Ethiop. J. Agric. Sci. 30(2) 31-53 (2020)

Genetic Variations in CIMMYT and Ethiopian Maize (*Zea mays* L.) Inbred Lines as Determined by Microsatellite Markers

Tsegaye Abebe[1], Leta Tulu[2], Kalkidan Tesfu[2] and Wosene Gebreselassie[3]

[1]Gambella University; P. O Box 126, Gambella, Ethiopia; [2]Ethiopian Institute of Agricultural Research, National Agricultural Biotechnology Research Center(NABRC) P. O. Box 249, Holetta, Ethiopia; [3]College of Agriculture and Veterinary Medicine, Jimma University, P. O. Box, 370, Jimma, Ethiopia

አህፅሮት

በዕፀዋት ማሻሻያ ምርምር ሀደት ውስጥ ተዳቅለው በቀጣይ ትውልዳቸው ዘረመልን መሰረት ያደረገ ተለያይነት የሚያሳዩ እናት ዝርያዎችን ለመምረጥና የተለያዩ ባህርያትን ለማሻሻል ሞለኩላር ማርከሮችን መሠረት ያደረገ መረጣ የዕፀዋትን ዉጪአዊ ዕይታን መሰረት ካደረገ መረጣ የበለጠ ጠቃሚ መሆኑ ይታመናል። በዚህ ጥናት 14 ከኢትዮጵያ እና 23 ከአለም አቀፍ የስንዴና በቆሎ ምርምር ማዕከል የተገኙትን ጨምሮ በአጠቃላይ 37 ልዕል የበቆሎ እናት ዝርያዎች 28 ማይክሮ ሳቴላይት ማርከሮችን በመጠቀም ተጠንተዋል። ከ23ቱ ከCIMMYT የተገኙ እናት ዝርያዎች 13ቱ የአፈርን ኮምጣጣነት በመቋቋም ባሕርያቸው የታወቁና በቅርብ ጊዜ ወደ ኢትዮጵያ የገቡ ናቸው። የጥናቱም ዋና ዓላማ በኢትዮጵያ ከዚህ ቀደም በምርምር በተገኙና የአፈርን ኮምጣጣነት በመቋቋም ባሕርያቸው በታወቁ የCIMMYT እናት ዝርያዎች መካከል ያለውን የጀኔቲክስ ርቀት በመወሰን በቀጣይ የአፈርን ኮምጣጣነት ለመቋቋም የሚያስችል የብዝሃ ተለያይነት የሚፈጥሩ እናት ዝርያዎችን ለመለየት ነው። በተደረገው ሞሌኩላር አናሊሰስ መሰረት በተጠኑት እናት ዝርያዎች መካከል ከፍተኛ (77%) ተለያይነት የታየ ሲሆን የሞለኩላር ማረከሮቹ እናት ዝርያዎችን የመለየት አቅም እንዳላቸው አሳይተዋል። በአጠቃላይ 107 አሌሎች እና በአማካይ 3.71 አሌሎች በሎሴስ ተለይተዋል። Expected hetrozygosity እና major allele frequency እያንዳንዳቸው ከ0.09 እስከ 0.61 እና ከ0.34 እስከ 0.95 ተለያይነት ያላቸው ሲሆን በእናት ዝርያዎቹ መሃል ሰፊ የብዝሃ ተለያይነት መኖሩንና ይህ ተለያነት ደግሞ ወደፊት ሰብሉን ለማሻሻል ትልቅ አቅም መሆኑን ለማወቅ ተችሏል። የማርከሮቹ Polymorphic information content ከ0.10 (Bnlg 1063) እስከ 0.74 (Umc 2205) በመለያየትና በአማካይ 0.50 በማስመዝገብ ዝርያዎቹን የመለየት አቅም እንዳላቸው አሳይተዋል። በአጠቃላይ የኒስ ጀኔቲክ ርቀት (Nei's genetic distance) በጥንድ እናት ዝርያዎች መካከል ከ0.16 እስከ 0.98 ተለያይነት ያሳየ ሲሆን ከኢትዮጵያ በተገኙ እናት ዝርያዎች መካከል ከ0.26 እስከ 0.98 ተላያይነት ሲኖር በኢትዮጵያና የአፈርን ኮምጣጣነት በመቋቋም ባሕርያቸው በታወቁ የCIMMYT እናት ዝርያዎች መካከል ከ0.42 እስከ 0.95 ሆኖ ከጥንዶቹ ከ41.6 በመቶ የሚሆኑ ጥንዶች ከ0.8 በላይ የሆነ የጀኔቲክስ ርቀት) በማሳየት የአፈርን ኮምጣጣነት የሚቋቋሙ የብዝሃ ተለያይነትን ለማስፋትና ዝርያዎችን ለማነልበት የላቀ አቅም እንዳለቸው ለማወቅ ተችሏል። በተለይም ከኢትዮጵያ የተገኘው እናት ዝርያ 124-b(109) የአፈርን ኮምጣጣነት የመቋቋም ባህርይ ካላቸው እናት ዝርያዎች ማለትም CML359, CML360, CML361, CML436 እና CML438 እንዲሁም ደግሞ ልላው ከኢትዮጵያ የተገኘው F-7215 የሚባል እናት ዝርያ ከCML361 ጋር ከ0.9 በላይ የሆነ የጀኔቲክስ ርቀት እንዳለው ታውቋል። እነዚህ ጥንድ እናት ዝርያዎች የአፈርን ኮምጣጣነት የመቋቋም ባህርይንና ሌሎች ተፈላጊ የሆኑ ባህርያትን በጣምራ ያላቸውን የበቆሎ ዝርያዎችን ለማነልበት ትልቅ አቅም ሲኖራቸው ይህ አቅማቸው በማዳቀልና በመምረጥ ጥቅም ላይ ሊውል ይችላል። ሆኖም የአፈርን ኮምጣጣነት የመቋቋም ባህርይ ያላቸውን እናት ዝርያዎች፤ ዝርያዎችንነ የማነልበት ሥራ ላይ ከመዋላቸው በፊት በኢትዮጵያ በቆሎ አብቃይ ሥነ-ምህዳር ተስማሚነታቸው መጠናት ይኖርበታል።

