GENETIC DIVERSITY AMONG ELITE MAIZE(*Zea mays* L.) INBRED LINES USING SIMPLE SEQUENCE REPEAT MARKERS

MSc. THESIS

BY

TSEGAYE ABEBE

DECEMBER, 2019

JIMMA, ETHIOPIA

Genetic Diversity among Elite Maize (*Zea mays L.)* Inbred Lines using Simple Sequence Repeat Markers

MSc. Thesis

Submitted to the School of Graduate Studies of Jimma University College of Agriculture and Veterinary Medicine Department of Horticulture and Plant Sciences in Partial Fulfillment of the Requirement for the Degree of Master of Science in Plant Biotechnology

Major Advisor: Leta Tulu (PhD)

Co- Advisor: Wosene Gebreselassie (PhD)

December, 2019

Jimma, Ethiopia

APPROVAL SHEET

Jimma University College of Agriculture and Veterinary Medicine

Thesis Submission for External Defense Request Form (F-07)

Name of student: TsegayeAbebe ID No: RM 1514/10

Program of the study: Plant biotechnology (MSc)

Title: Genetic Diversity among Elite Maize (*Zea mays L.)* Inbred Lines using Simple Sequence Repeat Markers

I have incorporated the suggestion and modification given during internal defense and got the approval of my advisors. Hence, I hereby kindly request the department to allow me to submit my thesis for external defense.

Name ______________ Signature _______________ Date _______________

We, the thesis advisors, have verified that the student has incorporated the suggestions and modifications given during the internal thesis defense and the thesis is ready to be submitted. Hence, we recommend the thesis to be submitted for external defense.

Major Advisors' Name: Leta Tulu (PhD) Signature __________Date ________

Co- Advisors' Name: Wosene Gebreselassie (PhD) Signature Date _________

Decision/suggestion of department graduate council (DGC) ____________________________

Chairperson of DGC Name ___________________Signature __________Date____________

Chairperson of CGC Name ___________________Signature ___________Date__________

DEDICATION

This thesis is dedicated to my beloved mother (AsegedechAyele), my father (Abebe Tumabo) my wife (Abebakassaye), my brother (AddisuAbebe) my child (Amen Tsegaye).

STATEMENT OF THE AUTHOR

First, I declare that this thesis is solely my original work with close supervision and guidance of my advisors. In addition to this all sources of materials used in the thesis have been duly acknowledged. This thesis has been submitted in partial fulfillment for the requirement of an advanced MSc degree at Jimma University and is deposited at the University Library to be made available for borrowers under rules of the Library. I solemnly declare that this thesis is not submitted to any other institution anywhere for the award of any academic degree, diploma, or certificate.

Brief quotation from the thesis is allowable without special permission provided that accurate acknowledgment of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the School of Graduate Studies when in his or her judgment the proposed use of the material is in the interests of the scholarship. However, in all other instances, permission must be obtained from the author.

Name: TsegayeAbebe Tumabo Signature: ___________________

Place: Jimma University, Jimma

Date of submission: December, 2019

BIBILOGRAPHICAL SKETCH

The author was born on June 20, 1991 in Gambella region GoderaworedaMetti town from his father Abebe Tumabo and his mother AsegedechAyele. He attended elementary School at Akashi Elementary and Junior School from 1998-2005 and He attended his secondary school in Teneshu Metti Secondary and Preparatory School from 2006-2007. Then He joined at (GATVT) Gambella agricultural technical vocational training college for the department of plant science in 2008. After three consecutive years of study he graduated with diploma in plant science in 2010 Then, He was employed at GATVT and served as technical assistance then he joined Haramaya University and graduated with B.Sc. Degree in Plant Sciences in 2016. After His graduation, the author was recruited as Graduated Assistance in Gambella University then in August 2018. He joined the School of Graduate Studies at Jimma University College of Agriculture and Veterinary to study MSc in plant biotechnology.

ACKNOWLEDGEMENT

First and foremost, I offer my obeisance to the 'Almighty God' for his boundless blessing, which accompanied me in all the endeavors. I would like to express my special gratitude and heartfelt thanks to my major research advisor Dr. Leta Tulu for his encouragement, constructive suggestions and moral support from the very beginning in executing the research work and in the write-up of this thesis. I wish to take this opportunity to extend special gratitude and indebtedness to my research co-advisor. Wosene Gebreselassie (PhD) for his inspiration, thoughtful guidance and suggestions throughout preparation of this thesis. I would like to appreciate their patience for close follow up of his work from the very beginning to the end. Without their dedication this work would not have been completed on time.

I would like to express my deepest and heartfelt gratitudeto Holeta National Agricultural Biotechnology research staff especially to Ms. KalikidanTesfu, Mr. Fedessakebede for helping me to put my hands on techniques from DNA extraction up to Polymerase chain reaction and their unfailing guidance in each and every necessary procedure of the entire laboratory activities.

I am extremely grateful to my mother, my father and brothers for their understanding and help in using whatever resources they have. I never forget their care for me that brought me to this success. I thank my wife, AbebaKassaye, for her general support, understanding, encouragement, patience and taking care of our son, Amen, during my study period. Finally, I wish my humble thanks to one and all who have directly and indirectly contributed to the conduct of the study.

.

LIST OF ABBREVATIONS

AMOVA	Analysis of Molecular Variance
BNMRC	Bako National Maize Research Center
CIMMYT	International Maize and Wheat Improvement Centre
CITAB	Cetyl Trimethyl Ammonium Bromide
CSA	Central statics Agency
DArT	Diversity Array Technology
EDTA	Ethylene Diamine Tetra Acetate
FAOSTAT	Food and Agriculture Organization Corporate Swastika Database
GD	Genetic diversity
GS	Genetic Similarity
HARC	Holeta Agricultural Research Center
IITA	International Institute of Tropical Agriculture
IPCC	Integrated professional competence course
IPGR	International plant genetics resource institute
MAS	Marker assisted selection
NABRC	National Agricultural Biotechnology Research Center
PCoA	Principle coordinates analysis
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Content
QTLs	Quantitative Treat Loci
RFLP	Restriction Fragment Length Polymorphism
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
UPGMA	Unweighted Pair Group Method with Arithmetic Average
USDA	United State Development of Agriculture

TABLE OF CONTENTS

DEDICATION ... i
STATEMENT OF THE AUTHOR ... ii
BIBILOGRAPHICAL SKETCH ... iii
ACKNOWLEDGEMENT ... iv
LIST OF ABBREVATIONS ... v
TABLE OF CONTENTS ... vi
LIST OF TABLES ... viii
LIST OF FIGURES ... ix
LIST OF APPENDICES ... x
ABSTRACT ... xi
1. INTRODUCTION ... 1
2. LITRATURE REVIEW ... 5
2.1. Origin and Distribution of Maize ... 5
2.2 Genetic diversity in maize ... 5
2.3. Genetic diversity assessment in maize using SSR markers ... 6
2.4 Genetic diversity evaluation ... 11
2.4.1 Quantification of genetic relationships ... 12
2.4.2 Determination of distance measures ... 13
2.4.3 Expressing relationships in a genetic diversity study ... 13
3. MATERIALS AND METHODS ... 14
3.1 Plant Material ... 14
3.2. Genomic DNA extraction ... 18
3. 3. Genomic DNA quality and quantity measurement ... 18
3. 4. SSR primer selection ... 20
3. 5. Amplification and detection of bands ... 23
3. 6. Scoring and data analysis ... 23
3. 7. Diversity parameters and analysis ... 23
3. 8. Analysis of molecular variance (AMOVA) ... 24
3. 9. Genetic distance and cluster analysis ... 24
3. 10. Principal coordinate analysis (PCoA) ... 24
4. RESULTS AND DISCUSSION ... 25
4.1. Polymorphism and Polymorphic information content ... 25

4.3. Analysis of molecular variance (AMOVA) ... 30
4.4. Genetic differentiation and gene flow between populations ... 31
4.5. Pattern of genetic diversity and relationship among populations ... 32
4.6. Genetic distance and relationship among populations ... 36
4.7. Allelic pattern across population ... 37
4.8. Principal coordinates analysis (PCoA) ... 38
4.9 Clustering analysis ... 39
5. SUMMARY AND CONCLUSSION ... 40
6. REFERENCE ... 42
7. APPENDEX ... 53

LIST OF TABLES

Table 1 Pedigree and origin of 37 maize inbred lines used in the study ---------------------------- 16

Table 2 SSR markers used for study genetic diversity in maize inbred lines ---------------------- 20

Table 3 Genetic parameter of 28 SSR markers used in the study of 37 elite maize inbred lines obtained from three sources (population) -- 26

Table 4 Analysis of molecular variance (AMOVA) based on standard permutation across the full data set of elite maize inbred lines from different source --- 30

Table 5 Pair wise genetic differentiation (FST) below the diagonal and pair wise estimate of gene flow (NM) above the diagonal between maize genetic sources ----------------------------------- 31

Table 6 Summary of genetic parameter for three population using 28 SSR markers ------------- 32

Table 7 Pair wise population Nei genetic distance -- 35

Table 8 Mean of important allelic value in three maize populations --------------------------------36

LIST OF FIGURES

Figure 1 Two dimensional plot of PCoA of 37 maize inbred lines and five populations based on SSR data 37
Figure 2 Unweighted Neighbor Joining (NJ) dendrogram showing genetic relationship of 37 maize inbred lines using 28 microsatellites 39

LIST OF APPENDICES

Appendix 1 maize inbred lines and population codes used for SSR data----------------------------54

Appendix 2 Genetic distance between 37 elite maize inbred lines ----------------------------------55

GENETIC DIVERSITY AMONG ELITE MAIZE (*Zea mays L.*) INBRED LINES USING SIMPLE SEQUENCE REPEAT MARKERS

ABSTRACT

In Ethiopia maize productivity still remain far below the world average due to several factors such as; lack of improved and wide adapted varieties, acidity, drought, and foliar disease of maize. Thus genetic improvement of the crop is vital to address some of the constraint. Therefore, the present study was conducted to assess the extent of genetic diversity among elite maize inbred lines using simple sequence repeat markers. Thirty seven elite maize inbred lines were obtained from Bako National Maize Breeding Center and CIMMIYT and twenty nine Simple Sequence Repeat markers with wider genomic coverage were obtained from the maize genomic data base. The research was conducted at molecular biotechnology laboratory of the National Agricultural Biotechnology Research Center (NABRC) at Holeta. Genomic DNA was extracted following DArT (Diversity arrays technology) protocol. Out of 29 SSR markers one showed monomorphic pattern and excluded from analysis. The rest 28 markers were polymorphic and revealed a total of 104 alleles. The number of alleles generated by each marker varied from 2 to 6 with average number of 3.71. The number of effective alleles ranged from 1.11 to 3.53 with a mean of 2.04. Fixation index varied from 0.70 to 1 with a mean of 0.96. The highest observed and expected heterozygosity was 0.17 and 0.626 respectively. The polymorphic information content of the markers ranged from 0.1 to 0. 74 with a mean of 0.5. Analysis of molecular variance showed highly significant (P<0.001) molecular variances. The highest (77%) of the variation was attributed to genetic variability among individuals (AI) within populations, while 23 % of the variation was observed in among population. The highest and lowest Fst value was observed between inbred lines sourced from lowland Mexico and South America, Bako and lowland Mexico respectively. The highest and lowest value of gene flow was observed between inbred line of Bako and lowland Mexico, South America and lowland Mexico population respectively. The highest genetic distance (0.42) was observed between inbred lines obtained from South America and lowland Mexico, whereas the lowest genetic distance (0.27) was showed from inbred lines obtained from Bako and (CIMMIYT) lowland Mexico.The highest number of private alleles was identified in Bako population. The principal coordinate analysis accounted for 69.27% of the total genetic variation. Cluster analysis revealed grouped the inbred lines in to three major groups by merging inbred lines from different populations in cluster I and III. Generally, the present study showed the existence of wide genetic variation among the studied materials. Thus, it provided precise information to use promising combination for exploitation of heterosis and establishment of heterotic group as source materials in maize breeding program.

Key words: Clustering, Genetic diversity, Maize inbred line, Simple sequence repeats (SSR)

1. INTRODUCTION

Maize (*Zea mays L.*) is a member of the grass family *Gramineae (Poaceae*). It is believed to have originated in Central America, specifically Mexico and introduced to Africa by the Portuguese traders in the 16th century (Gibson and Benson, 2002). It is grown in tropical, sub-tropical and temperate regions of the world (Shukla*et al.,* 2014). It is an annual, short day and the only monoecious crop among cereals having male and female inflorescences on separate branches of the same plant (USDA, 2005). It is predominantly cross pollinated crop showing the highest phenotypic and genetic variation (Molin*et al.,* 2013; *Li et al.,* 2014).

Maize is one of the most important staple food crops in the world. It is widely used for animal feed and industrial raw material in the developed countries where as in developing countries it is generally used for food and means of income and source of employment (OECD, 2003). In Ethiopia maize is used as staple food consumed as "*Injera*," Porridge, Bread and *"Nefro*." It is also consumed roasted or boiled as vegetables at green stage. In addition to the above, it is used to prepare *"Tella*" and "*Arekie*." The leaf and stalk are used for animal feed and dried stalk & cob are used for fuel. (MARD, 2014).

Globally maize is one of the highest ranking cereal crops in productivity. According to FAOSTAT, (2017) report the total area coverage of maize in the world was 197,185,936 ha, whereas its annual production was reaches 1.14 billion tone and productivity is 5.75 tone ha^{-1}.Maize is widely grown in Africa. According to FAOSTAT, (2017) report area coverage of maize in Africa was 40.6 million ha and annual production is about 84.2 million t.

In Ethiopia, among all other cereal crops maize ranks first in terms of total production and second in area coverage next to teff (CSA, 2018). Its total area coverage was estimated to be around 2.1 million hectares, whereas the estimated annual national production is about 8.4 million t and productivity is 3.944 t ha^{-1} (CSA, 2018). According to CSA (2017) report, out of 16.3 million private peasant holders growing major cereals, 10.9 million holders (67%) grew maize in 2017 cropping season.

In Ethiopia the average productivity of maize 3.94 ha^{-1} is far below the world average 5.7 ha^{-1}, this is due to limited availability of improved and wide adapted varieties, acidity, drought and Foliar disease of maize including maize streak virus (MSV), Turcicum leaf blight, grey leaf spot or common leaf rust significantly affect maize production (Keno *et al.,* 2018). Thus genetic diversity among maize inbred lines offers an opportunity to address some of these constraints. (Ertiro*et al.,* 2017).

Genetic diversity assessment among maize inbred lines is a pre-requisite to identify suitable parents for exploitation of heterosis (Azam*et al.,* 2018). It can be assessed by morphological and molecular differences existing among the population. Morphological assessment is relatively easy to carry out but it is labor intensive, time-consuming, epigenetics (non-heritable variability or temporary phenotypic change), Influenced by environment and does not show the entire genetic relationships between closely related genotypes (Shiri*et al.,* 2014; Govindaraj*et al.,* 2015). Additionally, morphological markers express limited polymorphism and late expression of traits (Smith and Smith, 1992). On the other hand, molecular markers are more efficient in genetic diversity study of maize because they are not influenced by the environment and the plant can be assessed at any stage of development and effective tools for differentiating between closely related genotypes (Garcia *et al.,* 2004).

The most common molecular markers that are used to assess genetic diversity in maize genotypes include, Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR), and Single Nucleotide Polymorphisms (SNP). Simple sequence repeats (SSRs) or Microsatellite markers are classes of repetitive sequences which are widely-distributed in all eukaryotic genomes. They consist of arrays of tandemly repeated short nucleotide motifs of 1-6 bases, and are called mono-, di-, tri- or tetra nucleotide repeats, respectively (Tautz*et al.,* 1986).

