GGE-Biplot and Genetic Diversity Analysis of Maize Hybrids and Inbred Lines from the breeding programme at UKZN

by

Mandisa Precious Makongwana
B Sc. Agric (Hons) (University of the Free State)

A dissertation submitted in partial fulfilment of the requirements for the degree of

Master of Science in Plant Breeding

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg, South Africa

November 2016

DISSERTATION ABSTRACT

Maize (Zea *mays* L.) is the most important cereal crop in southern Africa. It is classified as a major staple food for human consumption. It is also used for animal feed in the livestock industry. Therefore, maize plays a crucial role in ensuring food security. However, production of maize is outstripped by consumption. Therefore, there is a yield gap that needs to be closed by increasing maize yields. Unfortunately, the adequate production of the crop is affected by lack of highly stable and highly productive hybrids. Hybrids that combine these two attributes are highly desired in the small scale sector where resources are usually limiting. Hybrids can be exploited to increase productivity of maize; however, a study of diversity between the parent lines is required because hybrid heterosis is obtained when the lines are divergent and complimentary.

Therefore, the current study investigated genetic diversity and genetic gains that were realized by the breeding program at the UKZN. Thirty-one experimental hybrids from the program were tested alongside eleven commercial hybrids across 6 locations in South Africa. Hybrids 11C3201, 13C7082, 11C3417, 14XH149 and 14XH146 were among the best four hybrids across 6 locations. They combined high levels of stability and productivity, qualifying them as potential candidates for advancement. The study indicated that the program was successful at breeding new hybrids with the potential to compete with current commercial hybrids.

With respect to diversity, 51 inbred lines were genotyped with 396 SNP markers at the LGC genomics, UK. Therefore, PIC, genetic diversity, availability, inbreeding coefficient, heterozygosity and genetic distance were determined. SNP markers indicated there was large diversity between the lines as reflected by two major clusters at the truncation level of 0.14 in the coefficient scale. Under the second major cluster, there were eight sub-clusters (sub-cluster B-J) identified which indicated wide range of the genetic diversity within inbred lines. Genetic distance between lines ranged from 0.05 to 0.35.

This indicates the program was successful at generating new inbred lines that can be used to breed new hybrids.

DECLARATION

I, Mandisa Precious Makongwana declare that:

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

……

Mandisa P Makongwana (Candidate)

As the candidate's supervisors we agree to the submission of this dissertation

…….....

Professor John Derera (Supervisor)

ACKNOWLEDGEMENTS

Foremost, I am highly grateful to God for His blessing that continue to flow into my life, and because of You, I made it through against all odds.

I would like to express my gratitude to my supervisor, Professor John Derera for the invaluable support, guidance, encouragement and constructive criticism from the inception of the study, execution and up to the final write up of this dissertation.

A vote of thanks to Gordon Mabuyaye and Nyasha Chiurase for assistance with data analysis, support and assistance whenever needed.

Thank you to my family and friends for their continued motivation and support.

DEDICATION

I dedicate this thesis to my beloved brave mother Iris Nomangqika Makongwana for her hard work in raising me, guiding me and instilled the culture and importance of education from a very young age. Without your prayers, support, strength and wisdom, none of this was possible to me, Ncenjane.

To my only sister Bulelwa Mimi Makongwana for her tremendous support and for being my rock throughout my journey, Zizi Jama kaSijadu.

TABLE OF CONTENTS

LIST OF ACRONYMS

AFLP – Amplified fragment length polymorphism

AMMI - Additive main effects and multiplicative interaction model

ANOVA – Analysis of variance

ASI – Anthesis silking interval

CM – Centimeter

DC – Double cross

GCA – General combining ability

GD – Genetic distance

GEI – Genotype by environment interaction

GVC - Genotypic variance coefficient

h^2 – Heritability

MAS – Marker assisted selection

Mmt – Million metric

MSC – Modified single cross

MSV – Maize streak virus

OPV – Open pollinated variety

PCA – Principal component analysis

PIC - Polymorphic information content

QPM – Quality protein maize

RAPD - Random Amplified Polymorphic DNA

RFLP – Restriction fragment length polymorphism

SC – Single cross

SCA – Specific combining ability

SD – Selection differential

SG – Selection gain

SNP –Single nucleotide polymorphism

SSR – Simple sequence repeat

UKZN- University of KwaZulu- Natal

TABLE OF FIGURES

LIST OF TABLES

CHAPTER ONE

DISSERTATION INTRODUCTION

1.1 Significance of maize in South Africa

In South Africa, maize is recognized as a major staple food and significant source of carbohydrates, protein, iron, vitamin B, and minerals. Significantly, maize crop has a major economic value as livestock feed. Because of its ease of cultivation and nutritional qualities, it has became the largest locally-produced field crop (National Agriculture Marketing Council, 2009). Therefore, it ranks as one of the most crucial economic crops in relation to net production in metric tonnes value in South Africa (FAOSTAT, 2016). The exponentially growing South African population is increasingly depending on maize for food, feed and industrial usage. The demand for animal feed is expected to be 6.4 million tonnes by 2050 (Syngenta, 2013). The significance of the maize industry is reflected by the international exports to countries like Japan, Mexico, Italy and Korea, as a significant earner of foreign currency for the South African economy. Therefore, any production constraint of maize implies a negative impact on the South African economy, health and political stability at large.

1.2 Genetic Diversity

“Genetic diversity is the sum of the genetic characteristics within any species or genus” (Rao and Hodgkin, 2002). Genetic diversity in maize plays a crucial role in maize breeding (William and Michael, 2002). Information regarding the genetic diversity of breeding lines and populations is crucial for hybrid development purposes (Makumbi *et al.,* 2011). Additionally, enormous genetic diversity in maize presents a fundamental opportunity for germplasm enhancement (Prasanna, 2012), therefore impacts on the breeding strategy. As described by Frankham *et al.* (2002), genetic diversity is “the range of alleles and genotypes existing in a population, reflecting morphological, physiological and behavioral differences amongst individuals and populations”. The understanding of the amount, distribution and patterns of genetic variation within and among the various inbred lines provides a prediction guideline on the level of heterosis, variation and the degree of inheritance (Qi-Lun *et al.,* 2008). However, to improve the genetic diversity in local germplasm,

knowledge of the extent of already existing genetic diversity in the germplasm is vital (Ahmad *et al.,* 2011).

1.3 Heterosis

According to Birchler *et al.* (2010) heterosis as the phenomenon that progeny of a cross show superior performance to their parents. This phenomenon is also referred to as hybrid vigour (Baranwal *et al.,* 2012). Current ideas suggest that the combination of the parent pairs has a major impact on the level of heterosis manifested by the hybrid for different traits and total performance of the hybrid (Chen, 2013). Therefore the level of differences between parents is important because heterosis is the result of high degree of heterozygosis in the genome. Heterosis has been extensively studied in maize, because of its large expression in grain yield and exploitation in hybrid breeding of maize (Reif *et al.,* 2005). In maize, heterosis forms the principal basis for commercial hybrids (Garcia *et al.,* 2008) worldwide, including South Africa. The use of heterosis has critically contributed to the commercial success of plant breeding in many species and leads to the prevalent use of hybrids in numerous crops and horticultural species.

1.4 Genetic gains

As described by Condón *et al.* (2009) genetic gains through plant breeding are a function of the genetic variation within the breeding population and the efficiency of selection imposed by the breeder. This is usually used to refer to the increase after one generation has passed. Genetic advance illustrates the amount of gain attained by character under a particular selection pressure (Bello *et al.,* 2012). Future maize genetic gains are dependent on the employment of useful genetic diversity discovered in the public sector. For genetic gains to be significant and make impact, the inclusion of exceptional and useful genetic diversity to breeding programmes is required to actively improve germplasm and develop cultivars (Carena *et al.,* 2009). Breeding programmes exploit variability of traits among genotypes. Therefore heritability of traits is crucial and indicates whether traits can be improved via selection strategy (Bello *et al.,* 2012). While selection strategies may determine the rate of gain, it is the genetic variation within the breeding population and the number of recombinants generated that determine the potential gain that can be realized with optimal selection (Bernardo, 2002).

1.5. Heritability

Heritability is a predictive tool used by plant breeders to determine the amount of progress or genetic gain that can be made through certain breeding strategies. Heritability assists in explaining the degree to which genes control the expression of a trait (Cassell, 2009). This is expressed as a percentage from 0 to 100%. It generates information regarding the extent to which selection would be effective in improving the population for the trait under study. Heritability influences the selection methods and decisions in prediction of gains and determination of relative importance of genetic effects (Waqar-Ul-Haq *et al.,* 2008). The higher the heritability, the higher is the opportunity of making successful selections of the trait of interest (Wiggins, 2012). Highly heritable traits can easily be fixed with simple selection suggesting fast genetic progress (Bello *et al.,* 2012; Langade *et al.,* 2013).

Problem statement

Productivity of maize is very low which leads to food insecurity. In South Africa, maize production fails to meet the maize demand for human consumption. Therefore, breeders are faced with a challenge to address the maize yield gap through increasing maize yields. Hybrids can be exploited to increase productivity of maize. However, a study of diversity between the parent lines is required because hybrid heterosis is obtained when the lines are complimentary. The level of the genetic diversity between the advanced lines and the elite lines is not known which may compromise the breeding strategy. The breeding gains that have been realized by improving these lines over the past years have not been established.

Research objectives

The main objective of this study is to determine diversity and productivity among advanced and elite maize inbred lines with both tropical and temperate genetic background at UKZN.

The following specific objectives were pursued:

a) To determine the levels of genetic diversity in advanced inbred lines and elite maize lines from the programme at UKZN using SNP molecular markers.
b) To determine the level of genetic gain realized in breeding maize hybrids for productivity using advanced and elite lines.
c) To determine stability of experimental maize hybrids in South African production environments.

Research hypotheses

The following hypotheses were tested:

1. There is large genetic diversity among advanced and elite maize inbred lines in the UKZN breeding programmes.
2. High genetic gains can be realized by crossing the advanced and elite inbred lines in the programme.
3. Highly stable hybrids can be developed by crossing the advanced and elite inbred lines in the programme.

Research questions

a) Is there molecular diversity among the advanced inbred lines and elite maize lines in the breeding programme at UKZN?
b) Is there genetic gain realized in breeding maize hybrids for productivity using advanced and elite lines?
c) Is there high grain yield performance stability in the experimental maize hybrids in South African production environments?

Dissertation structure

- The dissertation is composed of five chapters:

Chapter one: Introduction to dissertation

- Presents the significance of maize globally, in Africa and in South Africa.
- Presents a brief definition of heterosis and genetic gains.

Chapter two: Literature review

- Presents reviewed literature on topics relevant to the study.

Chapter three: Genetic diversity analysis of the inbred lines

- Presents information on molecular diversity and genetic distance among the maize inbred lines used in the current study.

Chapter four: Assessment of genetic gain and genotype x environment interaction in experimental maize hybrids

- Presents information on the variability of yield performance in hybrids used in this study as influenced by the genotype and environment interaction (GXE).

Chapter five: General overview of the study

- Presents general discussion, conclusion and the way forward (Recommendation).

References

Ahmad, S.Q., S. Khan, M. Ghaffar and F. Ahmad. 2011. Genetic diversity analysis for yield and other parameters in maize (*Zea mays* L.) genotypes. *Asian Journal of Agricultural Science*, 3(5): 385-388.

Baranwal, V.K., V. Mikkilineni, U.B. Zehr, A.K. Tyagi and S. Kapoor. 2012. Heterosis: emerging ideas about hybrid vigour. *Journal of Experimental Botany*, 63(18): 6309-6314.

Bello, O., S. Ige, M. Azeez, M. Afolabi, S. Abdulmaliq and J. Mahamood. 2012. Heritability and genetic advance for grain yield and its component characters in maize *(Zea mays L.). International Journal of Plant Research*, 2(5): 138-145.

Bernardo, R. 2002. Breeding for quantitative traits in plants (No. 576.5 B523). *Stemma Press Woodbury*.

Birchler, J.A., H. Yao, S. Chudalayandi, D. Vaiman and R.A. Veitia. 2010. Heterosis. *The Plant Cell Online*, 22(7): 2105- 2112.

Cassell, B.G. 2009. Using heritability for genetic improvement. Va. Coop. Ext. Publ. 404-084.

Carena, M., G. Bergman, N. Riveland, E. Eriksmoen and M. Halvorson. 2009. Breeding maize for higher yield and quality under drought stress. *Maydica*, 54(2): 287- 296.

Chen, Z.J. 2013. Genomic and epigenetic insights into the molecular bases of heterosis. *Nature Reviews Genetics*, 14(7): 471- 482.

Condón, F., D.C. Rasmusson, E. Schiefelbein, G. Velasquez and K.P. Smith. 2009. Effect of advanced cycle breeding on genetic gain and phenotypic diversity in barley breeding germplasm. *Crop Science*, 49(5): 1751-1761.

FAOSTAT. (2016).Production. http://faostat3.fao.org/browse/Q/*/E (Accessed 20 August 2016).

Frankham, R., D.A. Briscoe and J.D. Ballou. 2002. Introduction to conservation genetics. *Cambridge University Press*.

Garcia, A.A.F., S. Wang, A.E. Melchinger and Z.-B. Zeng. 2008. Quantitative trait loci mapping and the genetic basis of heterosis in maize and rice. *Genetics*, 180(3): 1707-1724.

Langade, D.M., J. Shahi, K. Srivastava, A. Singh, V. Agarwal and A. Sharma. 2013. Appraisal of Genetic Variability and Seasonal Interaction for Yield and Quality Traits in Maize (*Zea mays L*.). Plant Gene and Trait, 4(18): 95-103.

Makumbi, D., J.F. Betrán, M. Bänziger and J.M. Ribaut. 2011. Combining ability, heterosis and genetic diversity in tropical maize *(Zea mays L.)* under stress and non-stress conditions. Euphytica, 180(2): 143-162.

National Agricultural Marketing Council (NAMC). Report on the investigation of the effects of deregulation on the maize industry. Section 7: Committee evaluating the deregulation process. Pretoria, South Africa 2003. http://www.grainmilling.org.za/2008%20MAIZE%20SECTION%207%20REPORT1.pdf. (Accessed August 2016).

Prasanna, B. 2012. Diversity in global maize germplasm: characterization and utilization. Journal of Biosciences, 37(5): 843-855.

Qi-Lun, Y., F. Ping, K. Ke-Cheng and P. Guang-Tang. 2008. Genetic diversity based on SSR markers in maize *(Zea mays L.)* landraces from Wuling mountain region in China. Journal of Genetics, 87(3): 287- 291.

Rao, V.R. and T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell, Tissue and Organ Culture, 68(1): 1-19.

Reif, J., A. Hallauer and A. Melchinger. 2005. Heterosis and heterotic patterns in maize. Maydica, 50(3/4): 215 - 223.

Syngenta. (2013). Maize and related products http://www.sygenta.com.country/za/en/crop. (Accessed 01 September 2014).

Waqar-ul-haq, M.M., F; Rashid, M; Munir, M and Akram, Z. 2008. Evaluation and estimation of heritability and genetic advancement for yield related attributes in wheat lines. Pakistan Journal of Botany, 40(4): 1699-1702.

Wiggins, B.T. 2012. Heritability and genetic gain of seed protein, oil, and yield among RIL of soybean. Master's Thesis, *University of Tennessee*. http://trace.tennessee.edu/utk_gradthes/1248.

William, S.K. and R. Michael. 2002. Essentials of genetics. Beijing: Higher Education Press, Pearson Education, 1: 16.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter reviews the literature on topics relevant to this study. Initially, the significance of maize globally and in Africa, genotype and environment interaction and genetic gains in maize is discussed. Heritability, genotypic and phenotypic coefficient of variance, genetic diversity in maize, estimation of genetic diversity and molecular markers are discussed in relation to factors influencing maize yield, increasing productivity and genetic diversity to a large extent. Conclusions drawn from this review are provided at the end of the chapter.

2.2 Global significance of maize

Maize (Zea *mays.L*) is one of the oldest human-domesticated and widely cultivated grain crop throughout the world with its origins centred in central Mexico. Maize is one of the major cereal crops of the world. It is projected that by 2020 maize will outperform both wheat and rice to become the number one cereal in the world. Maize is highly adapted from sea level to the highlands (above 3000 m altitude) and across latitudes from temperate to equatorial regions (Dowswell *et al.,* 1996).

Five years statistics for global average maize production indicate that United States of America had the highest production of 323,742,070 tonnes of maize (FAOSTAT, 2016) from 2010 to 2014. This was followed by China recording 201,991,080 tonnes for the same production period. South Africa had the least average maize production per ton over five years (Table 2.1).

Table 2.1: Global average maize production across countries (FAOSTAT, 2016)

Production Country	Average Production (Tonnes)	Production Year
South Africa	12494600	2010- 2014
Mexico	21788752	2010- 2014
Argentina	27076393	2010- 2014
Brazil	68449640	2010- 2014
United States of America	323742070	2010- 2014
China	201991080	2010- 2014

Maize, wheat and rice are the major crops accounting for more than 75% of grain production and 30% of energy (Shiferaw *et al.,* 2011). In addition it is a major industrial crop worldwide (Pingali, 2001). In 2004, global production was over 721 million metric tons (mmt) exceeding both wheat (627 mmt) and rice (605 mmt) (Meridian institute, 2014). Consequently, the global trends suggest the demand for maize continues to upsurge.

Projected use of maize by year 2020, indicate that animal feed will play a greater role in the demand of maize across most regions (Table 2.2), especially in East Asia where about 82% of maize demand will be for the animal feed sector. The scenario is however different for Africa south of the Sahara and South Asia. The total demand for Sub-Saharan Africa is projected to be 52 million metric tonnes of maize. This is because 76% of maize demand is for human consumption, while 10% is projected for animal feed. In South Asia, about 70% of the maize demand and use is for human consumption while only 13% accounts for animal feed (Table 2.2).

Table 2.2: Maize projected demand and use in 2020

Region	Area[1]	Demand[2]	% Food	% Feed	% Other	Net Trade[3]
Global	158	852	15%	69%	16%	
Industrial	50	344	5%	76%	19%	-67%
Developing	108	508	22%	64%	14%	-67%
East Asia	30	252	4%	82%	14%	-43%
Latin America	32	118	25%	60%	15%	+5%
Sub-Saharan Africa	26	52	76%	10%	14%	-6%
South East Asia	9	39	32%	58%	10%	-8%
WANA	2	28	28%	63%	9%	-14%
South Asia	9	19	70%	13%	17%	-<1

Source: IFPRI, 2003. [1]Million of hactares;[2] Millions of metric tonnes (MT); [3]Millions of MT,exports (+), imports (-)

In South Africa, five year's data indicated that 2011 was the least productive year with about 10.4 Million tonnes (Figure 1.1) of maize produces countrywide. However, from 2012 up to 2014, maize production trends indicated to be escalating to about 15 Million tonnes.

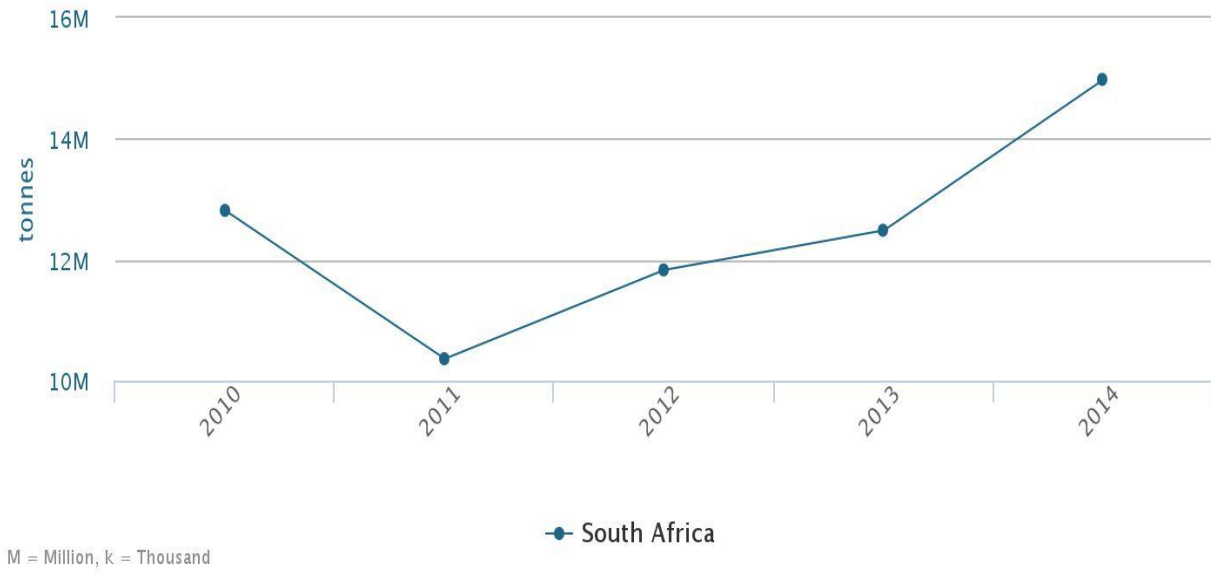

Figure 1.1: South African maize production (FAOSTAT, 2016)

2.3 Significance of maize in Africa

Among the twenty-two countries in the world where maize encompasses the majority of the diet, sixteen are in Africa (Meridian institute, 2014). In 2005, the leading maize producers were South Africa, Tanzania, Uganda, Zambia and Swaziland, whereas Zimbabwe, Angola, Ghana, Kenya and Mozambique had maize grain deficits which were filled by imports (Meridian Institute, 2014). Therefore these countries should boost grain production to achieve food security (Pingali, 2001). IFPRI (2003) reported that by 2020, the highest proportion (76%) of maize used for food will be in the countries of Sub Saharan Africa (Table 2.2). Because of increasing importance, maize has become a major staple food for over 300 million Africans (M'mboyi *et al*., 2010) and cash crop for small scale farmers. Abuali *et al*. (2014) indicated that maize was perceived as less important as human staple food when compared to sorghum, wheat and millet in Sudan, However, Sudan has more than doubled her grain production area from 17000 to 37000 ha in 39 years between 1971 and 2010 (Abuali *et al*., 2014).

Maize is reported to be a cheap staple food compared to other small grain cereals such as rice and wheat. It accounts for most of the carbohydrates, protein and mineral requirements (M'mboyi *et al*., 2010) especially in Africa, where it is the major crop. Therefore maize is identified as a strategic commodity for food security

in Africa (FARA, 2009). However, the yields of maize are extremely low by world standards with average estimation of 1.7 tons per hectare in 2006 compared to the global average of 5 tons per hectare (FARA, 2009). In Africa, virtually all plant parts of maize have economic value (M'mboyi *et al*., 2010). Maize grain, leaves, stalk, tassel and cob can all be used for production of food and non-food products. In developed African countries, maize is consumed as secondary products "in the form of meat, eggs and dairy product" (Du Plessis, 2003). Grain is used for livestock feed and production of industrial products for starch, oil and protein (M'mboyi *et al*., 2010). However, maize leaves, stalk and tassel are used as fodder or silage. In a processed form, maize is used as a source of fuel and is available as Ethanol (M'mboyi *et al*., 2010). Roots are burnt as fuel and used for mulching and manure.

In South and East Africa, maize accounts for 30-50% of low income household expenditures (www.iita.org) accessed 13/09/2014. For this reason, continued research in increasing maize grain yields to ensure future food security is imperative. Plant breeders are therefore confronted by the urgency of breeding for hybrids with increased vigour to close the gap between the growing population, consumption and maize supply.

2.4 Genotype and environment interaction (GEI)

The genotype and environment interaction describes the different response of a genotype towards the environmental effects observed across multiple locations (da Silva *et al.,* 2012). Genotypes that perform well in one environment may not do as well in another. Therefore, this suggests that the genotype and the environment effects are not independent of each other. The GEI proves to be an important consideration to plant breeders as it complicates the breeding and selection for quantitative traits, including yield (Nzuve *et al.,* 2013). This is because it results in change of rank of genotypes in different environments. Assessment of plant performance across various locations reduces the chance of misleading selection results and recommendations (Mendes *et al.,* 2012). According to Nzuve *et al.* (2013) the genotype and environment components contribute to noise, reducing heritability of traits and compromise breeding gains. In other words GEI compromises measurement of traits and reduce heritability (Grada and Ciulca, 2013). GEI studies provides an opportunity to identify stable and adapted genotypes

in specific targeted areas, as well as testing potential candidates for promising hybrid combinations (Nzuve *et al.,* 2013).

Several methods for multi environment trials data have been established to identify and show patterns of GE interaction and have been reported by various authors. These methods include the additive main effects and multiplicative interaction model (AMMI), principal component analysis (PCA) and linear regression analysis, analysis of variance (ANOVA) and GGE biplot analysis (Miranda *et al.,* 2009; Akcura *et al.,* 2011; Mitrovic *et al.,* 2012) . These statistical methods are briefly discussed.

2.4.1 Additive Main effect and Multiplicative Interaction (AMMI)

This analysis combines ANOVA and PCA where sources of genotype by environment interaction are partitioned by PCA. The interpretation of results obtained from AMMI analysis is performed with a biplot that relates genotypic means to the first or some of the principal interaction components (Grada *et al.,* 2013). Therefore, AMMI analysis has a capability of analyzing complex data. It uses a principal component (auto-vector) to interpret genotype performance by incorporating the use of ANOVA and PCA. It also allows for the graphic display of phenotypic stability, genotypic behavior of the cultivars and environments which optimize performance (Miranda *et al.,* 2009). Akcura *et al.* (2011) described AMMI as “useful in summarizing and estimating patterns of response which exist in the original data”. Abay *et al.* (2009) reported that AMMI has made a significant contribution to analysis of multi-environment data, especially yield.

According to Grada *et al.* (2013) the AMMI analysis is a crucial tool for multi-environment data analysis. This includes improvements in estimating yield of genotypes from different environments. It is an important tool for more diagnosis and interpretation of GEI and shows important trends between genotypes and environments.

Although the ANOVA method is used to identify the existence of GEI in multi – environmental experiments, it has limitations. According to Adu *et al.* (2013), the ANOVA has a limitation of “presumption of homogeneity of variance among environments required to determine genotype differences”. The ANOVA does not allow exploring the response of the genotypes in the non- additive term (Adu *et al.,*

2013). This limits its application in multi-environment data especially in tropical countries where the environment is variable.

Sousa *et al*. (2015) employed AMMI analysis to evaluate grain yield stability and adaptability of 27 soybean lines across 5 locations over 2 seasons. The plotted graph revealed genotype 23 as the most yielding and stable as closest to the zero score, followed by other 6 genotypes. The analysis was successful in identifying 3 most unstable genotypes that contributed more in the GEI. Two of the unstable genotypes generated higher yields compared to the stable genotype 23. It was concluded that in the study the highest yielding genotypes were the most unstable, therefore suggesting that if the environment was favorable, these genotypes behavior will be favorable. However, if the environment is unfavorable then the yield from these genotypes could be compromised.