Abstract

Molecular markers are more useful in identifying parental inbred lines crosses of which create genetic variation among progenies in the advanced generation more realistically than selection based on phenotypic data. In this study, 37 maize inbred lines including 14 of Ethiopian and 23 of CIMMYT origin were studied using 28 microsatellite markers. Among the 23 CIMMYT inbred lines, 13 were soil acidity tolerant lines and were recently introduced. The main objective was to identify distantly related Ethiopian maize inbred lines and soil acidity tolerant inbred lines

of CIMMYT origin that could be used in broadening the genetic basis of maize for tolerance to soil acidity. Analysis of molecular variance indicated higher genetic variability (77%) among the inbred lines indicating ability of the SSR markers in discriminating the inbred lines. One-hundred and seven alleles were identified in the inbred lines with average of 3.71 alleles per locus. Expected hetrozygosity (He) and major allele frequency ranged from 0.09 to 0.61 and from 0.34 to 0.95, respectively, indicating higher genetic variability among the inbred lines, which could be exploited in breeding program. Polymorphic information content (PIC) generated by each marker ranged from 0.10 (Bnlg 1063) to 0.74 (Umc 2205) with a mean of 0.50 indicating their power in discriminating the inbred lines. Nei's genetic distance among pairs of the 37 maize inbred lines varied from 0.16 to 0.98. Pairs of inbred lines of Ethiopian origin had genetic distance that ranged from 0.26 to 0.98 and the highest genetic distance was in this population. Pairs of inbred lines of Ethiopian and CIMMYT inbred lines developed for tolerance to soil acidity had genetic distance that varied from 0.42 to 0.95 and 41.6 % of the pairs had genetic distance of more than 0.8 indicating their potential in broadening the genetic basis of maize for developing varieties tolerant to soil acidity. Specifically, 124-b(109) was distantly related with five inbred lines, viz., CML359, CML360, CML361, CML436 and CML438, and F-7215 was also distantly related with CML361 with Nei's genetic distance of more than 0.9. These pairs of inbred lines could be exploited in pedigree breeding through which segregating populations can be selected for desirable traits in combination with tolerance to soil acidity. However, the soil acidity tolerant inbred lines have to be evaluated for adaptation to Ethiopian environment before embarking on exploiting their potential in breeding for tolerance to soil acidity.