Among from Hybridization and PCR based molecular markers simple sequence repeat marker is an ideal markers to assess the genetic diversity of maize because they have several advantages over the other molecular markers, such as high level of reproducibility, template DNA is needed in small quantity, hyper variability or highly polymorphic and give more genetic information even among very closely related varieties,multi-allelic

nature, co-dominant nature, large in number and present throughout the genome(Xu*et al.,* 2013; Sserumaga et al., 2014)

Several studies have been done on genetic diversity of maize using SSR markers (Gupta *et al.,* 2010; Pandit *et al.*, 2016; Sharma *et al.,* 2017). In Ethiopia very little work has been reported on genetic diversity in maize inbred lines at molecular level using simple sequence repeat marker. Beyene*et al.* (2006) studied the genetic diversity among 62 highland Ethiopian maize genotypes using 20 SSR markers. Who reported the existence of considerable amount of genetic variation in highland Ethiopian maize genotype. Legesse *et al.* (2007) investigated the genetic diversity of 35 maize inbred lines obtained from the highland maize breeding program in Ethiopia and 21 maize inbred lines obtained from CIMMYT Zimbabwe using 27 SSR loci. Who reported the existence of 59% of genetic distance in the inbred lines.Demissew (2014) investigated the genetic diversity of 30 quality protein maize (QPM) and 6 normal maize inbred lines adapted to highland agro-ecology of Ethiopia using 25 SSR markers. Who reported the present of genetic variation in the inbred lines.

The previous researcher studied maize inbred lines which are adapted to highland agro ecology of Ethiopia. Since, mid-altitude sub-humid agro-ecology is a high potential area for maize production in Ethiopia (Mosisa*et al.,* 2012). However, production and productivity of maize in this agro ecology is constrained by previously listed problem.The Ethiopian maize breeding program over long time accumulated several elite maize inbred lines that adapted to mid altitude agro ecological condition of Ethiopia through introduction from foreign sources and developing inbred lines from locally germplasm. Information on genetic diversity among these elite inbred lines has important in broadening the genetic base and genetic enhancement of the crop. Despite this fact, information on genetic variation of elite maize inbred lines introduced from foreign sources and developed from local germplasm source has inadequate. Lack of genetically diverse materials for developing improved high yielding and stress tolerant hybrids is the major limitation in the Ethiopian maize breeding program.

In the current study 24 elite maize inbred lines which is adapted to mid altitude agro ecology of Ethiopia. The 24 inbred lines were identified as a promising source of turcicum leaf blight and gray leaf spot and 13 soil acidity tolerant CIMMYT (South America) elite

maize inbred lines were used. The acid tolerant inbred lines were recently introduced from CIMMYT with the view of broadening the genetic bases of locally available elite Ethiopian maize inbred lines for tolerance to soil acidity. This study trying to examine the genetic distance or similarity among the acid tolerant lines and the Ethiopian elite inbred lines and the generate information from this research support effort on development of acid tolerant and some disease resistance hybrid for mid altitude agro ecology of Ethiopian. This calls for a more comprehensive genetic diversity study using molecular markers. Hence, this research was initiated with the following objectives:-

General objective

To study genetic diversity among elite maize inbred lines using simple sequence repeat markers

Specific Objective

- ✓ To determine the genetic distance and relationship among Ethiopian and CIMMIYT elite maize inbred lines using simple sequence repeat markers.

2. LITRATURE REVIEW

2.1. Origin and Distribution of Maize

Maize (*Zea mays* L), also known as corn, is a cereal grain crop that was originally domesticated in Mesoamerica (OECD, 2003). However, there are many thoughts about the progenitor of maize, it is generally accepted that maize originated from teosinte (*Zea mexicana*L) which is the closest known wild relative of maize (Aylor*et al.*, 2005). Mexico and Guattemala are the native countries of teosinte and it grows wild in cultivated maize fields in its natural habitat (Hallauer and Miranda 1988: Aylor*et al.*, 2005). This plant is similar to maize by having a monoecious flowering habit, same number of chromosomes and is readily crossed with maize (Poehlman and Sleper, 1995).

Maize spread throughout the world after European contact with the Americas in the late 15th and early 16th century (Gibson and Benson, 2002).The crop introduced to Africa by the Portuguese traders in the 16th century (Gibson and Benson, 2002). It has now become a principal cereal crop in the tropics and in the subtropical regions throughout the world. It is unknown what precipitated its domestication, because the edible portion of the wild variety is too small and hard to obtain to be eaten directly, as each kernel is enclosed in a very hard bi-valve shell (Hampl and Hampl, 1997).

2.2. Genetic diversity in maize

Maize is one of the domesticated crop species with the highest level of molecular polymorphism. Nucleotide diversity of more than 5% was reported in certain loci of the maize genome (Henry and Damerval, 1997). The average sequence divergence between any two individuals for a given locus is referred as nucleotide diversity (Buckler and Thornsberry, 2002). The polyploidy origin and the abundance of transposons in maize could make it possible to study genome size evolution (Bennetzen*et al.,* 2005).

Genetic diversity study of maize showed that it is highly variable both within and across populations. DNA sequencing of the *adh1* locus, in *Z. mays ssp. Parviglumis*(the maize progenitor) and Z*ea luxurians*(a distant maize relative) showed that maize retained 77% of the diversity of *parviglumis*and is more diverse than *Z. luxurians*(Eyre-Walker *et al.,* 1998).The molecular variation of maize is three to tenfold higher than that of other cereal

crops (Buckler *et al.,* 2001). Several factors are suggested as reasons for this diversity in maize, *viz*., (1) variability of environments, culture, production system and the type of consumption of maize (Aguirre *et al.,* 1998); (2) the high level of out-crossing in maize favors continuous gene exchange between neighboring plants and in some cases, with their wild relatives; (3) chromosomal duplications in maize are extensive providing new mutational opportunities for creating greater phenotypic variability (Helentjaris*et al.,* 1998); and (4) transposons and retro transposon elements also play a vital role in its genetic variation (Bennetzen*et al.,* 2005).

2.3. Genetic diversity assessment in maize using SSR markers

Molecular marker provide more reliable and consistent information about the genetic diversity of closely related genotypes as compared to morphological and biochemical markers, they directly determine the variation at DNA level. They were applied for the first time in 1980s. A typical DNA marker should be polymorphic, evenly distributed throughout the genome, it should be highly informative, simple, fast, and not expensive and it must need tissue and DNA in small quantity. The available molecular markers that used to assess genetic diversity among maize genotypes include Random Amplified Polymorphic DNA (RAPDs), Restriction Fragment Length Polymorphism (RFLPs), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR), and Single Nucleotide Polymorphisms (SNPs). (Wang *et al.*, 1994; Scott *et al.,* 2000; Manen*et al.,* 2003; Varshney*et al.,* 2005).

Among from Hybridization and PCR based molecular markers simple sequence repeats (SSRs) are preferred to assess the genetic diversity of maize because of the several advantages over the other molecular markers. Such as high level of reproducibility, require simple methodology, template DNA is needed in small quantity, highly polymorphic, gives more genetic information even among very closely related varieties, co-dominant nature, large in number and present throughout the genome (Sserumaga*et al.,* 2014).

Simple sequence repeats were found and were present in the whole eukaryotic genome (Moustacchi and Williamson, 1966), followed by the identification of satellite DNA with repeat motifs ranged from single to thousands basis such as centromeric DNA sequences consisted of 100 bp repeats (Pardue and Gall, 1970). Then satellites composed of 10-30 bp repeat motifs were identified in mammals (Jeffreys*et al.*,1985).

After that satellites of even more shorter length were found called microsatellites. It was in 1982 that Hamada and coworkers reported dinucleotide repeats. As these sequences were ranged from 1-6 bp, so named as simple sequence repeats (SSRs) or short tandem repeats (STRs) (Tautz and Renz, 1984). In plants SSRs were identified for the first time through the hybridization of probes possessing poly G-T and A-G on the phage libraries of genomes from tropical tree (Condit and Hubbell, 1991) and became one of the efficient marker system in present day plant genetics.

Comertpay*et al.* (2012) characterized 98 maize inbred lines using nineteen morphological traits and twenty eight SSR makers. The total alleles found were 172 with average 6.12 allele per locus. The genetic variation detected was minimum 0.18 and maximum 0.63, with average value of 0.35. It was suggested that because of the presence of high level of genetic diversity these genotypes can be proved a good source for maize breeding activities throughout the world

Kumar *et al*. (2012) characterized ninety one maize inbred lines for drought tolerance using six morphological and forty SSR markers. On the basis of SSR markers the number of total alleles found were 124. The numbers of alleles per locus were ranged from 2 to 5 with average value of 3.1 alleles. The PIC value ranged from 0.54 to 0.82, with average value of 0.55. On the basis of SSRs the 91 genotypes were divided into two main clusters and sub-clusters which indicated high genetic variability. The genotypes with more genetic diversity were found that will prove vital for breeding strategies to improve drought tolerant maize genotypes.

Reid *et al*. (2011) evaluated One hundred twenty nine maize genotypes on the basis of pedigree information and by using 105 SSR markers. Based on pedigree information they divided the genotypes into 8 groups which were Lowa Stiff Stalk Synthetics, European flint, Lancaster, Minnesota 13, Early Butler, Iodent, Poineer 3990 and Poineer 3994. Based on SSR markers a total of 380 allelic variants were found with mean value of 0.68 alleles per locus. Genetic similarity ranged from 0.53 to 0.93 with mean of 0.64. Based on 0.64 genetic similarities the genotypes were separated into ten main clusters. When the clusters produced by SSR analysis were compared with that of pedigree analysis, few differences were found i.e. the genotypes of similar pedigree groups were placed in different clusters

Pabendon*et al*. (2010) studied on genetic diversity of six quality protein maize and five normal maize inbred by using twenty four SSR markers. SSR marker result the total number of allele per locus was 94. The minimum number of allele was two and the maximum number allele per locus was six and means value was 3.9 number of allele per locus. The genetic distances varied from 0.55 to 0.9. Cluster analysis showed two main groups. Four QPM and one normal inbred were present in one cluster, while five normal and one QPM inbred in the other. This mixed allocation of inbred was due to the reason that most of the Indonesian inbred have obtained from (CIMMYT) Mexico, The two promising QPM hybrids were found which were crossed from different clusters. The genetic distances indicated by SSR markers and grain yield of F1 hybrids were negatively correlated, which was explained by the fact that environmental factors may have influenced the grain yield.

Shehata*et al*. (2009) analyzed eight maize inbred lines using six SSR markers. The markers result a total of 50 alleles and the minimum allele per locus were 2 and maximum numbers per locus were found 10 and average alleles per locus were 4.35. PIC value of the markers ranged from 0.42 to 0.88 with mean value 0.58

Ranatunga*et al*. (2009) Characterized Forty five maize inbred lines using forty two simple sequence repeat markers. The markers revealed a total of 132 alleles. The genetic diversity was 99.20% with mean polymorphic information content (PIC) value of 0.84. The cluster analysis produced two clusters.

Zheng*et al*. (2008) studied diversity assessment along with genetic relationship among thirty six maize inbred lines. Eighteen from distinct groups in Chinese and eighteen from more diverse American maize using 109 SSR markers. They found overall average polymorphic information content of 0.66 and an average 6 allele per locus. Upon clustering, the inbred line from America i.e. CIMMYT population Pool 41 did not group together with other lines, which means that generally the inbred lines from American population and particularly this line has more genetic diversity, which can be prove a good source for broadening the genetic diversity in Chinese maize.

Lun*et al*. (2008) investigated the genetic diversity of one hundred twenty four maize genotypes using 45 SSR markers. The markers revealed a total of 286 alleles and an

average of 6.4 alleles per locus were detected, while for within landrace genetic variation, a total of 357 alleles at the rate of average 7.93 per locus were noted. Within the landraces high level of genetic diversity was found as compared to genetic variation found among landraces in maize growing in Wuling mountain region in China. This high genetic diversity within the landrace was credited to the fact that usually seeds are renewed by farmers after each year and there is possibility to be affected by the genetic drift but farmers select seeds from the healthy ears, larger in length and diameter, means select the plants with more heterozygosity and diversity in their genetic makeup is maintained

Jambrovic*et al*. (2008) studied on genetic diversity assessment of 15 maize genotypes using 98 SSR markers. So based on the study the total number of alleles was 205. The minimum number of allele per locus was 2 and maximum number of allele per locus was found 8 with mean value of 4.18. The PIC value varied from 0.129 to 0.864 with mean of 0.65. The inbred lines of Osijek Agricultural Institute were identified to have wide range of genetic diversity and will be proved as good source for producing maize varieties.

Kostova*et al*. (2006) studied the genetic diversity of forty one Bulgarian and ten US maize inbred lines using 18 SSRs. The markers revealed a total of 163 alleles with average 9.1 alleles per locus. The high number of alleles per locus was considered possibly because of two reasons among 18 markers 16 were di-nucleotide motif and the investigation of the wide range of inbred. The cluster analysis divided the inbred into two main clusters and in all clusters Bulgarian inbred were present. It was concluded that Bulgarian inbred possess high genetic variability and could be effective for breeding approaches

Vigouroux*et al*. (2005) studied the genetic variation between maize and teosinte using 462 SSR loci present in the whole maize genome, their comparison and magnitude of genetic erosion took place during the domestication of maize. They reported the average number of allele per locus in teosinte was 11.8 with genetic diversity 0.74, while in maize genome the average number of alleles per locus was 9.0, with genetic diversity 0.64. The teosinte possessed 24% higher number of alleles than maize and 12% higher genetic diversity.

Kumari*et al*. (2005) studied the genetic diversity of ten maize inbred lines using forty SSRs which were distributed across 10 maize chromosomes. The result revealed out of 32

SSR markers 27 showed polymorphic pattern hence, included in the analysis, whereas the rest 5SSR markers noted monomorphic pattern and excluded from analysis. The polymorphic markers revealed a total of 82 alleles. The minimum number allele per locus was 2 and the maximum number allele per locus was 4 with a mean of 3.03 allele per locus. The PIC value ranged from 0.38 to 0.70. Only in 4 genotypes 8 unique alleles were found which clearly differentiated these genotypes. The cluster analysis showed high level of genetic diversity among the studied genotypes and determined a fractional agreement with pedigree information and breeding history.

Choukan and Warburton (2005) studied the genetic diversity among thirty eight maize genotypes. Out of 38 maize inbred line 37 inbred lines obtained from Iran and 2 inbred lines from CIMMYT using 43 SSRs. The SSRs revealed a total of 194 allele. The minimum number of allele per locus was 2 and the maximum number of allele per locus was 10 with a mean of 4.5 was detected. The Polymorphic information content value of the markers was range from 0.15 and 0.78 with a mean of 0.53. They found 44 unique alleles through 27 SSR loci that were present in one inbred line which supports that SSR markers are efficient tool to identify germplasm for breeding programs. The genetic diversity ranged from 0.16 to 0.80 with average value 0.58. The cluster analysis showed five groups in which the two inbred from CIMMYT were clustered completely separate from Iran genotypes..

Hoxha*et al*. (2004) investigated the genetic diversity of twenty Albanian local maize genotypes using 20 SSR markers distributed throughout maize genome. They reported an average number of alleles per locus was 9.1 and polymorphic information content value 0.71. They concluded that this high genetic diversity in these maize genotypes may be fruitful in future as a good source for planning well-organized breeding strategies. They also confirmed that SSR loci in the maize genome are a *robust* approach towards the conservation and management of maize germplasm

Liu *et al*. (2003) studied the genetic structure and diversity of 260 maize inbred lines using 94 SSR loci present throughout maize genome. They reported a total of 2039 allele with an average of 21.7 per locus. On the basis of model-based clustering 5 groups were obtained. The phylogenetic tree showed results parallel to the pedigree record and cluster analysis. They found that tropical and subtropical genotypes showed maximum genetic

diversity than temperate lines. It was inferred that maize possesses marvelous allelic variability but not present in lines on random. The extent of diversity is not equal in all groups i.e. the presence of more genetic diversity in tropical and subtropical maize as compared to temperate and within a large gene pool the genetic diversity present in one part is present in the present day inbred lines e. g. the absence of tropical inbred lines in the present day maize inbred which created the opportunity to utilized in developing new maize varieties

Li *et al.* (2002) studied the genetic diversity of 58 maize genotypes and one teosinte by forty SSR markers and differentiated the molecular based clustering from conventionally made heterotic groups. A total of 259 allelic variants were detected with mean of 6.475 allelic variants per locus. Polymorphic information content value ranged from 0.14 and 0.89 with mean value of 0.60 and genetic similarity range from 0.44 to 0.571. The cluster analysis conducted through UPGMA using NTSYS separated the genotypes into eight clusters, except four genotypes did not cluster, while the teosinte grouped in to cluster eight. These markers are also helpful in the collection and conservation of accessions for successful breeding programs.

