCR was analysed using Rotor-Gene 6000 Series Software (Version 1.7.87) (Corbett Research, Sydney, Australia).

3.3 Results

3.3.1 Comparison of DNA extraction by FTA^{TM} cards and that extracted by conventional methods

The melting profiles generated with the parental line LP23 using the SSR primer umc1917 for DNA extracted from FTA^{TM} elute cards and DNA extracted by the Edwards *et al*. (1991) protocol are shown in Fig. 3.1. The melting transitions generated appeared as peaks on a derivative plot, which displays the negative derivative (-dF/dT) of the normalised fluorescence with respect to temperature against temperature (Montgomery *et al*., 2007). The generated melting curves were similar in shape and differed by 0.08°C, with T_m for the crude DNA being 86.80°C and that for FTA^{TM} DNA being 86.88°C.

3.3.2 Comparison of HRM melt profiles generated when sufficient and insufficient sap is pressed onto the FTA cards.

The generated melting curves for the PCR amplicons using DNA eluted from FTA^{TM} cards when insufficient sap was pressed onto the FTA^{TM} cards were of similar shape but lower peak amplitude to that of the curves using DNA eluted from FTA^{TM} cards with sufficient sap. The melting temperature for the major peak of CML509 with sufficient and insufficient sap

was the same at 79.53°C. CML509 with insufficient sap had a very low dF/dT value of less than 0.05 (Fig. 3.2).

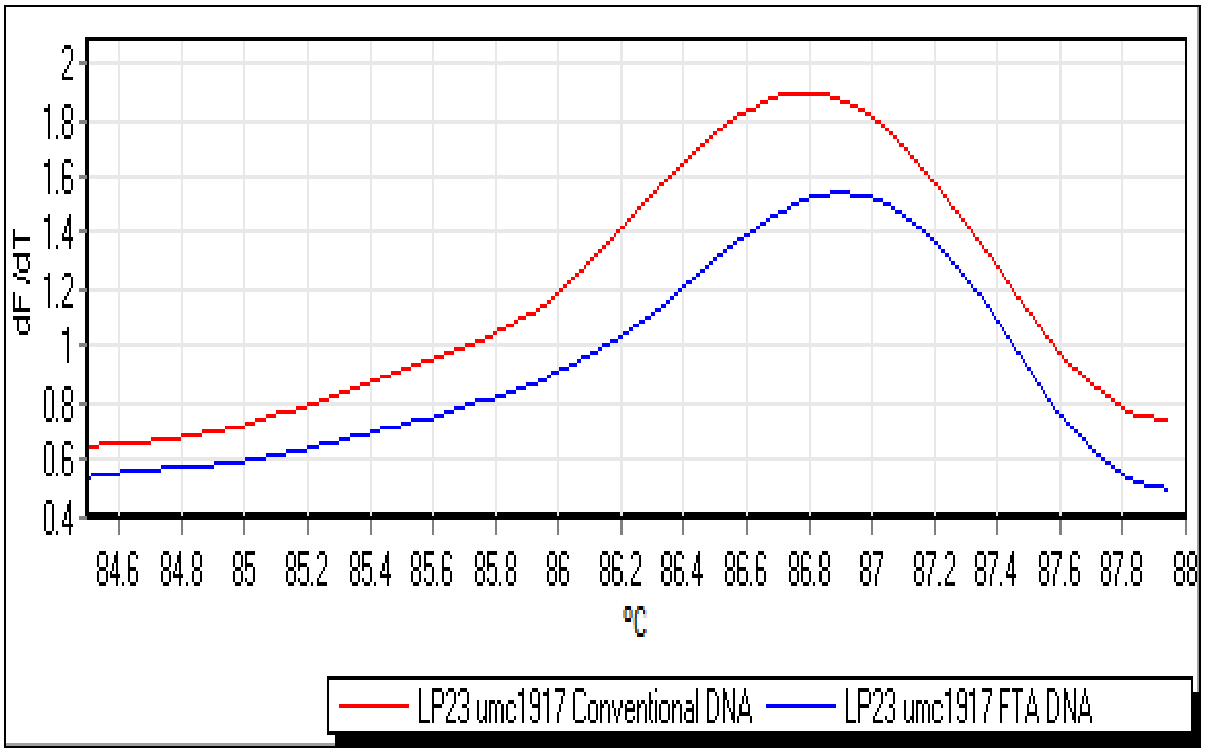

Figure 3.1: Melt profiles of genomic DNA extracted from FTA™ elute cards compared with crude DNA extracted from the parental line LP23 using SSR primer umc1917.

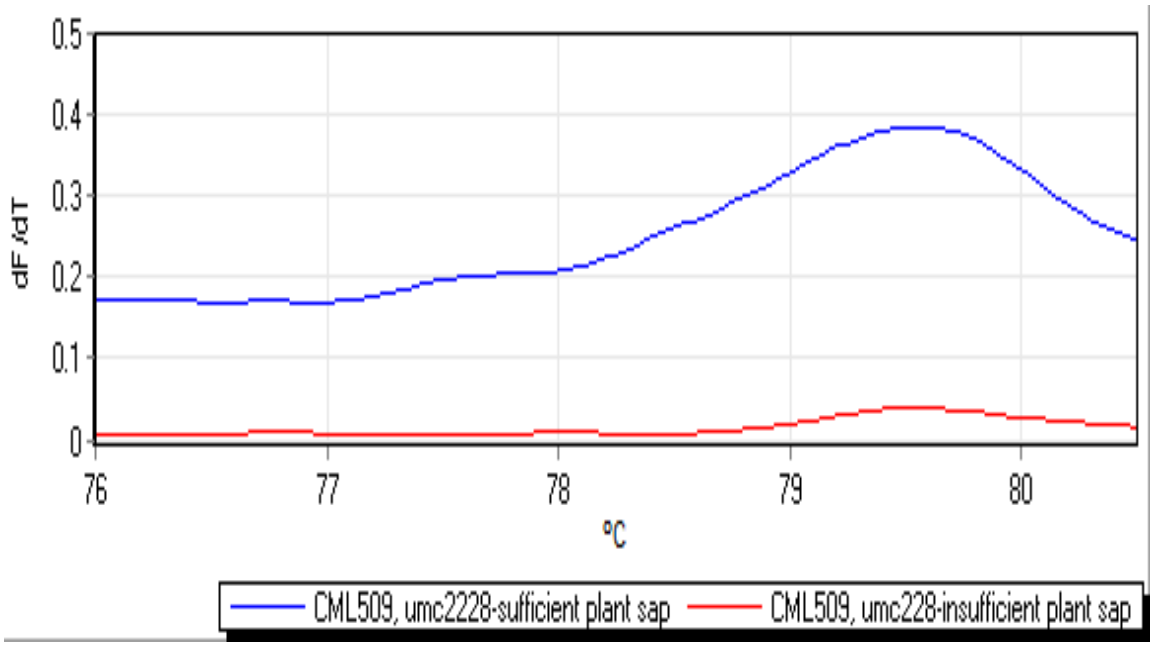

Figure 3.2: Melt profiles of genomic DNA extracted from FTA™ elute cards with sufficient and insufficient plant sap pressed onto the FTA™ cards from the parental line CML509 amplified using SSR primer umc2228.

3.3.3 HRM curve profiles for the screening for MSV resistance

The PCR products of the parental lines, CML505 and LP23 amplified by primer bnlg1811 were subjected to HRM analysis. The differences between the parental lines were visualised and quantified using melting transitions on a derivative plot. The melt profile for parental line CML505 exhibited a single peak at a T_m of 82.47°C while parental line LP23 showed two peaks at 79.97°C and 82.20°C, respectively (Fig. 3.3). The PCR amplicon melting profiles of the two parental controls were distinctively different and the genotypes were assigned MSV-resistant for CML505 and MSV-susceptible for LP23. The Rotor-Gene software was used for progeny genotype identification.

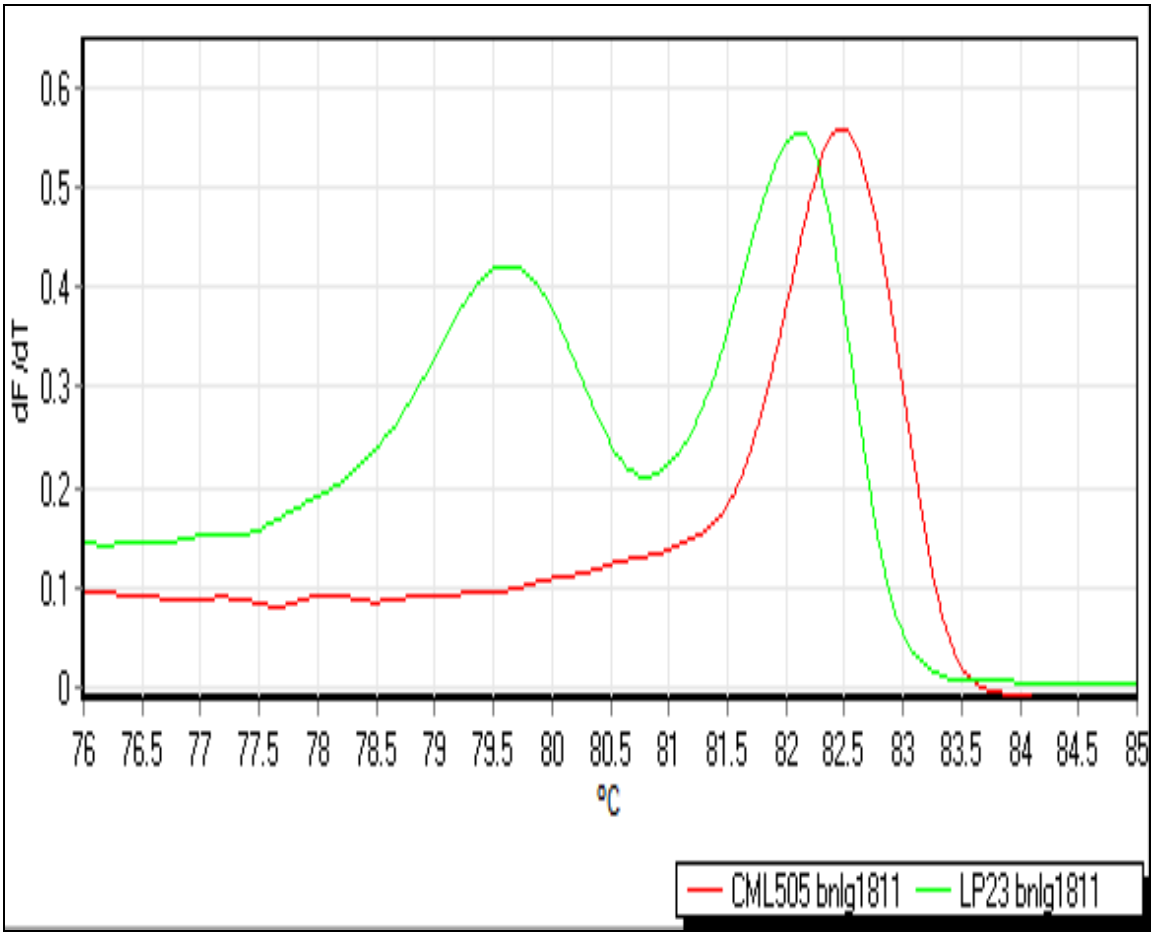

Figure 3.3: Melt profiles of parental lines CML505 (RED), MSV disease resistant and LP23 (GREEN), MSV disease susceptible with SSR marker bnlg1811.

The progeny lines from the cross CML505 x LP23 were characterised using difference curves. A difference plot is able to distinguish between the parental inbred lines as one of the parental/reference sample melting curves is normalised to the baseline. Fig. 3.4 shows a difference plot in which CML505 (RED), the MSV disease resistant parent is normalised to the baseline against which LP23 (GREEN) the MSV disease susceptible parent and CML505/LP23 progeny lines are plotted. The progeny lines CML505/LP23-1, CML505/LP23-11, CML505/LP23-15, CML505/LP23-18 and CML505/LP23-25 aligned themselves along the red baseline of MSV disease resistant parent, CML505 (Fig. 3.4). In total the marker bnlg1811 detected 16 samples out of the 28 of the CML505/ LP23 progeny that were MSV disease resistant (Table 3.3).

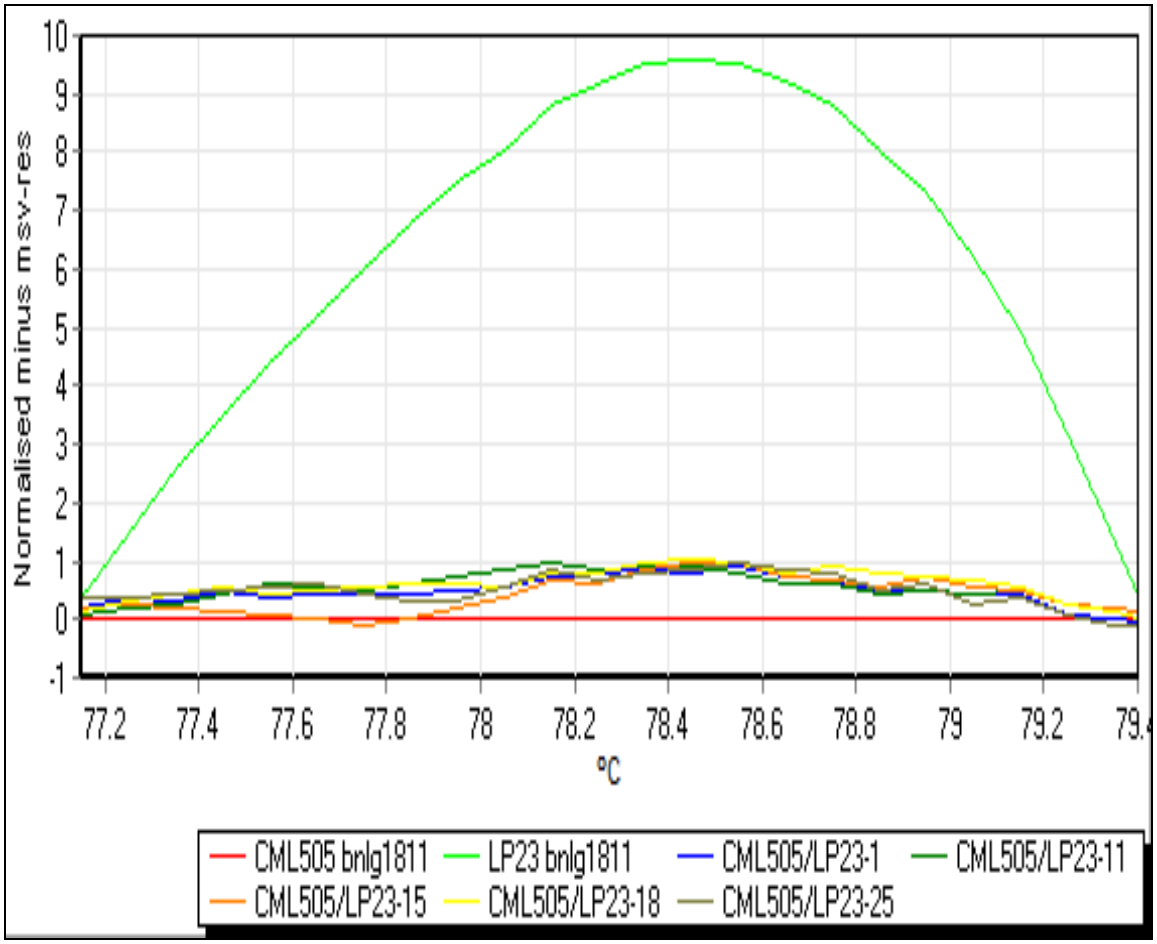

Figure 3.4: Difference plot of the parental LP23 line and CML505/LP23 progeny lines in relation to the parental line CML505 (baseline).

Fig. 3.5 shows a difference plot in which CML505 (MSV disease resistant parent) is normalised to the baseline and LP23 (MSV disease susceptible parent) and CML505/LP23 progeny lines are plotted on the difference graph. The CML505/LP23 progeny lines in Figure 3.5 aligned along the susceptible LP23 parent, and thus lines CML505/LP23-3, CML505/LP23-8, CML505/LP23-9, CML505/LP23-16, CML505/LP23-20 and CML505/LP23-23 were defined as MSV disease susceptible. The 28 CML505/LP23 progeny lines assayed with the SSR marker bnlg1811 are listed below in Table 3.3; those that showed MSV resistance had a temperature range that was between 82.03°C and 82.35°C. This marker detected 16 MSV disease resistant, 9 MSV disease susceptible and 3 heterozygous progeny lines. The confidence level values represent the likelihood that another sample will provide the same results in terms of the phenotypic trait. The confidence values ranged from 79.59% to 99.37% (Table 3.3).

Table 3.3: PCR-HRM melting-curve analysis using bnlg1811 SSR amplicons of parental lines CML505 and LP23 and the progeny from CML505/LP23. The genotype, T_m and the confidence values against the parental reference lines CML505 and LP23 values at 100% confidence were determined.

Name	Genotype	Peak 1	Peak 2	Confidence %
LP23 (PARENT)	**MSV-SUS**	**79.97** **82.47**	**82.20**	**100.00** **100.00**
CML505 (PARENT)	**MSV-RES**	79.85 79.77	82.18 82.18	95.84 98.69
CML505/LP23-3	msv-sus	79.43	82.15	89.91
CML505/LP23-4	msv-sus	80.43	83.05	97.77
CML505/LP23-5	msv-sus	79.75	82.15	97.63
CML505/LP23-7	het	79.58	82.12	99.37
CML505/LP23-8	msv-sus	79.45	82.10	98.70
CML505/LP23-9	msv-sus	80.45	83.25	83.89
CML505/LP23-16	msv-sus	79.48	82.15	99.31
CML505/LP23-19	het	79.70	82.22	97.98
CML505/LP23-20	msv-sus	80.25	83.25	84.51
CML505/LP23-23	msv-sus	79.30	81.85	92.00
CML505/LP23-26	het	82.23		97.68
CML505/LP23-27	msv-sus	82.35		93.99
CML505/LP23-1	msv-res	82.05		97.46
CML505/LP23-2	msv-res	82.15		95.94
CML505/LP23-6	msv-res	82.15		97.25
CML505/LP23-10	msv-res	82.10		79.59
CML505/LP23-11	msv-res	82.12		97.64
CML505/LP23-12	msv-res	82.20		98.15
CML505/LP23-13	msv-res	82.20		98.34
CML505/LP23-14	msv-res	82.15		80.57
CML505/LP23-15	msv-res	82.18		97.52
CML505/LP23-17	msv-res	82.27		91.92
CML505/LP23-18	msv-res	82.20		96.28
CML505/LP23-21	msv-res	82.20		91.82
CML505/LP23-22	msv-res	82.15		96.45
CML505/LP23-24	msv-res	82.03		80.42
CML505/LP23-25	msv-res			
CML505/LP23-28	msv-res			

sus=susceptible, res=resistant, het=heterozygous

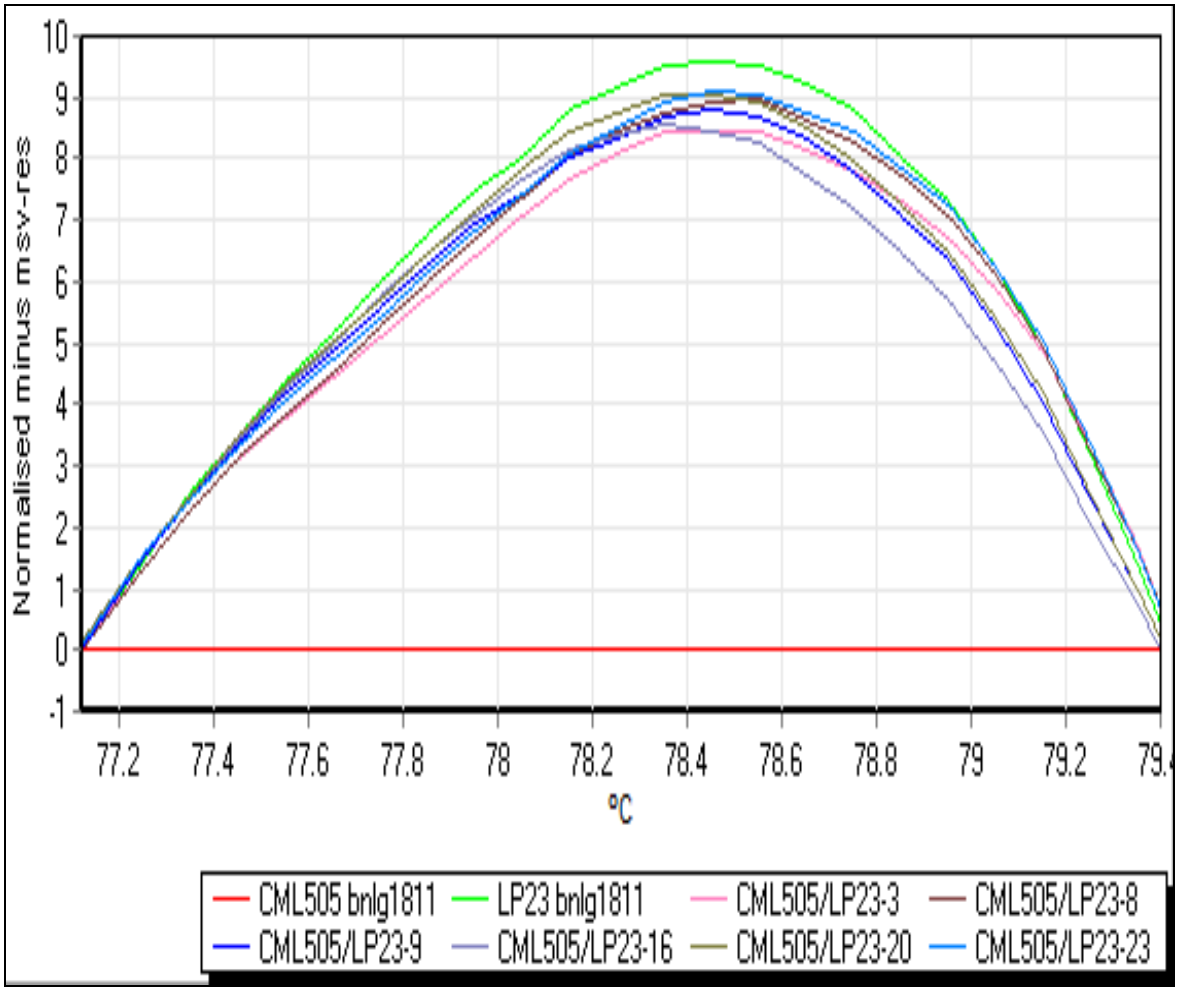

Figure 3.5: Difference plot of LP23 and the progeny lines with normalised CML505.

The SSR marker umc2228 was used to discriminate the progeny of CML509/LP23. Fig. 3.6 shows a difference plot with the MSV disease resistant parent CML509 normalised to the baseline. The progeny lines, CML509/LP23-63 and CML509/LP23-69 aligned with the LP23 (MSV disease susceptible line) and were therefore classified as MSV susceptible. The progeny line CML509/LP23-90 aligned with the MSV disease resistant parental line and was, therefore, classified as MSV resistant. High resolution melt analysis revealed that the resistant variety CML509 and the susceptible LP23 showed single melting peaks at 83.75°C and 85.06°C, respectively, using the SSR marker umc2228 (Table 3.4). Both parental lines were distinguished from each other in the melt profiles by a 1.31°C melt temperature difference. The marker bnlg1811 was unable to obtain clearly distinguishable profiles between these parental lines and was, therefore, not used to define this progeny population.