To determine stability, nature and magnitude of GEI effects on rice grain yield, Akter *et al*. (2014) cited using AMMI biplot analysis. The analysis was done using 12 rice genotypes, planted across 5 environments. The genotypes showed inconsistent performance across environments. Genotype 3 had an overall highest yield based on the genotype averaged yield values over environments while genotype 12 had the lowest yield. Genotype 1, 2, 3 and 4 had high yield with high main effects, showing positive IPCA1 scores. Environment 1 and 5 had positive IPCA1 score near zero, hence small interaction effects. This suggests that all genotypes performed well in these two locations. The analysis identified genotype 5 having an IPCA1 score closer to zero, indicating stability and less influenced by the environment. Genotype 10 had moderate stability across environments.

In Brazil, Silveira *et al*. (2013) reported using AMMI analysis to evaluate the adaptability and phenotypic stability of 15 cloned sugarcane genotypes. The genotypes were tested in 9 environments. The analysis indicated that 3 genotypes had high stability as they were positioned closer to the biplot origin. There were 4 additional genotypes including a check that exhibited high to intermediate stability. Furthermore, 4 genotypes were identified as unstable with specific adaptation. This was because they were very distant from the biplot origin, therefore contributing more to the GEI. Environment 8 was reported to be the largest contributor to the phenotypic stability of the genotypes. This was because in that environment there were differences found among genotypes when using individual ANOVA. The biplot

visually revealed positive association between genotypes and environments to create agronomic zones. Noticeably, 3 genotypes including a check were specifically adapted to environment 5. Based on the genotypes with the highest raw means in each environment, the study revealed that six mega environments were formed. Although based on the predicted means, only three mega environments were formed.

Currently, AMMI analysis is used across a wide range of crops to evaluate performance and interaction with the environment. This is because it is an important tool in breeding, which directs selection of most productive, adapted and stable hybrids or lines for specific locations and growing seasons. AMMI analysis' strength lies in the simple graphical representation of genotypes and environments in a multivariate dispersion diagram (Yokomizo *et al*., 2013).

2.4.2 GGE biplot analysis

The biplot has increasingly become a popular data visualization tool among plant breeders for the purpose of cultivar evaluation and mega – environment analysis. The GGE biplot analysis considers genotype main effects (G) and genotype by environment (GE) interaction as two sources of variation of the site regression model. "Sources of variation are found to be significant to genotype evaluation and must be factored concurrently for proper genotype and test environment evaluation" (Yan and Tinker, 2006) . It combines ANOVA and PCA by partitioning together of sums of squares of genotypes and sums of squares of GEI using PCA method (Mitrović *et al.*, 2012).

GGE biplot is created by plotting the first principal component scores (PC1) of the genotypes and environments against second principal component scores (PC2) obtained by decomposition of singular values of multi-location trials yield data (Mitrović *et al.*, 2012; Yan *et al.*, 2007).

The GGE biplot technique is used for the presentation and estimation of genotypes in different environments. The analysis enables identification of genotypes with highest yields in different environments, preferred genotypes and target environments (Mitrović *et al*., 2012). Therefore, the GGE biplot determines the which -won- where pattern analysis and does genotype and test environment evaluation (Asfaw *et al.*, 2009).

Kuchanur *et al*. (2015) successfully employed GGE biplot technique to test the efficacy of 6 testing environments and investigate stability performance of 27 single cross maize hybrids under stress and non-stress conditions. The analysis illustrated 6 environments fell into four mega environments with different high yielding hybrids. Nine hybrids were identified as unstable in performance and highly interactive with environments. Six hybrids were reported as vertex hybrids since they were the most responsive and farthest from the biplot origin.

In India, Rahnejat and Farshadfar (2015), reported on evaluating phenotypic stability in 15 Canola genotypes across 4 environments, using GGE biplot analysis. The analysis revealed that locations were clustered into two mega environments, with three locations belonging to one mega environment. It was found that two genotypes performed the poorest across all environments. In this study visualization of the "which-won-where pattern" of the multi environment trials data was achieved through the GGE biplot technique.

In Iran, Mortazavian *et al*. (2014), reported using GGE biplot analysis to test yield performance on 20 barley genotypes across 14 different environments. Results illustrated that locations fell into three mega environments. The GGE biplot analysis was successful in showing relative adaptation of genotypes with the highest grain yield across testing locations.

2.5 Genetic gains

Genetic gain is the difference in the mean value of the selection criterion between the original generation and the next generation of the selected population when compared in the same environment (www.passel.unl.edu) accessed 24/04/2015. Accordingly, genetic gain refers to the improvement of traits' genotypic value for the new population under one selection cycle at a given selection intensity. Genetic gain is calculated using a breeder's equation or formula:

$SG = h^2 \times SD$, whereby SG is a selection gain; h^2 is the heritability of the trait; and SD is the selection differential which is the average of the selection criterion of the selected individuals minus the average of the selection criterion of the original population (Bhering *et al.,* 2012). Selection criterion refers to the traits on which selection is based. Genetic gain results when there is a better gene combination that control traits of interest within the selected genotypes compared to the unselected genotypes (www.passel.unl.edu) accessed 24/04/2015.

Through genetic gain information, grain yield and other traits of high potential value can be identified. Additionally, genetic gain knowledge can influence alteration in breeding methodologies and strategy for further increased progress in future breeding efforts. Therefore, heritability and genetic advance of a trait reveals the extent to which a trait can be improved through selection (Sesay *et al*. 2016). However, genetic advance together with heritability estimates are more informative in the prediction of gain under selection compared to the use of heritability estimates alone (Anshuman *et al.,* 2013). Therefore, genetic advance is a prediction tool of genetic gains in the next generation (Najeeb *et al.,* 2009). Moreover, a character with high heritability does not necessarily have high genetic gains; however, to reach a more reliable conclusion, it is recommended that high heritability should be complemented by high genetic gains (Kumar *et al*., 2014).

To determine the levels of genetic advance in maize, Anshuman *et al.* (2013), tested 20 genotypes and observed 14 traits including grain yield per plant. Moderate values were recorded for grain yield genetic advance percentage (58.52 %) while high broad sense heritability values were observed at 99.8 %. This suggests that selection may lead to improvement of this trait.

Murtza *et al.* (2014) tested 14 maize genotypes to estimate heritability and genetic advance. He found high heritability estimates (93.08 %) coupled with high genetic advance. Therefore he concluded that for future yield enhancement, selection of genotypes may be based on traits like grain yield.

However, Munawar *et al.* (2013) found that maize grain yield is influenced and directly linked to other traits such as cob position, number of rows per cob and number of grain per cob. Therefore, these associations suggest that to improve grain yield further, selection of linked traits would result in an effective breeding programme.

Yusuf (2010) reported genetic advance of 95.1% in grain yield of single cross quality protein maize and a broader genetic base for grain yield was found in the hybrid population. Furthermore, grain yield was significantly and positively associated with thousand seed weight which recorded high genetic gain of 445.8%. It was concluded

that improvement of thousand seed weight will consequently lead to improvement of grain yield. Previously, Singhal *et al.* (2006) reported on maximum grain yield genetic advance of 15.69% coupled with 73.8% heritability in high quality protein maize. However, plant height, ear height and ear diameter were also directly associated with grain yield, therefore were crucial secondary traits for yield improvement (Singhal *et al*., 2006).

Reddy *et al*. (2013) found that grain yield as a complex trait is also dependant on other several contributing characters known as yield components. Grain yield was observed to have high genetic advance of 5%. However, ear height (43.93%), plant height (29.63%), number of kernels per row (31.33%) and ear length (26.41%) also recorded high genetic advance values coupled with high heritability. These traits are under the influence of the additive gene action. High direct effects were observed for these traits through path coefficient analysis, which indicates strong association of these traits with grain yield. Therefore, direct selection of these traits will be effective to improve grain yield.

Beyene *et al*. (2005) found maize yield to have low genetic advance estimate of 13.5% from the selected top 5% of 180 accessions in the Ethiopian highlands. This was also coupled with lowest heritability estimate of 17.0%. Therefore they concluded that selection of the yield trait would be more difficult. However, the results indicated that selection of the number of kernels per row can improve the trait up to 37.8% as genetic advance. Therefore, number of kernels per row was suggested to be used as selection criteria in improving grain yield, adding credence to the philosophy that yield can be improved through indirect selection of secondary traits.

Bekele and Rao (2014) reported high genetic advance values for plant height (14.957%), number of seed rows per cob (0.857%), 100 seed weight (15.026%), ear height (23.210%), grain yield per plot (36.199%) and protein percentage (15.015%).

2.6 Heritability

There are two types of heritability, broad and narrow sense heritability. Broad sense heritability (H^2) is the proportion of phenotypic variation (σ^2P) that is the result of the total genetic variation (σ^2G) including both dominance and epitasis effects ($H^2 = \sigma^2G/\sigma^2P$). Narrow sense heritability (h^2) is the proportion of phenotypic variance

(σ^2P) that is the result of additive genetic variation, σ^2A, ($h^2 = \sigma^2A/\sigma^2P$) (Nyquist, 1991). The heritability for any trait can be calculated.

Certain misconceptions regarding heritability have been reported (Visscher *et al.,* 2008). Phenotypes are not passed to the next generation but their genes. Each parent passes half of its information to the next generation. Numerous studies have been conducted to determine the heritability of traits in maize.

Soleri and Smith (2002) rapidly estimated broad- sense heritability (H) of farmer-managed (landraces) maize populations in Mexico through participatory plant breeding. This was relevant for the local low- resource farmers to generate initial heritability information specific to local genetic populations and their response to the local environments. These estimates were compared to the published estimates for similar traits and with estimates of narrow- sense heritability (h^2) from experimental plots in the study. However, heritability estimates are particular to the genetic and environmental populations on which they are based (Falconer *et al.,* 1996), but the average estimates for species and traits can be used by breeders as a common reference for initial evaluation. High mean H^2 estimates across populations and locations were recorded for ear height (0.74%), plant height (0.65%), stalk diameter (0.67%), ear leaf width and length (0.65%), ear length (0.63%), days to anthesis (0.65%) and 100 kernel weight (0.61%).

Mahmood *et al.* (2004) conducted a study to determine heritability and genetic advance estimates in maize populations. The type of maize populations used in the study comprised of five maize hybrids and five open pollinated varieties. They reported high heritability (99%) for grain yield per plant and plant height. Heritability estimates for grain yield was considered on a single plant basis, while plant height was calculated on a plot basis. In addition, the number of days taken to silking and number of days taken to tasseling showed high estimates of 97% and 91% respectively. According to the study, the broad sense heritability and higher genetic advancement in grain yield per plant, plant height, days taken to silking and tasseling provided evidence that these traits were under the control of additive gene effects. In consideration of these traits, the conclusion was high heritability coupled with high genetic gains provides better chances for selecting plant material.

Shakoor *et al.* (2007) evaluated thirty maize double crosses with two checks to determine the broad sense heritability, genetic advance, correlations and path coefficient analysis of morpho-physiological traits. Traits such as grain yield per plant (78.10%), days to 50% silking (49%) and days to 50% tasseling (45.7%) were reported as highly heritable traits for grain yield improvement. Additionally, grain yield per plant and ear height was reported to have comparatively better genotypic variability, broad sense heritability and genetic advance which are good combination for effective selection of a trait. Ears per plant trait was found to have low broad sense heritability estimate of 29.4% and a genetic advance of 2% as mean percentage.

In cowpea, Shimelis and Shiringani (2010) reported moderate to high broad sense heritability estimates for the days to flowering (50%), seed yield (55%), number of branches per plant (53%) and days to maturity (66%). However, low heritability estimates were found for pods per plant (23%) and 100- seed weight (11%). They found that low level of heritability of the two traits necessitate indirect selection via other agronomic traits.

In pearl millet, Govindaraj *et al.* (2011) reported high heritability estimates for days to 50% flowering, plant height, number of production tillers, days to maturity and grain yield per plant.

Aminu and Izge (2012) investigated heritability and correlation estimates in maize under drought conditions. In the study, high broad-sense heritability estimates were reported for number of stands per plot (61.54 %), anthesis-silking interval (60.78 %), plant height (60.61%), weight of cobs (67.44 %) and grain yield (60.73 %). However, moderate broad sense heritability were also recorded for days to 50% tasseling (47.91 %), days to 50% silking (50.03 %), ear height (58.45 %) and dehusked cobs (55.06 %). Additionally, low heritability estimates were observed for the number of cobs per plant (37.21%), number of cobs per plot (34.62%) and 100 seed weight (31.99%). They concluded that heritability estimation is important in selecting suitable segregating generations for exhibiting the best expression of gene of different studied traits.

Bello *et al.* (2012) evaluated maize populations to explain heritability and genetic advance for grain yield and its component characters. High broad sense heritability

estimates were recorded for seedling emergence (75.28%), days to 50% tasseling (77.54 %), days to 50% silking (84.32%), ASI (61.79%), plant height (98.64%), ear height (92.54%), number of grains per ear (96.45%), ear weight (89.54%) and grain yield (98.16%).However, days to 50% pollen shed had low broad sense heritability estimates of 8.54%. Traits which possess both high heritability estimates and genetic gains were suggestively controlled by preponderance of the additive gene action. However, traits with high heritability estimates and low genetic advance can be improved through hybridization and hybrid vigour as observed for seedling emergence, days to anthesis and silking. These traits were suggested to be under the influence of the non-additive genetic control.

2.7 Genotypic and Phenotypic coefficient of variance

Mahmood *et al.* (2004) reported that genetic variability was observed in traits such as grain yield per plant (21.24%), kernels per row (14.68%) and 100-grain weight (12.71%). More consistency was observed for kernel rows per ear with a minimal level of genotypic coefficient of variance (5.48%). High phenotypic variance was recorded for grain yield per plant (21.32%) followed by number of kernels per row (15.56%). However, remarkably the magnitude of the phenotypic variance coefficient was greater than those of genotypic variance coefficient suggesting that the environmental influence was large. The environmental influence was clearly indicated in the case of 100-grain weight, where the phenotypic coefficient was 14.20% and the genotypic coefficient was 12.71%. Genetic advance values ranged between 43.80% for grain yield per plant to 1.33% for number of kernel row per ear.

Govindaraj *et al.* (2011) reported phenotypic coefficient variation values which were slightly higher than the genotypic coefficient variation values in most measured characters which indicate least role of environmental influences for expression of the characters.

Bello *et al.* (2012) observed a high proportion of phenotypic variance and phenotypic variance coefficient (PVC) over the genotypic variance and genotypic variance coefficient (GVC) across traits such as seedling emergence, days to 50% silking, days to 50% tasseling, number of grains per ear, plant height, ear height, ASI and plant height. Therefore, this indicates low environmental control in the expression of these traits. High genetic advance as percentage of the mean was reported for

number of grains (59.87), plant height (32.48), ear height (31.65), ear weight (28.37) and grain yield (26.54).

Langade *et al.* (2013) evaluated maize inbred lines for the utility of quality and morphological traits as a selection criterion for yield improvement. For all characters studied, phenotypic coefficient of variation was slightly higher compared to genotypic coefficient of variation. This signifies that the environment played a crucial role in expression of these traits. High broad sense heritability estimates were recorded for days to tasseling (99.5%), days to silking (99.4%), protein content (84.7%), starch content (76.6%) and number of kernels per row (74.8%). They found that these characters were least influenced by the environment, however, selection may not be useful as broad sense heritability is based on total genetic variance which is inclusive of both fixable and non-fixable variances. They also found that number of kernel rows per ear and kernel weight had low heritability suggesting that improvement through selection will be difficult due to environmental masking effects on genotypic effects. They concluded that a high, direct and positive effect on grain yield was associated with days to tasseling, plant height, ear height, ear length, ear diameter, kernel weight, oil content, protein content and starch content, therefore suggesting the effectiveness of direct selection. However, negative and direct effects were shown by days to silking, number of kernels per row and sugar content, thus indirect selection effectiveness.

However, Bekele and Rao (2014) reported that plant height (0.591%), ear height (0.395%), number of seed per row (0.427%), 100 seed weight (0.416%), protein content (0.961%) and grain yield per plot (0.558 %) had high heritability values. However, both genetic advance and high heritability were observed for plant height, grain yield per plot and protein percentage.

Aminu *et al.* (2014) reported high heritability estimates for days to 50% silking (59.67%), days to 50% tasseling (56.61%), anthesis –silking interval (64.23%), plant height (67.26%), ear height (56.83%) and grain yield (55.57%). Moderate estimates of heritability were recorded for number of stands per plot (49.36%), number of cobs per plant (50.69 %), number of cobs per plot (45.54%) and 100 seed weight (45.89%). The cobs weight and dehusked cobs recorded 38.06% and 44.23% respectively which was below average heritability values of 52.82%. Additive genetic

variance was observed for many traits which emphasized the significance of selecting segregating material for best expression of genes of various characters. They concluded that the high to moderate heritability indicate a significant opportunity for the development of drought tolerant and high yielding varieties through selection of desirable plants. Careful selection of traits with below average heritability values can lead towards improvement of such traits.

Heritability provides information on the degree to which the trait is controlled by inheritance (Rani and Sumalini, 2013); therefore determination of possible genetic advancement under selection (Govindaraj *et al.,* 2011). However, low heritability estimates for various traits indicate that such traits are influenced by the environment to a greater magnitude (Ali *et al.,* 2012).

2.8 Genetic diversity in maize

Genetic diversity studies in maize are well documented by various authors, therefore providing a rationale on the importance of such studies (Dao *et al*., 2014; Legesse *et al*., 2007; Cholastova *at al*., 2011). Dao *et al.* (2014) cited that genetic diversity in different populations provides and strengthens the adaptability to changing environments and market requirements. Essentially, in crop improvement, genetic diversity is crucial for: (i) analysis of genetic variability; (ii) identification of diverse parental combinations for exploitation of heterosis and to generate segregating progenies with high genetic variability for further selection; (iii) sustenance and expansion of the genetic base of the elite germplasm; and (iv) introgression of desired genes from diverse germplasm into the available genetic base. Additionally, Legesse *et al.* (2007) cited the importance of genetic diversity in the formation of heterotic groups for use as source materials in a breeding programme. Genetic progress in yield and other traits of economic importance in any breeding programme is highly dependant and influenced by the genetic variability within the breeding population (Cholastova *et al.,* 2011). Therefore, selection of the improved breeding material depends on the level of available genetic variability (Cholastova *et al.,* 2011).

There are serious implications for compromising genetic diversity. For several decades, maize breeders have focused on short term breeding, resulting in a constricted genetic base for commercial maize hybrids (Darrah and Zuber, 1986). One of the major concerns arising with the massive use of uniform commercial

varieties in maize production is the loss of its genetic variability (Drinic *et al.,* 2012). Currently in the world, the main maize hybrids cultivated involve a limited amount and range of key inbred lines (Le Clerc *et al.,* 2005). Therefore, there is limited genetic diversity in current hybrids compared to the large genetic diversity in gene banks (Le Clerc *et al.,* 2005). Evidently there is a gap between available genetic resources and the breeding programme activities. Drinic *et al.* (2012) outlined that searching for superior genotypes which possess high yielding ability, disease and pest resistance, stress tolerance and better nutritional quality is hard, competitive, long term and expensive. This is the reason why breeders focus on the adapted and improved materials while avoiding efficiently exploring the wild parents, landraces and exotic collections in the germplasm banks (Drinic *et al.,* 2012). Vigouroux *et al.* (2005) cited domestication bottleneck and selective breeding through directional selection as some of the causes for wide loss of genetic diversity. This is because over time only a limited portion of the population contributed to each subsequent generation.

Goodman (1999) cited that the narrow germplasm base is due to deriving newer lines and varieties from intercrosses of existing elite breeding material. Li *et al.* (2002) indicated 91.6% of hybrids parenthood in China consists of approximately 20 elite inbred lines. There has been a serious loss of diversity in the USA such that less than 10 hybrids form the basis of hybrid production in the country (Carvalho *et al.,* 2004). Therefore, such a restrictive base indicates that maize may not contain all the desirable and favorable alleles for maintenance of selection progress (Qi-Lun *et al.,* 2008). However, Duvick (1990) outlined there is no indication in maize that improvement rates have been negatively affected by the narrowing germplasm base. There is a concern that bottlenecks may restrict breeding flexibility and slow response to new opportunities, pests, pathogens, and agronomic practices in the future (Duvick, 1990). Ahmad *et al.* (2011) suggested the use of exogenous breeding material, such as landraces (Drinic *et al.,* 2012) and other materials from other geographies with specific genetic background to aid in broadening the existing genetic base. This is because landraces are a reserve of ancestral genes and have vast genetic variability as they were not subjected to selection over period of time. The landraces have characteristics such as high tolerance to biotic and abiotic stresses, herbicide tolerance, low anti-nutritional components content and large grain content of proteins, oil and starch (Drinic *et al.,* 2012).

2.9 Estimation of genetic diversity

Various methods have been used for the estimation and assessment of genetic diversity in maize. These include pedigree analysis, quantitative genetic analysis, heterosis in crosses, morphological, biochemical, physiological and cytological markers (Reid *et al.,* 2011). Pedigree analysis is somehow more reliable for inbred heterotic grouping. However it cannot always be used because there are many inbred lines that do not have clear or known pedigrees (Reid *et al.,* 2011). Additionally, pedigree analysis has limitations in providing genetic distance information. The use of morphological traits as markers proved to be ineffective and unreliable because the markers are often influenced by the genotype x environment interaction (Cholastova *et al.,* 2011). Although morphological data is still widely used in selection of genetically diverse parents, the morphological differences are determined by a small limited number of genes. This is not representative of the genetic diversity in the total genome (Hoxha *et al.,* 2004). Although, cytological and biochemical markers are used to monitor genetic diversity, however, they are limited in number and therefore cannot be used to study the complete genome of a specie (Cholastova et al, 2011).

To overcome the limitations of the foregoing, DNA markers have been developed and widely used to accurately estimate the levels of genetic diversity and examining relationships between maize inbred genotypes. This is because the method is independent of the environment interaction and functional in detecting and elimination of duplicates that might occur during germplasm collection and exchange (Hoxha *et al.,* 2004). Genetic diversity can be quantified using polymorphic information content (PIC), heterozygosity and availability. PIC values estimates the discriminatory power of a marker by taking into consideration the number of alleles at the locus as well as the relative frequencies of these alleles (Xu, 2010). Heterozygosity values specify the average proportion of individuals which are heterozygous for a given trait (Nei, 1973). Availability values determine the number of times the molecular marker worked. It is calculated as the number of samples genotyped for each marker (Nei, 1973). Limited genetic diversity has a negative impact on future maize breeding. Therefore for this study, it was practical to investigate genetic diversity using DNA molecular markers. These are summarized (Table 2.3 and 2.4).

Table 2.3: Comparison of different molecular markers

Type of marker	Advantages	Disadvantages	Application in Diversity study
AFLPs	High genomic abundance	Very tricky due to changes in	Kinship studies in Soya beans (Nimnual *et al*., 2014)
	High polymorphism	patterns with respect to materials used	
	No need for sequence information	Not reproducible, results in non- consistent map	
	Can be employed across species	Need to have very good primers	
	Work with smaller RFLP fragments		
	Useful in preparing coting maps		
SNPs	Can be automated	High development costs	Association Mapping in Wheat (Mengistu *et al*., 2016)
	Very Robust	Require sequence information	
	Suitable for high throughput	Technically challenging	
SSRs	High genomic abundance	Need sequence information	DNA Fingerprinting in Maize (Sharma *et al*., 2014)
	Highly reproducible	Cannot be used across species	
	Has a good genome coverage	Not well tested	
	High polymorphism		
	No radio active labeling		
	Easy to automate		
	Multiple alleles		
RAPDs	High genomic abundance	No probe or primer information	Assessment of genetic diversity in Maize (Abuali *et al*., 2011)
	Has a good genome coverage	Dominant markers	
	No sequence information	Not reproducible	
	Ideal for automation	Cannot be used across species	
	Poor DNA acceptable (Requires less DNA)	Not well tested	
	No radio active labeling		
	Relatively faster		
RFLPs	High genomic abundance	Need large amount of good quality DNA	Association Mapping in Sugarcane (Bilal *et al*., 2015)
	Co- dominant markers	Laborious (compared to RAPD)	
	Highly reproducible	Difficult to automate	
	Can use filters many times	Need radioactive labeling	
	Good genome coverage	Cloning and characterization of probe are required	
	Can be used across species		
	No sequence information		
	Can be used in plants reliably (well-tested)		
	Needed for map- based cloning		

Source: Adjusted after Kumar et al. (2009)

Table 2.4: Comparison of features for five commonly used markers in plants

Features	RFLP[1]	RAPD[2]	AFLP[3]	SSR[4]	SNP[5]
Genomic abundance	High	Very high	Very high	Medium	High
Genomic coverage	Low copy coding region	Whole genome	Whole genome	Whole genome	Whole genome
DNA required	50-10µg	1-100ng	1-100ng	50-120ng	≥50ng
Type of polymorphism	Singe base changes, indels	Singe base changes, indels	Singe base changes, indels	Changes in length of repeats	Single base changes, indels
Level of polymorphism	Medium	High	High	High	High
Effective multiplex ratio	Low	Medium	High	High	Medium to high
Inheritance	Co-dominant	Dominant	Dominant/co-dominant	Co-dominant	Co-dominant
Ease of use	Labour intensive	Easy	Difficult initially	Easy	Easy
Automation	Low	Medium	High	High	High
Reproducibility (reliability)	High	Low to medium	High	High	High
Types of probes/primers	Low copy DNA or cDNA	Usually 10bp random nucleotides	Specific sequence	Specific sequence	Allele-specific PCR primers
Radioactive detection	Usually yes	No	Usually yes	Usually no	No
Time demanding	High	Low	Medium	Low	Low
Development/start-up cost	High	Low	Medium	High	High
Property rights required	No	Yes and licensed	Yes and licensed	Yes and some licensed	Yes and some licensed
Suitable utility in diversity, genetics and breeding	Genetics	Diversity	Diversity and genetics	All purposes	All purposes

Source: Adjusted after Xu (2010)

[1] Restricted Fragment Length Polymorphism are first generation molecular markers which detect polymorphisms due to changes in the nucleotide sequences in recognition sites of restriction enzymes or due to mutation of several nucleotides causing a shift in fragment size (Lateef, 2015).

[2] Restricted Amplified Polymorphic DNA is a molecular marker that detects nucleotide sequence polymorphism in DNA using a single primer of arbitrary nucleotide sequence (Jonah et al., 2011).