2.4. Genetic diversity evaluation

Genetic diversity analysis are important mechanisms of crop improvement programs (Mohammadi and Prasanna, 2003; Kitti*et al.,* 2012). Genetic diversity analysis is important to: (i) determine the extent of genetic variation present between germplasm (Smith, 1984; Cox *et al.,* 1986; Yunbi*et al.,* 2009), (ii) identify suitable parents for crosses and create segregating progenies with maximum genetic variability for further selection (Barrett and Kidwell, 1998; Bertan, 2007), (iii) introgress desirable genes from diverse germplasm sources into the available genetic base (Thompson *et al.,* 1998) and (iv) conserve unique and novel germplasm.. Critical assessment of the genetic relationships that may exist within and among the inbred lines or pure lines is useful to make successful cross combination, assigning lines according to heterotic groups, and for effective varietal protection (Melchinger*et al.,* 1990; Flint-Garcia *et al.,* 2009; Lu *et al.,* 2009).

Genetic diversity analysis enables grouping of a germplasm collections for specific breeding purposes (Hallauer and Miranda, 1988; Semagn*et al.,* 2012). Variation among individuals or groups of individuals or populations is analyzed using a specific method or

a combination of methods of genetic diversity study (Mohammadi and Prasanna, 2003). Usually data from phenotypic measurements and combinations of different types of variables are involved. Diverse data sets are collected to analyze genetic diversity in crop plants. These include pedigree data (Bernardo, 1993; Van Hintum and Haalman, 1994), morphological data (Bar- Hen et al., 1995), biochemical data obtained from isozymes analysis (Hamrick, 1997), analyses of storage proteins (Smith *et al.,* 1987) and DNA-based marker data allow more accurate differentiation of genotypes (Melchinger, 1999; Mohammadi and Prasanna, 2003; Demissew*et al.,* 2012). The choice of a particular technique(s) depends on the aim(s) of the study, the level of resolution required and availability of facilities (Mohammadi and Prasanna, 2003).

Genetic diversity analysis encompasses three levels of evaluation; it depends on the nature of the data collected and the genetic material that were used for the investigation. Normally, the data set may be morphological, isozyme and DNA-based markers assessed on diverse materials such as germplasm, accessions, inbred lines, or clones, populations and hybrids. Since each of these data sets provide different types of information. The choice of analytical method used depends on the objective(s) of the research, the level of resolution required, the resources and technological availability, facility, and operational time (Karp *et al.,* 1997).

The three levels of genetic diversity analyses are (i), quantification of genetic diversity (ii) and quantification of genetic relationships and (iii) expression of the relationships in terms of classification and/or ordination (Cavalli-Sforza and Bodmer, 1981). In different literature Various genetic diversity measurement are available that is number of alleles per locus or allele richness, level of polymorphism, observed and expected heterozygosity, effective population size (Cavalli-Sforza and Bodmer, 1981).

2.4.1. Quantification of genetic relationships

The basic measure of genetic diversity is genetic distance. Genetic distance is any quantitative measure of genetic difference at either the sequence or allelic frequency level that is calculated between individuals, populations or species on the basis of what they do not have in common. It may be expressed as dissimilarity measure arising from discrepancy in space and time through evolutionary changes, such as mutation, migration and genetic drift that have occurred since two populations existed as a single random

mating population. A small genetic distance indicates close relationship while large genetic distance indicates distant relationship (Beaumont *et al.,* 1998).

2.4.2. Determination of distance measures

2.4.2.1. Genetic similarity

There are numerous methods of genetic distance determination depending on the kind of data. That is morphological, allozyme, DNA based molecular markers data. Genetic distance measures include Euclidean distance; it is usually used for measuring agro morphological data. It is calculated as the root of the squared difference in traits between pairs of accessions i and j and Gower‟s distance for both quantitative and qualitative data (Gower‟s, 1971) and Bray-Curtis dissimilarity (Bray and Curtis, 1957). Various genetic distance measures are existed for gene frequency data, such as DNA bands or amplification products on SSR and RFLP gels. Few of these are the Cavalli-Sforza and Edwards (1967), Nei‟s distance (Nei, 1972) and Rogers (1972) distance. For nucleotide sequence data, the Jukes and Cantor (1969) distance is commonly used. In some instances where amplification products are converted to presence or absence binary data, the qualitative genetic distance measures frequently used include Nei and Li‟s (1979) coefficient, Jaccard‟s (1908) coefficient, simple matching coefficient (Sokal and Michener, 1958) and Modified Rogers‟ distance (Wright, 1978).

2.4.3. Expressing relationships in a genetic diversity study

Expressing relations in a genetic diversity study used classification methods, that assembly the entries genotype into clusters according to plant similarity and dissimilarity. Clustering methods can be hierarchical, non-hierarchical or overlapping. In hierarchical methods such as the Ward method (Ward, 1963), entries are organized into a tree or hierarchy where entries or groups are fused one at a time to entries or groups with the most similar patterns for all characters. In nonhierarchical methods such as the Gaussian Mixed model or Normix model (Wolfe, 1970), initial groups must be defined a priori and then the method improves the initial groups by an iterative process that results in a solution that corresponds to a maximum (global or local) of the likelihood function. With the Overlapping, individuals may belong to more than one group.

There are three methods of clustering which include clustering by simple linkage (or nearest neighbor), by complete linkage (or farthest neighbor) and clustering by average linkage (or UPGMA). Single linkage is fast which allows analyzing huge data sets. Besides it is also statistically reliable under many models of evolution. It produces a chaining effect which leads to poor resolution of individual groups complicating the interpretation of results. Complete Linkage produces very clear groups by using minimum values but sometimes tends to underestimate similarity between recognized clusters. UPGMA is the simplest method for constructing trees with a high level of accuracy. It is consistent in grouping biological data with relationships computed from different data sets. The greatest disadvantage of UPGMA is that it estimates the same evolutionary speed on all lineages, implying the rate of mutations is constant over time and for all lineages in the tree (Bodman*et al.,* 1981)

3. MATERIALS AND METHODS

3.1. Plant Material

Thirty seven elite maize inbred lines, including fourteen elite inbred lines developed by the National Maize Research Program (NMBRP) of Ethiopia and twenty three elite inbred lines developed by the International Maize and Wheat Improvement Center (CIMMYT) were used in this study. The fourteen inbred lines developed by the (NMBRP). ten of the CIMMYT inbred lines introduced previously by the NMBRP (Table 1) were received from Bako Agricultural Research Center (NMBRP) as part of the ongoing research program on evaluation of elite maize germplasm for tolerance to soil acidity *in vitro* being conducted at the National Agricultural Biotechnology Research Center (NABRC) at Holeta. The remaining thirteen CIMMYT inbred lines germplasm source from South America were developed for tolerance to soil acidity and were recently receive from CIMMYT-Mexico to use as standards in evaluating the Ethiopian germplasm for tolerance to soil acidity. Hence based on their pedigree, origin and adaptation the thirty seven inbred lines were divided into three populations; population 1 included 14 inbred lines developed by the NMBRP of Ethiopia, Bako. Population 2 included ten inbred lines of CIMMYT origin that were introduced long period ago and used in breeding program in Ethiopia and population 3 included thirteen elite inbred lines developed by CIMMYT from germplasm of South American origin for tolerance to soil acidity.

CML202, BKL004, 144-B, BKL001, 142-1-e, CML395, CML144, BKL003, CML444 MBRC5BCF108-2-3-1-B-B-B-BB, 35B-190-0-S10-2-1-2-2-1-2, A-7033 and PO, OOE-2-1-2-1 were identified as promising source of resistance against turcicum leaf blight in the mid agro ecology of Ethiopia. CML312, CML144, CML161, CML165, CML312 CML 334, CML536 were identified as promising source of resistance to turcicum leaf blight and gray leaf spot. BKL 004, 144-7-B and BKL 001 were identified for their resistance to gray leaf spot. CML202 and 142-1-e are resistance to turcicum leaf blight and gray leaf spot (Keno *et al.,* 2018). Inbred lines CML357, CML359, CML360, CML361, CML362, CML 363, CML364, CML365, CML366, CML435, CML436, CML438 and CML439 were obtained from CIMMIYT germplasm source from South America identified as promising acid tolerant lines. Thus, these inbred lines may be used in varietal development, disease management and to enhance productivity.

Table 1 : Pedigree and origin of the thirty seven maize inbred lines used in the study

No	Inbred lines	Pedigree	Origin / population
1	124-b(109)	Unknown (derived from Ecuador)	Bako (Ethiopia)
2	124-b (113)	Unknown (derived from Ecuador)	Bako (Ethiopia)
3	142-1-e	Unknown (derived from Ecuador _573)	Bako (Ethiopia)
4	144-7-b	Unknown (derived from Ecuador)	Bako (Ethiopia)
5	35B-190-0-S10-2-1-2-2-1-2	Unknown (derived from Ecuador)	Bako (Ethiopia)
6	A-7033	Unknown (derived from Ecuador)	Bako (Ethiopia)
7	BKL001	Unknown (derived from Ecuador)	Bako (Ethiopia)
8	BKL002	Unknown (derived from Ecuador)	Bako (Ethiopia)
9	BKL 003	Unknown (derived from Ecuador)	Bako (Ethiopia)
10	BKL 004	Unknown (derived from Ecuador)	Bako (Ethiopia)
11	F-7215	Unknown (derived from Kitale Syn. II)	Bako (Ethiopia)
12	MBRC5BCF108-2-3-1-B-	Unknown (derived from Ecuador)	Bako (Ethiopia)
13	PO,OOE3-2-1-2-1	Unknown (derived from Ecuador)	Bako (Ethiopia)
14	SC22	Unknown (derived from Ecuador)	Bako (Ethiopia)
15	CML144	P62C5F182-2-1-2BB-3-1	CIMMYIT (Mexico lowland)
16	CML161	G25QC18H520-1-1-1-25-3-B-1-BBBB	CIMMYT (Mexico lowland)
17	CML165	QF37SR-2-3SR-2-4-3-BBB	CIMMYT (Mexico lowland)
18	CML197	G34QH174-3-1-2-BB	CIMMYT (Mexico lowland)

19	CML 202	ZSR923-S4BULK-5-1-BBB	CIMMYT (Mexico lowland)
20	CML 312	S89500 F2-2-2-1-1-B*5	CIMMYT (Mexico lowland)
21	CML 334	S920-F47-2-1-2-1-BBBBB	CIMMYT (Mexico lowland)
22	CML 395	90323(B)-1-X-1-B-B-1-1-B-B-1-1-B	CIMMYT (Mexico lowland)
23	CML 444	P43C9-1-1-1-1-1-BBBBB	CIMMYT (Mexico lowland)
24	CML 536	ZM605C2F1-17-1-B-1-BB	CIMMYT (Mexico lowland)
25	CML 357	SA3-C4F5(6/24)-1-2-2-5-B	CIMMIYT (South America)
26	CML 359	SA3C4HC(16?25)-2-4-3-1-B	CIMMIYT (South America)
27	CML 360	SA4-C2-FS(21/26)-1-2-2-2-B	CIMMIYT (South America)
28	CML 361	SA4-C2-FS(21/26)-4-2-7-3-B	CIMMIYT (South America)
29	CML 362	SA5-FS1-3-9-1-3-4-B	CIMMIYT (South America)
30	CML 363	SA5-FS1-3-9-1-5-2-B	CIMMIYT (South America)
31	CML 364	SA5-FS1-5-1-1-5-3-B	CIMMIYT (South America)
32	CML 365	SA8- C1-FS(27/3)-1-1-4-8-B	CIMMIYT (South America)
33	CML 366	SA8-C2-FS(27/3)-1-3-6-1-B	CIMMIYT (South America)
34	CML 435	SA3-C4HC(16/25)-2-4-3-6-B-B-B-B-B	CIMMIYT (South America)
35	CML 436	SA3-C4-FS(19/25)-2-6-4-5-B	CIMMIYT (South America)
36	CML 438	SA4-C2-FS(21/26)-1-2-2-2-B	CIMMIYT (South America)
37	CML 439	SA5-C2HC(26/21)-4-3-7-5-B-B-B-B-B-B	CIMMIYT (South America)

3.2. Genomic DNA extraction

Five seeds from each inbred lines were grown in a greenhouse and fresh leaves were collected from 15 days old plants for gDNA extraction. Genomic DNA was extracted following DArT (Diversity array technology) DNA extraction method (https://www.diversityarrays.com/orderinstructions/plant-dna-extraction-protocol-for-dart/)

Mortar and pistil were washed and autoclaved at 121ºC for 20 minutes. The mortar was labeled with the name of the inbred line using permanent marker. After two week of planting the leave samples excised from five seedlings from each inbred line using sterile surgical blade and bulked. It was then placed in the mortar labeled with the name the inbred line. The mortar was then immediately firmly covered by aluminum foil and then placed in a refrigerator at -80ºc for overnight.

Thirty ml Extraction buffer, 30 ml lysis buffer, 10ml SDS and 1.2 gm of poly vinypyrrodine (PVP) were mixed in one flask. Then the mixed buffer solution was incubated in water bath at 65ºC. After the leaves were total dried the required amounts of plant materials were grounded and transferred to 2 ml sterile Eppendrof tube and incubated in water bath at 65ºC for 1hour. The samples were cool down for 5 min and 1 ml of chloroform isoamyl alcohol (24:1) mixture was added in each sample under biological safety cabinet and mixed well for 30 minute by gently inverting the tube. The samples were then centrifuged at10, 000 rpm for 20 minute and the supernatant was transferred to another 1.5 ml sterile Eppendrof tubes and waste was discarded. Then after the same volume of ice cold isopropanol was added in each tube, the samples were mixed by inverting the tube for ten times. The samples were then centrifuged again for 30 min at 10, 000 rpm. Finally the supernatant was discarded and the pellet was washed with 1.5 ml of 70% ethanol. Ethanol was then discarded by pouring slowly and the pellet was dried for 1 hour. Finally the DNA was dissolved in 250 µl of TE buffer.

3. 3. Genomic DNA quality and quantity measurement

The genomic DNA was assessed using 0.8% agarose gel electrophoresis. The concentration and quality of extracted genomic DNAwere also measured using Nano Drop 2000 UV Spectrophotometer (Thermo Scientific, Massachusetts, USA). The A260/A280

ratio was used to provide an estimate of DNA purity. The DNA samples with high band intensity, lesser smear, purity with 1.8 to 2 were used for further PCR analysis. Prior to PCR the genomic DNA was normalized by diluting the DNA concentration in to 10ng/µl.

3. 4. SSR primer selection

Twenty nine SSR markers having wider genomic coverage were obtained from the maize genomic data base (available at: http://www.maizegdb.org.php) were used in the study. These markers were selected based on genome coverage. The primer sequences, repeat motives and the bin of these markers are indicated in Table 2 below.