Introduction

Significant efforts have been made since the early 1950s (Tolessa *et al*., 1993) to develop well-adapted and high yielding hybrids and open-pollinated varieties of maize (*Zea mays* L.) for different agro ecologies of Ethiopia. Combined use of these varieties with improved agronomic practices have helped farmers to produce mean grain yield of 6 t/ha (Tolessa *et al*., 1997). As a result, the national maize production grew from 1.7 million tons in the early 1980s to the current 8.4 million tons (CSA, 2018 and several years report). Similarly, at the national level, the productivity of maize increased from 1.7 t/ha in the early 1980s to 4 t/ha of the current level. This is believed to have invariably contributed to better maize consumption and improved livelihood of households involved in maize production. There is robust effort still being exerted to genetically improve maize for tolerance to biotic and abiotic constraints that hinder further progresses in maize productivity. Among the abiotic constraints that hinder realization of the yield potential of the crop under Ethiopian condition is the increasing soil acidity that is covering the major maize belt of the country (Schlede, 1989).

In its hybrid development strategy, the Ethiopian maize breeding program has accumulated several elite inbred lines through introduction from foreign sources

and developing locally from germplasm of diverse genetic background. The positive contributions to the vision of developing stress tolerant and higher yielding hybrids are then expected from use of these inbred lines (Abate *et al*., 2015, Ertiro *et al*., 2017). Information on genetic diversity among these elite inbred lines has an important role in exploiting their contribution to the genetic enhancement of the crop. Despite this fact, information on molecular genetic variation of inbred lines introduced and/or developed locally from diverse sources has been inadequate. Information on genetic diversity or similarity among inbred lines is, therefore, essential for broadening the genetic basis and creating variation for selection.

Hitherto genetic variation among inbred lines used to be assessed focusing on morphological (phenotypic) variation with little emphasis on molecular tools. Genetic distances in most cases are inferred from grain yield performance of F_1 hybrids in surplus of their high yielding parents that is commonly referred to as high-parent heterosis (Worku *et al*., 1996, Tulu, 2004). In both cases environment has higher role to play in introducing errors that interfere with the accuracy of establishing reliable information on the genetic diversity in the target germplasm. In a more realistic manner, genetic variation within a population can be assessed based on: i) the number (and percentage) of polymorphic genes in the population, ii) the number of alleles for each polymorphic gene and iii) the proportion of heterozygous loci per individual. Recent techniques in DNA-based markers provide such information that then avail reliable insight in genetic diversity in target germplasm based on their genetic information.

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 1-4 bases, and are called mono-, di-, tri- or tetra nucleotide repeats, respectively (Tautz *et al.*, 1986). They are ideal markers to assess the genetic diversity of maize; because they have advantages such as high level of reproducibility, small quantity of DNA required to run polymerase chain reaction(PCR), high polymorphism and ability to generate more genetic information, co-dominant nature, presence in large number throughout the genome (Scott *et al.*, 2000, Manen *et al.*, 2003, Varshney *et al.*, 2005).

Numerous SSR markers have been identified in maize (Sharopova *et al*., 2014) and used to assess genetic diversity (Gupta *et al.*, 2010; Pandit *et al.*, 2016 and

Sharma *et al.*, 2017). However, in Ethiopia very little work has been reported on genetic diversity in maize inbred lines using SSR marker. The first report on use of SSR markers in Ethiopian maize was by Beyene *et al.* (2006) who studied the genetic diversity and association among sixty-two traditional Ethiopian highland maize genotypes using twenty SSR markers. They reported a mean of 4 alleles per locus and genetic diversity that ranged from 0.27 to 0.63 with a mean of 0.49. Later, Legesse *et al.* (2007) investigated the genetic diversity of 56 highland and mid-altitude maize inbred lines using 27 SSR markers. They reported 104 alleles with a mean of 3.85 alleles per locus. The average polymorphism information content (PIC) was 0.58. Genetic distance varied from 0.28 to 0.73 with an average of 0.59. More recently, Demissew (2014) reported genetic distances that ranged from 0.30 to 0.78 nearly in 98% of the pairwise comparisons in 30 quality protein maize (QPM) and 6 normal maize inbred lines adapted to highland agro-ecology of Ethiopia using 25 SSR markers. Ertiro *et al.*, (2017) investigated the genetic purity, relatedness and population structure of 265 maize inbred lines from the Ethiopian Institute of Agricultural Research (EIAR), the International Maize and Wheat Improvement Centre (CIMMYT) and the International Institute of Tropical Agriculture (IITA) using 220,878 single nucleotide polymorphic (SNP) markers obtained using genotyping by sequencing (GBS). They reported that of the tested inbred lines only 22% were considered pure with <5% heterogeneity, while the remaining 78% had a heterogeneity that ranged from 5.1 to 31.5%. Pairwise genetic distances among the inbred lines varied from 0.011 to 0.345, with 89% of the pairs falling between 0.301 and 0.345.