[3] Amplified Fragment Length Polymorphisms are PCR based markers visualized by selective PCR amplification of DNA restriction fragments using directed primers from restriction of genomic DNA (Jonah et al., 2011).

[4] Simple Sequence Repeats (Microsatellites) are random tandem repeats of short nucleotide motifs of 2 to 6 base pairs. Polymorphisms are based on the variation in the number of repeats in various genotypes (Lateef, 2015).

[5] SNP Is a single nucleotide base difference between two DNA sequences arising either due to substitution (point mutations) and deletions or (Insertions) of nucleotides (Lateef, 2015).

2.10 Molecular Markers

There are various types of molecular markers that have been developed since the emergence of RFLPs (Restricted Fragment Length Polymorphism) in the 1980s (Phillips and Vasil, 2001). However, the current generation of molecular markers is based on direct analysis of sequence variation in each assay compared to an indirect analysis using probes in RFLPs or primers in PCR (Polymerase chain reaction) based primers. The current available molecular markers are Simple Sequence Repeats (SSR), Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD) and Single Nucleotide Polymorphism (SNP). Different markers have different properties which affects their application in genetic diversity (Rao and Hodgkin, 2002). Cholastova *et al.* (2011) used RAPD and SSR markers to discriminate the marker efficacy and determining genetic diversity in 30 maize hybrids. The study indicated that SSRs revealed high average PIC values (0.71) ranging from 0.47 to 0.91 than RAPD (0.61) ranging from 0.44 to 0.82. High genetic similarity values were recorded for SSRs (26.3 to 88.5 %) compared to RAPDs (6.7 to 86.7%). This indicated that SSR markers were highly effective compared to RAPD markers.

The inconsistency between the marker analysis made may be related to the amount of genome coverage ability of the particular marker system in species and its efficiency in sampling variation in any given population (Staub *et al.,* 1997). Genetic diversity assessment methods differ in the way they resolve genetic differences, the type of data generated and taxonomic levels at which that can be most appropriately applied (Rao and Hodgkin, 2002).In this regard the SNPs have become the markers of choice (see Tables 2.3 and 2.4).

2.10.1 Single Nucleotide Polymorphisms (SNP)

Single nucleotide polymorphisms (SNP) are single base changes in the genome sequence and are the abundant source of variation in plant and animal genomes (Lateef, 2015). SNP genotyping is broadly classified into two groups, the gel- based assays and the non-gel-based assays, the latter being mostly preferred to economize on time and money (Gupta *et al.,* 2001). SNPs are the only new generation molecular markers for individual genotyping needed for molecular marker - assisted selection (MAS) (Gupta *et al.,* 2001). A large number of SNP markers are now available as developed from the DNA sequence of known genes (Dao *et al.,* 2014). This consequently allows SNP markers to be the assay of choice for genetic diversity analysis and other variety of task related to crop improvement. Dao *et al.* (2014) reported on successfully using SNP sets in genetic characterization of tropical maize germplasm.

Noticeably, more genetic diversity studies and literature are currently based on SSR (Simple sequence repeats) compared to SNP markers. For germplasm characterization SSR markers are reported to provide much better information compared to SNP markers (Hamblin *et al.,* 2007). However, direct sequence analysis is understood to be the most robust form of analyzing the genome variation (Xu and Crouch, 2008). In this regard, the SNP marker analysis possesses certain advantages compared to previous generation of markers. Various advantages have been cited by different authors (Table 2.3 and 2.4). Xu (2010) cited several advantages of SNP markers as: (i) high abundance and even distribution through the genome; (ii) provide high reproducible codominant information; and (iii) cost effective, high throughput genotyping system. Additionally, Syvänen, (2005) reported the high probability of finding a marker within the gene of interest due to high density of SNPs across the genome. Furthermore, the SNP marker detection can be automated, therefore scaling up the analysis throughput to levels appropriate for applications in plant breeding programs (Syvänen, 2005; Xu and Crouch, 2008). This is because automation allows for the handling of large segregating populations (Gupta *et al.,* 2001).

2.10.2 Genetic distance

Apart from genetic diversity, breeders would value genetic distance data. Molecular markers provide information on the genetic distance (GD) for prediction of genetic variability; identification of best parent combinations and to assign lines into heterotic groups (Melchinger *et al.,* 1990). Therefore, GD plays a crucial role in maize breeding programmes as it assist in identifying divergent genotypes that exploit heterosis (Leal et al., 2010). Various genetic markers such as SSRs, SNPs , RFLP and AFLPs can estimate the genetic distance (Abakemal *et al.,* 2014). Hamblin *et al.* (2007) compared SSR and SNP markers in elucidating the population structure and genetic relationships among genotypes. It was reported that SSRs performed better and provided more resolution in measuring the genetic distance.

The major tool used for estimating the GD is the multivariate analysis. The GD analysis allows gathering many variables into one analysis (Bertan *et al.,* 2007). Any genetic distance studies in plant species comprise of six steps: (i) selection of genotypes to be analyzed; (ii) data production and formatting; (iii) selection of the distance definition or measurement to be used for the estimates; (iv) Selection of the clustering or plotting procedure; (v) analysis of the degree of distortion caused by the clustering or plotting procedure; and (vi) Data interpretation (Cruz and Carneiro, 2003).

Various statistical methods are used to estimate genetic distance in maize breeding programmes depending on the data set. Methods commonly used are the Nei and Li coefficient, Simple matching coefficient, Modified Roger's distance, Jaccards coefficient, Mahalanobis (D^2) and Euclidean distance (Roy, 2000). Mahalanobis distance and Euclidean distance are the most utilized statistical procedures to estimate GD (Bertan *et al.,* 2007). Mahalanobis distance has more advantage than the Euclidean distance as it factors in the environment effects and allows correlation between characters. However, its use is limited to data with more than one replication to estimate distance (Bertan *et al.,* 2007).

Genetic distance estimates between genotypes can be presented in various cluster or plotting methods. Plant breeders have adopted using the hierarchal methods whereby genotypes are grouped by a process that repeats itself at various levels, thus forming a dendogram without concern of the number of groups formed (Bertan *et al.,* 2007). However, with the various clustering methodologies available, consideration of the most suitable method and the data set is vital. This is the case where Tocher's clustering leads to the formation of one large cluster, whereas the UPGMA best discriminates the closer genotypes (Bertan *et al.,* 2007). The UPGMA cluster analysis method uses average distances between all genotype pairs for the formation of each group .In this study, the UPGMA method was used. This is because of the ability of the cluster method to show genetic relationships among genotypes and cluster them according to genotypes similarity. Some genotypes in the study were bi-parental progenies and shared a parent which may result in close genetic relationship. Odong *et al.* (2011) cited advantages of using UPGMA hierarchal cluster analysis. It can be used in various types of applications and is easy to understand while it is also available in many statistical packages (Odong *et al*., 2011).

2.11 Conclusion

The literature review has shown challenges that a breeder should consider when aiming to breed new productive hybrids. These mainly include the impact of the environment on maize plants and level of genetic diversity in a breeding population. Literature review has shown that the success of the breeding programme is somewhat dependant on exploiting new sources for genetic variation such as landraces and exotic populations in gene banks. Due to scarcity of resources to breeders, more emphasis is placed on developing new products on an existing narrow genetic base, rather than a quest for new sources of genetic variation to expand the current genetic base. Previous studies have shown that high level of genetic diversity in a population provides and strengthens adaptability to the ever changing environment. Further, with genetic diversity, there is an advantage of diverse parental combinations and creating segregating progenies to exploit more

heterosis through selection. Breeding of new highly productive hybrids is therefore dependant on the better combination between the parental lines. Hence, determination and information on the level of genetic diversity in the breeding material can not be over emphasized. Numerous authors have shown the success in employing genetic markers across various crops. The use of genetic markers in identifying and quantifying levels of genetic diversity has been well recorded in previous studies.

Previous studies have widely reported on obtaining increased levels of genetic gain in maize hybrid yields. One study reported a genetic gain in yield of 5%. Association of grain yield with other traits of interest has been well documented. This is because in improving grain yield, selection of other associated traits may be necessary. Studies have also shown the levels of heritability of grain yield in maize. A single study reported grain yield heritability of about 98% in maize crop. The following chapter discusses the level of genetic diversity found in maize lines developed in the UKZN breeding programme. It also outlines the implications of the research findings with regards to the future of hybrid production.

References

Abakemal, D., S. Hussein, J. Derera and K. Semagn. 2014. Genetic purity and patterns of relationships among tropical highland adapted quality protein and normal maize inbred lines using microsatellite markers. Euphytica, 204(1): 49-61.

Abay, F. and A. Bjørnstad. 2009. Specific adaptation of barley varieties in different locations in Ethiopia. Euphytica, 167(2): 181-195.

Abuali, A.I., A.A. Abdelmula, M.M. Khalafalla. 2011. Assessment of genetic diversity in Sudanese maize (Zea mays L.) genotypes using random amplified polymorphic DNA (RAPD) markers. African Journal of Biotechnology,10 (42) :8245-8250.

Abuali, A.I., A.A. Abdelmula, M.M. Khalafalla, A.E. Idris and N.B. Hamza. 2014. Assessment of genetic variability of inbred lines and their F1-hybrids of grain maize *(Zea mays L.)* under drought stress conditions.International Journal of Agronomy and Agricultural Research (IJAAR), 5 (2) : 22-30.

Adu, G. B., R. Akromah, M. S. Abdulai, K. Obeng-Antwi, A. W. Kena, K. M. L. Tengan, and H. Alidu. 2013. Assessment of Genotype by Environment interactions and Grain Yield Performance of Extra-Early Maize (*Zea Mays L.*) Hybrids. Journal of Biology, Agriculture and Healthcare, 3(12): 7-15.

Ahmad, S.Q., S. Khan, M. Ghaffar and F. Ahmad. 2011. Genetic diversity analysis for yield and other parameters in maize (*Zea mays L.*) genotypes. Asian Journal of Agricultural Sciences, 3(5): 385-388.

Akçura, M., S. Taner and Y. Kaya. 2011. Evaluation of bread wheat genotypes under irrigated multi-environment conditions using GGE biplot analyses. Agriculture, 98(1): 35-40.

Akter, A., H. M. Jamil, M. R. Islam, K. Hossain, and R. M. Mamunur. 2014. AMMI biplot analysis for stability of grain yield in hybrid rice (*Oryza Sativa L.*). Journal of Rice Research, 2(2): 126-129.

Ali, F., M.M. Durrishahwar, W. Hassan, H. Rahman, M. Noor, T. Shah, I. Ullar, M. Iqbal, K. Afridi and H. Ullah. 2012. Heritability estimates for maturity and morphological traits based on testcross progeny performance of maize. Asian Research Publishing Network Journal of Agricultural and Biological Science, 7(5): 1-8.

Aminu, D. and A. Izge. 2012. Heritability and Correlation Estimates in Maize (*Zea mays L*.) Under Drought Conditions in Northern Guinea and Sudan Savannas of Nigeria. World Journal of Agricultural Sciences, 8(6): 598-602.

Aminu, D., S.G. Mohammed and M. M. Jaliya. 2014. Gene action and heterosis for yield and other agronomic traits in maize (*Zea Mays L*.), under drought conditions in the northern Guinea and Sudan savannas zones of Borno state, Nigeria. International Journal of Advanced Biological Research, 4(2): 130-136.

Anshuman, V., N.N. Dixit , Dipika, S.K. Sharma and S. Marker. 2013. Studies on heritability and genetic advance estimates in maize genotypes. Bioscience Discovery, 4(2): 165-168.

Asfaw, A., F. Alemayehu, F. Gurum and M. Atnaf. 2009. AMMI and SREG GGE biplot analysis for matching varieties onto soybean production environments in Ethiopia. Scientific Research and Essay, 4(11): 1322-1330.

Bekele, A. and T.N. Rao. 2014. Estimates of heritability, genetic advance and correlation study for yield and it's attributes in maize (*Zea mays L*.). Journal of Plant Sciences, 2(1): 1-4.

Bello, O., S. Ige, M. Azeez, M. Afolabi, S. Abdulmaliq and J. Mahamood. 2012. Heritability and genetic advance for grain yield and its component characters in maize (*Zea mays L*.). International Journal of Plant Research, 2(5): 138-145.

Bertan, I., F. Carvalho and A.d. Oliveira. 2007. Parental selection strategies in plant breeding programs. Journal of Crop Science and Biotechnology, 10(4): 211-222.

Beyene, T., A. Botha and A. Myburg. 2005. Phenotypic diversity for morphological and agronomic traits in traditional Ethiopian highland maize accessions. South African Journal of Plant and Soil, 22(2): 100-105.

Bhering, L.L., B.G. Laviola, C.C. Salgado, C.F.B. Sanchez, T.B. Rosado and A.A. Alves. 2012. Genetic gains in physic nut using selection indexes. Pesquisa Agropecuária Brasileira, 47(3): 402-408.

Bilal, M., M. Saeed, I.A. Nasir, B. Tabassum, M. Zameer, A. Khan, M. Tariq, M.A. Javed and T. Husnain. 2015. Association mapping of cane weight and tillers per plant in sugarcane. Biotechnology & Biotechnological Equipment, *29*(4): 617-623.

Carvalho, V.P., C.F. Ruas, J.M. Ferreira, R.M. Moreira and P.M. Ruas. 2004. Genetic diversity among maize (*Zea mays L*.) landraces assessed by RAPD markers. Genetics and Molecular Biology, 27(2): 228-236.

Cholastova, T., M. Soldanova and R. Pokorny. 2011. Random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) marker efficacy for maize hybrid identification. African Journal of Biotechnology, 10(24): 4794-4801.

Cruz, C.D. and P.C.S. Carneiro. 2003. Biometrical models applied to plant breeding. Viçosa: Editora UFV: 579.

Dao, A., J. Sanou, S.E. Mitchell, V. Gracen and E.Y. Danquah. 2014. Genetic diversity among INERA maize inbred lines with single nucleotide polymorphism (SNP) markers and their relationship with CIMMYT, IITA, and temperate lines. BMC Genetics, 15(1): 1.

Darrah, L. and M. Zuber. 1986. 1985 United States farm maize germplasm base and commercial breeding strategies. Crop science, 26(6): 1109-1113.

da Silva, T.R.B., P.R.A. da Luz, C.A.T. da Silva, L. Facchi, F.L.R. Pereira and J.P. Poletine. 2012. Maize genotypes expression in environments of southern Brazil. Journal of Food, Agriculture & Environment, 10(1): 313-315.

Dowswell, C., R. Paliwal and R. Cantrell. 1996. Maize in the third world. Boulder, CO, USA. Westview Press.

Drinic, S.M., D.I. Micic and V. Andjelkovic. 2012. Genetic Diversity of Maize Landraces as Sources of Favorable Traits in: Genetic Diversity, book 2. Kemal University, Department of Biology Hatay, Turkey, InTech (ISBN 987-953-307-1384-7).

Du Plessis, J. 2003. Maize production,1-38. Department of Agriculture. http://www.arc.agric.za/arc-gci/Fact%20Sheets%20Library/Maize%20Production.pdf (Accessed 14 August 2015).

Duvick, D.N. 1990. The romance of plant breeding and other myths. Gene Manipulation in Plant Improvement II. Springer: 39-54.

Falconer, D.S., T.F. Mackay and R. Frankham. 1996. Introduction to quantitative genetics (4th edn). Trends in Genetics, 12(7): 280.

FAOSTAT. (2016).Production. http://faostat3.fao.org/browse/Q/*/E (Accessed 20 August 2016).

Forum for Agricultural Research in Africa (FARA). 2009. Patterns of change in maize production in Africa: Implications for maize policy development. Regional Policies and Markets, Issue no. 3.

Goodman, M.M. 1999. Broadening the genetic diversity in maize breeding by use of exotic germplasm. The genetics and exploitation of heterosis in crops. ASA-CSSA-SSSA, Madison, WI: 139-148.

Govindaraj, M., B. Selvi, S. Rajarathinam and P. Sumathi. 2011. Genetic variability and heritability of grain yield components and grain mineral concentration in India's pearl millet (*Pennisetum glaucum (L) R. Br*.) accessions. African Journal of Food, Agriculture, Nutrition and Development, 11(3): 4758- 4771.

Gupta, P., J. Roy and M. Prasad. 2001. Single nucleotide polymorphisms (SNPs): a new paradigm in molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science, 80(4): 524-535.

Grada, F. and S. Ciulca. 2013. Analysis of genotype x environment interaction for yield in some maize hybrids. Journal of Horticulture, Forestry and Biotechnology, 17(2): 192-196.

Hamblin, M.T., M.L. Warburton and E.S. Buckler. 2007. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. Public Library of Science 2: e1367.

Hoxha, S., M. Shariflou and P. Sharp. 2004. Evaluation of genetic diversity in Albanian maize using SSR markers. Maydica, 49(2004): 97-104.

IFPRI (International Food Policy Research Institute). 2003. Cereal demands in 1997 and 2020. Washington, DC, USA.

Jonah, P.M., L. L. Bello, O. Lucky, A. Midau and S. M. Moruppa. 2011. Review : The Importance of Molecular Markers in Plant Breeding Programmes. Global Journal of Science Frontier Research, 11(5): 1-9.

Kuchanur, P. H., P. M. Salimath, M. C. Wali, and Channayya Hiremath. 2015. GGE biplot analysis for grain yield of single cross maize hybrids under stress and non-stress conditions. Indian Journal of Genetics and Plant Breeding, 75(4): 514-517.

Kumar, N., S. Markar and V. Kumar. 2014. Studies on heritability and genetic advance estimates in timely sown bread wheat (*Triticum aestivum L*.). Bioscience Discovery, 5(1): 64-69.

Kumar, P., V. Gupta, A. Misra, D. Modi and B. Pandey. 2009. Potential of molecular markers in plant biotechnology. Plant Omics Journal, 2(4): 141-162.

Langade, D.M., J. Shahi, K. Srivastava, A. Singh, V. Agarwal and A. Sharma. 2013. Appraisal of Genetic Variability and Seasonal Interaction for Yield and Quality Traits in Maize (*Zea mays L*.). Plant Gene and Trait, 4(18): 95-103.

Lateef, D.D. 2015. DNA Marker Technologies in Plants and Applications for Crop Improvements. Journal of Biosciences and Medicines, 3: 7-18.

Leal, A., C. Mangolin, A. do Amaral Junior, L. Goncalves, C. Scapim, A. Mott, I. Eloi,V. Cordoves and M. Da Silva. 2010. Efficiency of RAPD versus SSR Markers for determining genetic diversity among popcorn lines. Genetics and Molecular Research, 9(1):9-18.

Le Clerc, V., F. Bazante, C. Baril, J. Guiard and D. Zhang. 2005. Assessing temporal changes in genetic diversity of maize varieties using microsatellite markers. Theoretical and Applied Genetics, 110(2): 294-302.

Legesse, B.W., A. A Myburg, K. V Pixley and A. M Botha. 2007. Genetic diversity of African maize inbred lines revealed by SSR markers. Hereditas, 144(1): 10-17.

Li, Y., J. Du, T. Wang, Y. Shi, Y. Song and J. Jia. 2002. Genetic diversity and relationships among Chinese maize inbred lines revealed by SSR markers. Maydica, 47(2): 93-102.

Mahmood, Z., S.R. Malik, R. Akhtar and T. Rafique. 2004. Heritability and genetic advance estimates from maize genotypes in Shishi Lusht a valley of Krakurm. International Journal of Agriculture and Biology, 2(2): 253- 257.

Melchinger, A., M. Lee, K. Lamkey and W. Woodman. 1990. Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. Crop Science, 30(5): 1033-1040.

Mengistu, D.K., Y.G. Kidane, M. Catellani, E. Frascaroli, C Fadda, M.E Pè and M. Dell'Acqua. 2016. High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. Plant Biotechnology Journal, 14: 1800-1812.

Mendes, F.F., L.J.M. Guimarães, J.C. Souza, P.E.O. Guimarães, C.A.P. Pacheco, J.R.D.A. Machado, W.F. Meirelles, A.R. Silva and S.N. Parentoni. 2012. Adaptability and stability of maize varieties using mixed model methodology. Crop Breeding and Applied Biotechnology, 12(2): 111-117.

Meridian Institute, 2014. Science and innovation for African agricultural value chains (maize value chain overview). merid.org/…/Value_Chain_Innovations. (Accessed 15 May 2014).

Miranda, G.V., L.V.d. Souza, L.J.M. Guimarães, H. Namorato, L.R. Oliveira and M.O. Soares. 2009. Multivariate analyses of genotype x environment interaction of popcorn. Pesquisa Agropecuária Brasileira, 44(1): 45-50.

Mitrović, B., D. Stanisavljević, S. Treskić, M. Stojaković, M. Ivanović, G. Bekavac, *et al.* 2012. Evaluation of experimental maize hybrids tested in multi-location trials using AMMI and GGE biplot analyses. Turkish Journal of Field Crops, 17(1): 35-40.

M'mboyi, F., Mugo, S., Mwimali, M. and Ambani, L. 2010. Maize Production and Improvement in Sub-Saharan Africa 2. Available: www.absfafrica.org [Accessed 27/02/2015].

Mortazavian, S. M. M., H. R. Nikkhah, F. A. Hassani, M. Sharif-al-Hosseini, M. Taheri, and M. Mahlooji. 2014. GGE biplot and AMMI analysis of yield performance of barley genotypes across different environments in Iran. Journal of Agricultural Science and Technology, 16(3): 609-622.

Munawar, M.S., M. Shahbazb, G. Hammada and M. Yasirc. 2013. Correlation And Path Analysis Of Grain Yield Components In Exotic Maize (*Zea Mays L.*) Hybrids. International Journal of Sciences: Basic and Applied Research, 12(1): 22-27.

Murtza, N., G. Shabbir, T. Mahmood, M. Ansar, M. Tabassum and M. Ilyas. 2014. Criterion for the selection of high yielding maize (*Zea Mays*) genotypes. Journal of Agricultural Research, 52(2): 177-183.

Najeeb, S., A. Rather, G. Parray, F. Sheikh and S. Razvi. 2009. Studies on genetic variability, genotypic correlation and path coefficient analysis in maize under the high altitude temperate conditions of Kashmir. Maize Genetics Cooperation Newsletter, 83:46

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences, 70(12): 3321-3323.

Nimnual, A., J. Romkaew, E. Chukeatirote and S. Nilthong. 2014. Evaluation of genetic relationship among some important Japanese and Thai soybean varieties using AFLP analysis. Australian Journal of Crop Science, 8(4): 481-485.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Critical Review in Plant Sciences, 10(3): 235-322.

Nzuve, F., S. Githiri, D. Mukunya and J. Gethi. 2013. Analysis of genotype x environment interaction for grain yield in Maize hybrids. Journal of Agricultural Science, *5*(11) :75-85.

Odong, T. L., J. Van Heerwaarden, J. Jansen, T.J.L. van Hintum, and F. A. Van Eeuwijk. 2011. Determination of genetic structure of germplasm collections: are traditional hierarchical clustering methods appropriate for molecular marker data?. Theoretical and Applied Genetics, 123(2): 195-205.

Phillips, R.L. and I.K. Vasil. 2001. DNA-based markers in plants (Vol. 6). Springer Science & Business Media.

Pingali, P.L. 2001. CIMMYT 1999/2000 World maize facts and trends. Meeting world maize needs: technological opportunities and priorities for the public sector. International Maize and Wheat Improvement Center.

Qi-Lun, Y., F. Ping, K. Ke-Cheng and P. Guang-Tang. 2008. Genetic diversity based on SSR markers in maize (*Zea Mays L.*) landraces from Wuling mountain region in China. Journal of Genetics, 87(3): 287-291.

Rahnejat, S.S. and E. Farshadfar. 2015. Evaluation of phenotypic stability in canola *(Brassica napus)* using GGE-biplot. International Journal of Biosciences, 6(1): 350-356.

Rani, T. and K. Sumalini. 2013. Coheritability in wilt tolerant maize hybrids. Sky Journal of Agricultural Research, 2(11): 159 - 162.

Rao, V.R. and T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. Plant cell, Tissue and Organ Culture, 68(1): 1-19.

Reddy, V.R., F. Jabeen, M. Sudarshan and A.S. Rao. 2013. Studies on genetic variability, heritability, correlation and path analysis in maize (*Zea Mays L.*) Over locations. International Journal of Applied Biology and Pharmaceutical Technology, 4(1): 196-199.

Reid, L., K. Xiang, X. Zhu, B. Baum and S. Molnar. 2011. Genetic diversity analysis of 119 Canadian maize inbred lines based on pedigree and simple sequence repeat markers. Canadian Journal of Plant Science, 91(4): 651-661.

Roy, D. 2000. Plant Breeding—Analysis and Exploitation of Variation. Narosa, New Delhi.

Sesay, S., D. Ojo, O.J. Ariyo and S. Meseka. 2016. Genetic variability, heritability and genetic advance studies in topcross and three-way cross maize (*Zea mays L*) hybrids. Maydica,61(2): 1-7.

Shakoor, M.S., M. Akbar and A. Hussain. 2007. Correlation and path coefficients studies of some morphophysiological traits in maize double crosses. Pakistan Journal of Agricultural Sciences, 44(2): 213-216.

Sharma, J.K., A. Singh and S. Lata. 2014. DNA fingerprinting of commercial maize hybrids and their parental lines using simple sequence repeat markers. Journal of Crop Improvement , 41(1) : 69-75.

Shiferaw, B., B.M. Prasanna, J. Hellin and M. Bänziger. 2011. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Security, 3(3): 307-327.

Shimelis, H. and R. Shiringani. 2010. Variance components and heritabilities of yield and agronomic traits among cowpea genotypes. Euphytica, 176(3): 383-389.

Silveira, L.C.I., V. Kist, T.O.M. de Paula, M.H.P. Barbosa, L. Alexandre and E.D. Peternelli. 2013. AMMI analysis to evaluate the adaptability and phenotypic stability of sugarcane. Scientia Agricola, 70(1): 27-32.

Singhal, N., S.S. Verma, D.C. Baskheti and A. Kumar. 2006. Heritability, genetic advance, correlation and path-coefficient estimation in high quality protein maize (*Zea mays L.*). Asian Journal of Bio Science, 1(2): 54-56.

Soleri, D. and S. Smith. 2002. Rapid estimation of broad sense heritability of farmer-managed maize populations in the Central Valleys of Oaxaca, Mexico, and implications for improvement. Euphytica, 128(1): 105-119.

Sousa, L. B., O. T. Hamawaki, A. P. Nogueira, R. O. Batista, V. M. Oliveira, and R. L. Hamawaki. 2015. Evaluation of soybean lines and environmental stratification using the AMMI, GGE biplot, and factor analysis methods. Genetics and Molecular Research, 14(4): 12660-12674.