Table 2 :SSR markers used for study of genetic diversity in maize inbred lines

No	Marker	Primers(5'-------3')	Annealing	Repeat sequence	Bin	chromosom
1	Umc 2164	F- AGCACACAGACAAGAGAGACAACG R- GACCGACAACAGAGATCGAGTACA	58.2	(CGGC)8	5.05	5
2	Umc 1506	F- AAAAGAAACATGTTCAGTCGAGCG R- ATAAAGGTTGGCAAAACGTAGCCT	53.5	(AACA)4	10.05	10
3	Umc 1607	F- ACTAATTTTCGGTAGTCGTGTGCG R- GGAAAGAGAGAGGCTGTAGGTGGT	53.5	(TGC) 5	8.07	8
4	Umc 1137	F-TCAGTCACTCTTCTGCCTCCACT R-GGCTGGATAATGTTGTAGCTGGTC	52.5	(CT)15	9.08	9
5	Umc 2280	F- AAAAGAAGACGCTTTGTTTGTTGC R-TTTTCGTCAACTTGATGTTTATGAGAGT	58.3	(CATTA)4	4.03	4
6	Umc 1363	F- TGTTTAAGTGTTGGCAGAAAGCAA R- TCTCCCTCCCCTGTACATGAATTA	59.4	(ACG)4	1.01	1
7	Umc 1757	F- TTTTCTGCAGGGATAACATTTGTG R- ATAGGAGGTGAGGTGAGGAGGAAG	59.4	(TCC)7	4.01	4
8	Umc 1272	F- CTCTGACAGACCTGCAGATAGGGT R- ATCGAGGGGCTAATCAGCAAG	58.4	(CTAGC)4	10.04	10
9	Umc 1636	F- CATATCAGTCGTTCGTCCAGCTAA R- GTACTGGTACAGGTCGTCGCTCTT	58.4	(AGGC)4	9.02	9
10	Umc 1857	F- TTCCTTGCCAACAAATACAAGGAT R- GTTCATTGCTTCATCTTGGAACCT	55.8	(TAA)6	6.04	6
11	Umc 1470	F- AAAAACCTCAATAGCCGTTTCACA R- GATTCTTGTGTTGCATACTGGTGC	55.8	(TAA)7	8.04	8

12	Umc 2278	F- CTGACCTCCGTCATCAGCATC R- ATCACGGACAAAGAAAATTGAAGC	52.3	(TC)8	4.01	4
13	Umc 1003	F- AATAGATTGAATAAGACGTTGCCC R- TGTTCCAATGCTTTTGTACCTCTA	55.8	(TAAA)9	2.05	2
14	Umc 1913	F-GATCCTACCAAAATCTTATAGGC R- ACAGCTAGCCAAGATCTGATT	55,8	(TTG)6	8.02	8
15	Umc 2085	F- TGTACGACTTCTTCTGGACGCAC R- TAGATGTCGATGTCCTCCAGGG	57.6	(CGC)5	2.08	2
16	Umc 1075	F- GAGAGATGACAGACACATCCTTGG R- ACATTTATGATACCGGGAGTTGGA	57.6	(ATTGC)5	8.01	8
17	Blng 1063	F- GGAGACAACCCCGACGAC R- GGTACCAGAGCCACAGATCC	59.3	(AG)42	3.06	3
18	Umc 2080	F- GCCAAGGTGGGTCTGGCTAT R- ACCACCTTGTCCGTATCCTTCAC	59.3	(TGGCTC)4	1.08	1
19	Umc 1415	F- GTGAGATATATCCCCGCCTTCC R- AGACTTCCTGAAGCTCGGTCCTA	58.6	(GAC)10	8.04	8
20	Umc 2198	F- AGCCCAGAGAAGGGAAGCAG R- CTCTTCACTCGCTTCTCCCAGA	58.6	(CCCTC)4	5.06	5
21	Umc 2319	F- GATCCACGCGAGGTTCACTG R- GCTCTCACTAGCCTCGCATTCC	58.2	(GAGGAG)5	6.04	6
22	Bnlg 1927	F- TTTTTTTGTAAGCGATCCGG R- GATGAATCTGCGTCCGTCTT	55.5	(AG)41	4.07	4
23	Umc 1066	F- ATGGAGCACGTCATCTCAATGG R- AGCAGCAGCAACGTCTATGACACT	58,3	(GCCAGA)5	7.01	7
24	Umc 1904	F- CAGCCACTCGTTTATGGAGGTTTA R- TGTTACTAGTCGATCTGATGCCCA	58.7	(TAAGC)5	8.03	8
25	Umc 1639	F- CTAGCCAGCCCCCATTCTTC R- GCAAGGAGTAGGGAGGACGTG	58.7	(TGTCC)4	3.09	3
26	SSR 6	F- GATCCACGCGAGGTTCACTG R- GCTCTCACTAGCCTCGCATTCC	53.4	(CA)9	9.02	9
27	Umc 2294	F- ATTGGAGTGGCTCCATTGCTT R- CCCACCATTCTATATATTGTTGCCA	53.4	(TCCTG)4	5.03	5
28	Umc 2205	F- CATGATCATTTGGCGATGGTAAT	55	(TC)4	2.07	2

		R- ATGGTGAGCGAGTGAAAGAGAGAT			
29	SSR 14	F- AGGAGGTACCACAATGGAG R- GTGTACATCAAGGTCCAGATTT	52.7	(CA)16	8.09

3. 5. Amplification and detection of bands

Prior to normal or actual PCR all primers annealing temperature were optimized. Polymerase chain reaction (PCR) was carried out in total volume of 12.5 µl reaction containing 6.25 µl of one Taq 2x Master Mix (supplemented with all PCR reaction components such as MgCl2, PCR buffer, dNTPs, and TaqDNA polymerase), 3µl (10 ng/ µl) genomic DNA, 0.5 µl (10 Pico mole/ µl) of each of forward and reverse primer and 2 µl nuclease-free water. The reaction was carried out using thermal cycler Gene Amp® PCR system 9700 (Applied Bio system, USA) programmed at an initial denaturation of 94 °C for 4 minutes followed by denaturation at 94 °C for 30 seconds, annealing at 52.3 °C up to 59.3 °C (depending on the primer used) for 30 seconds, initial extension at 72 °C for 30 seconds and final elongation at 72 °C for 10 minutes. After amplification 8µl of the PCR product and 6µl of 100 + 50bp DNA marker ladder with known reference band was mixed with 2µl of loading dye in each well of the PCR plate and loaded in 4% (w/v) agarose gel and electrophoresis was run for three hours at 100 volts. Finally the gel was then visualized using gel documentation, 3UV-transilluminator (Bio-Doc).

3. 6. Scoring and data analysis

The amplified products were scored based on fragment band size comparing with 100+50 DNA ladder. Fragments with the same mobility were considered as identical fragments and treated as a unit character.

3. 7. Diversity parameters and analysis

Genetic diversity assessments of 37 elite maize inbred lines was analyzed by using GenAlex version 6.5 (Peakall and Smouse, 2015) software package. Genetic diversity parameters considered in this study were: number of allele per locus (Na), number of effective alleles per locus (Ne), Shannon information index (I), fixation index (F) (Nei's, 1978), gene flow (Nm) and percent polymorphism (% P). Observed Heterozygosity (Ho) and expected Heterozygosity (He) were analyzed using GenAlex version 6.5 (Peakall and Smouse, 2015) software package and polymorphic information content (PIC), was analyzed by using power marker v3, 25 (Liu and Muse, 2005) software.

3. 8. Analysis of molecular variance (AMOVA)

The analysis of molecular variance (AMOVA) was performed to estimate population genetic structure and differentiation among and within population based on their sources of origin. AMOVA uses the estimated F-statistics such as genetic differentiation (Fst), fixation index or inbreeding coefficient (Fis), and overall fixation index (Fit) to compare the genetic structure among and within populations. The AMOVA procedures were done using GeneAlex version 6.5 (Peakall and Smouse, 2015) software package

3. 9. Genetic distance and cluster analysis

To examine the degree of population differentiation among the study materials the Nei unbiased genetic distance were computed according to Nei's (1978) using GenAlex version 6.5 (Peakall and Smouse, 2015) software package. Pair-wise Fst and Nm values were also computed using the same software.

Cluster analysis was carried out by using neighbor-joining (NJ) algorithm in DARwin version 6 software (Perrier and Jacquemoud-Collet, 2010). A dendrogram for 37 elite maize inbred lines were generated based on the dissimilarity matrix to visualize pattern of cluster within and among the elite maize inbred lines.

3. 10. Principal coordinate analysis (PCoA)

Principal Coordinate Analysis (PCoA) was performed using the Jaccard's index to further validate the complementarity of clustering pattern revealed by the dendrogram using GenAlex version 6.5 (Peakall and Smouse, 2015) software package.

4. RESULTS AND DISCUSSION

4.1. Polymorphism and Polymorphic information content

The genetic diversity among 37 elite maize inbred lineswas evaluated using 29 SSR markers which were distributed across 10 maize chromosomes. Among 29 simple sequences repeat markers marker (Umc 1913) showed monomorphic patternand hence was excluded from the analysis and the rest 28 simple sequence repeat markers were showed polymorphic pattern and included in the analysis.Allele frequency (Table 3) varied from 0.34 to 0.95 with a mean of 0.55. Umc 1470 and Umc 2205 exhibited the highest and lowest allele frequency; respectively. This indicates the presence of high allelic variation in the marker loci. This variation happened might be due to evolutionary influencing forces such as mutation, migration, recombination, and selection and geneticDrift.

The polymorphic markers (Table 3) revealed a total of 104 alleles. The numbers of allele per locus generated by each marker varied from two to six with a mean of 3.71 alleles. The highest numbers of alleles (6) were detected in loci Umc 1066, Umc 1272 and Umc 2205, whereas the lowest number of allele (2) was detected in locus Umc 1415. It indicates the existence of allelic variation in the marker loci. The number of allele per locus obtained in the present study was higher than Lopes *et al.,* (2015) they reported 4 allele in Umc 1066 locus but in the present study Umc 1066 revealed 6 allele per locus. The difference might be due to genotype difference, the number of genotypes used in the study. Out of 28 SSR markers 12 showed four alleles each, 12 showed 3 alleles each, 3 showed six alleles each and 1 showed 2 alleles each.

The number of effective alleles (Ne) (is the number of alleles with equal frequency) detected varied from 1.11 to 3.53 with a mean of 2.04. The highest number of effective allele (3.53) was observed from marker Umc 2205, whereas the lowest number of effective allele (1.11) was observed from marker Bnlg1063. It indicates the existence of allele frequency variation in the marker loci, because the lowest allele frequencies have little contribution to effective number of allele and the highest allele frequency have greater contribution to effective number of allele.

Table 3: Genetic parameters of the 28 SSR markers used in the study of 37 Elite maize inbred lines obtained from threesources (populations).

No	Locus	Genetic Parameters								
		Major allele frequency	N_a	N_e	H_o	H_e	I	F_{is}	Nm	PIC
1	Umc 2164	0.51	3.00	2.32	0.00	0.55	0.87	1.00	2.24	0.55
2	Umc 1506	0.62	3.00	2.20	0.00	0.53	0.89	1.00	5.97	0.48
3	Umc 1607	0.40	3.00	1.81	0.00	0.45	0.64	1.00	0.54	0.58
4	Umc 1137	0.55	3.00	2.26	0.17	0.55	0.92	0.70	3.89	0.52
5	Umc 2280	0.41	4.00	2.33	0.02	0.47	0.84	0.95	0.52	0.63
6	Umc 1363	0.60	3.00	1.94	0.00	0.47	0.78	1.00	1.14	0.50
7	Umc 1757	0.57	4.00	1.95	0.05	0.48	0.72	0.90	1.24	0.51
8	Umc 1272	0.54	6.00	2.56	0.00	0.59	1.060	1.00	2.32	0.61
9	Umc 1636	0.43	4.00	2.43	0.00	0.58	0.99	1.00	5.08	0.54
10	Umc 1857	0.62	3.00	2.18	0.00	0.54	0.88	1.00	2.94	0.48
11	Umc 1470	0.35	3.00	1.39	0.00	0.24	0.37	1.00	0.14	0.59
12	Umc 2278	0.81	3.00	1.48	0.00	0.30	0.52	1.00	4.26	0.30
13	Umc 1003	0.65	4.00	1.55	0.00	0.32	0.53	1.00	0.53	0.41
14	Umc 2085	0.51	3.00	1.83	0.00	0.45	0.69	1.00	1.59	0.41
15	Umc 1075	0.65	4.00	1.97	0.00	0.47	0.73	1.00	2.01	0.50
16	Blng 1063	0.95	3.00	1.11	0.03	0.09	0.19	0.72	9.61	0.10
17	Umc 2080	0.35	4.00	1.72	0.00	0.38	0.66	1.00	0.31	0.63
18	Umc 1415	0.92	2.00	1.18	0.00	0.15	0.29	1.00	3.63	0.14
19	Umc 2198	0.65	4.00	1.87	0.10	0.44	0.75	0.76	1.63	0.47
20	Umc 2319	0.41	3.00	1.66	0.02	0.33	0.54	0.93	0.28	0.56
21	Bnlg 1927	0.57	4.00	2.10	0.00	0.52	0.81	1.00	16.85	0.44
22	Umc 1066	0.51	6.00	2.79	0.05	0.54	1.01	0.91	1.25	0.64

23	Umc 1904	0.38	4.00	2.11	0.03	0.52	0.81	0.95	0.86	0.60
24	Umc 1639	0.54	3.00	2.06	0.00	0.46	0.77	1.00	1.17	0.50
25	Umc 2319	0.65	4.00	2.06	0.00	0.45	0.76	1.00	1.87	0.50
26	Umc 2294	0.40	4.00	2.60	0.00	0.59	0.98	1.00	1.67	0.62
27	Umc 2205	0.34	6.00	3.53	0.03	0.61	1.24	0.95	1.05	0.74
28	SSR 14	0.50	4.00	2.19	0.00	0.53	0.88	1.00	1.24	0.60
Mean		0.55	3.71	2.04	0.02	0.45	0.75	0.96	2.71	0.50

N_a - number of observed alleles; N_e - Number of effective alleles; H_o - Observed Heterozygosity; H_e – expected Heterozygosity (Average gene diversity within genotypes); F_{is} - Fixation index; PIC - Polymorphic information content.

Gene flow (Nm) ranged between 0.14 to 16.85. The highest gene flow (Nm) was observed from marker Bnlg 1927 (16.85) and the lowest gene flow was observed from marker Umc 1470 (0.028) with a mean of (2.71). It indicates the presence of high gene flow in the marker loci. It leads to the present of high frequency of identical allele in the marker loci. Shannon's information (I) index ranged from 0.19 to 1.24. The highest value of Shannon's information (I) index was recorded for marker Umc 2205 (1.24), whereas the lowest value of Shannon information index was noted for marker Umc 1470 (0.19) with a mean value of (0.75). it indicated the present of species diversity in the marker loci.The observed heterozygosity (Ho) value varied from 0.00 to 0.17 with a mean of 0.02. The highest observed heterozygosity value observed from marker Umc 1137 (0.17), while the lowest observed heterozygosity (0.00) were detected from markers (Umc 2164, Umc 1506, Umc 1607, Umc 1363, Umc 1272, Umc 1636, Umc 1857, Umc 1470, Umc 2278, Umc 1003, Umc 2085, Umc 1075, Blng 1063, Umc 2080, Umc 1415, Bnlg 1927, Umc 1639, Umc 2319, Umc 2294, SSR14). This indicates that the loci reached at maximum homozygous state. This was expected based on the number of selfed generation in the development of the inbred lines, which was from five (S_5) to eight (S_8) generation and 98 per cent of homozugosity is expected in the inbred lines. The overall mean of observed heterozygosity (0.02) is lower than the average expected heterozygosity (0.45). This divergence is attributed by inbreeding that was applied in the development of the inbred lines.