In the current study 37 maize inbred lines including 24 elite maize inbred lines which are parental lines of commercial maize hybrids of Ethiopia and 13 soil acidity tolerant CIMMYT inbred lines were studied using 28 SSR markers. The study is part of the ongoing study on evaluation of elite maize germplasm for tolerance to soil acidity *in vitro* with the view of broadening the genetic bases of locally available elite inbred lines for tolerance to soil acidity. In this study we examined the genetic distance or similarity among the acid tolerant lines and the Ethiopian parental inbred lines, which we hope would generate information that will supplement the effort on development of acid tolerant germplasm under Ethiopian condition. Genetically distantly related inbred lines identified from Ethiopian and CIMMYT acid tolerant inbred lines will be used to create genetic variation in advanced generation of their progenies that could be exploited in developing acid soil tolerant maize germplasm for Ethiopia.

Materials and Methods

Plant materials

Thirty-seven elite maize inbred lines, including fourteen inbred lines developed by the National Maize Breeding Research Program (NMBRP) of Ethiopia and twenty-three developed by the International Maize and Wheat Improvement Center (CIMMYT) were used in this study. The fourteen inbred lines developed by the NMBRP and 10 of the CIMMYT inbred lines introduced previously by the NMBRP 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 Holetta. The remaining 13 CIMMYT inbred lines were developed for tolerance to soil acidity and were received from CIMMYT-Mexico in 2019 to use as standards in evaluating the Ethiopian germplasm for tolerance to soil acidity. Hence, based on their pedigree, origin and adaptation, the 37 inbred lines were divided into population 1 included 14 inbred lines developed by the NMBRP of Ethiopia, population 2 included 10 inbred lines of CIMMYT origin and population 3 included 13 inbred lines developed by CIMMYT from germplasm of Latin American origin for tolerance to soil acidity. Pedigree and origin of the 37 maize inbred lines are indicated in Table 1.

Genomic DNA extraction

Ten seeds of each inbred line were sown in plastic pots filled with soil medium composed of soil, sand and forest soil mixed in the ratio of 1:1:1 and grown in the greenhouse at NABRC, Holetta. Leaves were collected two to three weeks after planting and total genomic DNA was extracted from the bulked leaf sample per inbred line using the DArT (Diversity array technology) methods. The quality and concentration of the extracted genomic DNA were measured using Nano drop spectrophotometer and DNA samples were then diluted to concentration of 10 ng/µl.

SSR marker selection

Twenty-nine SSR markers were selected from the maize genomic database (http://www.maize gdb.orf/ssr.php) based on their distribution over the ten maize chromosomes (Table 2). Among these, one marker was excluded from the investigation due to lack of polymorphism across the inbred lines.

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

Inbred line	Pedigree	Origin
124-b(109)	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
124-b (113)	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
142-1-e	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
144-7-b	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
35B-190-0-S10-2-1-2-2-1-2	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
A-7033	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
BKL001	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
BKL002	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
BKL 003	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
BKL 004	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
F-7215	Unknown (derived from Kitale Syn. II)	Bako (Ethiopia)
MBRC5BCF108-2-3-1-B-5-2-B-B-B-B	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
PO,OOE3-2-1-2-1	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
SC22	Unknown (derived from Ecuador 573)	Bako (Ethiopia)
CML144	P62C5F182-2-1-2BB-3-1	CIMMYT
CML161	G25QC18H520-1-1-1-25-3-B-1-BBBB	CIMMYT
CML165	QF37SR-2-3SR-2-4-3-BBB	CIMMYT
CML197	G34QH174-3-1-2-BB	CIMMYT
CML 202	ZSR923-S4BULK-5-1-BBB	CIMMYT
CML 312	S89500 F2-2-2-1-1-B*5	CIMMYT
CML 334	S920-F47-2-1-2-1-BBBBB	CIMMYT
CML 395	90323(B)-1-X-1-B-B-1-1-B-B-1-1-B	CIMMYT
CML 444	P43C9-1-1-1-1-1-BBBBB	CIMMYT
CML 536	ZM605C2F1-17-1-B-1-BB	CIMMYT
CML 357	SA3-C4F5(6/24)-1-2-2-5-B	CIMMYT
CML 359	SA3C4HC(16?25)-2-4-3-1-B	CIMMYT
CML 360	SA4-C2-FS(21/26)-1-2-2-2-B	CIMMYT
CML 361	SA4-C2-FS(21/26)-4-2-7-3-B	CIMMYT
CML 362	SA5-FS1-3-9-1-3-4-B	CIMMYT
CML 363	SA5-FS1-3-9-1-5-2-B	CIMMYT
CML 364	SA5-FS1-5-1-1-5-3-B	CIMMYT
CML 365	SA8- C1-FS(27/3)-1-1-4-8-B	CIMMYT
CML 366	SA8-C2-FS(27/3)-1-3-6-1-B	CIMMYT
CML 435	SA3-C4HC(16/25)-2-4-3-6-B-B-B-B-B	CIMMYT
CML 436	SA3-C4-FS(19/25)-2-6-4-5-B	CIMMYT
CML 438	SA4-C2-FS(21/26)-1-2-2-2-B	CIMMYT
CML 439	SA5-C2HC(26/21)-4-3-7-5-B-B-B-B-B-B	CIMMYT