Staub, J.E., J. Box, V. Meglic, T.F. Horejsi and J. McCreight. 1997. Comparison of isozyme and random amplified polymorphic DNA data for determining intraspecific variation in Cucumis. Genetic Resources and Crop Evolution, 44(3): 257-269.

Syvänen, A.C. 2005. Toward genome-wide SNP genotyping. Nature Genetics, 37: S5-S10.

Vigouroux, Y., S. Mitchell, Y. Matsuoka, M. Hamblin, S. Kresovich, J.S.C. Smith and J. Doebley. 2005. An analysis of genetic diversity across the maize genome using microsatellites. Genetics, 169(3): 1617-1630.

Visscher, P.M., W.G. Hill and N.R. Wray. 2008. Heritability in the genomics era—concepts and misconceptions. Nature Reviews Genetics, 9(4): 255-266.

www.iita.org. Accessed 13/09/2014.

Xu, Y. 2010. Molecular plant breeding. CABI.

Xu, Y. and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: from publications to practice. Crop Science, 48(2): 391-407.

Yan, W., M.S. Kang, B. Ma, S. Woods and P.L. Cornelius. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Science, 47(2): 643-653.

Yan, W. and N.A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Canadian Journal of Plant Science, 86(3): 623-645.

Yokomizo, G.K.I., J.B. Duarte, N.A. Vello and J.R. Unfried. 2013. AMMI analysis of grain yield in soybean lines selected for resistance to Asian rust. Pesquisa Agropecuária Brasileira, 48(10) : 1376-1384.

Yusuf, M. 2010. Genetic variability and correlation in single cross hybrids of quality protein maize (*Zea Mays L.*). African Journal of Food, Agriculture, Nutrition and Development, 10(2): 2166- 2175.

CHAPTER THREE

Genetic diversity and genetic distances between UKZN advanced maize lines with founder parents and standard public lines based on SNP markers

Abstract

Genetic diversity is of crucial importance in exploitation of heterosis and has a direct impact on maize breeding strategies through resulting genetic gains thereof. Information on the genetic diversity existing between possible parental lines should be identified for the production of vigorous hybrids. Therefore this study reports genetic analysis of 51 inbred lines sub-grouped as elite , advanced and recombinant inbred lines using 365 single nucleotide polymorphism (SNP) markers to (i) establish genetic characterization of 51 experimental inbred lines, (ii) Establish line divergence through quantifying genetic distances between the inbred lines, and (iii) Establish and partition genotypes into clusters to different heterotic groups. The results revealed genetic diversity ranging from 0.00 to 0.50 with a mean of 0.24. The Polymorphism information content of the SNP markers ranged from 0.00 to 0.38 indicating the potential of markers to detect differences among genotypes. The highest genetic distance of 0.3526 was recorded between founder parent LN43 and elite line LN50 suggesting these lines are complementary to each other. The cluster analysis successfully discriminated genotypes into two major clusters.

Introduction

Genetic diversity in maize breeding programmes is of crucial importance. One of the important roles it plays it partitioning of germplasm according to heterotic groups (Legesse *et al.,* 2006) . In crop improvement, genetic diversity is of importance for identification of diverse parental combinations. Genetically divergent genotypes express high hybrid vigour, therefore exploitation of heterosis (Dao *et al.,* 2014).

The single nucleotide polymorphism (SNP) markers are of high preference due to lower genotyping errors (Foster *et al.,* 2010), high genomic abundance, even distribution in the genome, highly producible and throughput and highly cost effective (Xu *et al.,* 2009). Dao *et al.* (2014) reported successful use of 1057 informative SNPs revealing genetic variation among 96 inbred lines from different sources (temperate, CIMMYT and IITA) and between the INERA germplasm set. Laserna *et al.* (2015) successfully determined genetic diversity among transgenic and non-transgenic versions of a single cross hybrid using SNPs. It was reported 54% of genetic diversity was found within the hybrid version, however, 45.1% was reported for the genetic diversity among versions of the same hybrid.

In the current study, SNP markers were used to genotype 51 maize inbred lines. The study aimed to determine the genetic diversity and variation among 51 maize inbred lines which can be used for hybrid makeup, in the breeding programme at the University of KwaZulu – Natal (UKZN).

The objectives of this study are:

1. To genetically characterize the maize inbred lines of interest and determine levels of genetic diversity.
2. To determine the genetic distances between the maize inbred lines.
3. To determine the clustering of the maize inbred lines into heterotic groups.

3.1 Materials and methods

3.1.1 Germplasm

A total of 51 experimental maize inbred lines were genotyped in the study. The germplasm used in this study is divided into three groups. The first group comprise of 13 elite inbred lines namely; LN47, LN01, LN02, LN03, LN04, LN05, LN44, LN45, LN46, LN48, LN49, LN50, LN51. The second group consists 16 advanced inbred lines derived from 2 bi-parental populations with their 3 founder

parents: LN41, LN42 and LN43. The two populations share one parent (LN43) in common. The parents are in heterotic group A in the African tropical germplasm. Lines LN01, LN02, LN03, LN04, LN05, LN47, LN48, LN49, LN50 and LN51 are used as standard lines drawn from public programmes. LN01 is a tropical line while LN02 is South African adapted line. LN03 and LN49 are temperate checks, while LN05 is a CIMMYT tropical, mid – altitude line. Line LN47 is an African Iodant type. The third group comprise of 19 recombinant inbred lines LN19, LN21, LN22, LN23, LN24, LN25, LN26, LN27, LN28, LN29, LN30, LN31, LN32, LN33, LN34, LN36, LN37, LN38 and LN39.

The advanced inbred lines were derived by pedigree selection and were selfed until the F8 generation. The F1 population was derived from crosses between the 3 founder parents (LN43 X LN41) and (LN43 X LN42), therefore LN43 being a common parent of the two populations. Characteristics of the founder parent are:

LN43: This is a white grain, low land adapted and late flowering line. It has high ear prolificacy and possesses downy mildew resistance. However, it is susceptible to root lodging.

LN41: This line is a white grain, mid altitude adapted and early flowering type. It also has Maize Streak Virus (MSV) resistance, good standing ability, but with low yield potential.

LN42: It is a white grain line which is early maturing, maid altitude adapted, MSV resistance and has high yield potential and good general combining ability (GCA) for grain yield perfomance. However, it disadvantage is susceptibility to root lodging.

The F1 was advanced to F2 population then selection started thereof. Selection criteria used was based on good standing ability, pollination ability and kernel texture (flints) Pedigree selection was done in summer at Ukulinga Research Station in Pietermaritzburg, South Africa, located 810 metres above sea level (masl). In winter season, selection was done at Makhathini Research farm located 72 metres above sea level.

3.1.2 Genotyping

The SNP genotyping was done at the LGC Genomics platform in the United Kingdom. A total of 396 SNP markers (Appendix 1) were used to genotype the 51 maize inbred lines using the internal protocols of the KASP system at the LGC Genomics, United Kingdom (www.lgcgroup.com) accessed 17/02/2015.

3.1.3 Data analysis

The 396 SNP markers are listed in Appendix 1. Genotypic data was analysed using Power marker software version 3.25. The following statistics were determined: polymorphic information content (PIC), gene diversity, availability, inbreeding coefficient (f), heterozygosity and genetic distance for each SNP marker. Polymorphic information content values were calculated using the following formula (Xu, 2010):

$$\mathrm{PIC} = 1 - \sum_{j=1}^{n} P_{ij}^2$$

Where, P_{ij} is the frequency of the *j*th allele for marker *i* and the summation extends over *n* alleles.

PIC values provides an estimation of the discriminating power of a marker by taking into account not only the number of alleles at the locus but also the relative frequencies of these alleles.

Allelic diversity was calculated using the following formula (Singh, 1983):

$$Div = 1 - \sum_{u=1}^{k} P_u^{\sim 2}$$

Where *k* = number of alleles, $P_u^{\sim 2}$= frequency of the marker allele.

Allelic diversity values give an estimate of heritable characteristics present in a population of the same species. Heterozygosity was calculated using the following formula (Nei, 1973):

$$Ho = 1 - \sum_{i=1}^{n} (fi)^2$$

Assuming that a marker has 1,2,…*n* alleles, the *i* allele occurs by frequency f_i .
Heterozygosity values indicate the average proportion of individuals that are heterozygous for a given trait.
Genetic distance was calculated using the following formula (Nei, 1973):

$$GD = 1 - [2N_{11}/(2N_{11} + N_{10} + N_{01})]$$

Where, N_{11} = number of alleles in both genotypes *i* and *j*, N_{10} = number of alleles in genotype *I* and N_{01} = number of alleles in the genotype *j*.

Availability values determine the number of times the molecular marker worked. It is calculated as the number of observations/number of samples genotyped for each marker.
Dendograms were generated using Power marker software (v3.25) and displayed in MEGA5.2.
In-breeding coefficient (F) is the proportion of the homozygous genotypes for each marker.

3.2 Results

3.2.1 Marker characterization

3.2.1.1 SNP markers

The SNP markers were used to effectively genotype the inbred lines.

Precisely 396 SNP sets identified 748 alleles among 51 maize inbred lines. The number of alleles scored ranged between one and two, with mean of 1.89 (Figure 3.1). About 88% (352 out of 396) markers revealed 2 alleles. These markers include Fea2_1, Fea2_2, PHM12794_47, PHM4531_46, PZA00172_11 and PZA00334_2. Therefore, 12 % (44 out of 396) markers including PZA00444_5, PZA00489_1, PZA00565_3, PZA00587_6, PZA00600_11, PZA00793_2, PZA02949_26, PZA03012_10, PZA03366_2, PZA03384_1, PZA03411_3, PZA03431_1, PZA03668_4 and PZA03673_2 revealed 1 number of alleles respectively.

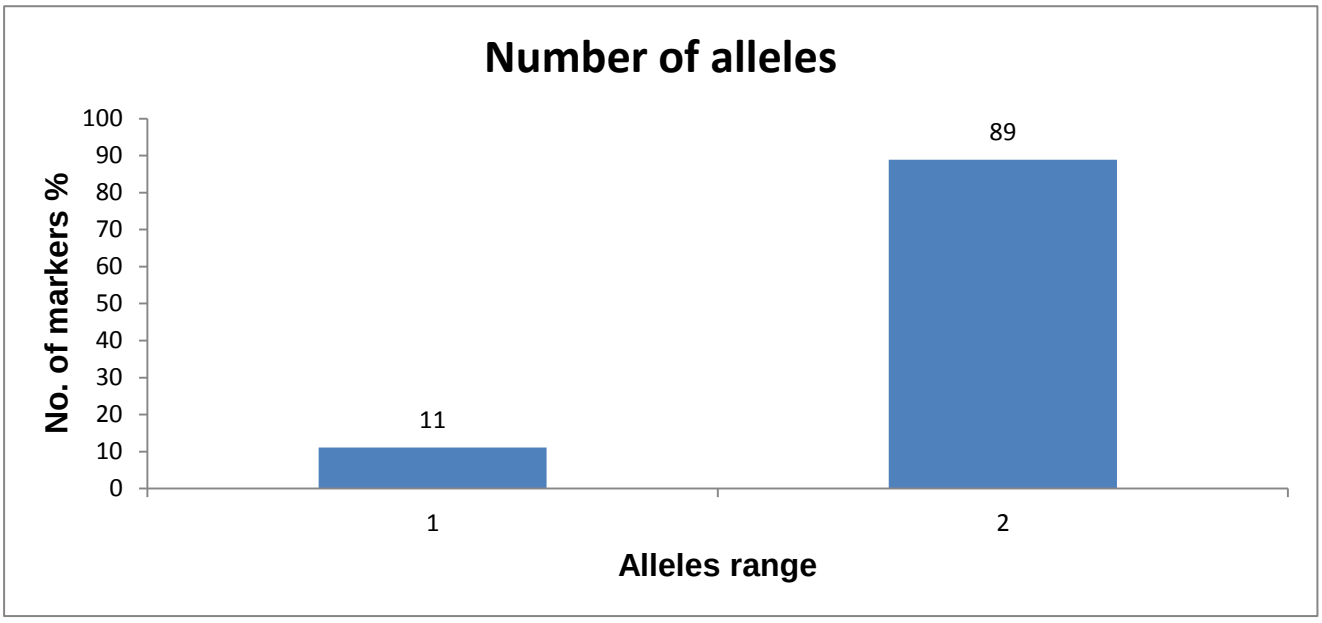

Figure 3.1: Distribution of the 396 SNP markers for number of alleles

The polymorphism information content (PIC) value of the SNPs ranged from 0.00 - 0.38 with a PIC mean of 0.19 (Figure 3.2). The PIC values indicate the potential information on the markers ability to detect differences among the lines. Approximately 12% (44 out of 396) of markers

PZA00444_5, PZA00489_1, PZA00565_3, PZA00587_6, PZA00600_11, PZA00793_2, PZA02949_26, PZA03012_10, PZA03366_2, PZA03384_1, PZA03411_3, PZA03431_1, PZA03668_4 and PZA03673_2 obtained PIC score of 0.00. While markers PHM12794_47, PZA03533_1, PZA03732_3, PZA02496_1 and PZB00087_1 obtained the highest PIC of 0.38. Approximately 50% of markers used (199 out of 396) constituted PIC values equal to, or greater than, 0.20. The ten SNPs (Appendix 1) exhibiting the highest PIC (0.3749 – 0.3750) and their potential to detect differences between the inbred lines were; PZB01457_1, PZA00334_2, PZA00562_4, PZA03116_2, PZA03716_1, PZA02496_1, PHM12794_47, PZA03533_1, PZA03732_3 and PZB00087_1. The high PIC value revealed in this study might be a pertinent indication confirming the potential for these SNP markers to discriminate between inbred lines from diverse origins. This was shown by the fact that markers were able to separate closely related lines, therefore indicating their usefulness for diversity analysis in maize under the current study. These markers will be equally useful for determining genetic diversity of any maize population especially if coupled with higher allelic frequencies which is an indication of genetic diversity in a population. Therefore 35% of the markers used in the study had allele frequency ranging from 0.90 to 0.99 %.

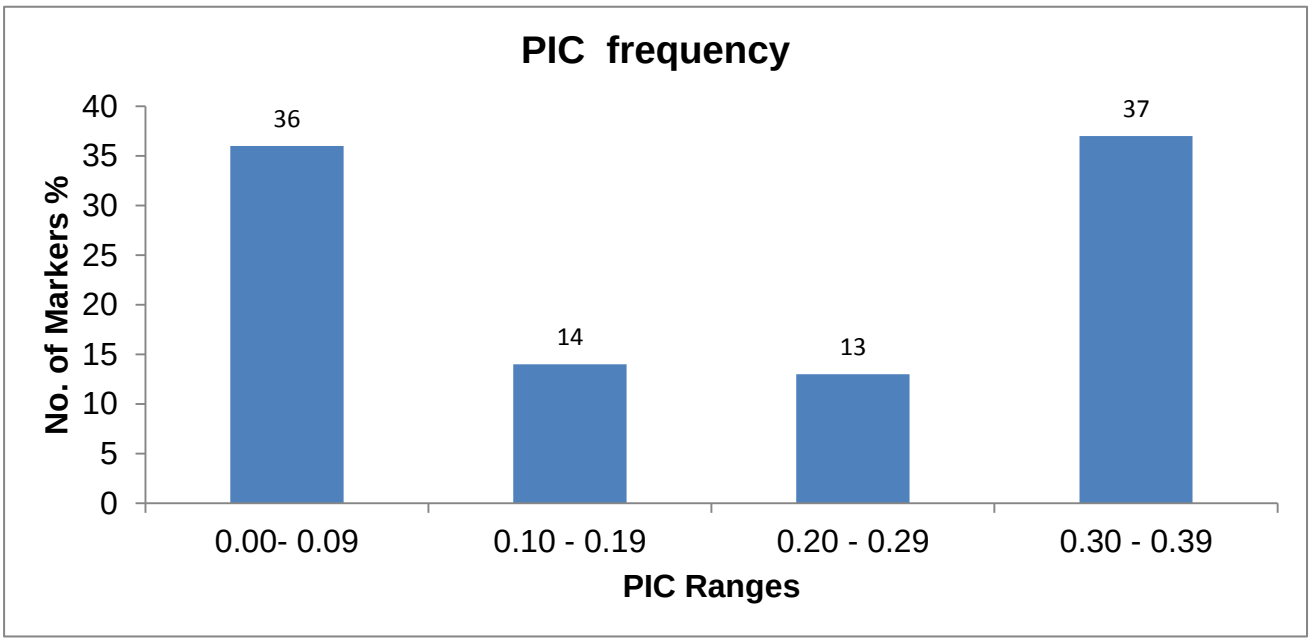

Figure 3.2: Distribution of the 396 SNP markers for Polymorphic Information Content

Genetic diversity ranged from 0.00 to 0.50 with a mean of 0.24 (Figure 3.3). Markers PZA00444_5, PZA00489_1, PZA00565_3, PZA00587_6, PZA00600_11, PZA00793_2, PZA02949_26, PZA03012_10, PZA03366_2, PZA03384_1, PZA03411_3, PZA03431_1, PZA03668_4 and

PZA03673_2 indicated diversity values of 0.00, respectively. However, 6% (25 out of 396) of markers including PHM12794_47, PZA00334_2, PZA00562_4, PZA00603_1, PZA02722_1, PZA03116_2 and PZA03477_1 obtained the highest diversity of 0.40. Approximately 162 markers revealed genetic diversity equal to, or greater than 0.30.

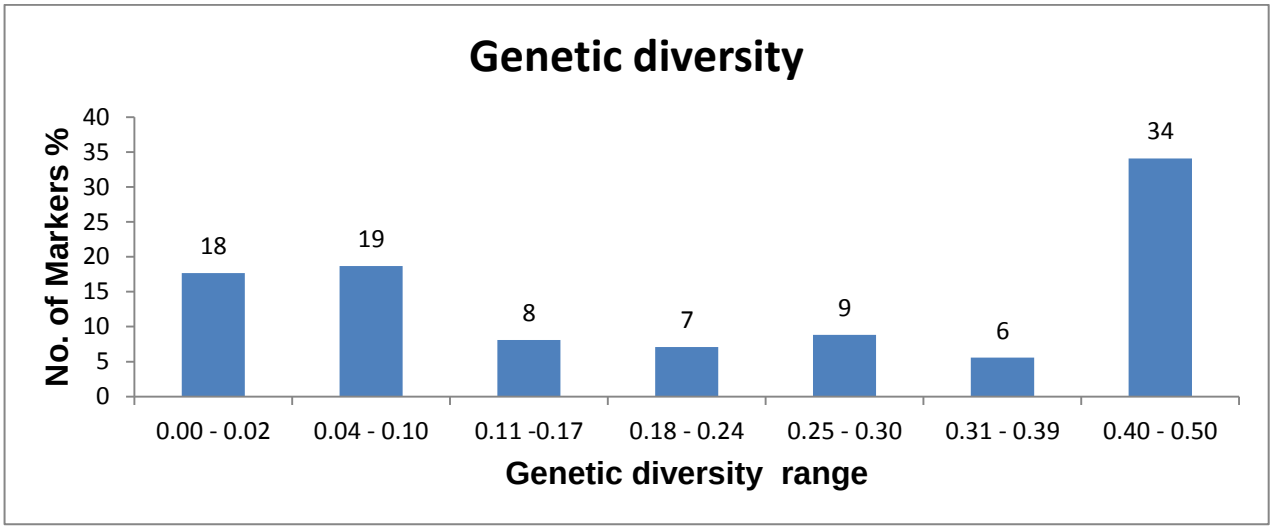

Figure 3.3: Distribution of the 396 SNP markers for genetic diversity

Availability scores ranged from 0.90 to 1.00, with a mean of 0.98 (Figure 3.4). Therefore, 6 markers PZA01607_1, PZA03692_1, PZA03758_1, PZB00153_3, PZB01115_1 and lac1_3 had a 90% call rate. About 231 markers scored 100% call rate, including Fea2_1, Fea2_2, PHM12794_47, PHM4196_27, PHM4348_16 and PHM4531_46 respectively.

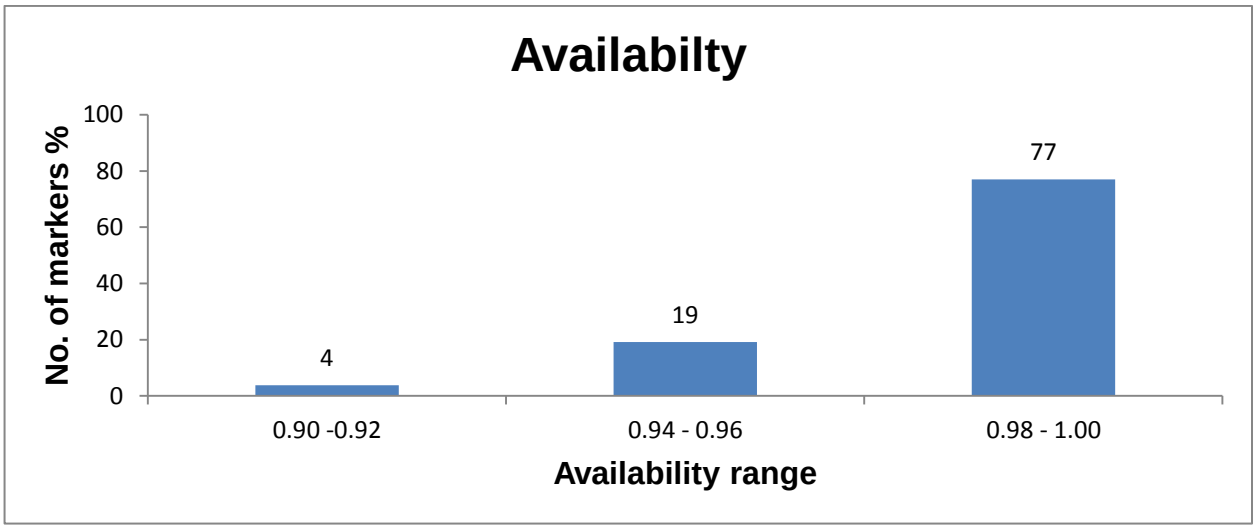

Figure 3.4: Distribution of the 396 SNP markers for availability

Heterozygosity values ranged between 0.00 and 1.00, with a mean of 0.11(Figure 3.5). About 18% markers had scores of heterozygosity equals to 0.00 and these are PZA00172_11, PZA00444_5, PZA00462_2, PZA00489_1, PZA00565_3, PZA00578_2 and PZA00587_6, while 2% (9 out of 396 markers) PZA03533_1, PZA03716_1, PZA03732_3, PZB01856_1, PZB01869_4, PZB02033_1, PZA02266_3, PZB00087_1 and PZB01446_1 had a range between 0.61 and 1.00 in heterozygosity.

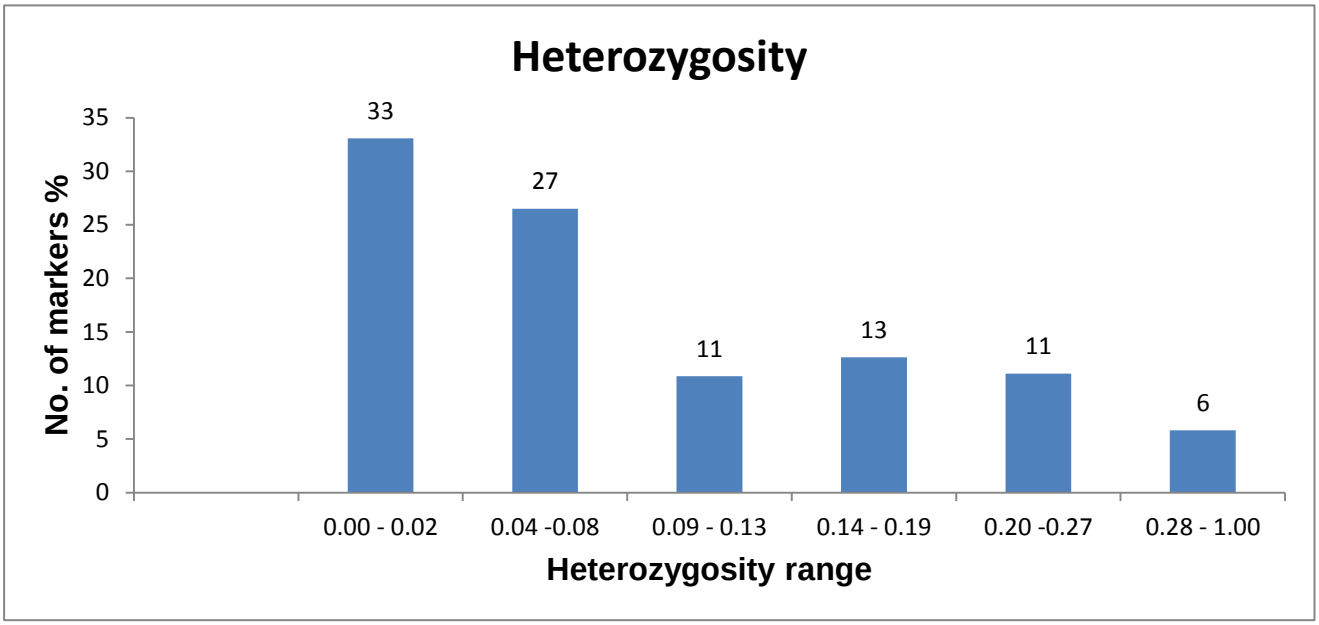

Figure 3.5: Distribution of the 396 SNP markers for heterozygosity

Inbreeding co-efficient (f) ranged between 0.00 and 1.00, with a mean of 0.56 (Figure 3.6).