Expected heterozygosity or gene diversity (He) of the markers ranged from 0.09 to 0.61 with a mean of 0.45.The highest expected heterozygosity (0.61) was observed from marker Umc 2205 whereas, the lowest expected heterozygosity (0.09) was observed from marker Bnlg 1063. The highest mean value of expected hetrozygosity observed in the current study indicates the presence of high allelic variation in the marker loci and their distribution in the inbred lines. Maize is highly cross pollinated crops as a result pollen or seed contamination during maintenance could be the probable reason of the presence of heterozygosity in the inbred lines. As a result of this, inbred lines tend to segregate for a few loci in spite of repeated cycles of selfing over many generations. In addition inbred lines used in the present study were S5 to S8 generation as result still there is segregation in few loci it leads to the present of heterozygosity in the inbred lines. The expected heterozygosity results obtained in

the present are closer to Nyaligwa*et al.* (2015) they reported an average of 0.51 expected heterozygosity in elite maize inbred lines.

Fixation index (Fis) varied from 0.70 to 1.00.The highest fixation index (1) were observed from markers Umc 2164, Umc 1506, Umc 1607, Umc 1363,Umc 1272, Umc 1636, Umc 1857, Umc 1470, Umc 2278, Umc 1003,Umc 2285, Umc 2085, Umc 1075, Umc 2080, Umc 1415, Bnlg 1927, Umc 1639, Umc 2319, Umc 2294 and SSR 14, whereas the lower fixation index (0.70) was observed from marker Umc 1137 with a mean value of 0.96.It indicates that the loci were fixed and the presence of high degree of inbreeding as a result of successive selfing over many generation or great reduction of heterozygosity and maximum enhancement of homozygosity in the marker loci.

Polymorphic information content (PIC) generated by each marker varied from 0.10 to 0.74 with a mean of 0.50. The highest PIC value (0.74) was observed from marker Umc 2205, whereas the lowest PIC value (0.10) was observed from marker Bnlg 1063. The present study was in line with Lopez *et al.* (2015) they reported PIC value of 0.71 for Umc 2205 marker. According to Botstein *et al.,* (1980), a marker with PIC value > 0.5 indicates highly informative of the SSR loci in detecting differences among the genotypes based on their genetic relationships. PIC value between 0.25 - 0.5 indicate moderate informative of the SSR loci and PIC value < 0.25 implies less informative of the SSR loci in detecting difference among the genotype. Based on this the present study (Table 3) out of 28 SSR markers two markers Bnlg 1063 and Umc 1415 revealed PIC value < 0.25. It indicates less discrimination ability of the markers for detecting genetic variation between the maize inbred lines. Thus, these markers have less application in genetic diversity study of maize in the future. Seven SSR markers revealed PIC value> 0.3.which indicate moderate informativeness of the markers and the rest 19 SSR markers revealed PIC value > 0.5. It shows the high discrimination ability of the markers for identifying genetic variation in maize inbred lines. As a result, these markers are convenient markers or highly applicable in genetic diversity study of maize in future. Moreover, it also suggests that the inbred lines were holding considerable amount of genetic diversity. The average PIC value observed in the current study was higher than Sharma *et al., (*2017) reported an average PIC value 0.36 in 33 maize inbred lines using 40 SSR markers and the average PIC value noted in this study was lower than

Gazal*et al.,* (2016) they reported an average PIC value of 0.78 by studying 19 Indian and 5 CIMMIYT maize inbred lines using 45 SSR markers. This difference happen might be due to the number of SSR loci, genotype difference and number of genotype used. Bantte and Prassana (2003) had reported that the overall PIC value could be influenced by several factors, mainly (1) the nature of germplasm used for the study, (2) number of SSR loci as well as inbred lines analyzed, (3) SSR loci assayed, in terms of the nature and type of repeats and (4) methodology employed for allele detection (agarose vs. PAGE).

4.3. Analysis of molecular variance(AMOVA)

Analysis of molecular variance (AMOVA) partitioned the total molecular variance into within and among the sets of varieties evaluated based on their source of genetic origins (Nyaligwa*et al.,* 2015). Analysis of molecular variance (Table 4) revealed that there were highly significant molecular variances (P<0.001) among population (AP) and among individual (AI) within populations. The highest percentage (77 %) of the variation was attributed to genetic variability among individuals (AI) within populations, Several studies have shown that mixed-mating and out crossing species have less than 25% of their genetic variation among populations or groups and the remaining within populations or groups (Hamrick and Godt 1997; Huff *et al.* 1993) The existence of greater variation within than between heterotic groups may also be attributed, in part, to the mix up of germplasm during the conversion process which, therefore, necessitates the establishment of new heterotic groups (Nesbitt *et al.,* 1995). While (23 %) of the variation was attributed to genetic variability among the population (AP). This result in agreement with Lopez *et al.* (2015) they reported 75 % of the genetic variation observed with in individuals in the population and 25 % of variation was observed from among population in 24 sweet corn inbred lines. The current study also support the result of Hinze*et al.* (2005) they reported that 78% of genetic variation was observed with in individuals in the population and 22% of the variation was noted from among the two population (Iowa Corn Borer Synthetic and Iowa Stiff Stalk Synthetic).The analysis of molecular variance result obtained in the present study also in agreement with the result of Nyaligwa*et al.* (2015) they reported 67 % of the genetic variation was attributed by genetic variation among individuals within the source of collection, while 33% of the variation was attributed to among population.

The magnitude between and within population differentiation were quantified using F-statistics (Fst, Fis and Fit), also known as fixation indexes. The latter term was coined by Wright (1951) to describe the properties of subdivided populations or with in the population. According to Wright (1951), when Fst value between 0 to 0.05 implies the existence of small genetic differentiation. Fst value between 0.05 to 0.15 indicates the presence of moderate genetic differentiation. Fst value between 0.15 to 0.25 indicate the presence of large genetic differentiation, Fst value > 0.25 indicate the presence of very large genetic differentiation. Based on this the result (Table 4) revealed high (Fst=0.197) genetic differentiation among the populations with the effect of high degree of inbreeding among individuals with in the population (Fis = 0.963). So it is very important for exploiting the variation for further genetic improvement in the crop in the future.

Table 4 : Analysis of molecular variance (AMOVA) based on standard permutation across the full data set of elite maize inbred linesfrom different source

Source	df	SS	MS	Est.	%	P	F-statistics
Among Populations	2	110.47	55.24	1.70	23%	>0.001	Fst =0.197
Among individuals with in populations	34	463.11	13.62	6.68	77%	>0.001	Fis = 0.963
Total	73	583.08		8.64	100%		Nm = 1.08

df = degree of freedom, SS=sum of squares and MS=mean squares

4.4. Genetic differentiation and gene flow between populations

The pair-wise genetic differentiation among inbred lines within the source of origin ranged from 0.146 to 0.272 (Table 5). The highest genetic differentiation (0.272) was observed between inbred lines developed at CIMMIYT germplasm source from lowland Mexico and inbred lines developed at CIMMIYT germplasm source from South America. It might be due to low gene flow between these two populations, while the lowest genetic differentiation (0.146) was observed between inbred lines of Bako and previously introduced inbred lines developed at CIMMYIT germplasm source of lowland Mexico. This low genetic differentiation among population may be due to gene flow that resulted from the movement of pollen during maintenance. This leads to an increase in the distribution of alleles among different populations regardless of their geographical distance.

The gene flow ranged from 0.325 to 1.12. According to Waples (1987), Nm values grouped into three categories: When Nm > 1.00 indicates high gene flow, Nm between 0.25-0.99 implies intermediate gene flow and Nm between 0.000 – 0.249 indicates low gene flow. Based on these inbred lines developed at Bako and inbred lines developed at CIMMIYT germplasm source from lowland Mexico possessed the highest gene flow (1.12).Because the two populations were maintained as the same breeding center as a result there is exchange of genetic materials from one population to the others. It clues thepresence of identical alleles between in these two populations,whereas the lowest gene flow (0.325) was observed between inbred lines developed at CIMMIYT germplasm source from lowland Mexico and inbred lines developed at CIMMIYT germplasm source from South America. It indicates the existence of less identical alleles and high genetic variation between these two populations.

Table 5 : Pair-wise genetic differentiation (FST) below the diagonal and pair-wise estimate of gene flow (Nm) above the diagonal between maize genetic sources

Inbred line source	Bako	CIMMIYT (lowland	CIMMIYT South
		1.12	0.445
Bako			0.325
Lowland Mexico	0.146		
South America	0.183	0.272	

4.5. Pattern of genetic diversity and relationship among populations

The mean numbers of allele (Table 6) in the populations were varied from 1.75 to 3.21. The highest (3.21) mean number of allele was recorded from Bako populations, whereas the lowest mean number of allele was detected from (CIMMIYT) lowland Mexico population . This might be happening due to different number of inbred lines were incorporated or considered in the three populations. i.e. Bako population contain 14 inbred lines while CIMMIYT (lowland Mexico population) contain 10 inbred lines) and CIMMIYT (South America population contain 13 inbred lines). Moreover Slight fluctuation in the number of alleles in the populations may be due to the selection of different SSR markers and the difference in the genetic diversity level of the experimental material.

The mean numbers of effective allele (Table 6) in the populations were ranged between 1.79 to 2.42. The highest mean number of effective allele was observed from inbred lines of Bako, whereas the lowest mean number of effective allele was recorded from inbred line of (CIMMIYT) lowland Mexico. It indicates the present of allele frequency variation in the populations. The highest number of effective allele observed in inbred lines of Bako indicate the existence of equally frequented alleles in the inbred lines because high equally frequented allele have a great contribution for effective number of allele and low allele frequency have little contribution to effective number of allele. The CIMMIYT (lowland Mexico) population possessed less equally frequented allele. The existence of different number of effective allele observed in the populations also indicated the present of genetic variation.

Table 6 : Summary of genetic parameters for three populations using 28 SSR markers

Populations	Genetic parameters							
	N	N_a	N_e	I	H_o	H_e	F	%
Bako	14	3.214	2.42	0.936	0.028	0.538	0.949	100 %
Mexico	10	2.321	1.792	0.624	0.014	0.390	0.957	92.86 %
South America	13	2.679	1.914	0.701	0.011	0.419	0.966	96.43 %
Mean	13.33	2.738	2.042	0.754	0.018	0.449	0.957	96.43%
SE	0.187	0.096	0.082	0.037	0.005	0.020	0.012	2.06

N=Number of inbred lines; N_a=Number of alleles; Ne=Number of effective alleles; I=Shannon's information index; H_o=Observed heterozygosity; H_e=Expected heterozygosity; F=Fixation index; %=Percentage of polymorphic loci; SE=Standard error

The mean numbers of Shannon information index in the populations were varied from 0.62 to 0.94. The highest (0.94) Shannon information index was noted from inbred lines of Bako, whereas the lowest (0.62) number Shannon information index was observed from inbred lines of CIMMIYT (lowland Mexico). It implies the present of high species diversity in the studied populations. In this study inbred lines obtained from Bako showed the highest mean number allele, the highest mean number of effective allele and the highest mean number of Shannon information index, comparatively, inbred lines sourced from low land Mexico revealed the lowest mean number of allele, lowest mean number of effective allele and the lowest mean number of Shannon information index. The differences could be due to small and large

number of inbred lines considered in the inbred lines of CIMMIYT (lowland Mexico) and inbred line of Bako respectively.

The highest (0.028) mean number of observed heterozygosis was observed from inbred lines sourced from Bako, while the lowest mean number (0.10) of observed heterozygosity was observed from inbred lines from South America germplasm source. The differences might due to small number of selfing (S_5) was applied in the development of inbred lines of Bako. As compared to inbred lines developed at CIMMIYT germplasm source obtained from South America, which was S8 generation. Inbred lines developed in Ethiopia are expected to experience less number of inbreeding cycles as compared to inbred lines developed at CIMMIYT. The overall mean value of observed heterozygosity (Ho) 0.02 was lower than the corresponding average of expected heterozygosity 0.449, which indicate the overall gain in homozygosity within the inbred lines. This can be confirmed by visual observation on the gel of specific markers which revealed single band.

The mean number of expected hetrozygosity (He) varied from 0.39 to 0.54. The highest mean number of expected heterozygosity was noted from inbred lines of Bako, because the developed inbred lines were S5 generation. While the lowest number of expected heterozygosity was revealed from inbred lines of CIMMIYT (lowland Mexico). This might be due to the number of selfed generation in the development of inbred lines. This was S8 generation. Moreover the difference existed in the populations indicate the presence of genetic variation. Some inbred lines revealed more than one band during amplifications, which may have resulted from the co-dominant nature of the SSR markers (Bantte and Prasanna 2003). Similar results have been previously reported in maize inbred lines (Helentjaris*et al.* 1988; Senior *et al.* 1998; Matsuoka *et al.* 2002; Liu *et al.* 2003). These investigators speculated a number of probable causes for the occurrence of double bands (heterozygosity) in maize inbred lines, including residual heterozygosity, pollen or seed contamination during maintenance, mutation at specific SSR loci, or amplification of similar sequences in different genomic regions due to duplication.

The present study found an average of 0.45 expected heterozygosity in the inbred lines. It shows higher than the observed heterozygosity values after 5 generations of selfing. The

present study closer to the finding of Semagn*et al.* 2012a). They reported 78% of the inbred lines showed high level of heterozygosity. This happened due to human errors (e.g. seed admixture, pollen contamination, mislabeling of seed sources and mixing of different seed stocks for planting) (Semagn*et al.,* 2012a). Small changes in allele frequencies may occur during seed Regeneration, bulking during maintenance breeding and possible contamination with Seeds or pollen of other samples (Warburton *et al.,* 2010). However, large proportions of heterogeneity can significantly change the Uniformity and performance of hybrids and in the worst case may result in the distribution of wrong hybrids. Consequently, additional generations of purification for all lines with higher proportion of heterogeneity are essential. (Heckenberger*et al.,* 2002

The mean number of Fixation index (F) in the populations was ranged from 0.949 to 0.966. The highest mean number of fixation index was observed from inbred lines of South American germplasm source, it is due to high number of inbreeding (S_8) was applied in the development of the inbred line as a result the loci were fixed. Whereas, inbred lines from Bako showed the lowest fixation index (0.949). It might be due to small number of inbred cycle was applied in Bako population with a mean of (0.957). The highest fixation of alleles observed in the present study may be the result of random changes in allele frequencies, where genetic drift acts on and leading to the fixation of alleles. The overall mean of fixation index indicate the existence of high degree of inbreeding and higher reduction of heterozygosity and maximum enhancement of homozygosity in the inbred lines due to selfing. Selfing is a powerful inbreeding system that allows the attainment of high level of homozygosity with few generations and simultaneously decreaseshetrozygosity in the inbred lines. The low mean number of observed heterozygosity and high mean number fixation index observed in inbred lines of South America germplasm indicated the occurrence of maximum level of inbreeding and high level of homozygosity in the inbred lines.

The degree of polymorphism varied from population to population. It ranged from 96.43 % to 100 %. Inbred lines sourced from Bako revealed the highest degree of polymorphism (100 %), followed by inbred lines obtained from South America (96.43 %), whereas the lowest degree of polymorphism noted from inbred lines obtained from CIMMIYT (Lowland Mexico)

(92.86%) with the average polymorphism of 96.43 %. The result indicates the existence of Wider genetic variations between the populations.

4.6. Genetic distance and relationship among populations

The genetic relationship among populations can be evaluated using genetic distance measurements. There are so many methods developed so far to examine genetic distance between populations (Cavalli-Sforza and Edwards, 1967; Nei, 1972, 1978; Takezaki and Nei, 1996). In the current study, Nei's (1978) method was used to estimate the genetic distance between the test maize inbred lines.