Table 2. Description of SSR markers used in the study. F: forward primer; R: reverse primer

Marker	Primer sequence (5'---3')	Annealing temperature (^{O}C)	Repeat motif	Bin
Umc 2164	F- AGCACACAGACAAGAGAGACAACG	58.2	(CGGC)8	5.05
	R- GACCGACAACAGAGATCGAGTACA			
Umc 1506	F- AAAAGAAACATGTTCAGTCGAGCG	53.5	(AACA)4	10.05
	R- ATAAAGGTTGGCAAAACGTAGCCT			
Umc 1607	F- ACTAATTTTCGGTAGTCGTGTGCG	53.5	(TGC) 5	8.07
	R- GGAAAGAGAGAGGCTGTAGGTGGT			
Umc 1137	F-TCAGTCACTCTTCTGCCTCCACT	52.5	(CT)15	9.08
	R-GGCTGGATAATGTTGTAGCTGGTC			
Umc 2280	F- AAAAGAAGACGCTTTGTTTGTTGC	58.3	(CATTA)4	4.03
	R- FAAAGAAGACGCTTTGTTTGTT AAAAGAAGACGCTTTGTTTGTTGC			
Umc 1363	F- TGTTTAAGTGTTGGCAGAAAGCAA	59.4	(ACG)4	1.01
	R- TCTCCCTCCCCTGTACATGAATTA			
Umc 1757	F- TTTTCTGCAGGGATAACATTTGTG	59.4	(TCC)7	4.01
	R- ATAGGAGGTGAGGTGAGGAGGAAG			
Umc 1272	F- CTCTGACAGACCTGCAGATAGGGT	58.4	(CTAGC)4	10.04
	R- ATCGAGGGGCTAATCAGCAAG			
Umc 1636	F- CATATCAGTCGTTCGTCCAGCTAA	58.4	(AGGC)4	9.02
	R- GTACTGGTACAGGTCGTCGCTCTT			
Umc 1857	F- TTCCTTGCCAACAAATACAAGGAT	55.8	(TAA)6	6.04
	R- GTTCATTGCTTCATCTTGGAACCT			
Umc 1470	F- AAAAACCTCAATAGCCGTTTCACA	55.8	(TAA)7	8.04
	R- GATTCTTGTGTTGCATACTGGTGC			
Umc 2278	F- CTGACCTCCGTCATCAGCATC	52.3	(TC)8	4.01
	R- ATCACGGACAAAGAAAATTGAAGC			
Umc 1003	F- AATAGATTGAATAAGACGTTGCCC	55.8	(TAAA)9	2.05
	R- TGTTCCAATGCTTTTGTACCTCTA			
Umc 1913	F-GATCCTACCAAAATCTTATAGGC	55,8	(TTG)6	8.02
	R- ACAGCTAGCCAAGATCTGATT			
Umc 2085	F- TGTACGACTTCTTCTGGACGCAC	57.6	(CGC)5	2.08
	R- TAGATGTCGATGTCCTCCAGGG			
Umc 1075	F- GAGAGATGACAGACACATCCTTGG	57.6	(ATTGC)5	8.01
	R- ACATTTATGATACCGGGAGTTGGA			
Blng 1063	F- GGAGACAACCCCGACGAC	59.3	(AG)42	3.06
	R- GGTACCAGAGCCACAGATCC			
Umc 2080	F- GCCAAGGTGGGTCTGGCTAT	59.3	(TGGCTC)4	1.08
	R- ACCACCTTGTCCGTATCCTTCAC			
Umc 1415	F- GTGAGATATATCCCCGCCTTCC	58.6	(GAC)10	8.04