In total, 13 of the 90 CML509/LP23 progeny lines were regarded MSV disease resistant, 33 as MSV disease susceptible and 42 as heterozygous progeny (Table 3.4). Two samples were

contaminated thus discontinued. Fig. 3.7 is a representation of the melt profile of the heterozygous CML509/LP23-66 progeny. The progeny line CML509/LP23 66 possessed a peak that lay almost midway between those of the parental lines with SSR marker umc2228, with a T_m value of 84.55. The single peak melting temperatures, confidence values and assigned genotypes are shown in Table 3.4 of the progeny of CML509 and LP23 and the parental lines. The confidence values ranged from 70.04% to 99.92%.

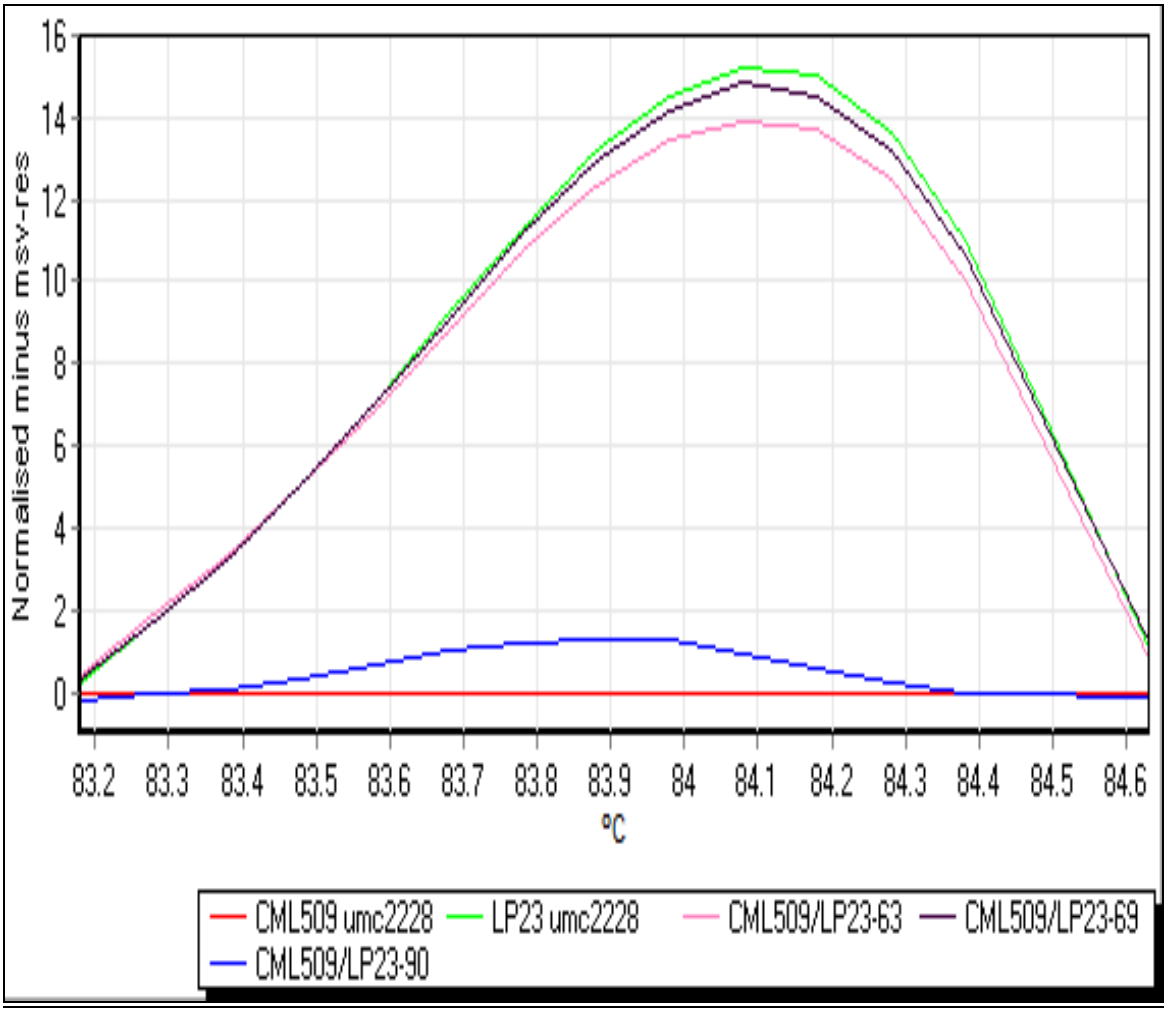

Figure 3.6: Representative profiles in the melting-curve analysis of parental lines CML509 MSV resistant (RED), LP23 MSV susceptible (GREEN) and progeny lines CML509/LP23 63, CML509/LP23 69 and CML509/LP23 90.

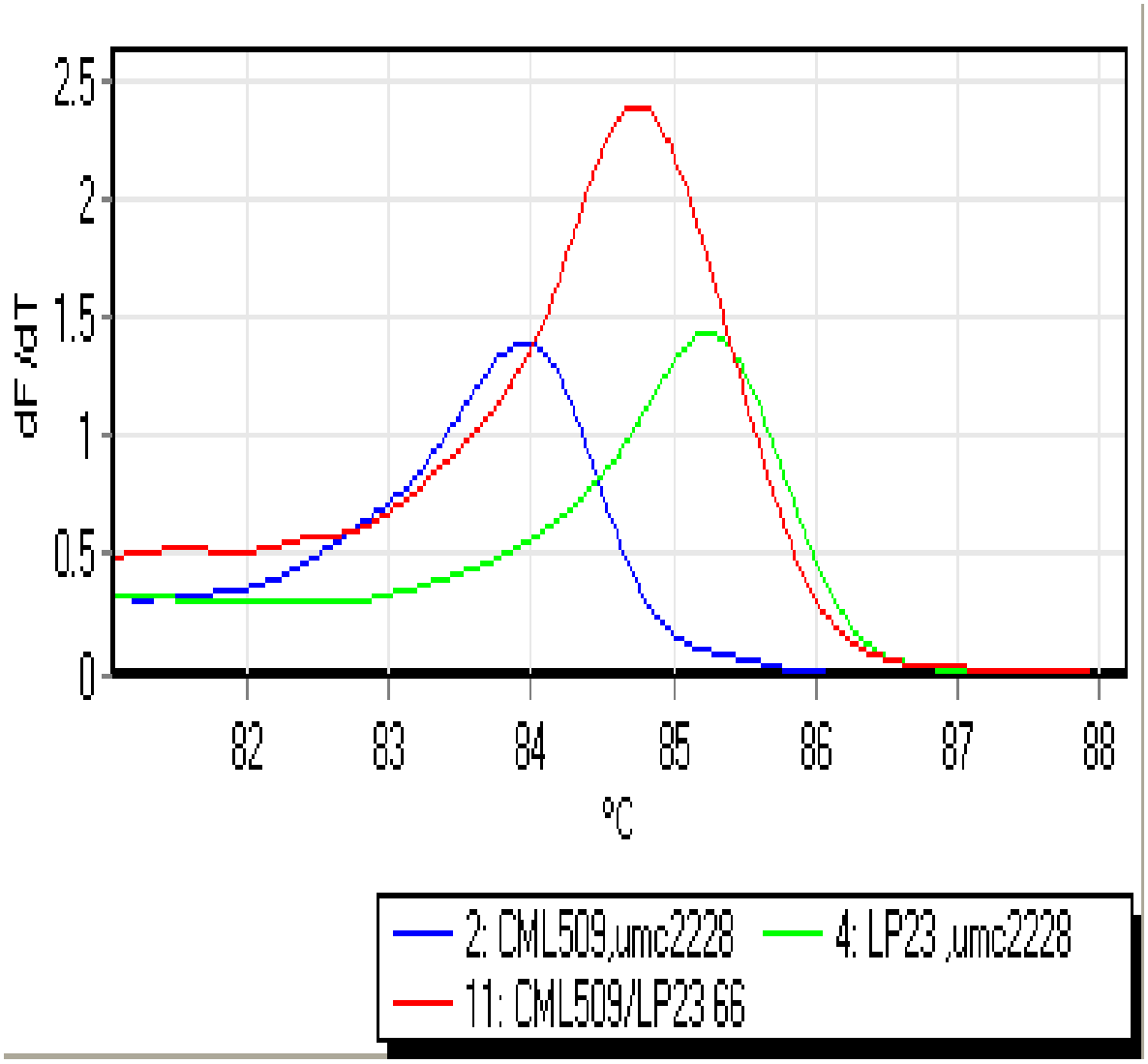

Fig 3.7: Representative melting profiles for the parental lines CML509 ($T_{m:}$ 83.85) and LP23 ($T_{m:}$ 85.05) and the heterozygous progeny line CML509/LP23 66 with a T_m value of 84.55.

Table 3.4: PCR-HRM melting-curve analysis using SSR umc2228 amplicons of parental lines CML509 and LP23 and their progeny. The genotype, T_m and the confidence values against the parental reference lines CML509 and LP23 values at 100% confidence were determined.

Name	Genotype	Peak	Confidence %
CML509 (Parent)	**MSV-Res**	**83.75**	**100**
LP23 (Parent)	**MSV-Sus**	**85.06**	**100**
CML509/LP23-1	msv-sus	85.35	87.11
CML509/LP23-2	msv-sus	85.45	83.93
CML509/LP23-3	msv-sus	85.33	83.57
CML509/LP23-4	msv-sus	85.20	93.92
CML509/LP23-5	msv-sus	85.28	90.66
CML509/LP23-6	msv-sus	85.28	89.53
CML509/LP23-7	msv-sus	84.92	99.83
CML509/LP23-8	msv-res	84.20	93.94
CML509/LP23-9	msv-res	84.18	98.05
CML509/LP23-10	het	84.65	93.58
CML509/LP23-11	mv-sus	85.23	93.91
CML509/LP23-12	msv-res	84.15	99.24
CML509/LP23-13	msv-sus	84.85	99.12
CML509/LP23-14	msv-sus	84.90	99.92
CML509/LP23-15	msv-res	83.24	97.39
CML509/LP23-16	msv-sus	85.22	90.97
CML509/LP23-17	msv-sus	85.12	96.85
CML509/LP23-18	msv-sus	84.80	98.84
CML509/LP23-19	msv-res	84.00	85.36
CML509/LP23-20	msv-sus	85.10	97.98
CML509/LP23-21	msv-res	84.17	99.75
CML509/LP23-22	msv-res	83.70	91.74
CML509/LP23-23	het	84.70	96.65
CML509/LP23-24	het	84.68	96.31
CML509/LP23-25	msv-res	84.00	96.03
CML509/LP23-26	msv-sus	85.23	92.93
CML509/LP23-27	msv-res	84.25	94.87
CML509/LP23-28	het	84.45	91.63
CML509/LP23-29	msv-sus	85.15	95.81
CML509/LP23-30	het	84.72	94.75
CML509/LP23-31	het	84.65	99.36
CML509/LP23-32	het	84.40	94.66
CML509/LP23-33	het	84.50	94.42
CML509/LP23-34	het	84.75	99.21
CML509/LP23-35	het	84.25	92.75
CML509/LP23-36	het	84.53	87.96
CML509/LP23-37	msv-sus	84.92	71.38
Name	**Genotype**	**Peak**	**Confidence %**

CML509 (Parent)	**MSV-Res**	**83.75**	**100.00**
LP23 (Parent)	**MSV-Sus**	**85.06**	**100.00**
CML509/LP23-38	het	84.40	94.61
CML509/LP23-39	het	84.50	94.98
CML509/LP23-40	het	84.42	96.97
CML509/LP23-41	het	84.15	87.26
CML509/LP23-42	het	84.57	98.99
CML509/LP23-44	het	84.45	95.05
CML509/LP23-45	het	84.63	91.91
CML509/LP23-46	het	84.53	95.56
CML509/LP23-47	het	84.40	92.63
CML509/LP23-48	het	84.42	94.84
CML509/LP23-49	het	84.65	98.12
CML509/LP23-50	msv-sus	85.10	98.45
CML509/LP23-51	het	84.55	98.88
CML509/LP23-52	het	84.45	96.94
CML509/LP23-53	het	84.65	83.45
CML509/LP23-54	het	84.43	97.43
CML509/LP23-55	het	84.45	97.42
CML509/LP23-56	het	84.55	95.46
CML509/LP23-57	het	84.55	93.81
CML509/LP23-58	het	84.53	93.78
CML509/LP23-59	het	84.40	87.01
CML509/LP23-60	het	84.43	97.39
CML509/LP23-61	msv-sus	84.85	86.68
CML509/LP23-62	msv-sus	84.80	85.14
CML509/LP23-63	msv-sus	84.90	90.56
CML509/LP23-64	msv-sus	84.98	91.09
CML509/LP23-65	msv-sus	84.90	91.49
CML509/LP23-66	het	84.55	79.54
CML509/LP23-67	msv-sus	84.92	92.59
CML509/LP23-68	msv-sus	84.85	90.17
CML509/LP23-69	msv-sus	84.97	95.31
CML509/LP23-70	het	84.75	70.04
CML509/LP23-72	het	84.78	80.49
CML509/LP23-73	msv-sus	84.86	87.21
CML509/LP23-74	msv-sus	84.80	87.29
CML509/LP23-75	msv-sus	84.90	91.46
CML509/LP23-76	msv-sus	84.88	90.40
CML509/LP23-77	msv-sus	84.85	90.38
CML509/LP23-78	msv-sus	84.95	94.27

Name	Genotype	Peak	Confidence %
CML509 (Parent)	**MSV-Res**	**83.75**	**100.00**
LP23 (Parent)	**MSV-Sus**	**85.06**	**100.00**
CML509/LP23 80	msv sus	84.82	87.40
CML509/LP23 81	msv-sus	84.70	83.91
CML509/LP23-82	msv-res	84.60	71.58
CML509/LP23-83	msv-sus	84.58	75.94
CML509/LP23-84	msv-res	84.52	74.91
CML509/LP23-85	msv-sus	84.63	73.67
CML509/LP23-86	msv-sus	84.57	74.38
CML509/LP23-87	msv-sus	84.55	74.22
CML509/LP23-88	msv-sus	84.70	83.11
CML509/LP23-89	msv-sus	84.70	84.19
CML509/LP23-90	msv-res	83.66	89.23

3.4 Discussion

In the present study, similar melting curve profiles for amplicons generated using DNA extracted from FTA™ cards and DNA extracted using a modified protocol from Edwards *et al*. (1991) were generated, indicating that the two DNA isolation methods were equally effective. The results are consistent with those of a study by Stoep *et al*. (2009) were the effects of DNA isolation methods on HRM results were examined. Results of several HRM tests using DNA samples purified using different extraction methods were compared. These methods included phenol extraction, Qiagen columns purification, and automated Chemagen and Gentra (Autopure) DNA isolations. Results of the study concluded that the DNA isolation method applied did not influence the HRM results.

The melt profile of the parental line LP23, in the current study obtained from amplification with genomic DNA isolated with the modified Edwards *et al*. (1991) protocol showed minimal differences in amplitude and curve shape from amplicons generated by genomic DNA isolated from FTA™ elute cards. The melting temperatures differed by 0.08°C, thus PCR amplicons generated using DNA extracted from FTA™ cards was equally effective as amplicons generated using DNA extracted by the Edwards *et al*. (1991) protocol for genotype characterisation. Diagnostic techniques can, thus be applied to DNA eluted from FTA™ cards in a manner equivalent to conventional DNA isolation methods, and the cost-effective technology significantly simplifies sampling and analysis of plants in both the laboratory and field environments.

FTA™ technology has been used for DNA extraction and storage in maize (Drescher and Graner, 2002; Ndunguru *et al*., 2005; Mbogori *et al*., 2006; Owor *et al*., 2007). These studies showed that the amount of plant material applied on FTA™ cards is critical and there has to be enough sap pressed onto the FTA™ cards until it is visible on both sides of the card. In the current investigation, the T_m value was the same when sap was either insufficient or sufficient since genomic DNA was extracted from the exact same sample material. However, pressing insufficient sap onto the FTA™ cards resulted in peaks with low dF/dT values, which could not be used for genotyping the samples. In the study by Mbogori *et al*. (2006), the age of the plant at the point of sampling was also critical, with older plants (over three months) being more difficult to sample compared to young plants (about one month old), increasing the time required per sample, thus in the present study, plants were sampled three weeks after emergence.

In the current study, the parental PCR amplicons of CML505, CML509 and LP23 when subjected to HRM analysis each achieved specific identification and differentiation profiles among the parental lines, thus enabling progeny differentiation. The melting profiles were reproducible and provided a unique profile for differentiation of genotypes. The observation that HRM is capable of distinguishing genotypes by the melting curves generated from PCR products is consistent with previous studies (Wittwer *et al*., 2003; Liew *et al*., 2004; Zhou *et al*., 2004). Detection of the presence of the *msv1* gene was carried out using two visualisations of the raw data (melting peak curve and difference curves). Clustering of melting or difference curves relative to the reference genotype identifies and groups similar variants (Montgomery *et al*., 2007). The melting peak curves are independent of normalisation and temperature shifting (Hondow *et al*., 2011). The difference curve view is, however, the most useful because the reference genotype forms a baseline with other groups displayed as positive differences (Montgomery *et al*., 2007).

Both homozygous and heterozygous sequence variants can be reliably differentiated using HRM analysis. Even those with very similar DNA profiles are distinguishable, for an example, genotypes of the neighbouring HbS and HbC single-base variants in β-globin (Gundry *et al*., 2003; Wittwer *et al*., 2003). However, in the current study, only homozygous *msv1* variants were differentiated for the purpose of achieving the specific objective of selecting for those lines with the homozygous *msv1* gene for further investigation. Field tests were done to verify the laboratory results and are presented in chapter four.

Simple sequence repeat molecular markers were used to track the MSV gene cluster. Melting profiles of the markers showed that two of the ten markers used, primer bnlg1811 which flanks the $(AG)_{16}$ repeat motif, and umc2228 which flanks the $(AGC)_4$ repeat motif, were able to distinguish all progeny lines. The polymorphism detected between the parental lines allowed for the classification of the progeny genotypes. These two markers gave the most distinguishable differences in melt profiles of the parental lines, as well as the largest melting temperature differences between the parental lines used in this study. The distance between T_m values allows for targeted primer flanking regions to be discriminated (Wilhelm and Pingond, 2003). The parental lines CML509 and LP23 were distinguished from each other in the melt profiles by 1.31°C melt temperature difference with marker umc2228.

In the present study, both parental lines, CML505 and LP23, were distinguished from each other by distinctly different melt profiles generated with primer bnlg1811. The parental line LP23 generated two peaks at T_m 79.97°C and 82.20°C, while CML505 generated a single peak with T_m at 82.47°C. The presence of multiple peaks adds further variation that can be used as a superior tool for species characterisation compared to single peaks (Rasmussen *et al*., 2007). Such variation allows for improved power of discrimination and genotyping of the HRM-curve analysis technique (Steer *et al*., 2008).

The CML505/LP23 progeny with melting temperatures close to the parental line CML505 were automatically designated by the Rotor-Gene software as MSV disease resistant progeny lines: CML505/LP23 1, CML505/LP23 2, CML505/LP23 6, CML505/LP23 10, CML505/LP23 11, CML505/LP23 12, CML505/LP23 13, CML505/LP23 14, CML505/LP23 15, CML505/LP23 17, CML505/LP23 18, CML505/LP23 21, CML505/LP23 22, CML505/LP23 24, CML505/LP23 25 and CML505/LP23 28. These families were therefore selected and advanced in the programme through self-pollination.

The DNA sequence of an amplicon determines the melt profile generated (Li *et al*., 2003). The T_m values of these peaks correspond to the alleles present in the sample and with the MSV disease resistant progeny, the allele is *msv1*. Additional peaks were generated by the rest of the progeny and the difference plot indicated that these aligned themselves away from the CML505 baseline.

Marker bnlg1811 detected 16 samples out of the 28 of the CML505/LP23 progeny as being MSV disease resistant. The confidence values ranged from 79.59% to 99.37%. The confidence value gives a value on the scale of 1 to 100% as to the relative probability of the sample belonging to a certain cluster (Bio-Rad, 2009). However, the choice of threshold will vary between different studies, but a threshold of 70% was recommended for investigations that aim to identify a large proportion of all SSRs and SNPs with only a moderate amount of manual review (Stephens *et al*., 2006). Marker umc2228 detected 13 MSV disease resistant CML509/LP23 progeny lines thus in total, of the 118 samples genotyped, 29 were MSV disease resistant. HRM analysis of the PCR products amplified by the SSR markers bnlg1811 and umc2228 proved to be a specific means of genotyping all the CML505 x LP23 and CML509 x LP23 progeny, respectively. The 29 MSV resistant families were verified in field trials. Lines that possessed field resistance as well were advanced in the programme through self-pollination to develop F_4 family lines.

3.5 Conclusions

From this study the following conclusions were drawn:

1) Maize DNA sampled using FTA^{TM} elute cards was suitable as a template for both PCR and HRM curve analysis. Diagnostic techniques can be applied to DNA eluted from FTA^{TM} elute cards in a manner equivalent to conventional DNA isolation methods. PCR-HRM analysis and FTA^{TM} technology were successfully used in this study to discriminate 118 F_3 maize family progeny genotypes according to susceptibility and resistance to MSV. The findings support that FTA technology is effective, which is in agreement with previous studies in the literature.
2) The *msv1* SSR markers bnlg1811 and umc2228 were able to successfully differentiate the MSV susceptible (LP23) and resistant parental maize inbred lines (CML505 and CML509) by the differences in their melt profiles due to HRM analysis. Thus results are consistent with previous findings that HRM analysis is an effective tool for use in MAS.
3) The application of MAS was successful for identifying 29 MSV resistant families which will be advanced in the breeding programme to develop new inbred lines.

References

Asins, M.J. 2002. Present and future quantitative trait locus analysis in plant breeding. Plant Breeding 121:281-291.

Bio-Rad. 2009. Precision Melt Analysis Software Instruction Manual. Bio-Rad Laboratories. http://www.bio-rad.com/webroot/web/pdf/lsr/literature/AW10014811-A.PDF (Accessed 16 Jun. 2011).

Bosque-Perez, N.A. 2000. Eight decades of maize streak virus research. Virus Research 71:107-121.

Chateigner-Boutin, A.L., and I. Small. 2007. A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Research 35:1-8.

CIMMYT. 2009. CIMMYT Maize Inbred Lines (CML). CIMMYT: International Maize and Wheat Improvement Center, Mexico City, Mexico. http://www.cimmyt.org/en/component/content/article/459-international-maize-improvement-network-imin/434-cimmyt-maize-inbred-lines-cml (Accessed 9 Jul. 2011).

Corbett Research. 2006. HRM™. High Resolution Melt Assay Design and Analysis. Corbett Life Science. http://www.gene-quantification.de/hrm-protocol-cls.pdf (Accessed 16 May 2011).

Cortes, A.L., E.R. Montiel, and I.M. Gimeno. 2009. Validation of Marek's disease diagnosis and monitoring of Marek's disease vaccines from samples collected in FTA cards. Avian Diseases 53:510-516.