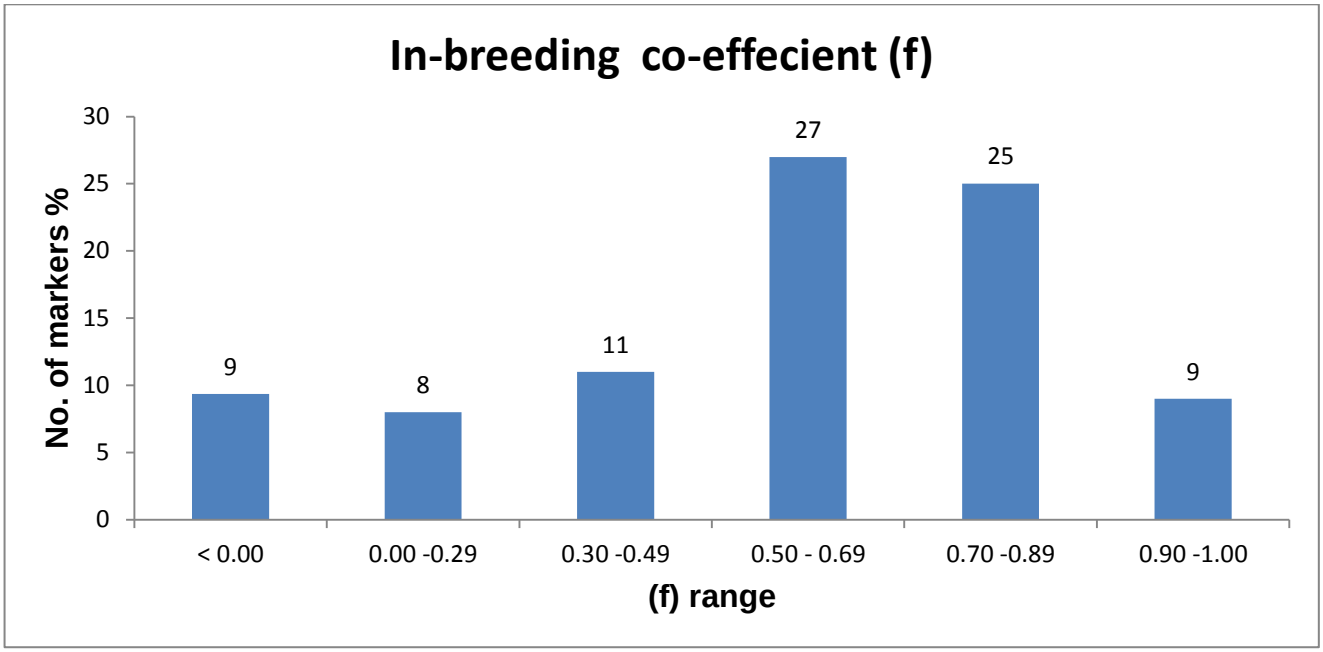

Figure 3.6: Distribution of the 396 SNP markers for inbreeding co-efficient

3.2.2 Cluster analysis

3.2.2.1 Cluster analysis of inbred lines based on SNP molecular markers

The cluster analysis was performed using the UPGMA method. Only SNP markers with a call rate above 90% were used. The SNP markers were effective for dividing the 51 maize inbred lines into six different genetic clusters (Figure 3.7). The results of clustering revealed the 51 genotypes could be classified into two major groups (I and II) at the truncation level of 0.14 in the coefficient scale. Cluster 2 is larger than Cluster 1, which contains only one inbred line LN47. At 0.13 cut off point two groups are identified (A and B), in which Group B is further subdivided into two sub-clusters (B1 and B2). The B1 sub- cluster comprised of genotypes LN50, LN45, LN48, LN02, LN44, LN03, LN49, LN04 and LN46. Sub-cluster B2 comprised of LN21. At 0.10 cut off point a sub-group comprising E and F were identified. Sub-group E was not subdivided therefore standing on its own; however, sub-group F was sub-divided into two sub- clusters (F1 and F2) at 0.07 cut off point. Additionally, F1 and F2 sub-cluster further sub-divided to more sub-groups (G, H, I and J). Inbred lines LN21, LN31, LN23 and LN29 were in the same group, under sub-group G, while LN28, LN38, LN37, LN25, LN30, LN39, LN22, LN27, LN32, LN24, LN26, LN33, LN34, LN19 and LN36 were clustered under sub-group H. Sub-group I comprised of inbreds LN35, LN14, LN43, LN17, LN07, LN06, LN15, LN16 and LN40, while sub-group J clustered together inbred LN41, LN08, LN20, LN11, LN13, LN09, LN10, LN12 and LN16, respectively. Noticeably, Inbred LN47 is in Cluster 1 on its own which indicates that it may be used to make productive hybrids. Seemingly all these lines originate from a source material developed in the UKZN breeding programme (Table 3.4).

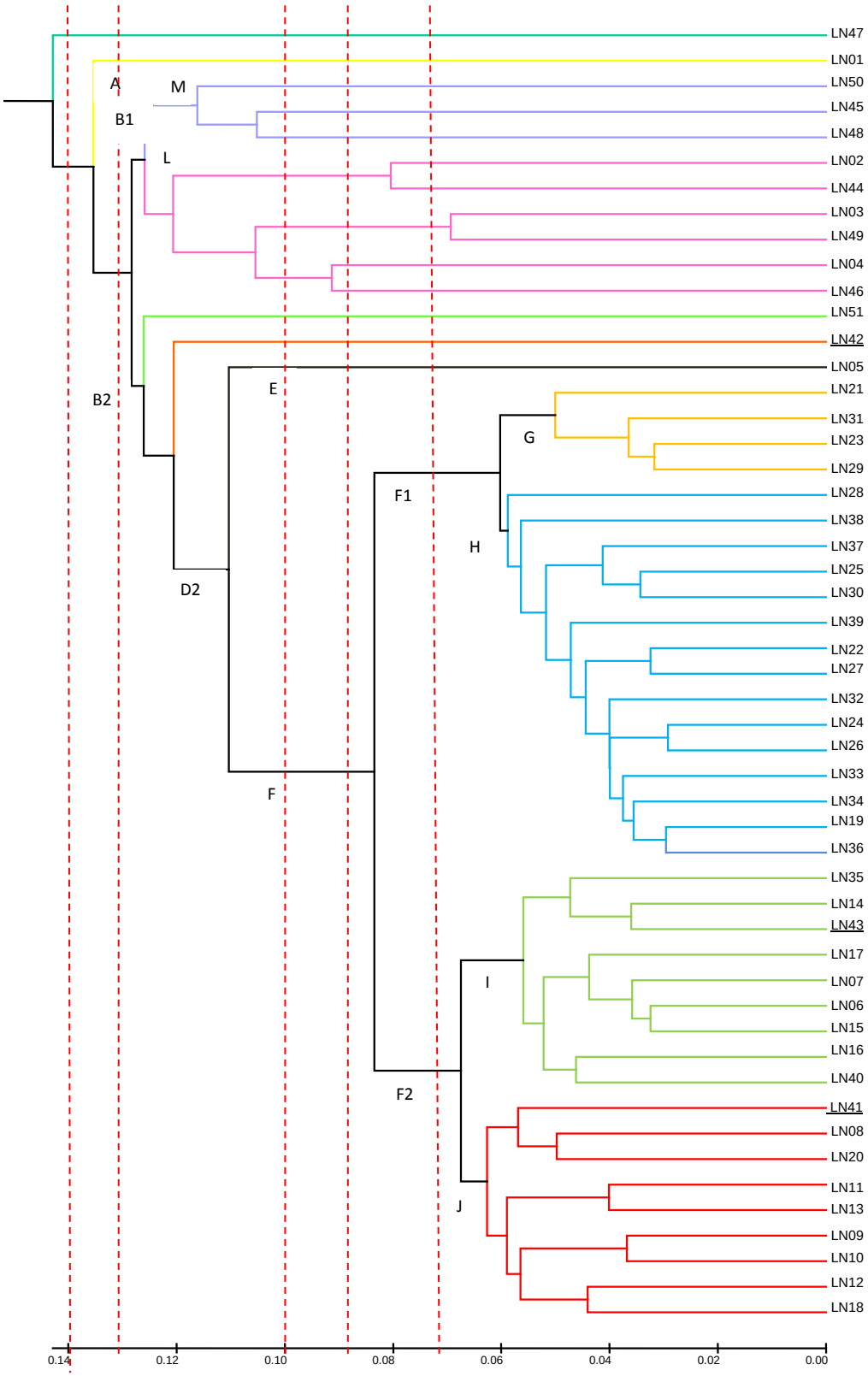

Figure 3.7: Dendogram of 51 maize experimental inbred lines based on 396 SNP molecular markers using UPGMA

Table 3.1: Diallel matrix of Elite inbred lines clustered under group A to D in the dendogram (Figure 3.7)

	LN01	LN02	LN03	LN04	LN05	LN44	LN45	LN46	LN47	LN48	LN50	LN51
LN01	0.0000											
LN02	0.2572	0.0000										
LN03	0.2504	0.2255	0.0000									
LN04	0.2727	0.2449	0.2007	0.0000								
LN05	0.3036	0.2723	0.2484	0.2555	0.0000							
LN44	0.2587	0.1609	0.2492	0.2397	0.2656	0.0000						
LN45	0.2616	0.2394	0.2486	0.2401	0.2558	0.2341	0.0000					
LN46	0.2046	0.1795	0.1603	0.1826	0.2210	0.1950	0.2152	0.0000				
LN47	0.2962	0.2729	0.2431	0.2741	0.2954	0.3006	0.2777	0.2112	0.0000			
LN48	0.2542	0.2577	0.2012	0.2490	0.2154	0.2380	0.2105	0.2044	0.2398	0.0000		
LN50	0.3019	0.2688	0.2709	0.2784	0.3153	0.2718	0.2121	0.2156	0.3008	0.2527	0.0000	
LN51	0.2678	0.2624	0.2313	0.2385	0.2436	0.2725	0.2775	0.2099	0.2769	0.2302	0.3105	0.0000

3.2.3 Genetic distance

3.2.3.1 Genetic distance between elite inbred lines

Genetic distances were revealed for 1275 inbred combinations in a diallel matrix. Some of the inbred combinations obtained are shown in Table 3.3, Table 3.4, Table 3.5 and Table 3.6. The total distances between all genotypes ranged from 0.0584 to 0.3526. The genetic distances partitioned in four ranges with 9% of the inbreds showing low genetic distances ranging from 0.05 to 0.09, while 5% of the inbred lines had genetic distances ranging between 0.30 – 0.35 (Table 3.2). Noticeably, 44% of the inbred lines had genetic distances in the range of 0.20 to 0.29 and another 41% had genetic distances ranging between 0.10 – 0.19 (Table 3.2).

Diallel matrix of the elite inbred lines clustered under group A to D in the dendogram, revealed the highest genetic distance of 0.3153 between elite inbred lines LN05 and LN50 (Table 3.1). The lowest genetic distance found between elite inbred lines LN02 and LN44 was 0.1609 (Table 3.1).

Table 3.2: Genetic distance frequency between inbred lines using 396 SNP markers.

GD Range	No. of genotypes	%
0.05 - 0.09	115	9
0.10 - 0.19	523	41
0.20 - 0.29	569	44
0.30 - 0.35	68	5

Some of the lowest genetic distances were noticed between the advanced recombinant inbred lines (Table 3.3). Noticeably, LN07 and LN06 had a genetic distance of 0.0722 suggesting that these lines are not that different from each other (Table 3.3). A similar observation was found between LN09 and LN10 with a genetic distance of 0.0736 (Table 3.3). These line combinations will not be ideal for new hybrid make up purposes.

Table 3.3: Inbred lines with the lowest genetic distances

Inbred lines	GD
LN07 X LN06	0.0722
LN10 X LN09	0.0736
LN34 X LN33	0.0802
LN33 X LN32	0.0816
LN35 X LN34	0.0821
LN26 X LN25	0.0850
LN13 X LN12	0.0938
LN09 X LN08	0.0978
LN27 X LN26	0.0985
LN11 X LN10	0.1161
LN08 X LN07	0.1220

The highest genetic distances recorded noticeably included inbred lines LN47, LN49, LN50 and LN51 (Table 3.4), which by location originate from RSA, USA, RSA and ZIM respectively.

Table 3.4: Origin of maize inbred lines with the highest genetic distances

		Origin				Origin	
Inbred lines	**GD**	**P1**	**P2**	**Inbred lines**	**GD**	**P1**	**P2**
LN50 X LN20	0.3000	RSA	UKZN	LN50 X LN21	0.3102	RSA	UKZN
LN47 X LN44	0.3006	RSA	UKZN	LN49 X LN12	0.3104	USA	UKZN
LN50 X LN47	0.3008	RSA	RSA	LN49 X LN32	0.3104	USA	UKZN
LN49 X LN45	0.3009	USA	UKZN	LN51 X LN50	0.3105	ZIM	RSA
LN47 X LN35	0.3016	RSA	UKZN	LN29 X LN01	0.3107	UKZN	ZIM
LN50 X LN33	0.3016	RSA	UKZN	LN50 X LN11	0.3110	RSA	UKZN
LN50 X LN01	0.3019	RSA	ZIM	LN50 X LN19	0.3117	RSA	UKZN
LN50 X LN09	0.3019	RSA	UKZN	LN49 X LN25	0.3135	USA	UKZN
LN49 X LN34	0.3021	USA	UKZN	LN50 X LN07	0.3141	RSA	UKZN
LN47 X LN24	0.3024	RSA	UKZN	LN50 X LN12	0.3144	RSA	UKZN
LN47 X LN28	0.3024	RSA	UKZN	LN50 X LN05	0.3153	RSA	CIMMYT
LN50 X LN16	0.3024	RSA	UKZN	LN50 X LN24	0.3159	RSA	UKZN
LN43 X LN01	0.3026	UKZN	ZIM	LN50 X LN32	0.3162	RSA	UKZN
LN49 X LN26	0.3026	USA	UKZN	LN49 X LN31	0.3173	USA	UKZN
LN49 X LN02	0.3029	USA	UKZN	LN50 X LN10	0.3173	RSA	UKZN
LN05 X LN01	0.3036	CIMMYT	ZIM	LN31 X LN01	0.3175	UKZN	ZIM
LN47 X LN39	0.3041	RSA	UKZN	LN50 X LN15	0.3176	RSA	UKZN
LN50 X LN13	0.3041	RSA	UKZN	LN49 X LN05	0.3187	USA	CIMMYT
LN49 X LN37	0.3045	USA	UKZN	LN49 X LN28	0.3204	USA	UKZN
LN47 X LN29	0.3046	RSA	UKZN	LN50 X LN39	0.3219	RSA	UKZN
LN49 X LN35	0.3056	USA	UKZN	LN49 X LN23	0.3222	USA	UKZN
LN45 X LN43	0.3060	UKZN	UKZN	LN50 X LN29	0.3225	RSA	UKZN
LN47 X LN31	0.3064	RSA	UKZN	LN50 X LN34	0.3232	RSA	UKZN
LN47 X LN23	0.3065	RSA	UKZN	LN50 X LN37	0.3234	RSA	UKZN
LN49 X LN08	0.3066	USA	UKZN	LN49 X LN43	0.3246	USA	UKZN
LN50 X LN25	0.3066	RSA	UKZN	LN50 X LN23	0.3254	RSA	UKZN
LN47 X LN09	0.3070	RSA	UKZN	LN50 X LN49	0.3302	RSA	USA
LN49 X LN29	0.3072	USA	UKZN	LN47 X LN43	0.3399	RSA	UKZN
LN50 X LN26	0.3074	RSA	UKZN	LN50 X LN43	0.3526	RSA	UKZN
LN51 X LN49	0.3080	ZIM	USA				
LN50 X LN30	0.3082	RSA	UKZN				
LN47 X LN32	0.3085	RSA	UKZN				
LN50 X LN35	0.3085	RSA	UKZN				
LN50 X LN18	0.3090	RSA	UKZN				
LN50 X LN22	0.3090	RSA	UKZN				
LN50 X LN08	0.3094	RSA	UKZN				

Genetic distances found between founder parents of the recombinant lines and standard lines ranged between 0.1387 and 0.3526 (Table 3.5). High genetic distances were found between LN43 and LN47 (0.3399), LN49 and LN50 (0.3302) and LN43 and LN50 (0.3526). The higher genetic distance is ideal for the production of hybrids. This is because the higher the genetic distance between lines, the more divergent the genes are and the more there is a possibility for better gene combination.

Diallel matrix between advanced recombinant inbred lines and two founder parents indicated that genetic distance between LN06 and LN15 was 0.0685 while genetic distance between LN16 and LN06 recorded 0.0706 (Table 3.6). The low genetic distances found between lines point toward genetic similarity between these lines. This directly influences combination of these lines for hybrid production purposes.

Table 3.5: Genetic distance between founder parents of the recombinant lines (UKZN) and standard lines from the public programmes

	LN41	LN42	LN43	LN01	LN02	LN03	LN04	LN05	LN47	LN48	LN49	LN50	LN51
LN 41	0.0000												
LN42	0.2519	0.0000											
LN43	0.2968	0.2699	0.0000										
LN01	0.2570	0.2674	0.3026	0.0000									
LN02	0.2627	0.2838	0.2880	0.2572	0.0000								
LN03	0.2398	0.2521	0.2784	0.2504	0.2255	0.0000							
LN04	0.2705	0.2991	0.2910	0.2727	0.2449	0.2007	0.0000						
LN05	0.2552	0.2488	0.2444	0.3036	0.2723	0.2484	0.2555	0.0000					
LN47	0.2624	0.2488	0.3399	0.2962	0.2729	0.2431	0.2741	0.2954	0.0000				
LN48	0.2155	0.2111	0.2540	0.2542	0.2577	0.2012	0.2490	0.2154	0.2398	0.0000			
LN49	0.2870	0.2900	0.3246	0.2822	0.3029	0.1387	0.2526	0.3187	0.2834	0.2677	0.0000		
LN50	0.2827	0.2837	0.3526	0.3019	0.2688	0.2709	0.2784	0.3153	0.3008	0.2527	0.3302	0.0000	
LN51	0.2507	0.2592	0.2798	0.2678	0.2624	0.2313	0.2385	0.2436	0.2769	0.2302	0.3080	0.3105	0.0000

The founder parents of recombinant lines are underlined. The rest are standard lines.

Table 3.6: Diallel matrix of genetic distance between advanced recombinant inbred lines and two founder parents (underlined) clustered in group I and J in the dendogram

	LN06	LN07	LN08	LN09	LN10	LN11	LN12	LN13	LN14	LN15	LN16	LN17	LN18	LN20	LN35	LN40	LN41	LN43
LN06	0.0000																	
LN07	0.0722	0.0000																
LN08	0.1191	0.1220	0.0000															
LN09	0.1054	0.1461	0.0978	0.0000														
LN10	0.0685	0.1170	0.1300	0.0736	0.0000													
LN11	0.1224	0.1040	0.1450	0.1425	0.1161	0.0000												
LN12	0.1134	0.0978	0.1056	0.1201	0.0948	0.1196	0.0000											
LN13	0.1116	0.0899	0.1073	0.1279	0.1014	0.0801	0.0938	0.0000										
LN14	0.0737	0.1096	0.1353	0.1156	0.1064	0.0906	0.1309	0.1280	0.0000									
LN15	0.0649	0.0713	0.1373	0.1527	0.1001	0.1396	0.1123	0.1223	0.1028	0.0000								
LN16	0.0706	0.1196	0.1345	0.1366	0.1383	0.1339	0.1253	0.1510	0.0834	0.1262	0.0000							
LN17	0.0723	0.0953	0.1497	0.1525	0.1309	0.0987	0.1399	0.1034	0.0929	0.0954	0.1112	0.0000						
LN18	0.1118	0.1379	0.1420	0.1120	0.1247	0.1359	0.0881	0.1059	0.1008	0.1231	0.0966	0.1325	0.0000					
LN20	0.1011	0.1120	0.0997	0.1058	0.1449	0.1235	0.1392	0.1118	0.0891	0.1518	0.1184	0.1102	0.1460	0.0000				
LN35	0.1108	0.1636	0.2067	0.1790	0.1480	0.1272	0.1828	0.1787	0.1057	0.1392	0.1409	0.1255	0.1614	0.1925	0.0000			
LN40	0.0911	0.1092	0.1120	0.1083	0.1043	0.0985	0.1047	0.1008	0.0905	0.1089	0.0924	0.0978	0.1021	0.1151	0.1429	0.0000		
LN41	0.1651	0.1314	0.1036	0.1112	0.1478	0.1603	0.1104	0.0911	0.1825	0.1901	0.1696	0.1739	0.1352	0.1243	0.2644	0.1409	0.0000	
LN43	0.0881	0.1286	0.1864	0.1797	0.1253	0.0966	0.1663	0.1702	0.0721	0.0991	0.1221	0.0928	0.1533	0.1493	0.0837	0.1080	0.2968	0.0000

3.3 Discussion

3.3.1 Molecular characterization

3.3.1.1 SNP markers

In this study, Single nucleotide polymorphism (SNP) markers were used effectively to genotype 51 experimental maize inbred lines. The highest PIC value reported in this study is 0.38. This is in line with Hao *et al.* (2011), who reported the PIC range between 0.01 to the highest PIC value of 0.38, using 1536 SNP markers on 95 maize inbred lines. Correspondingly, Yang *et al.* (2011) reported the PIC value range between 0.27 to the highest PIC value of 0.38 with an average PIC value of 0.34 while using 884 SNP markers. The mean PIC in the present study (0.19) was in harmony with the PIC of 0.24 reported by Hao *et al.* (2011). Lu *et al.* (2009) reported a PIC mean of 0.25, using 1034 SNP markers on 770 maize lines, which is again comparable with observations from the current study. According to Abakemal *et al*. (2014) “Polymorphic information content provides an estimate of how informative is a particular marker by taking into consideration both the number of alleles that are expressed, and the relative frequencies of those alleles. PIC values range from 0 which is monomorphic to 1 which is highly discriminative, with many alleles in equal frequencies” (Smith *et al*., 1997).

Genetic diversity ranged from 0.00 to 0.50, with a mean of 0.24, which is expected among the experimental inbred lines in the study, as they have similar gene frequency. Availability data ranged from 0.90 to 1.00, with a mean of 0.98, therefore indicating most markers were successful in genotyping the inbred lines. Heterozygosity ranged from 0.00 to 1.00 in the study. Heterozygosity values were low which is expected, as the inbred lines were advanced, ranging between F_7 and F_8 generations of self-pollination. The inbreeding coefficient ranged from 0.00 to 1.00, with a mean of 0.56. This suggests that some of the experimental inbred lines with a value of 1.00 were already fixed. However, inbreeding coefficient values of 0.00 indicate a non-inbred as a result of genetic contamination due to out-crossing in the nursery.

3.3.2 Cluster analysis of inbred lines based on molecular markers

The cluster analysis clearly discriminated the advanced recombinant lines derived from the two bi-parental populations from the elite inbred lines. Group A to group D comprise of all elite lines. The founder parents are lines LN43 crossed to line LN41 and line LN43 crossed to line LN42, thus the common parent being line LN43. Group F is a composite of the bi-parental crosses which are two populations which share one parent in common. The advanced recombinant inbred lines were clustered under sub-group I and J (Figure 3.7). Under sub-group J, lines LN18, LN12, LN10, LN09,

LN13, LN11, LN20, and LN08 clustered together with LN41 as one of the founder parents. This suggests these inbred lines are more genetically similar to the line LN41 as one of the base parents. Similarly, sub-group I indicated lines LN40, LN16, LN15, LN06, LN07, LN17, LN14 and LN35 clustered with one of the bi-parents LN43. Inbred lines under sub-group G (LN21, LN31, LN23 and LN29) and sub-group H (LN28, LN38, LN37, LN25, LN30, LN39, LN22, LN27, LN32, LN24, LN26, LN33, LN34, LN19 and LN36) are recombinants; therefore the lines did not cluster to any of the founder parents. These represent the new lines which may combine adaptation to the lowland tropics and MSV resistance.

The elite lines LN47, LN01, LN50, LN05, LN42 and LN51 were standing on their own indicating that they are different from the advanced inbred lines which were developed from the UKZN. They would belong to different heterotic groups. The founder parent LN42 was standing on its own and did not cluster with any of its progenies because of selection for standing ability which was applied at F_2 to F_3 generation of the progenies. Additionally, some of its progenies were eliminated due to poor pollination ability and poor standability. However, most of the LN42 progenies were found in group H and I as recombinants. Elite lines LN02 and LN44, LN03 and LN49, LN04 and LN46 were clustered together respectively suggesting close similarity between these lines. This indicated that they may belong to the same heterotic group.

The cluster analysis results proved reliable as the inbred lines were successfully discriminated into two heterotic groups. The advanced inbred lines are in group A in the African tropical germplasm while LN05 is in group B. Therefore, it might be possible to get highly productive hybrids when crossing LN05 to the advanced inbred lines because they are in different heterotic groups.

The low genetic distance between LN18 and LN12 (0.0881), LN09 and LN10 (0.0736) and LN11 and LN13 (0.0801) indicate some level of similarity between the lines. This may suggest the lines share a similar genetic background from the parents and inherited most genes from the base parent shared.

Inbred line LN47 is in Cluster 1 on its own which indicates that it may be used to make productive hybrids in combination with the rest of lines under study. This suggests LN47 is not closely related to any of these lines; therefore the phenomenon of heterosis can be exploited using this line. Similarly, this same trend was observed for LN01, LN05, LN42 and LN51 which originate from ZIM, CIMMYT, UKZN and ZIM respectively. Additionally, the cluster analysis indicated LN03 and LN49 which are of the American origin. However, it was observed that LN45 originating from UKZN was grouped with LN48 from Mozambique, while LN04 from UKZN was grouped with LN46

originating from RSA. This reflected the use of both Mozambique and RSA adapted germplasm in the breeding programme at UKZN.

The results from this study revealed correlation between the genotype clustering based on the dendogram and genetic distance matrix. For example line LN09 and LN10 originating from UKZN are closely clustered together in a dendogram and they are closer to each other with a genetic distance of 0.0736.

3.3.3 Genetic distance

3.3.3.1 Genetic distance between founder parents and standard lines (Table 3.5)

The SNP markers were used to approximate the genetic distance between the three founder parents and standard maize inbred lines. Genetic distance between parents LN41 and LN43 was 0.2968, while parents LN42 and LN43 was 0.2699. Therefore, parents LN41 and LN43 have the highest genetic distance compared to other parents. Overall it is indicated that the three founder parents had different gene frequencies and were genetically distant from the standard lines from the public programmes.

The genetic distance between the elite lines and founder parents varied. The lowest genetic distance of 0.138 was found between the line LN49 and LN03, thus indicating they are closely related to each other. Furthermore, this implies they will not be suitable for hybrid make-up as the cross will result in reduced heterosis. The elite lines LN49 and LN50 had a higher genetic distance of 0.3302 compared to other standard lines. However, the highest genetic distance was found between a founder parent LN43 and elite line LN50. This suggests that if these lines are complementary to each other and if they were crossed, highly productive hybrids could be obtained through hybrid vigour.

3.3.3.2 Diallel of founder parents and advanced recombinant inbred lines (Table 3.6)

The genetic distance between the advanced inbred lines derived from 2 bi-parental populations varied. However, the lowest genetic distance of 0.0649 was observed between LN06 and LN15. This indicates that these lines are genetically closely related and therefore will not be recommended for hybrid production. Additionally, LN06 was closely related to one of the founder parent LN41 with a genetic distance of 0.0881. Moreover, LN15 and founder parent LN43 obtained a low genetic distance of 0.0991, therefore indicating they have less genetic divergence between them. High genetic distance was found between LN35 and LN20. Noticeably, the genetic distance between the founder parents LN41 and LN43 was 0.2968 which is the highest value. This

suggests parents are divergent and vigorous hybrids which could result from the cross can be expected. However, it was observed that parents LN41 and LN42 had higher genetic distance to lines LN34, LN32, LN25 and LN22 compared to parent LN43 indicating they are most suitable to be considered for a hybrid cross. Most of the advanced inbred lines exhibited a genetic distance of ≥0.10 among them indicating that they exhibited similar gene frequencies. This is expected for some of the lines derived from a narrow based bi-parental population.

