IDENTIFICATION OF QUANTITATIVE TRAIT LOCI FOR ANTHESIS-SILKING INTERVAL UNDER FLOWERING WATER STRESS USING SSR MARKERS IN SELECTED KENYAN MAIZE (*Zea mays* L.) LANDRACES

DORCUS CHEPKESIS GEMENET

A Thesis Submitted to the Graduate School in Partial Fulfillment of the Requirements for the Award of the Master of Science Degree in Agronomy (Plant Breeding Option) of Egerton University

EULIB
038 679

EGERTON UNIVERSITY LIBRARY
- 3 NOV 2014
P.O. BOX 536,
EGERTON
AFRICANA SECTION

EGERTON UNIVERSITY

November, 2009

2013/194664

EGERTON UNIVERSITY LIBRARY

DECLARATION AND RECOMMENDATION

DECLARATION

I hereby declare that this is my original work and has not been previously presented in this or any other University for the award of a degree.

Signature..... Date...5·11·09.....

Dorcus Chepkesis Gemenet

RECOMMENDATION

This thesis has been submitted with our approval as the University Supervisors.

Signature...R S Pathak..... Date...5·11·09.....

Prof. R. S. Pathak
Department of Crops, Horticulture and Soils

Signature..... Date...5/11/2009.....

Dr. F. N. Wachira
Department of Biochemistry and Molecular Biology

2013/94664

COPYRIGHT

© 2009

DORCUS CHEPKESIS GEMENET

All rights reserved. No part of this thesis may be reproduced, stored in retrieval system or transmitted in any form or by any means be it electronic, mechanical, photocopying, recording or otherwise, without the permission of Dorcus Chepkesis Gemenet.

DEDICATION

To my parents, Mr. Musa Gemenet and Mrs. Eunice Gemenet

ACKNOWLEDGEMENTS

I would like with pleasure to acknowledge all the people who directly or indirectly offered assistance at all stages during my study for this degree. I express my sincere gratitude to my supervisors Prof. R.S. Pathak and Dr. F.N. Wachira both of Egerton University for their guidance and technical support offered during the study. They were always there and ready to listen even within the shortest notice.

Financial support from the Regional Universities Forum for Capacity Building in Agriculture (RUFORUM), to undertake this study is gratefully acknowledged. The RUFORUM Biotechnology Grant-2005CG13 awarded to Prof. R.S. Pathak (Principal Investigator) and Dr. F.N. Wachira (Co-investigator) paid part of my first year tuition fee, a 12- month stipend and all field and laboratory expenses.

I thank the Director, KARI-Katumani and Mr. Mutunga, the in-charge of KARI-Masongaleni sub-station for providing the field facilities for screening maize landraces for drought tolerance. The friendly working environment during field work made the harsh conditions of Masongaleni bearable. In this regard, I also sincerely thank Mr. Mike Kuria and Mr. Erick Mwembia of Greenland Agro-Producers Limited. Their company was highly appreciated.

I appreciate the assistance and the laboratory facilities provided by the Tea Research Foundation of Kenya (TRFK) where the first part of molecular work was done. I specifically thank Dr. Samson Kamunya for his technical support. Sincere regards to Mr. Richard Chalo, Mr. Robert Korir, Mr. Joseph Koske and Mr. Tony Maritim for their input in the molecular work and to the entire TRFK fraternity for offering a friendly working environment.

My special gratitude to staff of the Department of Biochemistry and Molecular Biology, Egerton university, through Dr. F.N. Wachira who enabled the completion of molecular analysis. I particularly thank Mr. John Ndung'u of KARI-Njoro and a fellow student for his assistance and company. I also acknowledge my other fellow students Mr. Godwin Macharia, Ms. Caroline Okoth, Mr. Mike Kamau, Mr. Bramwel Wanjala, Mr. Charles Kigen and Mr. Edward Mwangi. Their input into this study cannot be overestimated. The entire Egerton

University fraternity, the Faculty of Agriculture and the Dept of Crops, Horticulture and Soils are all sincerely acknowledged.

My thanks go to Dr. E. Mukisira, the Director KARI and Dr. F. Muyekho, the Centre Director KARI-Kakamega for allowing me time to complete this degree.

My great appreciation goes to my family, husband, Mr. Godfrey Masai and my two little angels Annette and Angela for their patience. When the going became tough, they became tougher. Thanks a lot for the time and financial support offered.

It is practically impossible to mention everybody in this thesis. I am however very grateful to all who inspired me to complete the study successfully. Their input will always be remembered and appreciated.

ABSTRACT

Drought is an important climatic phenomenon, which is the most severe limitation to maize production in developing countries. Maize (*Zea mays* L.) is particularly sensitive to water stress at reproductive stages, and breeding to improve drought tolerance has been a challenge. When drought occurs just before or during the flowering period, it results in delayed silking and a consequent increase in the length of anthesis-silking interval (ASI). Selection for a reduced ASI has shown correlated response to improved grain yields under drought. However, conventional selection has been limited by the difficulty in managing uniform experimental conditions to eliminate environmental effects and their possible interactions with the genotype. In this study, SSR (Microsatellite) molecular markers were used to identify genomic regions responsible for the expression of ASI using bulked segregant analysis (BSA). This would enable develop marker assisted selection (MAS) strategies. An F_2 population of 203 individuals was developed from a cross between drought susceptible (KCB) and drought tolerant maize landraces (GBK 032357) and they were screened under drought to categorize into drought tolerant and drought susceptible. Based on the BSA procedure, the DNA from 10 most drought tolerant and 10 most drought susceptible plants based on ASI values were used to make DNA pools. These DNA pools and DNA from parents were assayed at 109 loci. A confirmation of polymorphic candidate markers on these 20 individuals revealed four genomic regions associated with ASI. These were regions near markers p-umc2189, p-bnlg1179 and p-bnlg1014 on chromosome 1 and p-umc1542 on chromosome 2. Together, the candidate QTLs accounted for about 65% of the observed variation for ASI. Significant phenotypic correlations among flowering parameters, grain yield and yield components were observed. Overlaps between the corresponding candidate QTLs were observed. For instance, markers, p-umc2189 and p-bnlg1014 showed significant association with days to silking (DS) whereas markers, p-bnlg1179 and p-umc1542 showed significant association with both kernel number (KN) and grain yield (GY). This finding indicated pleiotropism between loci for ASI and DS, KN and GY. The generated information will enable formulation of Marker Assisted Selection for drought tolerance in maize.

TABLE OF CONTENTS

EGERTON UNIVERSITY LI

DECLARATION AND RECOMMENDATION ii
COPYRIGHT iii
DEDICATION iv
ACKNOWLEDGEMENTS v
ABSTRACT vii
TABLE OF CONTENTS viii
LIST OF TABLES xi
LIST OF FIGURES xii
LIST OF APPENDICES xiii
LIST OF ACRONYMS xiv
CHAPTER ONE 1
INTRODUCTION 1
1.1. Background 1
1.2. Statement of the problem 5
1.3. Objectives 6
1.4. Null Hypotheses 6
1.5. Justification 6
1.6. Expected outputs 7
CHAPTER TWO 8
LITERATURE REVIEW 8
2.1. Drought stress 8
2.2. Maize Landraces 9
2.3. Germplasm Characterization 10
2.3.1. Morphological Descriptors 10
2.3.2. Agronomic Descriptors 11
2.3.3. Isozyme Polymorphism 11
2.3.4. Molecular Characterization 11
2.4. Grain yield secondary traits and QTLs in maize under drought 21
2.5. General procedures in molecular analysis 22
CHAPTER THREE 26
MATERIALS AND METHODS 26

3.1. Experiment I: Field phenotyping for drought tolerance traits....26
3.1.1. Experimental site....26
3.1.2. Plant materials....26
3.1.3. Generating the F_2 segregating population....27
3.1.4. F_2 population development....27
3.1.5. Field screening of F_2 populations for drought tolerance....28
3.1.6. Field data collection....28
3.1.7. Field data analysis....29
3.2. Experiment II: Determining molecular markers linked to ASI....29
3.2.1. Experimental site....29
3.2.1. Genomic DNA extraction....29
3.2.2. DNA purification....31
3.2.3. Estimation of DNA quantity and quality....31
3.2.4. Polymerase chain reactions (PCR)....32
3.2.5. Quantitative trait analysis....34
CHAPTER FOUR....41
RESULTS AND DISCUSSIONS....41
4.1. Experiment I: Field phenotyping for drought tolerance traits....41
4.1.1. Correlation between traits in maize under drought....41
4.1.2. Days to anthesis (DA)....41
4.1.3. Days to silking (DS)....43
4.1.4. Anthesis-silking interval (ASI)....43
4.1.5. Leaf rolling (LR) and leaf senescence (LS)....44
4.1.6. Tassel size (TS) and plant height (PH)....45
4.1.7. Yield and yield components....45
4.18. Frequency distribution of traits measured in F_2 under drought....47
4.2. Experiment II: Determining molecular markers linked to ASI....49
4.2.1. Optimization of PCR conditions for SSR primers....49
4.2.3. Confirmation of candidate markers....57
4.2.4. ASI candidate QTLs....57
4.2.5. Co-segregation of candidate QTLs with ASI in F_2 progeny....60
4.2.6. Phenotypic correlation and pleiotropism....62
4.2.7. Genetic effects....64

CHAPTER FIVE **67**
CONCLUSIONS AND RECOMMENDATIONS **67**
5.1. Conclusisons 67
5.2. Recommendations 67
CHAPTER SIX **69**
REFERENCES AND APPENDICES **69**
6.1 REFERENCES 69
6.2 APPENDICES 81

LIST OF TABLES

Table 1. Flowering data and ASI values in days of plants selected for BSA 35
Table 2. List of primers screened on parental and F_2 progeny DNA 36
Table 3. Pearson's correlation coefficients (r) of F_2 under drought 42
Table 4. QTL data for ASI from F_2 of KCB x GBK 032357 59
Table 5. Multivariate QTL analysis for ASI 59
Table 6. Co-segregation of candidate QTLs with ASI 61
Table 7. Linear regression QTL data for DS, KN and GY 63
Table 8. Multiple regression QTL data for DS, GY and KN 63
Table 9. Genetic characteristics of QTLs for ASI under drought 65

LIST OF FIGURES

Figure 1. Relationship between ASI and GY in F_2 under drought.....44
Figure 2. Relationship between KN and GY under drought.....46
Figure 3. Frequency distribution of measured traits 48
Figure 4. Results from optimization for $MgCl_2$ and agarose 50
Figure 5. PCR products for template DNA optimization 50
Figure 6. PCR products for *Taq* DNA polymerase optimization.....52
Figure 7. PCR products for dNTP optimization experiment.....52
Figure 8. Primer optimization PCR products.....54
Figure 9. PCR products from the DNA samples selected for analysis. 54
Figure 10. PCR products of 10 primers at BSA.....56
Figure 11. PCR products generated by marker *p-umc1542* 58
Figure 12. PCR products generated by primer *p-umc2189.*.....58

LIST OF APPENDICES

Appendix 1. Primer screening plate layout 81
Appendix 2. Primer screening plate photos 82
Appendix 3. Primer Confirmation Photos 84
Appendix 4. Scored molecular data 86
Appendix 5. ANOVA for ASI, DS, KN and GY QTL. 90

LIST OF ACRONYMS

AFLP	Amplified Fragment Length Polymorphisms
ASI	Anthesis-Silking Interval
BSA	Bulked Segregant Analysis
CIA	Chloroform and Isoamyl Alcohol
CIMMYT	International Centre for Wheat and Maize Improvement
CTAB	Cetyltrimethylammonium Bromide
DA	Days to Anthesis
ddH_2O	Double Distilled Water
DH	Doubled Haploids
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleoside Triphosphates
DS	Days to Silking
EDTA	Ethylenediaminetetra-acetic Acid
EPP	Ears Per Plant
EtBr	Ethidium Bromide
FAO	Food and Agricultural Organization
GoK	Government of Kenya
GxE	Genotype by Environment Interaction
GY	Grain Yield
HCl	Hydrochloric Acid
ISSR	Inter-Simple Sequence Repeats
KARI	Kenya Agricultural Research Institute
KN	Kernel Number
LR	Leaf Rolling
LS	Leaf Senescence
M.E	Mercaptoethanol
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
OD	Optical Density
PCR	Polymerase Chain Reaction
PH	Plant Height (cm)

PVP	Polyvinylpolypyrrolidone
QTL	Quantitative Trait Loci
RAPD	Randomly Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphisms
RILs	Recombinant Inbred Lines
RPM	Revolutions Per Minute
SDS	Sodium Dodecyl Sulfate
SDW	Sterile Distilled Water
SSR	Simple Sequence Repeats
TBE	Tris-Borate-EDTA Buffer
TE	Tris-EDTA Buffer
TRIS	Tris (hydroxymethyl) Amino Methane
TS	Tassel Size
UV	Ultra Violet Light

CHAPTER ONE

INTRODUCTION

1.1. Background

Maize is a member of the grass family Poaceae (Gramineae) which represents a range of complexities in structure and genomes. Members of this family range from diploid species with a genome size of 415 Mb in rice to 16000 Mb in hexaploid wheat (Beyene *et al.*, 2005). Maize belongs to the genus *Zea* which has four species (*Zea diploperennis, Zea perennis, Zea luxurians* and *Zea mays*). Cultivated maize is *Zea mays* whereas the other three species are its wild relatives commonly known as teosintes. There are four sub-species identified in *Zea mays* i.e. *Z. mays ssp. huehuetenangensis, Z. mays ssp. mays, Z. mays ssp. mexicana* and *Z. mays ssp. parviglumis* which is thought to be the progenitor of cultivated maize (*Z. mays ssp. mays*). All the members of the genus *Zea* are native to Mexico and Central America (Beyene *et al.*, 2005)

Maize is a diploid (2n=20) and, its chromosomes contain 2.5 billion base pairs thus lying in the middle of grass genome size and complexity (Gaut *et al.*, 2000) . It is a tall grass having between 8 to 44 leaves and exhibiting diversity in terms of number, size, shape and colour of ear (seed bearing structure). Great diversity also exists in plant seed characteristics, resistance to diseases, insect pests and tolerance to various environmental stresses (Dowswell *et al.*, 1996). This is because maize is one of the crop species with the highest level of molecular polymorphisms. According to Henry and Damerval (1997), nucleotide diversity (the average sequence of divergence between any two individuals for a given locus) for certain loci has been reported to be over 5%. Maize molecular diversity has been shown to be between 3-10 fold higher than that of other domesticated grass crops. This, according to Beyene *et al.* (2005) can be attributed to several factors. First is the diversity of environments, culture, production systems and the type of consumption, which result in the development of different maize types throughout the world. Secondly, being monoecious, maize has a high level of out-crossing thus favoring continuous gene exchange between neighboring plants and or wild relatives. Thirdly, there are extensive chromosomal duplications in maize, which provide new mutational opportunities for creating greater phenotypic variability. Lastly, the presence of transposons and retrotransposons result in rearrangements or insertions and deletions thus changing the amount of DNA in the genome.

This natural diversity enhances selection for improvement of economically important traits and the polymorphisms make maize a model crop for studying the biology of cereals. However, genetic erosion occurs because of breeding and this reduces the genetic variability within a population making it susceptible to biotic and abiotic stresses. In such cases, landraces and wild relatives have been shown to offer novel genes for breeding programs (Perales, 2003).

Maize is cultivated in both temperate zones such as the US Corn Belt, the Mediterranean area and in the tropics (Passioura, 1996). It grows between 50° N and S of the equator; from sea level to 3000 m asl; under heavy rainfall and in semi-arid conditions; in cool and in hot climates. Its growing cycles range from 3 to 13 months (CIMMYT, 2000). Maize, which is cultivated mainly for its grain, is the third most important crop worldwide after wheat and rice and accounts for about 20% of the total area under cereals (FAO, 2005). The world production is about 720 million Mt from 150 million hectares of which about 100 million hectares is in the developing world (Pandey, 2000). Maize is used as food and feed and sometimes as a cash crop for industrial purposes such as wet milling (e.g. starch and oil), dry milling (e.g. meal and flour), and fermentation and distillation (e.g. alcohol and whisky) in both developed and developing countries. It dominates the diets of the rural and urban-poor, providing over 50% of staple calories within Eastern and Southern Africa (FAO, 2002).

In Kenya, maize is the leading staple cereal for about 90% of the population (Kitavi *et al.*, 2005), growing over a wide range of agro-ecological zones, from sea level to about 2100 m asl. The average annual rainfall varies from 250 mm to more than 2000 mm (Mutinda *et al.*, 2004). Up to 82% of land in Kenya is classified as arid and semi-arid lands (ASALs) with only 18% being of high potential (GoK, 2004). Maize is produced on about 1.6 million hectares of which 70-90% is grown by small-scale farmers (Wekesa *et al.*, 2003; FAO, 2005) with an estimated annual production of about 2.38 million mt. According to Corbett (1998), the maize production area in Kenya has been classified into six agro-ecologica! zones based on growing season variables crucial to maize production such as altitude and climate. First, the lowland tropics (coastal strip and adjoining inland area) with elevation between 0->700 m asl, Mar-Aug rainfall between 300-550 mm and Mar-Aug mean temperatures between 25.4-25.8 ^{0}C. Second is the dry mid altitude (in the south east) with elevation between 700-1400 m asl March-Aug rainfall of between 300-550 and mean temperature of 22.0 ^{0}C. Third is the dry transitional zone (in the south east) with elevation between 1100-1700 m asl, mean

rainfall (Mar-Aug) of <550 and mean temperature of 19.7 ^{0}C. Fourth is the moist transitional zone (to the east and west of highland tropics) with elevation between 1200-2000 m asl, mean rainfall (Mar-Aug) of >500 and mean temperatures of 19.7 ^{0}C. Fifth is the highland tropics zone with elevation between 1600-2900 m asl, average rainfall (Mar-Aug) ranging from 550-1000 mm and mean temperature between 13.8-16.6 ^{0}C. Lastly, sixth is moist mid altitude (around Lake Victoria) elevated at 1200-2000 m asl, mean rainfall (Mar-Aug) >500 and mean temperature of 22.1 ^{0}C (Corbett, 1998).

Across the six maize agro-ecological zones, farmers' actual realization of yield is about 2 mt/ha. The moist transitional zone is the most productive accounting for 85% of maize produced in Kenya, with an average yield of about 2.7 mt/ha up to 5 mt/ha (Hassan *et al.*, 1998a). The dry transitional and the moist mid-altitude are the medium yielding areas with average yields of a little over 1 mt/ha, whereas the lowland tropics and the dry mid altitude zones are regarded as the low yielding areas with average yields of around 1 mt/ha. The mean yields in national performance trials range between 6.0-8.0 mt/ha (FAO, 2005). The yields from small-scale farmers' fields averaging about 1.2 mt/ha are therefore too low a comparison to the expected yield. This yield disparity has been attributed to several constraints on maize production (Republic of Kenya, 2001). Surveys done by gathering farmers' views (IRMA, 2001; De Groote *et al.*, 2001), indicate that inadequate and erratic or unfavourable distribution of rainfall is a priority problem in all agro-ecological zones. Other constraints are low soil fertility; parasitic weeds (*striga* species); diseases e.g. streak virus, gray leaf spot, blight, rusts and head smut; insect and animal pests and storage pests (Bonhof *et al.*, 2001; Songa *et al.*, 2002). Also important are other production factors such as lack of cash, credit and tools; prohibitive cost of fertilizers; lack of technical know-how; poor extension services and problems with hybrid maize seed such as high cost, poor quality and low availability (IRMA, 2002). The significance of these factors however varies across the maize agro-ecological zones.

According to Tuberosa (2006), drought is the most devastating of all the abiotic stresses that curtail crop productivity. Coupled with low soil fertility, they form the most important limiting factors to maize production. Average yield losses due to drought are estimated to be 17% but may go up to about 70% (Edmeades *et al.*, 1994; 2000). However, losses of up to 100% have been recorded in some parts of Kenya due to frequent and unpredicted droughts, leading to periodic famines in such areas (KARI, 2004). The use of drought tolerant cultivars is therefore the only affordable option for small-scale farmers (Byrne *et al.*, 1995). However,

in most developing countries, farmers in high yielding and high input conditions have been a more attractive target for the private seed sector than resource-poor farmers. As a result, breeding for abiotic stress tolerant maize is ignored (Banziger *et al.*, 2000). In addition, breeding for tolerance to drought has been considered difficult because of the tremendous variability in the intensity and timing of drought events and a diversity of drought tolerance mechanisms used by the plant (Nguyen *et al.*, 2004). Consequently, very few varieties have been developed for moisture stressed areas and even these few varieties have failed to perform well under stress conditions. For instance, the existing maize varieties (Katumani Composite B and the Dry-Land Composite) developed for the dry mid-altitude ecologies of Kenya, are not very productive under conditions of moisture stress and low soil fertility (KARI, 2004). This has led to the continued cultivation of the local maize landraces.

Landraces are populations that are the product of selection carried out by farmers, either deliberately or not, continuously over many generations. Their importance as major components in plant genetic diversity is well established worldwide. In Kenya, farmers especially in the moisture stressed areas of low land tropics and dry mid-altitude zone prefer local landraces for their yield stability (Hassan *et al.*, 1998b). For instance, farmers on the Kenyan coast grow five to six varieties. Studies by CIMMYT (2003) showed that Kanjerenjere, a yellow landrace is grown because "even when rains are variable, you get something" though it yields poorly and is susceptible to storage pests. Another landrace, a dark purple Mdzihana was found to be a good yielder and resistant to storage pests but vulnerable to erratic water conditions. Therefore, farmers would plant both along with four or five other landraces having different agronomic and consumption characteristics. These unique traits and adaptations are attributed to the wide genetic base of landraces, which acts as a buffer against biotic and abiotic stresses. These local varieties are preferred by farmers not only for their good agronomic qualities such as drought tolerance, resistance to pests and diseases, performance under low inputs, early maturity, resistance to shattering, among others, but also for their end use qualities. These include attributes such as superior storage characteristics, and superior taste (Hassan *et al.*, 1998b; Songa *et al.*, 2002). Such landraces also offer important genes for drought tolerance that can be incorporated in maize breeding programmes.

Despite the importance of crop germplasm found in landraces, few of the world's landrace collections have been systematically evaluated (Perales, 2003). In Kenya, scientists from

International Centre for Wheat and Maize Improvement (CIMMYT) and Kenya Agricultural Research Institute (KARI) have collected and stored some of the local landraces at the Gene Bank of Kenya. These have however not been characterized. Information on both qualitative and quantitative morphological traits of these landraces may be useful in maintaining their genetic variability. Characterization of these landraces and refining selection procedures is therefore an important step.

Phenotypic selection for quantitative traits has been limited by the increase in environmental variance relative to genetic variance, which decreases yield heritability (Campos *et al.*, 2004). Therefore, selection for secondary traits such as anthesis-silking interval (ASI) which is correlated with grain yield in maize and has a relatively high heritability, has been shown to increase selection efficiency under drought conditions (Gebre, 2005). In maize, male flowering (anthesis) and female flowering (silking) are easy to measure in the field and ASI can be calculated with precision. However, the major problem of classical selection is the difficulty in management of experimental conditions (Ribaut *et al.*, 1996). Genomics, (study of the function and structure of specific genetic sequences) accompanied by high through-put laboratory-based analysis of DNA (Tinker, 2002) has therefore been considered key to comprehending gene-phenotype associations at the level of candidate genes and sequences. This is critically important for quantitative traits such as drought tolerance, whose performance is regulated by many loci and is subject to multiple genotype by environment (GxE) and gene by gene interactions (epistasis) (Campos *et al.*, 2004). Therefore, the use of molecular markers to improve the efficiency of breeding towards better drought tolerance may provide a working alternative.

1.2. Statement of the problem

There is low productivity of maize in Kenya, with the majority small-scale farmers realizing yield as low as 1.2 mt/ha as opposed to the expected 6.0-8.0 mt/ha. The low maize productivity and production has largely been associated with frequent droughts and lack of superior drought tolerant cultivars (Mutinda *et al.*, 2004), among other factors. There is also increased settlement in marginal areas, which has led to the introduction of maize in these areas where hybrid varieties do not perform well. This has therefore resulted in a fall in the production volumes leading to periodic imbalances between food production and demand since mid 1970's (Mati, 2000). The local landraces in the dry mid-altitude and the coastal agro-ecological zones of Kenya have shown great potential for drought tolerance. However,

information on these landraces is lacking. It is therefore necessary to investigate the molecular basis of such valuable traits exhibited by these maize landraces.

1.3. Objectives

1.3.1. General Objective

To contribute towards food security through the identification of molecular markers that would expedite the selection process for drought tolerance in maize.

1.3.2. Specific objectives

1. To identify SSR polymorphisms for ASI between the selected parents and the F_2 population
2. To determine correlations between morphological traits such as ASI, grain yield and SSR markers
3. To estimate co-segregation of SSR candidate markers and ASI

1.4. Null Hypotheses

1. There were no SSR polymorphisms for ASI between the selected parents and F_2 population.
2. There was no significant correlation between SSR markers and morphological traits.
3. There was no cosegregation between SSR markers and ASI

1.5. Justification

Due to global warming, droughts are getting longer and more frequent hence the patterns and trends of rainfall can no longer be predicted (Hillel and Rosenz-weig, 2002). These erratic rainfall patterns significantly affect maize production. The increasing population has also forced settlement in marginal areas, which are untraditional for maize production. Low and erratic rainfall, high temperatures and high evapo-transpiration rates characterize such areas. The majority of the inhabitants of these marginal areas are resource-poor farmers whose major daily calorie source is maize.

Most of the improvement work done in the past emphasized the high potential areas and as a result, many improved varieties/hybrids with improved management practices have been made available to farmers in these areas. In contrast, in the drought stressed areas there are only afew improved varieties that are grown, and under low input levels and management. These varieties have also been shown to fail in performance as a result of the changing climatic conditions. Breeding activities for moisture stressed areas are handicapped by a

narrow genetic base of available germplasm and not well established selection criteria for drought tolerance.

Unavailability of suitable drought tolerant varieties has contributed significantly towards such a yield gap as there is between the national research trials and the actual yields realized by small-scale farmers. Studies have revealed that maize landraces have additional genetic diversity for maize breeding (Perales, 2003). The landraces also perform better under stress conditions. There is therefore an urgent need to refine selection strategies for drought tolerance and to convert promising genotypes into (drought tolerant) varieties. Characterization of the available germplasm therefore is important.

Most characterization and selection procedures have however mainly been based on the phenotype. In this case, it is difficult to separate the genetic effects from the environmental effects for many plant traits because the testing procedures may be tedious, lengthy or expensive. Besides, many important traits including drought tolerance are quantitatively inherited. In this case, it is difficult or impractical to determine at phenotypic level, whether or not a specific gene, or genes, is present (Campos *et al.*, 2004). This therefore calls for a more precise way of determining the genotype using molecular markers through marker assisted selection (MAS). Molecular characterization is therefore crucial as it forms the first step towards MAS. In this study, simple sequence repeat (SSR) molecular markers have been used to identify genomic regions responsible for anthesis-silking interval under drought through bulked segregant analysis (BSA), with an aim towards MAS.

1.6. Expected outputs

1. SSR markers linked to ASI in local maize landraces established and documented
2. A basic marker assisted selection protocol for ASI under drought
3. Dissemination of results through peer reviewed publications
4. MSc Plant Breeding degree

CHAPTER TWO

LITERATURE REVIEW

2.1. Drought stress

Drought is one of the major limitations to food production worldwide and is endemic particularly in the semi-arid tropics (Vinod *et al.*, 2006). It is a common phenomenon in the developing countries and after low soil fertility represents the second most important cause of yield loss in maize (Ribaut *et al.*, 1996). Crops experience drought stress when their water requirements at one or more growth stages exceed the available water within the root zone. According to Frova *et al.* (1999), drought exerts a multidimensional stress that affects plants at various levels of their organization over space and time so that the responses to it are often unpredictable. It is evenly distributed across the world's major maize production regions and is a particularly severe problem for slightly more than one fifth of the tropical and subtropical maize planted in the developing world (Pingali and Pandey, 2000).

Drought stress reduces yields through three mechanisms; first, through reduced canopy absorption of photosynthetically active radiation (PAR) by temporary leaf wilting, leaf rolling or early leaf senescence; secondly, by reducing the efficiency of PAR utilization and thirdly, through the reduction of harvest index (Blum, 2000). According to Ribaut *et al.* (1996), the plant's responses to water stress depend on its metabolic activity, morphology and stage of growth. Water stress occurring at any plant developmental stage is undesirable. However, in crop plants, drought stress during flowering is considered catastrophic in terms of productivity (Bolanos and Edmeades, 1996; Ribaut *et al*, 1996; Laffite, 2002).

Maize has been shown to be unusually susceptible to drought around the flowering period and this leads to a delay in the silk emergence resulting in the increase in anthesis-silking interval (ASI) (Li *et al.*, 2002). This delay in silking is as a result of reduced photosynthesis due to drought (Banzinger *et al.*, 2000). ASI has been shown to be highly negatively correlated with grain yield, in particular kernel number and ear number per plant. This correlation is however weak under well watered conditions. This indicates that variation in grain yield under stress is due to variation in ear setting processes related to biomass partitioning at flowering (Monneveux, 2006). Frova *et al.* (1999) also indicated that a halt in embryo growth under water stress is due to a decrease in sucrose flux and an altered carbohydrate metabolism in ovaries. Earl and Davis (2003) further suggested that the failure to utilize available sugars at a low ovary water potential would severely inhibit assimilate flux to the ears and render kernel set highly vulnerable to water deficit during pollination.