Danson, J., M. Lagat, J. Ininda, and M. Kimani. 2006. Application of simple sequence repeats (SSRs) markers to study the resistance of locally adapted maize hybrids to damaging maize streak virus disease. African Journal of Biotechnology 5:1430-1434.

Denic, M. 2005. Maize breeding program: Breeding activities and national trials. National Agriculture Research Institute Annual Progress Report. Department of Agriculture and Farming Systems. Maputo, Mozambique. http://apps.cimmyt.org/english/docs/IPBS06-pdf (Accessed 1 Oct. 2010).

Drescher, A., and A. Graner. 2002. PCR-genotyping of barley seedlings using DNA samples from tissue plants. Plant Breeding 121:228-231.

Edwards, K., C. Johnstone, and C. Thompson. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research 19:1349.

Farooq, S., and F. Azam. 2002. Molecular markers in plant breeding 1: Concepts and characterisation. Pakistan Journal of Biological Sciences 5:1135-1140.

Fato, P. 2010. Investigation of heterotic patterns and genetic analysis of downy mildew resistance in Mozambican lowland maize (*Zea mays* L.) germplasm. PhD thesis. University of KwaZulu-Natal, Republic of South Africa. pp. 134.

Gady, A.L.F., F.W.K. Hermans, M. Van de Waal, E.N. Van Loo, R.G.F. Visser, and C.W.B. Bachem. 2009. Implementation of two high-throughput techniques in a novel application: detecting point mutations in large EMS mutated plant populations. Plant Methods 15:13-26.

Gundry, C.N., J.G. Vandersteen, G.H. Reed, R.J. Pryor, J. Chen, and C.T. Witter. 2003. Amplicon melting analysis with labelled primers: A closed-tube method for differentiating homozygotes and heterozygotes. Clinical Chemistry 3:396-406.

Hofinger, B.J., H.C. Jing, K.E.H. Kosack, and K. Kanyuka. 2009. High-resolution melting analysis of cDNA-derived PCR amplicons for rapid and cost-effective identification of novel alleles in barley. Theoretical and Applied Genetics 119:851-865.

Hondow, H.L., S.B. Fox, G. Mitchell, R.J. Scott, V. Beshay, S.Q. Wong, and A. Dobrovic. 2011. A high-throughput protocol for mutation scanning of the BRCA1 and BRCA2 genes. BMC Cancer 11:265-276.

Lehmensiek, A., M.W. Sutherland, and R.B. McNara. 2008. The use of high resolution melting (HRM) to map single nucleotide polymorphism markers linked to a covered smut resistance gene in barley. Theoretical and Applied Genetics 117:721-728.

Li, W., X. Biao, W. Yang, M. Hawkins, and U.K. Schubart. 2003. Complex DNA melting profiles of small PCR products revealed using SYBR Green I. Biotechniques 35:702-706.

Liew, M., R.J. Pryor, R. Palais, C. Meadows, M. Erali, E. Lyon, and C.T. Witter. 2004. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clinical Chemistry 50:1156-1164.

Lin J.H., C.P. Tseng, C.H. Chen, C.Y. Lin, S.S. Chang, H.S. Wu, and J.C Cheng. 2008. Rapid differentiation of influenza a virus subtypes and genetic screening for virus variants by high-resolution melting analysis. Journal of Clinical Microbiology 46:1090-1097.

Luo, Z.W. 1998. Detecting linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. Heredity 80:198-208.

Martin, D.P., and D.N. Shepherd. 2009. The epidemiology, economic impact and control of maize streak disease. Food Security 1:305-315.

Mbogori, M.N., M. Kimani, A. Kuria, M. Lagat, and J.W. Danson. 2006. Optimization of FTA technology for large scale plant DNA isolation for use in marker-assisted selection. African Journal of Biotechnology 5:693-696.

Montgomery, J., C.T. Wittmer, R. Palais, and L. Zhou. 2007. Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2:59-66.

Moscoso, H., E.O. Raybon, S.G. Thayer, and C.L. Hofacre. 2005. Molecular detection and serotyping of infectious bronchitis virus from FTA filter paper. Avian Diseases 49:24-29.

Ndunguru, J., N.J. Taylor, J. Yadar, H. Aly, J.P. Legg, T. Areling, G. Thompson, and C.M. Fauquet. 2005. Application of FTA technology for sampling, recovery and molecular characterisation of viral pathogens and virus-derived transgenes from plant tissues. Virology Journal 2:45.

Orlandi, P.A., and K.A. Lampel. 2000. Extraction-free, filter-based template preparation for rapid and sensitive PCR detection of pathogenic parasitic protozoa. Journal of Clinical Microbiology 38:2271-2277.

Owor, B.E., D.N. Shepherd, N.J. Taylor, R. Edema, A.L. Monjane, J.A. Thomson, D.P. Martin, and A. Varsani. 2007. Successful application of FTA^{R} Classic Card Technology and use of bacteriophage lambda29 DNA polymerase for large-scale field sampling and cloning of complete maize streak virus genomes. Journal of Virological Methods 140:100-105.

Pangasa, A., A.R. Jex, B.E. Campbell, N.J. Bott, M. Whipp, G. Hogg, M.A. Stevens, and R.B. Gasser. 2009. High resolution melting curve (HRM) analysis for the diagnosis of Cryptosporidiosis in humans. Molecular and Cellular Probes 23:10-15.

Rasmussen, J.P., C.P. Saint, and P.T. Monis. 2007. Use of DNA melting simulation software for in silico diagnostic assay design: targeting regions with complex melting curves and confirmation by real-time PCR using intercalating dyes. BMC Bioinformatics 8:107.

Reed, G.H., J.O. Kent, and C.T. Wittwer. 2007. High resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics 8:597-608.

Saxena, K.M.S., and A.L. Hooker. 1968. On the structure of a gene for disease resistance in maize. Genetics 61:1300-1305.

Sharma, K., and B.S. Misra. 2011. Molecular approaches towards analyzing the viruses infecting maize (*Zea mays* L.). Journal of General and Molecular Virology 3:1-17.

Shepherd, D.N., T. Mangwende, D.P. Martin, M. Bezuidenhout, J.A. Thomson, and E.P. Rybicki. 2007. Inhibition of maize streak virus (MSV) replication by transient and transgenic expression of MSV replication-associated protein mutants. Journal of General Virology 88:325-336.

Steer, P.A., N.C. Kirkpatrick, D. O'Rourke, and A.H Noormohammadi. 2008. Classification of fowl adenovirus serotypes by use of high-resolution melting-curve analysis of the hexon gene region. Journal of Clinical Microbiology 47:311-321.

Stephens, M., J.S. Sloan, P.D. Robertson, P. Scheet, and D.A. Nickerson. 2006. Automating sequence-based detection and genotyping of single nucleotide polymorphisms (SNPs) from diploid samples. Nature Genetics 38:375-81.

Stoep, N., C. Paridon, T. Janssens, P. Krenkova, A. Stambergova, M. Macek, G. Matthijs, and E. Bakker. 2009. Diagnostic guidelines for High-Resolution Melting Curve (HRM) analysis: An interlaboratory validation of BRCA1 mutation scanning using the 96-Well LightScanner™. Human Mutation 30:899-909.

Studer, B., L.B. Jensen, A. Fill, and T. Asp. 2009. ‘‘Blind’’ mapping of genetic DNA sequence polymorphisms in *Lolium perenne* L. by high resolution melting curve analysis. Molecular Breeding 24:191-199.

Sultan, D.M., M.M. Khalil, A.S. Abdouh, W.F. Poleh, and A.A.M. Al Muthanna. 2009. Imported malaria in United Arab Emirates: Evaluation of a new DNA extraction technique using nested PCR. Korean Journal of Parasitology 47:227-233.

Talmi-Frank, D., A. Nasereddin, L.F. Schnur, G. Schonian, S.O. Toz, C.L. Jaffe, and G. Baneth. 2010. Detection and identification of old world *Leishmania* by high resolution melt analysis. Neglected Tropical Diseases 4:581-586.

Tefera, T., S. Mugo, Y. Beyene, H. Karaya, and R. Tende. 2011. Grain yield, stem borer and disease resistance of new maize hybrids in Kenya. African Journal of Biotechnology 10:4777-4783.

Thottappilly, G., N.A. Bosque-Perez, and I.I.W. Rossel. 1993. Viruses and virus diseases of maize in Tropical Africa. Plant Pathology 42:494-509.

Wilhelm, J., and A. Pingond. 2003. Real-time polymerase chain reaction. Journal of Biochemistry 4:1120-1128.

Wisser, R.J., P.J. Kurti-Balint, and R.J. Nelson. 2006. The genetic architecture of disease resistance in maize: A synthesis of published studies. Phytopathology 96:120-129.

Wittwer, C.T., G.H. Reed, C.N. Gundry, J.G. Vandersteen, and R.J. Pryor. 2003. High resolution genotyping by amplicon melting analysis using LC Green. Clinical Chemistry 49:853-860.

Yu, R.H., X.H. Shan, S. Wang, X.H. Li, and Y. Jiang. 2011. A screening method for detecting simple sequence repeat (SSR) polymorphism of *Zea mays* using high-resolution melting curve analysis. African Journal of Biotechnology 10:16443-16447.

Zhou L., J.G. Vandersteen, T.Y. Wang, T. Fuller, M. Taylor, R. Palais, and C.T. Witter. 2004. High-resolution DNA melt analysis to establish HLA genotypic identity. Tissue Antigens 64:156-164.

CHAPTER FOUR

Phenotypic characterisation of progeny maize lines for maize streak virus (MSV) and downy mildew (DM) resistance

Abstract

Downy mildew (DM) and maize streak virus (MSV) are major diseases of maize which compromise yield in Mozambique. Pyramiding of resistance genes for DM and MSV was investigated in F_3 progeny derived from F_2 crosses between the commercial Mozambican line LP23 (DM resistant but MSV susceptible) and two known sources of resistance to MSV (maize lines CML505 and CML509). The objective was to create a single inbred line with combined resistance to both MSV and DM diseases through simultaneous selection for MSV and DM resistance in one base population. A total of 118 individuals of the $F_{2:3}$ progeny were screened for MSV resistance under artificial infestation with leaf hoppers (*Cicadulina mbila*, Naude) in the greenhouse at the International Maize and Wheat Improvement Centre (CIMMYT)-Zimbabwe. Trials were conducted simultaneously at Chokwe Research Station in Mozambique for both MSV and DM disease assessment under natural and artificial disease infestation, respectively. Downy mildew disease pressure was enhanced through the use of spreader rows. Forty-three of the progeny family lines obtained had moderate to high MSV disease resistance, and 75 lines were MSV disease susceptible under artificial MSV inoculation conditions. Seventy-one were resistant to DM and 47 were found to be DM susceptible at Chokwe. A negative correlation was observed between plant height and MSV disease infection. Individual plants from a total of 41 progeny lines that mostly exhibited superior disease tolerance to both DM and MSV disease and other desirable morphological traits were selected based on field screening and the presence of the *msv1* gene based on SSR data. It is from these lines that some individual plants were recommended for advancement to the F_4 generation through self-pollination. These will be subsequently fixed by continued self-pollination to obtain homozygous F_9 generation seeds.

Keywords: Maize streak virus (MSV); downy mildew (DM); molecular marker-assisted selection (MAS); gene pyramiding.

4.1 Introduction

Maize (*Zea mays L.*) is grown in all the agro-ecological zones of Mozambique and is the major staple food (Wulffe and Torp, 2005). Productivity of the crop is threatened by the DM and MSV disease (Denic *et al*., 2001). The *mastrevirus* MSV is transmitted by viruliferous leafhoppers of the genus *Cicadulina* (Bosque-Perez, 2000; Harkins *et al*., 2009) that suck sap from the plant leaves, concurrently injecting MSV into the plant (Hill, 2008). The symptoms of MSV disease on maize plants begin as spherical chlorotic spots, which later coalesce, causing chlorosis of the entire leaf lamina (Thottappilly *et al*., 1993; Magenya *et al*., 2008). Necrotic leaf areas can later develop from the chlorotic streaking in highly sensitive varieties (Thottappilly *et al*., 1993; Bosque-Perez, 2000; Fajemisin, 2003; Magenya *et al*., 2008) whereas partially or highly resistant varieties form few or no streaks (Efron *et al*., 1989). Chlorosis is caused by failure of chloroplasts to develop, thus impairing photosynthesis, which leads to reduced yields and premature death of the plant (Thottappilly *et al*, 1993; Bosque-Perez, 2000; Mawere *et al*., 2006). Yield losses in susceptible maize varieties due to MSV disease range from a trace to virtually 100% when the crop is infected in the first three weeks after planting (Magenya *et al*., 2008).

Downy mildew disease of maize caused by *Perenosclerospora sorghi* (Weston and Uppal) C.G. Shaw, can devastate the crop and result in yield losses ranging from 10-100% (Gowda *et al*., 1987). Symptoms include a characteristic pale green colour that begins at the base of the leaf, white powdery conidia, stunted plants that have upright stiff leaves and generally fail to produce seed and cobs (Craig and Frederiksen, 1983; Jewell *et al*., 1995; Jeger *et al*., 1998; Ajala *et al*., 2003). *Perenosclerospora sorghi* reproduces asexually by means of conidia (Jeger *et al*., 1998). Once infection is established in susceptible cultivars, the spread to neighbouring plants occurs via asexual conidia that are released from lower leaf surfaces following periods of high relative humidity (Bock *et al*., 1998; Jeger *et al*., 1998). The fungus requires high relative humidity (>85%) and cool temperatures (20-21ºC) to sporulate (Jewell *et al*., 1995).

Using chemicals such as carbofuran to eliminate the MSV-transmitting leafhopper (Magenya *et al*., 2008) and phosphonic acid for DM (Panicker and Gangadharan, 1998) are among the agrochemical control measures that can be employed. However, chemical control strategies such as these are relatively expensive, and may be harmful to the environment. Thus breeding

resistant cultivars is the most effective way of controlling these diseases especially for small-scale farmers (Sriwatanapongse *et al*., 1993; Asea *et al*., 2008). Combining multiple different resistance genes into a single genotype (gene pyramiding) broadens the expression of resistance in a new cultivar for sustainable durable control of biotic stress in crops (Eibach *et al*., 2007; Joshi and Nayak, 2010). Though the efficiency of plant breeding is improved, the process is tedious and costly, thus marker-assisted selection (MAS) based gene pyramiding can be employed to facilitate effective introgression of genes into a single genetic background (Joshi and Nayak, 2010). Molecular markers allow breeders to identify desirable traits at DNA level, thus they can be employed in young plants at each generation, increasing the speed of the pyramiding process (Joshi and Nayak, 2010). The goal of several international and regional maize breeding programmes like CIMMYT and the International Institute of Tropical Agriculture (IITA) has been to develop germplasm with resistance to multiple foliar diseases with good agronomic traits and adaptation across a variety of agro-climatological zones (Kim *et al*., 1989; Ceballos *et al*., 1991; DeVries and Toenniessen, 2001; Pratt *et al*., 2003). Progress has been made in the development of genetic resistance to many maize diseases such as MSV (Welz *et al*., 1998; Bosque-Perez, 2000; Mawere *et al*., 2006; Asea *et al*., 2008), with the resistant varieties playing a major role in hunger and poverty alleviation in many African countries (Bosque-Perez, 2000). However, large areas are still planted with varieties that lack resistance to important pathogens (Paliwal, 2000; Pratt *et al*., 2003).

Previous field evaluations carried out in Zimbabwe identified resistance to the MSV disease in CIMMYT inbred lines CML505 and CML509 (CIMMYT, 2009). These lines were used in the present study as the MSV resistant donor parental lines for the development of MSV disease resistant hybrids. Downy mildew disease is as problematic as the MSV disease in Mozambique, thus concurrent development of resistance to DM is essential. A commercially grown inbred line LP23 possesses resistance to DM and is adapted to various ecological production zones of Mozambique (Denic, 2005). However, because this line is susceptible to MSV disease it was crossed to the MSV resistant lines CML505 and CML509. Resistance to MSV and DM diseases can be combined in a single inbred line and such lines can be obtained through simultaneous selection for MSV resistance and DM resistance in one base population. The objectives of the present study were, therefore to: 1) develop maize inbred lines with combined MSV and DM disease resistance, 2) assess the levels of MSV and DM disease expression in the F_3 progeny population derived from three maize inbred lines,

CML505, CML509 and LP23; and 3) determine the effects of MSV disease on growth with emphasis on height of the maize plants. In this regard, a breeding scheme based on self-fertilisation, and selection (pedigree breeding method) of the progeny of the crosses between the two resistant varieties (CML505 and CML509) and the commercial MSV susceptible inbred line, LP23 was adopted.

4.2 Materials and methods

4.2.1 Locations

Simultaneous trials for the investigation of MSV disease expression were run at CIMMYT, Zimbabwe and from the Institute of Agriculture of Mozambique (IIAM) at Chokwe Research Station (40 m above sea level; latitude 24° 31´ S and longitude 33° 00´ E; average temp 23°C, min 17°C, max 30°C). Maize streak virus disease expression of the cross progenies at the F_3 level were assessed under artificial inoculation conditions at CIMMYT and under natural field conditions at Chokwe Research Station. At Chokwe there was also the artificial inoculation of spreader rows for DM disease assessment.

4.2.2 Germplasm and generations

The MSV resistant parents used in this study were the inbred lines CML505 and CML509, developed by CIMMYT, Zimbabwe. The susceptible parent was the inbred line LP23 developed at IIAM, Mozambique. This is a high grain yielding elite Mozambican line with a high level of resistance to DM disease but it is highly susceptible to MSV (Denic, 2005). In a pedigree breeding scheme, single plant selections were carried out at the F_2 stage through to the F_6 generations. In this study, two crosses were made: LP23 x CML505 and LP23 x CML509 to generate F_1 hybrids. The F_1 hybrids from the crosses were self-pollinated to develop F_2 populations segregating for MSV and DM disease resistance. Single plants were selected from amongst the segregating F_2 population and were further selfed at Makhathini Agricultural Research Station (27° 38´ 15; 32° 10´ E) in the KwaZulu-Natal Province of South Africa. Selection was based on agronomic superiority such as plant height, seed quality, grain texture, stay-green trait and vigour. The F_3 population was harvested, and 118

F_3 lines were selected from the two populations based on the above mentioned agronomic superiority traits and were therefore used for this study (Table 4.1).

Table 4.1: The 118 F_3 maize inbred lines selected.

Designated name	Pedigree	Origin	No. of plants observed	$F_{2:3}$ ears Selected
GCPMOZ7	CML505/LP23-F2B	09MAK28-7	49	28
GCPMOZ12	CML509/LP23-F2B	09MAK28-12	188	90
Total no. of plants				118

4.2.3 Establishment of screening nurseries and artificial inoculation for maize streak virus (MSV) disease in Zimbabwe

4.2.3.1 Planting

A total of 121 lines (eight pots/line): the 118 $F_{2:3}$ lines and the three parental lines (CML505, CML509 and LP23) were planted out in a greenhouse kept at 26-30^{o}C for mass production at CIMMYT-Harare Research Station. Three seeds per pot were planted in sterilised clay loam soil in order to minimise fungal disease. Germination occurred 5-8 days after planting and seedlings were thinned from three to one plant per pot. The maize fertiliser nitrogen: phosphorus: potassium (NPK) was applied by hand as basal application at the equivalent rate of 500 kg per ha. Additional nitrogen (N) fertiliser was applied as topdressing at 4 weeks after emergence at the same rate of 500 kg per ha. The plants were irrigated regularly as and when needed to avoid abiotic stress.

4.2.3.2 Inoculation

Plants were exposed to the MSV disease via infected viruliferous leafhoppers at the three leaf stage, approximately 8-10 days after planting as described by Bosque-Perez and Alam (1992). Non-viruliferous leafhoppers were reared on pearl millet (*Pennisetum americanum* [L] Leeke) in CIMMYT greenhouses for use as vectors in the experiments (Fig. 4.1A). The reared virus free leafhoppers were allowed to acquire the virus from stocks of infected MSV susceptible maize plants for two days. After the insects acquired the virus, the cages containing leafhoppers were covered using dark sheets with small openings for light

penetration. All adult insects that moved towards the light source were removed with a modified hand-operated vacuum cleaner (Fig. 4.1B). Recovered insects were anaesthetised with carbon dioxide prior to dispensing them into the whorls of plants (Bosque-Perez and Alam, 1992) (Fig. 4.1C). Three to five viruliferous, anaesthetised leafhoppers were transferred into the leaf whorl of each plant 10-12 days after planting and allowed to feed on maize seedlings for one week (Fig. 4.1D).

4.2.3.3 Disease assessment

MSV disease nurseries were established for evaluation of MSV disease severity and expression of the 118 F_3 family progeny lines along with the parental lines CML505, CML509 and LP23. Maize streak virus disease severity was scored twice during the growth period; first at four weeks after emergence and five days after the crop had flowered. Disease severity was scored on a 1-5 MSV disease rating scale (Rodier *et al*., 1995; Ngwira and Khonje, 2002; Ininda *et al*., 2006). This scale is adapted from the IITA, where 1 = no streaking to very light streaking (specks with no subsequent development); 2 = light streaking on old leaves gradually decreasing on young leaves; 3 = moderate streaking on old and young leaves; 4 = severe streaking on 60% of leaf area, plants stunted; 5 = severe streaking on all leaves (≥75%), plants severely stunted, dying or dead. The IITA's best known accomplishment in maize improvement has been the development of the practical resistance screening system used in this study for large-scale field use in MSV disease evaluation (Bosque-Perez and Alam, 1992). It is also the standard scale used at CIMMYT-Zimbabwe (Jewell *et al*., 1995).

Figure 4.1: Artificial inoculation of maize lines for maize streak virus disease screening **A**: Leafhoppers fed on pearl millet plants. The arrows indicate leafhoppers on pearl millet leaves, **B**: Leafhoppers within the cages were attracted to the light. Small plastic collection vials connected to modified vacuum cleaners with one end having a narrow tube were used to collect the insects, **C**: To ease infestations the leafhoppers were anesthetised with carbon dioxide immediately before dispensing them, **D**: Insects were then dispensed into the leaf whorl at a rate of 3-4 leafhoppers per plant. Photos by: Nothando F. Mafu, PMB, UKZN. Taken at CIMMYT, Zimbabwe. Date: 06/12/10.

Visual estimates were made on the 118 rows and the mean disease rating of each progeny row was established after which individual plant assessments were made within the selected

rows. The inbred parental lines CML505 and CML509 served as resistant controls, while LP23 served as the susceptible control. Plant height was also measured eight weeks after emergence to evaluate the effects of the MSV disease. The heights were recorded during the second MSV scoring.

4.2.4 Establishment of screening nurseries and artificial inoculation of spreader rows for downy mildew (DM) disease in Chokwe.

The F_3 progenies were screened for DM resistance under artificial infestation at Chokwe Research Station, Mozambique. Seed was planted in non-replicated blocks on the 21st of January 2011 and harvested on the 18th of May 2011. Thirty plants were planted per family. Spreader rows were planted to enhance disease infection as suggested in previous studies (Cardwell *et al*., 1997; Abalo *et al*., 2009). Agronomic data such as the days to mid-pollen shed (DMP), the days to mid-silking (DMS) and the number of ears selected in each family was collected. The DMP value was measured as the number of days after planting when 50% of the plants were shedding pollen. The DMS value was measured as the number of days after planting when silks emerged.