3.3.3.3 Genetic distance between founder parents, advanced inbred lines and elite inbred lines

Parent LN41 is white grain, resistant to maize streak virus and with a low yielding potential. Low genetic distance was observed between LN41 and lines LN07, LN08, LN09, LN12, LN13, LN18, LN20 and LN40. This indicates these lines are closely related to LN41. Furthermore, as the lines are genetically less dissimilar to founder parent LN41, this suggests that they can adapt in environments where LN41 is well adapted. Parent LN42 is resistant to maize streak virus but susceptible to root lodging. The genetic distance between LN42, elite lines and advanced inbred lines exhibited a high genetic distance between 0.20 to 0.30. Parent LN43 is resistant to downy mildew and is prolific, but susceptible to root lodging. Lines with lower genetic distance between them and LN43 include LN06, LN07, LN11, LN14, LN15, LN16, LN17, LN19, LN22, LN23, LN24, LN25, LN26 and LN27. These lines are closely related to LN43 and their combinations will not be recommended for hybrid production. These lines also belong to the same genetic clusters (Figure 3.7) with less potential to produce superior hybrids. The SNP markers were useful in determining the genetic distances to enable discrimination of closely related genotypes.

3.4 Conclusion

From the current study the following conclusions were drawn:

The SNP markers were useful and were able to discriminate and determine the level of genetic similarity between founder parents. High diversity of 0.30 was found between the three founder parents using SNP markers.

Three hundred and ninety six SNP markers were successful in discriminating the 51 inbred lines according to genetic distances. Additionally, the genotypes were successfully assigned to various genetic clusters accordingly.

Genetic variation among elite inbred lines and advanced inbred lines was broad ranging from 0.1-0.3.Inferences about the heterotic groups of the germplasm lines developed at UKZN could be made, using this set of SNP markers.

References

Abakemal, D., S. Hussein, J. Derera and K. Semagn. 2015. Genetic purity and patterns of relationships among tropical highland adapted quality protein and normal maize inbred lines using microsatellite markers. Euphytica, 204(1): 49-61.

Dao, A., J. Sanou, S.E. Mitchell, V. Gracen and E.Y. Danquah. 2014. Genetic diversity among INERA maize inbred lines with single nucleotide polymorphism (SNP) markers and their relationship with CIMMYT, IITA, and temperate lines. BMC Genetics, 15(127): 1-14.

Foster, J.T., G. J. Allan, A. P. Chan, P. D. Rabinowicz, J. Ravel, P. J. Jackson and P. Keim. 2010. Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis)*. BMC Plant Biology, 10(13): 1-11.

Hao, Z., X. Li, C. Xie, J. Weng, M. Li, D. Zhang, X. Liang, L. Liu, S. Liu and S. Zhang. 2011. Identification of functional genetic variations underlying drought tolerance in maize using SNP markers. Journal of Integrative Plant Biology, 53(8): 641-652.

Laserna, M., C. López, M. Aulicino and G. Maddonni. 2015. Genetic diversity among plants of non-transgenic and transgenic versions of a single cross maize hybrid. Field Crops Research, 176: 56-60.

Legesse, B., A. Myburg, K. Pixley and A. Botha. 2006. Genetic Diversity analysis of CIMMYT-Mid-Altitude maize inbred lines using AFLP markers. South African Journal of Plant and Soil, 23: 49-53.

LGC Genomics. www.lgcgroup.com (accessed 17/02/2015).

Lu, Y., J. Yan, C.T. Guimaraes, S. Taba, Z. Hao, S. Gao, S. Chen, J. Li, S. Zhang, B.S. Vivek and C. Mogorokosho. 2009. Molecular characterization of global maize breeding germplasm based on genome-wide single nucleotide polymorphisms. Theoretical and Applied Genetics, 120(1): 93-115.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences, 70(12), 3321-3323.

Singh, B. D. (1983). Plant Breeding Principles and Methods. Kalyani Publishers.

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

Xu, Y. 2010. Molecular plant breeding .CABI.

Xu, Y., D.J. Skinner, H. Wu, N. Palacios-Rojas, J.L. Araus, J. Yan, S. Gao, M.L. Warburton and J.H. Crouch. 2009. Advances in maize genomics and their value for enhancing genetic gains from breeding. International Journal of Plant Genomics, 2009.

Yang, X., Y. Xu, T. Shah, H. Li, Z. Han, J. Li and J. Yan. 2011. Comparison of SSRs and SNPs in assessment of genetic relatedness in maize. Genetica, 139(8): 1045-1054.

CHAPTER FOUR

GGE- Biplot Analysis and genetic gain of maize hybrids developed in the program at UKZN

Abstract

Grain yield is a quantitative trait in maize which is affected by genotype x environment (GXE) effects. This complicates selection of superior hybrids because different hybrids behave differently in various locations. The GXE interaction effects of experimental maize hybrids which were developed by the breeding programme at the University of KwaZulu-Natal (UKZN) has not been determined. Therefore, fourty-two maize hybrids were evaluated, across six locations during the 2013 seasons representative of the maize growing environments of South Africa. These trial entries were laid out in a randomized complete block design with two replications. The trials were conducted under dry land conditions. The additive main effects and multiplicative interaction (AMMI) analysis and GGE-biplot analyses were conducted to determine the effects of hybrid x environment interaction on grain yield. There were highly significant differences among genotypes, environments and their interactions, accounting for 25%, 26% and 29% of the total variance, respectively. The AMMI-2 model explained 68% of the GXE interaction. The GGE biplot revealed that hybrid 14XH082 was the most stable across South Africa. AMMI model revealed that hybrid 11C3201, 13C7082 and 10HDTX11 were the most stable when compared to commercial hybrids PAN6Q445B, PAN6611 and DKC80-40BRGEN. This represents significant progress in availability of stable and well adapted new hybrids for the South African market.

4.1 Introduction

Hybrid performance in relation to traits such as yield is influenced by genotype, environment and genotype by environment interaction (Tiwari *et al*., 2014). Large amounts of yield variation are a result of change in environmental stresses such as climatic, soil conditions and disease prevalence (Asfaw *et al*., 2009; Grada and Ciulca, 2013). For this reason, genotype by environment interaction has widely become a crucial factor for breeding programmes (Adu *et al*., 2013). This is because large GXE interaction impairs accuracy of yield estimation due to varying response patterns among the hybrids across environments. The response of genotypes to the effect of GXE directly impacts on the ranking of hybrids (Mitrović *et al*., 2012). The choice of hybrids for any given environment is hindered by cross interactions (Mitrović *et al*., 2012). These cross interactions account for the different hybrid ranking from one environment to another, making it more challenging to draw reliable conclusions (Stojaković *et al.*,2012). This means one hybrid can be high yielding and rank high in one environment and low yielding and rank lower in another. Moreover, the breeding progress is affected as the GXE effect complicates the demonstration of superior genotypes across multiple locations (Rashidi *et al.*, 2013). Selection of experimental hybrid genotypes is based on the assessment of their phenotypic value in a number of locations or environments (İlker *et al.*, 2009). Additionally, it reduces the relationship between the phenotypic and genotypic values (Grada and Ciulca, 2013), thereby reducing progress from selection. Therefore, it has a direct impact on the adoption of the hybrid, its productivity and total production of the crop (Lule *et al.*, 2014).

Plant breeders are more interested in selecting hybrids which perform consistently superior across most testing environments therefore GXE effects are a crucial factor in hybrid development. According to Kandus *et al*. (2010), genotypes can be classified in two ways in accordance with their behavior, either as stable genotype or adapted to a particular environment. Hybrid adaptability refers to consistent high performance of the genotype across diverse sets of environments due to adjustment to its environment (Farshadfar *et al*., 2013). Genotype adaptability is evident where a genotype yields high in specific environmental conditions and produces poor yields in another set of environment (Kandus *et al*., 2010, Rashidi *et al*., 2013). According to Lule *et al*. (2014), adaptability is the function of genotype, environment and genotype by environment interaction. Lule *et al.* (2014) cited that adaptability generally falls into two classes: (i) the ability to perform at an acceptable level in a range of environments, known as general adaptability, and (ii) the ability to perform well only in desirable environments known as specific adaptability. On the

other hand, static stability refers to genotype with a capacity to remain constant, both with high or low yield levels in various environments. However breeders desire to select hybrids with dynamic stability, that is the product which responds to improvements in resources or management of the crop in the farmers field. This is the type of stability which is desirable.

Given the foregoing assessment of hybrid performance, adaptability and stability across varying environments and prior to commercial release is paramount in maize hybrid development. Assessment of hybrids across environments generates crucial and valuable information for identifying adaptation and stability of these hybrids (Crossa, 1990). Most importantly, the same information allows maize breeders to make specific selections of hybrids for particular locations and for seed companies to make appropriate recommendations to maize growers (İlker *et al.*, 2009). Tiwari *et al.* (2014) cited that the cause of yield stability is unclear. However, mechanisms that may confer stability of maize hybrids may include genetic heterogeneity, yield component compensation, stress tolerance and capacity to rapidly recover from stress. According to Basford and Cooper (1998), GXE effects influences contribution of genes governing grain yield resulting in different performance of hybrids in different environments. This makes it mandatory for maize breeders to conduct multi- environmental trials (MET) which present the largest cost in maize breeding research.

There are various statistical tools and methods used to quantify and graphically represent various levels of GXE effects in maize hybrid development. However, no single statistical method can sufficiently describe maize hybrid performance across environments (Rashidi *et al.*, 2013). These methods can be classified into two groups namely univariate and multivariate. The multivariate model includes Principal Component Analysis (PCA), cluster analysis and Additive Main Effects and Multiplicative Interaction Models (AMMI). According to Crossa *et al.* (1990), the purpose of multivariate methods is to eliminate “noise” found in the data set. This enables separation of systematic and non-systematic variation. Multivariate approaches facilitate summarization of the information and easily reveal data structure.

Among them is the AMMI model which considers the effect of genotype and environment as the additive effects plus the genotype x environment interaction effects as the multiplicative component and submits it to principal component analysis. The AMMI models are usually called AMMI-1, AMMI-2 to AMMI- n, depending on the number of principal components used to study the interaction (Kandus *et al.*, 2010).

The modified AMMI model is the GGE biplot analysis this combines the genotype effects (G) with genotype x environment (GE) multiplicative effect. These effects are submitted to the principal component analysis (Adu *et al.*, 2013). The graphical presentation produced by the GGE biplot is very appealing for plant breeding and has become a crucial tool in product evaluation and advancement. In the same GGE graph, it is possible to observe the following: (i) genotype (points) and environment (vectors); (ii) the exploration of patterns attributable to the effects of GXE interaction. The specific interaction between the genotypes and the environments is clearly visualized through the distances from the origin of the graph. According to Yan and Tinker (2006), high performance of a genotype in an environment when compared to average is realized if the angle between its vector and the environment vector is <90º, it is poorer than average if the angle is >90º, and it is near average if the angle is about 90º.

As a result the GGE-biplot is now widely used by maize breeding programmes. There are more features encompassed in the GGE biplot analysis which attract plant breeders and scientist. For example, recently Badu-Apraku *et al.*(2012) and previously Yan and Tinker (2006), cited the following functions as the most appealing: discrimination and ranking of hybrids, head-to-head analyses, differentiation and therefore selection of suitable environments for METs, and ultimately the capability of the tool to select winning hybrids in each environment :

The objective of this study was to determine the GXE effects among experimental hybrids to see if progress has been made by the programme at the UKZN, by identifying and obtaining hybrids which are adapted to environments in South Africa. The AMMI and GGE-biplot analysis tools were therefore applied.

4.2 Materials and Methods

4.2.1 Germplasm

Fourty-two hybrids listed in Table 4.1(b), were evaluated in summer of 2013- 2014 season across six locations. Thirteen of the hybrids were commercial hybrids used as positive control (standard checks). The experimental single cross hybrids originated from the UKZN breeding programme. These were developed as crosses between tropical adapted lines with South African adapted lines.

4.2.2 Experimental design, Locations and field management

The experimental design was a 6x7 alpha lattice design with two replications across environments. The row length for each plot was 6.6 m. The row width was 0.90 m for sites in the east and 1.5 m

for the western sites, which is consistent with the hybrid production culture in South Africa. The higher row width in the west is aimed to reduce competition and compensate for soil moisture due to drier climate conditions in the west. Therefore, the western sites have low planting density of 28 000 plants per hectare. However the eastern sites have high planting density of 44 000 plants per hectare (Table 4.1 (a)). The planting was done at one row plot for each hybrid with 22 plant stations per plot.

Six locations listed in Table 4.1(a) were used as testing environments. The testing environments were based in three provinces which are KwaZulu–Natal (KZN), Mpumalanga and North West provinces which are among the maize production areas in the country. . The trial sites were prepared and planted according to each farmer's practices. Fertilizer application and field management were done according to the requirements of each farmer. Pre-emergent and post emergent weeds were controlled using Basagran and roundup herbicide. Pesticides were used to control stalk borer.

Table 4.1 (a): Experimental locations (environments) used in the study

Loc #	Loc Name	Latitude	Longitude	Province	Density	Long term Average annual rainfall (mm)
E1	BETHAL	26.4579° S	29.4667° E	Mpumalanga	44000	710
E2	CAROLINA	26.0731° S	30.1070° E	Mpumalanga	44000	614
E3	CEDARA	29.5478º S	30.2667º E	KwaZulu-Natal	44000	900
E4	POTCHEFSTROOM	26.7145° S	27.0970° E	North West	28000	615
E5	WINTERTON	28.8166° S	29.5296° E	KwaZulu-Natal	44000	789
E6	MOOI RIVER	29.2106° S	30.0030° E	KwaZulu-Natal	44000	900

Table 4.1(b): List of genotypes used in the study

Hybrid code	Genotype no.	Developing sector	Hybrid code	Genotype no.	Developing sector
03C475	1	Experimental hybrid	11C1511	22	Experimental hybrid
11C3417	2	Experimental hybrid	11C2242	23	Experimental hybrid
11C3201	3	Experimental hybrid	11C1483	24	Experimental hybrid
11C2974	4	Experimental hybrid	11C2243	25	Experimental hybrid
02C3156	5	Experimental hybrid	1OHDTX11	26	Experimental hybrid
11C6363	6	Experimental hybrid	14XH050	27	Experimental hybrid
11C2557	7	Experimental hybrid	14XH146	28	Experimental hybrid
13C7060	8	Experimental hybrid	14XH149	29	Experimental hybrid
13C7065	9	Experimental hybrid	14XH082	30	Experimental hybrid
13C7071	10	Experimental hybrid	14XH065	31	Experimental hybrid
13C7082	11	Experimental hybrid	PAN6Q445B	32	Commercial Check hybrid
13C7083	12	Experimental hybrid	PAN6611	33	Commercial Check hybrid
13C7109	13	Experimental hybrid	PAN6Q308B	34	Commercial Check hybrid
13C7110	14	Experimental hybrid	DKC78-45BRGEN	35	Commercial Check hybrid
13C7119	15	Experimental hybrid	DKC80-40BRGEN	36	Commercial Check hybrid
13C7122	16	Experimental hybrid	PAN53	37	Commercial Check hybrid
11C1774	17	Experimental hybrid	PAN67	38	Commercial Check hybrid
11C1579	18	Experimental hybrid	SC633	39	Commercial Check hybrid
11C1566	19	Experimental hybrid	SC506	40	Commercial Check hybrid
11C2245	20	Experimental hybrid	SC301	41	Commercial Check hybrid
11C1350	21	Experimental hybrid	SC403	42	Commercial Check hybrid

4.2.2 Statistical analysis

A combined analysis of variance (ANOVA) was done for the genotypes across locations using GENSTAT statistical software version 14.1 (GENSTAT, 2011). The analysis was based on the mean data obtained from each location. The GGE biplot analysis for genotypes against environments was conducted to ascertain performance of hybrids across various test environments. The biplot was conducted to explain the genotype x environment interaction, adaptation and which-won-where pattern of hybrids towards the environments. The AMMI-2 model was used to plot IPCA scores against the mean yield of genotypes.

4.3 Results

4.3.1 Genotype x environment interaction

The genotypes and environments as main effects were highly significantly different from each other and their interactions were also highly significant (Table 4.2). The IPCA1 and IPCA2 were highly significant (p >0.01) and contributed 9.55% and 8.14% respectively to the total GXE interaction (Table 4.2).

Table 4.2: AMMI-2 model ANOVA for grain yield in maize genotypes across environments

Source	df	SS	MS	F	F_prob	Variance explained (%)
Total	503	1852.2	3.68	*	*	
Treatments	251	1483.6	5.91	4.22	0.00000**	80.09
Genotypes	41	457.6	11.16	7.97	0.00000**	24.7
Environments	5	489.6	97.92	24.32	0.00000**	26.43
Block	6	24.2	4.03	2.88	0.00999*	1.3
Interactions (GXE)	205	536.4	2.62	1.87	0.00000**	28.96
IPCA 1	45	177	3.93	2.81	0.00000**	9.55
IPCA 2	43	150.8	3.51	2.51	0.00001**	8.14
Residuals	117	208.6	1.78	1.27	0.05944	11.26
Error	246	344.4	1.4	*	*	18.6

* , ** Significant difference at 0.05 and 0.01 probability level, respectively.

Environment means for Cedara were the lowest with a score of 6.404 (Table 4.3). Cedara is a low yielding site established in Natal midlands along the mist belt. Therefore it is a high disease pressure site with which no spraying was done for disease control. Winterton was the highest yielding site with means of 9.555 followed by Mooi River. Winterton site was under irrigation and spraying for disease control therefore contributing to resultant yield.

Potchefstroom and Mooi River locations had negative IPCA1 scores suggesting the magnitude of interaction with hybrids were more similar (Table 4.3).

Table 4.3: Environment scores sorted by means

Environment	NE	Em	IPCAe[1]	IPCAe[2]
Cedara	3	6.404	0.24321	-0.25185
Carolina	2	7.016	2.18493	0.61992
Bethal	1	7.101	0.13815	0.18887
Potchefstroom	4	7.523	-1.98831	1.58567
Mooi River	6	7.632	-0.75531	-2.3733
Winterton	5	9.555	0.17732	0.2307

IPCA scores for hybrids were determined (Table 4.4). A higher IPCA1 score suggests a high interaction of a hybrid with the environment of similar IPCA1 scores. Hybrids with similar IPCA1 scores are specifically adapted to similar environments. Ideally, hybrids with IPCA scores closer to zero are more desirable as a measure of stability. High IPCA1 scores were found for the hybrids 14XH065, 14XH146, PAN6Q445B, PAN67 and 11C2557 (Table 4.4). On the contrary hybrids 11C2243, 14XH082, 11C2242, PAN6Q308B, SC403 and 13C7110 had IPCA1 scores closest to zero indicating these hybrids are more adapted and stable.

AMMI results indicated genotype means as predicted yields ranged from 4.23 to 10.405 tonnes per hectare (Table 4.4). Among the top five hybrids with high predicted means, there are two experimental hybrids 11C3201 and 13C7082. This means hybrid 11C3201 is the second best hybrid after PAN6Q445B, and it outperformed PAN611 and DKC80-40BRGEN commercial hybrids. Six experimental hybrids appear in the top ten and these were superior to commercial hybrids such as DKC78-45BRGEN, SC403, PAN53, PAN67, SC633, SC301 and PAN6Q308B which did not appear in the top ten performing hybrids.

Table 4.4: IPCA1 and IPCA2 scores for the 42 hybrids sorted by mean yield and evaluated at 6 locations

Genotype	NG	Gm	IPCAg[1]	IPCAg[2]	Genotype	NG	Gm	IPCAg[1]	IPCAg[2]
14XH065	27	4.23	-0.99567	-0.76164	11C1511	5	7.611	0.39628	-0.30395
PAN6Q308B	37	5.65	-0.07497	0.82104	13C7060	17	7.614	-0.36516	-0.48574
14XH050	26	6.07	-0.64029	-0.08905	03C475	2	7.618	0.72474	-0.42532
SC301	39	6.345	0.62309	-0.42463	11C1483	4	7.646	0.44992	0.11368
SC633	42	6.412	0.14996	-0.80558	11C2974	13	7.728	0.20476	0.00578
13C7119	24	6.588	0.54541	0.03812	1OHDTX11	31	7.735	-0.60419	0.44806
PAN67	36	6.938	0.66204	-0.06684	SC403	40	7.756	0.04827	-1.01821
11C2242	9	6.962	-0.05197	-0.15389	13C7065	18	7.861	0.41202	0.50882
11C1350	3	7.006	-0.24476	0.284	13C7122	25	7.897	-0.23757	-0.13444
14XH146	29	7.019	-0.79676	-0.63822	DKC78-45BRGEN	32	7.929	0.33586	0.55029
11C2557	12	7.192	0.70333	0.05444	13C7110	23	8.005	0.04228	-0.3313
14XH149	30	7.363	-1.14265	-0.55208	SC506	41	8.091	-0.155	0.00091
11C1579	7	7.404	0.10294	-0.22023	11C1774	8	8.105	0.11822	0.42012
14XH082	28	7.419	0.01504	0.15171	11C2245	11	8.152	-0.54357	-0.08996
11C1566	6	7.457	-0.45375	0.47135	11C3417	15	8.169	0.52327	0.20658
11C2243	10	7.476	0.01325	-0.0794	02C3156	1	8.285	0.31846	0.23565
13C7109	22	7.504	0.38626	-0.01024	13C7082	20	8.42	0.34709	0.00812
13C7083	21	7.522	-0.10417	-0.15291	DKC80-40BRGEN	33	8.642	-0.4761	0.82784
13C7071	19	7.53	0.57452	0.24448	PAN6611	35	8.815	0.17302	0.1538
11C6363	16	7.56	0.2802	-0.59502	11C3201	14	8.907	-0.3436	0.77468
PAN53	34	7.582	-0.16381	0.0254	PAN6Q445B	38	10.405	-0.75624	0.99379

AMMI-2 model revealed four hybrid selections per environment based on mean yield (Table 4.5). Results revealed that hybrids 14XH149 and 14XH146 were specifically adapted to Mooi River location only. Experimental hybrids 11C3201 together with commercial hybrid PAN6Q445B and PAN6611 were selected as best hybrids across four locations namely Cedara, Winterton, Bethal and Potchefstroom. This suggests that hybrid 11C3201 has a high potential to compete with the commercial hybrids in terms of high stability and high productivity.

Table 4.5: AMMI-2 model's best 4 hybrid selections for mean yield in relation to environments evaluated

Environment	Mean	Score	Hybrid 1	Hybrid 2	Hybrid 3	Hybrid 4
Carolina	7.016	2.1849	11C3417	PAN6Q445B	PAN6611	13C7082
Cedara	6.404	0.2432	PAN6Q445B	PAN6611	11C3201	13C7082
Winterton	9.555	0.1773	PAN6Q445B	11C3201	PAN6611	DKC80-40BRGEN
Bethal	7.101	0.1381	PAN6Q445B	11C3201	PAN6611	DKC80-40BRGEN
Mooi River	7.632	-0.7553	SC403	14XH149	14XH146	13C7060
Potchefstroom	7.523	-1.9883	PAN6Q445B	DKC80-40BRGEN	11C3201	10HDTX11

Results indicated that the lowest yielding environments were Cedara, Carolina and Bethal (Table 4.6). Twenty best genotypes were also ranked in these three locations. In Bethal, the twenty best genotypes' yield ranged from 8.599 to 7.230 t/ha, while in Cedara, best performing genotypes' yield range was 8.518 to 6.455 t/ha

Table 4.6: Low to moderate yielding environments based on mean yield with ranked 20 best genotypes

Bethal Carolina Cedara

Genotypes	YLD (t/ha)	Genotypes	YLD (t/ha)	Genotypes	YLD (t/ha)
DKC80-40BRGEN	8.599	PAN6611	9.182	PAN6Q445B	8.518
PAN6Q445B	8.496	11C3417	9.137	02C3156	7.688
PAN6611	8.233	13C7082	8.913	PAN53	7.676
DKC78-45BRGEN	8.207	03C475	8.694	13C7071	7.671
10HDTX11	7.874	02C3156	8.569	13C7082	7.466
11C1483	7.781	11C3201	8.489	11C3201	7.349
11C1511	7.766	PAN6Q445B	8.446	11C3417	7.158
11C2974	7.703	11C2557	8.416	PAN67	7.040
13C7110	7.694	SC506	8.158	SC506	7.002
PAN53	7.558	DKC78-45BRGEN	8.116	DKC80-40BRGEN	6.971
13C7060	7.457	11C1483	8.061	14XH149	6.944
13C7122	7.455	13C7065	8.061	SC301	6.871
11C1774	7.429	13C7071	7.908	11C1774	6.780
11C2243	7.422	11C1774	7.860	11C6363	6.756
13C7083	7.420	13C7109	7.518	11C1566	6.705
11C3417	7.309	SC403	7.497	11C2974	6.619
PAN67	7.308	11C1511	7.311	14XH082	6.520
SC403	7.250	13C7110	7.213	13C7109	6.502
14XH082	7.245	11C2243	7.125	13C7065	6.497
13C7065	7.230	11C2245	7.123	11C2557	6.455

Table 4.7 depicts twenty best hybrids in three high yielding environments namely Potchefstroom, Winterton and Mooi River. In Potchefstroom, the yield ranged from 7.534 to 11.723 t/ha with PAN6Q445B commercial hybrid identified as the highest yielding hybrid followed by 11C3201 and DKC80-40BRGEN. However, in Winterton and Mooi River, yield ranged from 9.64 to 11.46 t/ha and 7.666 to 8.840 t/ha respectively. In Winterton, the top three high yielding hybrids were PAN6Q445B, 11C1579 and 13C7065, while in Mooi River the three high yielding hybrids were SC403 followed by 11C2245 and 14XH149.