	R- AGACTTCCTGAAGCTCGGTCCTA			
Umc 2198	F- AGCCCAGAGAAGGGAAGCAG	58.6	(CCCTC)4	5.06
	R- CTCTTCACTCGCTTCTCCCAGA			
Umc 2319	F- GATCCACGCGAGGTTCACTG	58.2	(GAGGAG)5	6.04
	R- GCTCTCACTAGCCTCGCATTCC			
Bnlg 1927	F- TTTTTTTGTAAGCGATCCGG	55.5	(AG)41	4.07
	R- GATGAATCTGCGTCCGTCTT			
Umc 1066	F- ATGGAGCACGTCATCTCAATGG	58.3	(GCCAGA)5	7.01
	R- AGCAGCAGCAACGTCTATGACACT			
Umc 1904	F- CAGCCACTCGTTTATGGAGGTTTA	58.7	(TAAGC)5	8.03
	R- TGTTACTAGTCGATCTGATGCCCA			
Umc 1639	F- CTAGCCAGCCCCCATTCTTC	58.7	(TGTCC)4	3.09
	R- GCAAGGAGTAGGGAGGACGTG			
SSR 6	F- GATCCACGCGAGGTTCACTG	53.4	(CA)9	9.02
	R- GCTCTCACTAGCCTCGCATTCC			
Umc 2294	F- ATTGGAGTGGCTCCATTGCTT	53.4	(TCCTG)4	5.03
	R- CCCACCATTCTATATATTGTTGCCA			
Umc 2205	F- CATGATCATTTGGCGATGGTAAT	55	(TC)4	2.07
	R- ATGGTGAGCGAGTGAAAGAGAGAT			
SSR 14	F- AGGAGGTACCACAATGGAG	52.7	(CA)16	8.09
	R- GTGTACATCAAGGTCCAGATTT			

PCR amplification of SSR loci

Polymerase chain reaction (PCR) was performed in 12.5 µl reaction volume containing 6.25 µl one Taq Master Mix (containing $MgCl_2$, PCR buffer, dNTPs, and Taq DNA polymerase), 3µl genomic DNA (10 ng/µl), 0.5 µl of each forward (10 pico mole/µl) and reverse primer (10 pico mole/µl) and 2.25 µl nuclease-free water. PCR was carried out by using thermal cycler Gene Amp® PCR system 9700 (Applied Bio system, USA) programed 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 melting temperature of primer pairs) for 30 seconds, extension at 72 °C for 30 seconds and final elongation at 72 °C for 10 minutes. The PCR products were loaded on a 4% (w/v) agarose gel and run on electrophoresis for three hours at 100 volts. A 100+ 50bp DNA ladder with known reference bands was used to quantify the size of the amplified bands. The gel was visualized using gel documentation, 3UV-transilluminator (Bio-Doc).

Fragment scoring and data analysis

The amplified products were scored based on band size that was established with reference to the corresponding bands of the 100+50 base pair ladder. Fragments with the same mobility were considered as identical fragments and treated as a

unit character. The data from all entries were combined for statistical analysis. Genetic diversity among the 37 maize inbred lines was analyzed using GenAlex version 6.5 software package (Peakall and Smouse, 2015). The same software was used to analyze genetic diversity parameters such as number of alleles 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 hetrozygosity (Ho) and expected hetrozygosity (He), Genetic distance between inbred lines, principal coordinate analysis and analysis of molecular variance. Polymorphic information content (PIC) was analyzed using power marker v3, 25 software package (Liu and Muse, 2005). Clustering analysis was carried out using DARwin.