4.2.4.1 lnoculation

For screening purposes, inoculum in the form of debris must be collected and applied directly to test material because DM is caused by an obligate parasite and therefore cannot be cultured. The pathogen is only able to infect, grow and produce spores in living host tissue. Conidia germinate and lose viability within hours after sporulation (Jewell *et al*., 1995). Downy mildew diseased leaves were harvested from local fields (Fig. 4.2A). Leaves were sterilised by washing with the disinfectant Javel™ (3.0% sodium hypochlorite) (Fig. 4.2B) and then rinsed with tap water to remove old conidia (Fig. 4.2C). Seeds of a DM susceptible local landrace variety were pre-germinated for 96 hours and washed with Javel for 1 min, and then washed with tap water. Inoculation of seed from the local landrace variety was done by placing a layer of seedlings over the layer of diseased leaves in clean containers and left to sporulate overnight in a 21^{o}C incubator (Fig. 4.2D).

4.2.4.2 Planting

The infected seedlings were planted in black clay loam soils at both ends of the field as spreader rows with 0.5 m intra-row spacing (Fig. 4.2E). The technique was developed by a plant pathology team led by Dr Bobby Renfro (Sriwatanapongse *et al*., 1993). The field was channel irrigated regularly to achieve 600 mm of precipitation which is equivalent to the annual rainfall for the station ([1]Fato, 2011; pers. comm.).

The insecticide cypermethrin was applied two weeks before planting. The pathogen was already sporulating on the spreader plants at planting of the experimental progeny rows. The 118 F_3 populations together with the three parental lines (CML505, CML509 and LP23) were planted 30 seeds per population two weeks after the first planting of spreader rows. Each entry was planted in non-replicated 5 m single row plots each, with spacing of 0.80 m between rows and 0.25 m between plants within each row. The experiment was laid out as an alpha lattice design comprising of 12 plots in 10 blocks (10 x 12 alpha lattice). The experiment was augmented by replicating the control plots in each block. With respect to DM screening the lines CML505 and CML509 were the DM susceptible checks, while LP23 was the resistant control.

To ensure higher levels of infection, seven days after planting, a second planting of spreader rows was conducted in between the initial plantings using non-infected seed from the same susceptible maize local landrace variety. This gave a resultant intra-row spacing of 0.25 m. Agronomic practices included fertiliser application at the rate of 300 kg ha^{-1} compound (NPK 12:24:12) at planting and supplemented with 69 kg N ha^{-1} 6 wk after planting followed by weeding and top dressing with urea at the rate of 150 kg ha^{-1} during the vegetative stage. The fields were maintained clean through regular hand weeding as and when necessary.

[1]Dr Pedro Fato, Chokwe Research Station, Institute of Agriculture of Mozambique (IIAM).

Figure 4.2: Steps of spreader rows establishment for downy mildew artificial infestation. **A**: DM diseased leaves were harvested after which, **B**: leaves were sterilised by washing with a disinfectant, and then **C**: rinsed with tap water to remove old conidia, **D**: Susceptible seed was inoculated by placing a layer of seedlings over the layer of diseased leaves in clean containers and left to sporulate overnight in a 21^{o}C incubator. **E**: The infected seedlings were

then planted at both ends of the field as spreader rows. Photos by: Pedro Fato, IIAM Chokwe, Mozambique.

4.2.4.3 Disease assessment

Visual estimates of DM disease severity were made on the whole plots and the mean disease rating of each progeny line was established. Thereafter, individual plant assessment of disease incidence was done by counting the number of diseased plants (plants showing symptoms of DM disease) in each plot six weeks after emergence. In the 2010/11 season the level of natural occurrence of MSV in Chokwe was high, allowing for the evaluation of material for MSV resistance under natural field conditions, using the scale described above. The score for the severity of DM symptoms was rated on a scale from 1 = clean, no infection to 5 = severely diseased which is used at CIMMYT for the regional trials (Vivek *et al*., 2005). Planting and harvesting of experimental materials were done manually on the 21st of January 2011 and the 18th of May 2011, respectively.

4.2.5 Statistical analysis

The data for MSV disease ratings and plant heights for all plants and the parental lines LP23, CML505 and CML509 were analysed using a one-way analysis of variance (ANOVA) followed by a Newman-Keuls Multiple Comparison Test to detect significant differences amongst the lines. Relationships between MSV disease ratings and plant height, and DM disease ratings and plant height, were determined using Pearson's phenotypic correlation. All statistical analyses were conducted using GraphPad Prism Statistical Package, Version 5.04 (GraphPad Software, Inc., 2010).

4.3 Results

4.3.1 Disease expression and effect of MSV disease on growth of parents

At CIMMYT, artificial MSV inoculations were successful as disease development was as expected in the parental inbred lines. The differences in MSV disease expression between the

MSV resistant parents, CML505 and CML509, and the susceptible LP23 were significant (Table 4.2). The resistant parents CML505 and CML509 had mean MSV scores of 1.5 ± 0.1 and 2 ± 0.2, respectively; however, this difference was not statistically significant (P>0.05). The line LP23 had a mean score of 4 ± 0.3, significantly higher than both CML505 and CML509. Line LP23 was the most affected by MSV of the three parents as evidenced by plant height (650 ± 20.5 mm), significantly lower than CML505 and CML509 (750 ± 28.5 and 900 ± 11.9, respectively) (Table 4.2). Although there were no significant differences in MSV scores between CML505 and CML509, the latter had a significantly higher plant height than the former, which is attributed to genotypic differences.

Table 4.2: Mean height and MSV scores and reaction for the three parental lines.

Parent	Mean MSV score*	MSV reaction	Mean height (mm)*
CML505	1.5 ± 0.1^{a}	Resistant	750 ± 28.5^{a}
CML509	2 ± 0.2^{a}	Resistant	900 ± 11.9^{b}
LP23	4 ± 0.3^{b}	Susceptible	650 ± 20.5^{c}

* Dissimilar alphabet characters denote a statistical significance (One-way ANOVA; MSV score F pr. < 0.0001; ……Height F pr. < 0.0001; n = 8, mean ± standard error (SE)).

4.3.2 Disease expression and effect of MSV on growth of the progeny

Based on visual MSV disease assessment of the 118 F_3 progeny lines under artificial MSV inoculation with the virus at CIMMYT, Harare, nine lines were rated 1 (highly resistant), 13 were rated 2, 21 were rated 3, 53 were rated 4 and 22 were rated 5 (highly susceptible). Of the 28 progeny lines from CML505 and LP23, 15 (54%) were rated as resistant, while of the 90 progeny lines from crosses between CML509 and LP23, 28 (31%) were resistant. There was a differential reaction of the family lines ranging from resistant to susceptible. The majority of the progeny population was skewed towards susceptibility with 75 progeny lines having a score of 4 and 5 (Fig. 4.3A). There was an insignificant negative correlation ($r = -0.096$, $p > 0.05$) between MSV disease rating and plant height.

Of the 118 progeny lines under natural MSV infestation field conditions in Chokwe, 10 lines were rated 1 (highly resistant), 47 were rated 2, 21 were rated 3, 19 were rated 4 and 21 were rated 5 (highly susceptible) to MSV (Fig. 4.3B). There was significant negative correlation ($r = -0.187$, $p < 0.05$) between MSV disease rating and plant height.

4.3.3 Disease expression of DM in the progeny lines

Of the 118 progeny lines, 39 lines were rated 1 (highly resistant), eight were rated 2, 24 were rated 3, 28 were rated 4 and 19 were rated 5 (highly susceptible). Amongst the 28 progeny lines from CML505 and LP23, 26 (93%) were resistant, while of the 90 progeny lines from crosses between CML509 and LP23, 45 (50%) were resistant in the field. Forty-seven of all the progeny lines were DM disease susceptible having a score of 4 and 5 (Fig. 4.4). There was a significant negative correlation ($r = -0.4086$, $p < 0.05$) between DM disease rating and plant height. Individual plants with low disease scores of 1, 2 and 3 from each plot were self-pollinated to advance seed to the F_4 generation and seeds from individual cobs were packed separately and used in further investigations.

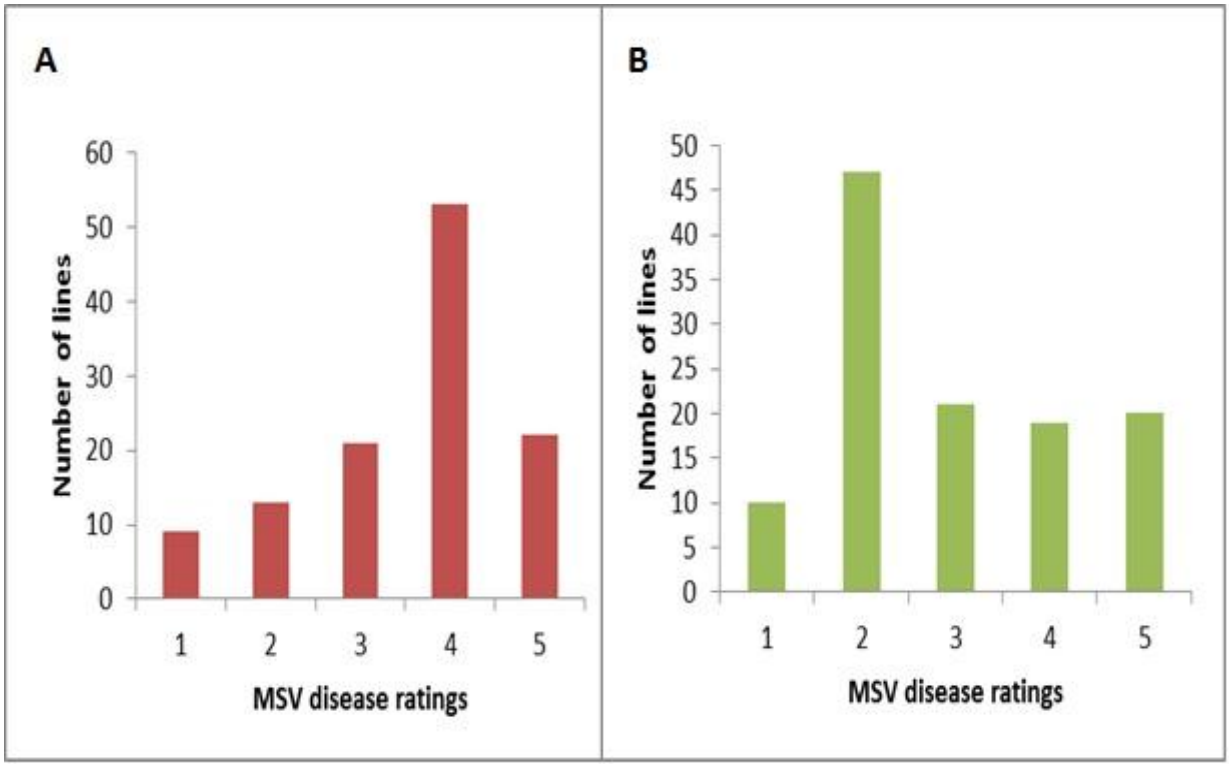

Figure 4.3: Maize streak virus disease expression scores for the 118 F_3 progeny lines in A) Harare and B) Chokwe.

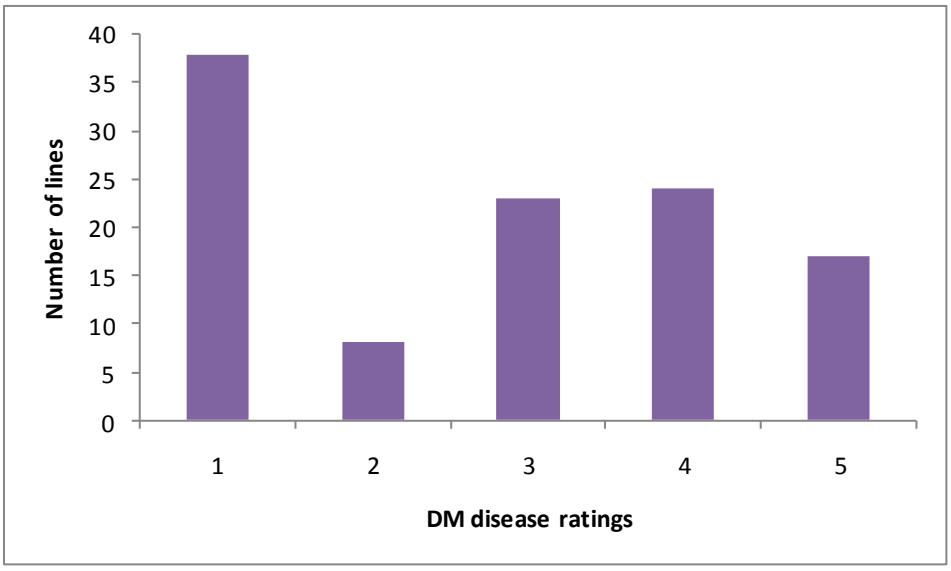

Figure 4.4: Downy mildew disease expression scores for the 118 F_3 progeny lines in Chokwe.

Figure 4.5 is a representation of height distribution among the 118 F_3 progeny lines. The most prominent height range was 51-60 cm with the least being 111-120 cm. The graph is skewed to the left towards the lower height values thus infection with MSV and DM results in overall reduced plant height.

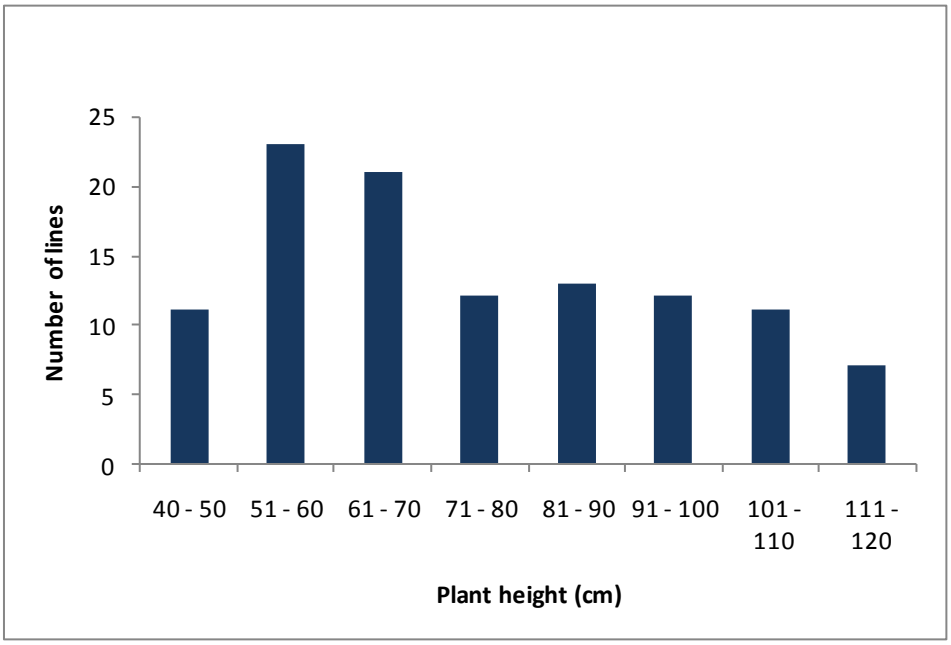

Figure 4.5: Plant heights of the 118 F_3 progeny lines infected with maize steak virus and downy mildew disease.

Selection was conducted both between families and within families at the F_3 population level and subsequent generations. Firstly, selection was of families with the best MSV and DM mean scores (1-2.5) and then the best individuals with high disease resistance within those families. Individual plants from a total of 41 progeny lines, that exhibited MSV disease severity ratings of 2.5 or less in both locations within each of the F_3 family lines, were subsequently selected based on the presence of the *msv1* gene as detected using SSR data (Chapter three) or field DM disease resistance in the field thus for instance NM-4 and NM-5 though revealed on marker were selected because of superiority interms of DM resistance. These were selected and advanced to the F_4 generation to be fixed for use to improve maize hybrids in Mozambique for MSV resistance. The number of individuals that were selected in each family is indicated in the Table 4.3 below. Table 4.3 also shows the phenological features of the selected $F_{3:4}$ progenies.

Table 4.3: Pollen shedding and silking dates, number of ears and the MSV and DM scores of 41 progeny lines.

Entry	Stock ID	Pedigree	DMP	DMS	MSV score	MSV marker present	DM incidence (%)	No. of ears selected
1	NM-4	CML505/LP23-F2B-4	55	57	2	No	0	1
2	NM-5	CML505/LP23-F2B-5	58	57	1	No	0	2
3	NM-6	CML505/LP23-F2B-6	54	53	1	Yes	9	1
4	NM-8	CML505/LP23-F2B-8	55	54	1	No	0	3
5	NM-9	CML505/LP23-F2B-9	57	58	2	No	0	2
6	NM-10	CML505/LP23-F2B-10	53	52	1	Yes	0	1
7	NM-12	CML505/LP23-F2B-12	56	57	1	Yes	25	1
8	NM-13	CML505/LP23-F2B-13	52	53	1	Yes	0	1
9	NM-15	CML505/LP23-F2B-15	53	54	1	Yes	0	2
10	NM-17	CML505/LP23-F2B-17	56	58	1	Yes	0	1
11	NM-18	CML505/LP23-F2B-18	52	53	2	Yes	0	2
12	NM-22	CML505/LP23-F2B-22	55	55	1	Yes	0	3
13	NM-24	CML505/LP23-F2B-24	53	56	1	Yes	0	1
14	NM-25	CML505/LP23-F2B-25	52	54	1	Yes	0	2
15	NM-27	CML505/LP23-F2B-27	52	52	1	No	0	3
16	NM-28	CML505/LP23-F2B-28	54	54	1	Yes	0	1
17	NM-30	CML509/LP23-F2B-2	55	52	2.5	No	36	1
18	NM-31	CML509/LP23-F2B-3	55	55	2	No	38	1
19	NM-33	CML509/LP23-F2B-5	55	58	2	No	0	1
20	NM-38	CML509/LP23-F2B-10	55	56	2	No	21	1
21	NM-39	CML509/LP23-F2B-11	55	57	2	No	17	2
22	NM-42	CML509/LP23-F2B-14	56	55	1	No	13	2
23	NM-43	CML509/LP23-F2B-15	56	54	1	Yes	64	1
24	NM-49	CML509/LP23-F2B-21	54	56	1.5	Yes	80	1
25	NM-51	CML509/LP23-F2B-23	56	58	1	No	33	1
26	NM-53	CML509/LP23-F2B-25	52	51	1.5	Yes	6	1
27	NM-54	CML509/LP23-F2B-26	55	56	1	No	20	2
28	NM-55	CML509/LP23-F2B-27	56	55	1	Yes	0	2
29	NM-58	CML509/LP23-F2B-30	55	53	1	No	22	2
30	NM-59	CML509/LP23-F2B-31	58	56	1	No	60	1
31	NM-63	CML509/LP23-F2B-35	55	52	1	No	10	1
32	NM-68	CML509/LP23-F2B-40	56	55	1.5	No	22	1
33	NM-69	CML509/LP23-F2B-41	56	53	1	No	33	1
34	NM-72	CML509/LP23-F2B-44	54	52	1.5	No	10	1
35	NM-87	CML509/LP23-F2B-59	53	51	1	No	0	1
36	NM-93	CML509/LP23-F2B-65	56	58	1	No	0	1

Entry	Stock ID	Pedigree	DMP	DMS	MSV	MSV	DM	No. of

					score	marker present	incidence (%)	ears selected
39	NM-102	CML509/LP23-F2B-74	58	65	1	No	0	1
40	NM-107	CML509/LP23-F2B-79	57	56	1	Yes	0	1
41	NM-112	CML509/LP23-F2B-84	63	58	1.5	Yes	0	1

4.4 Discussion

The adoption of MSV disease resistant lines is of great importance to improving maize yields in Africa where the disease is a serious biotic constraint (Thottappilly *et al*., 1993; Martin and Shepherd, 2009). In the present study, the CIMMYT inbred donor lines CML505 and CML509 exhibited high MSV disease resistance, confirming previous observations made at CIMMYT (CIMMYT, 2009). The negative correlation between MSV disease rating and plant height under artificial and natural infestation conditions in Harare and Chokwe, respectively, indicated the negative effect of MSV disease on growth of the infected susceptible progeny lines. Stunting was also observed in the susceptible parent, LP23, which had the lowest mean plant height amongst the three parental lines. This was consistent with the findings of Bosque-Perez *et al*. (1998) who concluded that MSV disease scores negatively correlated with plant height, ear length and diameter, dry weight and grain weight per plot. Furthermore, a study by Barrow (1992) from Pannar Seeds, Greytown also noted that highly MSV susceptible hybrids were markedly stunted in height.

The field study aimed at the identification of new stable sources of high-yielding MSV disease resistant lines in two different locations to simultaneously enhance the efficiency of the breeding program. MSV disease expression however, differed between CIMMYT, Zimbabwe and Chokwe, Mozambique. Genotype x environment (GxE) interaction is the change is a cultivar's relative performance across environments due to different responses to numerous biotic, climatic and edaphic factors (Dixon *et al*., 1991).

In the absence of MSV disease, resistant hybrids should possess stable disease expression in diverse environments to make them superior hybrids (Barrow, 1992; KARI; 1995). Results from the present study were collected under artificial MSV disease pressure and natural MSV disease infestation conditions at CIMMYT and Chokwe, respectively. A study by Ngwira and Khonje (2002) evaluated gray leaf spot (GLS) infected nurseries under artificial inoculation pressure in the greenhouse and in the field and concluded that natural disease pressure was

not enough to make conclusive evaluations. In the current study, the MSV disease severity score results recorded at Chokwe were not used to make any evaluations. Maize streak virus infection levels were much higher in the greenhouse than in the field in this study, probably because temperature and humidity were more conducive to streak development in the greenhouse. This is consistent with a study by Álvarez-Alfageme *et al*. (2011) where powdery mildew infection was analysed under greenhouse and field conditions.

Selection for superior plants with both MSV and DM disease resistance involved both visual assessment for esistance to the diseases and laboratory marker tests for the MSV disease resistance trait. Marker assisted selection is an indirect selection process where a trait of interest is selected for based on a QTL 'tagged' marker linked to the QTL (Ribaut and Hoisington 1998; Rosyara 2006). Markers should be tightly linked to target loci, preferably less than 5cM genetic distance to reduce the possibility of recombination between the marker and QTL (Langridge *et al*., 2001; Collard and Mackill 2008). Recombination explains the lack of correlation between the marker and field data in the present study. In other words, a marker assay may not predict phenotype with 100% reliability (Collard and Mackill 2008). However, the markers enabled plant selection in order to select a subset of plants thus reduced the number of plants that needed to be phenotypically evaluated in the study (Chapter 3). Phenotypic screening can be strategically combined with MAS. In the first instance, 'combined MAS' (coined by Moreau *et al*., 2004) may have advantages over phenotypic screening or MAS alone in order to maximize genetic gain (Thormann *et al*., 1994; Danson *et al*., 2006; Karanja *et al*., 2009). Karanja *et al.* (2009) stated that compared to most crops "maize exhibits a wider range of morphological and molecular dynamism", thus the need to use both molecular markers and morphological markers in maize breeding programmes.