Table 4.7: High yielding environments based on mean yield with ranked 20 best genotypes

Potchefstroom		Winterton		Mooi River	
Genotypes	Yield (t/ha)	Genotypes	Yield (t/ha)	Genotypes	Yield (t/ha)
PAN6Q445B	11.723	PAN6Q445B	11.46	SC403	8.840
11C3201	10.223	11C1579	10.44	11C2245	8.324
DKC80-40BRGEN	9.596	13C7065	10.42	14XH149	8.304
11C2245	8.908	11C1566	10.26	14XH146	8.188
1OHDTX11	8.898	11C2974	10.26	PAN6611	8.179
SC506	8.761	11C2245	10.21	13C7060	8.120
PAN6611	8.705	11C3201	10.19	SC506	8.112
11C1566	8.638	13C7071	10.17	13C7110	8.086
14XH149	8.406	DKC80-40BRGEN	10.16	PAN6Q445B	8.011
11C1774	8.208	13C7122	10.11	03C475	8.007
02C3156	8.091	11C1774	10.05	11C6363	7.988
13C7082	8.050	11C6363	10.03	13C7082	7.981
13C7122	7.847	13C7083	9.98	13C7122	7.963
11C3417	7.840	13C7109	9.92	SC633	7.776
DKC78-45BRGEN	7.819	13C7082	9.87	13C7083	7.756
11C1350	7.775	13C7110	9.84	11C1511	7.755
13C7065	7.609	13C7060	9.77	11C2243	7.751
PAN53	7.568	DKC78-45BRGEN	9.71	11C3201	7.698
14XH146	7.542	11C1511	9.68	11C2242	7.686
14XH082	7.534	PAN6611	9.64	11C3417	7.666

GGE biplot was employed to determine and visualize which hybrid won in which environment (Figure 4.1). Vertex hybrids were SC301 (hybrid 41), 03C475 (hybrid 1), 11C3417 (hybrid 2), PAN6Q445B (hybrid 32), 14XH149 (hybrid 29), 14XH065 (hybrid 31). Hybrid 11C3417 (hybrid 2), PAN6Q445B (hybrid 32) were similarly nominated as best hybrids in the AMMI model of best hybrid selection in Carolina and Potchefstroom respectively (Table 4.5). Hybrid 14XH149 (hybrid 29) was nominated as second hybrid in AMMI best four selections in Mooi river (Table 4.5). However, vertex hybrids SC301 (hybrid 41), 03C475 (hybrid 1) and 14XH065 (hybrid 31) were not found in the AMMI model four hybrid selections across environments.

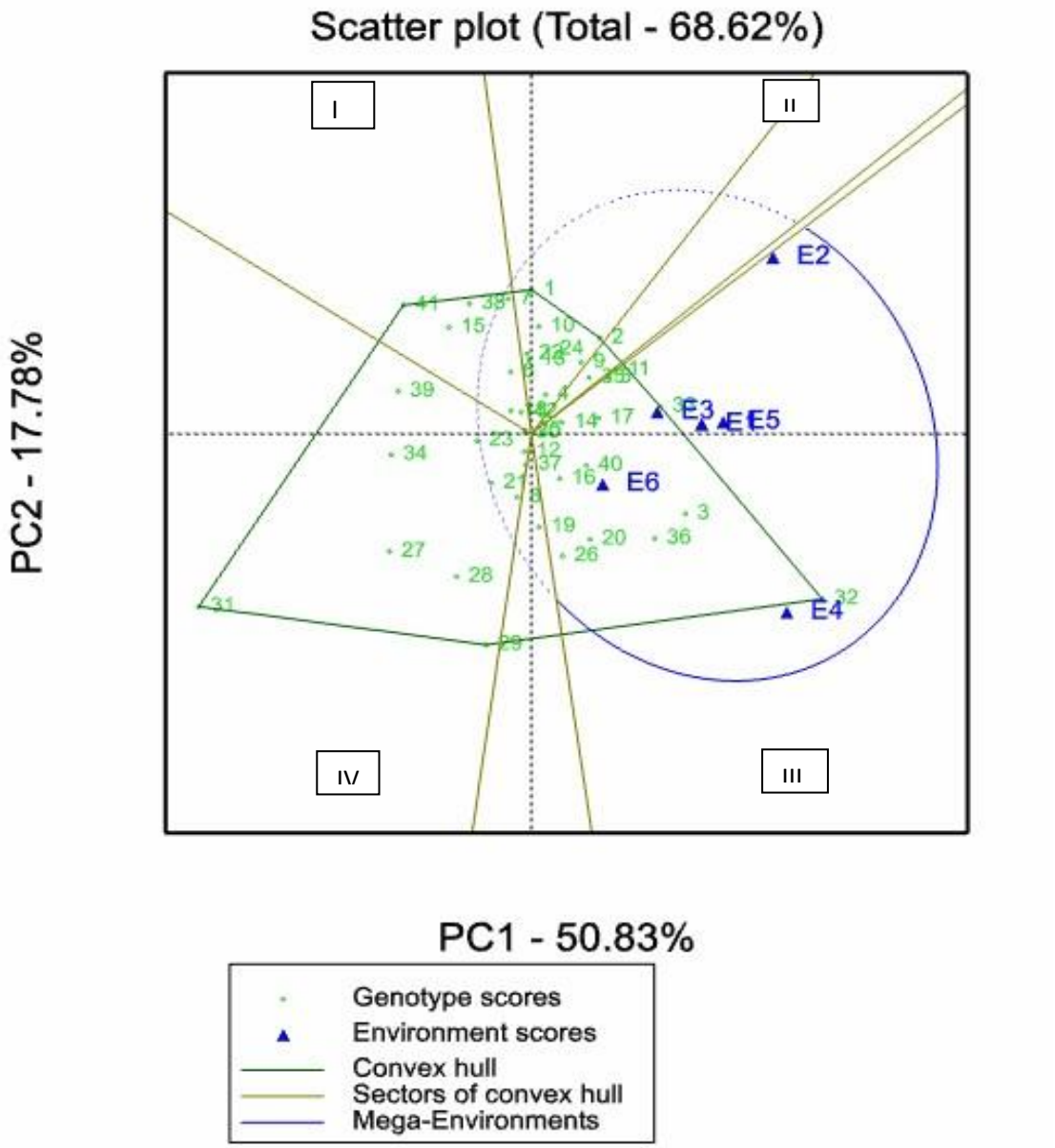

Figure 4.1: GGE biplot for grain yield indicating the which-won-where analysis of genotypes[6]

[6] Genotype hybrid names are listed in Table 4.1(b).

4.4 Discussion

4.4.1 Genotype and environment interaction

Analysis of variance showed that grain yield was influenced by testing environments, genotypes and their interactions (Table 4.2) which was expected for grain yield data given that the test locations represented mega-environments in South Africa. Results indicated fourty-two genotypes tested in six locations and showed that 80.09% of the total sum of squares was attributable to treatments (GEI) effects, 26.43% to environment effect and 24.70% to genotypic effects (Table 4.2), indicating the high precision of the experiment. In wheat, Verma *et al*. (2015) reported results on GEI effects accounting for 89.56% of the total sum of squares, 3.18 % for genotypes and 68.81% for environments.

The first principal component analysis (IPCA1) of the AMMI analysis conducted accounted for 9.55% of the total sum of squares, while the second IPCA2 accounted for 8.14% of the sum of squares and both were significant while the residual was not significant. Hence the AMMI-2 model was adopted for the data. The sum of squares for the genotype by environment interaction was higher than of environments and genotypes indicating that GXE was indeed very important. This suggests that there were substantial differences in environmental response towards genotypes. Means of sum of squares revealed genotypes behaved differently towards adaptation to various environments. When testing stability for grain yield in durum wheat, Mohammadi *et al*. (2015) found high GEI effects (11%) accounting for total sum of squares, which were greater than the genotype effects. They reported that the magnitude of the GEI sum of squares were two times larger than that for genotypes. They argued that in mega environmental trials there can be a mixture of crossover and non-crossover types of GEI which cause more dissimilarity in the genetic systems controlling physiological processes that confer yield stability in different environments. Therefore, genotypes may be selected for adaptation to specific environments. Additionally, there was high level of genetic diversity among genotypes as indicated by differences among the genotypic means, accounting for the variation in grain yield. There were highly significant differences as revealed by the ANOVA for genotypes as main effects, again signifying diversity within hybrids. Therefore, this provides an opportunity for selection of suitable hybrids for the different environments.

About 57% of the hybrids IPCA1 scores indicated positive interaction effects (Table 4.3). Amongst many, these included SC633, SC301, 13C7119, PAN67, 11C1579, 14XH082, 11C2243, 11C6363, 11C1774, 13C7110, SC403, 11C2974, 11C1483, 03C475 and 11C1511 indicating their alignment

towards locations with positive IPCAs. This is similar to the findings reported by Verma *et al*., (2015) concerning positive and negative IPCA scores for both genotypes and environments. They found that two of the environments had high positive IPCA scores. The environments displayed positive interaction with about 8 genotypes which had positive IPCA scores. The remainder of 43% hybrids including 11C2242, PAN53, SC506, DKC80-40BRGEN and 13C7083 had negative interaction effects, indicating that they were inclined towards locations with negative IPCAs. Therefore the study supports observation of specific adaptation of certain hybrids to particular environments which were represented by these locations. However, hybrid 14XH082, 11C2243, 13C7083, PAN53, SC403, 13C7110 and SC506 had IPCA1 scores closer to zero, indicating that there were also hybrids that exhibited general adaptation to many locations. Therefore, this indicates small interaction effects meaning the hybrids are stable and well adapted as less influenced by the environments. These results are similar to Ilker *et al*., (2009) report, who found a genotype with smaller absolute IPCA score closer to zero to be less responsive compared to other genotypes which were further from the plot of origin. The genotype was then considered as most stable. Similar findings in wheat were reported by Ali *et al.,* (2015)

The IPCA scores of a genotype in the AMMI analysis indicate the stability or adaptation over environments. The higher the IPCA score (either positive or negative), the more the genotype becomes specific adapted to certain environments with similar IPCAs. If a genotype or an environment has an IPCA score closer to zero, it has a small interaction effects and considered as stable. This means the genotype is stable or adapted across all the locations. When both a genotype and environment possess an equivalent sign on the PCA axis, their interaction effects is positive, however, if different, their interaction effects is negative (Akter *et al.*, 2014). Table 4.3 and 4.4 presents the AMMI analysis data with the IPCA1 and IPCA2 scores for the genotypes and environments respectively. The positive main effects were recorded by the environments E2, E1, E4 and E5 with a mean of 7.016, 7.101, 7.523 and 9.555 respectively. However, environment E3 and E6 had negative specific main effects with mean of 6.404 and 7.632 respectively.

4.4.2 AMMI model best four hybrid selections

Table 4.5 presents the results of the best four hybrids selected by the AMMI-2 model. These hybrids selections are based on mean yield and indicate the best adapted hybrids in relation to different environments. From the results, hybrid PAN6Q445B was the best and most adapted across five environments (E1, E2, E3, E4 and E5). The other hybrids that were adapted across other environments were PAN611 (E1, E2, E3 and E5), 11C3201 (E1, E3, E4 and E5) and

DKC80-40BRGEN (E1, E4 and E5). The highest environment mean of 9.555 was for environment E5 and the lowest mean of 6.404 was for E3.

The AMMI model of the four best hybrid selection revealed that environment E5 was the highest yielding environment. However, environment E3 was the lowest yielding of all the environments. Hybrid PAN6Q445B had the highest yield in environment E5; however, the hybrid seems to be well adapted in high yielding environments. Nonetheless, the same hybrid performed well above average in environment E3 which is the low yielding environment. Interestingly environment 6 top four selections were hybrid SC403, 14XH149, 14XH146 which were never selected in other environments, therefore revealing no definite pattern towards other environments. This is suggestive of specific adaptation of these hybrids only to environment E6.

4.4.3 Winning genotypes and environments

The biplot analysis was conducted using genotypic and environmental scores of the first two AMMI components explaining 68.62% of the GEI variation (Figure 4.1). Amongst various features of the GGE biplot, the greater one is its ability to display the which-won-where pattern of genotype by environment dataset (Ilker *et al*., 2009; Solonechnyi *et al*., 2015). The “which-won –where” graph is constructed by joining genotypes farthest from the line of origin, therefore forming a polygon (Verma *et al*., 2015). Perpendicular lines are drawn from the centre origin of the biplot to each side of the polygon, forming several sectors with one genotype at the vertex of the polygon (Yan and Tinker, 2006). The lines are known as equality lines (Rakshit *et al.*, 2012). Genotypes found at the vertices of the polygon are either the best or poorer in one or more environments. The results indicated hybrid 03C475 (1), SC301 (41), 14XH065 (31), 14XH149 (29), PAN6Q445B (32) and 11C3417 (2) were on the vertices of the biplot polygon. This suggests these genotypes were the furthest from the biplot origin line and could be either good or poorer performers in their respective environments. Therefore, hybrid 03C475 11C3417, 13C7071, 11C2557, 11C1483 in the second quadrant had good yields, however, less stable. Genotypes located near the plot origin were less responsive than the genotypes farther from the centre of the biplot. However, hybrid 14XH082 (30) was found on the line of origin which indicates stability across environments. Hybrid 11C3201 (3), 02C3156 (5), 13C7082 (11), 13C7109 (13), 13C7122 (16), 11C1774 (17), 11C2245 (20), 11C1483 (24), 10HDTX11 (26), DKC80-40BRGEN (36) were better adapted in environment E1, E2, E3, E5 and E6. However, hybrid PAN6Q445B (32) was mostly and specifically adapted in environment E4, although still performed well in other environments.

The equality lines effectively divided the biplot into seven sectors. All six environments in which the hybrids were tested are located in one large sector in the biplot. This indicates the environments

are partitioned in one mega-environment. Tonk *et al*., (2011) reported that seven locations in the study were divided into 2 sectors, forming two mega environments. The first mega environment consists of five environments while two environments form part of second mega environment. Most hybrids such as 13C7119 (15), 14XH050 (27), 14XH146 (28), 14XH149 (29), 14XH065 (31), PAN6Q308B (34) and SC633 (39) did not fall in any of the environment sectors, indicating that there was no environment in which they produced the highest grain yield. This is in line with the results reported by Munawar *et al.* (2013). Environment E4 and E2 was the best discriminating environments. Interestingly in environment E2, there was no genotype that was specifically adapted to the environment.

4.5 Conclusion

In this study, the following conclusions can be established:

- There were highly significant differences for the genotype by environment interaction which suggests the importance of testing the genotypes in various environments. There is an opportunity for selection of the genotypes due to the high significant differences within genotypes indicating existence of genetic diversity. Therefore, genotypes and environments did influence the total grain yield output.

- All the six environments were partitioned into one mega environment. However, the environments were divided into two sectors which are high and low yielding environments. This provides an opportunity for selection of hybrids in the two environments and will allow the breeders to make recommendations on the hybrid material and test locations. Environment 5 was the highest yielding environment with the highest mean yield.

- The GGE biplot was effective in identifying the most responsive hybrids in each location. It was also effective in discriminating hybrids that did not perform well in any of the locations. Hybrid PAN6Q445B was mostly responsive in environment E4. None of the locations fell in the sector with hybrids 14XH149, 14XH065, SC301 and 03C475, suggesting these hybrids were not performing best in any of the locations.

- GGE biplot analysis was effective in revealing hybrids with the most adaptation across environments with regards to minor effects to total grain yield. Therefore the which- won-where patterns were identified successfully.

- AMMI analysis employment was successful in revealing experimental hybrids appearing among best 4 hybrid selections in each environment. In Carolina, experimental hybrids

11C3417 and 13C7082 were selected alongside PAN6Q445B and PAN6611 commercial hybrids. In Cedara, experimental hybrids 11C3201 and 13C7082 were selected alongside PAN6Q445B and PAN6611. Winterton and Bethal had 11C3201 as the only experimental hybrid selected alongside three commercial hybrids PAN6Q445B, PAN6611 and DKC80-40BRGEN. Mooi River had three experimental hybrids 14XH149, 14XH146 and 13C7060 alongside commercial hybrid SC403 as the best four hybrids selected. Potchefstroom had 11C3201 and 10HDTX11 experimental hybrids with PAN6Q445B and DKC80-40BRGEN commercial hybrids as the best performing hybrids.

- Experimental hybrid 11C3201 emerged four times as the best hybrid selection in four locations which were Cedara, Winterton, Bethal and Potchefstroom. This is suggestive of a high potential experimental hybrid with high stability and productivity. The results indicate the high potential of this hybrid to compete with commercial hybrids which are in the market therefore warrants the experimental hybrid as a candidate for advancement.

Therefore the study was successful in determining the GXE effects among experimental hybrids to assess it progress has been realized by the programme at the UKZN in developing hybrids which are adapted to environments in South Africa.

References

Adu, G., R. Akromah, M. Abdulai, K. Obeng-Antwi, A. Kena, K. Tengan and H. Alidu. 2013. Assessment of Genotype by Environment interactions and Grain Yield Performance of Extra-Early Maize (*Zea Mays L.*) Hybrids. Journal of Biology, Agriculture and Healthcare, 3(12): 7-15.

Akter, A., M. Jamil Hassan, M. Umma Kulsum, M. Islam and K. Hossain. 2014. AMMI Biplot Analysis for Stability of Grain Yield in Hybrid Rice (*Oryza sativa L.*). Journal of Rice Research, 2(2): 126-129.

Ali, M. B., A. N. El-Sadek, M. A. Sayed and M. A. Hassaan. 2015. AMMI biplot analysis of genotype× environment interaction in wheat in Egypt. Egyptian Journal of Plant Breeding, 19(6): 1889-1901.

Asfaw, A., F. Alemayehu, F. Gurum and M. Atnaf. 2009. AMMI and SREG GGE biplot analysis for matching varieties onto soybean production environments in Ethiopia. Scientific Research and Essays, 4(11): 1322-1330.

Badu-Apraku, B., M. Oyekunle, K. Obeng-Antwi, A. Osuman, S. Ado, N. Coulibay, C. Yallou, M. Abdulai, G. Boakyewaa and A. Didjeira. 2012. Performance of extra-early maize cultivars based on GGE biplot and AMMI analysis. The Journal of Agricultural Science, 150(4): 473-483.

Basford, K. and M. Cooper. 1998. Genotype x environment interactions and some considerations of their implications for wheat breeding in Australia. Australian Journal of Agricultural Research, 49(2): 153-174.

Crossa, J. 1990. Statistical analyses of multilocation trials. Advances in Agronomy, 44: 55-85.

Crossa, J., H. Gauch and R.W. Zobel. 1990. Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. Crop Science, 30(3): 493-500.

Farshadfar, E., M. Rashidi, M.M. Jowkar and H. Zali. 2013. GGE biplot analysis of genotype× environment interaction in chickpea genotypes. European Journal of Experimental Biology, 3(1): 417-423.

GENSTAT(2011). Statistical data analysis programme. UKZN.

Grada, F. and S. Ciulca. 2013. Analysis of genotype x environment interaction for yield in some maize hybrids. Journal of Horticulture, Forestry and Biotechnology, 17(2): 192-196.

İlker, E., F.A. Tonk, Ö. Çaylak, M. Tosun and İ. Özmen. 2009. Assessment of Genotype X Enviroment Interactions for Grain Yield in Maize Hybrids Using Ammi and GGE Biplot Analyses. Turkish Journal of Field Crops, 14(2): 123-135.

Kandus, M., D. Almorza, R. Boggio Ronceros and J. Salerno. 2010. Statistical models for evaluating the genotype-environment interaction in maize (*Zea mays L.*). Phyton-Revista Internacional de Botanica Experimental, 79: 39-46.

Lule, D., M. Fetene, S. de Villiers and K. Tesfaye. 2014. Additive Main Effects and Multiplicative Interactions (AMMI) and genotype by environment interaction (GGE) biplot analyses aid selection of high yielding and adapted finger millet varieties. Journal of Applied Biosciences, 76(1): 6291–6303.

Mitrović, B., D. Stanisavljević, S. Treskić, M. Stojaković, M. Ivanović, G. Bekavac and M. Rajković. 2012. Evaluation of experimental maize hybrids tested in multi-location trials using AMMI and GGE biplot analyses. Turkish Journal of Field Crops, 17(1): 35-40.

Mohammadi, M., P. Sharifi, R. Karimizadeh, J.A. Jafarby, H. Khanzadeh, T. Hosseinpour, M.M. Poursiabidi, M. Roustaii, M.H. Hosni and P. Mohammadi. 2015. Stability of grain yield of durum wheat genotypes by AMMI model. Agriculture and Forestry, 61(3): 181-193.

Munawar, M., G. Hammad and M. Shahbaz. 2013. Evaluation of maize (*Zea Mays L.*) hybrids under different environments by GGE biplot analysis. American-Eurasian Journal of Agriculture & Environmental Sciences, 13(9): 1252-1257.

Rakshit, S., K. Ganapathy, S. Gomashe, A. Rathore, R. Ghorade, M.N. Kumar, K. Ganesmurthy, S. Jain, M. Kamtar and J. Sachan. 2012. GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. Euphytica, 185(3): 465-479.

Rashidi, M., E. Farshadfar and M.M. Jowkar. 2013. AMMI analysis of phenotypic stability in chickpea genotypes over stress and non-stress environments. International Journal of Agricultural and Crop Sciences, 5(3): 253-260.

Solonechnyi, P., N. Vasko, A. Naumov, O. Solonechnaya, O. Vazhenina, O. Bondareva and Y. Logvinenko. 2015. GGE biplot analysis of genotype by environment interaction of spring barley varieties. Žemdirbystė (Agriculture), 102(4): 431-436.

Stojaković, M., M. Ivanović, G. Bekavac, A. Nastasić, B. Purar, B. Mitrović and D. Stanisavljević. 2012. Evaluation of new NS maize hybrids using biplot analysis. Genetika, 44(1): 1-12.

Tiwari, R., A.K. Sharma and B. Kumar. 2014. Genotype-Environment interaction for yield and its attributes in maize (*Zea Mays L*.). Plant Archives, 14(2): 841-845.

Tonk, F.A., E. Ilker and M. Tosun. 2011. Evaluation of genotype x environment interactions in maize hybrids using GGE biplot analysis. Crop breeding and applied biotechnology, 11(1): 01-09.

Verma, A., R. Chatrath and I. Sharma. 2015. AMMI and GGE biplots for G× E analysis of wheat genotypes under rain fed conditions in central zone of India. Journal of Applied and Natural Science, 7(2): 656-661.

Yan, W. and N.A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Canadian Journal of Plant Science, 86(3): 623-645.

CHAPTER FIVE

General overview of the study

This chapter outlines an overview of the completed study by giving a summary of the major objectives, major findings, drawing out implications of the study findings and recommendations for the future.

The specific objectives of the study were:

i) To investigate and establish the genetic diversity using SNP molecular markers in a set of 51 maize inbred lines from the UKZN breeding programme.

ii) To determine the level of genetic distances between the founder parents and their progenies (advanced inbred lines).

iii) To determine the clustering of the inbred lines into heterotic groups.

iv) To determine the genotype by environment interaction in white maize hybrids.

v) To determine hybrid performance in terms of which hybrid performed better and stable in which environment.

5.2 Summary of main findings

The main findings are presented below:

5.2.1 Genetic diversity

- Three hundred and ninety six SNP markers were successful in discriminating the 51 inbred lines according to genetic distances and genetic clusters. Therefore, two clusters were observed (I and II) with ten sub-clusters (A - J).
- The SNP markers indicated the founder parents are divergent, which explained the significant variation in their progenies.
- The highest genetic distance of 0.3526 was recorded between founder parent LN43 and elite line LN50 indicating these lines complement each other and are potentially important for making a hybrid cross.

- The genetic distance between founder parents and standard lines ranged from 0.1380 to 0.3302. The variation in progeny lines was wide, thus suggesting that the inbred lines are different regardless of the similarities in gene frequency.

- Inbred lines which were clustered on the same group could be assigned to one heterotic group and inbred lines on different clusters could be allocated to other different heterotic groups.

5.2.2 Hybrid performance

The study identified hybrids which can be advanced in the breeding programme:

- The study revealed significant differences between hybrids and environments as main effects and their interaction. The IPCA1 and IPCA2 were found to be highly significant.

- GGE biplot analysis was used successfully to discriminate performance of hybrids in various locations (which-won-where).

- AMMI model selected four best hybrids per environment, however, hybrid PAN6Q445B performed better in five environments, suggesting high adaptability.

- Hybrid 11C3417 and SC403 was the first selection in environment E2 and E6 respectively and out performed the check hybrid PAN6Q445B indicating a new useful product from the UKZN breeding programme. Hybrid 11C3201 emerged in four locations as part of the AMMI best four hybrid selections.

- The significant level of GXE indicates the importance of testing of genotypes over multiple seasons and locations to identify superior and stable hybrids.

5.2.3 Major outcomes from literature review

The following conclusions could be deduced from the literature.

- Maize is a major crop in South Africa and is staple food for millions of people in Africa. Because of increasing population, the demand for maize also increases with the environment becoming a great challenge in the production of the crop thereof.
- Molecular markers can be used to estimate genetic diversity of a given breeding population. There are various factors to consider in choosing the type of genetic markers to employ. Molecular markers are compared to each other by outlining their parameters.

- Genetic diversity is an important criterion used to select parental combinations for future development of progenies with high genetic variability.
- Genotype by environment interaction is an important consideration in crop breeding as it complicates selection of important traits like yield because of cross interactions among hybrid ranks across environments. Therefore testing hybrid performance across various testing locations reduces a chance of misleading selection results and recommendations.
- Identification of stable and adapted genotypes to specific target areas is crucial in obtaining potential candidates to be used for hybrid combinations and production.

5.3 Closing remarks: Breeding implications and future directions

- The study revealed high genetic diversity among the progeny lines, suggesting there is a high potential for producing new and superior hybrids.
- The identified heterotic groups in the diversity analysis indicate the opportunity of exploiting heterosis and developing high performing hybrids.
- The study revealed new hybrids such as 11C3417; SC403 and 14XH149 which out performed the check hybrids in Carolina, and Mooi River respectively. This implies the breeding programme pipeline has hybrids to penetrate the local hybrid product market for the benefit of the farmer in the enviroments which are represented by these locations.

6. Appendices

Appendix 1: Details of the 396 SNP markers used to genotype 51 inbred lines in this study.