The synchrony between male and female flowering dates is therefore strongly associated with yield increase under drought (Edmeades *et al.*, 2004). The devastating effects of drought on crop species could be overcome by exploring existing genetic variation for drought tolerance in order to develop genotypes better adapted to cope with water stress (Blum, 1996). Traditional cultivars (landraces) have been shown to have unique stress tolerance strategies which may be appropriate to specific growing conditions (Blum, 2000). For instance, Skovmand *et al.* (2001) indicated that some landraces have very high chlorophyll concentration, which may increase leaf photosynthetic rate even under stress conditions.

2.2. Maize Landraces

According to Perales (2003), a complete understanding of the dynamics affecting traditional (landrace) crop populations is needed if conservation and utilization of genetic resources is to be successful. Landraces are the cultivated maize material with the highest genetic variation as well as with the best adaptation to the natural and anthropological environment where they have evolved and represent an irreplaceable bank of highly co-adapted genotypes (FAO, 2006). Over the last few decades, the agricultural scenery has profoundly changed and the subsistence mixed farming unit is now transformed into an intensive monoculture (Bertolini *et al.*, 1998). At present, populations of landraces are found growing in farmers' fields under peculiar agricultural situations such as heterogeneous soils, economic isolation, marginal agricultural conditions, and small fields traditionally managed according to low input agronomic practices with production exclusively addressed to human consumption (Lucchin *et al.*, 2003).

Whereas agricultural environments together with traditional diets allow for preservation of some landraces *in-situ*, many locally cultivated populations have been lost before it was realized that they were important sources of germplasm (Perales, 2003). Many of the existing landraces have also undergone both natural and artificial selection since their introduction. For instance, a local cultivar "Machakos local white" cultivated in the dry mid-altitude zone of Kenya is thought to be a mix of genetic material from Katumani Composite B (an improved open pollinated variety developed for the zone) and local landraces (Songa *et al.*, 2002). After years of lack of interest in the "old local varieties", this valuable source of maize germplam has been rediscovered and exploited as a niche crop suitable for cultivation in marginal lands. Maize breeders have recently become more aware of the need to both maintain genetic diversity among hybrid varieties and improve the management of genetic resources through the conservation of landraces. This is not only to preserve important

sources of genetic material for breeding, but also to allow their use as essential components of sustainable agriculture (Lucchin, *et al.*, 2003).

In spite of their importance, little information is available on landraces (Pallotini, 2002). Research is therefore needed on the conservation biology of the remaining landrace populations in order to plan both *in-situ* and *ex-situ* conservation (Perales, 2003) and their use in maize breeding programs. Genetic characterization of the existing landraces therefore represents an essential step (Pallotini, 2002), with several countries embarking on characterizing their local landraces using molecular markers. For instance, molecular characterization of an Italian landrace of flint maize (*Zea mays* var. indurata) "Nostrino di storo" was carried out using different PCR- based markers (Barcaccia *et al.*, 2003) and a linkage map for the same maize landrace constructed based on a pseudo-testcross strategy using dominant multi-locus PCR- derived markers (Barcaccia *et al.*, 2000; 2005).

2.3. Germplasm Characterization

Germplasm characterization refers to the observation, measurement and documentation of heritable plant traits in a collection, thus allowing for identifying and classifying accessions and building a catalogue of descriptors with biological information for management or direct use in agriculture (FAO, 2006). Germplasm characterization of plant accessions deposited in gene banks has been limited and is probably a major cause for the limited use of accessions in plant breeding programs. Characterization can be achieved through:

- Morphological descriptors
- Agronomic descriptors
- Isozyme polymorphisms
- Molecular marker polymorphism

Morphological, agronomic and isozymic descriptors reveal differences at gene expression level (protein level) hence are influenced by the environment whereas molecular markers reveal differences at the DNA level and are not influenced by the environment (de Vicente *et al.*, 2006).

2.3.1. Morphological Descriptors

Most characterization relies strongly on morphological descriptors. Though these are reliable, easy to study and are relatively low cost to evaluate, the use of these descriptors however presents some limitations. According to Ferreira (2006), these include limited

polymorphism that lower the potential success of an extended classification approach and thus would require a high number of descriptors in order to compensate for the small number of morphotypes; potential influence on the phenotype, makes the process of evaluation and information exchange even more complex. This increases the chances of false positives when the environment affects specific morphotypes

2.3.2. Agronomic Descriptors

Germplasm characterization based on agronomic traits on the other hand is particularly useful in crops of economic importance. The amount of data related to agronomic traits that is available by crop germplasm evaluation is however limited. Due to reasons that include relatively high costs and difficulties of large-scale experimental trials, the use of agronomic evaluation to characterize germplasm collection is far from the actual need of uncovering the phenotypes of agronomic interests in accessions of a collection (Ferreira, 2006).

2.3.3. Isozyme Polymorphism

Isozymes are enzymes with multiple molecular forms, which occur within a single organism or even within a cell (Muoki *et al.*, 2008). These isomeric structural forms have the same molecular weight and specificity for the substrate, catalyzing the same biochemical reaction but differing in structural configuration and properties (Muoki *et al.*, 2008). Isozymes have been widely used in the determination of genetic diversity and identification of plant varieties (Ilarslan *et al.*, 2001). They have been used in various aspects of breeding such as marker assisted selection (Ghesquiere *et al.*, 2000) and confirming hybridity (Isshiki and Taura, 2003). However, like phenotypic markers, these protein-based marker systems only sample actively expressed regions of the genome and therefore their use in certain aspects of plant breeding and genetics as co-dominant neutral genetic markers and their potential use in genetic mapping is often limited by lack of adequate polymorphism (Wachira, 1996).

2.3.4. Molecular Characterization

This refers to the detection of variation due to differences in either DNA sequences or specific genes or modifying characters (FAO, 2006). Germplasm collections can undergo either structural molecular characterization, which is based on the building blocks of the DNA sequences, or functional molecular characterization, which is based on the identification of genes and their functions. Molecular characterization therefore permits

access to the raw materials-genes for nearly all the objectives of today's and tomorrow's breeding programs. According to de Vicente *et al.* (2006), the information gathered from structural characterization not only provides increased clarity on existing genetic diversity and its organization in individuals, but also determines sample and population organization that may ultimately form the basis for functional characterization.

DNA-based assays have therefore revolutionized and modernized the ability to characterize genetic variation. The first advantage of molecular techniques is their capacity to detect genetic diversity at a higher level of resolution than other methods. Furthermore, DNA-based assays are robust, speedy and information may be obtained from small amounts of plant material at any stage of development. In addition, they are not affected by environmental conditions (Lanteri and Barcaccia, 2006). The use of nuclear DNA markers potentially explains the portion of the genome that can be examined (Lanteri and Barcaccia, 2006; Muoki *et al.*, 2008). Many different molecular techniques are presently available and each differs in the information content generated (Wachira *et al.*, 2001).

2.3.4.1. Molecular Markers

Molecular markers are identifiable DNA sequences found at specific chromosomal locations on the genome and transmitted by the standard laws of inheritance from one generation to the next (de Vicente *et al.*, 2006). They could be expressed DNA regions (genes) or DNA segments that have no known coding functions and therefore do not have any biological effects and normally assumed to represent neutral loci unaffected by selection (de Vicente *et al.*, 2006). These markers reveal variation in genomic DNA sequences.

New technological developments in molecular biology have expanded the range of DNA polymorphic assays for genetic mapping, marker assisted plant breeding, genome fingerprinting and for investigating genetic relatedness (FAO/IAEA, 2002). These technologies include restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs), inter-simple sequence repeats (ISSRs), sequence characterized amplified regions (SCARs), cleaved amplified polymorphic regions (CAPs), single strand and double conformational polymorphisms (SSCPs; DSCPs), single-tagged sites (STSs), allele-specific amplicons (ASAs), sequence taq sites (STS) and expressed sequence taqs (ESTs).

Ideal molecular markers should be co-dominant in expression, economical to use, highly polymorphic, easily assayed, multifunctional, readily available, can be multiplexed and be

automated easily (FAO/IAEA, 2002). Dominant markers identify only one allele thus offering no possibility of discriminating between homozygous and heterozygous individuals whereas co-dominant markers identify both alleles in heterozygous individuals and are therefore more flexible and supply more robust and comparable data (Karp, 2002). These markers can be classified as hybridization-based or PCR-based markers.

2.3.4.1.1. Hybridization-based markers

Restriction fragment length polymorphisms (RFLPs)

According to FAO/IAEA (2002), genetic variability can be readily assessed using restriction enzymes. These are nucleases produced by varieties of microorganisms that cleave DNA at particular restriction (target) sites that are made up of specific base sequences in the DNA. RFLPs refer to the variation in the length of DNA fragments produced by a specific restriction endonuclease from genomic DNAs of two or more individuals in a species (Kahl, 2001). The RFLP procedure involves digesting DNA from two genetically distinct individuals with restriction enzymes such as Pst1, EcoR1 Mse1 followed by separation through gel electrophoresis. The DNA is then transferred to a membrane by Southern blotting and hybridized to a radioactively labeled DNA clone or probe (Wachira, 1996).

These markers have a co-dominant mode of inheritance, minimal or lack of environmental and developmental influences on their expression, large numbers, reusable filters and a high level of genetic variation compared to morphological and biochemical markers. However, RFLPs are expensive, labour-intensive, have low throughput and are not omnipotent solutions to the limitations of classical breeding, although they can be used to accelerate many breeding endevours (FAO/IAEA, 2002). RFLPs were the first generation of molecular markers, however, the invention of PCR to amplify short segments of DNA gave rise to a second generation of faster and less expensive PCR-based markers.

2.3.4.1.2. Polymerase chain reaction (PCR)-based markers

PCR is an *in-vitro* procedure for the enzymatic synthesis of DNA by which a particular segment of DNA can be specifically replicated (FAO/IAEA, 2002). The process involves use of two oligonucleotide primers that flank the DNA fragment of interest. Amplification is achieved by a series of repeated cycles of heat denaturation of the DNA, annealing of primers to their complementary sequences on the target template DNA and extension of the annealed primers using a thermo-stable DNA polymerase enzyme (Holton *et al.*, 1999; Zidenga, 2004).

The efficiency of PCR has been tremendously improved in the advent of identification of thermo-stable DNA polymerases like the Taq polymerase isolated from the bacterium *Thermus aquaticus* which is now extensively used (Rahman *et al.*, 2002). The use of such thermo-stable polymerases eliminates the need to add fresh enzyme after the denaturation step, which is characterized by a high temperature (90-95°C). In addition, the higher temperature optimum of this enzyme (70-75°C) significantly increases the specificity, yield and length of targets that can be amplified. The method has also been automated making it a useful tool for genetic analysis because it takes comparatively shorter time than Southern blotting, requires extremely small quantities of crude DNA for amplification and its products can be visualized on ethidium bromide stained agarose gels (Wachira, 1996).

Several PCR-based marker systems have been described (Zidenga 2004). However, those commonly used in plant breeding include RAPDs, AFLPs, SRRs and ISSRs (FAO/IAEA, 2002).

Random amplified polymorphic DNAs (RAPDs)

This assay is based on the amplification of genomic DNA with single primers of arbitrary nucleotide sequence (Kahl, 2001).

The RAPDs procedure involves binding a single species of primer to the genomic DNA at two different sites on opposite strands of the DNA template. If these priming sites are within an amplifiable distance (2000 bp) of each other, a discrete DNA product is produced through thermocyclic amplification. The presence of each amplification product identifies complete or partial nucleotide sequence homology between genomic DNA and the oligonucleotide primer at each end of the amplified product.

As molecular markers, RAPDs have the advantages of giving quick results, being fairly cheap, not requiring sequence information, having good polymorphism and being easily automated. However, RAPDs are dominant in expression and therefore cannot distinguish between heterozygous and homozygous individuals such as those likely to be found in an F_2 population. They also have low reproducibility within and between laboratories and cannot be used across populations or species (FAO/IAEA, 2002). RAPDs have been widely used in the assessment of the degree of polymorphism and cultivar identification, genetic diversity, phylogenetic and pedigree studies, genetic linkage mapping and identification of markers linked to genes controlling important traits in crops, clone identification, among others (Hacket *et al.*, 2000).

Amplified fragment length polymorphisms (AFLPs)

AFLPs combine the specificity of restriction analysis with PCR amplification (FAO/IAEA, 2000). The sequence variation detected is primarily the same as that detected with RFLP analysis but has a higher multiplex ratio (higher number of polymorphisms detected per analysis) (Wachira *et al.*, 2001).

The AFLP procedure involves the cleavage of DNA with two restriction enzyme; a rare 6-base recognition sequence restriction enzyme e.g. EcoR1 or Pst1 and a frequent 4-base recognition sequence restriction enzyme e.g. Mse1. Two double-stranded oligonucleotide adapters with overlapping ends specifically designed for annealing to the restricted DNA are then ligated. This produces a library of fragments with known DNA sequences at their termini. Two rounds of PCR amplification are then used to amplify the restriction fragments. Amplification fragments are thereafter separated on sequencing gels and are either silver stained or radioactively detected. In general, 75-150 fragments are amplified with each primer combination and as each fragment theoretically represents a unique site, the proportion of the genome assayed with each primer combination is much higher than any other DNA analysis approach (Vos et al., 1995; FAO/IAEA, 2002).

When used as molecular makers, AFLPs have the advantages of requiring small DNA quantities, not requiring sequence information and being easily automated. However, these markers are also expressed in a dominant manner and therefore unable to distinguish between heterozygotes and homozygotes. AFLPs have nonetheless found wide use in all aspects of germplasm management and crop breeding (Wachira *et al.*, 2001).

Inter-simple sequence repeats (ISSRs).

This is a relatively new approach with the potential to serve as a powerful tool for genome analysis. The ISSR procedure involves the PCR amplification of regions between adjacent, inversely oriented microsatellites using a single simple sequence repeat (SSR)-containing primer (Kahl, 2001). Specificity of the primers is increased by anchoring some non-repeat nucleotides at the 3' or 5'ends. ISSR molecular markers can be applied for any species that contains a sufficient number and distribution of SSR motifs (Mondal, 2002). Due to the greater length of ISSR primers, they may show greater repeatability and stability of map position in the genome when comparing genotypes of closely related individuals (Muoki *et al.*, 2008).

When used as molecular markers, ISSRs have the advantages of not requiring genomic sequence data, being polymorphic, robust and easily automated. These markers are however

species specific and are expressed in a dominant manner thus are unable to distinguish between homozygous and heterozygous individuals in a segregating population (FAO/IAEA, 2002). These markers have been widely used to genetically characterize various plant species (Mondal, 2002) and to study the reproductive biology of plant species (Muoki *et al.*, 2008).

Simple sequence repeats (SSRs)

Simple sequence repeats also called microsatellites are a class of repetitive elements consisting of motifs 2-4 base pairs (bp) in length (FAO/IAEA, 2002). They are abundant, occurring on average after 6-7 kb (Cardle *et al.*, 2000), dispersed throughout the fungal, plant and human genomes and show high levels of polymorphisms than any other genetic markers (Kahl, 2001). The DNA sequences flanking such microsatellites are normally highly conserved within individuals of the same species and from these conserved sequences, forward and reverse primers can be designed for PCR amplification of the repeat motifs. This makes SSR analysis amenable to automation and multiplexing thus allowing genotyping to be performed on large numbers of lines and enhancing simultaneous analysis of multiple loci. The markers are also inherited in a co-dominant manner hence are able to distinguish between homozygotes and heterozygotes in a segregating population (FAO/IAEA, 2002). This therefore makes SSRs markers of choice in germplasm characterization. It is the variation in the number of times that these elements are repeated that provides the basis for polymorphisms that can be used in genetic studies.

Among the different classes of molecular markers available for evaluating genetic diversity, microsatellites are well known for their potentially high information content and versatility (Wachira *et al.*, 2001). These markers have been associated with high levels of polymorphisms in both animals and plants and have been applied widely in all aspects of genetic studies for example, in DNA fingerprinting, construction of highly informative saturated genetic maps and population genetic studies (Ngugi and Muthamia, 2004; Li *et al.*, 2003).

2.3.4.1.3. Information content of different molecular markers

According to Powell *et al.* (1996), the utility of a given marker system is a function of its marker index; (MI) = H x E; where H= diversity index and E= effective multiplication ratio (portion of polymorphic loci times the number of loci analyzed per gel). A comparison of the informativeness of the various marker systems in the tea species revealed the following descending order: AFLP>SSR>RAPD>RFLP (Wachira *et al.*, 2001; Muoki *et al.*, 2008). The

SSRs, AFLPs and RAPDs can therefore be classified as ideal genetic markers based on their informativeness, abundance and ease of use.

2.3.4.2. Importance of molecular characterization in plant breeding

Molecular characterization by itself, or in conjunction with other data, provides reliable information for assessing among other factors, the amount of genetic diversity (Perera *et al.*, 2000) and the structure of diversity in samples and populations (Shim and Jorgensen, 2000; Figliuolo and Perrino, 2004). This has also been applied in the assessment of rates of genetic divergence among populations (Maestri *et al.*, 2002) and the distribution of diversity in populations found in different locations (Perera *et al.*, 2000; Ferguson *et al.*, 2004). For instance, AFLP markers have been used to analyze diversity in Abyssinian banana (*Ensete ventricosum*) (Negash *et al.*, 2002) and SSR markers in analyzing diversity in Taro (*Colocasia esculenta*).

Molecular characterization also helps to determine the breeding behavior of species, individual reproductive successes and the existence of gene flow (the movement of alleles within and between populations of the same or related species and its consequences) (Papa and Gepts, 2003; Muoki *et al*, 2008). It also improves or even allows the elucidation of phylogeny and provides the basic knowledge for understanding taxonomy, domestication and evolution (Nwakanma *et al.*, 2003). For instance, Dean *et al.* (1999) used microsatellite markers to analyze the genetic diversity and structure of 19 sorghum accessions. Molecular markers have also been applied in examining genetic identities and relationships of Malus accessions (Hokanson *et al.*, 1998). A special and increasingly important role of molecular characterization is however in the identification of useful genes in germplasm. This is important because the valuable genes help develop varieties that will be able to meet the challenges of current and future agriculture.

Molecular characterization has benefited from several approaches resulting from advances in molecular genetics such as QTL mapping and gene tagging (Kelly *et al.*, 2003; Yamada *et al.*, 2004). Research in this field has led to the acknowledgement of the value of wild relatives in which modern techniques have discovered useful variation that could contribute to varietal improvement (Xiao *et al.*, 1996). Knowledge of molecular information in major crop species has opened up prospects for identifying important genes or variants in other crop types, particularly those that receive little attention from formal research (FAO, 2006).

2.3.4.3. Genetic linkage maps

One of the most important applications of molecular markers is in the development of linkage maps, which determine linkage relationships between genes or between genetic markers and traits of economic importance (Wachira *et al.*, 2001). Since linkage is based on meiotic recombination, genetic mapping can be accomplished in all plant species that undergo sexual reproduction. The construction of linkage maps is based on the estimation of recombination frequencies between genetic loci representing different alleles and determination in linear order of loci in linkage groups.

In deriving a linkage map, a mapping population segregating at desired loci is first developed (CIMMYT, 1996). Many types of populations can be used in linkage mapping. For developing molecular linkage maps in maize and wheat, backcross BC_1 and F_2 populations are adequate. However, for more resolution and/or to locate genes for complex quantitative traits (QTLs), more advanced populations such as recombinant inbred lines (RILs), doubled haploids (DHs) or backcross derived-isolines can be used (CIMMYT, 1996). According to Wachira, (1996) and CIMMYT (1996), five main steps normally accomplish linkage map construction:

- Analysis of segregation at individual loci
- Detection of linkage between loci
- Estimation of recombination frequencies and map distances between loci
- Assembly of loci into linkage groups
- Estimation of the linear order of loci within linkage groups

Analysis of segregation is done through a quick comparison of the alleles segregating in a population and those found in parents, then determining the genotype of the locus in each individual. Departures from Mendelian ratios are then tested by applying a Chi-square test. Based on Mendel's first law of inheritance, dominant markers segregate at a ratio of 3:1 or 1:3, whereas co-dominant markers segregate at 1:2:1 in F_2 populations (CIMMYT, 1996). Based on Mendel's second law of independent assortment, linkage is detected when the proportion of progeny inheriting recombinant combinations of any given pair of markers falls significantly below a chosen threshold. This is done through the application of the Chi-square test or the logarithm of "odds" (LOD) score test (Wachira, 1996; Xiao *et al.*, 2005).

Estimation of recombination frequencies and map distances can be achieved by counting the number of cross-overs and dividing this by the total number of meiotic events. However, this is only an approximation and more accurate methods based on maximum likelihood

methods are available which are not affected by segregation or phase ambiguous data. Map distances conventionally expressed in centimorgans (cM), are derived from recombination frequencies with the help of several mapping functions (CIMMYT, 1996).

Loci are assembled into linkage groups based on recombination frequencies so that every locus within a group is linked to at least one other locus in that group. The estimation of the linear order of loci within linkage group is then done by determining the best fitting order and a comparison of alternative possible orders. This could be achieved through the sum of pair-wise LOD scores for each pair of loci adjacent in a given order or through log-likelihood in MAPMAKER or through the unweighted least squares or weighted residual sum of squares in JOINMAP (Wachira, 1996).

2.3.4.3.1. QTL mapping

Traits of economic importance in plants mostly involve a complex genetic control, generally determined by several genes and a strong interaction with the environment (Ferreira, 2006). The genetic basis of drought tolerance for instance, which is a trait of great importance for genetic improvement of most crop plants, is typically quantitative. These polygenes or QTLs differ little from genes affecting simply inherited traits and follow the fundamental Mendelian laws of inheritance. However, individual genes involved in the control of a quantitative trait cannot be monitored directly by conventional methods (Wachira, 1996).

Genetic mapping allows for the breakdown and examination of the genetic control of a complex trait. This in turn makes it possible to identify the number of genes that are involved in the expression of that trait, to locate these genes or genomic regions on the chromosomes, measure their impact on the phenotypic variation and understand gene interaction in the manifestation of the trait (Risch, 2000). These maps make ample use of molecular markers and are constructed in order to locate genomic regions (QTLs) associated with economically important traits. The procedure involves examination of several populations segregating for alleles associated with the control of the trait using markers that cover the whole genome. The segregating populations are then submitted to extensive phenotypic evaluation through different bio-assays and replicated experiments in the field. The genotypic and phenotypic information is used to identify the genomic regions that control the traits of interests (QTL) (Li *et al.*, 2003; Zhang *et al.*, 2003; Guo *et al.*, 2008). The QTL x Environment interactions are then estimated using data obtained from different locations (Ferreira, 2006).

Whereas genetic dissection of polygenic traits has been hailed as a promising application of DNA markers, few concrete results have been published in the past years (Ribaut and Hoisington, 1998). According to Ribaut *et al.* (2002), the difficulty of manipulating quantitative traits is related to their genetic complexity, principally the number of genes involved in their expression and interaction between the genes (epistasis). This implies that several regions or QTLs must be manipulated simultaneously to have a significant impact and that the effect of individual regions is not easily identified. Consequently, in many crop species with large and complex genomes such as maize, it may be demanding to locate new markers linked to traits of agronomic importance in chromosomes. To enhance the speed and efficiency of identification of important QTLs, the bulked segregant analysis (BSA) technique can therefore be used to quickly identify potentially linked markers.

2.3.4.3.2. Bulked segregant analysis (BSA)

This is a method, which involves screening two pooled DNA samples of individuals from a segregating population originating from a single cross (Tabor *et al.*, 2000). Each pool or bulk contains individuals selected to have identical genotypes for a particular genomic region but random genotypes at loci unlinked to the selected region. Therefore, the two resultant DNA bulks differ genetically only in the selected region and are seemingly heterozygous and monomorphic for all other regions. Markers that are polymorphic between the pools will be genetically linked to the loci determining the trait used to construct the pools.

The two DNA bulks are constructed using equal amounts of DNA from 5-10 plants, from which a linkage between a polymorphic marker and a target locus is confirmed and quantified. This approach is based on linkage disequilibria that exist for markers surrounding a target locus in the individuals used to construct the pools. Thus, all individuals in a pool will share only markers that are closely linked to the trait (Brauer *et al.*, 2006).

Bulked segregant analysis has found several applications in plant breeding. Grattapaglia *et al.* (1996) used BSA in mapping of QTLs for growth and wood quality in *Eucalyptus grandis*. Mackay and Caligari (2000) used BSA to study the efficiency of F_2 and backcross populations in mapping. Brauer *et al.* (2006) used BSA for mapping several traits in yeast. BSA has also been used in positional cloning of mutations in Medaka by Martinez-Morales *et al.* (2004). Tabor *et al.* (2000) identified two RAPD markers linked to *Melampsora medusae* resistance in *Populus deltoides*. BSA therefore offers a faster way of determining linkage between molecular markers and traits of interest, thus allowing the application of such markers in marker assisted selection (MAS).

2.3.4.4. Marker assisted selection (MAS)

Marker assisted selection involves selection of plants carrying genomic regions that are involved in the expression of traits of interest through the use of molecular markers (Babu *et al.*, 2004). This avoids relying on the phenotype, which may largely be determined by environment rather than genotypic factors (Wachira, 1996; Campos *et al.*, 2004). This technique has been considered as a promising application of molecular genetics (Dekkers and Hospital, 2002; Kraakman *et al.*, 2006) that would greatly enhance the effectiveness of plant breeding programs. An important step is to find molecular markers that are located close to genes that govern the trait phenotypes of interest. If a gene of interest is tightly linked (lies near) to a particular DNA marker, it can be predicted whether the gene is present or absent even without actually visualizing the trait (Wachira, 1996).

The work of Yousef and Jurik (2001) designed to compare the efficiency of MAS and phenotypic selection in enhancing economically important quantitative traits showed that by incorporating DNA markers to traditional breeding, the programs could expedite selection and be cost effective. In the context of MAS, DNA based markers can be effectively used for two main purposes: first, in tracing favourable alleles (dominant or recessive) across generations and secondly, in identifying the most suitable individuals among segregating progeny based on allelic composition across part or the entire genome (Babu *et al.*, 2004). Molecular markers therefore provide a powerful selection tool for plant breeding particularly in situations where the phenotype is difficult to score, recessive or impossible to monitor in the presence of other genes.

2.4. Grain yield secondary traits and QTLs in maize under drought

The genetic mechanism of drought tolerance in maize is very complex. Therefore, the yield, which is the primary trait, is strongly influenced by both genotype and environment, and the interaction of genotype-by-environment is remarkable (Guo *et al.*, 2008). Secondary traits are those traits that are correlated with grain yield under stress, highly heritable, easy to measure, and stable over time, which can be used to indirectly select for grain yield (Campos *et al.*, 2004). Identification and measurement of secondary traits associated with grain yield provides a guide to specific mechanisms that contribute to grain yield under drought. Some secondary traits are associated with specific developmental stages such as flowering, while others, such as photosynthetic rate, are indicative of plant growth throughout the life cycle of the crop (Guo *et al.*, 2008). Such traits include the root phenology, male and female

EGERTON UNIVERSITY LIBRARY

flowering (days to anthesis and days to silking), leaf rolling, leaf senescence, plant height, tassel size, ears per plant, ear length, kernel number, among others (Barker *et al.*, 2005).

During the past decade, major efforts have been dedicated to the genetic dissection of drought tolerance components in maize under water stress conditions (Guo *et al.*, 2004; Jeanneau *et al.*, 2002; Quan *et al.*, 2004; Guo *et al.*, 2008). Many QTLs involved in the yield components and secondary morphological traits of interest, such as anthesis-silking interval (ASI), have been identified (Campos *et al.*, 2004; Xiao *et al.*, 2004). QTLs that control grain yield under drought stress had been studied and 5 QTLs were detected (Xiao *et al.*, 2004). Ribaut *et al.* (1997) examined the QTLs that govern the yield of tropical maize under three irrigation regimes: 5 QTLs were detected under intermediate and 4 under severe stresses, respectively (Ribaut *et al.*, 1997). However, most of the QTL mapping studies listed were done using populations developed from inbred parents but little has been done to explore the valuable traits in maize landraces. The discovery of drought tolerant genes and their closely linked markers in the landrace populations can therefore provide useful information in setting up breeding strategies.

2.5. General procedures in molecular analysis

Genomic DNA extraction

The aim of isolating plant genomic DNA is to produce high molecular weight DNA for restriction analysis, cloning or PCR amplification process. The protocols used have therefore to be rough enough to release DNA from rigid plant cell walls but gentle enough to produce DNA of high integrity (i.e. not sheared). Certain key features must therefore be observed in genomic DNA extraction (Ochanda *et al.*, 1996):

- Breaking of cell walls to release cell contents which can be achieved by grinding the plant tissue under liquid nitrogen or acid washed sand,
- Disruption of the cell membranes using a detergent e.g. SDS or CTAB so that DNA can be released into the extraction buffer,
- Protecting the DNA from the released endogenous nucleases which is achieved by use of chelating agents such as EDTA which protect DNA during isolation,
- Shortening the time between thawing of the frozen, ground tissue and the addition of the extraction buffer to prevent nucleolytic degradation of DNA,
- Careful and gentle handling of tubes in order to minimize shearing and to achieve high molecular weight DNA.

Precipitation of nucleic acids

DNA can then be concentrated by use of ethanol by allowing precipitation at low temperature (-20 or less) in the presence of moderate concentrations of monovalent cations. It is then recovered by centrifugation and is re-dissolved in an appropriate buffer at the desired concentration. Isopropanol is however mostly used in the place of ethanol though it is less volatile than ethanol and could easily be co-precipitated with solutes such as sucrose or sodium chloride. The use of isopropanol offers an advantage of requiring less volume and limitations associated with it are circumvented by washing the DNA pellet in 70-75% ethanol (Ochanda *et al.*, 1996).