Genetic distance range from 0.27 to 0.42 (Table 7). The highest genetic distance (GD=0.42) was noted between inbred lines developed at CIMMIYT germplasm source from low land Mexico and inbred lines developed at CIMMIYT germplasm source from South America. The second largest genetic distance (GD=0.34) was observed between inbred lines obtained from Bako and inbred lines developed at CIMMIYT germplasm source from South America, whereas the smallest genetic distance (GD=0.27) was observed between inbred lines from Bako and inbred lines developed at CIMMYT germplasm source from lowland Mexico, which have been used in breeding program in Ethiopia. The result indicates the existence of wider genetic distance between inbred lines obtained from the two CIMMIYT populations. Therefore it could be important for genetic improvement of the crop by selecting parental lines from these two distant populations.

Table 7: Pair-wise population Nei's genetic distance showing the magnitude of genetic Differentiation between elite maize inbred lines populations

Genotype Source	Bako	CIMMIYT(Mexico)	South America
Bako	0.00		
CIMMIYT (Mexico)	0.27	0.00	
South America	0.34	0.42	0.00

4.7. Allelicpattern across population

Number of different allele frequency > = 5% ranged from 2.32 -3.18. The highest (2.32) number of different allele frequency > = 5% (Table 8) was noted from inbred lines of Bako. It indicates the existence of minor allele (less common allele) in Bako population, while the lowest number of different allele frequency > = 5% was observed from CIMMIYT (lowland Mexico) population. It indicates the present of minor or less common allele in the CIMMIYT (lowland Mexico) population.

Simple sequence repeat markers used in the current study showed tremendous discriminative performance during the estimation of genetic diversity among maize inbred lines and resulted in the identification of private allele in the populations. The private alleles ranged from 0.036 to 0.5. Bako populations showed the highest number (0.5) of private alleles (Table 8). These populations could be used as a source of important traits in the future maizebreeding programs because the private alleles provide a unique genetic variability in certain loci. In contrast the smallest number (0.036) of private allele noted from CIMMIYT (lowland Mexico) maze inbred lines. The present of private allele

Number of local common allele frequency < = 25% were ranged from 0.05 to 0.62. The highest number of local allele frequency < = 25% was observed from Bako population, while the lowest number of local allele frequency < = 25 % was noted from CIMMIYT (lowland Mexico) population. These results indicate the present of genetic variation in the populations. Bako population possessed the highest (2.10) local common allele frequency < = 50, while CIMMIYT (lowland Mexico) population exhibited the lowest (0.5) local number of allele frequency < = 50% .it indicate the existence of common allele frequency in the populations. It leads to the present of genetic variation in the populations. The present finding indicates that Bako Populations showed an observable variation in loci carrying private alleles indicating the existence of a high genetic uniqueness.

Table 8 : Mean of important allelic values recorded in three maize population groups

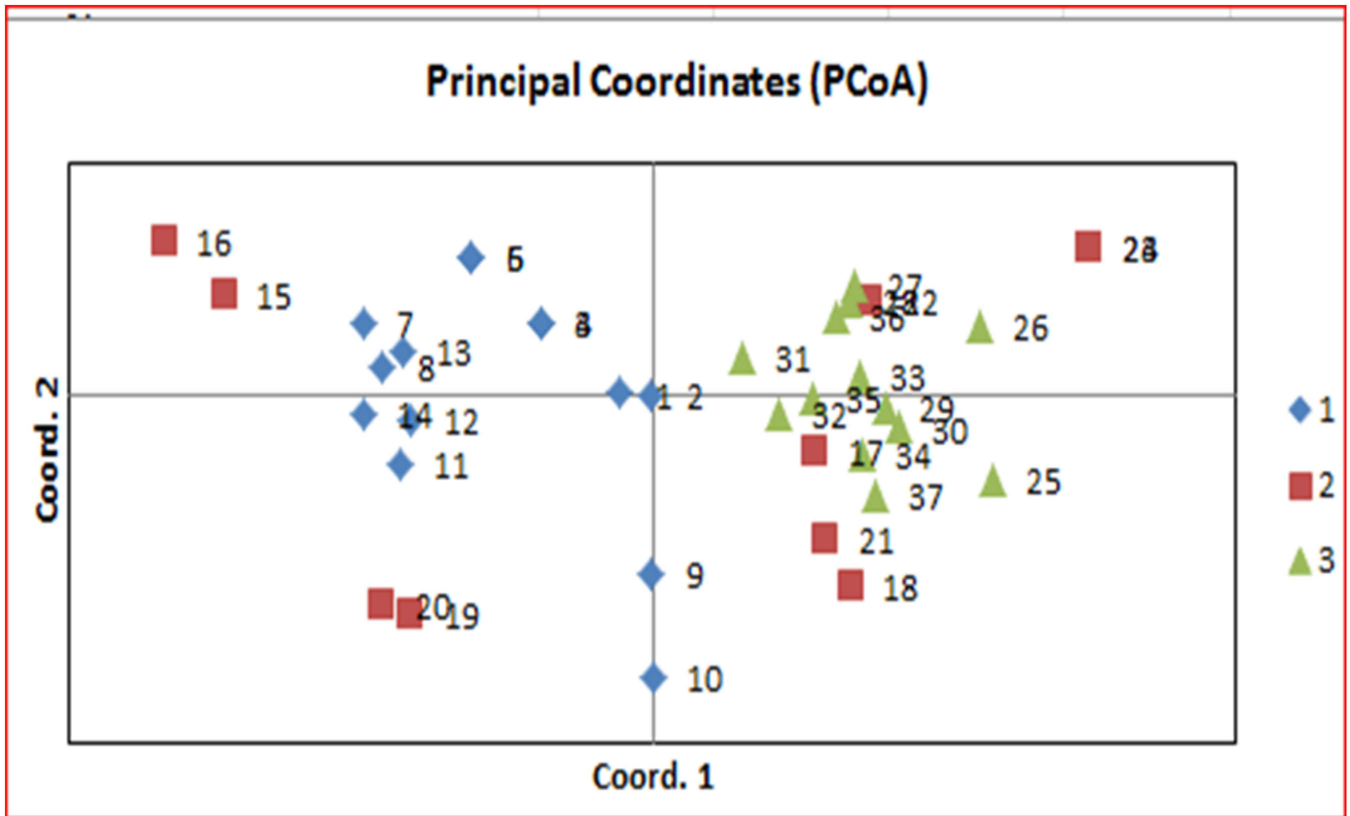
Parameter	Populations		
	Bako	CIMMIYT(lowland Mexico)	CIMMIYT(South America)
Na freq.> = 5%	3.18	2.32	2.57
Number of private allele	0.5	0.036	0.36
No of local common allele (<=25%)	0.62	0.05	0.5
No of local common allele (<=50%)	2.1	0.5	1.75

Na = No. of Different Alleles; Na (Freq>= 5%) = No. of Different Alleles with a Frequency >= 5%; Ne = No. LComm Alleles (<=25%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 25% or Fewer Populations; No. LComm Alleles (<=50%) = No. of Locally Common Alleles (Freq. >= 5%) Found in 50% or Fewer Populations

4.8. Principal coordinates analysis (PCoA)

Principal Co-ordinate Analysis (PCoA) (Table 9) showed that the three most informative principal coordinates accounted about 69.27% of the genetic variation present in SSR molecular data derived from maize inbred lines in the study. The first, second, and third principal coordinates explained about 42.69 %, 13.99 % and 12.59 %, respectively of the gross variation. The pattern of distribution of the inbred lines in the PCoA plot revealed three major clusters in the two-dimensional coordinates (Figure 1).

The PCoA analysis in the two-dimensional plot (Figure 1) showed that inbred lines from different source often clustered together. There was no separate group formed by a single population. This, in turn, agrees with the results of the dendrogram.In principal coordinate analysis inbred lines scattered in the same coordinate indicated the presence of closer genetic relationship and narrow genetic distance between them and inbred lines scattered in different coordinate indicates the presence of wider genetic variation. So, based on these the result of principal coordinate analysis showed the existence of wide genetic variation in the inbred lines and importance for further genetic improvement.



1 (BAKO)2 (CIMMIYT) lowland Mexico 3 (CIMMIYT) South America

Figure 1: Two dimensional plot of PCoA of 37 maize inbred lines and five populations based on SSR data

4.9 Clustering analysis

Dendrogram revealed 37 elite maize inbred lines were grouped into three major clusters Consisting of 32.43%, 35.14% and 32.43% of the total populations in clusters I, II and III respectively and forming different hierarchical sub-group (Figure 2). The distribution of inbred lines in the three clusters was not homogenous. Cluster I comprise inbred lines from Bako and CIMMIYT, while Clusters II constituted inbred lines from the same genetic sources or inbred lines sourced from South America. So, this population is genetically distant from other population and cluster III contain inbred lines from Bako and CIMMIYT. This is due to the presence of gene flow between these two populations. Cluster I was further divided into two subgroups. It comprised inbred lines of CML 444,CML 536,F-7215,CML 395, CML 161,CML 165,CML 197,CML 312,MBRC5BCF, PO,OOE3-2-1-1-2-1(subgroup CIA) from CIMMIYT and Bako population and CML 334, SC 22 (sub group CI B) from CIMMIYT

and Bako population. Cluster II contain inbred lines sourced from South America alone. These populations are genetically distinct from other population and important for further improvement. It comprised CML 357, CML 395, CML 360, CML 361, CML 362, CML 363, CML 364, CML 365, CML 366, CML 435, CML 436, CML 438, and CML 439. Cluster III also divided into two sub groups. It comprised inbred lines of 124-b(109),124-b(113),142-1-e,144-7-b,35B-190-0510-2-1-2-2-1-2 and A-7033 from Bako population(subgroup CIII A) and CML 161,BKL 001, BKL 002,CML 144,BKL 003 and BKL 004 (subgroup CIII B) from CIMMIYT and Bako population.

The distribution of Ethiopian as well as CIMMIYT inbred lines into the same clusters in the light of SSR markers based information verified that inbred lines obtained from different source have no common ancestors in their own location. It may also be inferred that these inbred lines have evolved in environments with slightly different climatic conditions. Such distribution also indicated that gene flow among these inbred lines might have occurred (Hoxha*et al.*, (2004). The cluster analysis indicated genetic diversity identified in the present study is independent of the effect of geographical distribution because inbred lines from the two population were distributed in two clusters, which showed that the studied materials possessed enough genetic variation.

Moreover, inbred lines from different populations clustered together, which may imply the existence of gene flow between and within populations. Bako populations were grouped with geographically distant inbred lines in Cluster II. Inbred lines from CIMMIYT (lowland Mexico) also grouped in all Clusters with other geographically distant populations. This indicates inbred lines in one cluster might be evolved from different lines of ancestry.

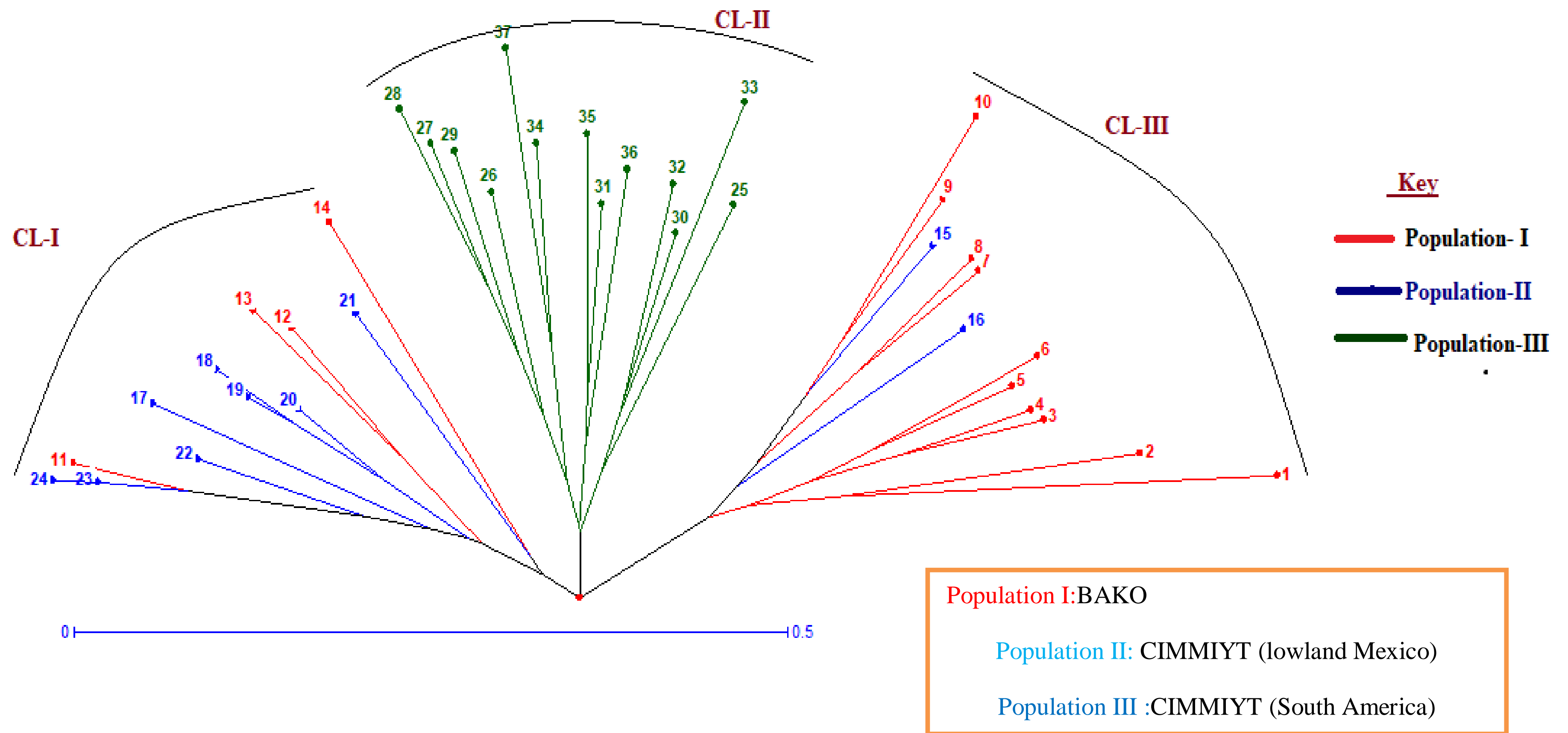


Figure 2: Unweighted Neighbor Joining (NJ) dendrogram showing genetic relationship of 37 *maize inbred lines* using 28 SSR markers

5. SUMMARY AND CONCLUSSION

Maize is one of globally top ranking cereal crops in productivity and has worldwide significance as a human food, animal feed and as source of large number of industrial products. More than 1.2 billion people in Sub-Saharan Africa and Latin America consume maize. In Ethiopia it is the second most popular staple food crop next to teff. However, maize production in Ethiopia is far below the potential due to several factors such as, limited availability of improved and wide adapted varieties, drought, acidity, insect and disease. Therefore, the present study was conducted to determine genetic distance among Ethiopian and CIMMIYT (lowland Mexico) maize inbred lines and recently introduced acid soil tolerant South America maize inbred lines using simple sequence repeat markers. The research was conducted in molecular biotechnology laboratory at Holeta Agricultural Research Center (HARC). Thirty seven elite maize inbred lines were used in the study.

For the current study 29 simple sequence repeat markers distributed in 10 maize chromosomes were used from maize genomic data base. Out of 29 simple sequences repeat markers one marker showed monomorphic pattern hence excluded and the rest 28 markers identified a total of 104 alleles with a mean of 3.71 alleles per locus. The overall mean of polymorphic information content (0.5) of the marker showed the potential discriminate ability of the SSR loci in detecting difference among the inbred lines. The overall mean of observed heterozygosity (Ho) 0.02 was lower than the average of expected heterozygosity 0.449, which indicate the presence of overall achievement in homozygosity within the tested maize inbred lines

The analysis of molecular variance (AMOVA) showed the existence of high molecular variance in the studied materials. The highest genetic variation (77%) was observed among individual within population. Inbred lines of lowland Mexico and inbred lines of South American populations showed the existence of wide genetic variation and inbred lines of Bako and CIMMIYT (lowland Mexico) showed narrow genetic variation. The highest Fst value was observed in inbred lines of CIMMIYT (lowland Mexico) and inbred lines of South America whereas the lowest Fst value was observed in inbred lines of Bako and inbred lines

of CIMMIYT (lowland Mexico) population. The highest genetic distance was observed between inbred lines of CIMMIYT (lowland Mexico) and inbred lines of South America, while the lowest genetic distance was observed between inbred lines of Bako and inbred lines of CIMMIYT (lowland Mexico).