A total of 41 progeny lines exhibited disease severity ratings of 2.5 or less in both locations. Single plant selections were made on the basis of superiority of MSV and DM resistance expression in the field or on the presence of the *msv1* gene with field resistance to MSV. These plants had minimal incidence and severity of maize streak and the varieties were also superior to other agronomic traits like plant height. It is these single plants that were advanced to the F_4 generation. Estimates of heritability for any trait are less precise for F_2 and F_3 populations than they might be if populations were homozygous (McGrath *et al*., 2005). Hence, in the initial stages where there is greater segregation only single plants are

advanced. Single plant selection is repeated, but with an increasing shift from individual plants to row performance, until plants are near homozygous (e.g. F_5). Further testing of the selected lines is in progress.

Results from the present study showed that CML505 x LP23 progeny lines had a greater percentage of MSV resistant lines with 15 out of 28 (54%) lines compared with 28 out of 90 (31%) from CML509 x LP23 under artificial conditions at CIMMYT. This is probably due to the fact the maize donor parent CML505 has a significantly lower MSV disease rating of 1.5 as compared to 2 for CML509 as identified in this study and previous screen house and field evaluations carried out at the International Maize and Wheat Improvement Centre, Zimbabwe (CIMMYT, 2009). This suggests that the parental line CML505, which showed higher levels of resistance, imparts MSV disease resistance to a greater extent than CML509. MSV is a quantitative trait so it is likely that CML505 has more resistance QTL than CML509.

Downy mildew disease pressure was enhanced through the spreader rows in the trial run at Chokwe. The spreader row technique, which is not labour intensive (Ajala *et al.*, 2003), has also been used in screening maize germplasm in a Mozambican programme at Umbeluzi Research Station (Denic *et al*., 2001). In the present study DM susceptibility was also associated with stunted growth.

4.5 Conclusions

From this study the following conclusions were drawn:

1) The MSV resistance levels ranged between highly resistant (score of 1) and highly susceptible (score = 5), which provides the opportunity for selection. Of the 118 progeny lines from both the CML505 x LP23 and CML509 x LP23 crosses, 43 (36%) were MSV disease resistant with resistance scores less or equal to 2.5 under artificial and natural infestation in the greenhouse and field, respectively. Thus the objective of developing MSV resistant lines was achieved.

2) It was concluded, as was by a study by Bosque-Perez *et al*. (1998), that MSV disease affects the agronomic trait of plant height. MSV disease susceptible lines were significantly shorter than the MSV disease resistant progeny in the study, and in general disease scores were negatively correlated with plant height indicating that disease infection affected plant growth and development.
3) Using an adapted lowland tropical germplasm with DMSR (downy mildew and maize streak virus resistance) and a MSR (maize streak resistant) background, the genes for MSV and DM disease resistance were successfully stacked in the breeding populations (CML505 x LP23, and CML509 x LP23). Incorporation of these genes in the MSV susceptible but locally adapted LP23 has enhanced the levels of MSV and DM resistance.
4) The 41 plants advanced to F_4 generation will be fixed by continued self-pollination to obtain F_9 generation seeds which are homozygous (F, coefficient of inbreeding about 99.99). They will be used to improve the maize hybrids in Mozambique for MSV and DM resistance with positive implications on food security in Mozambique. These will be valuable to the maize breeding programme in Mozambique.

References

Abalo, G., P. Tongoona, J. Derera, and R. Edema. 2009. A comparative analysis of conventional and marker-assisted selection methods in breeding maize streak virus resistance in maize. Crop Science 49:509-520.

Ajala, S.O., J.G. Kling, S.K. Kim, and A.O. Obajimi. 2003. Improvement of maize populations for resistance to downy mildew. Plant Breeding 122:328-333.

Álvarez-Alfageme, F., S. Burg, and J. Romeis. 2011. Infestation of transgenic powdery mildew-resistant wheat by naturally occurring insect herbivores under different environmental conditions. PLoS ONE 67: e22690. doi:10.1371/journal.pone.0022690.

Asea, G., B.S. Vivek, G. Bigirwa, P.E. Lipps, and R.C. Pratt. 2008. Validation of consensus quantitative trait loci associated with resistance to multiple foliar pathogens of maize. Phytopathology 99:540-547.

Barrow, M.R. 1992. Development of maize hybrids resistant to maize streak virus. Crop Protection 11:267-271.

Bock, C.H., L.K. Mughogho, M.J. Jeger, E. Mtisi, and K.F. Cardwell. 1998. Production of conidia by *Peronosclerospora sorghi* on sorghum crops in Zimbabwe. Plant Pathology 47:243-251.

Bosque-Perez, N.A. 2000. Eight decades of maize streak virus research. Virus Research 71:107-121.

Bosque-Perez, N.A., and M.S. Alam. 1992. Mass Rearing of *Cicadulina* leafhoppers to Screen for Maize Streak Virus Resistance: A Manual. International Institute of Tropical Agriculture, Ibadan, Nigeria. pp. 22.

Bosque-Perez, N.A., S.O. Olojede, and I.W. Buddenhagen. 1998. Effect of maize streak virus disease on the growth and yield of maize as influenced by varietal resistance levels and plant stage at time of challenge. Euphytica 101:307-317.

Cardwell, K.F., J.G. Kling, and C.H. Bock. 1997. Methods for screening maize against downy mildew *Peronosclerospora sorghi*. Plant Breeding 116:221-226.

Ceballos, H., J.A. Deutsch, and H. Guttierrez. 1991. Recurrent selection for resistance to *Exserohilum turcicum* in eight subtropical maize populations. Crop Science 31:964-971.

CIMMYT. 2009. CIMMYT Maize Inbred Lines (CML). CIMMYT: International Maize and Wheat Improvement Center. Mexico City, Mexico. http://www.cimmyt.org/en/component/content/article/459-international-maize-improvement-network-imin/434-cimmyt-maize-inbred-lines-cml (Accessed 9 Jul. 2011).

Collard, B.C.Y., and D.J. Mackill. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philosophical Transactions of the Royal Society of Biological Sciences 363: 557-572.

Craig, J., and R.A. Frederiksen. 1983. Differential sporulation of pathotypes of *Peronosclerospora sorghi* on inoculated sorghum. Plant Disease 67:278-279.

Danson J., M. Lagat, J. Ininda, and M. Kimani. 2006. Application of simple sequence repeats (SSRs) markers to study the resistance of locally adapted maize hybrids to damaging maize streak virus disease. African Journal of Biotechnology 5:1430-1434.

Denic, M. 2005. Maize breeding program: Breeding activities and national trials. National Agriculture Research Institute Annual Progress Report. Department of Agriculture and Farming Systems, Maputo, Mozambique.

Denic, M.P., C. Chauque, M. Jose, D. Langa, D. Mariote, P. Fato, and W. Haag. 2001. Maize screening for multiple stress tolerance and agronomic traits. Proceedings of the Eastern and Southern Africa Regional Maize Conference 7:88-91.

DeVries, J., and G. Toenniessen. 2001. Securing the Harvest: Biotechnology, Breeding and Seed Systems for African Crops. CABI, Wallingford, UK. pp. 208.

Dixon, A.G.O., R. Asiedu, and S.K. Hahn. 1991. Genotypic stability and adaptability: Analytical methods and implications for Cassava breeding for low-input agriculture. pp. 130-137. *In* Ofori F and Hahn SK (eds.) Tropical Root Crops in Developing Economy. Proceedings of the 9th Symposium of the International Society for Tropical Root Crops, 20-26. October. Accra, Ghana.

Efron, Y., S.K. Kim, J.M. Fajemisin, J.H. Mareck, C. Tang, Z.T Dabrowski, H.W. Rossel, G. Thottappilly, and I.W. Buddenhagen. 1989. Breeding for resistance to maize streak virus: A multidisciplinary team approach. Plant Breeding 103:1-36.

Eibach, R., E. Zyprian, L. Welter, and R. Topfer. 2007. The use of molecular markers for pyramiding resistance genes in grapevine breeding. Vitis 46:120-124.

Fajemisin, J.M. 2003. Overview of maize viruses in sub-Saharan Africa. pp. 158-171. *In* J.H. Hughes and J. Odu (eds.) Plant Virology in Sub-Saharan Africa. International Institute of Tropical Agriculture Conference, Ilesa, Nigeria.

Gowda, K.T.P., S. Rajashekaraiah, B.J. Gowda, and B.S. Naidu. 1987. Mandya maize hybrid, a boon for the farmers in downy mildew disease endemic areas. Current Research 16:168-169.

Graphpad Prism statistical package, version 5.04. 2010. http:download.cnet.com/windows/graphpad-software/3260-20_4-52733.html (Accessed 18 Jul. 2011).

Harkins, G.W., D.P. Martin, S. Duffy, A.L. Monjane, D.N. Shepherd, O.P. Windram, B.E. Owor, L. Donaldson, T.V. Antwerpen, R.A. Sayed, B. Flett, M. Ramusi, E.P. Rybicki, M. Peterschmitt, and A. Varsani. 2009. Dating the origin of the maize-adapted strain of maize streak virus, MSV-A. Journal of General Virology 90:3066-3074.

Hill, D.S. 2008. Pests of Crops in Warmer Climates and Their Control. Springer, the Netherlands.

Ininda, J., L. Gichunu, J.G.M. Njuguna, and P. Lorroki. 2006. Performance of three-way cross hybrids for Agronomic traits and resistance to maize streak virus in Kenya. African Journal of Crop Science 14:287-296.

Jeger, M.J., E. Gilijamse, C.H. Bock, and H.D. Frinking. 1998. The epidemiology, variability and control of the downy mildews of pearl millet and sorghum, with particular reference to Africa. Plant Pathology 47:544-569.

Jewell, D.C., S.R. Waddington, J.K. Ransom and K.V. Pixley (eds.). 1995. Maize Research for Stress Environments. Proceedings of the Fourth Eastern and Southern Africa Regional Maize Conference, held at Harare, Zimbabwe, 28 March - 1 April 1994. CIMMYT, Mexico D.E. pp. 306.

Joshi, R.K., and S. Nayak. 2010. Gene pyramiding - A broad spectrum technique for developing durable stress resistance in crops. Biotechnology and Molecular Biology Review 5:51-60.

Karanja, J., N.O. Amugune, J. Ininda, J.N. Kimatu, and J.W. Danson. 2009. Microsatellite analysis of the correlation between molecular and morphological traits in assorted maize inbred lines. African Crop Science Journal 17:133-144.

KARI. 1995. Kenya Agricultural Research Institute: *In*: K. Njoroge and S.N. Mugo (eds.) Maize National Performance Trial Handbook. Kenya Agricultural Research Institute, Nairobi, Kenya. pp. 88.

Kim, S.K., Y. Efron, J.M. Fajemisin, and I.W. Buddenhagen. 1989. Mode of gene action for resistance in maize to maize streak virus. Crop Science 29:890-894.

Langridge, P., E. Lagudah, T. Holton, R. Appels, P. Sharp, and K. Chalmers. 2001. Trends in genetic and genome analyses in wheat: a review. Australian Journal of Agricultural Research 52: 1043-1077.

Magenya, O.E.V., J. Mueke, and C. Omwega. 2008. Significance and transmission of maize streak virus disease in Africa and options for management: A review. African Journal of Biotechnology 7:4897-4910.

Martin, D.P., and D.N. Shepherd. 2009. The epidemiology, economic impact and control of maize streak disease. Food Security 1:305-315.

Mawere, S., V. Vincent, J. De Meyer, and K.V. Pixley. 2006. Resistance of four inbred maize lines to inoculation with 20 isolates of maize streak virus from Zimbabwe. Plant Disease 90:1485-1489.

McGrath, J.M., T.K. Koppin, and T.M. Duckert. 2005. Breeding for genetics: Development of recombinant inbred lines (RILs) for gene discovery and deployment. USDA-ARS 1:124-132.

Moreau, L., A. Charcosset, and A. Gallais. 2004. Experimental evaluation of several cycles of marker-assisted selection in maize. Euphytica 137, 111–118.

Ngwira, P., and P.T. Khonje. 2002. Managing maize diseases through breeding under Malawi field conditions. Malawi Journal of Agricultural Sciences 1:16-22.

Paliwal, R.L. 2000. Hybrid maize breeding. *In*: R.L. Paliwal., G. Granados, H.R. Lafitte, A.D. Violic, and J.P. Marethee (eds.) Tropical Maize: Improvement and Production. Food and Agriculture Organisation of the United Nations, Rome, Italy.

Panicker, S., and K. Gangadharan. 1998. Controlling downy mildew of maize caused by *Peronosclerospora sorghi* by foliar sprays of phosphonic acid compounds. Crop protection 18:115-118.

Pratt, R., S. Gordon, P. Lipps, G. Asea, G. Bigirwa, and K. Pixley. 2003. Use of IPM in the control of multiple diseases in maize: Strategies for selection of host resistance. African Crop Science Journal 11:189-198.

Ribaut, J.M., and D.A. Hoisington. 1998. Marker assisted selection: new tools and strategies. Trends in Plant Science 3: 236-239.

Rodier, A., J. Assie, J.L. Marchand, and Y. Herve. 1995. Breeding maize lines for complete and partial resistance to maize steak virus (MSV). Euphytica 81:57-70.

Rosyara, U.R. 2006. Requirements of robust molecular marker technology for plant breeding applications. Journal of Plant Breeding 1: 67 - 72.

Sriwatanapongse, S., S. Jinahyon, and S.K. Vasal. 1993. Suwan-l: Maize/rom Thailand to the World. Mexico, D.F.: CIMMYT.

Thormann, C.E., M.E. Ferreira, L.E.A. Camargo, J.G. Tiranga, and T.C. Osborn.1994. Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. Theoretical and Applied Genetics 88:973-980.

Thottappilly, G., N.A. Bosque-Perez, and I.I.W. Rossel. 1993. Viruses and virus diseases of maize in tropical Africa. Plant Pathology 42:494-509.

Vivek, B., M. Banziger, and K.V. Pixley. 2005. Characterisation of maize germplasm grown in Eastern and Southern Africa: Results of the 2004 regional trials coordinated by CIMMYT. Harare, Zimbabwe. pp. 68.

Welz, H.G., A. Schechert, A. Pernet, K.V. Pixley, and H.H. Geiger. 1998. A gene for resistance to the maize streak virus in the African CIMMYT maize inbred line CML202. Molecular Breeding 4:147-154.

Wulffe, E.G., and J. Torp. 2005. Overview of Seed Supply Systems and Seed Health Issues. Seed Sector Country Profile: Mozambique. The Royal Veterinary and Agricultural University, Frederiksberg. pp. 8.

CHAPTER FIVE

Genetic diversity among maize inbred lines selected for downy mildew and maize streak virus resistance as determined by SNP markers

Abstract

The maize inbred line LP23 is an important parent of hybrids in Mozambique due to its high productivity and its resistance to downy mildew (DM) disease. Its productivity, however, is compromised by its susceptibility to the maize streak virus (MSV) disease. F_4 progeny lines derived from crosses LP23 x CML505 and LP23 x CML509 identified as possessing resistance to both MSV and DM were subjected to genotypic characterisation using 400 publicly available single nucleotide polymorphism (SNP) markers. This was to identify lines that could be effectively used to create new hybrids in lieu of LP23. The SNP genotyping was performed using a Sequenom MassArray genotyping platform. The genetic similarity of 50 maize lines was estimated using the Jaccard coefficient and similarity coefficients ranged from 52.45% to 87.52%. The progeny lines with LP23 as a common parent were placed in two clusters, with those F_4 lines derived from cross CML509 x LP23 clustering with LP23 and those from cross CML505 x LP23 clustering with CML505. The standard lines and CML509 were fitted into different clusters. The potential tester lines displayed large genetic distances from the new progeny lines, which has implications for breeding MSV resistant hybrids. There were also a few progeny lines with ≥85% similarity with the common parent LP23, qualifying them as suitable candidates for use in an accelerated backcross programme to recover the full genome of LP23 without having to conduct six generations of backcrossing.

Keywords: dendrogram, downy mildew (DM), genetic similarity, maize genetic diversity, maize streak virus (MSV), single nucleotide polymorphisms (SNPs).

5.1 Introduction

The characterisation of genetic diversity and relatedness within and between elite breeding populations is important for the selection of lines with the desired traits for use as breeding parents (Marti *et al*., 2012). Characterisation has moved from the conventional use of morphological features to molecular markers that examine diversity directly at the DNA sequence level (Ajmone-Marsan *et al*., 1992; Hagdorn *et al*., 2003; Diniz *et al*., 2005). Molecular markers have an advantage as cultivar descriptors because they are unaffected by environmental factors (Akter *et al*., 2008).

Molecular markers used in marker-assisted breeding programmes for the analysis of genetic diversity in plants include the hybridisation-based marker known as restriction fragment length polymorphisms (RFLPs) and PCR-based molecular markers: random amplified polymorphic DNA (RAPDs), simple sequence repeats (SSRs) and amplified fragment length polymorphisms (AFLPs) (Farooq and Azam, 2002; Akter *et al*., 2008; Yan *et al*., 2009a). Simple sequence repeats are highly polymorphic and informative and thus are one of the most extensively used DNA marker types in maize for characterisation of germplasm collections (Warburton *et al*., 2001; Vigouroux *et al*., 2005; Kostova *et al*., 2006). Maize is one of the most diverse crop species and compared to most crops exhibits a wider range of variation in morphological traits and molecular dynamism (Smith and Smith, 1989; Zhao *et al*., 2006; Karanja *et al*., 2009).

The high level of polymorphism in maize facilitates single nucleotide polymorphism (SNP) identification (Rafalski, 2002a). With advances in technology there has been a move toward use of SNPs (Hamblin *et al*., 2007), which are highly abundant, with an estimate of over 20 million polymorphisms available in maize (Ching *et al*., 2002; Rafalski, 2002b) and have the potential for high throughput low cost genotyping (Yan *et al*., 2009b). Simple sequence repeats are neither as abundant nor widely distributed as SNPs, making them unsuitable for association studies (Ching *et al*., 2002; Agarwal *et al*., 2008). Furthermore, they are expensive to evaluate, and their analysis is difficult to scale up (Yan *et al*., 2009b). The high density of SNPs makes them suitable for high-throughput methods, such as genotyping arrays, at lower error rates (Ching *et al*., 2002; Rafalski, 2002b; Kennedy *et al*., 2003). Genome wide scans with SNPs are more cost-effective and have since replaced the use of microsatellites (Abecasis and Wigginton, 2005; Yan *et al*., 2009b) in some crops.

Next-generation sequencing (NGS) technologies make use of SNPs for high throughput in marker-assisted breeding allowing for accelerated genetic analysis of traits and the development of high-density linkage maps for map-based gene discovery (Hakim *et al*., 2010; Trick *et al*., 2012). Several techniques have been developed to assay SNPs, including the Sequenom MassARRAY iPLEX genotyping platform (Gabriel *et al*., 2009). In this approach a short section of DNA containing a SNP is amplified by PCR and then a high-fidelity single-base primer extension is performed using nucleotides of modified mass. The different alleles therefore produce oligonucleotides with mass differences that can be detected using highly accurate Matrix-Assisted Laser Desorption/Ionisation Time Of-Flight (MALDI-TOF) mass spectrometry (Gabriel *et al*., 2009; Meyer and Ueland, 2011). In the present study SNP genotyping was performed using the Sequenom MassArray genotyping platform (Sequenom, San Diego, USA), following the standard protocols provided by Gabriel *et al*. (2009).

The objectives of the study were to assess genetic diversity and distance among the LP23 derived progeny lines and other standard germplasm lines which can be used as testers to develop new hybrids and to investigate potential utilisation of the LP23 x CML505 and LP23 x CML509 derived maize lines in a maize improvement programme. This was to categorise the lines into cluster groups based on their molecular profiles, as a first step toward creating viable heterotic groups, and identify the lines that were most closely related to the elite Mozambican LP23 parent line, but showing resistance not only to DM but to MSV disease as well. These lines could be used to make hybrids in lieu of the LP23, and will be subjected to further improvement through a backcross programme.

5.2 Materials and methods

5.2.1 Plant material

A total of 50 maize inbred lines were used in this study. Thirty-five of these were $F_{4:5}$ lines derived from F_3 progeny of the crosses CML505 x LP23 and CML509 x LP23 and were selected because they exhibited resistance to both DM and MSV diseases and were designated as “DMSR” lines. The pedigrees of the lines are presented in Table 5.1. Breeding of these lines was done using pedigree selection and marker-assisted selection (MAS) and is

described in Chapter 4. The parental lines LP23 (MSV susceptible), CML505 and CML509 (MSV resistant) together with 12 standard public lines which are parents of commonly grown hybrids in Southern Africa and designated as follows: B73WX, CML202, DXL37, DXL59, H24W, I137TN, LP19, M017WX, M162W, PA1, PA2 and 8CED67 were included in the current study. The 12 lines are potential testers for making crosses with the progeny lines. Lines Mo17WX and B73WX are derivatives of members of important heterotic groups, A and B, which are used in the USA (Gethi *et al*., 2002; Lai *et al*., 2010), while the rest represent important heterotic groups used in Southern Africa. The lines I137TN and M162W represent a sample of widely used germplasm lines in South Africa. The CML202 was developed at CIMMYT under mid-altitude environment in Zimbabwe and is widely used in Africa (Welz *et al*., 1998; Warburton *et al*., 2002). The lines PA1 and PA2 were also developed in Zimbabwe under mid altitude conditions. Lines DXL37, DXL59, H24W and 8CED67 represent a sample of advanced inbred lines from the programme at the University of KwaZulu-Natal. Line LP19 is a Mozambican line that is adapted to the lowland environment. These control lines were used in the study as standards for the heterotic group classifications, which would form inference about the use of new progeny lines in the hybrid programme in Mozambique.