Purification of nucleic acids

Removal of proteins is often carried out simply by extracting aqueous solutions of nucleic acids with phenol and/or chloroform to inactivate and remove enzymes. The most commonly used extractant is chloroform isoamyl alcohol (CIA, 24:1 v/v). However, additional measures may be required to purify nucleic acids from complex mixtures such as cell lysates. These are digested by proteolytic enzymes such as pronase or proteinase K, which are active against a broad spectrum of native proteins before extracting with organic solvents. The standard way is however to extract once with phenol, once with 1:1 mixture of phenol and chloroform and once with chloroform. The final extraction with chloroform serves to remove lingering traces of phenols from the nucleic acids (Ochanda *et al.*, 1996).

DNA Quantification

Two methods are widely used to measure the amount of DNA. For pure samples (without significant amounts of contaminants e.g. protein, phenol or other nucleic acids), spectrophotometric measurements of UV irradiation absorbed by the bases is simple and accurate. However, if the amount of DNA is so small or if the sample contains significant quantity of impurities, the amount of DNA can be estimated from the intensity of fluorescence emmited by ethidium bromide (EtBr) on agarose gels with uncut, unmethylated lambda DNA standards (Ochanda *et al.*, 1996).

Spectrophotometric determination of DNA quantity

Optical density readings are taken at a wavelength of 260 nm to quantify DNA. This reading allows calculation of the concentration of nucleic acids in the sample. An OD of 1 corresponds to approximately 50 ng/ml for double stranded DNA, 40 ng/ml for single

stranded DNA and 20 ng/ml for oligo-nucleotides (Ochanda *et al.*, 1996). The ratio between the readings at 260 nm and 280 nm (OD260/OD280) provides an estimate for the purity of nucleic acids. Pure preparations of DNA have an OD260/OD280 of 1.8. If there is contamination with protein or phenol, the OD260/OD280 reading will be significantly less than the values given above hence accurate quantification of DNA will not be possible (Ochanda *et al.*, 1996)

Ethidium bromide quantification

Where there is no sufficient DNA (<250 ng/ml) to assay spectrophotometrically or DNA may be heavily contaminated with other UV absorbing substances that impede accurate analysis, a rapid way to estimate the amount of DNA in such samples is to utilize the UV induced fluorescence emitted by ethidium bromide molecules intercalated into the DNA. Because the amount of fluorescence is proportional to the total mass of DNA, the quantity in the sample can be estimated by comparing the fluorescent yield of the sample with that of a series of standards. As little as 1-5 ng of DNA can be detected by this method and the migration of DNA in these agarose gels is determined by factors such as molecular size of DNA, agarose concentrations, DNA conformations, voltage applied and temperature (Ochanda *et al.*, 1996).

Polymerase Chain Reaction

PCR is an enzymatic reaction for *in-vitro* amplification of a specific genomic region exponentially. PCR applications employ a heat-stable DNA polymerase, mostly Taq polymerase which was isolated from bacterium *Thermus aquaticus*. The amplified region could be a single gene, a part of a gene or a non-coding sequence. Most PCR methods typically amplify DNA fragments up to 10kb, though some techniques allow amplification of fragments up to 40 kb. A basic PCR set-up requires the following components: (Ahmed, 2006)

- DNA template that contains the DNA region to be amplified
- Two primers, which are complementary to the DNA regions at the 5' and 3' ends of the DNA region of interest
- Taq DNA polymerase
- Deoxynucleoside triphosphates (dNTPs), which are the building blocks from which the DNA polymerase synthesizes a new DNA strand

- PCR buffer solution that provides a suitable chemical environment for optimum activity and stability of the DNA polymerase
- Divalent cations, mostly Mg^{2+}

PCR consists of a series of 20-40 repeated temperature cycles, which involve;

- Initialization step which heats the reaction to a temperature of 94-96 ^{0}C and holds for 1-9 min. This is mainly for Hot-start PCR
- Denaturation step where the reaction is heated to 94-98 ^{0}C for 20-30 sec to enhance melting of DNA template and primers by disrupting the hydrogen bonds between complementary bases of the DNA strands. This yields single stranded DNA
- Annealing step where the reaction temperature is lowered to 50-60 ^{0}C for 20-40 sec to allow the annealing of the primers to the single stranded DNA template. The polymerase then binds to the primer-template complex and begins DNA synthesis.
- Extension step where the temperature used depends on the DNA polymerase used. Taq polymerase has optimal activity at 72-80 ^{0}C and a temperature of 72 ^{0}C is commonly used with this enzyme. In this step, DNA polymerase synthesizes a new DNA strand complementary to the template in the 5'-3' direction. At each extension, the amount of DNA target is doubled thus leading to exponential (geometric) amplification of the specific DNA fragment.
- Final extension is a single step performed at a temperature of 70-74 ^{0}C to ensure that any remaining single-stranded DNA is fully extended.
- Final hold at 4-15 ^{0}C for an indefinite time for short storage of the PCR amplicons

The PCR amplicons are then size separated through electrophoresis to check whether the PCR generated the anticipated DNA fragment. The size of the PCR products is determined by comparison with a DNA ladder, which contains DNA fragments of known size and run on the gel alongside PCR products. In PCR, the temperatures used and the length of time they are applied in each cycle depend on the parameters such as the enzyme used for DNA synthesis, the concentration of $MgCl_2$, the concentration of dNTPs and the melting temperature of the primers (Tm). There is therefore need to optimize PCR conditions for the project (Ahmed, 2006).

CHAPTER THREE

MATERIALS AND METHODS

3.1. Experiment I: Field phenotyping for drought tolerance traits

3.1.1. Experimental site

The field experiments were carried out at the Kenya Agricultural Research Institute (KARI) - Masongaleni farm in Kibwezi division of Makueni district in Eastern province. The farm is located at latitude 2^0 21.6' south and longitude 38^0 7.3' east, at an elevation of approximately 650 m asl. The area falls under the lowland semi-arid zones of Kenya, under Kenya's agro-ecological zone VI and receives 400 mm of rainfall per year with an estimated annual evapo-transpiration potential of between 1650-2300 mm (KARI, 2004). Average maximum temperature is between 24 ^{0}C -30 ^{0}C described as fairly hot to very hot. The soils are well drained sandy loams (KARI, 2004).

The purpose of this experiment was to develop and characterize the F_2 segregating population into drought tolerant and drought susceptible individuals based on ASI and other traits. This was done to guide laboratory analysis at the molecular level and the correlations between the two.

3.1.2. Plant materials

3.1.2.1. Selection of parents

The parents used in this study were selected based on results from a prior project seeking to characterize the Kenyan maize landraces based on phenotypic descriptors (Munyiri, 2008) whose original experimental material consisted of 64 maize genotypes which consisted of 37 local maize landraces obtained from the gene bank of Kenya, Muguga. These landraces had been collected from certain environmental niches of Western, Eastern, Coast and Nyanza provinces by CIMMYT/Gene Bank of Kenya. The other genotypes were ten (10) inbred lines from KARI-Kitale, seven (7) inbred lines from KARI-Katumani and seven (7) from CIMMYT-Nairobi.

These 64 genotypes were subjected to field screening for drought through a first season experiment arranged in an 8 x 8 lattice design replicated three times followed by a second season confirmatory experiment arranged in RCBD design under well watered and water stressed conditions (Munyiri, 2008). Variables such as days to anthesis, days to silking, leaf rolling, leaf senescence, among others were measured in the field during screening however;

the anthesis-silking interval (ASI) was considered to be the most important character in estimating the level of drought tolerance. Those genotypes with a shorter ASI were considered drought tolerant whereas those with a longer ASI were considered as being drought susceptible. Based on these ASI estimates, the 64 genotypes were categorized into three groups

- 1-5 days- drought tolerant
- 6-11 days- medium tolerant
- 12-16 days- drought susceptible.

Based on these results, two parents: the most susceptible parent, KCB (ASI = 12) and the most tolerant parent, GBK032357 (ASI= 3) were selected to develop the F_2 mapping population.

3.1.3. Generating the F_2 segregating population

Using the selected genotypes, the cross KCB (P_1) X GBK 032357 (P_2) was made to give F_1s. The crosses were made at KARI- Masongaleni during the period Aug-Nov, 2006. The susceptible parent (P_1) was used as the female parent whereas the tolerant parent (P_2) was used as the male parent. The crossing was achieved through staggered planting to synchronize on the flowering dates of the female and male parent, (temporal separation). The parents were planted at 1 row of 10 plants spaced at 0.75 x 0.3 m, replicated 4 times at intervals of 4 days for each parent. The planting dates were determined based on the estimated number of days from emergence to silking of the P_1s and the estimated number of days from emergence to anthesis for the P_2s (Munyiri, 2008). A 1.5 m path was used to separate each of the four crossing blocks. The crop was grown under normal rainfall supplemented with irrigation as no drought stress was required at this stage. The agronomic management followed was as follows: DAP fertilizer applied at planting at the rate of 5gms per hill. The crop was weeded three times and top dressing done twice with CAN at 5gms per plant. Bulldock was applied at one-foot height for protection against stalk borer attack. After harvesting, at least 200 seed of the F_1 was saved.

3.1.4. F_2 population development

The F_1 population from the cross was planted during the following season (Dec-Mar, 2007) at the same site. The plot consisted of 10 rows of 10 plants, spaced at 0.9 x 0.45m, with a 1.5 m path separating the plots. This F_1 population was also raised under normal rainfall

supplemented with irrigation, following the same agronomic practices mentioned above. At least 200 F_1 plants were then selfed to produce the F_2 seeds.

3.1.5. Field screening of F_2 populations for drought tolerance

The F_2 seed was planted in 10 rows, alongside a row each of P_1, P_2, F_1 and the dry land composite (DLC) as the standard check. This was done at KARI- Masongaleni during June-October, 2007. The spacing was 0.75 x 0.3 m with each row having 21 hills. Three seeds were planted per hill and later thinned to one plant per hill. Stressing for drought was done between August to September, 2007 over a period of 50 days.

The crop was managed as follows; DAP fertilizer applied at planting at the rate of 5gms per hill, weeded thrice to avoid stress from weed competition, two top-dressings with CAN at 5gms per plant, Bulldock application at one-foot height for protection against stalk borer attack. The crop was raised under controlled line source sprinkler irrigation.

Drought stress was achieved by stopping irrigation at one week before flowering in the absence of rain. The field screening for drought tolerance of the F_2 plants was done to categorize the plants into tolerant and susceptible based on ASI and other parameters. It also served to monitor the segregation of the population for these parameters to be used as supplementary data to the laboratory results. Irrigation was resumed a week after silk emergence of the last plant to enhance some grain formation.

3.1.6. Field data collection

The parameters measured in the field were, days to anthesis (DA), days to silking (DS), leaf rolling, leaf senescence, tassel size, plant height, ears per plant, seed set per plant (kernel number per plant) and yield per plant.

Days to anthesis (DA) and days to silking (DS), were recorded as the number of days from emergence to anther extrusion from tassel and to visible silks respectively, per plant. Anthesis-silking interval (ASI) was calculated as the difference between DS and DA of each plant. Leaf rolling was taken just before flowering on a scale of 1-5, with 1 indicating less rolling of leaves and 5 indicating extreme rolling. Leaf senescence was recorded as the percentage (%) of prematurely dried leaves against the total leaf area per plant. Plant height was taken by measuring each plant in centimeters from the ground to the first flag leaf. Tassel size was measure on a scale of 1-5 where 1 indicated a tassel with few branches (small tassel) whereas 5 indicated a tassel with many branches (big tassel). Ears per plant data was recorded

as the number of cobs per plant with at least one fully developed grain whereas seed set was recorded as the number of grains on each cob. Yield per plant was found by weighing grain from each plant, at 13% moisture (Dickey John moisture tester) and recorded as g/plant. All data was collected on the plant basis (data collected from each individual plant).

3.1.7. Field data analysis

Simple correlation analysis was done for all measured traits using Pearson's correlation applying the PROC MIXED procedure in SAS (SAS institute 2001) version 8.2. Each of the measured traits was plotted in histograms to give an indication of the distribution of the traits in the F_2 population.

3.2. Experiment II: Determining molecular markers linked to ASI

3.2.1. Experimental site

Part of this work was done from Tea Research Foundation of Kenya (TRFK) Molecular Biology Laboratory, (DNA extraction and optimization for PCR conditions) while the PCR analysis was done at Egerton University (Dept. of Biochemistry and Molecular Biology Laboratory).

The aim of this experiment was to identify molecular markers linked to anthesis-silking interval (ASI) which is a secondary trait determining grain yield in maize. According to Banziger *et al.* (2000), secondary traits are those traits that are genetically associated with grain yield and having high heritability under stress. Such traits should also be stable and fast to measure. They should also be reliable estimators of grain yield potential. ASI has been shown to have high heritability, is relatively cheap to measure with precision and has a strong negative correlation with grain yield under drought.

Fresh leaf samples were taken from (P_1, P_2, and F_1) and from each individual F_2 plant at four-leaf stage and stored frozen at -80 ^{0}C while awaiting completion of field phenotyping.

3.2.1. Genomic DNA extraction

3.2.1.1. Optimization of DNA extraction protocol for maize

This experiment was designed to help determine whether to do micro-extraction or macro-extraction, whether to use DTT or β-mercaptoethanol as reducing agents, and whether adapting the extraction protocol as originally described by Gawel and Jarret, (1991). 100 ml 2xCTAB solution was prepared (20 ml 1M sterile Tris (pH 7.5), 28 ml 5M sterile Nacl, 4 ml 0.5M sterile EDTA (pH 8.0), 47.8 ml ddH_2o and 2g CTAB) (2% CTAB, 100 mM Tris-HCl

pH 8.0, 1.4 M NaCl, 20 mM EDTA). This was preheated in a water bath at 65 ^{0}C for 30min. The buffer was then divided into two 50 ml volumes. To the first volume, 0.5 ml mercaptoethanol was added whereas to the second half, 0.05g of DTT was added. Each of these was then used to extract 3 random samples using both micro- (small amount of leaf in a 1.5 ml eppendorf tube) and macro- (extraction in a pre-chilled mortar and pestle) extraction methods (a 3x2x2 factorial experiment). The resultant DNA was then checked for quality on a 1.5% agarose gel. There was no difference in the quality and quantity of DNA on extraction with either DTT or mercaptoethanol (data not presented here). Both micro- and macro-extraction protocols yielded high molecular weight DNA. It was decided to adopt micro-extracton as it required less consumables while the amount of DNA needed for PCR is also very small.

3.2.1.2. 2xCTAB micro-DNA extraction

The leaf samples were coded and the eppendorf tubes used in extraction were well labeled based on these codes. 100 ml of 2xCTAB extraction buffer was prepared (20 ml 1M Tris (pH 7.5), 28 ml 5M NaCl, 4 ml 0.5M EDTA (pH 8.0), 47.8 ml ddH_2o, 1 ml β-mercaptoethanol and 2g CTAB) and pre-heated at 65 ^{0}C for 30 min before extraction. Six discs made by closing the lid of a 1.5 ml eppendorf tube on a sample leaf, were used in extraction. The closed eppendorfs were then dropped into a container with liquid nitrogen and let for a few minutes. The eppendorf tubes were then picked one by one from the container using forceps and the now brittle leaf sample quickly crushed using a glass rod. A half a spatula of polyvinyl polypyrrolidone (PVP) was added during grinding (to remove polyphenols). 800 µl of extraction buffer was added in two portions of 400 µl each with further grinding and mixing (The time between grinding and addition of extraction buffer was shortened to avoid oxidation of the samples). The already extracted samples were stored on ice as the others were being extracted. After extraction, the samples were incubated in a water bath at 60 ^{0}C for 30 min with constant shaking. 600 µl of chloroform: isoamyl alcohol (CIA 24:1) was added to each eppendorf tube after incubation and thereafter, thoroughly mixed through careful shaking for 15 min. The samples were then centrifuged in a CHERMLE Z252M centrifuge at 7000 rpm for 20 min. The supernatant (acqueous phase) was carefully aliquoted into fresh-labeled 1.5 ml eppendorf tubes. To these, 900 µl (approx. one and a half times volume of supernatant) of ice-cold isopropanol was added and mixed through gentle inversions and then centrifuged at 7000 rpm for 20 min. Isopropanol was then discarded and the resultant pellet rinsed thrice in 1000 µl of 70% ethanol and centrifuged at 7000 rpm for 5

min. The alcohol was then drained off and the pellets air-dried on clean soft tissue for 3 hours. The pellets were afterwards re-suspended in 100 µl double distilled sterile water (ddH_2O) and treated for RNA using 1µl of pre-boiled RNAse A (10 mg/ml) enzyme, incubated overnight at 55 ^{0}C and stored at 4 ^{0}C.

3.2.2. DNA purification

The stored DNA was taken through a second purification procedure where 100 µl of CIA was added to each sample and shaken for 15 min until properly mixed. The samples were then centrifuged at 7000 rpm for 20 min and the resulting supernatant transferred to new and sterile 0.5 ml eppendorf tube. 40 µl of 1M NaCl-TE (1M NaCl, 10 mM Tris-HCL (pH 8.0), 1 mM EDTA) was added to the old tube, centrifuged at 7000 rpm for 20 min and the supernatant transferred to the corresponding tube already containing the previous supernatant and shaken to mix. 200 µl of ice-cold isopropanol was then added and mixed through gentle inversions to precipitate the DNA, followed by 10 min of centrifugation at 7000 rpm. The isopropanol was discarded and 400 µl of ice-cold 75% ethanol added to the pellets for a gentle upside-down inversion wash. This was followed by centrifugation at 7000 rpm for 2 min, careful discarding of alcohol and another centrifugation at 7000 rpm for 1min followed by the complete removal of the alcohol. The pellets were then air-dried on clean soft tissue for 3 hours, re-suspended in 50 µl ddH_2O, incubated at 55 ^{0}C overnight and stored at 4 ^{0}C.

3.2.3. Estimation of DNA quantity and quality

DNA yield and quality (intactness) was done by visual estimation using the mini-gel method. 5 µl of DNA was drawn from test samples, 2 µl of gel loading dye (50% of glycerol, 250 mM EDTA pH8.0, 0.01 % Bromophenol blue) added to it, and the mixture run on a 1.5% mini-agarose gel in 1x TBE buffer (89 mM Tris-HCl pH 8.3, 89 mM Boric acid, 2.5 mM EDTA), which was also used as the running buffer. Standards of uncut, unmethylated Lambda (λ) DNA (50, 100 and 200 ng/µl standards) were run alongside DNA samples. Electrophoresis was done at 50 volts for 30 min. The 1.5% gel was made by dissolving 0.75g of agarose (molecular biology grade) in 50 ml of 1xTBE. The minigel was then stained in ethidium bromide (EtBr) and visualized under UV light (312 mm). EtBr intercalates between DNA molecules, which consequently fluoresce under UV light. Comparison of band size and staining intensity was made between the test samples and the standards. DNA concentration was inferred from these comparisons. The results were then documented using a digital camera for later analysis.

3.2.4. Polymerase chain reactions (PCR)

3.2.4.1. Optimization of PCR conditions using SSR markers

Optimization involved setting up of factorial experiments involving various levels of PCR reagents to establish the best level of each that would work best with the samples in question. The SSR amplification involved the following thermacycling profiles: initial denaturation at 94 ^{0}C for 5 min followed by 35 cycles of 94 ^{0}C for 30 sec; 55 ^{0}C for 1 min; 72 ^{0}C for 1.5 min, then a final extension step at 72 ^{0}C for 10min and holding at 4 ^{0}C as suggested by SIGMA. Among the factors optimized for are DNA concentration required for a PCR reaction of 10 µl total volume (this was achieved by making serial dilutions of the DNA in ddH_2O); $MgCl_2$ concentration; dNTP mix concentration; primer concentration; concentration of *Taq polymerase* and agarose concentration for resolution of amplified products. For a total PCR volume of 10 µl the effect of a single parameter was determined by varying the levels. The amount of template DNA, (2.6 (dil.1:15), 5.7 (dil.1:7), 13.3 (dil.1:3) and 40 ng (stock)), enzyme (1, 5 and 25, magnesium chloride (0.5, 2.0, 5 mM), primer (0.5, 2, 4 and 10 µM), agarose (1.5, 2.0 and 2.5%), and dNTPS (0.1, 0.2 and 0.5 mM). A given reagent was varied while keeping the rest constant. These levels were all established before bulked segregant analysis was begun.

3.2.4.2. Polymerase chain reaction (PCR) amplification for bulked segregant analysis (BSA)

Ten most tolerant and ten most susceptible F_2 individuals of the cross KCB x GBK 032357 were selected from the F_2 population based on ASI values in days (Table 1). Two DNA bulks were constructed by mixing equimolar amounts of genomic DNA from these individuals, for bulked segregant analysis (BSA). These DNA bulks were then assayed for simple sequence repeat (SSR) polymorphisms alongside parental DNA (P_1, P_2, B_1 and B_2) (Appendix 1), at 109 Loci (Table 2). A total PCR reaction volume of 10 µl was used containing 1x PCR reaction buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTP mix, 0.4 µM of each primer set 0.5 U *Taq* DNA polymerase (SIGMA) and adjusted to 10 µl using ddH_2O. Amplifications were carried out in a MyGenie 96 thermocycler (BIONEER) with the following program: 1 initial denaturation step at 94 ^{0}C for 5 min followed by 35 cycles at 94 ^{0}C for 30 sec, 55 ^{0}C for 1 min and 72 ^{0}C for 1 min, 30 sec, and a final extension cycle at 72 ^{0}C for 10 min. Holding was done at 4 ^{0}C. The amplified products were separated by electrophoresis in a 2.0% agarose in 1 x TBE buffer (89 mM Tris, 89 mM Boric acid and 2.5

mM EDTA pH 8.0), containing 0.15 µg/µl of Ethidium Bromide. The gel was made by dissolving 4g of agarose (Molecular Biology Grade) in 200 ml of 1x TBE (which was also used as the running buffer). The gels were visualized under UV light and images captured using a digital camera and transferred to a computer for analysis (Appendix 2). Primers used in the BSA screening were obtained from SIGMA ALDRICH in a maize SSR Primer set. Only the first plate of the set was screened due to limited resources. Those primers that showed polymorphism among the four samples were selected as candidate markers for further analysis and confirmation

3.2.4.3. PCR for candidate marker confirmation of polymorphism in F_2 progeny

All the primers that showed polymophism between the bulks and/or parents in BSA were selected as candidate markers for further screening on parents and individual F_2 plants making up each bulk (with asterics* in Table 2). A total PCR reaction volume of 10 µl was used containing 1x PCR reaction buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTP mix, 0.4 µM of each primer set 0.5 U *Taq DNA polymerase* and adjusted to 10 µl using ddH_2O. Amplifications were carried out in a MyGenie 96 thermocycler (BIONEER) with the following program: 1 initial denaturation step at 94 ^{0}C for 5 min followed by 35 cycles at 94 ^{0}C for 30 sec, 55 ^{0}C for 1 min and 72 ^{0}C for 1 min, 30 sec, and a final extension cycle at 72 ^{0}C for 10 min and holding at 4 ^{0}C. For each primer, a master mix enough for 25 samples was made because the primers were being confirmed on 22 samples each (P_1, P_2, 10 tolerant and 10 susceptible F_2 progeny). The PCR amplimers were separated on a 3% MetaPhor Agarose in 1x TBE buffer (89 mM Tris, 89 mM Boric acid and 2.5 mM EDTA pH 8.0), containing 0.15 µg/µl of ethidium bromide.

MetaPhor agarose is a high resolution agarose that challenges polyacrilamide and able to resolve small DNA fragments that differ in size by 2 bp. It is mainly used in high resolution separation of 20 bp-800 bp DNA fragments, recovery of fragments under 800 bp, fine analysis of PCR products, and in AMPFLP, STR and tri- and tetranucleotide repeat analysis. The gel was made by weighing 6g of MetaPhor agarose and sprinkling it in a conical flask containing pre-chilled 1x TBE electrophoresis buffer with rapid stirring. The agarose was then soaked for 15 min to reduce the foaming tendency of MetaPhor agarose during heating. The flask was weighed before heating and was then covered using and aluminium foil which was pierced for ventilation. The flask was thereafter heated on medium power for 2 min in a microwave oven. The flask was then swirled gently to resuspend any settled powder and gel pieces and heated on high power to boiling. This was then held at high power for 1 min to

ensure dissolving of all the particles. The flask was removed from the oven, re-weighed and the deficit in weight was compensated for by addition of hot distilled water and thorough mixing. EtBr (0.15 μg/μl) was then added followed by careful swirling to mix thoroughly. The gel solution was then cooled to about 60 °C prior to casting. Upon casting, the molten agarose was left to gel at room temperature. Thereafter, the gel was placed in a fridge at 4 °C for 20 min to obtain optimal resolution and gel handling characteristics (CAMBREX). The amplimers were then mixed with 2 μl of 1x loading dye (50% glycerol, 0.01% bromophenol blue and 250 mM EDTA (pH 8.0)) and carefully loaded into the wells sandwiched by 100 bp DNA ladder and/or SIGMA marker (50-2000 bp) for estimation of band sizes. The gel was then electrophoresed at 80 volts for 120 mins with data documentation being done by digital photographing of the gel on a UV transilluminator (312 m) (Appendix 3). Data was then scored as presence (1) or absence (0) of a DNA band. Missing data was recorded as (-) (Appendix 4). The total number of amplified bands were scored though only the data from intensely stained, unambiguous polymorphic bands was used for linkage analysis. Scoring was done three times to minimize scoring errors.

3.2.5. Quantitative trait analysis

Association between markers and quantitative traits in the F_2 individuals selected for the bulked segregant analysis was determined by regression analysis using EXCEL. A $p < 0.05$ value based on the F-test significance, was used as the threshold for considering the likely presence of a QTL near a marker. The proportion of the total phenotypic variance attributable to each QTL was estimated via linear regression. The total percentage of phenotypic variance explained for each trait by all QTLs was calculated with a multiple regression analysis using the trait as dependent variable and the previously identified markers, linked to a QTL, as treatments (Crouzillat *et al.*, 2000).