The fixation index (F) was ranged from 0.949 to 0.966. The overall mean of fixation index indicate high level of inbreeding and higher reduction of heterozygosity and maximum enhancement of homozygosity due to selfing over many generation. The highest gene flow (0.652) was observed between inbred lines of Bako and inbred lines of CIMMIYT (lowland Mexico). It indicates the presence of identical alleles between two populations, whereas the lowest gene flow (0.325) was observed between inbred lines of CIMMIYT (lowland Mexico) and inbred lines of South America. It shows this two population are different. The highest number of private alleles was identified in Bako population. Thus this population is important for genetic improvement.

The principle coordinate analysis revealed that inbred lines from different population clustered together and also showed inbred lines from the same population scattered in different coordinate. It indicates the presence of genetic variation in the inbred lines. The dendrogram showed the admixtures of inbred lines in cluster I and Cluster III. Generally, this study indicated the existence of wide genetic variation among the inbred lines. Thus, it provided precise information to use promising combination for exploitation of heterosis and establishment of heterotic group as source materials in maize breeding program.

Further studies need to focus on the following points

- For deepest genetic diversity study use more robust molecular marker like SNP is recommended.

More number of maize inbred lines should be used to see the potential application of the marker system in genetic diversity of maize in the future.

6. REFERENCE

Arumuganathan, K. and Earle, E.D., 1991. Nuclear DNA content of some important plant species. *Plant molecular biology reporter*, ***9*****(3):**208-218.

Aylor, D.E., Baltazar, B.M. and Schoper, J.B., 2005. Some physical properties of teosinte Zea mays subsp. parviglumis) pollen. *Journal of Experimental Botany*, ***56*****(419):** 2401-2407.

Azam, M.G., Islam, S., Hossain, M.G. and Rohman, M.M., 2018. Molecular assessment of maize inbred lines (Zea mays L.) Using microsatellite markers.*Bangladesh Journal of Agricultural Research*, ***43*****(4):**533-542.

Bantte K, and Prasanna, B.M., 2003. Simple sequence repeats polymorphism in Quality Protein Maize (QPM) lines. *Euphytica*, ***129*****(3):**337-344.

Barrett, B.A. and Kidwell, K.K., 1998. AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest.*Crop science*, ***38*****(5):**1261-1271.

Beaumont, M.A., Ibrahim, K.M., Boursot, P. and Bruford, M.W., 1998. Measuring genetic distance.In *Molecular tools for screening biodiversity* ***16*****:**315-325.

Bennetzen, J.L., Ma, J. and Devos, K.M., 2005. Mechanisms of recent genome size variation in flowering plants. *Annals of botany*, ***95*****(1):**127-132.

Bernardo, R., 1993. Estimation of coefficient of co ancestry using molecular markers in maize.*Theoretical and Applied Genetics*, ***85*****(8):**1055-1062.

Bertan, I., Carvalho, F.I.F. and Oliveira, A.D., 2007.Parental selection strategies in plant breeding programs.*J Crop SciBiotechnol*, ***10*****(4):**211-22.

Beyene, Y., Botha, A.M. and Myburg, A.A., 2006. Genetic diversity among traditional Ethiopian highland maize accessions assessed by simple sequence repeat (SSR) markers. *Genetic resources and crop evolution*, ***53*****(8):**1579-1588.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, ***32*****(3):**314-317

Bray, J.R. and Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs* ***27*****(4):**325-349.

Buckler IV, E.S. and Thornsberry, J.M., 2002. Plant molecular diversity and applications to genomics. *Current opinion in plant biology*, ***5*****(2):**107-111.

Buckler, E.S., Thornsberry, J.M. and Kresovich, S., 2001.Molecular diversity, structure and domestication of grasses. *Genetics Research*, ***77*****(3):**213-218.

Cavalli-Sforza, L.L. and Edwards, A.W., 1967. Phylogenetic analysis: models and estimation procedures. *Evolution*, ***21*****(3):**550-570.

Choukan, R. and Warburton, M.L., 2005. Use of SSR data to determine relationships among early maturing Iranian maize inbred lines. *Maydica*, ***50*****(2):**163-170.

Cömertpay, G., Baloch, F.S., Kilian, B., Ülger, A.C. and Özkan, H., 2012. Diversity assessment of Turkish maize landraces based on fluorescent labeled SSR markers. *Plant Molecular Biology Reporter*, ***30*****(2):**261-274.

Condit, R. and Hubbell, S.P., 1991. Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome*, ***34*****(1):**66-71.

Cox, T.S., Murphy, J.P. and Rodgers, D.M., 1986.Changes in genetic diversity in the red winter wheat regions of the United States.*Proceedings of the National Academy of Sciences*, ***83*****(15):**5583-5586.

CSA (Central Statics Agency). 2018. The Federal Democratic Republic of Ethiopia Central Statistical Agency report on area and production of major crops. Addis Ababa, Ethiopia.

DemissewD.A., 2014. Genetic diversity and combining ability of selected quality protein maize (QPM) inbred lines adapted to the highland agro-ecology of Ethiopia (Doctoral dissertation, University of KwaZulu-Natal, Pietermaritzburg

Ellneskog-Staam, P., Henry Loaisiga, C. and Merker, A., 2007. Chromosome C-banding of the teosinteZeanicaraguensis and comparison to other Zea species.*Hereditas*, ***144*(3):** 96-101.

Ertiro, B.T., Semagn, K., Das, B., Olsen, M., Labuschagne, M., Worku, M., Wegary, D., Azmach, G., Ogugo, V., Keno, T. and Abebe, B., 2017. Genetic variation and population structure of maize inbred lines adapted to the mid-altitude sub-humid maize agro-ecology of Ethiopia using single nucleotide polymorphic (SNP) markers. BMC genomics, **18(1):**777.Ethiopian commodity exchange.

Eyre-Walker, A., Gaut, R.L., Hilton, H., Feldman, D.L. and Gaut, B.S., 1998.Investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences*, ***95*(8):**4441-4446.

FAOSTAT.(2017). Production of commodity in selected country, production share by region and production of top 5 producers. [Available online] http://faostat3.fao.org[Date accessed: 26/7/2016 at 5:30 pm].

Flint-Garcia, S.A., Buckler, E.S., Tiffin, P., Ersoz, E. and Springer, N.M., 2009.Heterosis is prevalent for multiple traits in diverse maize germplasm. *PloS one*, ***4*(10):**7421-7433.

Garcia, A. F., Benchimol, L. L., Barbosa, A. M. M., Geraldi, I. O., Souza, C. L. and de Souza, A. P. (2004). Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. Genetics and Molecular Biology **(4):** 579–588

Gibson, J. and Benson, B., 2002. Origin and History of Corn: The Botany of Tropical Crops.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. (2015). Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. Genetics Research International, **3:**1-14

Gower, J.C., 1971. A general coefficient of similarity and some of its properties. *Biometrics*, pp.857-871.

Gupta, A.S.T.H.A. and Singh, A.K., 2010. Studies on molecular diversity in maize inbred using SSR markers. Trends Biosci, **3(2):**106-109.

Hake, S. and Walbot, V., 1980.The genome of Zea mays, its organization and homology to related grasses.*Chromosoma*, ***79*****(3):**251-270.

Hallauer, A.R. and Miranda, J.B., 1988. Quantitative genetics in maize breeding, 2nd éd.*Iowa State Uni. Press, Ames, IA*.

Hampl, J.S. and Hampl, W.S., 1997. Pellagra and the origin of a myth: evidence from European literature and folklore. *Journal of the Royal Society of Medicine*, ***90*****(11):**636-639.

Hamrick, J.L., and M.J.W. Godt. 1997. Allozyme diversity in cultivated crops. Crop Science **37:**26-30.

Heckenberger, M., M. Bohn, J.S. Ziegle, L.K. Joe, J.D. Hauser, M. Hutton, and A.E. Melchinger. 2002. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. Molecular Breeding **10:**181-191

Helentjaris, T., Weber, D. and Wrights, S. 1988. Identification of genetic locations of duplicate nucleotide sequence in maize by analysis of restriction fragment length polymorphisms. Genetic **118:** 353:363.

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

Hinze, L.L.; Kresovich, S.; Nason, J.D.; Lamkey, K.R. 2005.Population genetic diversity in a maize reciprocal recurrent selection program. Crop Science **45:** 2435-2442.

https://www.diversityarrays.com/orderinstructions/plant-dna-extraction-protocol-for-dart

Huff, D.R. R. Peakall, P.E. Smouse. 1993. RAPD variation within and among natural populations of outcrossing buffalograss [(*Buchloedactyloides*(Nutt.) Engelm]. Theoretical and Applied Genetics **86:** 927-934

Jaccard, P. (1908). Nouvellesrecherchessur la distribution florale. Bulletin de la SocieteVaudoise des Sciences Naturelles**44:**223-270.

Jambrović, A., Šimić, D., Ledenčan, T., Zdunić, Z. and Brkić, I., 2008. Genetic diversity among maize (Zea mays, L.) inbred lines in Eastern Croatia. *Periodicumbiologorum*, ***110*****(3):**251-255.

Jeffreys, A.J., Wilson, V. and Thein, S.L., 1985.Individual-specific 'fingerprints' of human DNA. *Nature*, ***316*****(6023):**76.

Karp, A., Edwards, K.J., Bruford, M., Funk, S., Vosman, B., Morgante, M., Seberg, O., Kremer, A., Boursot, P., Arctander, P. and Tautz, D., 1997. Molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature biotechnology*, ***15*****(7):**625.

Keno, T., Azmach, G., Gissa, D.W., Regasa, M.W., Tadesse, B., Wolde, L., Deressa, T., Abebe, B., Chibsa, T. and Mahabaleswara, S., 2018. Major biotic maize production stresses in Ethiopia and their management through host resistance. *African Journal of Agricultural Research*, ***13*****(21):**1042-1052.

.

Kostova, A., Todorovska, E., Christov, N., Sevov, V. and Atanassov, A.I., 2006.Molecular characterization of Bulgarian maize germplasm collection via SSR markers. *Biotechnology & Biotechnological Equipment*, ***20*****(2):**29-36.

Kumar, A., Rakshit, A., Mangilipelli, N.K., Varalaxmi, Y., Vijayalakshmi, T., Vanaja, J.M., Yadav, S.K., Venkateswarlu, B. and Maheswari, M., 2012. Genetic diversity of maize genotypes on the basis of morpho-physiological and simple sequence repeats (SSR) markers. *African Journal of Biotechnology*, ***11*****(99):**16468-16477.

Kumari, J., Gadag, R.N. and Prasanna, B.M., 2005. Molecular profiling of maize (Zea mays L.) inbred lines using SSR markers. *The Indian Journal of Genetics and Plant Breeding*, ***65*****(4):**249-252.

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

Li, L., Eichten, S.R., Shimizu, R., Petsch, K., Yeh, C.T., Wu, W., Chettoor, A.M., Givan, S.A., Cole, R.A., Fowler, J.E. and Evans, M.M., 2014. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome biology*, ***15*****(2):**40.

Li, Y., Du, J., Wang, T., Shi, Y., Song, Y. and Jia, J., 2002. Genetic diversity and relationships among Chinese maize inbred lines revealed by SSR markers [Zea mays L.-Simple Sequence Repeat]. *Maydica (Italy)*.

Liu, K., Goodman, M., Muse, S., Smith, J.S., Buckler, E.D. and Doebley, J., 2003. Genetic structure and diversity among maize inbred lines as inferred from DNA microsatellites. *Genetics*, ***165*****(4):**2117-2128.

Lopez, A., Scapim, C., (2015). Genetic diversity assessed by microsatellite markers in sweet corn cultivars

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

Lun YQ, Ping F, Cheng KK and Tang PG (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.

Manen, J.F., Bouby, L., Dalnoki, O., Marinval, P., Turgay, M. and Schlumbaum, A., 2003. Microsatellites from archaeological Vitisvinifera seeds allow a tentative assignment of the geographical origin of ancient cultivars. *Journal of Archaeological Science*, ***30*****(6):**721-729.

Matsuoka, Y., Mitchell, S. E. and Kresovich, S. 2002. Microsatellites in Zea variability, patterns of mutations and use for evolutionary studies.Theor. Appl. Genet. **104:** 436:450.

Melchinger, A.E., 1999. Genetic diversity and heterosis.*The genetics and exploitation of heterosis in crops*, **10**: 99-118.

Melchinger, A.E., Lee, M., Lamkey, K.R., Hallauer, A.R. and Woodman, W.L., 1990. Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallele sets of maize inbreds. *Theoretical and applied genetics*, ***80*****(4):**488-496.

Mohammadi, S.A. and Prasanna, B.M., 2003. Analysis of genetic diversity in crop plants salient statistical tools and considerations. *Crop science*, ***43*****(4):**1235-1248.

Molin, D., Coelho, C.J., Máximo, D.S., Ferreira, F.S., Gardingo, J.R. and Matiello, R.R., 2013. Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers. *Genet Mol Res*, ***12*****(1):**99-114.

Mosisa, W., W. Legesse, T. Berhanu, D. Girma, A. Girum, A. Wende, K. Tolera, B. Gezahegn, W. Dagne, A. Solomon, Z. Habtamu, Y. Kassa, C. Temesgen, J. Habte, N. Demoz, and B. Getachew. 2012. Status and future direction of maize reseach and production in Ethiopia. p. 17-23.

Moustacchi, E. and Williamson, D.H., 1966. Physiological variations in satellite components of yeast DNA detected by density gradient centrifugation. *Biochemical and biophysical research communications*, ***23*****(1):**56-61.

Nei, M. and Li, W.H., 1979.Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, ***76*****(10):**5269-5273.

Nei, M., 1972.Genetic distance between populations. *The American Naturalist*, ***106*****(949):**283-292.

Nei, M., 1978.Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, ***89*****(3):**583-590.

Nesbitt, K.A., B.M. Potts, R.E. Vaillancourt, and A.K. West. 1995. Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus*(Myrtaceae). Heredity **74:** 628-637

Nyaligwa**,** L.,Hussein, S., Amelework, B., Ghebrehiwot, H., (2015). Genetic diversity analysis of elite maize inbred lines of diverse sources using SSR markers

OECD (Organization for Economic Cooperation and Development), 2003. Consensus Document on the Biology of Zea mays subsp. mays (Maize). *Series on Harmonization of Regulatory Oversight in Biotechnology (ENV/JM/ **11*** **(27):**1-49.

Pabendon, M.B., Mejaya, M.J., Koswara, J. and Aswidinnoor, H., 2010. Korelasijarakgenetikberbasismarkamikrosatelitinbridajagungdenganbobotbiji F1.*J. PenelitianPertanianTanamanPangan*, ***29*****(1):**11-17.

Pandit, M., Chakraborty, M., Haider, Z.A., Pande, A., Sah, R.P. and Sourav, K., 2016. Genetic diversity assay of maize (Zea mays L.) inbreds based on morphometric traits and SSR markers. African Journal of Agricultural Research, **11(24):**2118-2128.

Pardue, M.L. and Gall, J.G., 1970. Chromosomal localization of mouse satellite DNA. *Science*, ***168*****(3937):**1356-1358.

Poehlman, J.M. and Sleper, D.A., 1995. Breeding field crops.Iowa State Univ. Press, Ames, IA. *Breeding field crops.4th ed. Iowa State Univ. Press, Ames, IA.*

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A., 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular breeding*, ***2*****(3):**225-238.