Marker	Availability	He	PIC	Gene Diversity
Fea2_1	1.0000	0.1765	0.3492	0.4508
Fea2_2	1.0000	0.0588	0.3419	0.4377
PHM12794_47	1.0000	0.3333	0.3750	0.5000
PHM4196_27	1.0000	0.1569	0.3524	0.4567
PHM4348_16	1.0000	0.0588	0.1197	0.1278
PHM4531_46	1.0000	0.0588	0.1197	0.1278
PZA00005_8	1.0000	0.0980	0.1977	0.2224
PZA00071_2	1.0000	0.1373	0.2572	0.3032
PZA00136_2	1.0000	0.1373	0.3555	0.4623
PZA00160_3	1.0000	0.1569	0.2484	0.2907
PZA00172_11	1.0000	0.0000	0.0377	0.0384
PZA00210_9	1.0000	0.0196	0.0192	0.0194
PZA00270_1	1.0000	0.0588	0.2194	0.2509
PZA00270_3	1.0000	0.0392	0.0725	0.0754
PZA00297_7	1.0000	0.0392	0.1861	0.2076
PZA00332_7	1.0000	0.0784	0.2656	0.3153
PZA00334_2	1.0000	0.0784	0.3749	0.4998
PZA00363_7	1.0000	0.0980	0.3016	0.3700
PZA00403_5	1.0000	0.1961	0.3290	0.4152
PZA00442_5	1.0000	0.1176	0.3524	0.4567
PZA00444_5	1.0000	0.0000	0.0000	0.0000
PZA00455_16	1.0000	0.0392	0.1612	0.1769
PZA00460_8	1.0000	0.0784	0.2951	0.3599
PZA00462_2	1.0000	0.0000	0.1612	0.1769
PZA00466_2	1.0000	0.0196	0.0555	0.0571
PZA00486_2	1.0000	0.4118	0.2735	0.3270
PZA00489_1	1.0000	0.0000	0.0000	0.0000
PZA00499_12	1.0000	0.1373	0.3715	0.4931
PZA00527_10	1.0000	0.0392	0.2295	0.2645
PZA00562_4	1.0000	0.2745	0.3749	0.4998
PZA00565_3	1.0000	0.0000	0.0000	0.0000
PZA00578_2	1.0000	0.0000	0.0377	0.0384
PZA00582_4	1.0000	0.0196	0.0555	0.0571
PZA00587_6	1.0000	0.0000	0.0000	0.0000
PZA00600_11	1.0000	0.0000	0.0000	0.0000
PZA00603_1	1.0000	0.2157	0.3735	0.4969
PZA00606_10	1.0000	0.0000	0.0377	0.0384
PZA00606_3	1.0000	0.2549	0.3608	0.4723
PZA00721_4	1.0000	0.1765	0.3608	0.4723
PZA00755_2	1.0000	0.1569	0.3290	0.4152
PZA00770_1	1.0000	0.1569	0.3582	0.4675

Marker	Availability	He	PIC	Gene Diversity
PZA00793_2	1.0000	0.0000	0.0000	0.0000
PZA00878_2	1.0000	0.0588	0.1480	0.1609
PZA00881_1	1.0000	0.0784	0.2811	0.3383
PZA00902_1	1.0000	0.0392	0.0377	0.0384
PZA00925_2	1.0000	0.1961	0.1612	0.1769
PZA00944_2	1.0000	0.0000	0.0377	0.0384
PZA00948_1	1.0000	0.1176	0.2951	0.3599
PZA01038_1	1.0000	0.0588	0.1480	0.1609
PZA01230_1	1.0000	0.0196	0.0192	0.0194
PZA01292_1	1.0000	0.0588	0.1197	0.1278
PZA01410_1	1.0000	0.0392	0.0377	0.0384
PZA01557_1	1.0000	0.0196	0.0192	0.0194
PZA01570_1	1.0000	0.0392	0.1612	0.1769
PZA01588_1	1.0000	0.0784	0.2484	0.2907
PZA01591_1	1.0000	0.1765	0.3608	0.4723
PZA01597_1	1.0000	0.0392	0.0725	0.0754
PZA01619_1	1.0000	0.2157	0.2883	0.3493
PZA01715_2	1.0000	0.1176	0.3190	0.3983
PZA01877_2	1.0000	0.0000	0.0725	0.0754
PZA01964_29	1.0000	0.0392	0.0377	0.0384
PZA02012_7	1.0000	0.0392	0.3631	0.4767
PZA02029_21	1.0000	0.0196	0.1739	0.1924
PZA02094_9	1.0000	0.1176	0.3582	0.4675
PZA02113_1	1.0000	0.0784	0.2484	0.2907
PZA02129_1	1.0000	0.2353	0.3631	0.4767
PZA02148_1	1.0000	0.1569	0.3290	0.4152
PZA02247_1	1.0000	0.1176	0.3457	0.4444
PZA02264_5	1.0000	0.0784	0.2656	0.3153
PZA02291_1	1.0000	0.1961	0.3702	0.4906
PZA02396_14	1.0000	0.1569	0.3457	0.4444
PZA02408_2	1.0000	0.0392	0.2088	0.2368
PZA02426_1	1.0000	0.0000	0.0725	0.0754
PZA02436_1	1.0000	0.2549	0.3336	0.4231
PZA02509_14	1.0000	0.0784	0.1861	0.2076
PZA02514_1	1.0000	0.0196	0.0192	0.0194
PZA02549_3	1.0000	0.0392	0.0725	0.0754
PZA02564_2	1.0000	0.0588	0.1480	0.1609
PZA02585_2	1.0000	0.1176	0.3457	0.4444
PZA02606_1	1.0000	0.0196	0.0889	0.0932
PZA02683_1	1.0000	0.0392	0.1612	0.1769
PZA02722_1	1.0000	0.1765	0.3735	0.4969
PZA02817_15	1.0000	0.0392	0.2484	0.2907
PZA02817_3	1.0000	0.0784	0.2295	0.2645
PZA02949_26	1.0000	0.0000	0.0000	0.0000

Marker	Availability	He	PIC	Gene Diversity
PZA02984_10	1.0000	0.0588	0.0555	0.0571
PZA03012_10	1.0000	0.0000	0.0000	0.0000
PZA03034_1	1.0000	0.1373	0.3688	0.4877
PZA03116_2	1.0000	0.2745	0.3749	0.4998
PZA03182_5	1.0000	0.0588	0.3336	0.4231
PZA03191_2	1.0000	0.0784	0.0725	0.0754
PZA03243_4	1.0000	0.1176	0.1046	0.1107
PZA03244_4	1.0000	0.0196	0.0555	0.0571
PZA03255_4	1.0000	0.0196	0.1197	0.1278
PZA03289_4	1.0000	0.0000	0.0377	0.0384
PZA03359_4	1.0000	0.0196	0.0192	0.0194
PZA03366_2	1.0000	0.0000	0.0000	0.0000
PZA03384_1	1.0000	0.0000	0.0000	0.0000
PZA03385_1	1.0000	0.1569	0.3290	0.4152
PZA03385_2	1.0000	0.3137	0.2295	0.2645
PZA03388_1	1.0000	0.2157	0.3555	0.4623
PZA03411_3	1.0000	0.0000	0.0000	0.0000
PZA03431_1	1.0000	0.0000	0.0000	0.0000
PZA03445_1	1.0000	0.1569	0.3379	0.4306
PZA03452_6	1.0000	0.0392	0.1046	0.1107
PZA03461_1	1.0000	0.0196	0.0555	0.0571
PZA03470_1	1.0000	0.0196	0.0555	0.0571
PZA03477_1	1.0000	0.2353	0.3741	0.4983
PZA03478_1	1.0000	0.0392	0.1612	0.1769
PZA03484_1	1.0000	0.0588	0.1480	0.1609
PZA03490_1	1.0000	0.1765	0.2572	0.3032
PZA03498_1	1.0000	0.0196	0.0555	0.0571
PZA03504_1	1.0000	0.0588	0.1197	0.1278
PZA03505_1	1.0000	0.1176	0.3524	0.4567
PZA03520_3	1.0000	0.0196	0.0192	0.0194
PZA03533_1	1.0000	1.0000	0.3750	0.5000
PZA03568_1	1.0000	0.0196	0.0192	0.0194
PZA03569_2	1.0000	0.0196	0.0192	0.0194
PZA03573_1	1.0000	0.0588	0.0555	0.0571
PZA03573_3	1.0000	0.2157	0.3608	0.4723
PZA03587_1	1.0000	0.0196	0.0555	0.0571
PZA03598_1	1.0000	0.1176	0.3290	0.4152
PZA03629_1	1.0000	0.0588	0.0889	0.0932
PZA03637_3	1.0000	0.0588	0.0889	0.0932
PZA03638_1	1.0000	0.0588	0.0889	0.0932
PZA03645_2	1.0000	0.0196	0.1197	0.1278
PZA03668_4	1.0000	0.0000	0.0000	0.0000
PZA03673_1	1.0000	0.0196	0.0192	0.0194

Marker	Availability	He	PIC	Gene Diversity
PZA03673_2	1.0000	0.0000	0.0000	0.0000
PZA03677_1	1.0000	0.0588	0.1739	0.1924
PZA03686_1	1.0000	0.0000	0.1341	0.1446
PZA03695_2	1.0000	0.0196	0.0192	0.0194
PZA03696_2	1.0000	0.0000	0.0377	0.0384
PZA03696_3	1.0000	0.0196	0.0192	0.0194
PZA03700_3	1.0000	0.0000	0.0000	0.0000
PZA03706_1	1.0000	0.0196	0.0555	0.0571
PZA03714_1	1.0000	0.0392	0.0725	0.0754
PZA03716_1	1.0000	0.9804	0.3749	0.4998
PZA03719_1	1.0000	0.1373	0.3419	0.4377
PZA03728_1	1.0000	0.0392	0.0377	0.0384
PZA03731_2	1.0000	0.0000	0.0000	0.0000
PZA03732_2	1.0000	0.0000	0.0000	0.0000
PZA03732_3	1.0000	1.0000	0.3750	0.5000
PZA03733_1	1.0000	0.0784	0.2295	0.2645
PZA03735_1	1.0000	0.0980	0.2392	0.2778
PZA03742_2	1.0000	0.2745	0.3582	0.4675
PZA03747_1	1.0000	0.0784	0.2088	0.2368
PZA03750_2	1.0000	0.0196	0.1480	0.1609
PZA03760_3	1.0000	0.0392	0.0377	0.0384
PZB00001_2	1.0000	0.1765	0.3419	0.4377
PZB00054_3	1.0000	0.0000	0.0000	0.0000
PZB00062_10	1.0000	0.0000	0.0000	0.0000
PZB00062_9	1.0000	0.0392	0.0377	0.0384
PZB00104_1	1.0000	0.0588	0.2572	0.3032
PZB00114_1	1.0000	0.0588	0.2194	0.2509
PZB00125_1	1.0000	0.0196	0.0192	0.0194
PZB00165_6	1.0000	0.1373	0.2735	0.3270
PZB00175_6	1.0000	0.0196	0.0555	0.0571
PZB00207_3	1.0000	0.0000	0.0000	0.0000
PZB00235_1	1.0000	0.2157	0.3608	0.4723
PZB00425_1	1.0000	0.0196	0.0192	0.0194
PZB00592_1	1.0000	0.1373	0.3555	0.4623
PZB00607_2	1.0000	0.0000	0.0377	0.0384
PZB00677_3	1.0000	0.0000	0.0377	0.0384
PZB00677_4	1.0000	0.0196	0.0889	0.0932
PZB00746_1	1.0000	0.0000	0.0000	0.0000
PZB00772_4	1.0000	0.0196	0.0192	0.0194
PZB00895_3	1.0000	0.1176	0.1046	0.1107
PZB00963_2	1.0000	0.0588	0.0889	0.0932
PZB00963_3	1.0000	0.0000	0.0000	0.0000
PZB01021_5	1.0000	0.0196	0.0889	0.0932
PZB01051_1	1.0000	0.0000	0.0377	0.0384

Marker	Availability	He	PIC	Gene Diversity
PZB01057_4	1.0000	0.0000	0.0000	0.0000
PZB01086_2	1.0000	0.0000	0.0000	0.0000
PZB01103_4	1.0000	0.0000	0.0000	0.0000
PZB01107_8	1.0000	0.0196	0.0192	0.0194
PZB01110_6	1.0000	0.0392	0.1861	0.2076
PZB01111_3	1.0000	0.0392	0.0725	0.0754
PZB01111_6	1.0000	0.3333	0.3715	0.4931
PZB01112_1	1.0000	0.0196	0.1739	0.1924
PZB01114_2	1.0000	0.0196	0.0889	0.0932
PZB01186_1	1.0000	0.0784	0.1861	0.2076
PZB01186_4	1.0000	0.0000	0.0000	0.0000
PZB01261_2	1.0000	0.1765	0.2392	0.2778
PZB01301_6	1.0000	0.1373	0.2392	0.2778
PZB01370_1	1.0000	0.2157	0.3555	0.4623
PZB01412_2	1.0000	0.0000	0.0377	0.0384
PZB01460_2	1.0000	0.0000	0.0000	0.0000
PZB01463_2	1.0000	0.0000	0.0000	0.0000
PZB01463_7	1.0000	0.0196	0.0192	0.0194
PZB01500_1	1.0000	0.0196	0.0192	0.0194
PZB01617_2	1.0000	0.0000	0.1046	0.1107
PZB01642_2	1.0000	0.0000	0.0000	0.0000
PZB01683_2	1.0000	0.0000	0.0000	0.0000
PZB01689_3	1.0000	0.1176	0.3524	0.4567
PZB01730_3	1.0000	0.0392	0.0725	0.0754
PZB01856_1	1.0000	0.9216	0.3735	0.4969
PZB01869_4	1.0000	0.6078	0.3336	0.4231
PZB01963_2	1.0000	0.0392	0.0377	0.0384
PZB01963_4	1.0000	0.0000	0.0000	0.0000
PZB01964_5	1.0000	0.0196	0.0889	0.0932
PZB01977_11	1.0000	0.0392	0.3457	0.4444
PZB01977_4	1.0000	0.1176	0.1341	0.1446
PZB01977_9	1.0000	0.3137	0.2295	0.2645
PZB02017_2	1.0000	0.1176	0.2484	0.2907
PZB02020_2	1.0000	0.0000	0.0000	0.0000
PZB02033_1	1.0000	0.6275	0.3379	0.4306
PZB02033_2	1.0000	0.0196	0.0555	0.0571
PZB02122_1	1.0000	0.2157	0.2194	0.2509
PZB02179_1	1.0000	0.0392	0.2484	0.2907
PZB02227_2	1.0000	0.0980	0.2572	0.3032
PZB02448_1	1.0000	0.2353	0.3524	0.4567
PZB02516_1	1.0000	0.1569	0.3671	0.4844
PZB02534_3	1.0000	0.0392	0.1046	0.1107
PZB02542_1	1.0000	0.0000	0.0000	0.0000
PZB02542_3	1.0000	0.0000	0.0000	0.0000

Marker	Availability	He	PIC	Gene Diversity
PZB02544_1	1.0000	0.1176	0.1046	0.1107
PZD00016_4	1.0000	0.0392	0.1046	0.1107
PZD00022_6	1.0000	0.0784	0.3190	0.3983
PZD00027_5	1.0000	0.0196	0.0192	0.0194
PZD00043_2	1.0000	0.0000	0.0000	0.0000
PZD00043_4	1.0000	0.0000	0.0000	0.0000
PZD00056_1	1.0000	0.1176	0.2951	0.3599
PZD00066_5	1.0000	0.0196	0.0192	0.0194
PZD00072_2	1.0000	0.1961	0.3457	0.4444
Ra1_1	1.0000	0.0196	0.0192	0.0194
ba1_6	1.0000	0.0000	0.0000	0.0000
fea2_3	1.0000	0.0196	0.0192	0.0194
sh2_3	1.0000	0.0392	0.2484	0.2907
zb27_1	1.0000	0.0392	0.0725	0.0754
zb7_2	1.0000	0.0000	0.0377	0.0384
PZA00031_5	0.9804	0.0600	0.0905	0.0950
PZA00047_2	0.9804	0.2200	0.3626	0.4758
PZA00210_8	0.9804	0.0200	0.0196	0.0198
PZA00237_8	0.9804	0.1000	0.0905	0.0950
PZA00297_4	0.9804	0.0200	0.0565	0.0582
PZA00326_18	0.9804	0.1600	0.3481	0.4488
PZA00498_5	0.9804	0.1800	0.3444	0.4422
PZA00516_3	0.9804	0.0400	0.2327	0.2688
PZA00523_2	0.9804	0.0200	0.0196	0.0198
PZA00587_4	0.9804	0.1800	0.2225	0.2550
PZA00616_13	0.9804	0.1400	0.3444	0.4422
PZA00726_8	0.9804	0.0600	0.2768	0.3318
PZA00740_1	0.9804	0.0000	0.0384	0.0392
PZA01029_1	0.9804	0.1000	0.3165	0.3942
PZA01216_1	0.9804	0.1200	0.2983	0.3648
PZA01315_1	0.9804	0.0200	0.3444	0.4422
PZA01652_1	0.9804	0.2400	0.3602	0.4712
PZA01726_1	0.9804	0.0400	0.0384	0.0392
PZA02011_1	0.9804	0.0400	0.2118	0.2408
PZA02197_1	0.9804	0.0600	0.2604	0.3078
PZA02203_1	0.9804	0.2200	0.3741	0.4982
PZA02266_3	0.9804	0.7600	0.3602	0.4712
PZA02296_1	0.9804	0.1200	0.3481	0.4488
PZA02388_1	0.9804	0.1600	0.3219	0.4032
PZA02423_1	0.9804	0.1800	0.3668	0.4838
PZA02478_7	0.9804	0.3200	0.3648	0.4800
PZA02496_1	0.9804	0.0800	0.3750	0.5000
PZA02589_1	0.9804	0.2800	0.3685	0.4872
PZA02616_1	0.9804	0.2400	0.3602	0.4712

Marker	Availability	He	PIC	Gene Diversity
PZA02746_2	0.9804	0.2600	0.3668	0.4838
PZA02890_4	0.9804	0.0600	0.2006	0.2262
PZA02981_2	0.9804	0.0600	0.3047	0.3750
PZA03069_4	0.9804	0.1000	0.2424	0.2822
PZA03120_1	0.9804	0.1400	0.1766	0.1958
PZA03243_7	0.9804	0.0000	0.0000	0.0000
PZA03329_1	0.9804	0.1400	0.2424	0.2822
PZA03381_2	0.9804	0.0400	0.1638	0.1800
PZA03388_2	0.9804	0.0000	0.0000	0.0000
PZA03398_2	0.9804	0.1000	0.2006	0.2262
PZA03442_1	0.9804	0.0800	0.1889	0.2112
PZA03462_1	0.9804	0.2000	0.3746	0.4992
PZA03474_1	0.9804	0.0200	0.1504	0.1638
PZA03519_2	0.9804	0.0200	0.1504	0.1638
PZA03528_1	0.9804	0.1600	0.3108	0.3848
PZA03583_2	0.9804	0.0400	0.0384	0.0392
PZA03607_1	0.9804	0.0800	0.2327	0.2688
PZA03632_2	0.9804	0.0000	0.0384	0.0392
PZA03650_1	0.9804	0.0400	0.2516	0.2952
PZA03663_1	0.9804	0.3000	0.3725	0.4950
PZA03668_1	0.9804	0.1200	0.2688	0.3200
PZA03676_2	0.9804	0.0000	0.0000	0.0000
PZA03714_3	0.9804	0.0600	0.1217	0.1302
PZB00068_1	0.9804	0.2400	0.3685	0.4872
PZB00087_1	0.9804	1.0000	0.3750	0.5000
PZB00092_2	0.9804	0.4200	0.2768	0.3318
PZB00686_2	0.9804	0.1400	0.3626	0.4758
PZB00859_1	0.9804	0.1000	0.1766	0.1958
PZB01186_3	0.9804	0.1400	0.3165	0.3942
PZB01403_1	0.9804	0.1400	0.3444	0.4422
PZB01403_3	0.9804	0.0400	0.0384	0.0392
PZB01403_4	0.9804	0.0600	0.2424	0.2822
PZB01647_1	0.9804	0.0600	0.1217	0.1302
PZB01881_11	0.9804	0.0200	0.0565	0.0582
PZB01919_1	0.9804	0.2200	0.3668	0.4838
PZB01963_1	0.9804	0.0400	0.0739	0.0768
PZB01963_3	0.9804	0.0400	0.0739	0.0768
PZB02155_1	0.9804	0.1400	0.3515	0.4550
PZB02480_1	0.9804	0.0800	0.3318	0.4200
PZD00010_3	0.9804	0.0400	0.0384	0.0392
PZD00067_2	0.9804	0.1800	0.3444	0.4422
ae1_7	0.9804	0.1200	0.3714	0.4928
csu1171_2	0.9804	0.3200	0.3746	0.4992
d8_3	0.9804	0.2600	0.3741	0.4982

Marker	Availability	He	PIC	Gene Diversity
su1_9	0.9804	0.0000	0.0384	0.0392
PZA00040_19	0.9608	0.1633	0.1387	0.1499
PZA00130_9	0.9608	0.0408	0.3724	0.4948
PZA00273_5	0.9608	0.0204	0.0921	0.0968
PZA00310_5	0.9608	0.0000	0.1387	0.1499
PZA00315_6	0.9608	0.0204	0.0200	0.0202
PZA00963_3	0.9608	0.0408	0.3139	0.3898
PZA01238_2	0.9608	0.0408	0.0752	0.0783
PZA01301_1	0.9608	0.2449	0.3505	0.4531
PZA01332_2	0.9608	0.1020	0.3299	0.4167
PZA01427_1	0.9608	0.0816	0.3568	0.4648
PZA01473_1	0.9608	0.0204	0.0200	0.0202
PZA01677_1	0.9608	0.0204	0.0200	0.0202
PZA01714_1	0.9608	0.1224	0.3505	0.4531
PZA01883_2	0.9608	0.0408	0.1083	0.1150
PZA02260_2	0.9608	0.0816	0.3346	0.4248
PZA02385_6	0.9608	0.0408	0.1083	0.1150
PZA02398_2	0.9608	0.2857	0.3724	0.4948
PZA02450_1	0.9608	0.0204	0.1529	0.1668
PZA02673_1	0.9608	0.0612	0.2456	0.2868
PZA02824_3	0.9608	0.1837	0.3712	0.4925
PZA02825_8	0.9608	0.2245	0.3390	0.4325
PZA02916_5	0.9608	0.1224	0.3505	0.4531
PZA02921_4	0.9608	0.1224	0.3249	0.4082
PZA03012_12	0.9608	0.1429	0.3712	0.4925
PZA03046_4	0.9608	0.2041	0.3749	0.4998
PZA03243_3	0.9608	0.0000	0.0000	0.0000
PZA03247_1	0.9608	0.0000	0.1665	0.1833
PZA03367_1	0.9608	0.2449	0.3664	0.4831
PZA03507_1	0.9608	0.2653	0.3682	0.4867
PZA03527_1	0.9608	0.0816	0.3249	0.4082
PZA03610_1	0.9608	0.0408	0.2149	0.2449
PZA03624_2	0.9608	0.1224	0.3505	0.4531
PZA03638_2	0.9608	0.1633	0.3346	0.4248
PZA03651_1	0.9608	0.0000	0.1918	0.2149
PZA03742_1	0.9608	0.2245	0.3538	0.4592
PZA03745_1	0.9608	0.2041	0.3431	0.4398
PZB00104_3	0.9608	0.1837	0.3733	0.4967
PZB00123_1	0.9608	0.2245	0.3469	0.4467
PZB00750_1	0.9608	0.2041	0.3724	0.4948
PZB00865_2	0.9608	0.1224	0.3568	0.4648
PZB01114_1	0.9608	0.0612	0.2257	0.2593
PZB01457_1	0.9608	0.2041	0.3749	0.4998
PZB01913_2	0.9608	0.1633	0.3698	0.4898

Marker	Availability	He	PIC	Gene Diversity
PZB01977_5	0.9608	0.0000	0.0000	0.0000
PZB02017_3	0.9608	0.0816	0.2149	0.2449
PZB02058_2	0.9608	0.2653	0.3746	0.4992
PZB02059_1	0.9608	0.1020	0.3746	0.4992
PZB02425_3	0.9608	0.0204	0.1239	0.1327
d8_2	0.9608	0.1224	0.1083	0.1150
PZA00305_2	0.9412	0.2500	0.3639	0.4783
PZA00356_8	0.9412	0.0208	0.1261	0.1352
PZA00447_6	0.9412	0.1875	0.3560	0.4633
PZA01062_1	0.9412	0.1042	0.3226	0.4043
PZA01289_1	0.9412	0.1458	0.3615	0.4737
PZA01383_1	0.9412	0.0000	0.0767	0.0799
PZA01447_1	0.9412	0.1875	0.3615	0.4737
PZA01530_1	0.9412	0.0208	0.1823	0.2029
PZA01672_1	0.9412	0.0000	0.0400	0.0408
PZA02923_7	0.9412	0.0000	0.0400	0.0408
PZA03300_2	0.9412	0.0417	0.0400	0.0408
PZA03394_1	0.9412	0.0833	0.3639	0.4783
PZA03440_4	0.9412	0.1042	0.2980	0.3644
PZA03446_1	0.9412	0.2917	0.3528	0.4575
PZA03451_5	0.9412	0.0000	0.2181	0.2491
PZA03602_1	0.9412	0.0625	0.3661	0.4824
PZA03603_1	0.9412	0.0833	0.3733	0.4965
PZA03711_3	0.9412	0.2292	0.3661	0.4824
PZA03723_1	0.9412	0.1042	0.3417	0.4373
PZB00178_3	0.9412	0.0625	0.2289	0.2637
PZB00607_1	0.9412	0.0417	0.0767	0.0799
PZB01352_3	0.9412	0.2083	0.3733	0.4965
PZB01446_1	0.9412	0.7917	0.3680	0.4861
PZB01454_6	0.9412	0.1250	0.2392	0.2778
PZB02002_1	0.9412	0.0000	0.0400	0.0408
PZB02425_2	0.9412	0.1458	0.3615	0.4737
PZB02544_2	0.9412	0.3333	0.2392	0.2778
PHM2885_31	0.9216	0.0000	0.0408	0.0416
PZA00425_11	0.9216	0.1915	0.3582	0.4674
PZA00587_5	0.9216	0.0000	0.0000	0.0000
PZA01688_3	0.9216	0.0213	0.0957	0.1007
PZA02733_1	0.9216	0.1702	0.3657	0.4817
PZA03244_3	0.9216	0.0000	0.0000	0.0000
PZA03632_3	0.9216	0.1277	0.3551	0.4617
PZA03648_2	0.9216	0.0851	0.3609	0.4726
PZA03669_1	0.9216	0.0000	0.0408	0.0416
PZA01607_1	0.9020	0.2391	0.3542	0.4601
PZA03692_1	0.9020	0.0217	0.0611	0.0631

Marker	Availability	He	PIC	Gene Diversity
PZA03758_1	0.9020	0.1087	0.3542	0.4601
PZB00153_3	0.9020	0.4783	0.3113	0.3856
PZB01115_1	0.9020	0.0652	0.3720	0.4941
lac1_3	0.9020	0.2391	0.3691	0.4884
Mean	0.9842	0.1067	0.1940	0.2406

He= heritability. PIC = polymorphic information content