Table 1. Flowering data and ASI values in days of plants selected for BSA

Susceptible plant samples				Tolerant plant samples			
Sample	DA	DS	ASI	Sample	DA	DS	ASI
P_1	42	53	11	P_2	57	59	2
F_2 1	56	71	15	F_2 1	53	57	4
2	55	68	13	2	53	54	1
3	56	71	15	3	48	50	2
4	48	62	14	4	53	55	2
5	48	69	21	5	48	51	3
6	51	72	21	6	53	54	1
7	57	71	14	7	56	58	2
8	55	72	17	8	55	58	3
9	58	74	16	9	51	54	3
10	55	74	19	10	50	51	1

DA= Days to anthesis, DS= Days to silking, ASI= Anthesis-silking interval

Table 2. List of primers screened on parental and F_2 DNA

No.	LINKAGE GROUP	Row	PRIMER NAME	SSR PRIMER SEQUENCES (F5'-3'//R5'-3')
1	1	A1	p-umc1354	5'GATCAGCCCGTTCAGCAAGTT3'//5'GAGTGGAGGCGGAGGATCTG3'
2	1	A2	p-umc1566*	ATCTCGTCTACCTAACCCACCCTC//CAGGTGAAGAATCTGGTGAGGTC
3	1	A3	p-umc1292*	GAAGTGAATGTGGGAACATGGTTC//TCACGGTTCAGACAGATACAGCTC
4	1	A4	p-bnlg1124*	TCTTCATCTCTCTCTATCAAACTGACA//TGGCACATCCACAAGAACAT
5	1	A5	p-bnlg1179*	GCGATTCAGTCCGCAGTAGT//GTACTGAACAAACCGGTGGGC
6	1	A6	p-bnlg1014*	CACGCTGTTTCAGACAGGAA//CGCCTGTGATTGCACTACAC
7	1	A7	p-umc1685	TAGTTTGAGGGATCAAGAACCACC//GCTCAAAGGCAAGGCAGTATTTTA
8	1	A8	p-umc2225*	TCGGCTGACATAATAAAACCATAGC//ATGCGAATTTTACCGGGTTTTT
9	1	A9	p-bnlg1429*	CTCCTCGCAAGGATCTTCAC//AGCACCGTTTCTCGTGAGAT
10	1	A10	p-umc2226*	TGCTGTGCAGTTCTTGCTTCTTAC//AGCTTCACGCTCTTCTAGACCAAA
11	1	A11	p-bnlg109*	GCCAGCTGATGTCTGATGAACAGCACA//GATCGGGCCAGATTTCTCAAGTCGTCA
12	1	A12	p-umc1073	CACCAACGCCAATTAGCATCC//GTGGGCGTGTTCTCCTACTACTCA
13	1	B1	p-umc1397*	GTTACACTTGCAGACAAACAACCG//GTCATGATCCGGGAGTAAATCT
14	1	B2	p-bnlg2204	AGGCGACTTAGCTGCAGAAG//CGACTTTCGGTTTGGAAAAG
15	1	B3	p-bnlg2180*	ACAAGGGCGTACCAACCAC//TGACCAGAGGCTTCCATACC
16	1	B4	p-umc1598	CACGTAATAATCCTTTTCCCCAGA//TTTTCTGCAAATGCTCAGAATGAT
17	1	B5	p-bnlg2238	TGCCACTCAAGCCTTCTTTT//TTCTGATTGCAGTGCAGACC
18	1	B6	p-umc1472	TTTTTCTTCTCACCATCACCTTCA//TGGCTTCAAAGAAGAGGAAACATC
19	1	B7	p-umc1144	ATGGCCCACTCATCATATCTCTGT//TGTGTTGATTAGCAGCGGATAAAA
20	1	B8	p-umc2228	ACCATACCTCTCTGAACATGAGCC//GTGAGGTGAAAATGAAGCTGGAAC
21	1	B9	p-umc2025	CGCCGTAGTATTTGGTAGCAGAAG//TCTACCGCTCCTTCGTCCAGTA
22	1	B10	p-umc1461	GGTCACTTCTTGCACCTTTCTGAG//GTCTTTGGGTTGTGATCCCCTT

23	1	B11	p-umc1611*	TACTACCAGCAGCTTGCTTCAACA//CTTCTTGTTCTTCACGCAGTTGTC
24	1	B12	p-umc1323	TTGCACCCCGTTATTATCCTACAG//CCAGACTAGAGTGCCATGATCCTT
25	1	C1	p-umc1812	TACAAGGAAGGCAAGTTCATCCTC//ATGCAGGTGACATTCATCATCATC
26	1	C2	p-umc2234*	GACGGACTATAGAGGGCGATGAG//CAAGATCGTTAGGTTCTAGGCGTC
27	1	C3	p-umc2235	TCGTCCCAGTACCATGCCTC//GAACCCTCTAGGCTCCGGTTC
28	1	C4	p-umc1664	AATTGTTTACTGCGCTGAAACTCC//CCTCTTTGCCTGTACCGTGTATTC
29	1	C5	p-umc1035	CTGGCATGATCACGCTATGTATG//TAACATCAGCAGGTTTGCTCATTC
30	1	C6	p-umc1925	CCTCCGGGTTGTCATACTGTAAAA//GATGGTTGTGATTAGTAGCGCAGA
31	1	C7	p-umc1374	CTACTCCAGCAGTTCCGCTCC//TTCAGATGGTGTAGTGTGGAGAGC
32	1	C8	p-umc1358*	AGAACCTCCCGCTTGACGAC//ACCTCAACCTCGACCTCTGCAT
33	1	C9	p-umc2064	CGAGAGGGGAGAGAGAGAGAGAGA//AGTCGCTGAGTCTACGGTCCCTAC
34	1	C10	p-bnlg1025	TGGTGAAGGGGAAGATGAAG//CCGAGACGTGACTCCTAAGC
35	1	C11	p-umc1245	TGGTTATGTGCATGATTTTTCCTG//CATGCGTCTGATCTTCAGAATGTT
36	1	C12	p-umc1998	GCCTCCCAAGTGCAATATTAAATAGA//AGAGCACAAGAACCACAACAACAA
37	1	D1	p-umc2080	ACCACCTTGTCCGTATCCTTCAC//GCCAAGGTGGGTCTGGCTAT
38	1	D2	p-mmc0041*	AGGACTTAGAGAGGAAACGAA//TTTATCCTTGCAGTTGC
39	1	D3	p-umc1914	CAACATGAGCGTGCTAAATACTCG//ACAGGAACACATGAGGTCATCAAA
40	1	D4	p-umc1715	TTCATTGGGTCTCTAGCCAAGAAG//GGGGAGTCACAGATCTCATCAACT
41	1	D5	p-umc1298	AGCTGAACAAAATAAACGGAACGA//AGGACAAGAAAAAGAAGAAGCACG
42	1	D6	p-umc1184	CTTCCTTACGTGTCACCGCTCT//GTGGAGTGATGTGATCGATGATG
43	1	D7	p-umc2189*	CGTAAGTACAGTACACCAATGGGC//ACACCGACTACAAGCCTCTCAACT
44	1	D8	p-umc1534	GACCCCAAATCTCTCCTTCCTC//TAGCTAAGCTTGTGCTTGCTCG
45	1	D9	p-bnlg1347	GTGGTCACGACGAAATCCTT//TTGCAATCACACAGGTGGTT
46	1	D10	p-phi265454	CAAGCACCTCAACCTCTTCG//TCCACGCTGCTCACCTTC
47	1	D11	p-umc1500	TCTCTGACTATTCCACGAGCTCAA//CTGGTGCGTGCTACAACTGTG
48	1	D12	p-umc1553	TGAATGGAAGAGAAGGGAAATCTG//GCTCTGTACATCCTTAGCGACACA

49	1	E1	p-umc1737	ATGCTTCTCTTCAGAAGCCATCC//TAGCTAGGTATGATGTGCGTGCT
50	1	E2	P-umc1744	ATGAAGCTGGCTGATACGGAAT//CGAGCCCTTAATTAGTTCTCGGTT
51	1	E3	p-bnlg2123	TGATGCAGACAAGTCCTTCG//ACAAATCTCACCTCTGCGCT
52	1	E4	p-umc1331*	TTATGAACGTGGTCGTGACTATGG//ATATCTGTCCCTCTCCCACCATC
53	1	E5	p-umc2100	TGACGTGCAAACAACCTTCATTAC//AAAGGCATTATGCTCACGTTGATT
54	2	E6	p-umc1622	CGCTACAAATCCTACTGGTGCTTT//CCTCGGATTTTCCAAAACATTTCT
55	2	E7	p-umc1419	CTCATCACAACTAGCGCCACTCTA//ATAGTGCAGAGGTCATCGTGGC
56	2	E8	p-phi96100	AGGAGGACCCCAACTCCTG//TTGCACGAGCCATCGTAT
57	2	E9	p-umc2363	TTGACTCGAAAGACTTGTGAGCTG//GTGTGTCAGAAGGAGGACTGATGA
58	2	E10	p-umc1542*	TAAAGCTATGATGGCACTTGCAGA//CATATTTGCCTTTGCCCTTTTGTA
59	2	E11	p-bnlg1017	ATTGGAAGGATCTGCGTGAC//CAGCTGGTGGACTGAATCTA
60	2	E12	p-umc1265	GCCTAGTCGCCTACCCTACCAAT//TGTGTTCTTGATTGGGTGAGACAT
61	2	F1	p-umc1824*	ATCGTGCTTAAGCGGTTATAGGAAT//TGCACATGCTTTGTATAAGATGCC
62	2	F2	p-umc1756	ATCTCAGGTACTCTGCCTACGGG//AACAGAGGGTAGCTTGTGGCCT
63	2	F3	p-umc1262	GTGAAGCTCTGCACCACGCT//ATCGTCCAAAGAAGAAGAGGGAGA
64	2	F4	p-bnlg125	GGGACAAAAGAAGAAGCAGAG//GAAATGGGACAGAGACAGACAAT
65	2	F5	p-mmc0231	GAGCGACTGCGAGACGG//AGATCGCGCCACCGCTC
66	2	F6	p-umc1776*	AAGGCTCGTGGCATACCTGTAGT//GCTGTACGTACGGGTGCAATG
67	2	F7	p-bnlg2248*	CCACCACATCCGTTACATCA//ACTTTGACACCGGCGAATAC
68	2	F8	p-umc1185	AGTAAAAGAGGCAAGGACTACGGC//GCGGCGATATATACGGGTTGT
69	2	F9	p-bnlg1064	CTGGTCCGAGATGATGGC//TCCATTTCTGCATCTGCAAC
70	2	F10	p-phi109642	CTCTCTTTCCTTCCGACTTTCC//GAGCGAGCGAGAGAGATCG
71	2	F11	p-umc1026	TCGTCGTCTCCAATCATACGTG//GCTACACGATACCATGGCGTTT
72	1	G1	p-phi083	CAAACATCAGCCAGAGACAAGGAC//ATTCATCGACGCGTCACAGTCTACT
73	1	G2	p-umc2249	GCATAGACTCCCTGACAGCCAC//AGAAGGTCGTCGTCCTGGAAC
74	1	G3	p-umc1259	CTCTTTGGTGGCAGAAGCAGAAT//TAGCTAAACTTGAGTGCTCTGCCC

75	1	G4	p-umc2079	ATGATCACGTCGTGCTGGTAGTG//CGGCCTCGCTGTCTTCTAGC
76	1	G5	p-umc2007	GTGCATATGCGCACGGAAGT//AGTAGTTACACAACGCAACACGAGG
77	1	G6	p-umc1922*	CTTTTCCGCTGCTCTTCCTTTT//TTCCCTTGTTCCAGATCCTCAATA
78	1	G7	p-umc2252*	GTCTTTGACCCCTTCCTCTTCTTG//CACTGCACTGCAAGGTACATACG
79	1	G8	p-umc2253	TCACGTCCAGGAGGATCACC//ACACCACGGGCTACTGCTCAC
80	1	G9	p-umc2372*	CACCAGGCGTAGTGAGACAGC//ACCCCTTGCGTTCTCTTCTGTT
81	1	G10	p-umc1108	TAAAGAGAGAGGAAACAAGCTGCG//ATGATGAAGCCTCACATACGAA
82	1	G11	p-umc2129	AAGGAGGAGCGTTCTCGTGG//ACGTGGTCACTCACCGC
83	1	G12	p-umc1890	TGAAGGGCCTCCAAGTTAATCTAAA//GCTAATGTGGAACTGGAAGAACGA
84	1	H1	p-mmc0271*	CGTAATGCGTAGCAACATAG//CAACATCCTTTCCACCG
85	1	H2	p-phi251315	CCAGTCCAATGGAGAGGG//GAGATTCCCCTGCAGGACT
86	1	H3	p-bnlg2077*	GACCAGAGGATGGGGAAATT//GTAGGCACATGCACATGAGG
87	1	H4	p-umc1745	GGCACTTACAACTTCTTGCCTTCTT//GCCGTCAAGCGATGATCAAG
88	1	H5	p-umc1126	CAACAGGGTGAACCCTCTGTACTT//AATATGGTGTTGTGATTTGCATCG
89	1	H6	p-phi435417*	CTGACGCCACTGTTGCTTG//AAAAGTAGCCAATCTGCCACG
90	1	H7	p-umc2085	TAGATGTCGATGTCCTCCAGGG//TGTACGACTTCTTCTGGACGCAC
91	1	H8	p-umc1992	TGTACGACTTCTTCTGGACGCAC//TAGATGTCGATGTCCTCCAGGG
92	1	H9	p-mmc0381	GTGGCCCTGTTGATGAG//CGACGAGTACCAGGCAT
93	1	H10	p-umc1516	CAATAACAGATCAGCATGGAGACG//TAGCGACCTGTTATCCGAAAGAAC
94	1	H11	p-umc1256	AGAAGACGATGATGATGATGCAGA//CATCTCGACCTTTGACATTCTCCT
95	1	H12	p-umc1525	TTTGTGCCGAATATAAATGTGACG//AATAATATCAAATGGCGCCAAGC
96	1	IL1	Umc1122	CACAACTCCATCAGAGGACAGAGA//CTGCTACGACATACGCAAGGC
97	1	IL2	phi339017*	ACTGCTGTTGGGGTAGGG//GCAGCTTGAGCAGGAAGC
98	2	IL3	phi101049*	CCGGGAACTTGTTCATCG//CCACGTCCATGATCACACC
99	2	IL4	phi127*	ATATGCATTGCCTGGAACTGGAAGGA//AATTCAAACACGCCTCCCGAGTGT
100	3	IL5	phi053	CTGCCTCTCAGATTCAGAGATTGAC//AACCCAACGTACTCCGGCAG

101	3	IL6	phi029	TTGTCTTTCTTCCTCCACAAGCAGCGAA//ATTTCCAGTTGCCACCGACGAAGAACTT
102	5	IL7	phi331888*	TTGCGCAAGTTTGTAGCTG//ACTGAACCGCATGCCAAC
103	6	IL8	phi423796	CACTACTCGATCTGAACCACCA//CGCTCTGTGAATTTGCTAGCTC
104	7	IL9	phi328175	GGGAAGTGCTCCTTGCAG//CGGTAGGTGAACGCGGTA
105	7	IL10	phi034	TAGCGACAGGATGGCCTCTTCT//GGGGAGCACGCCTTCGTTCT
106	8	IL11	phi100175	TATCTGACGAATCCCATTCCC//GTACGTAACGGACGGACGG
107	8	IL12	phi233376	CCGGCAGTCGATTACTCC//CGAGACCAAGAGAACCCTCA
108	10	IL13	phi059	AAGCTAATTAAGGCCGGTCATCCC//TCCGTGTACTCGGCGGACTC
109	10	IL14	phi041	TTGGCTCCCAGCGCCGCAAA//GATCCAGAGCGATTTGACGGCA

* Primers screened for confirmation

F 5'-3'= forward primer in the 5' to 3' direction; R 5'-3'= reverse primer in the 5'to 3' direction

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1. Experiment I: Field phenotyping for drought tolerance traits

4.1.1. Correlation between traits in maize under drought

Correlation refers to the relationship between two metric characters which can either be positive or negative, with the cause of this correlation in crop plants being either genetic or environmental (Gebre, 2005). Genetic correlation in particular determines the degree of association between traits and how they may enhance selection. According to (Edmeades *et al.,* 1997), the magnitudes of selection and correlated responses depend on heritabilities and phenotypic standard deviations as well as genetic correlations. The main criteria for drought tolerant trait selection are the associations of each trait with grain yield under stress conditions (Edmeades *et al.,* 1997; Banziger and Lafitte, 1997). In this study, the various secondary traits measured under drought showed variations in their relationships with each other and with the primary trait, the grain yield (GY) (Table 3).

4.1.2. Days to anthesis (DA)

Days to anthesis (DA) was significantly and positively correlated ($r = 0.69^{**}$) with days to silking (DS). Similar observations were reported by Ribaut *et al.* (1996) and Munyiri (2008) ($r = 0.7^{**}$, $r = 0.87^{**}$) respectively. DA also showed a weak positive correlation ($r = 0.2^{*}$) with anthesis-silking interval (ASI), an observation also reported by Munyiri (2008) ($r = 0.19^{*}$). Ribaut *et al.* (1996) reported no correlation between DA and ASI but a high positive correlation between DS and ASI. The significant positive correlation between ASI and DA in this experiment could be attributed to non uniform response to drought by the segregating plants that were being observed with the varying water deficit intensities over time, coupled by the possibility of genetic diversity among the individuals as the parents in question were non-inbred. DA however did not show any correlation with grain yield (GY) as was reported also by Frova *et al.* (1999). According to Monneveux *et al.* (2006), severe stress just before flowering reduces stomatal conductance thereby reducing assimilate partitioning to the tassel. This results in delayed anthesis (DA). DA however did not show any significant correlation with grain yield (GY) as also reported by Frova *et al.* (1999).

Table 3. Pearson's correlation coefficients (r) of F_2 under drought

	DA	DS	ASI	LR	LS	PH	TS	EPP	KN	GY
DA	1.00									
DS	0.69**	1.00								
ASI	0.20*	0.84**	1.00							
LR	0.19**	0.19*	0.29**	1.00						
LS	0.09	0.06	0.02	0.43**	1.00					
PH	-0.10	-0.01	-0.03	-0.25**	-0.16*	1.00				
TS	0.08	0.39**	0.40**	0.11	0.04	0.05	1.00			
EPP	-0.10	-0.36**	-0.41**	-0.13	-0.18	-0.07	-0.24*	1.00		
KN	0.07	-0.30**	-0.47**	-0.16	-0.003	-0.05	0.19	0.60**	1.00	
GY	0.04	-0.31**	-0.45**	-0.24*	-0.02	-0.04	-0.18	0.54**	0.91**	1.00

*significant at 5% probability level, **significant at 1% probability level

DA= Days to anthesis, DS= Days to silking, ASI= anthesis-silking interval, LR= leaf rolling,LS=leaf senescence, PH= plant height, TS= tassel size, EPP= ears per plant, KN= kernel weight per plant, GY= grain yield

4.1.3. Days to silking (DS)

DS was significantly and positively correlated ($r = 0.84^{**}$) with ASI (Table 3). The strong positive association between DS and ASI was also reported by Bolanos and Edmeades (1996; 2000), Frova *et al.* (1999) and Munyiri (2008). This trait also had significant and positive associations with DA ($r = 0.69^{**}$), leaf rolling (LR) ($r = 0.19^{*}$), and tassel size (TS) ($r = 0.39^{**}$). The trait was however significantly and negatively correlated with EPP ($r = -0.36^{**}$), KN ($r = -0.30^{**}$) and GY ($r = -0.31^{**}$). The significantly negative correlation between DS and GY ($r = -0.31^{**}$) and its high positive correlation with ASI demonstrate that ASI is determined largely by variation in DS under drought (Bolanos and Edmeades, 1996; Frova *et al.*, 1999). Mugo *et al.* (1998) and Frova *et al.* (1999) noted that delays in silking were due to less partitioning of assimilates into the ear and this was closely associated with barrenness.

4.1.4. Anthesis-silking interval (ASI)

The ASI had positive and significant associations with DA ($r = 0.20^{*}$), DS ($r = 0.84^{**}$), LR ($r = 0.29^{**}$) and TS ($r = 0.40^{**}$) but significant and negative associated with EPP ($r = -0.41^{**}$), KN ($r = -0.47^{**}$) and GY ($r = -0.45^{**}$) (Figure 1). Two factors have been shown to lead to yield losses when drought occurs before or during flowering. First, the lack of synchronization between anthesis and silking leads to a decrease in ears per plant (EPP). Secondly, photosynthesis is severely affected (Mugo *et al.*, 1998; Frova *et al.*, 1999; Bruce *et al.*, 2002; Richards, 2005), and, thus, accumulation of carbohydrate is reduced, resulting in a shortage in the flux of assimilates to the developing ear (Edmeades *et al.*, 2003). These result in barrenness and abortion, thus reducing GY (Mugo *et al.*, 1998). Frova *et al.* (1999) suggested that the high negative correlation between ASI and GY under drought stress indicated that the variation in GY under drought is due to variation in ear setting processes that are related to biomass partitioning at flowering. Bänziger and Lafitte (1997) also reported that under drought stress conditions, GY increases were strongly associated with reduced ASI, reduced barrenness and increased harvest index, but mildly associated with delayed LS and reduced TS. Andrade *et al.* (2002) noted that a short ASI is indicative of general tolerance to reduced photosynthesis at flowering.

According to Monneveux (2006) and Bruce *et al.* (2002), the ovule abortion rate is positively correlated with ASI period implying that ASI determines the kernel set. Munyiri (2008) found high associations between GY and ASI, and GY and EPP under drought stress. The findings of this study are therefore of significance. Negative correlations between ASI

and EPP, KN and GY (r = -0.41, r = -0.47 and r = -0.45, respectively), suggest that a delay in silking under drought was one of the major causes of GY losses in maize under drought.

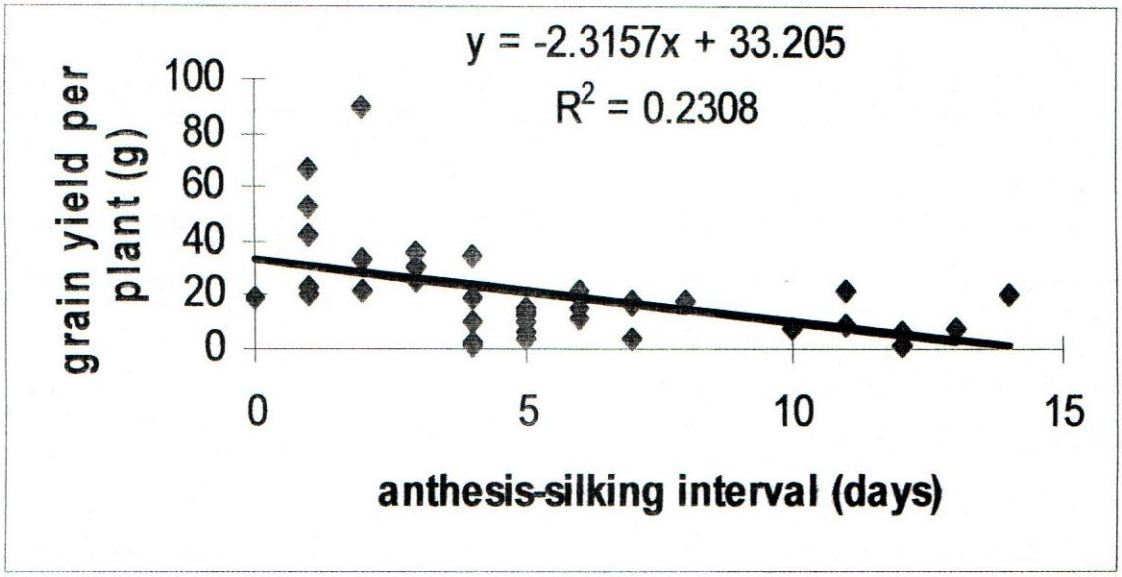


Figure 1. Relationship between ASI and GY in an F_2 maize population evaluated under drought stress

4.1.5. Leaf rolling (LR) and leaf senescence (LS)

Positive associations were observed between leaf rolling (LR) and DA, DS, ASI and leaf senescence (LS) (Table 3). However, this trait was significant and negatively associated with plant height (PH), and GY. Leaf senescence (LS) had a significant and negative association with PH, though it showed a weak negative association with GY. The effects of LR in crop productivity are debatable. Bolanos and Edmeades (1996) established that leaf rolling is a drought adaptive trait. Barker *et al.* (2005) also inferred that leaf rolling increased water use efficiency through the reduction of water loss at critical times of the day (when there is high water pressure deficits), while allowing photosynthesis when water pressure deficits are less. Mugo *et al.* (1998) and Bruce *et al.* (2002) note that the GY loss under drought is attributable to strong leaf rolling and the associated abscisic acid (ABA) which accumulates when a plant wilts. This in turn inhibits leaf expansion, thus reducing radiation interception. The significant negative association between LR and GY in this study (r = -0.24*) therefore support earlier suggestions by Monneveux *et al.* (2006), Mugo *et al.* (1998) and Bruce *et al.* (2002) that LR reduces radiation interception thus reducing photosynthesis and consequently

GY. Barker *et al.* (2005) also noted that drought at or immediately after flowering accelerated leaf senescence thus reducing leaf area, reduced intercepted radiation and photosynthesis that result in a reduction in photo-assimilate flux to the spikelets. In this study, the significant positive association between LS and LR (r = 0.43**) also supports the idea of reduced assimilates due to a reduced photosynthetic area, though no significant association was observed between LS and GY. This may also explain the significant and negative associations between PH with LR and LS (r = -0.25** and r = -0.16*, respectively).

4.1.6. Tassel size (TS) and plant height (PH)

Tassel size (TS) was significantly and positively correlated with DS and ASI but negatively with EPP (Table 3). This trait was however weakly associated with GY. Whereas Bolanos and Edmeades (1996) had reported negative correlation between TS and GY, Munyiri (2008) reported no correlation between the two traits. The results of this study on TS could be attributed to the landrace origin of the parents used in this study. Munyiri (2008) reported that most landraces had larger tassels. According to Chapman and Edmeades (1999), reduced tassel size culminated into reduced competition for assimilates thus resulting into increased ear size. Banziger *et al.* (2000), who suggested that genotypes with small tassels allowed allocation of the scarce assimilates during drought to the developing ears, also supported this inference. PH showed negative correlation with LR and LS. According to Monneveux (2006), PH was unaffected by selection under stress.

4.1.7. Yield and yield components

Ears per plant (EPP) was significantly and positively associated with kernel number (KN) and grain Yield (GY) (r = 0.60** and r = 0.54**, respectively) (Table 3). The association between KN and GY was highly significant and positive (r = 0.91**) (Figure 2). Based on evaluation of S1 to S3 progenies from six elite tropical maize populations, Bolaños and Edmeades (1996) reported high correlations (r = 0.7** to 0.8**) between GY and KN, and between GY and EPP respectively under drought. These associations increased when the stress levels intensified. A strong phenotypic association between GY and KN in both water-stressed and well- watered environments (r = 0.96**; r = 0.87**) was reported by Chapman and Edmeades (1999). Bolanos and Edmeades (1996) indicated that variation in KN has a more pronounced effect on GY. In tropical populations improved for drought tolerance, a strong positive relationship was reported between harvest index and GY under drought conditions (Edmeades *et al.*, 1999; 2004). They also reported that GY, EPP and KN were

strongly correlated with ASI under drought conditions, but not under well watered conditions. Andrade *et al.* (1999) reported that EPP and KN directly influence the GY and that GY is more highly correlated with KN than EPP. In this study therefore, the correlation between GY and KN (r = 0.91) was much stronger than that between GY and EPP (r = 0.54) thus supporting results reported by Xiao *et al.* (2005) and earlier studies already discussed, hence further demonstrating that KN is a very important yield component for maize than EPP and 100K weight.

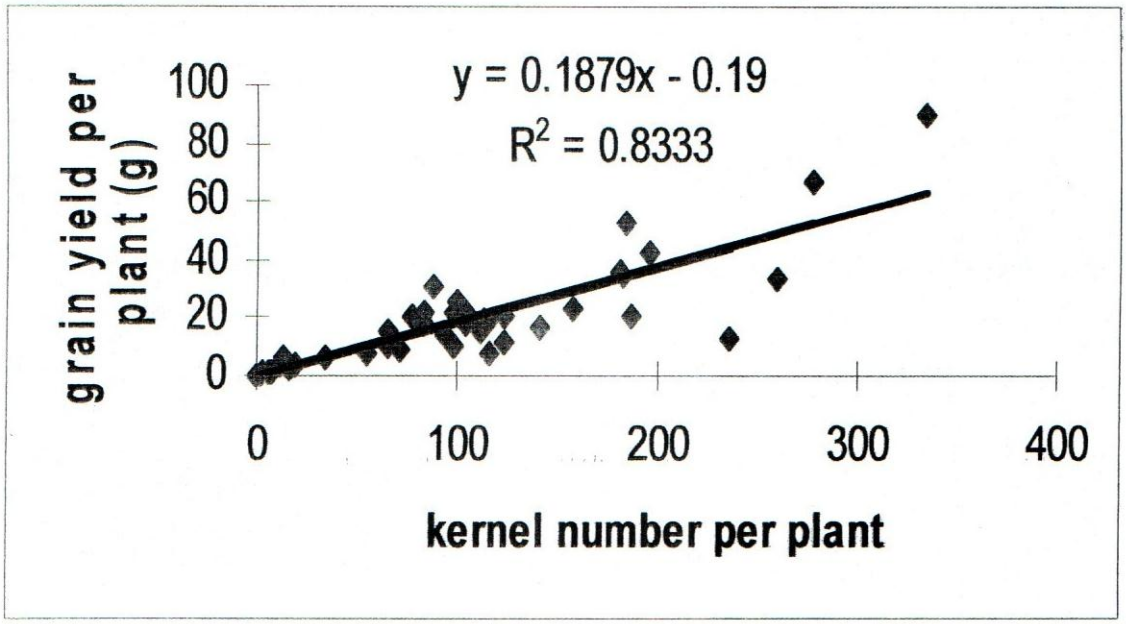


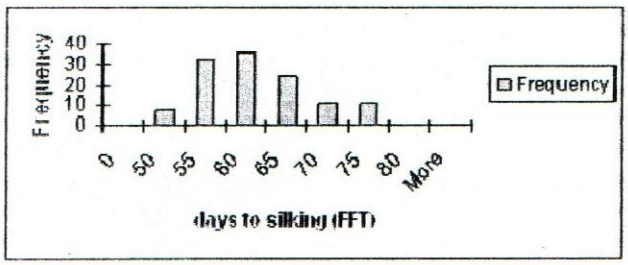
Figure 2. Relationship between kernel number per plant and grain yield per plant under drought.

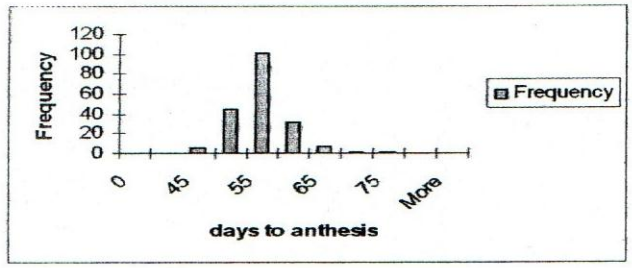
Several studies have attributed GY losses to EPP and KN. Bolanos and Edmeades (1996) attributed more than 75% of variation in GY under drought to variation in ears and kernels per plant. Monneveux *et al.* (2006) noted that 85% of kernel abortion under water stress occurred during the first 20 days after DS and that this was associated with the level of barrenness. This suggestion agrees with the earlier report by Tollenaar and Wu (1999) that grain abortion occurred during the first 2-3 weeks after silking and gets worse as the photosynthetic canopy is reduced by stress leading to reduced assimilates into the ear. According to Boyer and Westgate (2004), water deficit later in development reduces kernel size rather than kernel number and this is largely determined by the available photosynthetic

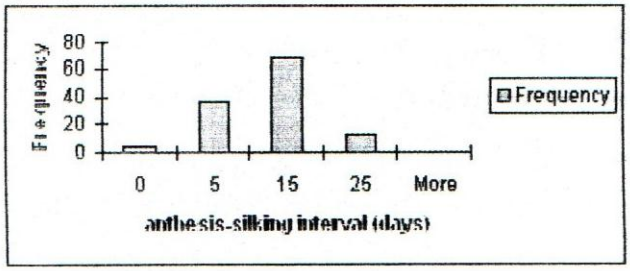
assimilates to the ear. Bolanos and Edmeades (1996) reported that ASI accounted for about 76% of variation in GY, with almost 10% reduction per day increase of ASI. The relatively lower percentage of GY variation attributable to the ASI in this project (45%) can be attributed to the difficulty in managing uniform stress in the segregating population in which data was collected on the plant basis. This therefore means that most of these plants underwent both flowering and grain-filling stress such that some plants that had a relatively short ASI and were therefore able to be pollinated lost grain through abortion in the grain filling stress. GY therefore strongly depended on the ears that could support some grain formation and this therefore explains the strong correlation between KN and GY.