Ranatunga, M.A.B., Meenakshisundaram, P., Arumugachamy, S. and Maheswaran, M., 2009. Genetic diversity analysis of maize (Zea mays L.) inbreds determined with morphometric traits and simple sequence repeat markers. *Maydica*, ***54*****(1):**113.

Reid, A., Hof, L., Felix, G., Rücker, B., Tams, S., Milczynska, E., Esselink, D., Uenk, G., Vosman, B. and Weitz, A., 2011.Construction of an integrated microsatellite and key morphological characteristic database of potato varieties on the EU common catalogue.*Euphytica*, ***182*****(2):** 239.

Roger, J.S., 1972. Measure of genetic similarity and genetic distance studies in genetics VII. *University of Texas publication*, ***7213*****:**145-153.

Semagn, K., Y. Beyene, D. Makumbi, S. Mugo, B.M. Prasanna, C. Magorokosho, and G. Atlin. 2012a. Quality control genotyping for assessment of genetic identity and purity

in diverse tropical maize inbred lines. Theoretical and Applied Genetics **125:**1487-1501

Semagn, K., Magorokosho, C., Vivek, B.S., Makumbi, D., Beyene, Y., Mugo, S., Prasanna, B.M. and Warburton, M.L., 2012. Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. *BMC genomics*, ***13*****(1):**113

Scott, K.D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E.M., Lee, L.S. and Henry, R.J., 2000. Analysis of SSRs derived from grape ESTs. *Theoretical and applied genetics*, ***100*****(5):**723-726.

Senior, M. L., Murphy, J. P., Goodman, M. M. et al. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using agarose gel system. Crop Sci. **38:** 1088:1098.

Sharma, S., Mishra, D.P., Kumar, B., Kaur, H. and Jewlia, H.R., 2017. Genetic Diversity of Maize (Zea mays L.) Inbred Lines Revealed by Simple Sequence Repeat Markers. Int. J. Curr. Microbiol. App. Sci, **6(12):**543-550.

Shehata AI, Al-Ghether HA and Al-Homaidan AA (2009). Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. Saudi Journal of Biological Sciences; **16:** 57-62.

Shiri, M. R., Choukan, R. and Aliyev, R. T. (2014). Study of genetic diversity among maize hybrids using SSR markers and morphological traits under two different irrigation conditions. Crop Breeding Journal **4:**65-72.

Shukla N, Mishra DK, Chavan A, Singh S. (2014) Genetic divergence and heterosis among maize genotypes as inferred from DNA microsatellites, The Bioscan.**10:**1753-1757.

Slavov, G.T., Howe, G.T. and Adams, W.T., 2005. Pollen contamination and mating patterns in a Douglas-fir seed orchard as measured by simple sequence repeat markers. *Canadian Journal of Forest Research*, ***35*****(7):**1592-1603.

Smith, J. S. C. and Smith, O. S (1992).Fingerprinting crop varieties.Advances in Agronomy.**47:** 85-140.

Smith, J.S., Paszkiewicz, S., Smith, O.S. and Schaffer, J., 1987.Electrophoretic, chromatographic, and genetic techniques for identifying associations and measuring genetic diversity among corn hybrids.In *Report of annual Corn and Sorghum Research Conference (USA)*.

Smith, J.S.C., O.S. Smith, and K.R. Lamkey. 2005. Maize breeding. Maydica**50:**185-187

Sokal, R.R., 1958.A statistical method for evaluating systematic relationship. *University of Kansas science bulletin*, ***28*:**1409-1438.

Sserumaga JP, Makumbi D, Ji H, Njoroge K, Muthomi JW, Chemining wa GN *2014,* Molecular characterization of tropical maize inbred lines using microsatellite DNA markers. **59:**267-274.

Tautz, D., and Reno, M., (1984). Simple sequences are repetitive components of eukaryotic genomes. Nucleic Acids Research; 12(10): 4127-4138

Tautz D, Schlotterer C. 1994; Concerted Evolution, Molecular Drive and Natural-Selection - Reply.*Curr Bio* **4:** 1166-1166.

Thompson, J.A., Nelson, R.L. and Vodkin, L.O., 1998. Identification of diverse soybean germplasm using RAPD markers.*Crop Science*, ***38*(5):**1348-1355.

Tito, C.M., Poggio, L. and Naranjo, C.A., 1991. Cytogenetic studies in the genus Zea. *Theoretical and Applied Genetics*, ***83*(1):**58-64.

United States Department of Agriculture (2005).Germplasm Resources Information Network.United States Department of Agriculture, Agricultural Research Service, Beltsville.Available from http://www.ars-grin.gov/cgi-bin/npgs/html/index.pl (Accessed on 25.03.2014).

Van Hintum, T.J. and Haalman, D., 1994. Pedigree analysis for composing a core collection of modern cultivars, with examples from barley (Hordeumvulgares. lat.).*Theoretical and applied genetics*, ***88*(1):**70-74.

Varshney, R.K., Graner, A. and Sorrells, M.E., 2005. Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology*, ***23*(1):**48-55.

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

Waples, R.S., 1987. A multi species approach to the analysis of gene flow in marine shore fishes.*Evolution*.**41(2):** 385-400.

Warburton, M.L., P. Setimela, J. Franco, H. Cordova, K. Pixley, M. Banziger, S. Dreisigacker, C. Bedoya, and J. MacRobert. 2010. toward a cost-effective finger printing methodology to distinguish maize open-pollinated varieties. Crop Science **50:**467-477.

Ward Jr, J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, ***58*****(301):**236-244.

Wolfe, J.H., 1970. Pattern clustering by multivariate mixture analysis. *Multivariate Behavioral Research*, ***5*****(3):**329-350.

Wright, S., 1951.The genetic structure of populations. *Annual Eugenics***.15:** 323-354

Xu J, Liu L, Xu Y, Chen C, Rong T, Alli F, Zhou S, Wu F, wang LYJ, Cao M, 2013. Development and characterization of simple sequence repeat markers providing genome- wide coverage and high resolution in maize. DNA Res 20: 497-509

Zheng, D.H., Van, K. and Lee, S.H., 2008. Molecular diversity and relationships among elite maize inbreds from US and CIMMYT populations and current heterotic groups in China.*Hereditas*, ***145*****(4):**182-193.

7. APPENDEX

Appendix 1: maize inbred lines and population codes used for SSR data

Maize inbred lines	Inbred lines code	population group	population code
124- B (109)	1	Bako	1
124 - B (113)	2	Bako	1
142 – 1- E	3	Bako	1
144 - 7- B	4	Bako	1
35B-190-0-110-2-1-2-2-1-2	5	Bako	1
A-7033	6	Bako	1
BKL 001	7	Bako	1
BKL 002	8	Bako	1
BKL 003	9	Bako	1
BKL 004	10	Bako	1
F-7215	11	Bako	1

MBRC5BCF 108-2-3-1	12	Bako	1
PO, OOE3-2-1-2-1	13	Bako	1
SC22	14	Bako	1
CML 144	15	CIMMIT	2
CML 161	16	CIMMIT	2
CML 165	17	CIMMIT	2
CML 197	18	CIMMIT	2
CML 202	19	CIMMIT	2
CML 312	20	CIMMIT	2
CML 334	21	CIMMIT	2
CML 395	22	CIMMIT	2
CML 444	23	CIMMIT	2
CML 536	24	CIMMIT	2
CML 357	25	South America	3
CML 359	26	South America	3

CML 360	27	South America	3
CML 361	28	South America	3
CML 362	29	South America	3
CML 363	30	South America	3
CML 364	31	South America	3
CML 365	32	South America	3
CML 366	33	South America	3
CML 435	34	South America	3
CML 436	35	South America	3
CML 438	36	South America	3
CML 439	37	South America	3

Appendix 2. Genetic distance between 37 elite maize inbred lines

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	
0																																					1
57	0																																				2
51	36	0																																			3
64	49	21	0																																		4
75	56	32	25	0																																	5
70	61	41	46	25	0																																6
75	52	58	43	34	33	0																															7
79	52	56	53	48	41	22	0																														8
83	56	48	47	44	53	46	44	0																													9
71	68	68	63	64	61	54	44	32	0																												10
91	84	84	81	84	89	70	72	80	92	0																											11
98	89	73	72	69	72	71	69	81	91	37	0																										12
84	81	73	72	69	76	77	85	81	83	55	26	0																									13
79	76	88	79	88	77	66	68	76	60	60	63	61	0																								14
90	69	53	52	49	48	51	33	29	39	73	60	82	71	0																							15
79	68	68	57	56	53	46	48	52	52	64	73	79	52	33	0																						16
88	83	85	82	81	78	73	69	69	73	53	62	54	49	62	41	0																					17
91	82	72	69	60	69	68	76	72	84	56	45	41	60	69	64	39	0																				18
91	80	72	65	68	77	74	72	76	80	56	41	39	56	65	60	37	12	0																			19
87	72	64	53	56	65	58	60	68	72	56	49	47	56	53	48	45	28	16	0																		20
74	61	63	60	71	64	61	47	65	63	61	50	62	49	46	45	58	61	49	45	0																	21
86	69	73	70	73	80	71	65	73	79	37	36	50	63	52	41	38	45	37	37	36	0																22
91	80	88	81	84	85	62	72	80	88	16	45	51	56	77	60	49	52	48	48	57	33	0															23
95	76	92	85	88	89	66	76	84	92	20	49	55	60	81	64	53	56	52	52	61	37	4	0														24
79	80	68	63	72	61	66	68	72	64	76	63	57	44	71	64	65	60	56	52	53	71	76	80	0													25
95	68	60	63	52	57	66	76	68	84	76	45	47	80	65	72	77	56	52	52	63	65	72	68	60	0												26
90	73	69	74	61	64	67	73	61	75	69	56	62	71	60	69	66	73	69	69	58	64	69	73	63	33	0											27
91	80	76	75	68	67	76	76	64	68	92	69	67	68	61	60	61	84	72	72	67	73	88	92	60	40	25	0										28
83	78	68	67	68	71	78	76	76	84	80	57	53	68	73	80	59	68	60	60	63	69	76	80	52	40	37	28	0									29
80	67	57	50	59	70	67	71	67	75	75	52	42	67	72	71	70	59	51	39	56	60	67	71	37	45	58	57	37	0								30
83	64	64	55	64	73	62	68	68	80	72	53	53	52	65	68	77	68	56	48	57	65	64	68	60	44	53	56	48	31	0							31
77	74	56	55	64	69	74	72	76	80	88	53	45	72	73	80	65	52	44	48	55	65	80	84	44	52	65	64	44	27	44	0						32
75	74	68	75	80	77	88	88	76	64	88	67	57	76	83	88	75	68	64	68	71	63	80	84	48	64	63	60	52	37	60	40	0					33
73	70	64	75	72	67	80	80	80	84	68	47	53	72	79	84	73	68	68	64	67	67	76	80	48	44	51	48	36	45	52	44	44	0				34
91	80	72	75	72	69	74	84	76	88	60	43	45	56	67	72	77	64	64	64	61	63	64	68	60	48	55	68	68	55	32	60	68	48	0			35
92	87	67	66	67	62	65	71	67	81	55	45	49	69	69	77	70	67	63	51	72	65	55	59	49	51	57	65	59	42	39	47	45	45	37	0		36
80	85	75	78	71	70	81	87	87	95	75	58	58	71	82	71	74	55	63	63	68	70	83	87	63	47	66	63	59	58	51	59	71	35	39	60	0	37

Appendix 3: visualized PCR products and DNA ladder using gel documentation 3UV Trans illuminator at 4% agarose gel

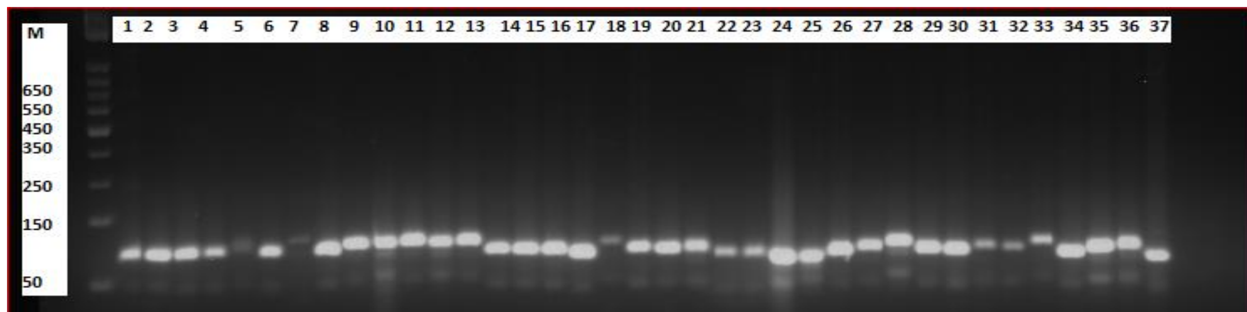


Figure 3. Microsatellite marker using Umc 1857, inbred line 1(124-b (109), 2(124-b(113), 3(142-1-e), 4(144-7-b), 5(35-190-0-510),6(A-7033), 7(BKL 001) 8(BKL 002), 9(BKL 003), 10 (BKL 004), 11(F-7215), 12(MBRC5BCF) 13(PO,OOE3-2-1), 14(SC22), 15(CML 144), 16(CML 161), 17(CML 165), 18(CML 197), 19(CML 202), 20(CML 312), 21(CML 334), 22(CML 395), 23(CML 444), 24 (CML 536) 25(CML 357), 26(CML 359), 27(CML 360), 28(CML 361), 29(CML 362), 30(CML 363), 31(CML 364), 32(CML 365), 33(CML 366), 34(CML 435), 35(CML 436), 36(CML 438), 37(CML 439)

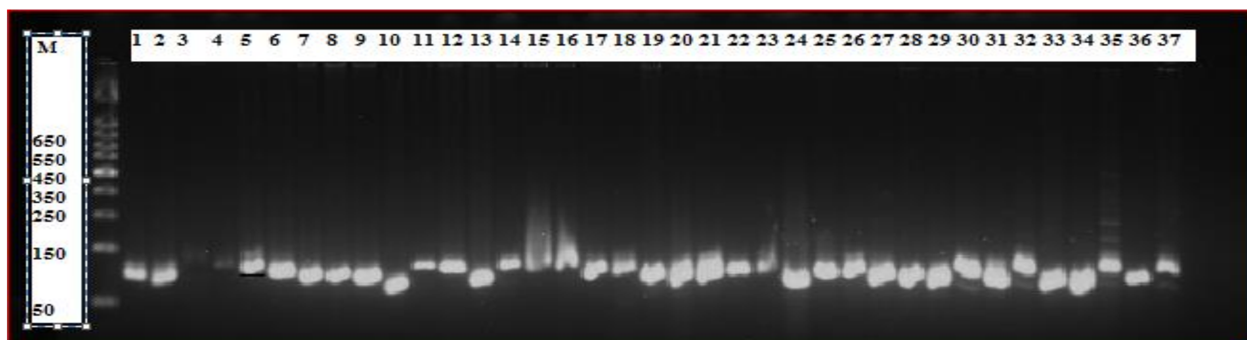


Figure 4. Microsatellite marker using Umc 2164, inbred line 1(124-b (109), 2(124-b (113), 3(142-1-e), 4(144-7-b), 5(35-190-0-510),6(A-7033), 7(BKL 001) 8(BKL 002), 9(BKL 003), 10 (BKL 004), 11(F-7215), 12(MBRC5BCF) 13(PO,OOE3-2-1), 14(SC22), 15(CML 144), 16(CML 161), 17(CML 165), 18(CML 197), 19(CML 202), 20(CML 312), 21(CML 334), 22(CML 395), 23(CML 444), 24 (CML 536) 25(CML 357), 26(CML 359), 27(CML 360), 28(CML 361), 29(CML 362), 30(CML 363), 31(CML 364), 32(CML 365), 33(CML 366), 34(CML 435), 35(CML 436), 36(CML 438), 37(CML 439)