Table 5.1: Name, pedigree and origin data of the 35 F_4 maize lines

*Designated Name	Pedigree**	Origin
DMSR-1	CML505/LP23-F2B-1-1	CERU-10CR1-1-1
DMSR-2	CML505/LP23-F2B-2-1	CERU-10CR1-2-1
DMSR-4	CML505/LP23-F2B-3-1	CERU-10CR1-3-1
DMSR-8	CML505/LP23-F2B-6-1	CERU-10CR1-6-1
DMSR-10	CML505/LP23-F2B-10-1	CERU-10CR1-10-1
DMSR-12	CML505/LP23-F2B-11-1	CERU-10CR1-11-1
DMSR-13	CML505/LP23-F2B-12-1	CERU-10CR1-12-1
DMSR-16	CML505/LP23-F2B-13-1	CERU-10CR1-13-1
DMSR-18	CML505/LP23-F2B-15-1	CERU-10CR1-15-1
DMSR-21	CML505/LP23-F2B-16-1	CERU-10CR1-16-1
DMSR-23	CML505/LP23-F2B-17-1	CERU-10CR1-17-1

*Designated Name	Pedigree**	Origin
DMSR-26	CML505/LP23-F2B-18-1	CERU-10CR1-18-1
DMSR-30	CML505/LP23-F2B-21-1	CERU-10CR1-21-1
DMSR-34	CML505/LP23-F2B-22-1	CERU-10CR1-22-1
DMSR-35	CML505/LP23-F2B-25-1	CERU-10CR1-25-1
DMSR-39	CML509/LP23-F2B-16-1	CERU-10CR1-44-1
DMSR-40	CML509/LP23-F2B-25-1	CERU-10CR1-53-1
DMSR-43	CML509/LP23-F2B-27-1	CERU-10CR1-55-1
DMSR-46	CML509/LP23-F2B-28-1	CERU-10CR1-56-1
DMSR-47	CML509/LP23-F2B-29-1	CERU-10CR1-57-1
DMSR-51	CML509/LP23-F2B-30-1	CERU-10CR1-58-1
DMSR-55	CML509/LP23-F2B-37-1	CERU-10CR1-65-1
DMSR-56	CML509/LP23-F2B-41-1	CERU-10CR1-69-1
DMSR-57	CML509/LP23-F2B-57-1	CERU-10CR1-85-1
DMSR-60	CML509/LP23-F2B-61-1	CERU-10CR1-89-1
DMSR-62	CML509/LP23-F2B-64-1	CERU-10CR1-92-1
DMSR-64	CML509/LP23-F2B-65-1	CERU-10CR1-93-1
DMSR-65	CML509/LP23-F2B-66-1	CERU-10CR1-94-1
DMSR-66	CML509/LP23-F2B-67-1	CERU-10CR1-95-1
DMSR-69	CML509/LP23-F2B-74-1	CERU-10CR1-102-1
DMSR-71	CML509/LP23-F2B-75-1	CERU-10CR1-103-1
DMSR-73	CML509/LP23-F2B-83-1	CERU-10CR1-111-1
DMSR-74	CML509/LP23-F2B-84-1	CERU-10CR1-112-1
DMSR-75	CML509/LP23-F2B-85-1	CERU-10CR1-113-1
DMSR-77	CML509/LP23-F2B-86-1	CERU-10CR1-114-1

*DMSR, downy mildew and MSV resistant lines.

**Pedigrees of the lines are the same for the F_3 generation lines in Chapter 4

5.2.2 DNA Extraction

The 50 lines were planted in the tunnel at the University of KwaZulu-Natal, Pietermaritzburg. At the four leaf stage, sampling from individual plants using scissors that was sterilised between samples using 100% ethanol. DNA extraction was done using a Sarkosyl based method (Hasan *et al*., 2008) at the DNA LandMarks laboratory in Quebec, Canada. The leaf material was ground into a fine powder in liquid nitrogen after which 3 ml of DNA extraction buffer (100 mM Tris-HCl at pH 8.0, containing 0.35 M sorbitol, 5 mM EDTA at pH 8.0, and 1% 2-mercaptoethanol) and 1 ml of phenol was added in a test tube and homogenised. Another 2 ml of phenol was added and centrifuged at 12 000 rpm for 5 min. The supernatant was transferred into an equal volume of 200 µl of ice-cold 95% ethanol and centrifuged at 12 000 rpm for 5 min to precipitate the DNA. The precipitated DNA was washed in 70% ethanol, dissolved in 0.5 ml of Tris EDTA (TE) with 2µg of RNAase and incubated at 37^{o}C for 30 min. Then further 0.25 ml phenol and 0.25 ml chloroform was added and centrifuged and the upper phase was transferred into a fresh tube with an equal volume of 95% ethanol for DNA precipitation. The precipitated DNA was again washed with ice-cold 70% ethanol and dissolved in 0.2 ml of TE.

5.2.3 SNP selection and amplification

The 50 maize leaf samples were genotyped using 400 SNPs on the MassARRAY platform of Sequenom at the BASF Plant Science Centre, DNA LandMarks, 84 Rue Richelieu, Quebec, Canada. The SNPs that were selected are Panzea (public SNPs) and were chosen based on their even distribution on the 10 maize chromosomes (~40 SNPs/chromosome). The SNP mapping positions on the chromosomes were based on the nested association mapping (NAM) populations involving the maize inbred line B73 as a common parent. The SNP genotyping was performed using the Sequenom MassArray genotyping platform following standard protocols. The primer information for Sequenom genotyping is presented in Appendix 1. PCR reaction mixes were prepared for each sample containing 5 × PCR Buffer, 2.5 mM dNTPs, 25 mM MgC_{12}, 10 mM of each primer, 5 U of HotStar*Taq* DNA polymerase and 25 ng μl^{-1} of genomic DNA. Cycling parameters were as follows: 94°C for 5 min. followed by 45 cycles of 94°C for 20 s, annealing step for 30 s at 56°C , 72°C for 1 min and a polishing step of 72°C for 3 min (Gabriel *et al*., 2009). The DNA quality was evaluated

carefully before genotyping by screening each sample on a 0.8% (w/v) agarose gel. Once the DNA quality passed the quality control, the DNA samples were used for SNP genotyping by a commercially available Sequenom MassARRAY platform following the standard protocols described by Gabriel *et al*. (2009) at DNA Land Marks Inc., Quebec, Canada. The protocol for this assay recommended using 2.5 ng μl^{-1} DNA per sample.

5.2.4 Statistical analysis

SNP data was scored on the basis of presence or absence of marker alleles and this data was used to estimate the genetic similarity (GS) between any pair of lines based on the Jaccard coefficient using the NTSYSpc v2.1 software package (Exeter Software Setauket, NY, USA).The dendrogram showing the genetic relatedness among the lines was constructed using the Unweighted paired group method using arithmetic averages (UPGMA) method.

For each SNP, number of alleles, allele frequency, number of genotypes, genotype frequency, observed heterozygosity, gene diversity, and polymorphic information content (PIC) were computed using PowerMarker version 3.25. Observed heterozygosity was calculated by dividing the number of heterozygous individuals by the number of individuals scored. Polymorphism information content (PIC) for the SSR markers in the sample DNA was calculated as:

PIC = $1- \Sigma p_i^2$ where p_i is the frequency of the i^{th} allele in a locus for individual p.

5.3 Results

5.3.1 Genetic diversity levels

For the 400 SNP markers, a total of 752 alleles with an average of 1.88 alleles per locus were observed. The PIC value ranged from 0.00 to 0.389 with a mean of 0.1807, gene diversity mean was 0.2217 and that for heterozygosity was 0.1052. Table 5.2 summarizes the range and mean values for PIC, heterozygosity, gene diversity and number of alleles per locus with the 400 SNPs. The individual values for all 400 SNP markers including the major allele frequencies are presented in Appendix 1.

Table 5.2: Average (minimum-maximum) of polymorphism for all lines assayed with 400 SNPs.

Parameter	Mean and Range
No. of alleles per locus	1.88 (1-2)
PIC	0.18 (0.00-0.389)
Gene diversity	0.22 (0.00-0.509)
Heterozygosity	0.11 (0.00-1.00)

The SNP call rate was 98.3%. Estimates of genetic similarities based on the SNP markers among the 50 maize lines are presented in Appendix 2. The genetic similarity coefficients among the lines ranged from 52.45% to 87.52%. The lowest similarity value of 52.45% was between the standard line M162W and the common parent for progeny lines, LP23. Table 5.2 gives the genetic similarity coefficients for each of the 35 DMSR lines against the three parental lines (LP23, CML505 and CML509) used in the breeding programme.

The similarity percentages among all the DMSR progeny lines ranged between 71% and 87% (Table 5.3). Wider genetic distances were observed between all the DMSR progeny lines and the MSV resistant donor parent CML505, as similarity of the lines to CML505 ranged from 60% to 83%. Similarity values of the DMSR progenies with the LP23 and CML509 ranged from 71% to 86% and 61% to 68% respectively. On the other hand, the DMSR lines with CML509 background are 60% to 65% similar to CML505 and DMSR lines with the CML505 background also displayed almost the same level of similarity (61-68%) with parent CML509. It is clearly shown that the progeny lines were more distantly related with the CML509 than with the other parental lines.

The highest similarity of 88% was observed between DMSR55 and DMSR47 (Appendix 2). The similarity values of lines DMSR47 and DMSR55 are 0.80 and 0.81 against LP23, 0.60 for both against CML505 and 0.65 and 0.66 against CML509, respectively. The progeny line DMSR69 had the highest similarity coefficient of 86% with LP23, DMSR21 with CML505 at 83% and DMSR23 with CML509 at 68%. All progeny had ≥70% similarity percentage to LP23. Overall results indicate that the DMSR lines were more closely related to the elite LP23 Mozambican line than their respective CML parents.

Table 5.3: Similarity percentage index of the progeny lines (DMSR lines) against the parental lines (LP23, CML505 and CML509)

Progeny lines	LP23 (common parent)	CML505 (MSV donor)	CML509 (MSV donor)
DMSR1	0.83	0.73	0.64
DMSR2	0.73	0.81	0.61
DMSR4	0.75	0.78	0.65
DMSR8	0.80	0.75	0.64
DMSR10	0.77	0.77	0.62
DMSR12	0.71	0.82	0.64
DMSR13	0.72	0.80	0.62
DMSR16	0.78	0.75	0.62
DMSR18	0.82	0.74	0.65
DMSR21	0.73	0.83	0.62
DMSR23	0.84	0.70	0.68
DMSR26	0.83	0.70	0.62
DMSR30	0.80	0.73	0.63
DMSR34	0.82	0.71	0.62
DMSR35	0.75	0.78	0.65
DMSR39	0.81	0.62	0.64
DMSR40	0.75	0.64	0.63
DMSR43	0.84	0.63	0.64
DMSR46	0.82	0.61	0.64
DMSR47	0.80	0.60	0.65
DMSR51	0.83	0.61	0.64
DMSR55	0.81	0.60	0.66
DMSR56	0.80	0.61	0.63
DMSR57	0.73	0.62	0.67
DMSR60	0.81	0.61	0.64
DMSR62	0.80	0.65	0.61
DMSR64	0.79	0.61	0.65
DMSR65	0.80	0.60	0.64
DMSR66	0.80	0.63	0.64
DMSR69	0.86	0.61	0.61
DMSR71	0.85	0.61	0.62
DMSR73	0.79	0.64	0.66
DMSR74	0.79	0.60	0.63
DMSR75	0.71	0.64	0.66
DMSR77	0.77	0.61	0.64

5.3.2 Patterns of genetic diversity

A dendrogram (Fig. 5.1) was generated to further assess the genetic diversity of the maize inbred lines. The dendrogram analysed all 50 maize lines, 35 being DMSR lines, three parental lines and the 12 maize control lines. The dendrogram based on UPGMA cluster analysis of genetic similarities showed that all maize lines were grouped into 13 major clusters at 72% similarity coefficient (Fig. 5.1).

Most of the clusters were consistent with the origin and the pedigree information of the inbred lines. For example, Cluster 6 contained all the CML505/LP23 progeny lines and the parental line CML505. The LP23 parental line was in Cluster 7 with all the CML509/LP23 progeny lines. The parental line CML509 was not placed in the same cluster with any of its progenies, but was in its own cluster 8. At a 60% similarity coefficient the temperate (B73WX and MO17WX derivatives) from the USA and subtropical control lines (DXL37, 8CED67 and PA2) from South Africa (Lai *et al*., 2010) clustered together. There were four clusters. B73WX to PA2 – Cluster 1; CML 202 to PA1 – Cluster 2; DXL59 and M162W – Cluster 3; I137TN – Cluster 4. The controls M162W and DXL59 at a 65% similarity coefficient were grouped in one cluster and so were lines LP19 and CML509. I137TN and PA1 stood each in their own cluster each. Overall, the clusters corresponded to pedigree breeding groups. Line DMSR55 and DMSR47 were the most genetically related with an 88% similarity value which was also confirmed in the matrix (Appendix 2) as 0.88.

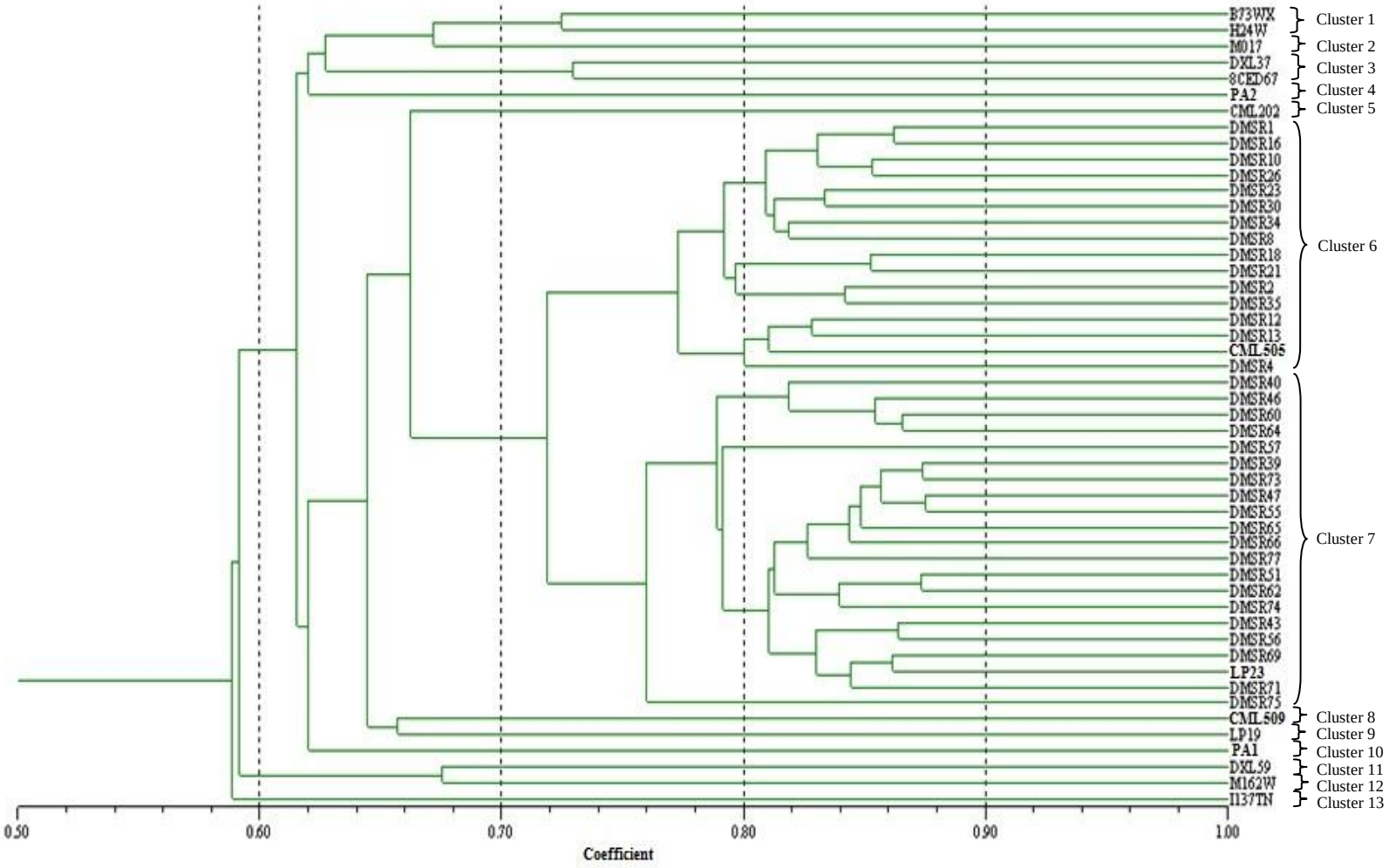

Figure 5.1: UPGMA dendrogram deciphering the genetic relatedness of maize streak virus and downy mildew resistant progeny lines based on Jaccard distances calculated using 400 SNP markers.

5.4 Discussion

The levels of similarity among the 35 progeny lines varied between 71% and 83%, indicating that there was ample genetic variation for further selection to obtain new productive inbred lines.

In this study, SNP markers were successfully used to genotype inbred maize germplasm. Higher levels of genetic diversity existed between the CML505 x LP23 and CML509 x LP23 progeny lines than within the populations which are expected as lines within a population have a similar gene frequency. DMSR lines clustered together according to their pedigree with all the CML505 x LP23 and CML509 x LP23 progeny in clusters 6 and 7, respectively. This implies that productive hybrids can be developed by crossing DMSR lines from cluster 6 with those from cluster 7. The LP23 parent was in cluster 7 mostly with the lines that were derived from the combination CML509 x LP23 implying that these lines had a higher percentage of the LP23 genome than CML505 x LP23 progeny lines. The MSV donor parent CML509 was placed in a different cluster from its progenies indicating that it may be used to make productive hybrids in combination with its progenies and obviously with lines derived from the CML505 x LP23 population.

The new DMSR lines were divergent from the regional testers (CML202, 1137TN, LP19, DXL59, 8CED67, PA1 and PA2) indicating that they can be crossed with these testers to make hybrids. In addition, they were also divergent from the widely used temperate testers such as MO17WX and B73WX. This indicated that there exists a high utility for these lines. The genetic difference between the lines generally ranged between 12% and 48%. However, based on the analysis of check inbred testers, we found that lines with 60-70% similarity combined well to produce excellent hybrids. For example, MO17WX x B73 is a great hybrid combination in USA history (Glover *et al*., 2005; Nelson, 2010) whereas PA1 x PA2 is also an excellent hybrid combination used commercially in Southern Africa. In addition, the two MSV resistance donor inbred lines CML505 and CML509 when crossed, produced high levels of heterosis, although they have 63% similarity. Therefore, lines observed in this study can be exploited in combination with existing inbred testers to make new hybrids with potential to form excellent heterotic patterns.

With respect to the control tester lines, the inbred testers from Zimbabwe PA1 and PA2 were in different clusters, which was consistent with breeders information ([1]Fato, 2011; pers. comm.). These two lines form good heterotic patterns. There were high levels of variation between the control lines however, with the lines in Clusters 1 to 5, consisting of temperate lines B73, MO17 and South African lines (DXL37 and 8CED67; H24W), which were derived from temperate and subtropical material. Clusters 8 to 13 consisted of singletons of lowland (CML509 and LP19) and subtropical adapted lines (PA1, M167W) from Southern Africa. This suggests that management of the maize programme can be simplified by working with four heterotic groups which can be designated as A for Clusters 1 to 5, B for Cluster 6, C for Cluster 7 and D by pooling lines in Clusters 8 to 13.

The highest genetic similarity was 87.52% between the new progeny lines DSMR47 and DMSR55 suggesting that heterosis between the two lines would be minimum; depending on other economic factors breeders can select one of the lines for use in hybrid breeding. Lines DMSR69 and LP23 had an 86% similarity percentage. Line LP23 was the elite line from Mozambique, which is susceptible to MSV disease and required improvement for resistance to the disease. The new line DMSR69, which was 86% similar to LP23, expressed resistance to both DM and MSV, and thus qualified as an adequate replacement for LP23 as an inbred parent for Mozambican hybrid breeding programmes. However a backcross breeding programme would be recommended to recover more of the LP23 genome so that adaptation to the lowland is enhanced.

The genetic distances between the lines ranged from 12% to 48% (Table 5.3) and this can be explained by the fact that only 400 SNPs were used in the current study. A greater number of markers would be required to obtain greater discrimination between the lines. Only a few SNP markers were publicly available at the time of designing the experiment. However, the use of many markers also has implications on the costs of conducting such an experiment.

---

[1]Dr Pedro Fato, Chokwe Research Station, Institute of Agriculture of Mozambique (IIAM)

Given the trend towards increased use of SNP markers, Rafalski (2002b) concluded that in the future, SNP assays would become inexpensive, and there would be a greater demand for SNPs, with the result that other marker technologies such as SSRs will co-exist with SNP analysis. The small distance between the standard lines is explained by the fact that breeders were selecting in the same environment, hence they tended to select lines with similar gene frequencies. Vigouroux *et al*. (2005) identified small genetic distance values in a study using SSR markers as being the result of a 'domestication bottleneck effect', which is the result of the breeder using a limited genetic pool of wild founder plants. The observation that there was genetic variation between the new progeny line families indicates that there is an opportunity to select the best lines for productivity in hybrids.

5.5 Conclusions

From this study the following conclusions were drawn:

1) Four hundred SNP markers were successful in discriminating between 50 maize inbred lines, according to genetic distances, and which resulted in four clusters that are consistent with pedigree information and the origin of the parent lines.
2) The study identified two new progeny lines (DMSR69 and DMSR71) which are at least 85% similar to LP23 (the candidate for further improvement) and which were resistant to both MSV and DM. These lines will be used as parents in lieu of LP23 to improve the hybrids in Mozambique for MSV resistance, with positive implications for yield and food security.
3) Genetic variation among the new progeny lines was revealed which indicates potential for further selection to come up with the best lines for use in developing new hybrids.
4) Observation of different genetic clusters for the new progeny lines and the standard testers can form the basis for allocating the lines into different heterotic groups with positive implication for the hybrid maize breeding programme in Mozambique.
5) It is recommended that the actual crosses should be formed between lines in different clusters, and within clusters, and that they should be evaluated in

different environments in Mozambique to confirm the levels of heterosis among the lines.

References

.Abecasis, G.R., and J.E. Wigginton. 2005. Handling marker-marker linkage disequilibrium: Pedigree analysis with clustered markers. The American Journal of Human Genetics 77:754-767.

Agarwal, M., N. Shrivastava, and H. Padh. 2008. Advance in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27:617-631.

Ajmone-Marsan, P., C. Livini, M.M. Messmer, A.E .Melchinger, and M. Motto. 1992. Cluster analysis of RFLP data from related maize inbred lines of the BSSS and LSC heterotic groups and comparison with pedigree data. Euphytica 60:139-148.

Akter, J., M.S. Islam, A.A. Sajib, N. Ashraf, S. Haque, and H. Khan. 2008. Microsatellite markers for determining genetic identities and genetic diversity among jute cultivars. Australian Journal of Crop Science 1:97-107.

Ching, A., K.S. Caldwell, M. Jung, M. Dolan, O.S. Smith, S. Tingey, M. Morgante, and A.J. Rafalski. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. Genetics 3:19-26.

Diniz, L.E.C., C.F. Ruas, V.P. Carvalho, F.M. Torres, E.A. Ruas, M.O. Santos, T. Sera, and P.M. Ruas. 2005. Genetic diversity among forty coffee varieties assessed by RAPD markers associated with restriction digestion. Brazilian Archives of Biology and Technology 48:511-521.

Farooq, S., and F. Azam. 2002. Molecular markers in plant breeding: Concepts and characterisation. Pakistan Journal of Biological Sciences 10:1135-1140.

Gabriel, S., L. Ziaugra, and D. Tabbaa. 2009. SNP genotyping using the Sequenom MassARRAY iPLEX Platform. Current Protocols in Human Genetics 60:1-18.

Gethi, J.G., J.A. Labate, K.R. Lamkey, M.E., Smith, and S. Kresovich. 2002. SSR variation in important U.S. maize inbred lines. Crop Science 42:951-957.

Glover, M.A., D.B. Willmot, L.L. Darrah, B.E. Hibbard, and X. Zhu. 2005. Diallel analyses of agronomic traits using Chinese and U.S maize germplasm. Crop Science 6:1192-1102.

Hagdorn, S., K.R. Lamkey, M. Frisch, P.E.O. Guimara, and A.E. Melchinger. 2003. Molecular genetic diversity among progenitors and derived elite lines of BSSS and BSCB1 maize populations. Crop Science 43:474-482.

Hakim, I.R., N.G. Kammoun, E. Makhloufi, and A. Rebai. 2010. Discovery and potential of SNP markers in characterisation of Tunisian olive germplasm. Diversity 2:17-27.

Hamblin, M.T., M.L. Warburton, and E.S. Buckler. 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. PLoS ONE 12:1367.