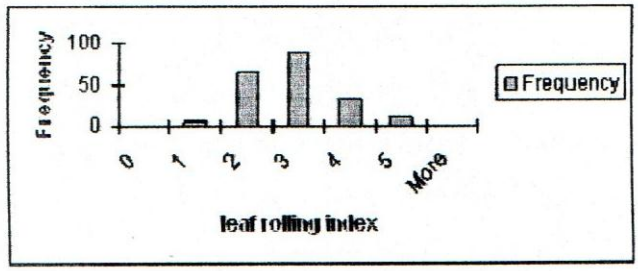
4.1.8. Frequency distribution of traits measured in F_2 under drought

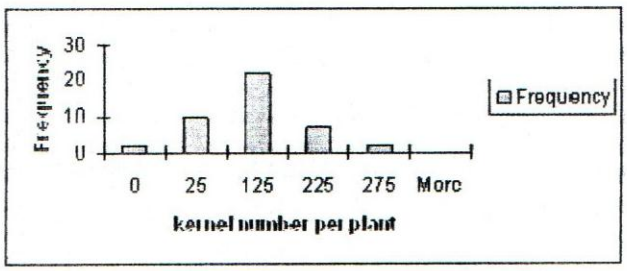
All the traits showed a large quantitative variability and a near normal frequency distribution under the drought conditions (Figure 3). The two parents were clearly differentiated for all traits. Thus, this F_2 population was suitable for QTL analysis of drought tolerance.

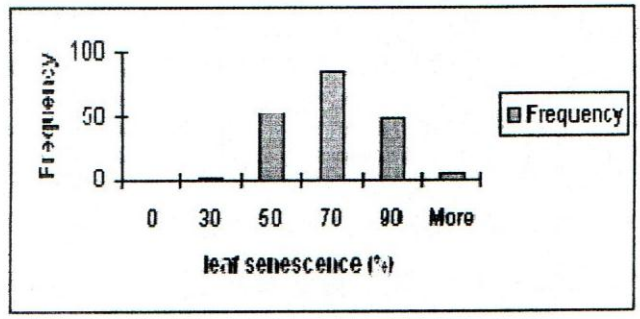


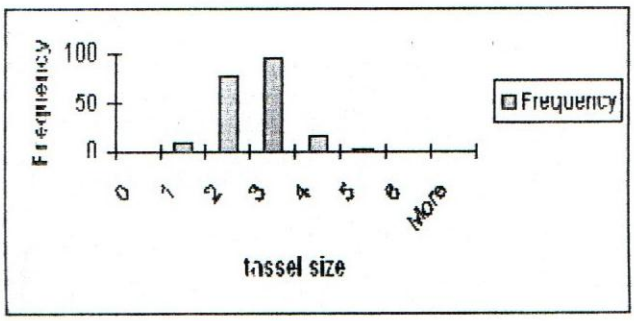


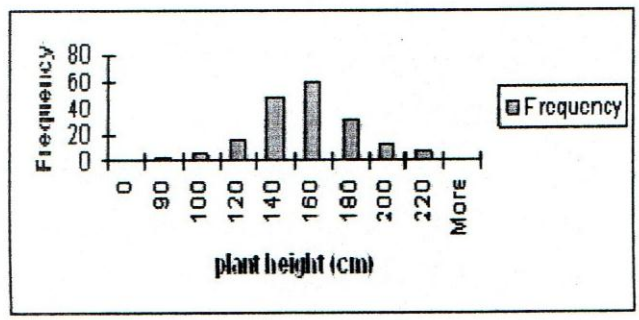












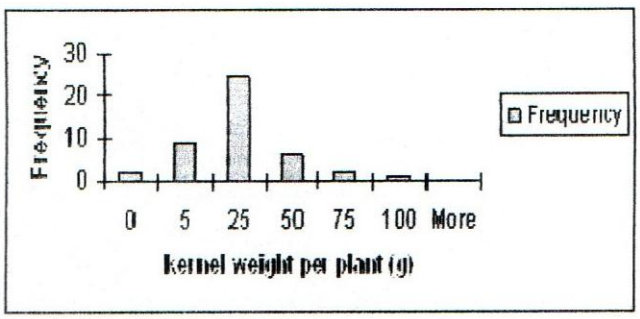


Figure 3. Frequency distribution of the traits measured in the field under flowering water stress in the F_2 population used in the cosegregation QTL analysis

4.2 Experiment II: Determining molecular markers linked to ASI

4.2.1 Optimization of PCR conditions for SSR primers

According to Ahmed (2006), a PCR set-up requires several components such as the DNA template that contains the DNA region to be amplified and two primers, which are complementary to the DNA regions at the 5' and 3' ends. Also required are *Taq* DNA polymerase, deoxynucleoside triphosphates (dNTPs), which are the building blocks from which the DNA polymerase synthesizes a new DNA strand, PCR buffer solution that provides a suitable chemical environment for optimum activity and stability of the DNA polymerase, and divalent Mg^{2+} cations. The concentration used for each of the above affects the results. Furthermore, the temperature cycles used largely depend on these components.

Before conducting any genome mapping study, optimization of concentration of different reagents and temperature regimes used in polymerase chain reaction (PCR) is necessary (Rahman *et al.*, 2002). In this study, the thermocycling profiles were kept constant whereas all the other reagents were optimized. From the study, it was observed that high concentrations of template DNA failed to give bands or resulted in smears (Figure 5). This could be because of enzyme-substrate imbalance when the level of the rest of the reagents was held constant. These results therefore support the finding by Rahman *et al.* (2002) that the amount of template DNA strongly influences the outcome of the reaction.

A 1x2x2 factorial experiment involving 1 primer, 2 levels of $MgCl_2$ (5 mM and 2 mM) and 2 levels of agarose (2.0 and 2.5%) (Figure 4), established that a higher level of $MgCl_2$ in a PCR mix gave many but non-specific bands. Such bands cannot reveal the desired polymorphisms. Rahman *et al.* (2002) reported that lower $MgCl_2$ concentration (0.5 mM) failed to yield visible bands, and higher magnesium concentration (1.5 and 2.0 mM) yielded adequate but somewhat less amplification products. They also noted that with more than 3.5 mM $MgCl_2$, no specific bands of microsatellite product could be seen. The importance of optimum magnesium concentration for PCR is well recognized (Rahman *et al.*, 2002). The Mg^{2+} ions bind tightly to the phosphate sugar backbone of nucleotides and nucleic acids, and variation in the $MgCl_2$ concentration has strong effects on nucleic acid interactions.

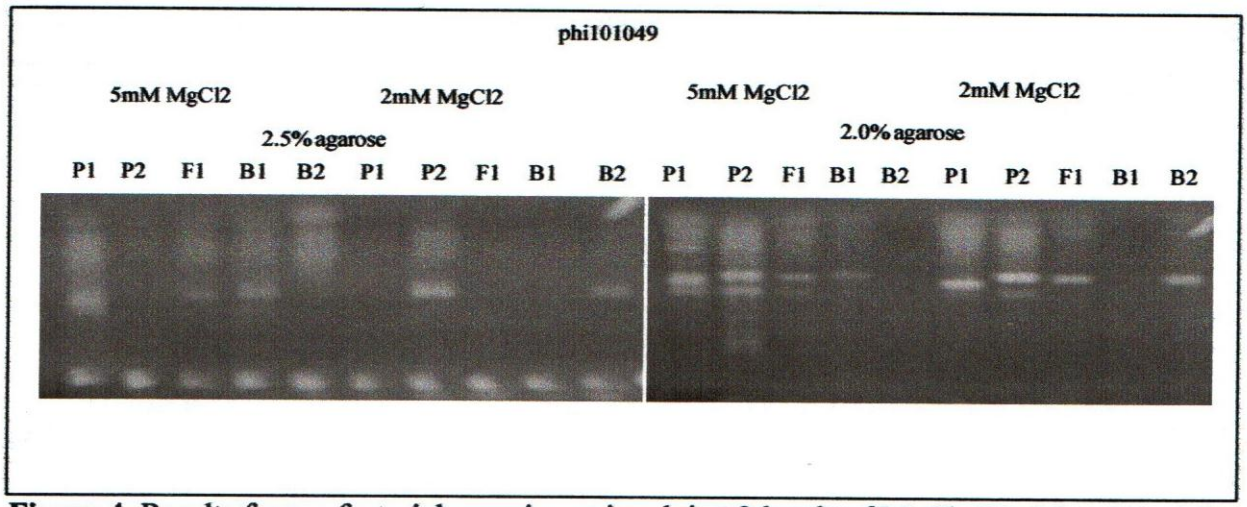


Figure 4. Results from a factorial experiment involving 2 levels of $MgCl_2$ (5 mM and 2 mM) and two levels of agarose (2.0 and 2.5%) with one primer. P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), F_1= Hybrid individual, B_1= bulk 1 (from 10 drought susceptible individuals), B_2= bulk 2 (from 10 drought tolerant individuals)

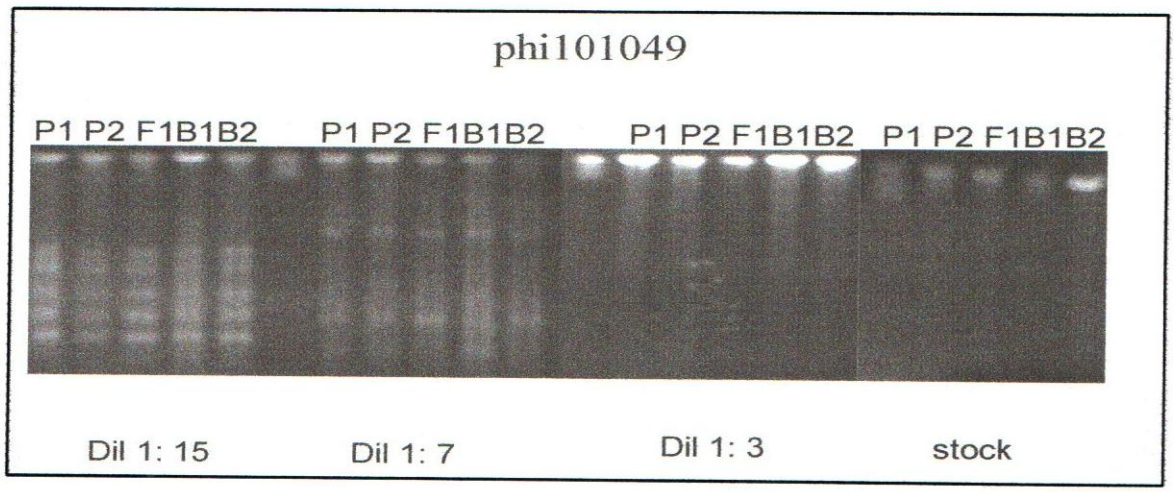


Figure 5. PCR products generated in a template DNA concentration optimization experiment using marker *phi101049*. P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), F_1= Hybrid individual, B_1= bulk 1 (from 10 drought susceptible individuals), B_2= bulk 2 (from 10 drought tolerant individuals). Dil 1:15, 1:7, 1:3 are the serial dilutions made between Stock DNA: ddH_2O.

According to Rahman (2000) variations in $MgCl_2$ concentration below 4 mM can improve performance of PCR by affecting specificity as they also observed that higher concentrations lowered the specificity whereas lower concentrations raised specificity.

In this study, the different enzyme (*Taq* DNA polymerase) concentrations gave variation in the PCR products (Figure 6). Optimization for the same samples using the same primer showed that very low levels of *Taq* DNA polymerase (1U/reaction) did not give any clear bands whereas a very high level (25 U/reaction) gave very many monomorphic bands. 5 U/reaction of *Taq* polymerase gave the best results in this case. Rahman *et al.* (2000) found out that when the amount of the enzyme (*Taq* polymerase) was decreased to 0.1 unit/reaction, there was no visible amplification product whereas more enzyme units upto 10 units/reaction resulted in the formation of amplification products. They however also noted that with higher enzyme levels, the resolution of the bands decreased. In addition, this resulted in the formation of non-specific products of higher molecular weight, clearly visible on the ethidium-bromide stained agarose gels. Padmalatha and Prasad (2006) also reported that lower concentration of *Taq* DNA polymerase did not show proper amplification whereas high concentration showed decreased specificity. This could be due to the presence of excess enzyme molecules which results in non-specific priming.

In this study, very low (0.1 mM) and very high (0.5 mM) concentration of dNTPs gave smears with almost invisible bands (Figure 7). Ahmed (2006) reported that lower concentrations of dNTPS resulted in additional smears and artifacts with low intensity of desired bands. Padmalatha and Prasad (2006) also observed that increased concentrations of dNTPS reduce free Mg^{2+} thus interfering with the enzyme, *Taq DNA polymerase*. According to Rahman *et al.* (2000) the concentration of dNTPs in reaction mixture is strongly correlated to the Mg ions concentration due to the interaction between mononucleotides and the Mg^{2+}. A higher concentration of Mg^{2+} allows amplification with a higher concentration of dNTPs, which is not seen at lower Mg^{2+} concentrations. Results from this study therefore support the fact that the correlation between the level of $MgCl_2$ and dNTPs affects the activity of *Taq* DNA polymerase, thereby affecting the final PCR products.

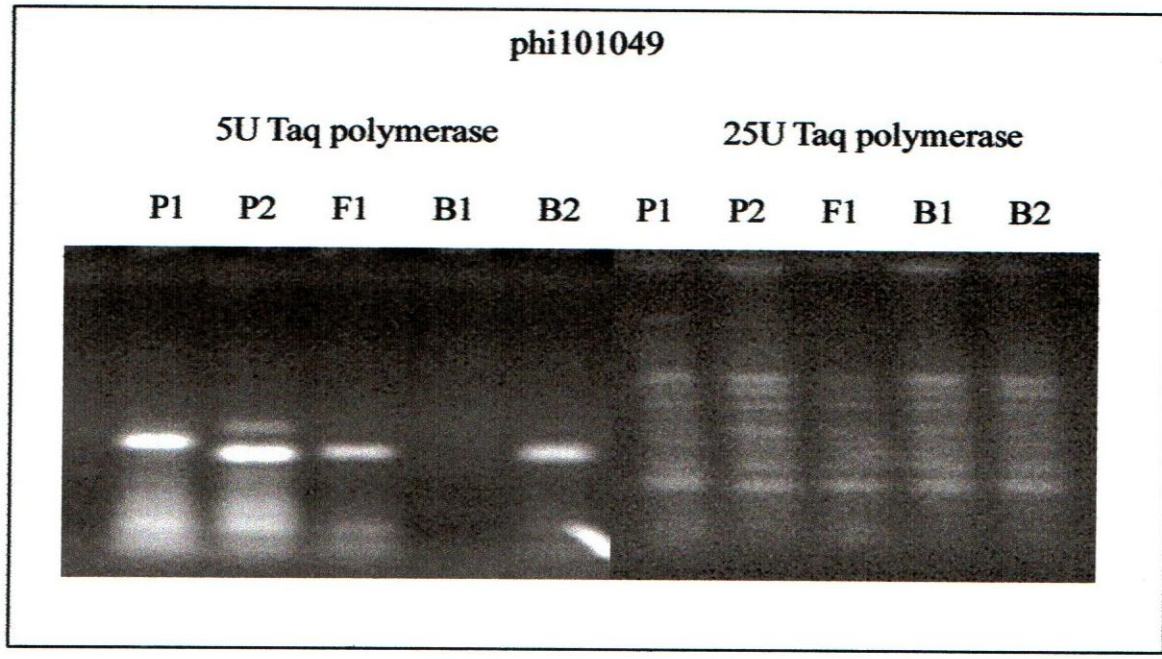


Figure 6. PCR products generated by marker *phi101049* using two levels of *Taq* DNA polymerase (5U and 25U). P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), F_1= Hybrid individual, B_1= bulk 1 (from 10 drought susceptible individuals), B_2= bulk 2 (from 10 drought tolerant individuals).

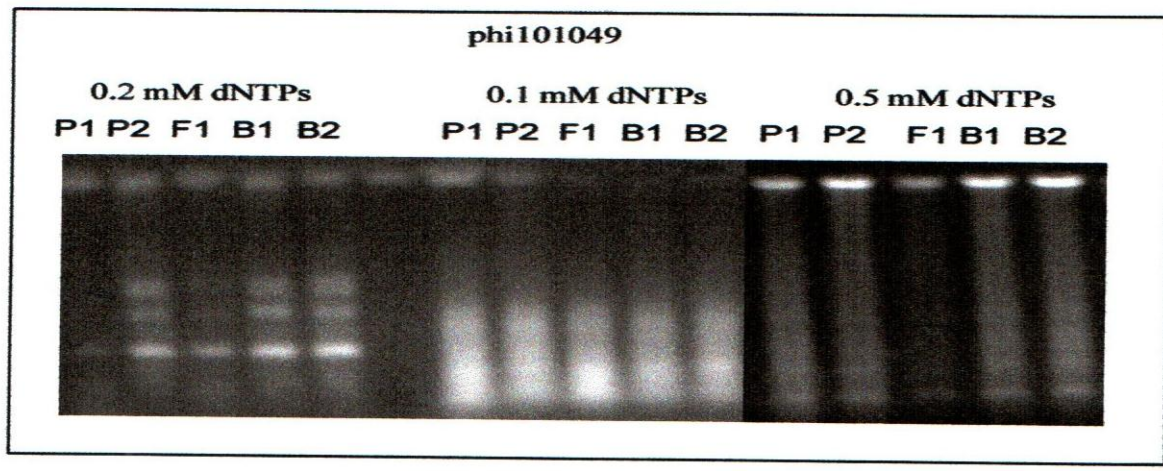


Figure 7. PCR products generated by marker *phi101049* in a dNTP optimization experiment with all other factors constant, at dNTP levels 0.1 mM, 0.2 mM and 0.5 mM. P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), F_1= Hybrid individual, B_1= bulk 1 (from 10 drought susceptible individuals), B_2= bulk 2 (from 10 drought tolerant individuals).

Primer concentration was established to have effects on PCR products (Figure 8). Very low concentration (0.5 µM) did not give any bands. This, according to Rahman *et al.* (2000) could be due to the very low concentration of the primer, which becomes limiting and results in the formation of very little or no amplification product. Ahmed (2006) also reported reducing band intensity with decreasing primer concentrations. Very high concentration (10 µM) gave several bands together with smears and dimers. Padmalatha and Prasad (2006) observed that lower and higher concentrations lead to absence of amplification and primer dimer formation, respectively.

It was also established that a 2.5% agarose gel gave more smears and missing data than 2.0% (Figure 4). This therefore implies that a higher percentage of agarose reduces the pore size thereby making it difficult for larger fragments to pass through. DNA is a negatively charged molecule and which is moved by electric current through a matrix of agarose with the distance traveled depending on the fragment sizes and the concentration of agarose, which determines the pore sizes (Ahmed, 2006).

From these results therefore it was apparent that concentrations of the various components are very important in determining the PCR products. For this study, it was established that for a total PCR reaction volume of 10 µl, 1x PCR reaction buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTP mix, 0.4 µM of each primer set, 0.5 U of *Taq* DNA polymerase and 2.5 ng template DNA, gave the best results (Figure 9). These levels were, therefore adopted for use in both bulked segregant analysis and candidate markers confirmation.

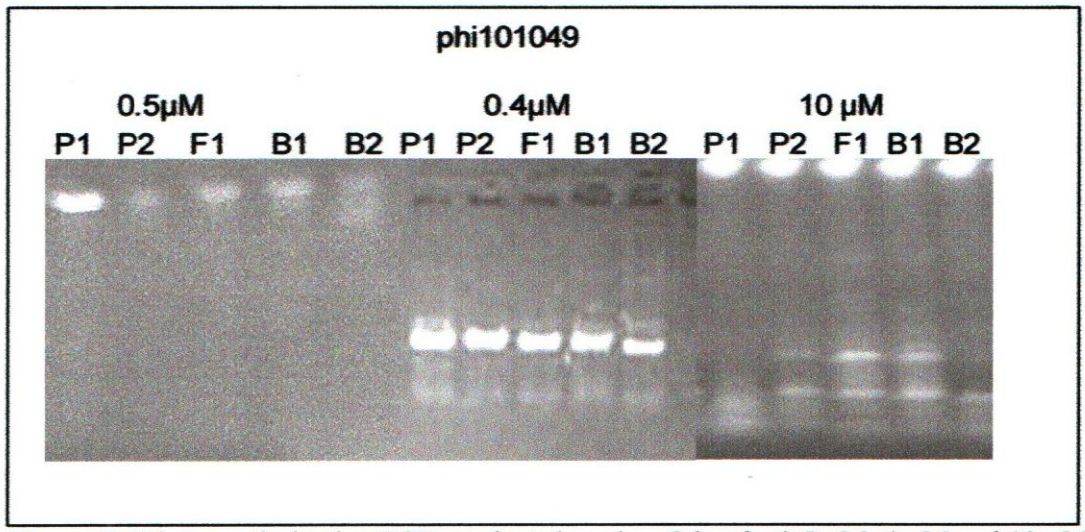


Figure 8. Primer optimization PCR products based on 3 levels, 0.5 µM, 4 µM and 10 µM.

P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), F_1= Hybrid ndvidual, B_1= bulk 1 (from 10 drought susceptible individuals), B_2= bulk 2 (from 10 drought tolerant individuals).

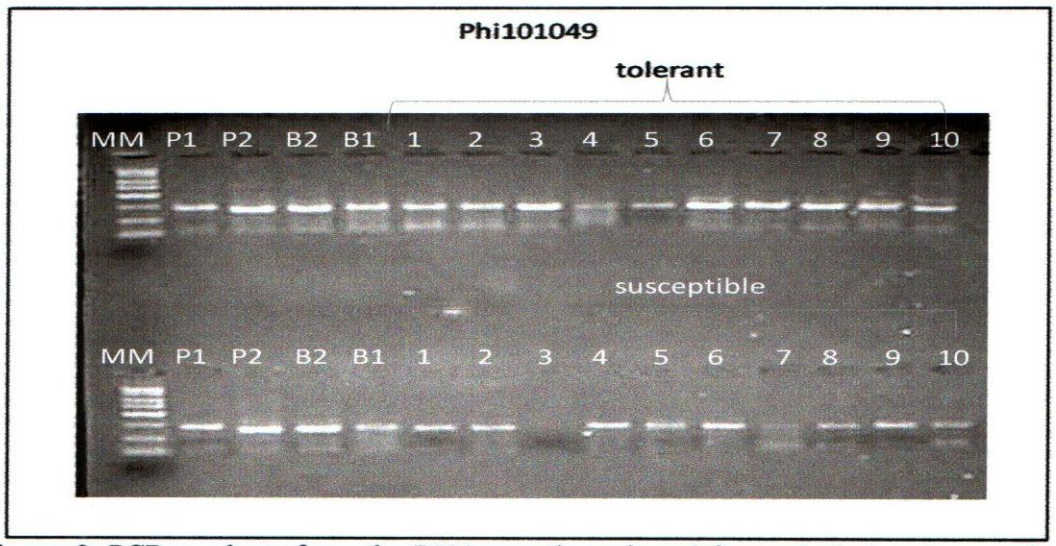


Figure 9. PCR products from the DNA samples selected for analysis based on optimized PCR conditions. On gel 1, MM= Molecular marker (100bp), P_1= drought susceptible parent, P_2= drought tolerant parent B_1 is drought susceptible bulk, B_2= drought tolerant bulk, 1-10= drought tolerant individual F_2 plants. On gel 2, MM= Molecular marker, P_1= drought susceptible parent, P_2= drought tolerant parent B_1 is drought susceptible bulk, B_2= drought tolerant bulk, 1-10= drought susceptible individual F_2 plants.

4.2.2 Bulked segregant analysis

Because each pool contains individuals selected to have identical genotypes for a particular genomic region but random genotypes at loci unlinked to the selected region, the two resultant DNA bulks differ genetically only in the selected region and are seemingly heterozygous and monomorphic for all other regions. Markers that were polymorphic between the pools were therefore considered genetically linked to the loci determining the trait used to construct the pools. In this study, 27.5% of the 109 screened primers showed polymorphism between the parents and/or the bulks selected for anthesis-silking interval (ASI) (Appendix 2). Such primer pairs were taken as candidate markers for further analysis e.g. primers *p-umc1358*, *p-umc2225* among others (Figure 10). The parents used in this study were not inbred hence the possibility of genetic diversity for the trait among the selected F_2 individuals. A lot of polymorphism would, therefore be expected.

The low propotion of polymorphic markers observed in this study can therefore be attributed to possibilities of past introgression events or a common geneology between the parents. KCB is an improved open pollinated cultivar for the dry mid-altitude areas especially in Eastern Province of Kenya whereas GBK032357 is a landrace cultivar from Kwale, a neighbouring coastal district. Both these cultivars are therefore adapted to almost similar environments. According to Carvalho *et al.* (2002), different landrace accessions are developed and selected for different environments and morphological characteristics. Perales (2003) identified several ways through which farmer management of landraces results in continued changes. First is through purposeful hybridization between traditional and modern maize types. Secondly, there is a possibility of creation of new maize landraces by directional selection of the progeny of hybridization between two traditional landraces. Third is through the displacement of a local landrace by the introduction of a modern variety or a non-local landrace. Fourth is through the maintenance of stable populations of locally dominant landraces at the expense of the weak ones and lastly is through the market driven selection for a minor variety.

The low level of polymorphism in this study can be attributed to the fact that BSA when applied to QTL identification, has been shown to only detect the loci with large genetic effects (Grattapaglia *et al.*, 1996) (see section 4.2.8). Furthermore, the normal Molecular Biology Grade agarose, which was used at 2% to separate the PCR products at this level, has been shown to fail to detect polymorphisms between bands that differ in only a few base pairs. This may also have contributed to the low level of polymorphism observed.

Though some polymorphism was observed, these results were not conclusive because at this level, the co-segregation between the suspected candidate markers with the trait could not be established. The candidate markers had therefore to be confirmed and the PCR products separated on a gel with better resolution.

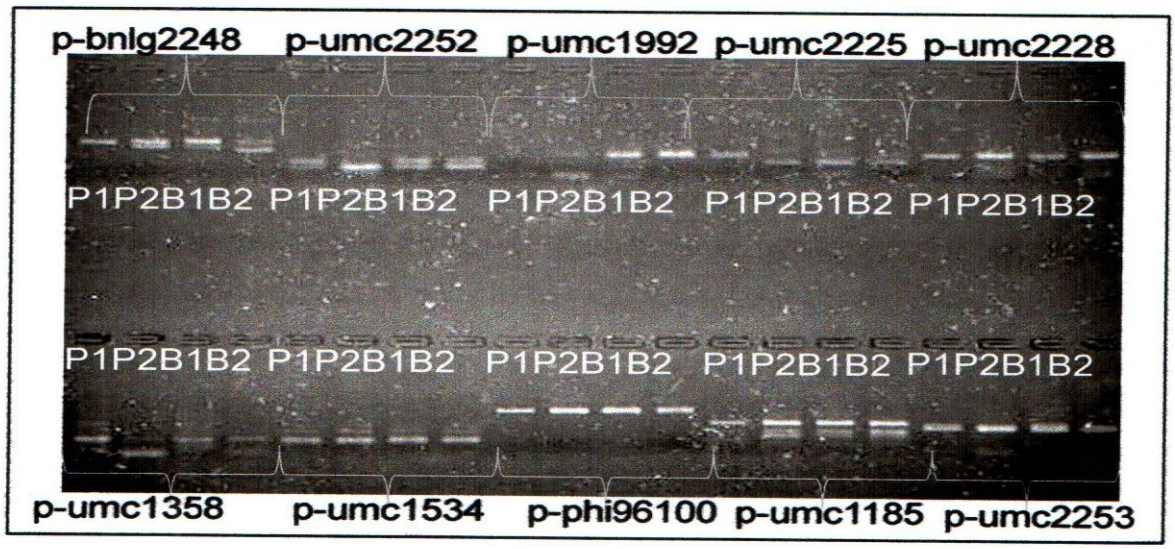


Figure 10. PCR products generated by 10 primers with P_1, P_2, B_1 and B_2 respectively, during bulked segregant analysis. P_1=drought susceptible parent, P_2= drought tolerant parent, B_1= drought susceptible bulk, B_2= drought tolerant bulk.

4.2.3 Confirmation of candidate markers

Grattapaglia *et al.* (1996) noted that with a large sample size, bulked segregant analysis (BSA) followed by genotyping of a random sample of individuals, would offer a quick and less expensive way for identification of genomic regions controlling quantitative traits of interest. The 30 primers used for confirmation of polymorphisms gave a total of 71 scorable bands (alleles) with band sizes ranging from 70- 350 bp (Appendix 4). The PCR products generated by the candidate markers *p-umc1542* (Figure 11) and *p-umc2189* (Figure 12) with P_1, P_2 and the 20 F_2 individuals used to constitute the two bulks used in BSA screening are presented.

4.2.4 ASI candidate QTLs

Using linear regression, 4 out of 71 F tests were found to be significant ($P \leq 0.05$) (Table 4; Appendix 5) and these markers were therefore assumed to be linked to ASI. The regions near these markers have been referred to as candidate QTLs in this study. These were regions near markers *p-umc2189*, *p-umc1542*, *p-bnlg1179* and *p-bnlg1014* and the individual variance explained by each candidate QTL ranged from 26- 39 % (Table 4). The candidate QTLs together could explain 65% of the total observed variation in ASI as calculated with multiple regression analysis using the trait as dependent variable and the closest markers linked to the candidate QTL as treatments (Table 5). In their study on identification of QTLs for flowering parameters and anthesis silking interval under drought conditions, Ribaut *et al.* (1996) observed that all QTLs detected by MAPMAKER/QTL under intermediate stress and severe stress were also consistently detected with a one-way ANOVA.

Flowering parameters have been studied under both well-watered and water stress conditions in maize (Guo *et al.*, 2008; Li *et al.*, 2003; Ribaut *et al.*, 1996) and QTLs for several traits identified. According to Zhang *et al.* (2006) and Li *et al.* (2005), the source of the parents, the size of the progeny and the genotype by environment interaction effects are responsible for the difference in the localization of QTLs. Most genetic studies have been carried out in populations with inbred parents or recombinant inbred lines (RILs) and such inbred lines have fixed characters (Guo *et al.*, 2008; Li *et al.*, 2003; Zhang *et al.*, 2006; Li *et al.*, 2005). Guo *et al.* (2008) using RILs, found four QTLs for ASI on chromosomes 3, 6 and 7 under well-watered conditions and two QTLs on chromosomes 5 and 6 under water stress. Zhang *et al.* (2003) reported two QTLs for ASI on chromosomes 2 and 6 under water stress.

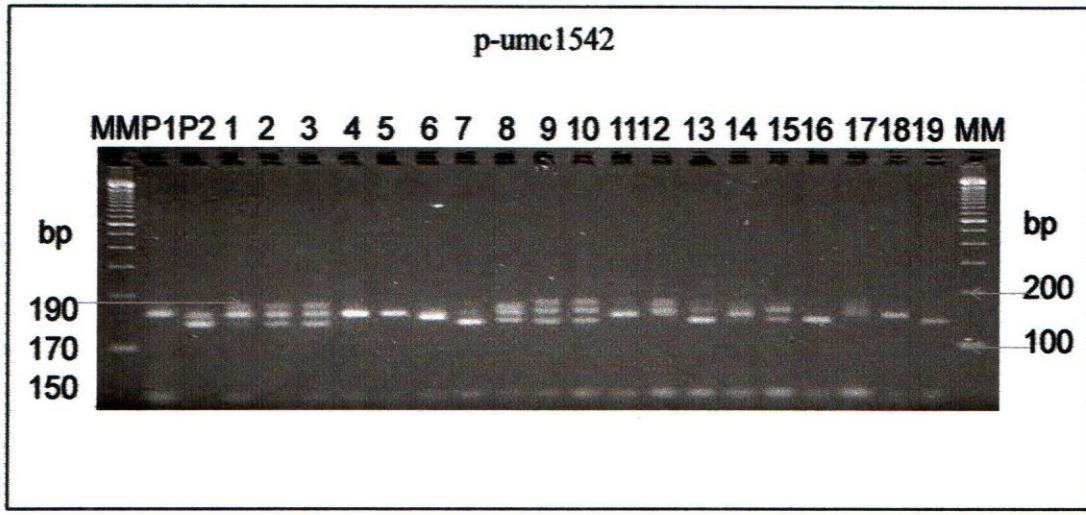


Figure 11. PCR products generated by marker *p-umc1542* on parents, 10 drought tolerant and 9 drought susceptible F_2 individuals separated on MetaPhor agarose (3%). MM= 100 bp DNA ladder, P_1= parent 1 (drought susceptible), P_2= parent 2 (drought tolerant), 1-10= F_2 plants constituting the drought tolerant bulk, 11-19= F_2 individuals constituting the drought susceptible bulk used in BSA. Right bp= the fragment sizes of the two standard marker bands flanking bands generated by marker p-umc1542. Left bp= the estimated allele sizes of the three scored bands.

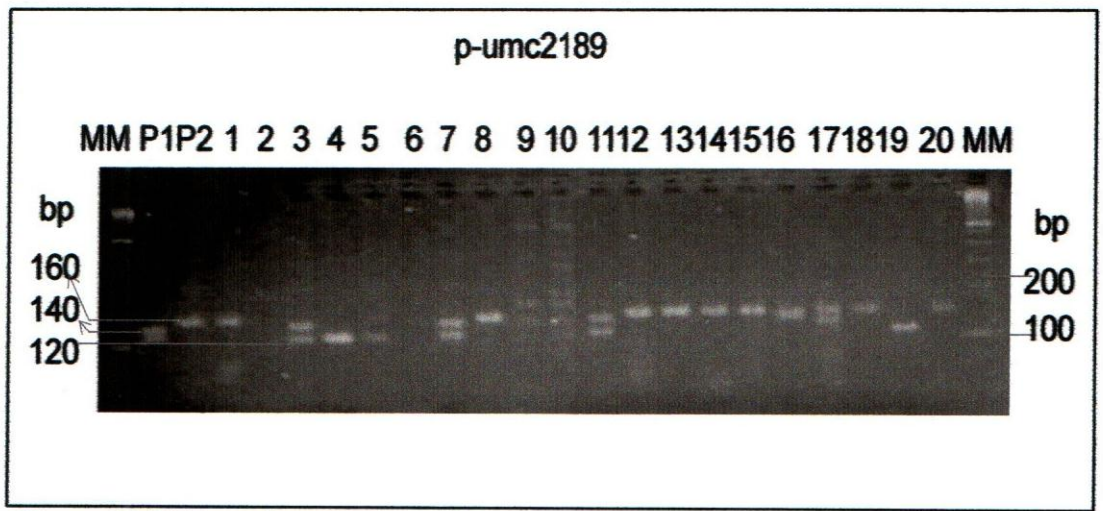


Figure 12. PCR products generated by primer *p-umc2189* on parents, 10 drought tolerant and 10 drought susceptible individuals and separated on MetaPhor agarose (3%). MM= 100 bp DNA ladder, P_1= parent 1 (S), P_2= parent 2 (T), 1-10= F_2 plants constituting the drought tolerant bulk, 11-19= F_2 individuals constituting the drought susceptible bulk used in BSA. Right bp= the fragment sizes of the two standard marker bands flanking bands generated by marker *p-umc2189*. Left bp= the estimated allele sizes of the three scored bands.

Table 4. QTL data for ASI from F_2 progeny of KCB x GBK 032357

Anthesis-silking interval (days)

Marker	Progeny no.	R^2 (%)	p-value	Linkage group
p-bnlg1179	20	28.3	0.033749*	1
p-bnlg1014	20	38.7	0.01006*	1
p-umc2189	20	34.3	0.0107*	1
p-umc1542	20	26.2	0.025014*	2

*significant at 5% level of probability

QTL= quantitative trait loci; ASI= anthesis-silking interval

Table 5. Multivariate QTL analysis for ASI

Trait	QTL no	Linkage group	Variance (%)	p-value
ASI	4	1 and 2	65.3	0.002108**

**significant at 1% level of probability

QTL= quantitative trait loci; ASI= anthesis-silking interval

Ribaut *et al.* (1996) also reported six QTLs for ASI on chromosomes 1, 2, 5, 6, 8 and 10 under water stress conditions. Guo *et al.* (2008) reported QTLs on chromosomes 1, 2, 6, 8 and 10 in all water regimes.

According to the maize genome data base, the candidate markers identified for ASI under water stress conditions in this study can be mapped on chromosomes 1 and 2 i.e. *p-umc2189*, *p-bnlg1179* and *p-bnlg1014* on chromosome 1 and *p-umc1542* on chromosome 2. QTLs for ASI have also been mapped on chromosome 1 (Ribaut *et al.*, 1996; Guo *et al.*, 2008) and chromosome 2 (Ribaut *et al.*, 1996; Zhang *et al.*, 2003). These results also confirm the observation by Ribaut *et al.* (1996) that ASI QTLs are relatively consistent in the maize genome. Another important observation from this study is that three out of the four identified candidate QTLs are located on linkage group 1 (chromosome 1) (Table 4). Grattapaglia *et al.* (1996) noted that several of the significant markers associated with trait expression were linked on the same linkage group. The presence of the three candidate QTLs on the same chromosome in this study therefore implies that these markers are linked and possibly inherited together. They therefore could offer a powerful tool for indirect selection of grain yield in maize under drought.

4.2.5 Co-segregation of candidate QTLs with ASI in F_2 progeny

Genetic linkage mapping is based on meiotic recombination and based on Mendel's second law of independent assortment; linkage is detected when the proportion of progeny inheriting recombinant combinations of any given pair of markers falls significantly below a chosen threshold. Because the population used in this analysis was a subset of the F_2 population derived from the two tails, a segregation ratio of 1:1 is expected for each marker in this case because the SSRs are co-dominant markers. Any deviations from this in the two subsets were treated as recombinants. The candidate QTLs identified from this study i.e. *p-umc2189*, *p-umc1542*, *p-bnlg1179* and *p-bnlg1014* had 22%, 26%, 25% and 18% recombination frequencies with ASI (Table 6), respectively. A lower recombination frequency indicates close linkage between the marker and the genomic region conferring the trait. In this case therefore, crossing –over during meiosis is limited hence the marker is co-inherited with the trait of interest.

Table 6. Co-segregation of markers *p-umc1542*, *p-umc2189*, *p-bnlg1179* and *p-bnlg1014* with ASI

p-umc 1542				
	Progeny	Present	Absent	R.F
Short ASI	10	7	3	
Long ASI	9	2	7	26%
p-umc 2189				
	Progeny	Present	Absent	R.F
Short ASI	8	6	2	
Long ASI	10	2	8	22%
p-bnlg 1179				
	Progeny	Present	Absent	R.F
Short ASI	9	6	3	
Long ASI	7	1	6	25%
		p-bnlg 1014		
	Progeny	Present	Absent	R.F
Short ASI	7	4	3	
Long ASI	9	0	9	18%

ASI=anthesis-silking interval, R.F= recombination frequency

Unlike in other QTL studies, the parents used in developing the F_2 population in this study were not inbred but landrace. The possibility of genetic variability in the population cannot therefore be ruled out. The relatively high recombination frequency observed in the candidate markers could be attributed to this fact. Grattapaglia *et al.* (1996) noted that multiple alleles as well as multiple QTLs could be expected from a genetically heterogeneous population as compared to homogeneous populations. The absence of the marker of interest in some of the individuals of the same bulk in this study may not necessarily mean the presence of recombination as treated in this case. This could imply that not all individuals in a bulk derived their trait of interest from the same loci and especially so for a quantitave trait like ASI. The presence of candidate markers with close association with ASI regardless of possible heterogeneity of the population as observed in this study therefore agrees with the observation by Ribaut *et al.* (1996) that ASI QTLs are relatively consistent in the maize genome.

4.2.6 Phenotypic correlation and pleiotropism

Two of the four markers associated with ASI (*p-umc2189* and *p-bnlg1014*) (Table 7), both located on chromosome 1, showed significant association with days to silking (DS) implying a possible co-localization of QTLs for the two traits (ASI and DS). The two markers could together account for 39 % of the total observed variation (Table 8). The other two markers, (*p-bnlg1179* and *p-umlc1542*) which are located on chromosome 1 and 2 repectively, showed significant association with both KN and GY (Table 7). These two could account for 36.1% of the total observed variation in kernel number (KN) and 40.6% of the observed variation in Grain yield (GY) (Table 8). Guo *et al.* (2008) reported QTLs for grain yield on chromosome 1 and 2 under well watered conditions and on chromosome 1 and 9 under water stress conditions. Li *et al.* (2003) also reported 5 QTLs for GY on chromosomes 1, 2, 4 and 8 under drought conditions.

The co-location of QTLs for several traits has been reported in previous QTL studies. Ribaut *et al.* (1996) reported four common QTLs between ASI and DS on chromosomes 1, 2, 5 and 8. Guo *et al.* (2008) also reported co-localization of QTLs for ASI and Grain yield (GY) on chromosome 10 while Xiao *et al.* (2005) reported two co-localized QTLs for KN and GY on chromosome 1. The co-location of QTLs for different traits has been taken to imply the likely presence of pleiotropic or close linkage between the QTLs that control the traits (Agrama and Moussa, 1996; Tuberosa *et al.*, 2002). Guo *et al.* (2008) suggested that the co-localization of QTLs for several traits is associated with a correlation in the phenotypic data. In this study, the phenotypic correlation between ASI and DS was very high (r = 0.84**) (Table 3). ASI was also significantly but negatively correlated with KN (r = -0.47**) and GY (r = -0.45**) (Table 3). The expected association between the level of phenotypic correlation and the linkage of QTLs is therefore verified by the results of this study.

Although pleiotropic gene action may be the most likely explanation for the potential co-localization of these candidate QTLs the current data are insufficient to establish with certainty the presence of pleiotropic effects. This is more so considering the fact that ASI is calculated as the difference between DS and DA data thus making it correlated with DS and/or DA. Again, the data used here is based on a subset of the entire F_2 population which may necessitate further confirmation of this association on the entire F_2 population. Furthermore, these observations were made by regressing data from markers found by pooling DNA based on ASI phenotypic data, on the phenotypic data of the other measured traits such as DS, KN and GY.

These results therefore are just indicative of the possibility of co-localization of genomic regions responsible for several traits.

The presence of two common QTLs for ASI and both KN and GY in this study, coupled with the high negative correlation between ASI with grain yield (GY) and other yield components phenotypically therefore imply that grain yield (GY) can be effectively and indirectly selected for using ASI. These results also confirm the genetic basis of the phenotypic findings by Andrade *et al.* (1999) and Xiao *et al.* (2005) that whereas EPP and KN directly influence the GY, GY is more highly correlated with KN than EPP thereby demonstrating that KN is a very important yield component for maize than EPP and 100K weight.

4.2.7 Genetic effects

The percent phenotypic variance explained by each of the four markers (*p-bnlg1179 p-umc2189, p-umc1542* and *p-bnlg1014*) individually (ranging between 26.2- 38.7%) (Table 4) was significantly lower than the total variance explained by all of them together (65%) (Table 9). This therefore could imply additive genetic effects where each marker was contributing a small amount of effect to the total observed variation. The directions of their effects (Table 9) could imply that the first marker (*p-bnlg1179*) which had a positive value may have been inherited from the P_1 whereas the other three (*p-umc2189, p-umc1542* and *p-bnlg1014*) with negative effects, from the P_2. However, individually, marker *p-bnlg1179* had a negative genetic effect implying that its presence led to reduced ASI values. Furthermore, the sum of the individual effects of these markers is far much more than the total variation explained by all of them together. This could therefore imply the presence of both additive and dominant effects (Sibov *et al.*, 2003).

In this study, the candidate, co-localized QTLs (areas near markers *p-umc2189* and *p-bnlg1014*) for ASI and DS had the same sign for the additive component for the two traits, showing that at these loci, the genetic contribution came from the same parental line for both traits. The co-localized QTLs (areas near markers *p-umc1542* and *p-bnlg1179*) for ASI with both KN and GY did not show the coincidence of the same sign for the additive component with ASI, which may imply that, close linkage of the loci rather than pleiotropy was responsible. However, the markers had the same sign for the additive components for the two traits, KN and GY. Crouzillat *et al.* (2000) noted that a coincidence of both the location of the related traits and the direction of their genetic effects implied pleitropic effects, and otherwise, close linkage of the loci. However, as already mentioned these results are not enough to confirm pleiotropic effects but are rather indicative of the possibility of this fact.

Table 9. Genetic characteristics of QTLs involved in the expression anthesis-silking interval (ASI) under stressed conditions

Marker	Multiple effects	Total variance
p-bnlg1179	0.379163	
p-umc2189	-3.70469	65%
p-umc1542	-3.52095	
p-bnlg1014	-4.25868	

A positive value means that it increases the numeric value of the trait

4.2.8 Effectiveness of bulked segregant analysis in identifying ASI QTLs

Whereas BSA has been said not to be adequate in identifying all the genetic factors controlling a trait, it has been shown to be successful in tagging QTL of very large effects (Mackay and Caligari, 2000). The results obtained from this study showing the effects attributable to each candidate marker as ranging from 25-39% (Table 4), and their total effects of 65% (Table 9), therefore support this fact. In the same regard, Grattapaglia *et al.* (1996) reported an agreement between the results of interval mapping analysis using MAPMAKER/QTL and those of linear regression as used in this study. They however noted that interval mapping detected more QTLs, which BSA could not due to the possibility of sampling effects in the bulk composition. Ribaut *et al.* (1996) also observed that all QTLs detected by MAPMAKER/QTL under intermediate stress and severe stress were also consistently detected with a one-way ANOVA.

In this study, BSA was successful in identifying four candidate markers linked to ASI QTLs. Though 30 markers had been selected as polymorphic after BSA in this study, only a small percentage of the total scorable markers (5.4%) were found to co-segregate with ASI after screening the 30 markers on individuals constituting each bulk. Similar results have been reported in earlier studies. In their work, Grattapaglia *et al.* (1996) *a posteriori* assayed individual samples that composed the bulks for markers found to be in QTL regions and noted that one individual in 10 carrying the marker of interest could "contaminate" an otherwise null genotype bulk thus leading to "false positives".

However, the results as reported here may not be all conclusive as they are based on a subset of the F_2 population selected from the two tails of the segregation continuum. Map positions of polymorphic loci and the localization of QTLs for the trait cannot be established

at this level, and so is the definite determination of allelic effects. These would require genetic analysis and linkage mapping of the data from the whole F_2 population with the selected candidate markers using the available computer programs such as the MAPMAKER and MAPMAKER/QTL (Ribaut *et al.*, 1996; Grattapaglia *et al.*, 1996; Guo *et al.*, 2008), among others. Mendelian segregation ratios for the candidate markers also need to be established using assays on the whole F_2 populations.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

Attaining a high grain yield is the ultimate objective for all maize breeding programs but its evaluation and improvement are complicated by the complex maize biology, environmental interactions and low heritability. In maize, DA and DS are relatively easy to measure in the field hence ASI can be calculated with precision. The heritability of ASI under drought is also high. However, the major problem of conventional selection is in the management of uniform environmental conditions. DNA-based molecular markers allow for the identification of QTLs that determine the phenotypic value of the trait and allow for the analysis and interpretation of the cause-effect relationship among traits. The use of such markers in breeding programs could therefore offer a more powerful tool for selection. In view of this and based on the results presented in this study, the following conclusions and recommendations are made.

5.1 Conclusisons

1. The population used in QTL identification was developed from local landraces (KCB and GBK032357). The relatively normal distribution of drought tolerance traits in F_2 population measured in the field indicated that the local landraces contain novel genes for drought tolerance. Such genes could therefore be useful if incorporated in maize breeding programs. The landraces could also be adopted for use as open pollinated cultivars thus making them essential components of sustainable agriculture.
2. Only four candidate QTLs (*p-umc2189, p-bnlg1179* and *p-bnlg1014* on chromosome 1 and *p-umc1542* on chromosome 2) with sharp effects were identified. These candidate QTLs may be useful in developing molecular marker assisted selection stragegies to transfer drought tolerance traits from the local landraces into the elite varieties.
3. BSA is effective in identifying genomic regions linked to traits of interest.

5.2 Recommendations

1. The candidate markers identified should be screened in the entire F_2 populations and further analysis be carried out to confirm and Map these QTLs.
2. Further screening should be carried out in the population using more primers to identify more QTLs linked to anthesis-silking interval.
3. The candidate QTLs should be tested in varied environments to establish their stability.

4. The mapping population used in this study should be screened for QTLs linked to other traits of interest like yield and its components, male flowering time, female flowering time, among others.

CHAPTER SIX

REFERENCES AND APPENDICES

6.1 REFERENCES

Agrama, H.A.S. and Moussa, M.E. (1996). Mapping QTLs in breeding for drought tolerance in maize (*Zea mays* L.). *Euphyt* **91**: 89-97.

Ahmed, Z. (2006). Optimization of PCR conditions in-vitro for maximum amplification of DNA from *Xanthomonas campestris*. *J. Appl. Sci. Res.* **2**(3): 112-122

Andrade, F.H., Echarte, L., Rizalli, R., Della Maggiora, A. and Casanovas, M., (2002). Kernel number prediction in maize under nitrogen or water stress. Crop Sci. 42, 1173–1179.

Andrade, F.H., Vega, C., Uhart, S., Cirilo, A., Cantarero, M. and Valentinuz, O (1999). Kernel number determination in maize. *Crop Sci.* **39**: 453-459.

Babu, R., Sudha, K., Nair, B., Prasanna, M. and Gupta, H.S. (2004). Integrating marker-assisted selection in crop breeding- prospects and challenges. *Current Sci.* **87** (5): 607-619.

Bänziger, M. and Lafitte, H.R. (1997). Efficiency of secondary traits for improving maize for low-nitrogen target environments. *Crop Sci.* **39**, 1035-1040

Banziger, M., Edmeades, G.O., Beck, D. and Bellon, M. (2000). Breeding for Drought and Nitrogen Stress Tolerance in Maize: From Theory to Practice. Mexico DF CIMMYT

Barcaccia, G., Lucchin, M and Parrini, P. (2003). Characterization of a flint maize (*Zea mays* var. indurata) Italian landrace: II . Genetic diversity and relatedness assessed by PCR-based markers. *Gen. Res. Crop Evol.* **50**(3): 253-271

Barcaccia, G., Pallottini, L., Arzenton, F. and Parrini, P. (2000). Comparing linkage maps of flint maize (*Zea mays* var.indurata) Italian landrace based on an F_1 and BC_1 populations. Atti del XLIV convegno societa italiana di genetic agravia, Bologna, Italy, 20-23 September. pp 120-121.

Barcaccia, G., Pallottini, L., Parrini, P. and Lucchin, M. (2005). A genetic linkage map of flint maize (*Zea mays* var. indurata. L.) Italian landrace using one-way pseudo-testcross strategy and multilocus PCR-based markers. *Maydica* **50** (3): 253-271

Barker, T., Campos, H., Cooper, M., Dolan, D., Edmeades, G.O., Habben, J., Schussler, J., Wright, D., and Zinselmeier, C. (2005). Improving drought tolerance in maize. *Plt. Breed. Rev.* **25**: 173-253

Bertolini, M., Bianchi, A., Lupotto, E., Salamini, F., Verderio, A., and Motto, M. (1998). Maize. In: Scarascia, G., Mugnozza, T. and Pagnotta, M.A. (eds). Italian contribution to plant genetics and breeding.viterbo, Italy. Quantrini, A. and F. Publisher. pp. 209-229

Beyene, Y., Botha, A. and Alexander, A.M. (2005). A comparative study of molecular and morphological methods of describing genetic relationships in traditional Ethiopian highland maize. *Afr. J. Biotech.* **4**: 586-595

Blum, A. (1996). Crop Responses to Drought and the Interpretation of Adaptation. *Plant Growth Res.* **20**:135-148.

Blum, A. (2000). Drought Resistance in Complex Traits. Proc of an Intl. Workshop of Field Screening for Drought Tolerance in Rice. Saxena, N.P. & O'Toole, J.C. (eds.) pp 17-22

Bolanos, J. and Edmeades, G.O. (1996). The Importance of the Anthesis – Silking Interval in Breeding for Drought Tolerance in Tropical Maize. *Field Crops Res.* **48**: 65-80

Bonhof, M.J., Hulls, A., Kiros, F.G. and Dibogo, N (2001). Farmers' Perceptions of Importance, Control Methods and Natural Enemies of Maize Stem Borers at the Kenya Coast. *Insects Sci. Appl.* **21**: 33-42.

Boyer, J.S. and Westgate, M.E. (2004). Grain yields with limited water. *J. Expt. Bot.* **55**(407): 2385-2394.

Brauer, J.M. Christianson, M.C., Pai, A.D. and Dunham, J.M. (2006). Mapping novel traits by array-assisted bulk segregant analysis in *Saccharomyces cerevisiae*. *Genet.* **173**: 1813-1816.

Bruce, W.B., Edmeades, G.O, and Barker, T. (2002). Molecular and physiological approaches to maize improvement for drought tolerance. *J. Exp. Bot.* **53**: 13-25.

Brush, S.B. (1995). *In situ* Conservation of Landraces in Centers of Crop Diversity. *Crop Sci.* **35**(2): 346-354.

Byrne, P.F., Bolanos, J., Edmeades, G.O. and Eaton, D.L. (1995). Grains from Selection Under Drought versus Multillocation Testing in Related Tropical Maize Populations. *Crop Sci.* **35**: 63-69.

Campos, H., Copper, M., Habben, J.E., Edmeades, G.O., and Schussler, J.R (2004). Improving drought tolerance in maize: A view from industry. *Field Crops Res.* **90**: 19-34

Cardle, L., Ramsay, L., Melbourne, D., Macaulay, M., Marshall, D. and R. Waugh. (2000). Computational and Experimental Characterization of Physically Clustered Simple Sequence Repeats in Plants. *Genet.* **156**: 847-854.

Carvalho, V.P., Ruas, P.M., Ruas C.F., Ferreira, J.M. and Moreira, R.M.P. (2002). Assessment of genetic diversity in maize (*Zea mays* L.) landraces using inter simple sequence repeat (ISSR) markers. *Crop Breed. Appl. Biotech.* **2**: 557-568

Chapman, S.C. and Edmeades, G.O. (1999). Selection improves drought tolerance in tropical maize population: II. Direct and correlated responses among secondary traits. *Crop Sci.* **39**, 1315-1324.

CIMMYT, (1996). Molecular Marker Application to Plant Breeding. Workshop Manual, CIMMYT HQs, ElBatan, Mexico.

CIMMYT, (2003). Conservation by numbers. Reducing genetic drift in crop gene bank collections. CIMMYT news Vol. 2 no. 4

CIMMYT, (2006). Conservation by numbers. Reducing genetic drift in crop gene bank collections. CIMMYT news Vol. 3 no. 1

Corbett, J.D (1998). Classifying Maize Production Zones in Kenya through Multivariate Cluster Analysis. In: Hassan R.M (ed) Maize Technology Development and Transfer A GIS Application For Research Planning In Kenya. CAB International wallingford, UK, pp, 15-25.

Crouzillat D., Menard, B., Mora, A., Phillips, W. and Petiard, V. (2000). Quantitative trait analysis in *Theobroma cacao* using markers. Yield QTL detection and stability over 15 years. *Euphy.* **114**: 13-23.

De Groote, H., Bett, C., Mose, L., Mugo, S., Odendo, M. Okuro, J.O and Wekesa, E. (2001) Assessing the Demand for Insect Resistant Maize Varieties through Participatory Rural Appraisals (PRA). In: Insects Resistant Maize for Africa Project Annual report 2000. KARI/CIMMYT IRMA Project, IRMA Project documents No.4 Nairobi

de Vicente, M.C. Guzman, F.A., Engels, J. and Rao, V.R. (2006). Genetic characterization and its use in decision making for the conservation of crop germplasm. In: Ruane, J. and Sonnino, A. (eds.). The role of biotechnology in exploring and protecting agricultural genetic resources. FAO, Rome. pp87-95

Dean, R.E., Dahlberg, J.A., Hopkins, M.S., Mitchell, S.E. and Kresovich, J. (1999). Genetic redundancy and diversity among "orange" accessions in the US national sorghum collections as assessed with SSR markers. *Crop Sci.* **39**: 1215-1221

Dekkers, J.C.M and Hospital, F (2002). The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* **3**: 22-32

Dowswell, C.R. Paliwal, R.L. and Cantrell, R.P. (1996). Maize in the Third world. West view Press, a div. of Harper Collins Publishers pp 108-126

Earl., H.J and Davis, R.F. (2003). Effects of drought stress on leaf and whole canopy radiation use efficiency and yield of maize. *J. Agron.* **95**(3): 688-696.

Edmeades, G.O., Banziger, M., Schussler, J.R., Campos, H., (2004). Improving abiotic stress tolerance in maize: a random or planned process? In: Proceedings of the Arnel R. Hallauer International Symposium on Plant Breeding. Mexico City, 17–22 August (2003), Iowa State University Press, in press.

Edmeades, G.O., Bolan˜os, J., Elings, A., Ribaut, J.-M., Ba¨nziger, M., Westgate, M.E., (2000). The role and regulation of the anthesis-silking interval in maize. In: Westgate, M.E., Boote, K.J. (Eds.), Physiology and Modeling Kernel Set in Maize. CSSA, Madison, WI, CSSA Special Publication No. 29, pp. 43–73.

Edmeades, G.O., Bolaños, J. and Chapman, S.C. (1997). Value of secondary traits in selecting for drought tolerance in tropical maize. *In* Edmeades, G.O., Bänziger, M., Mickelson, H.R. & Pena-Valdiva, C.B. (Eds.), Developing Drought and Low-N Tolerant Maize. *Proceedings of a Symposium*, March 25-29, 1996, CIMMYT, El Batan, Mexico. Mexico, D.F.: CIMMYT, pp. 222-234.

Edmeades, G.O., Bolanos, J., Chapman, S.C., Lafitte, H.R. and Banziger, M. (1999). Selection improves drought tolerance in tropical maize populations. II. Direct and correlated responses among secondary traits. *Crop Sci.* **39**: 1315-1324

Edmeades, G.O., Chapman, L., Bellon, M., Banziger, M. and Lafitte, H.R. (1994). Recent Evaluation of Progress in Selection for Drought in Tropical Maize. Proc. the 4th Eastern and Southern Africa Regional Maize Conference. Harare, Zimbabwe. April, 1994 pp 94-100

Edmeades, G.O., Schussler, J., Campos, H., Zinselmeier, C., Habben, J., Collinson, S., Cooper, M., Hoffbeck, M., Smith, O., (2003). Increasing the odds of success in selecting for abiotic stress tolerance in maize. In: Birch, C.J., Wilson, S.R. (Eds.), In: Proceedings of the 5th Australian Maize Conference (Versatile Maize – Golden Opportunities), Toowoomba, Maize Assoc. of Australia, February 18–20, pp. 16–28.