Hasan, S.M.Z., M.S.B. Shafie, and R.M. Shah. 2008. Efficient method for the extraction of genomic DNA from wormwood (*Artemisia capillaris*). African Journal of Biotechnology 7:3211-3216.

Karanja, J., N.O. Amugune, J. Ininda, J.N. Kimatu, and J.W. Danson. 2009. Microsatellite analysis of the correlation between molecular and morphological traits in assorted maize inbred lines. African Crop Science Journal 17:133-144.

Kennedy, G.C., H. Matsuzaki, S. Dong, W.M. Liu, and J. Huang. 2003. Large scale genotyping of complex DNA. Nature Biotechnology 21:1233-1237.

Kostova, A., E. Todorovska, N. Christov, V. Sevov, and A.I. Atanassov. 2006. Molecular characterisation of Bulgarian maize germplasm collection via SSR markers. Biotechnology and Biotechnology Equipment 20:29-36.

Lai, J., R. Li, X. Xu, W. Jin., M. Xu, H. Zhao, Z. Xiang, W. Song, K. Ying, M. Zhang, Y. Jiao, P. Ni, J. Zhang, D. Li, X. Guo, K. Ye, M. Jian, B. Wang, H. Zheng, H. Liang, X. Zhang, S. Wang, S. Chen, J. Li, Y. Fu, N.M. Springer, H. Yang, J. Wang, J. Dai, P.S. Schnable, and J. Wang. 2010. Genome-wide patterns of genetic variation among elite maize inbred lines. Nature Genetics 42:1027-1030.

Marti, A.F., B. Athanson, T. Koepke, C. F. Forcadas, A. Dhingra, and N. Oraguzie. 2012. Genetic diversity and relatedness of sweet cherry (*Prunus avium* L.) cultivars based on single nucleotide polymorphic markers. Plant Science 3:1-7.

Meyer, K., and P.M. Ueland. 2011. Use of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for multiplex genotyping. Advances in Clinical Chemistry 53:1-29.

Nelson, P.T. 2010. Evaluation of Elite Exotic Maize Inbreds for Use in Long-term Temperate Breeding. MSc thesis, University of North Carolina, United States of America.

Rafalski, A. 2002a. Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5:94-100.

Rafalski, A.J. 2002b. Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Science 162:329-333.

Smith J.S.C., and O.S. Smith. 1989. Comparison of heterosis among hybrids as a measure of relatedness with that to be expected on the basis of pedigree. Maize Genetics 63:86-87.

Trick, M., N.M. Adamski, S.G. Mugford, C. Jiang, M. Febrer, and C. Uauy. 2012. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploidy wheat. Plant Biology 12:14-31.

Vigouroux, Y., S. Mitchell, Y. Matsuoka, M. Hamblin, S. Kresovidi, J.S.C. Smith, J. Jaqueth, O.S. Smith, and J. Doebley. 2005. An analysis of genetic diversity across the maize genome using microsatellites genetics. Genetics Society of America 169:1617-1630.

Warburton, M., X. Xianchun, S. Ambriz, L. Diaz, E. Villordo, and D. Hoisington. 2001. Use of molecular markers in maize diversity studies at CIMMYT. Proceedings of the Eastern and Southern Africa Regional Maize Conference 7:130-133.

Warburton, M.L., X. Xianchun, S. Ambriz, L. Diaz, S.K. Vasal, and F. Gonzalez. 2002. Genetic relationships among the CIMMYT maize lines developed for Asia and from other CIMMYT breeding programs. Proceedings of the 8th Asian Regional Maize Workshop, Bangkok, Thailand. http://apps.cimmyt.org/english/docs/proceedings/armw/armw.pdf (Accessed 9 Dec. 2011).

Welz, H.G., A. Schechert, A. Pernet, K.V. Pixley, and H.H. Geiger. 1998. A gene for resistance to the maize streak virus in the African CIMMYT maize inbred line CML202. Molecular Breeding 4:147-154.

Yan, J., T. Shah, M.L. Warburton, E.S. Buckler, M.D. McMullen, and J. Crouch. 2009a. Genetic characterisation and linkage disequilibrium estimation of a global maize collection using SNP markers. PLoS One 24:8451.

Yan J., X. Yang, T. Shah, H. Sa´nchez-Villeda, J. Li, M. Warburton, Y. Zhou, J.H. Crouch, and Y. Xu. 2009b. High-throughput SNP genotyping with the Golden Gate assay in maize. Molecular Breeding 25:441-451.

Zhao, W., P. Canaran, R. Jurkuta, T. Fulton, J. Glaubitz, E. Buckler, J. Doebley, B. Gaut, M. Goodman, J. Holland, S. Kresovich, M. McMullen, L. Stein, and D. Ware. 2006. Panzea: A database and resource for molecular and functional diversity in the maize genome. Nucleic Acids Research 34:D752-D757.

OVERVIEW OF THE RESEARCH FINDINGS

Introduction

This chapter summarises the main objectives of the study, highlights the major findings and the implications and recommends the direction for future research.

The specific objectives of the study were:

- To determine the genetic diversity in 25 elite maize inbred lines to aid in the selection of suitable lines for the introgression of the *msv1* gene that confers maize streak virus (MSV) disease resistance, thus enabling the production of the best high yielding hybrid possible for Mozambique.
- To evaluate the effectiveness of using marker-assisted selection (MAS) to transfer MSV resistance genes from CIMMYT donor lines (CML505 and CML509) into the Mozambican adapted line, LP23 by evaluating the F_3 progeny.
- To determine the effects of MSV disease on growth with emphasis on height of the infected maize plants.
- To identify progeny lines that combine both MSV and downy mildew (DM) disease resistance for potential use in developing MSV and DM resistant hybrids for Mozambique.

Summary and implications of the major findings

Selection of suitable lines for the introgression of the *msv1* gene for improved heterosis in a cultivar for Mozambique

- Genetic diversity was observed among the 25 candidate lines. The information was used to fit the lines into three heterotic groups, A, B and C. Breeding new MSV resistant hybrids from this set of germplasm would be viable, the identified groups can be exploited to develop new varieties.
- The identified heterotic patterns would be subjected to further testing in multi-location trials to confirm their yield potential and stability.

- Inbred lines CML505, CML509 and LP23 were identified as the best possible parental lines for maximum heterosis in hybrid combinations. Lines CML505 and CML509 were selected as the *msv1* gene donor lines and Line LP23 as the MSV disease susceptible recipient parent.

Detection of MSV disease resistance with SSR markers and high resolution melt (HRM) analysis of F_3 maize population samples stored on Whatman FTA™ elute cards

- It was demonstrated that maize DNA sampled using FTA™ cards was suitable as a template for both PCR and HRM curve analysis. Diagnostic techniques can be applied to DNA eluted from FTA™ cards in a manner equivalent to conventional DNA isolation methods. PCR-HRM analysis and FTA™ technology were successfully used in this study to identify progeny genotypes.
- Crudely extracted DNA was as effective as DNA extracted by the FTA™ technology for melt curve analysis.
- High resolution melt analysis with the markers bnlg1811 and umc2228 successfully differentiated between the two parental lines CML505 and CML509, respectively. The markers then distinguished alleles of the *msv1* trait based on melting curves and difference plots.

Phenotypic characterisation of progeny maize lines for MSV and DM resistance

- The infection of maize with MSV disease resulted in significant stunting.
- The study showed that it is possible through simultaneous selection to develop MSV and DM resistance in one base population. Incorporation of these genes in the MSV susceptible but locally adapted LP23 enhanced the levels of MSV and DM resistance.
- The implication is that yield in hybrids can be improved by exploiting these lines to make disease resistant lines for possible deployment in Mozambique.

Genetic diversity among maize lines selected for downy mildew and maize streak virus resistance as determined by SNP markers

- Four hundred SNP markers were successful in discriminating 50 maize inbred lines according to genetic distances, and resulted in clusters which were consistent with pedigree information and origin of the lines.
- The study identified two new progeny lines (DMSR69 and DMSR71) which were at least 85% similar to LP23 (the candidate for further improvement) and were resistant to both MSV and DM. These lines will be used as parents in lieu of LP23 to improve lowland maize hybrids in Mozambique for MSV resistance, with positive implications on yield and impact on food security.
- Observation of different genetic clusters for the new progeny lines and the standard testers can form the basis for allocating the lines into different heterotic groups with positive implication for the hybrid breeding programme in Mozambique. The identified groups can be exploited to develop new varieties and subjected to further testing in multi-location trials to confirm their yield potential and stability.

Recommendations

As the diversity data alone is not adequate to predict performance of hybrids, new lines may be crossed to produce hybrids and tested for yield under DM and MSV disease resistance in Mozambique.

APPENDICES

Appendix 1: The Ho, He, PIC and mean values for the 400 SNP markers used in the study

Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC
Fea2_1	0.4422	0.1800	0.3444	**PZA00499_12**	0.4950	0.1400	0.3725	**PZA01588_1**	0.2952	0.0800	0.2516
Fea2_2	0.4278	0.0600	0.3363	**PZA00527_10**	0.2688	0.0400	0.2327	**PZA01591_1**	0.4662	0.1800	0.3575
PHM12794_47	0.4998	0.3400	0.3749	**PZA00562_4**	0.4992	0.2800	0.3746	**PZA01597_1**	0.0768	0.0400	0.0739
PHM4196_27	0.4550	0.1400	0.3515	**PZA00565_3**	0.0000	0.0000	0.0000	**PZA01619_1**	0.3542	0.2200	0.2915
PHM4348_16	0.1128	0.0400	0.1064	**PZA00578_2**	0.0392	0.0000	0.0384	**PZA01715_2**	0.3942	0.1000	0.3165
PHM4531_46	0.1302	0.0600	0.1217	**PZA00582_4**	0.0582	0.0200	0.0565	**PZA01877_2**	0.0768	0.0000	0.0739
PZA00005_8	0.2262	0.1000	0.2006	**PZA00587_6**	0.0000	0.0000	0.0000	**PZA01964_29**	0.0392	0.0400	0.0384
PZA00071_2	0.2822	0.1400	0.2424	**PZA00600_11**	0.0000	0.0000	0.0000	**PZA02012_7**	0.4800	0.0400	0.3648
PZA00136_2	0.4608	0.1200	0.3546	**PZA00603_1**	0.4950	0.2200	0.3725	**PZA02029_21**	0.1800	0.0000	0.1638
PZA00160_3	0.2952	0.1600	0.2516	**PZA00606_10**	0.0392	0.0000	0.0384	**PZA02094_9**	0.4608	0.1200	0.3546
PZA00172_11	0.0392	0.0000	0.0384	**PZA00606_3**	0.4712	0.2400	0.3602	**PZA02113_1**	0.2952	0.0800	0.2516
PZA00210_9	0.0198	0.0200	0.0196	**PZA00721_4**	0.4712	0.1600	0.3602	**PZA02129_1**	0.4758	0.2200	0.3626
PZA00270_1	0.2262	0.0600	0.2006	**PZA00755_2**	0.4200	0.1600	0.3318	**PZA02148_1**	0.4032	0.1600	0.3219
PZA00270_3	0.0768	0.0400	0.0739	**PZA00770_1**	0.4608	0.1600	0.3546	**PZA02247_1**	0.4352	0.1200	0.3405
PZA00297_7	0.2112	0.0400	0.1889	**PZA00793_2**	0.0000	0.0000	0.0000	**PZA02264_5**	0.2952	0.0800	0.2516
PZA00332_7	0.2952	0.0800	0.2516	**PZA00878_2**	0.1302	0.0600	0.1217	**PZA02291_1**	0.4928	0.2000	0.3714
PZA00334_2	0.5000	0.0800	0.3750	**PZA00881_1**	0.3432	0.0800	0.2843	**PZA02396_14**	0.4352	0.1600	0.3405
PZA00363_7	0.3542	0.1000	0.2915	**PZA00902_1**	0.0392	0.0400	0.0384	**PZA02408_2**	0.2112	0.0400	0.1889
PZA00403_5	0.4118	0.1800	0.3270	**PZA00925_2**	0.1638	0.1800	0.1504	**PZA02426_1**	0.0768	0.0000	0.0739
PZA00442_5	0.4488	0.1200	0.3481	**PZA00944_2**	0.0392	0.0000	0.0384	**PZA02436_1**	0.4118	0.2600	0.3270
PZA00444_5	0.0000	0.0000	0.0000	**PZA00948_1**	0.3432	0.1200	0.2843	**PZA02509_14**	0.1800	0.0800	0.1638
PZA00455_16	0.1800	0.0400	0.1638	**PZA01038_1**	0.1638	0.0600	0.1504	**PZA02514_1**	0.0198	0.0200	0.0196
PZA00460_8	0.3648	0.0800	0.2983	**PZA01230_1**	0.0198	0.0200	0.0196	**PZA02549_3**	0.0768	0.0400	0.0739
PZA00462_2	0.1800	0.0000	0.1638	**PZA01292_1**	0.0950	0.0600	0.0905	**PZA02564_2**	0.1302	0.0600	0.1217
PZA00466_2	0.0582	0.0200	0.0565	**PZA01410_1**	0.0392	0.0400	0.0384	**PZA02585_2**	0.4352	0.1200	0.3405
PZA00486_2	0.3318	0.4200	0.2768	**PZA01557_1**	0.0198	0.0200	0.0196	**PZA02606_1**	0.0950	0.0200	0.0905
PZA00489_1	0.0000	0.0000	0.0000	**PZA01570_1**	0.1472	0.0400	0.1364	**PZA02683_1**	0.1800	0.0400	0.1638

Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC
PZA02722_1	0.4982	0.1800	0.3741	**PZA03490_1**	0.2822	0.1800	0.2424	**PZA03728_1**	0.0392	0.0400	0.0384
PZA02817_15	0.2688	0.0400	0.2327	**PZA03498_1**	0.0582	0.0200	0.0565	**PZA03731_2**	0.0000	0.0000	0.0000
PZA02817_3	0.2550	0.0600	0.2225	**PZA03504_1**	0.1302	0.0600	0.1217	**PZA03732_2**	0.0000	0.0000	0.0000
PZA02949_26	0.0000	0.0000	0.0000	**PZA03505_1**	0.4608	0.1200	0.3546	**PZA03732_3**	0.5000	1.0000	0.3750
PZA02984_10	0.0582	0.0600	0.0565	**PZA03520_3**	0.0198	0.0200	0.0196	**PZA03733_1**	0.2688	0.0800	0.2327
PZA03012_10	0.0000	0.0000	0.0000	**PZA03533_1**	0.5000	1.0000	0.3750	**PZA03735_1**	0.2688	0.0800	0.2327
PZA03034_1	0.4838	0.1400	0.3668	**PZA03568_1**	0.0198	0.0200	0.0196	**PZA03742_2**	0.4662	0.2600	0.3575
PZA03116_2	0.4998	0.2600	0.3749	**PZA03569_2**	0.0198	0.0200	0.0196	**PZA03747_1**	0.2262	0.0600	0.2006
PZA03182_5	0.4278	0.0600	0.3363	**PZA03573_1**	0.0582	0.0600	0.0565	**PZA03750_2**	0.1302	0.0200	0.1217
PZA03191_2	0.0768	0.0800	0.0739	**PZA03573_3**	0.4758	0.2200	0.3626	**PZA03760_3**	0.0198	0.0200	0.0196
PZA03243_4	0.0950	0.1000	0.0905	**PZA03587_1**	0.0582	0.0200	0.0565	**PZB00001_2**	0.4278	0.1800	0.3363
PZA03244_4	0.0582	0.0200	0.0565	**PZA03598_1**	0.4118	0.1000	0.3270	**PZB00054_3**	0.0000	0.0000	0.0000
PZA03255_4	0.1302	0.0200	0.1217	**PZA03629_1**	0.0950	0.0600	0.0905	**PZB00062_10**	0.0000	0.0000	0.0000
PZA03289_4	0.0392	0.0000	0.0384	**PZA03637_3**	0.0950	0.0600	0.0905	**PZB00062_9**	0.0392	0.0400	0.0384
PZA03359_4	0.0000	0.0000	0.0000	**PZA03638_1**	0.0950	0.0600	0.0905	**PZB00104_1**	0.3078	0.0600	0.2604
PZA03366_2	0.0000	0.0000	0.0000	**PZA03645_2**	0.1302	0.0200	0.1217	**PZB00114_1**	0.2262	0.0600	0.2006
PZA03384_1	0.0000	0.0000	0.0000	**PZA03668_4**	0.0000	0.0000	0.0000	**PZB00125_1**	0.0198	0.0200	0.0196
PZA03385_1	0.4200	0.1600	0.3318	**PZA03673_1**	0.0198	0.0200	0.0196	**PZB00165_6**	0.3200	0.1200	0.2688
PZA03385_2	0.2688	0.3200	0.2327	**PZA03673_2**	0.0000	0.0000	0.0000	**PZB00175_6**	0.0198	0.0200	0.0196
PZA03388_1	0.4550	0.2200	0.3515	**PZA03677_1**	0.1638	0.0600	0.1504	**PZB00207_3**	0.0000	0.0000	0.0000
PZA03411_3	0.0000	0.0000	0.0000	**PZA03686_1**	0.1472	0.0000	0.1364	**PZB00235_1**	0.4758	0.2200	0.3626
PZA03431_1	0.0000	0.0000	0.0000	**PZA03695_2**	0.0198	0.0200	0.0196	**PZB00425_1**	0.0000	0.0000	0.0000
PZA03445_1	0.4200	0.1600	0.3318	**PZA03696_2**	0.0392	0.0000	0.0384	**PZB00592_1**	0.4662	0.1400	0.3575
PZA03452_6	0.1128	0.0400	0.1064	**PZA03696_3**	0.0198	0.0200	0.0196	**PZB00607_2**	0.0392	0.0000	0.0384
PZA03461_1	0.0582	0.0200	0.0565	**PZA03700_3**	0.0000	0.0000	0.0000	**PZB00677_3**	0.0392	0.0000	0.0384
PZA03470_1	0.0198	0.0200	0.0196	**PZA03706_1**	0.0582	0.0200	0.0565	**PZB00677_4**	0.0950	0.0200	0.0905
PZA03477_1	0.4982	0.2200	0.3741	**PZA03714_1**	0.0768	0.0400	0.0739	**PZB00746_1**	0.0000	0.0000	0.0000
PZA03478_1	0.1800	0.0400	0.1638	**PZA03716_1**	0.5000	1.0000	0.3750	**PZB00772_4**	0.0198	0.0200	0.0196
PZA03484_1	0.1638	0.0600	0.1504	**PZA03719_1**	0.4278	0.1400	0.3363	**PZB00895_3**	0.0950	0.1000	0.0905

Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC
PZB00963_2	0.0950	0.0600	0.0905	**PZB01869_4**	0.4278	0.6200	0.3363	**ba1_6**	0.0000	0.0000	0.0000
PZB00963_3	0.0000	0.0000	0.0000	**PZB01963_2**	0.0392	0.0400	0.0384	**fea2_3**	0.0198	0.0200	0.0196
PZB01021_5	0.0950	0.0200	0.0905	**PZB01963_4**	0.0000	0.0000	0.0000	**sh2_3**	0.2688	0.0400	0.2327
PZB01051_1	0.0392	0.0000	0.0384	**PZB01964_5**	0.0950	0.0200	0.0905	**zb27_1**	0.0768	0.0400	0.0739
PZB01057_4	0.0000	0.0000	0.0000	**PZB01977_11**	0.4422	0.0200	0.3444	**zb7_2**	0.0392	0.0000	0.0384
PZB01086_2	0.0000	0.0000	0.0000	**PZB01977_4**	0.1472	0.1200	0.1364	**PZA00031_5**	0.0950	0.0600	0.0905
PZB01103_4	0.0000	0.0000	0.0000	**PZB01977_9**	0.2550	0.3000	0.2225	**PZA00047_2**	0.4700	0.2245	0.3596
PZB01107_8	0.0198	0.0200	0.0196	**PZB02017_2**	0.2822	0.1000	0.2424	**PZA00210_8**	0.0198	0.0200	0.0196
PZB01110_6	0.2112	0.0400	0.1889	**PZB02020_2**	0.0000	0.0000	0.0000	**PZA00237_8**	0.0968	0.1020	0.0921
PZB01111_3	0.0768	0.0400	0.0739	**PZB02033_1**	0.4352	0.6400	0.3405	**PZA00297_4**	0.0582	0.0200	0.0565
PZB01111_6	0.4950	0.3400	0.3725	**PZB02033_2**	0.0392	0.0000	0.0384	**PZA00326_18**	0.4398	0.1633	0.3431
PZB01112_1	0.1958	0.0200	0.1766	**PZB02122_1**	0.2408	0.2000	0.2118	**PZA00498_5**	0.4422	0.1800	0.3444
PZB01114_2	0.0950	0.0200	0.0905	**PZB02179_1**	0.2952	0.0400	0.2516	**PZA00516_3**	0.2688	0.0400	0.2327
PZB01186_1	0.2112	0.0800	0.1889	**PZB02227_2**	0.3078	0.1000	0.2604	**PZA00523_2**	0.0202	0.0204	0.0200
PZB01186_4	0.0000	0.0000	0.0000	**PZB02448_1**	0.4608	0.2400	0.3546	**PZA00587_4**	0.2732	0.2041	0.2359
PZB01261_2	0.2822	0.1800	0.2424	**PZB02516_1**	0.4872	0.1600	0.3685	**PZA00616_13**	0.4398	0.1224	0.3431
PZB01301_6	0.2688	0.1200	0.2327	**PZB02534_3**	0.0950	0.0200	0.0905	**PZA00726_8**	0.3318	0.0600	0.2768
PZB01370_1	0.4608	0.2000	0.3546	**PZB02542_1**	0.0000	0.0000	0.0000	**PZA00740_1**	0.0000	0.0000	0.0000
PZB01412_2	0.0392	0.0000	0.0384	**PZB02542_3**	0.0000	0.0000	0.0000	**PZA01029_1**	0.3992	0.1020	0.3195
PZB01460_2	0.0000	0.0000	0.0000	**PZB02544_1**	0.1128	0.1200	0.1064	**PZA01216_1**	0.3698	0.1224	0.3015
PZB01463_2	0.0000	0.0000	0.0000	**PZD00016_4**	0.1128	0.0400	0.1064	**PZA01315_1**	0.4467	0.0204	0.3469
PZB01463_7	0.0198	0.0200	0.0196	**PZD00022_6**	0.3942	0.0600	0.3165	**PZA01652_1**	0.4648	0.2449	0.3568
PZB01500_1	0.0000	0.0000	0.0000	**PZD00027_5**	0.0000	0.0000	0.0000	**PZA01726_1**	0.0392	0.0400	0.0384
PZB01617_2	0.1128	0.0000	0.1064	**PZD00043_2**	0.0000	0.0000	0.0000	**PZA02011_1**	0.2449	0.0408	0.2149
PZB01642_2	0.0000	0.0000	0.0000	**PZD00043_4**	0.0000	0.0000	0.0000	**PZA02197_1**	0.2868	0.0612	0.2456
PZB01683_2	0.0000	0.0000	0.0000	**PZD00056_1**	0.3648	0.1200	0.2983	**PZA02203_1**	0.4992	0.2245	0.3746
PZB01689_3	0.4488	0.1200	0.3481	**PZD00066_5**	0.0198	0.0200	0.0196	**PZA02266_3**	0.4700	0.7551	0.3596
PZB01730_3	0.0768	0.0400	0.0739	**PZD00072_2**	0.4422	0.1800	0.3444	**PZA02296_1**	0.4531	0.1224	0.3505
PZB01856_1	0.4968	0.9200	0.3734	**Ra1_1**	0.0198	0.0200	0.0196	**PZA02388_1**	0.3898	0.1633	0.3139

Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC	Marker	Gene Diversity	Heterozygosity	PIC
PZA02423_1	0.4792	0.1837	0.3644	**PZB00087_1**	0.5000	1.0000	0.3750	**PZB02542_1**	0.0000	0.0000	0.0000
PZA02478_7	0.4748	0.3265	0.3621	**PZB00092_2**	0.3367	0.4286	0.2800	**PZB02542_3**	0.0000	0.0000	0.0000
PZA02496_1	0.5098	0.0800	0.3897	**PZB00686_2**	0.4700	0.1429	0.3596	**PZB02544_1**	0.1128	0.1200	0.1064
PZA02589_1	0.4831	0.2857	0.3664	**PZB00859_1**	0.1833	0.0816	0.1665	**PZD00016_4**	0.1128	0.0400	0.1064
PZA02616_1	0.4648	0.2449	0.3568	**PZB01186_3**	0.3898	0.1224	0.3139	**PZD00022_6**	0.3942	0.0600	0.3165
PZA02746_2	0.4913	0.2653	0.3802	**PZB01403_1**	0.4467	0.1429	0.3469	**PZD00027_5**	0.0000	0.0000	0.0000
PZA02890_4	0.1993	0.0612	0.1794	**PZB01403_3**	0.0402	0.0408	0.0398	**PZD00043_2**	0.0000	0.0000	0.0000
PZA02981_2	0.3750	0.0600	0.3047	**PZB01403_4**	0.2868	0.0612	0.2456	**PZD00043_4**	0.0000	0.0000	0.0000
PZA03069_4	0.2822	0.1000	0.2424	**PZB01647_1**	0.0968	0.0612	0.0921	**PZD00056_1**	0.3648	0.1200	0.2983
PZA03120_1	0.1833	0.1224	0.1665	**PZB01881_11**	0.0582	0.0200	0.0565	**PZD00066_5**	0.0198	0.0200	0.0196
PZA03243_7	0.0000	0.0000	0.0000	**PZB01919_1**	0.4838	0.2200	0.3668	**PZD00072_2**	0.4422	0.1800	0.3444
PZA03329_1	0.2868	0.1429	0.2456	**PZB01963_1**	0.0768	0.0400	0.0739	**Ra1_1**	0.0198	0.0200	0.0196
PZA03381_2	0.1833	0.0408	0.1665	**PZB01963_3**	0.0783	0.0408	0.0752	**ba1_6**	0.0000	0.0000	0.0000
PZA03388_2	0.0000	0.0000	0.0000	**PZB02155_1**	0.4467	0.1429	0.3469	**fea2_3**	0.0198	0.0200	0.0196
PZA03398_2	0.2262	0.1000	0.2006	**PZB02480_1**	0.4200	0.0800	0.3318	**sh2_3**	0.2688	0.0400	0.2327
PZA03442_1	0.2112	0.0800	0.1889	**PZB01964_5**	0.0950	0.0200	0.0905	**zb27_1**	0.0768	0.0400	0.0739
PZA03462_1	0.4992	0.2000	0.3746	**PZB01977_11**	0.4422	0.0200	0.3444	**zb7_2**	0.0392	0.0000	0.0384
PZA03474_1	0.1685	0.0204	0.1573	**PZB01977_4**	0.1472	0.1200	0.1364	**PZA00031_5**	0.0950	0.0600	0.0905
PZA03519_2	0.1638	0.0200	0.1504	**PZB01977_9**	0.2550	0.3000	0.2225	**PZA00047_2**	0.4700	0.2245	0.3596
PZA03528_1	0.3800	0.1429	0.3078	**PZB02017_2**	0.2822	0.1000	0.2424	**PZA00210_8**	0.0198	0.0200	0.0196
PZA03583_2	0.0392	0.0400	0.0384	**PZB02020_2**	0.0000	0.0000	0.0000	**PZA00237_8**	0.0968	0.1020	0.0921
PZA03607_1	0.2449	0.0816	0.2149	**PZB02033_1**	0.4352	0.6400	0.3405	**PZA00297_4**	0.0582	0.0200	0.0565
PZA03632_2	0.0392	0.0000	0.0384	**PZB02033_2**	0.0392	0.0000	0.0384	**PZA00326_18**	0.4398	0.1633	0.3431
PZA03650_1	0.2952	0.0400	0.2516	**PZB02122_1**	0.2408	0.2000	0.2118	**PZA00498_5**	0.4422	0.1800	0.3444
PZA03663_1	0.4967	0.3061	0.3733	**PZB02179_1**	0.2952	0.0400	0.2516	**PZA00516_3**	0.2688	0.0400	0.2327
PZA03668_1	0.3200	0.1200	0.2688	**PZB02227_2**	0.3078	0.1000	0.2604	**PZA00523_2**	0.0202	0.0204	0.0200
PZA03676_2	0.0000	0.0000	0.0000	**PZB02448_1**	0.4608	0.2400	0.3546	**PZA00587_4**	0.2732	0.2041	0.2359
PZA03714_3	0.1302	0.0600	0.1217	**PZB02516_1**	0.4872	0.1600	0.3685	**PZA00616_13**	0.4398	0.1224	0.3431
PZB00068_1	0.4872	0.2400	0.3685	**PZB02534_3**	0.0950	0.0200	0.0905	**PZA00726_8**	0.3318	0.0600	0.2768

Marker	Gene Diversity	Heterozygosity	PIC
PZA00740_1	0.0000	0.0000	0.0000
PZA01029_1	0.3992	0.1020	0.3195
PZA01216_1	0.3698	0.1224	0.3015
PZA01315_1	0.4467	0.0204	0.3469
PZA01652_1	0.4648	0.2449	0.3568
PZA01726_1	0.0392	0.0400	0.0384
PZA02011_1	0.2449	0.0408	0.2149
PZA02197_1	0.2868	0.0612	0.2456
PZA02203_1	0.4992	0.2245	0.3746
PZA02266_3	0.4700	0.7551	0.3596
PZA02296_1	0.4531	0.1224	0.3505
PZA02388_1	0.3898	0.1633	0.3139
PZA02423_1	0.4792	0.1837	0.3644
PZA02478_7	0.4748	0.3265	0.3621
PZA02496_1	0.5098	0.0800	0.3897
PZA02589_1	0.4831	0.2857	0.3664
PZA02616_1	0.4648	0.2449	0.3568
PZA02746_2	0.4913	0.2653	0.3802
PZA02890_4	0.1993	0.0612	0.1794
PZA02981_2	0.3750	0.0600	0.3047
PZA03069_4	0.2822	0.1000	0.2424
PZA03120_1	0.1833	0.1224	0.1665
PZA03243_7	0.0000	0.0000	0.0000
PZA03329_1	0.2868	0.1429	0.2456
PZA03381_2	0.1833	0.0408	0.1665
PZA03388_2	0.0000	0.0000	0.0000
PZA03398_2	0.2262	0.1000	0.2006
PZA03442_1	0.2112	0.0800	0.1889
PZA03462_1	0.4992	0.2000	0.3746
PZA03474_1	0.1685	0.0204	0.1573
PZA03519_2	0.1638	0.0200	0.1504
PZA03528_1	0.3800	0.1429	0.3078
PZA03583_2	0.0392	0.0400	0.0384
PZA03607_1	0.2449	0.0816	0.2149
PZA03632_2	0.0392	0.0000	0.0384
PZA03650_1	0.2952	0.0400	0.2516
PZA03663_1	0.4967	0.3061	0.3733
PZA03668_1	0.3200	0.1200	0.2688
PZA03676_2	0.0000	0.0000	0.0000
PZA03714_3	0.1302	0.0600	0.1217
PZB00068_1	0.4872	0.2400	0.3685
PZB00087_1	0.5000	1.0000	0.3750
PZB00092_2	0.3367	0.4286	0.2800
PZB00686_2	0.4700	0.1429	0.3596
PZB00859_1	0.1833	0.0816	0.1665
PZB01186_3	0.3898	0.1224	0.3139
PZB01403_1	0.4467	0.1429	0.3469
PZB01403_3	0.0402	0.0408	0.0398
PZB01403_4	0.2868	0.0612	0.2456
PZB01647_1	0.0968	0.0612	0.0921
PZB01881_11	0.0582	0.0200	0.0565
PZB01919_1	0.4838	0.2200	0.3668
PZB01963_1	0.0768	0.0400	0.0739
PZB01963_3	0.0783	0.0408	0.0752
PZB02155_1	0.4467	0.1429	0.3469
PZB02480_1	0.4200	0.0800	0.3318
MEAN	**0.2215**	**0.1051**	**0.1805**

Appendix 2: Similarity matrix of the 50 maize inbred lines based on 400 SNP markers

	B73WX	CML202	DMSR1	DMSR10	DMSR12	DMSR13	DMSR16	DMSR18	DMSR2	DMSR21	DMSR23	DMSR26	DMSR30	DMSR34
B73WX	1.00													
CML202	0.60	1.00												
DMSR1	0.64	0.67	1.00											
DMSR10	0.63	0.64	0.86	1.00										
DMSR12	0.60	0.63	0.79	0.78	1.00									
DMSR13	0.61	0.63	0.80	0.74	0.83	1.00								
DMSR16	0.62	0.63	0.86	0.79	0.77	0.85	1.00							
DMSR18	0.62	0.64	0.78	0.80	0.76	0.76	0.80	1.00						
DMSR2	0.60	0.66	0.79	0.81	0.81	0.78	0.82	0.78	1.00					
DMSR21	0.63	0.62	0.79	0.82	0.81	0.78	0.81	0.85	0.82	1.00				
DMSR23	0.64	0.67	0.85	0.80	0.76	0.78	0.81	0.82	0.78	0.77	1.00			
DMSR26	0.61	0.64	0.86	0.85	0.77	0.74	0.82	0.76	0.79	0.78	0.81	1.00		
DMSR30	0.61	0.64	0.85	0.79	0.77	0.77	0.76	0.77	0.79	0.75	0.83	0.79	1.00	
DMSR34	0.62	0.63	0.85	0.82	0.76	0.75	0.76	0.81	0.75	0.81	0.81	0.82	0.80	1.00
DMSR35	0.59	0.68	0.79	0.76	0.76	0.80	0.78	0.77	0.84	0.81	0.81	0.79	0.83	0.77
DMSR39	0.64	0.66	0.79	0.74	0.70	0.69	0.76	0.75	0.70	0.74	0.81	0.80	0.72	0.77
DMSR4	0.61	0.64	0.81	0.79	0.81	0.80	0.75	0.78	0.77	0.80	0.83	0.74	0.80	0.79
DMSR40	0.61	0.68	0.75	0.71	0.65	0.68	0.69	0.72	0.70	0.69	0.76	0.72	0.73	0.71
DMSR43	0.62	0.69	0.82	0.74	0.67	0.71	0.75	0.76	0.71	0.72	0.82	0.78	0.77	0.78
DMSR46	0.59	0.68	0.74	0.74	0.68	0.68	0.72	0.77	0.70	0.71	0.75	0.75	0.69	0.73
DMSR47	0.63	0.70	0.76	0.70	0.66	0.69	0.75	0.74	0.68	0.70	0.78	0.74	0.69	0.75
DMSR51	0.59	0.68	0.76	0.74	0.67	0.68	0.70	0.75	0.69	0.70	0.81	0.76	0.74	0.77
DMSR55	0.61	0.67	0.76	0.73	0.69	0.68	0.72	0.75	0.68	0.73	0.79	0.75	0.70	0.77
DMSR56	0.61	0.69	0.76	0.70	0.65	0.69	0.73	0.72	0.69	0.70	0.81	0.74	0.74	0.74
DMSR57	0.60	0.68	0.70	0.67	0.62	0.65	0.68	0.70	0.66	0.66	0.76	0.67	0.69	0.70
DMSR60	0.59	0.66	0.76	0.74	0.67	0.69	0.72	0.76	0.69	0.72	0.77	0.77	0.70	0.75
DMSR62	0.60	0.68	0.75	0.72	0.69	0.68	0.69	0.73	0.67	0.70	0.79	0.74	0.72	0.76
DMSR64	0.60	0.69	0.77	0.74	0.68	0.67	0.71	0.75	0.69	0.70	0.77	0.77	0.71	0.75
DMSR65	0.61	0.68	0.74	0.68	0.63	0.68	0.73	0.74	0.68	0.68	0.78	0.73	0.71	0.73
DMSR66	0.62	0.67	0.76	0.73	0.68	0.67	0.73	0.75	0.72	0.72	0.78	0.75	0.72	0.74
DMSR69	0.61	0.66	0.78	0.75	0.67	0.70	0.75	0.76	0.72	0.72	0.79	0.78	0.75	0.78
DMSR71	0.61	0.67	0.80	0.73	0.68	0.71	0.75	0.78	0.70	0.71	0.79	0.76	0.76	0.78
DMSR73	0.64	0.70	0.79	0.74	0.68	0.70	0.74	0.75	0.69	0.72	0.80	0.76	0.74	0.75
DMSR74	0.61	0.66	0.79	0.74	0.69	0.67	0.71	0.72	0.67	0.70	0.77	0.77	0.74	0.76
DMSR75	0.73	0.68	0.73	0.72	0.67	0.67	0.69	0.72	0.69	0.70	0.76	0.70	0.70	0.72
DMSR77	0.60	0.69	0.72	0.70	0.64	0.64	0.68	0.73	0.69	0.69	0.79	0.72	0.68	0.73
DMSR8	0.63	0.64	0.83	0.81	0.79	0.80	0.80	0.81	0.80	0.80	0.82	0.80	0.82	0.82
CML505	0.61	0.63	0.73	0.77	0.82	0.80	0.75	0.74	0.81	0.83	0.70	0.70	0.73	0.71
CML509	0.60	0.64	0.64	0.62	0.62	0.62	0.62	0.65	0.61	0.63	0.68	0.62	0.63	0.62
LP23	0.57	0.64	0.83	0.77	0.71	0.72	0.78	0.82	0.73	0.73	0.84	0.83	0.80	0.82
DXL37	0.60	0.62	0.63	0.59	0.59	0.62	0.63	0.61	0.62	0.61	0.63	0.61	0.64	0.61
DXL59	0.61	0.62	0.62	0.61	0.62	0.62	0.61	0.60	0.62	0.60	0.62	0.61	0.61	0.61
H24W	0.72	0.64	0.68	0.69	0.63	0.64	0.67	0.67	0.65	0.67	0.70	0.65	0.64	0.67
I137TN	0.61	0.58	0.58	0.59	0.58	0.57	0.58	0.59	0.58	0.61	0.59	0.59	0.59	0.58
LP19	0.67	0.65	0.65	0.66	0.66	0.65	0.64	0.66	0.65	0.67	0.67	0.63	0.64	0.64
M017	0.72	0.56	0.60	0.60	0.57	0.58	0.58	0.59	0.57	0.59	0.62	0.58	0.60	0.58
M162W	0.58	0.56	0.58	0.55	0.56	0.57	0.55	0.56	0.56	0.57	0.58	0.56	0.58	0.58
PA1	0.63	0.63	0.62	0.63	0.61	0.62	0.62	0.62	0.64	0.61	0.64	0.63	0.62	0.63

Appendix 2.....continued

	DMSR35	DMSR39	DMSR4	DMSR40	DMSR43	DMSR46	DMSR47	DMSR51	DMSR55	DMSR56	DMSR57	DMSR60	DMSR62	DMSR64
DMSR35	1.00													
DMSR39	0.71	1.00												
DMSR4	0.76	0.71	1.00											
DMSR40	0.73	0.76	0.68	1.00										
DMSR43	0.74	0.85	0.72	0.82	1.00									
DMSR46	0.71	0.79	0.69	0.82	0.76	1.00								
DMSR47	0.70	0.86	0.68	0.72	0.81	0.79	1.00							
DMSR51	0.73	0.78	0.74	0.79	0.79	0.82	0.81	1.00						
DMSR55	0.71	0.87	0.72	0.74	0.80	0.84	0.88	0.83	1.00					
DMSR56	0.74	0.86	0.69	0.81	0.86	0.76	0.82	0.83	0.82	1.00				
DMSR57	0.68	0.80	0.65	0.78	0.78	0.74	0.80	0.77	0.77	0.82	1.00			
DMSR60	0.71	0.83	0.70	0.81	0.82	0.86	0.78	0.81	0.81	0.78	0.76	1.00		
DMSR62	0.70	0.82	0.74	0.80	0.83	0.79	0.79	0.87	0.84	0.83	0.82	0.81	1.00	
DMSR64	0.71	0.81	0.69	0.83	0.80	0.85	0.78	0.80	0.80	0.76	0.76	0.87	0.80	1.00
DMSR65	0.72	0.84	0.68	0.75	0.81	0.78	0.87	0.82	0.82	0.83	0.80	0.79	0.77	0.77
DMSR66	0.71	0.87	0.72	0.79	0.83	0.80	0.84	0.80	0.83	0.82	0.81	0.81	0.82	0.80
DMSR69	0.74	0.85	0.71	0.79	0.86	0.81	0.82	0.82	0.82	0.85	0.80	0.82	0.81	0.80
DMSR71	0.73	0.80	0.69	0.82	0.83	0.80	0.79	0.76	0.79	0.81	0.78	0.74	0.77	0.79
DMSR73	0.72	0.87	0.71	0.78	0.84	0.78	0.85	0.82	0.86	0.85	0.82	0.80	0.82	0.84
DMSR74	0.70	0.82	0.71	0.79	0.80	0.80	0.80	0.85	0.84	0.80	0.77	0.77	0.83	0.82
DMSR75	0.69	0.78	0.71	0.75	0.76	0.76	0.76	0.77	0.80	0.77	0.72	0.74	0.76	0.75
DMSR77	0.69	0.84	0.69	0.77	0.82	0.76	0.83	0.82	0.85	0.84	0.81	0.77	0.82	0.77
DMSR8	0.81	0.76	0.79	0.71	0.78	0.73	0.72	0.73	0.73	0.75	0.70	0.73	0.74	0.72
CML505	0.78	0.62	0.78	0.64	0.63	0.61	0.60	0.61	0.61	0.61	0.62	0.61	0.62	0.61
CML509	0.65	0.64	0.65	0.63	0.64	0.64	0.65	0.64	0.66	0.63	0.67	0.64	0.65	0.65
LP23	0.75	0.81	0.75	0.75	0.84	0.82	0.80	0.83	0.81	0.80	0.73	0.81	0.80	0.79
DXL37	0.65	0.64	0.61	0.64	0.64	0.62	0.65	0.61	0.63	0.66	0.64	0.61	0.61	0.61
DXL59	0.61	0.61	0.61	0.62	0.61	0.60	0.61	0.61	0.61	0.61	0.59	0.60	0.62	0.60
H24W	0.64	0.67	0.65	0.64	0.68	0.65	0.66	0.65	0.67	0.67	0.65	0.65	0.65	0.63
I137TN	0.60	0.59	0.60	0.59	0.59	0.57	0.57	0.58	0.58	0.58	0.57	0.57	0.59	0.57
LP19	0.66	0.66	0.66	0.66	0.65	0.66	0.65	0.64	0.67	0.64	0.65	0.65	0.65	0.65
M017	0.58	0.60	0.59	0.59	0.60	0.55	0.58	0.57	0.57	0.60	0.56	0.57	0.58	0.56
M162W	0.57	0.56	0.57	0.57	0.57	0.55	0.56	0.57	0.56	0.58	0.59	0.55	0.56	0.56
PAA1	0.62	0.62	0.62	0.63	0.62	0.62	0.62	0.60	0.61	0.61	0.63	0.61	0.61	0.61
PA2	0.62	0.61	0.62	0.58	0.62	0.57	0.62	0.59	0.61	0.61	0.59	0.56	0.58	0.56
8CED67	0.62	0.61	0.63	0.61	0.65	0.58	0.61	0.59	0.59	0.63	0.60	0.59	0.60	0.59

Appendix 2.....continued

	DMSR65	DMSR66	DMSR69	DMSR71	DMSR73	DMSR74	DMSR75	DMSR77	DMSR8	CML505	CML509	LP23
DMSR65	1.00											
DMSR66	0.85	1.00										
DMSR69	0.83	0.84	1.00									
DMSR71	0.76	0.79	0.84	1.00								
DMSR73	0.86	0.84	0.84	0.79	1.00							
DMSR74	0.77	0.83	0.81	0.81	0.85	1.00						
DMSR75	0.76	0.77	0.76	0.74	0.79	0.78	1.00					
DMSR77	0.81	0.81	0.80	0.75	0.83	0.81	0.77	1.00				
DMSR8	0.74	0.74	0.77	0.75	0.76	0.72	0.71	0.71	1.00			
CML505	0.60	0.63	0.61	0.61	0.64	0.60	0.65	0.61	0.75	1.00		
CML509	0.64	0.64	0.61	0.62	0.66	0.63	0.67	0.64	0.64	0.63	1.00	
LP23	0.80	0.80	0.86	0.85	0.79	0.80	0.72	0.77	0.80	0.58	0.61	1.00
DXL37	0.65	0.64	0.62	0.63	0.64	0.60	0.67	0.61	0.62	0.59	0.61	0.61
DXL59	0.59	0.62	0.58	0.60	0.60	0.60	0.67	0.60	0.61	0.63	0.61	0.57
H24W	0.64	0.66	0.66	0.66	0.68	0.64	0.75	0.65	0.69	0.64	0.63	0.64
I137TN	0.57	0.59	0.58	0.58	0.61	0.58	0.61	0.57	0.60	0.62	0.64	0.54
LP19	0.63	0.66	0.64	0.64	0.67	0.63	0.72	0.64	0.65	0.66	0.66	0.61
M017	0.57	0.58	0.58	0.57	0.60	0.57	0.63	0.58	0.59	0.60	0.59	0.55
M162W	0.56	0.57	0.55	0.57	0.58	0.55	0.64	0.55	0.58	0.59	0.60	0.52
PAA1	0.60	0.61	0.61	0.62	0.64	0.61	0.68	0.62	0.62	0.62	0.62	0.59
PA2	0.60	0.61	0.59	0.59	0.61	0.58	0.64	0.60	0.62	0.63	0.62	0.58
8CED67	0.61	0.61	0.60	0.60	0.61	0.58	0.66	0.59	0.64	0.61	0.59	0.59

Appendix 2.....continued

	DXL37	DXL59	H24W	I137TN	LP19	M017	M162W	PAA1	PA2	8CED67
DXL37	1.00									
DXL59	0.65	1.00								
H24W	0.67	0.65	1.00							
I137TN	0.58	0.60	0.65	1.00						
LP19	0.63	0.66	0.68	0.63	1.00					
M017	0.59	0.58	0.63	0.60	0.60	1.00				
M162W	0.61	0.68	0.64	0.58	0.61	0.55	1.00			
PAA1	0.60	0.59	0.66	0.59	0.63	0.57	0.56	1.00		
PA2	0.63	0.62	0.65	0.58	0.61	0.60	0.58	0.60	1.00	
8CED67	0.73	0.63	0.69	0.60	0.61	0.57	0.60	0.61	0.62	1