FAO (2006). The Role of Biotechnology in Exploring and Protecting Agricultural Genetic Resources. FAO, Rome

FAO, 2002. FAOSTAT data 2002, faostat.fao.org. Nov. 2005

FAO, 2005. FAOSTAT data 2005, faostat.fao.org. Feb. 2006

FAO/ IAEA, (2002). Mutant Germplasm Characterization Using Molecular Markers. A Manual Prepared by Joint FAO/ IAEA Division of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna Austria. pp1-87

Ferguson, M.E., Bramel, P.J. and Chandra, S. (2004). Gene diversity among botanical varieties in peanut (*Arachis hypogaea*. L.). *Crop Sci.* **44**: 1847-1854

Ferreira, M.E. (2006). Molecular analysis of genebanks for sustainable conservation and increased use of crop genetic resources. In Ruane, J. and Sonnino, A. (eds.). The role of biotechnology in exploring and protecting agricultural genetic resources. FAO, Rome. pp105-114

Figliuolo, G. and Perrino, P. (2004). Genetic diversity and intra specific phylogeny of *Triticum turgidum* L. sub sp. dicoccon (Schrank Thell.) revealed by RFLPs and SSRs. *Genet. Res. Crop. Envt.* **51**: 519-527

Frova, C., Krajewski, P., Fonzo, N.D., Villa, M, and Sari-Gorla, M. (1999). Genetic analysis of drought tolerance in maize by molecular markers. I. Yield components. *Theor Appl. Genet.* **99**: 280-288.

Gaut, B.S., d'Ennequin, M.L.T., Peek, A.S. and Sawkins, M.C. (2000). Maize as amodel for the evolution of plant nuclear genomes. *PNAS* **97**: 7008-7015.

Gawel N.J. and Jarret, R.L (1991). A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. *Pl. Mol. Bio. Reporter*. **9**: 262-266.

Gebre, G.B. (2005). Genetic variability and inheritance of drought and plant density adaptive traits in maize. PhD thesis, University of Free State, South Africa.

Ghesquiere, M., Barre, P.H., Marhadour, S and Kerlan, M.C. (2000). Estimation of introgression rate of a fescue isozymic marker into tetraploid Italian rye grass at early generations of backcross. *Euphyt.* **114**(3): 223-231.

GOK, (2004). Strategy for Revitalizing Agriculture 2004-2014.

Grattapaglia, D., Bertolucci, F.L.G., Penchel, R. and Sederoff, R. R. (1996). Genetic Mapping of Quantitative Trait Loci controlling growth and wood quality traits in *Eucalyptus grandis* using maternal half-sib family and RAPD markers. *Genet.* **144**: 1205-1214.

Guo, J., Su, G., Zhang, J., and Wang, G. (2008). Genetic analysis and QTL mapping of maize yield and associate agronomic traits under semi-arid land condition. *Afr. J. Biotech.* 7(12): 1829-1838.

Guo, M., Rupe, M.A., Zinselmeier, C., Habben, J., Bowen, B.A., Smith, O.S. (2004). Allelic variation of gene expression in maize hybrids. *Pl. Cell*, **16**: 1707-1716.

Hackett, C.A., Wachira, F.N., Paul, S., Powell, W and Waugh, R. (2000). Construction of a genetic linkage map for *Camellia sinensis* (tea). *Heredity*. **85**: 346-355

Hassan, R.M., Corbett, J.D and Njoroge, K. (1998a). Combining Geo-referenced Survey Data with Agro-climatic Attributes to Characterize Maize Production Systems in Kenya .In: Hassan,R.M. (ed).Maize Tech. Devt. And Transfer: a GIS Aplication for Research Planning in Kenya. CAB International, Wallingford, UK, pp. 43 –68.

Hassan, R.M., Njoroge, K., Njore, M., Ostyula, R. and Laboso, A., (1998b). Adoption Patterns and Performances of Improved Maize in Kenya. In: Hassan R.M (ed) Maize Technology Development and Transfer A GIS Application for Research Planning in Kenya. CAB International wallingford, UK, pp 107-136.

Henry, A.M. and Damerval, C. (1997). High rates of polymorphism and recombination at *Opaque-2* locus in cultivated maize. *Mol. and General Genet*. **256:** 147-157

Hillel, D. and Rosenzweig, C. (2002). Desertification in relation to climate variability and change. *Adv. Agron.* **77**: 1-38

Hokanson, S.C., Szewc-McFadden, A.K., Lamboy, W.F and McFerson, J.R. (1998). Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a mallus x domestica borkh core subset collection. *Theor. Appl. Genet.* **97**: 671-683.

Holton, T.A., Cross, M., Karakousis, A. and Henry, R.J. (1999). Genetic Screening of Australian Barley Varieties Using Microsatellite Markers. In: Proc.of the 9th Australian Barley Technical Symposium. Pp64-76

Ilarslan, R., Kaya, Z., Tolun, A.A., and Bretting, P.K. (2001). Genetic variability among Turkish pop, flint and dent corn (*Zea mays* L. spp. mays) races: Enzyme polymorphisms. *Euphyt*. **122**(1): 171-179.

IRMA, (2001). IRMA updates, December 2001, volume 2, issue 4. www.cimmyt.org/ABC--February, 2006.

IRMA, (2002). IRMA updates August 2002 volume 3, issue2. www.cimmyt. org /ABC--2006, February.

Isshiki, S. and Taura, T. (2003). Fertility restoration of hybrids between *Solanum melongena* L. and *S. aethiopicum* L. Gilo group by chromosome doubling and cytoplasmic effect on pollen fertility. *Euphyt*. **134**(2): 195-201

Jeanneau, M., Gerentes, D., Foueillassar, X., Zivy, M., Vidal, J. and Toppan, A. (2002). Improvement of drought tolerance in maize: towards the functional validation of the Zm-Asr1 gene and increase of water use efficiency by over-expressing C4-PEPC. *Biochimie*, **84**: 1127-1135.

Kahl, G. (2001). The dictionary of gene technology. Wiley-VCH, Weinheim. Pp95-104

KARI. (2004). Masongaleni Commercialization Plan. Agricultural Research Investments and Services, Nairobi.

Karp, A. (2002). The new genetic era: will it help us in managing genetic diversity? In. Engels, J.M.M., Ramanatha, R.V., Brown, A.H.D. and Jackson, M.T. (eds.) Managing plant genetic diversity. Rome, IPGRI. pp 43-56.

Kelly, J.D., Gepts, P., Miklas, P.N. and Coyne, D.P. (2003). Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. *Field. Crops Res.* **82**(2/3): 135-153

Kitavi, M., Ininda, J. Muluvi, G. and Danson, J. (2005). Application of Molecular Markers for the Improvement of Maize Germplasm of the Kenya Agricultural Research Institute for Fungal Disease Resistance. In: Biotechnology, Breeding and Seed Systems for African Crops- The Rockefeller Foundation and KARI. Nairobi, Kenya. 24th-27th Jan. 2005.

Kraakman, A.T.W., Martinez, F., Mussiraliev, B., VanEeuwijk, F.A. and Niks, R.E (2006). Linkage disequilibrium mapping of morphological resistance and other agronomically relevant traits in modern spring barley cultivars. *Mol. Breed.* **17**: 41-58.

Laffite, R. (2002). Relationship between Leaf Relative Water Content during Reproductive Stage, Water Deficit and Grain Formation in Rice. *Field Crop Res.* **76**: 165-74

Lanteri, S. and Barcaccia, G. (2006). Molecular marker based analysis for crop germplasm preservation. In: Ruane, J. and Sonnino, A. (eds.). The role of biotechnology in exploring and protecting agricultural genetic resources. FAO, Rome. pp 105-120

Li, X., Liu, X., Li, M and Zhang, S. (2003). Identification of quantitative trait loci for anthesis-silking interval and yield components under drought stress in maize. *Acta Bot. Sci.* **45**(7): 852-857

Li, X-H., LI, X-H., Hao, Z-F., Tian, Q-Z., and Zang, S-H. (2005). Consensus map of the QTL relevant to drought tolerance of maize under drought conditions. *Sci. Agric. Sinic.* **38**: 882-890.

Lucchin, M. Barcaccia, G., and Parrini, P. (2003). Characterization of a flint maize (Zea mays var. indurata) Italian landrace. I. Morphological and agronomic traits. *Genet. Res. Crop Evol.* **50**(3): 315-327.

Mackay, I.J. and Caligari, P.D.S (2000). Effieciencies of F2 and Backcross generations for bulked segregant analysis using dominant markers. *Crop Sci.* **40**: 625-630

Maestri, E., Malcervschi, A., Massari, A. and Marmiroli, N. (2002). Genome analysis of cultivated barley (*Hordeum vulgare*) using sequence-tagged molecular markers. Estimates of divergence based on RFLP and PCR markers derived from stress responsive genes and simple-sequence repeats (SSRs). *Mol. Genet. Genomics* **267**(2): 186-201.

Martinez-Morales, J., Naruse, K., Mitani, H., Shima, A and Wittbrodt, J. (2004). Rapid chromosomal assignment of Medaka mutants by bulk segregant analysis. *Gene* **329**: 159-165.

Mati, B.M. (2000). The Influences of Climate Change on Maize Production in the Semi-humid Semi-arid Areas of Kenya. *Arid Envts. J.* **46**: 333-344.

Mondal, T.K. (2002). Assessment of genetic diversity of tea (*Camellia sinensis* (L.) O. Kuntze) by inter-simple sequence repeat polymerase chain reaction. *Euphyt.* **128**: 307-315.

Monneveux, P., Sanchez, C., Beck, D and Edmeades, G.O. (2006). Drought tolerance improvement in tropical maize source populations: evidence of progress. *Crop Sci.* **46**: 180-191.

Mugo, S. N., Smith, M.E., Banziger, M., Setter, T.I., Edmeades, G.O and Elings, A. (1998). Performance of early maturing Katumani and Kito maize composites under drought at the seedling and flowering stages. *Afr. Crop Sci.* J. **6**(4). 329-334.

Munyiri, S.W. (2008). Phenotypic characterization of selected local maize landraces for drought tolerance based on flowering stress and their resistance to larger grain borer in Kenya. M. Sc. Thesis, Egerton University.

Muoki, R.C., Wachira, F.N,. Pathak, R.S. and Kamunya, S.M. (2008). Assessment of mating system of *Camellia sinensis* in biclonal seed orchards based on PCR markers. *J. Hort. Sci. Biotech.* **82**: 733-738

Mutinda, C., Diallo, A.O. and Manyara, F. (2004). Meeting Challenges of Breeding for Improved Drought Tolerance and Other traits in Maize in Kenya. In: Resilient Crops for Water Limited Environments Workshop May 24th-28th, 2004. Cuernavaca Mexico. (Abs.). p 120

Negash, A., Tsegaye, A., Van Treren, R. and Visser, B. (2002). AFLP analysis of enset clonal diversity in south and south western Ethiopia for conservation. *Crop Sci.* **42**: 1105-1111

Ngugi, K. and Muthamia, Z. (2004). Molecular Marker Assisted Selection for Drought Tolerance in Maize Population Bred for Semi-arid Kenya. In: Resilient Crops for Water Limited Environments. Workshop Proc. May 24th-28th, 2004. Cuernavaca, Mexico. (Abs.). p 265

Nguyen, T.T., Klueva, N., Chamareck, V., Aarth, A., Magpantay, G., Millena, AC., Pathan, M.S. and Nguyen, H.T (2004). Saturation mapping of QTL regions and identification of putative candidate genes for drought tolerance in rice. *Mol. Genet. Genomics.* **272**:35-46

Nwakanma, D.C., Pillay, M., Okoli, B.E. andTenkouano, A. (2003). Sectional relationships in the genus *Musa*. L. inferred from the PCR-RFLP of organelle DNA sequences. *Theor. Appl. Genet.* **107**: 850-856.

Padmalatha, K. and Prasad, M.N.V. (2006). Optimization of DNA isolation and PCR protocol for RAPD analysis of selected medicinal and aromatic plants of conservation concern from Peninsular India. *Afr.J. Biotech.* **5**(3): 230-234.

Pallotini, L. (2002). Variation of the Genetic Structure of Populations in Relation to Different Strategies of Germplasm Conservation in Local Varieties of Maize. Pandua, Italy, University of Pandua, pp.316 (PhD thesis)

Pandey, S. (2000). Maize Research Highlights Bulletin. Mexico D.F. CIMMYT

Papa, R. and Gepts, P. (2003). Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris*. L.) from Meso-America. *Theor. Appl. Genet.* **106**: 239-250.

Passioura, J.B. (1996). Drought and Drought tolerance. In: Belhassen, E. (ed) Drought Tolerance in Higher Plants: Genetical, Physiological and Molecular Bio-analysis. Kluwer Academic Publisher, Dordrecht, the Netherlands pp 1-7

Perales, R.H (2003). Dynamic Management of Maize Landraces in Central Mexico. *Econ. Bot.* **57**(1): 21-34

Perera, L., Russell, J.R., Provan, J. and Powell, W. (2000). Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera*. L.). *Genome* **43**: 15-21

Pingali, P.L. and Pandey, S. (2000). Meeting World Maize Needs. Technological Opportunities for the Public Sector. CIMMYT World Maize Facts and Trends Mexico D.F CIMMYT. p 60.

Powell, W., Morgante, M., Andre, C., Hanafay, M., Vogel, J., Tingey, S, and Rafalsk, A. (1996). The utility of RFLP, RAPD, AFLP and SSRP (microsatellite) markers for germplasm analysis. *Mol. Breed.* **2**: 225-238.

Quan, R., Shang, M., Zhang, H., Zhao, Y., Zhang, J. (2004). Engineering of enhanced glycine betaine synthesis improves drought tolerance in maize. *Plant Biotechnol. J.* **2**: 477-486.

Rahman, M.H., Jaquish, B. and Khasa, P.D. (2000). Optimization of PCR protocol in microsatelite analysis with silver and SYBR. *Pl. Mol. Bio. Reporter* **18**: 339-348

Rahman, M.U., Malik, T.A, Aslan, N., Asif, M., Ahmad, R., Khan, I.A. and Zafar, Y. (2002). Optimization of PCR conditions to to amplify microsatellite loci. *Intl. J. Agric. and Bio*. **4**(2): 282-284

Republic of Kenya, (2001). Agriculture Sector Review. Ministry of Agriculture, Nairobi pp23-25

Ribaut, J., Jiang, C. and Gonzalez-de-Leon, C. (1997). Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. *Theor. Appl. Genet.* **94**: 887-896.

Ribaut, J.M. and Hoisington, D. (1998). Marker Assisted Selection: New Tools and Strategies. *Tren. Pl. sci.* **3**: 236- 239.

Ribaut, J.M., Banziger, M., Bertran, J., Jiang, C., Edmeades, G., Dreher, K. and Hoisington, H. (2002). Use of Molecular Markers in Plant Breeding: Drought Tolerance Improvement in Tropical Maize. In: Manjit, S.K. (ed.). Quantitative Genetics, Genomics and Plant Breeding. Wallingford, UK: CABI Publishing pp 85-99.

Ribaut, J.M., Hoisington, D.A., Deutsch, J.A., Jian, C. and Gonzalez de Leon. (1996). Identification of QTL under Drought Conditions in Tropical Maize. 1. Flowering Parameters and the ASI. *Theor.Appl. Genet.* **92**(7): 905-914.

Richards, A.R. (2005). Physiological traits used in the breeding of new cultivars for water-scarce environments. *Agric.Water Mgt.* **80**: 197-211.

Risch, N.J. (2000). Searching for genetic determinants in the new millennium. *Nature.* **405**(15). 847-856.

SAS Institute, (2001). SAS Language Guide for Personal Computers. Edition 8.2 Cary, N.C. (ed.).

Schussler, J.R. and Westgate, M.E. (1995). Assimilate flux determine kernel set at low water potential in maize. *Crop Sci.* **35**: 1074-1080.

Shim, S.I. and Jorgensen, R.B. (2000). Genetic structure in cultivated wild carrots (*Daucus carota.* L.) revealed by AFLP analysis. *Theor. Appl. Genet.* **101**: 227-233.

Sibov, S.T., de, S.C., Garcia, A.A., Silva, A.R., Garcia, A.F. and Mangolin, C.A. (2003). Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 2. Quantitative trait loci (QTL) for grain yield, plant height, ear height and grain moisture. *Hered.* **139**: 107-115.

Skovmand, B., Reynolds, M.P. and Relay, I.H. (2001). Mining Wheat Germplasm Collections for Yield Enhancing Traits. *Euphyt.* **119**(1-2): 2001 25-32

Songa, J.M., Overholt, W.A. Mueke, J. M. and Okello, R.O. (2002). Farmers' Perceptions of Aspects of Maize Production Systems and Pests in Semi-arid Eastern Kenya: Factors Influencing Occurrence and Control of Stem borers. *Intern. J. Pest Mgt.* **48**: 1-11.

Songa, J.M., Overholt, W.A. Mueke, J. M. and Okello, R.O. (2002). Farmers' Perceptions of Aspects of Maize Production Systems and Pests in Semi-arid Eastern Kenya: Factors Influencing Occurrence and Control of Stem borers. *Intern. J. Pest Mgt.* **48**: 1-11.

Tabor, G. M., Kubisiak, T. L., Klopfensteio, N. B., Hall, R. B., and McNabb, H. S., Jr. (2000). Bulked segregant analysis identifies molecular markers linked resistance in *Populus deltoides. Phytopathology* **90**: 1039-1042.

Tinker, N.A (2002). Why quantitative genetics should care about bioinformatics. In: Kang M.S (ed.). quantitative genetics, genomics and plant breeding. CABI, Wallingford, pp. 33-44

Tollenaar, M. and Wu, J., (1999). Yield improvement in temperate maize is attributable to greater stress tolerance. *Crop Sci.* **39**: 1597-1604.

Tuberosa, R. and Salvi, S. (2006). Genomics-based approaches to improve drought tolerance of crops. *Trends in Plt. Sci.* **11**(8): 1360-1385

Tuberosa, R., Sanguineti, M.C., Landi, P., Giuliani, M.M., Salvi, S. and Conti, S. (2002). Identification of QTLs for root characteristics in maize grown in hydroponics and analysis of their overlap with QTLs for grain yield in the field at two water regimes. *Plant Mol. Biol* **48**: 697-712.

Vinod, M.S., Sharma, N., Anjunatha, K.M.A., Kanbar, A., Rakash, N.B.P and Hashidhar, H.E.S (2006). Candidate genes for drought tolerance and improved productivity in rice (*Oryza sativa* L.). *J. Bio. Sci* **31**(1): 69-74

Vos, P., Hogers, R., Bleaker, M., Reijans, M., Lee, T. and Homes, M (1995). AFLP: A New Technique for DNA Fingerprinting, *Nucl. Acids Res.* **23**:4407-4414

Wachira, F.N. (1996). Development of Molecular Markers in *Camellia.* Ph.D Thesis, Dundee University, UK.

Wachira, F.N., Tanaka, J. and Takeda, Y. (2001). Genetic variation and differentiation in tea (*Camellia sinensis*) germplasm by RAPD and AFLP variation. *J. Hort. Sci. Biotech.* **76**: 557-563

Wekesa, E., Mwangi, W., Verkuiji, H., Donda, K., and De Groote, H. (2003). Adoption of Maize Production Technologies in the Coastal Lowlands of Kenya. International Maize and Wheat Improvement Center (CIMMYT). Pp17-22

Xiao, J., Grandillo, S., Ahn, S.N., McCouch, S.R., Tanksley, S.D., Li, J.M. and Yuan, L.P. (1996). Genes from wild rice improve yield. *Nature*. (London). **384**(6606): 223-224.

Xiao, Y., Xin-Hai, L., Shi-Huang, Z., Xiang-Dong, W., Mng-Shun, L. and Yong-Lian, Z. (2004). Identification of Quantitative Trait Loci (QTLs) for Flowering Time Using SSRMarker in Maize under Water Stress. *Korean J. Genet*. **26**: 405-413.

Xiao, Y.N., Li, X.H., George, M.L., Li, M.S., Zhang, S.H. and Zheng, Y.L. (2005). Quantitative Trait Locus Analysis of Drought Tolerance and Yield in Maize in China. *Plant Mol. Bio. Repr.* **23:** 155–165, June 2005.

Yamada, T., Jones, E.S., Cogan, N.O.I., Vecchies, A.C. Nomura, T., Hisano, H., Shimamoto, Y., Smith, K.F., Hayward, M. D. and Forster, J.W. (2004). QTL analysis of morphological, developmental and winter hardiness-associated traits in perennial ryegrass. *Crop. Sci.* **44**: 925-935

Young, N.D. (1999). A cautiously optimistic vision for marker assisted breeding. *Mol. Breed.* **5**: 505-510.

Yousef, G.G. and Jurik, J.A. (2001). Comparison of Phenotypic and Marker Assisted Selection for Quantitative Traits in Sweet Corn. *Crop Sci.* **41** (3): 645- 655.

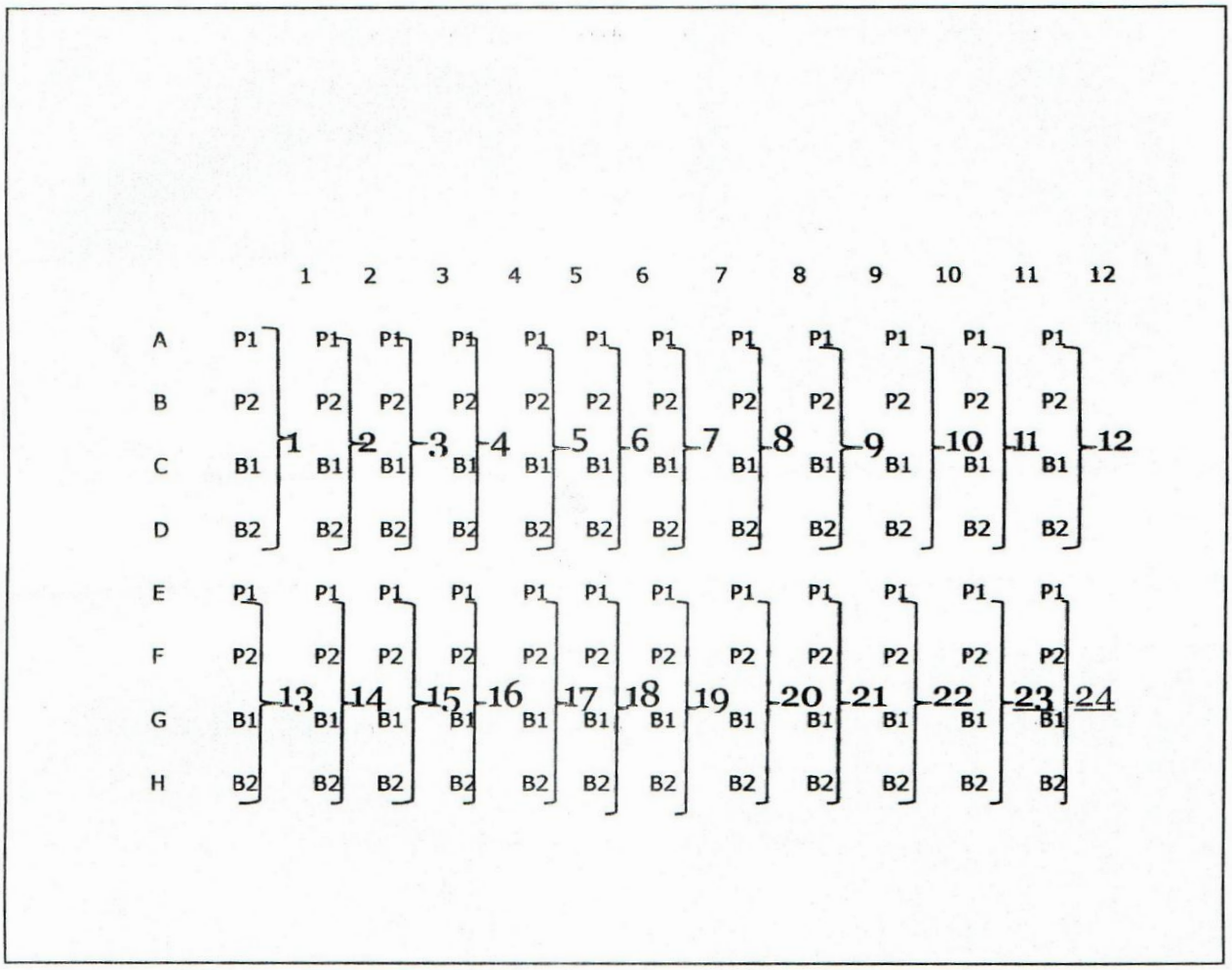
Zhang, S., Li, X., Liu, X., and Li, M. (2003). Identification of quantitative trait loci for anthesis-silking interval and yield components under drought stress in maize. *Acta. Bot. Sinica* **45**(7): 852-857.

Zhang, Z.M., Zhao, M.J., Ding, H.P., Rong, T.Z. and Pan, G.T. (2006). QTL mapping analysis of plant height and ear height of maize (*Zea mays* L.). *Genetika*, 42: 391-396.

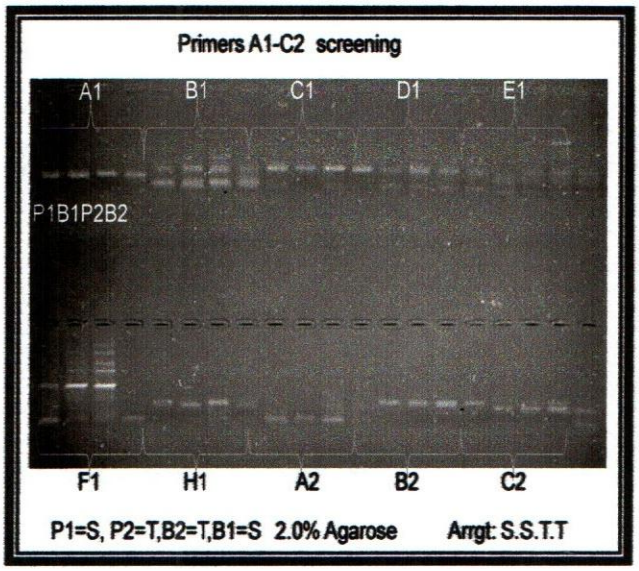
Zidenga, T. (2004). DNA Based Methods in Sorghum Diversity Studies and Improvement. *Pl. cell, Tissue and Organ Cult.* **75**:1-18.

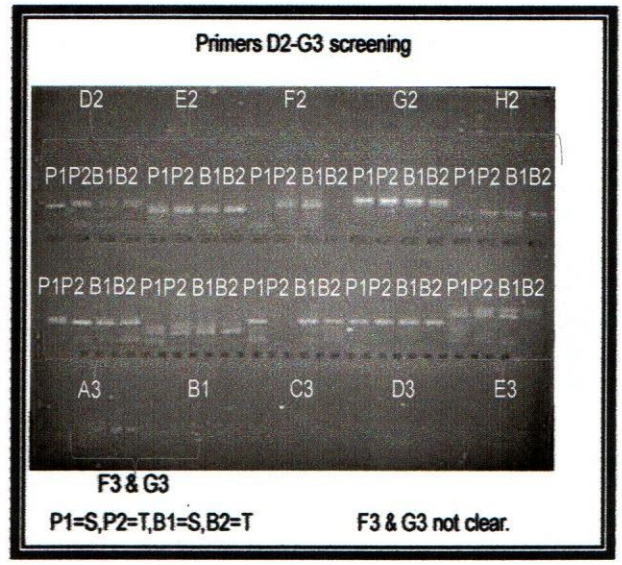
6.2 APPENDICES

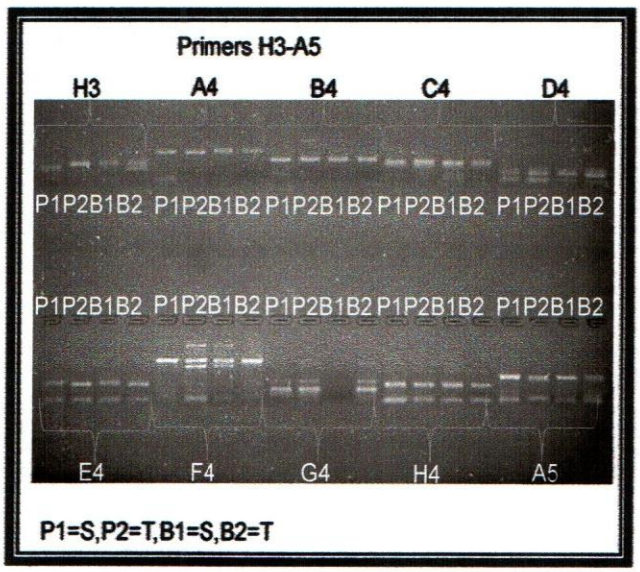
Appendix 1. Primer screening plate layout

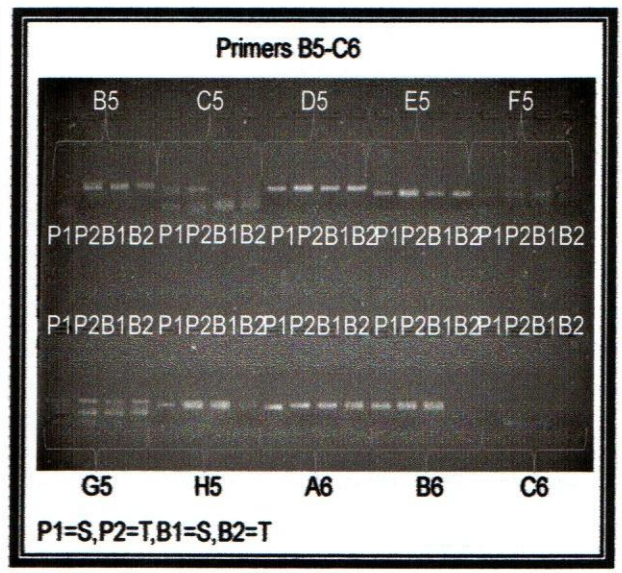


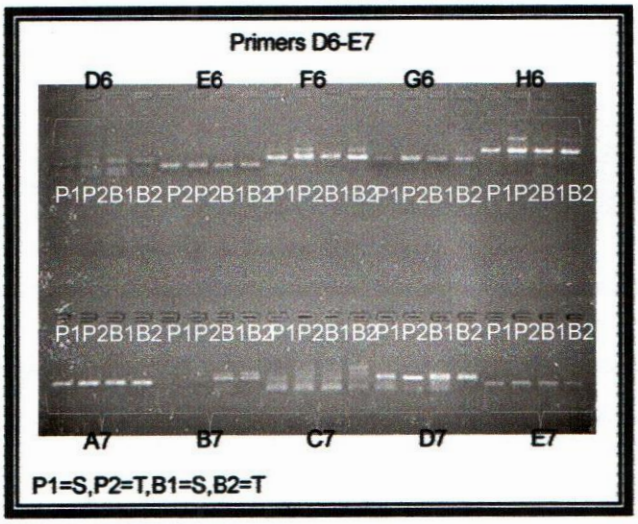
Appendix 2. Primer screening plate photos

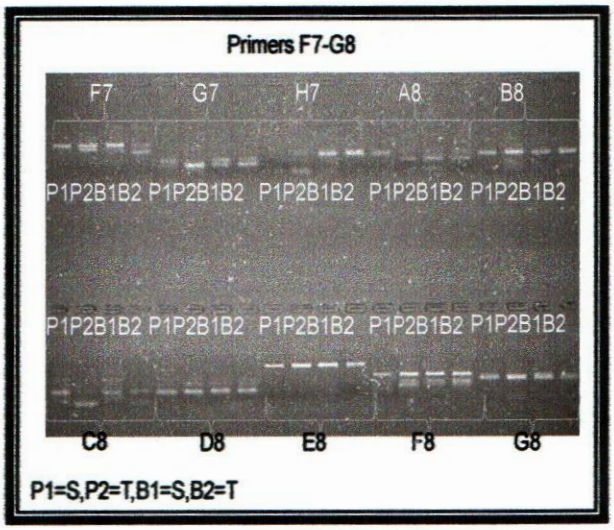


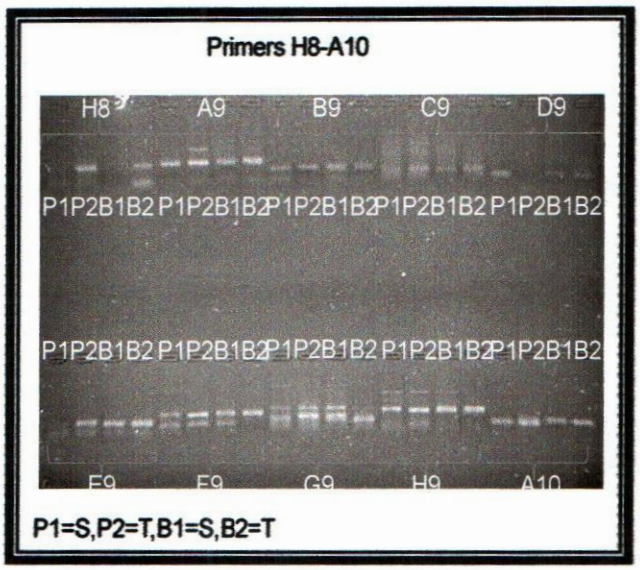


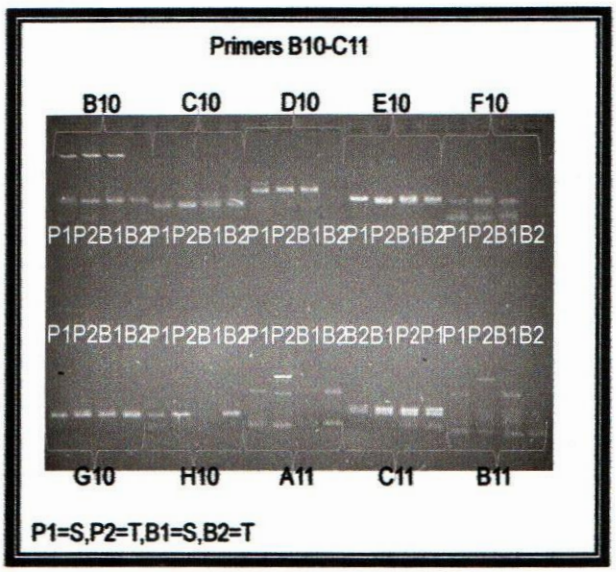




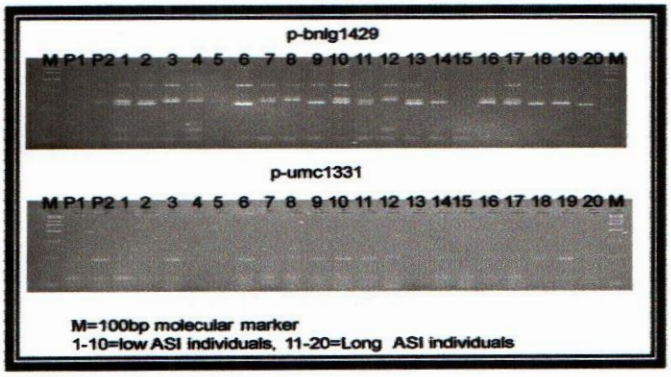

Primers D6-E7
D6 E6 F6 G6 H6
P1P2B1B2 P2P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
A7 B7 C7 D7 E7
P1=S,P2=T,B1=S,B2=T

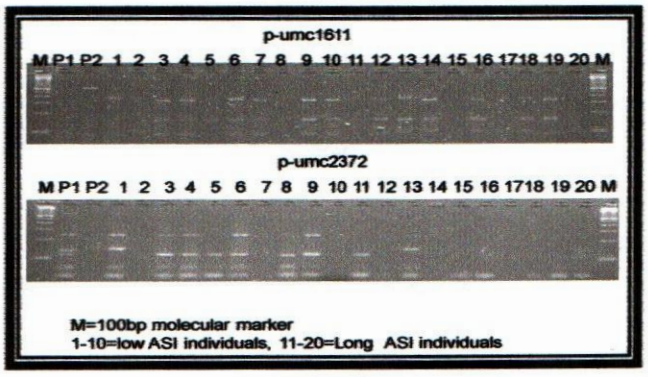

Primers F7-G8
F7 G7 H7 A8 B8
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
C8 D8 E8 F8 G8
P1=S,P2=T,B1=S,B2=T

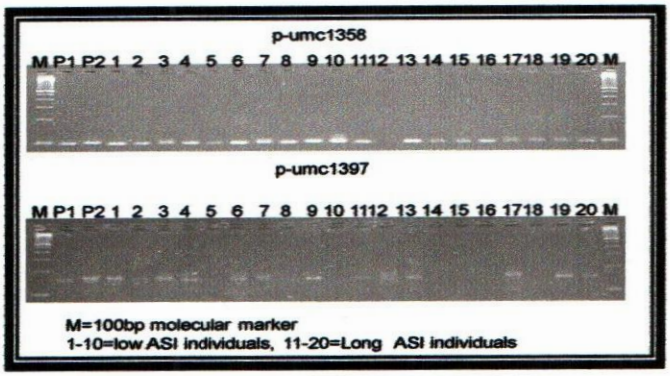

Primers H8-A10
H8 A9 B9 C9 D9
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
P1=S,P2=T,B1=S,B2=T

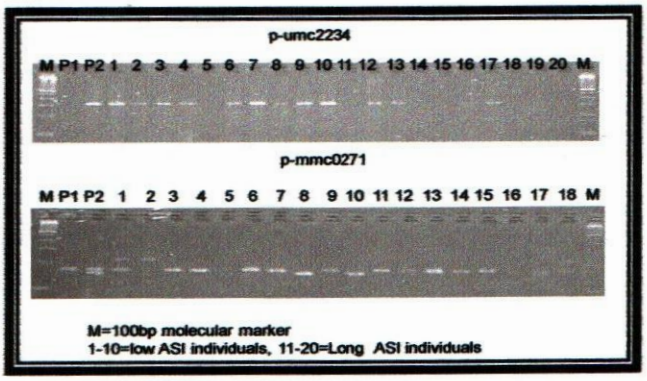

Primers B10-C11
B10 C10 D10 E10 F10
P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2 P1P2B1B2
P1P2B1B2 P1P2B1B2 P1P2B1B2 B2B1P2P1 P1P2B1B2
G10 H10 A11 C11 B11
P1=S,P2=T,B1=S,B2=T

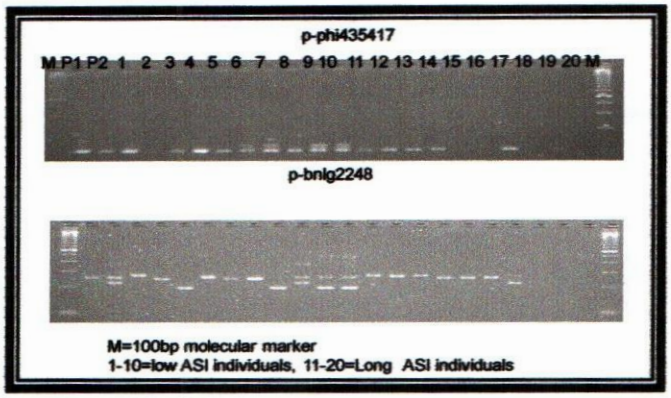
Appendix 3. Primer Confirmation Photos

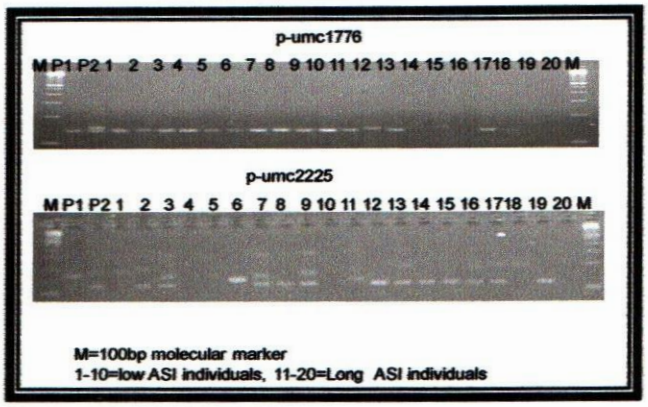


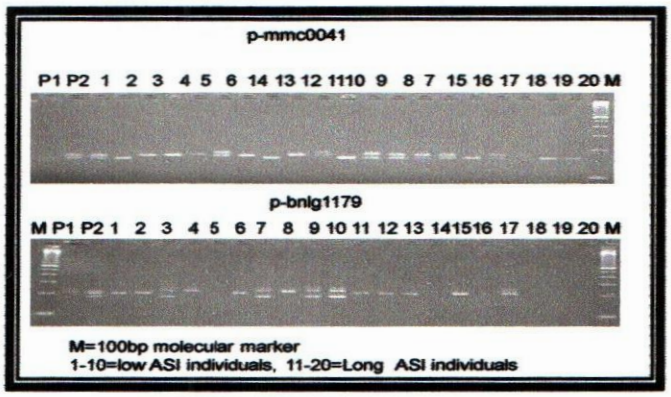


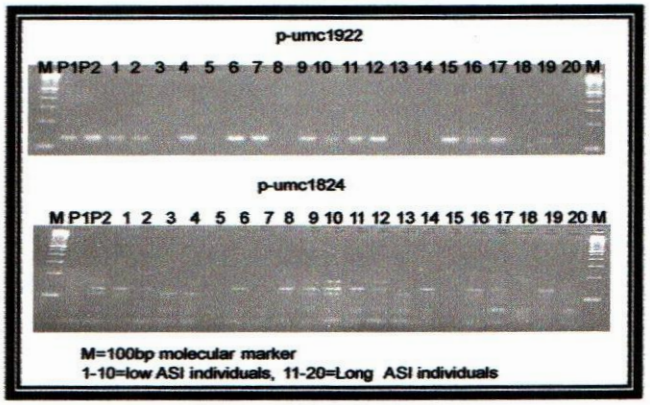


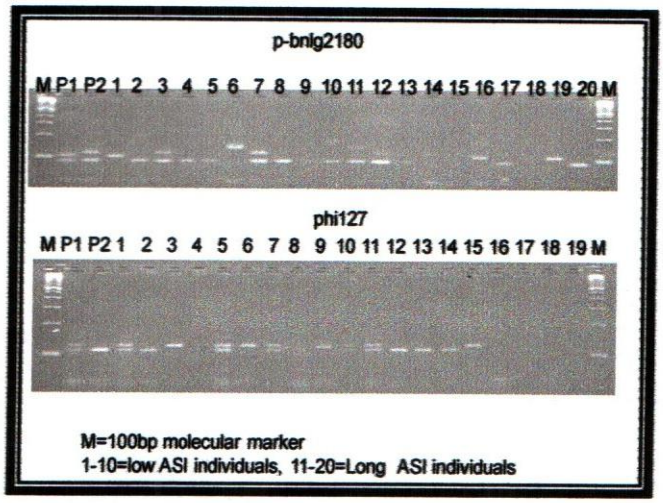


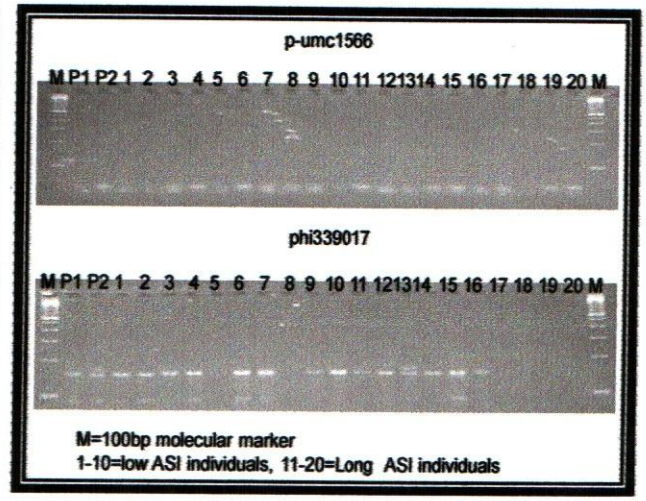


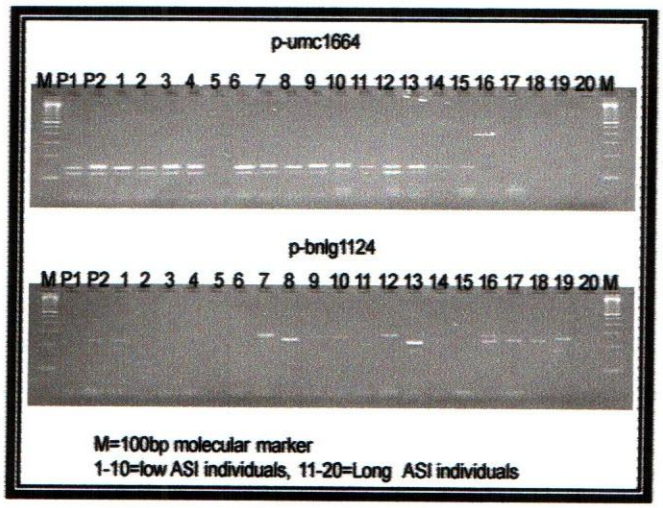


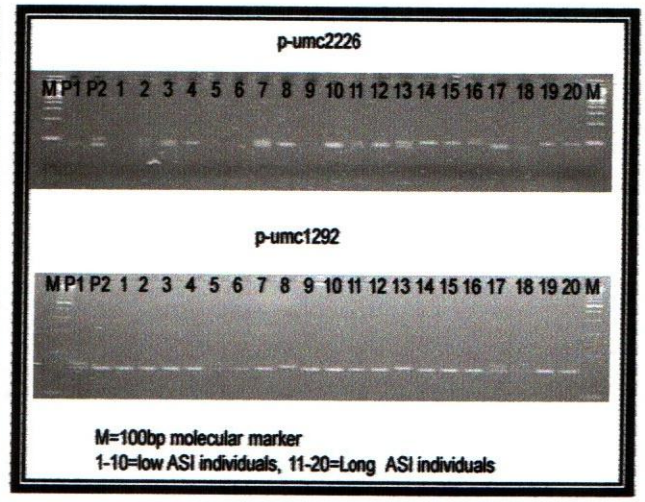


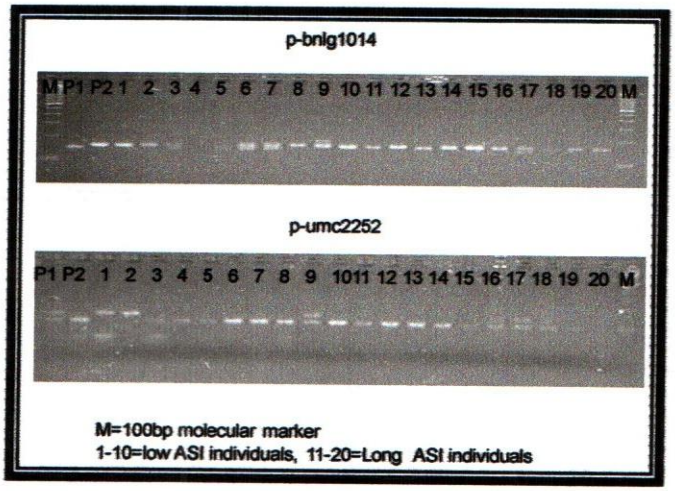


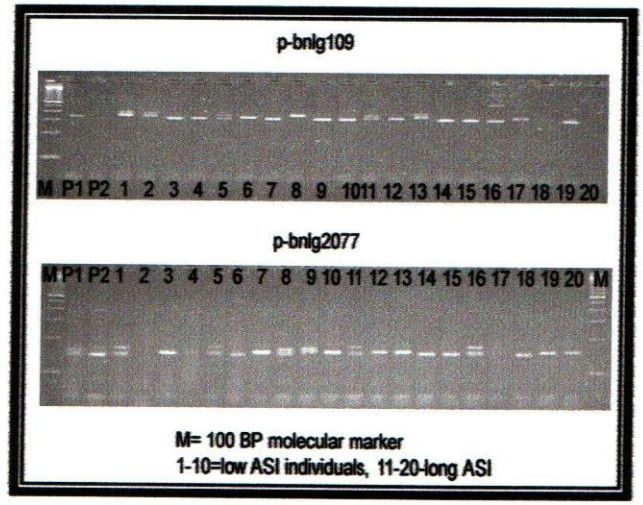

p-bnlg2180
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
phi127
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 M
M=100bp molecular marker
1-10=low ASI individuals, 11-20=Long ASI individuals


p-umc1566
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 121314 15 16 17 18 19 20 M
phi339017
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 121314 15 16 17 18 19 20 M
M=100bp molecular marker
1-10=low ASI individuals, 11-20=Long ASI individuals


p-umc1664
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
p-bnlg1124
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
M=100bp molecular marker
1-10=low ASI individuals, 11-20=Long ASI individuals


p-umc2226
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
p-umc1292
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
M=100bp molecular marker
1-10=low ASI individuals, 11-20=Long ASI individuals


p-bnlg1014
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
p-umc2252
P1 P2 1 2 3 4 5 6 7 8 9 1011 12 13 14 15 16 17 18 19 20 M
M=100bp molecular marker
1-10=low ASI individuals, 11-20=Long ASI individuals


p-bnlg109
M P1 P2 1 2 3 4 5 6 7 8 9 1011 12 13 14 15 16 17 18 19 20
p-bnlg2077
M P1 P2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 M
M= 100 BP molecular marker
1-10=low ASI individuals, 11-20-long ASI

Appendix 4. Scored molecular data

Marker	Size	No	T1	T0	T3	T4	T5	T6	T7	T8	T9	T10	S1	S0	S3	S4	S5	S6	S7	S8	S9	S10
p-bnlg109	310	1	0	0	1	1	1	0	1	0	1	1	0	0	0	1	1	1	0	0	1	-
	340	2	1	1	0	0	1	1	0	1	0	0	1	1	1	0	0	0	1	1	0	-
	370	3	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	-
p-bnlg1014	150	4	0	0	1	-	1	1	1	0	0	0	0	0	0	0	0	0	1	-	0	0
	170	5	1	1	1	-	1	1	1	0	1	1	1	1	1	1	1	1	1	-	1	1
	180	6	0	0	0	-	0	1	1	1	0	0	0	0	0	0	0	0	0	-	0	0
p-bnlg2077	140	7	1	-	0	0	1	0	0	0	0	0	0	0	0	1	1	0	-	1	0	0
	150	8	0	-	1	1	0	0	1	1	1	1	1	1	1	0	0	1	-	0	1	1
	160	9	1	-	0	0	1	0	0	1	1	0	1	0	0	0	0	1	-	0	0	0
	170	10	0	0	1	0	-	0	1	0	1	1	0	0	0	0	0	0	0	-	-	-
p-bnlg1179	190	11	1	1	0	0	-	1	0	0	0	0	1	1	1	1	1	1	1	-	-	-
	230	12	0	0	1	1	-	0	1	1	1	1	0	0	0	0	0	0	1	-	-	-
p-umc1824	130	13	1	1	1	1	0	0	0	1	1	1	0	0	1	0	-	1	1	0	1	0
	150	14	1	1	0	0	0	1	0	1	1	1	1	1	0	1	-	1	1	0	1	1
	170	15	0	0	1	1	0	0	0	1	1	1	0	1	0	0	-	0	0	0	0	0
p-mmc0041	180	16	1	1	0	0	0	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1
	190	17	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	0	0

p-umc2189	100	18	0	-	1	1	1	-	1	0	1	1	1	0	0	0	0	0	0	0	1	0
	130	19	0	-	1	0	0	-	1	0	1	1	1	0	0	0	0	0	1	0	0	0
	150	20	1	-	1	0	1	-	1	1	0	0	0	1	1	1	1	1	1	1	0	1
p-umc2226	70	21	1	1	1	0	1	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0
	80	22	0	0	0	1	0	1	0	1	1	1	0	1	0	0	0	0	0	0	1	1
	100	23	1	1	1	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0	0	0
p-umc2252	130	24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-
	150	25	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	-
p-bnlg1429	190	26	1	1	0	0	-	1	0	0	0	0	1	0	1	0	-	0	1	1	1	1
	210	27	1	1	1	1	-	0	1	0	0	1	1	0	1	1	-	1	1	0	0	0
	220	28	0	0	1	1	-	0	1	1	0	1	0	1	0	0	-	1	1	0	0	0
p-umc1542	140	29	1	1	1	0	0	0	1	1	1	1	0	0	1	0	1	1	0	0	1	-
	150	30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	-
	160	31	1	1	1	0	0	0	1	1	1	1	0	1	1	0	0	0	0	0	0	-
p-bnlg2180	80	32	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	0	0	1
	110	33	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
	120	34	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
p-umc2372	130	35	0	-	1	1	1	1	-	1	0	-	0	-	0	-	-	-	-	-	-	1
	150	36	0	-	0	0	0	0	-	0	1	-	1	-	0	-	-	-	-	-	-	1
	170	37	1	-	0	0	0	0	-	0	0	-	0	-	1	-	-	-	-	-	-	0
p-umc1331	80	38	-	-	1	-	-	1	1	0	1	1	1	0	-	0	-	0	-	0	1	1

	100	39	-	-	0	-	-	0	0	1	0	0	0	1	-	1	-	1	-	1	0	0
p-umc1611	350	40	0	0	1	1	-	1	1	-	1	1	-	1	0	1	-	1	1	1	1	-
	370	41	1	1	0	0	-	0	0	-	0	0	-	0	1	0	-	0	0	0	0	-
p-umc1397	160	42	0	0	0	1	-	1	0	-	0	-	0	1	0	-	-	-	0	-	1	1
	180	43	1	1	1	0	-	0	1	-	1	-	1	0	1	-	-	-	1	-	0	0
	190	44	0	0	1	1	-	1	1	-	1	-	0	1	1	-	-	-	0	-	0	0
p-umc1358	100	45	1	1	0	1	1	1	0	1	0	0	0	-	0	1	1	1	0	0	0	1
	120	46	0	0	1	0	0	0	1	0	1	1	1	-	1	0	0	0	1	1	1	0
p-bnlg2248	210	47	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	-	-	-
	230	48	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	-	-	-
	270	49	0	1	0	0	1	1	0	1	1	1	0	0	0	1	1	1	1	-	-	-
p-phi435417	210	50	1	-	1	1	1	1	1	1	1	1	1	1	1	1	-	-	1	-	1	-
	220	51	0	-	0	0	0	1	1	0	1	1	1	0	0	0	-	-	0	-	0	-
p-mmc0271	160	52	0	0	0	0	-	0	0	1	0	1	0	0	0	0	0	0	0	1	-	-
	180	53	1	0	1	1	-	1	1	0	1	0	1	1	1	1	1	1	1	0	-	-
p-umc2234	310	54	1	1	1	1	-	1	1	1	1	1	-	1	1	-	-	1	1	-	1	-
p-umc2225	130	55	0	1	1	-	-	0	1	1	1	-	0	1	1	1	1	1	1	0	1	-
	160	56	0	0	1	-	-	1	0	0	0	1	0	0	0	0	0	0	0	0	0	-
	180	57	1	1	0	-	-	0	1	0	1	-	1	0	1	0	0	1	1	1	0	-
p-umc1776	150	58	1	1	1	1	1	1	1	1	1	1	1	0	1	-	0	-	1	1	-	-
	160	59	0	0	0	0	0	0	0	0	0	0	0	1	0	-	1	-	0	0	-	-

p-umc1922	130	60	1	1	-	1	-	1	1	-	1	1	1	1	-	-	1	1	1	-	1	1
phi127	120	61	1	1	0	-	1	0	1	-	0	1	1	1	1	1	0	-	-	-	-	-
	130	62	1	0	1	-	1	1	1	-	1	1	1	0	0	0	1	-	-	-	-	-
phi339017	160	63	1	1	1	1	-	1	1	-	1	1	1	1	1	1	1	-	-	-	-	-
	170	64	0	0	0	0	-	0	0	-	0	1	0	1	0	0	0	-	-	-	-	-
p-umc1292	120	65	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1
	130	66	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0
p-umc1566	170	67	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
p-bnlg1124	220	68	0	-	1	1	-	-	1	0	1	1	-	1	0	-	-	1	0	0	0	-
	230	69	1	-	0	0	-	-	0	1	0	0	-	0	1	-	-	1	1	1	1	-
p-umc1664	110	70	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	-	-	-	-	-
	150	71	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	-	-	-	-	-

Appendix 5. ANOVA for ASI, DS, KN and GY QTL.

ANOVA for ASI and p-umc2189

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	6.092278	6.092278	8.341659	0.0107*
Residual	16	11.6855	0.730344		
Total	17	17.77778			

*significant at 5% level of probability

ANOVA for ASI and p-umc1542

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.967524	4.967524	6.040691	0.025014*
Residual	17	13.97984	0.822344		
Total	18	18.94737			

*significant at 5% level of probability

ANOVA for ASI and p-bnlg1179

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.464384	4.464384	5.538145	0.033749*
Residual	14	11.28562	0.806115		
Total	15	15.75			

*significant at 5% level of probability

Appendix 5 Cont'd

ANOVA for ASI and p-bnlg1014

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.645639	4.645639	8.843589	0.01006*
Residual	14	7.354361	0.525312		
Total	15	12			

*significant at 5% level of probability

ANOVA for DS and p-umc2189

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.673475	4.673475	5.706187	0.029573*
Residual	16	13.1043	0.819019		
Total	17	17.77778			

*significant at 5% level of probability

ANOVA for DS and p-bnlg1014

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.831118	3.831118	6.56585	0.022568*
Residual	14	8.168882	0.583492		
Total	15	12			

*significant at 5% level of probability

ANOVA for DS with p-umc2189 and p-bnlg1014

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	2	721.8932	360.9466	5.566592	0.013818*
Residual	17	1102.307	64.84157		
Total	19	1824.2			

*significant at 5% level of probability

ANOVA for KN and p-bnlg1179

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.987156	4.987156	6.973061	0.019379*
Residual	14	10.01284	0.715203		
Total	15	15			

*significant at 5% level of probability

ANOVA for KN and p-umc1542

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.46141	4.46141	5.235689	0.035209*
Residual	17	14.48596	0.852115		
Total	18	18.94737			

*significant at 5% level of probability

Appendix 5 Cont'd

ANOVA for KN with p-bnlg1179 and p-umc1542

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	46734.71	23367.35	4.790757	0.022381*
Residual	17	82919.04	4877.591		
Total	19	129653.8			

*significant at 5% level of probability

ANOVA for GY and p-bnlg1179

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.922573	3.922573	4.957471	0.042908*
Residual	14	11.07743	0.791245		
Total	15	15			

*significant at 5% level of probability

ANOVA for GY and p-umc1542

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	6.313878	6.313878	8.496141	0.009655**
Residual	17	12.63349	0.743147		
Total	18	18.94737			

**significant at 1% level of probability

ANOVA for GY with p-bnlg1179 and p-umc1542

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1891.687	945.8436	5.799271	0.01202*
Residual	17	2772.649	163.097		
Total	19	4664.336			

*significant at 5% level of probability