Results and Discussion

All 29 SSR markers amplified DNA fragments in the 37 maize inbred lines without any missing data. However, only 28 markers were polymorphic across the lines. Thus, one monomorphic marker was excluded from the analyses. The remaining 28 SSR markers identified a total of 104 alleles in 37 elite maize inbred lines. The number of alleles per polymorphic locus varied from 2.0 to 6.0 with the mean of 3.71 alleles. The highest number of alleles (6.0) was detected from loci Umc 1066 and Umc 2205, while the lowest (2) was detected from the locus Umc 1415. The number of alleles and the number of effective alleles (Ne) per polymorphic microsatellite locus, mean observed (Ho) and expected (He) hetrozygosity for the 28 loci are shown in Table 3. The average number of alleles obtained in the present study was lower than those reported from earlier maize diversity studies. Lu and Bernardo (2001) reported 4.9 alleles per SSR locus in 40 US maize inbred lines using 83 SSR markers. Warburton *et al.* (2002) also reported similar number of alleles per locus in 57 CIMMYT lines using 85 SSR markers. On the other hand, Senior *et al.* (1998) found five alleles per locus in 94 elite US maize inbred lines using 70 SSR markers while Vaz Patto *et al.* (2004) reported 5.33 alleles per locus in 104 Portuguese maize inbred lines using 15 SSR markers.

In the present study, the total number of alleles was similar with the findings of Legesse *et al*. (2007) who reported 104 alleles with a mean of 3.85 alleles per locus in 56 inbred lines using 27 SSR loci. In contrast, it revealed the highest total and average number of alleles compared to Sharma *et al.* (2017) who reported 40 alleles with a mean of 2.22 alleles on 33 maize inbred lines using 40 SSR markers. These differences in the number of alleles from different studies might be due to number of SSR loci, genotype differences and the number of genotypes used in the study and methodologies employed for detection of polymorphic marker (Agarose vs polyacrylamide). For example, using twenty SSR markers, Beyene *et al.* (2006) reported an average of 4.9 alleles per locus with a range of 3-10 alleles

in 62 traditional Ethiopian highland maize accessions indicating that the genetic basis of the germplasm also influence the number of alleles and average alleles per locus. From the total of 28 SSR markers used in current study, 12 markers showed 4.0 alleles each, 12 of them three alleles each, three of them six alleles each and one marker revealed 2 alleles. Based on this, the total number of alleles from the 28 SSR markers was 104.

The number of effective alleles (Ne) ranged from 1.11 to 3.53 with a mean of 2.04 (Table 3). Effective number of alleles refers to the number of alleles that would take to achieve the same expected hetrozygosity as in the studied population. The highest number of effective allele was obtained from marker Umc 2205 (3.53), whereas the lowest number was from marker Bnlg1063 (1.11). The observed hetrozygosity (Ho) value ranged from 0.00 to 0.17 with a mean of 0.02. The highest observed hetrozygosity value was from marker Umc 1137 (0.17), while the lowest hetrozygosity of nil was recorded from 20 markers. These were 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 and SSR14. This shows that the 20 markers listed above were fixed and reached at the maximum homozygous state. On the other hand, expected hetrozygosity or gene diversity (He) of the markers ranged from 0.09 to 0.61 with a mean of 0.45. While the highest expected hetrozygosity was observed from marker Umc 2205, the lowest was from marker Bnlg 1063. The highest mean value of expected hetrozygosity indicates the presence of high allelic variation (diversity) in the marker loci and reveals the presence of high level of polymorphism in the studied materials. This was observed with locus Umc 2205, which had the highest expected hetrozygosity of 0.61 and PIC value of 0.74. The overall mean of observed hetrozygosity was lower than that of expected hetrozygosity. This substantial differences between the means of observed and expected hetrozygosity might be due to the level of inbreeding in maize populations used in the current study.

In inbred population, proportion of heterozygous loci not exceeding 5% indicates genetic purity or genetic fixation (Semagn *et al*., 2012). In the current study, 32% of the loci had expected hetrozygosity of greater than 5% indicating lack of genetic purity or genetic fixation in maize inbred lines used for the investigation. This low hetrozygosity in the population might be due to early generation of the inbred lines used in the study and hence lack of genetic fixation, or contamination by pollen from nearby sources or seed of other germplasm during maintenance or seed increase (Ertiro *et al*., 2017). The Ethiopian maize breeding program prefers to use early generation inbred lines in the interest of reducing the cost of hybrid seed production. Ertiro *et al*., (2017) reported remarkably higher heterogeneity of 31.5% in the Ethiopian maize inbred lines used in their study compared to the

CIMMYT and IITA inbred lines that had heterogeneity of 21 and 30%, respectively, as revealed by SNP markers. This was ascribed to use of inbred lines with inbreeding level of less than fourth generations (S_4).

Table 3. Genetic parameters estimated for each microsatellite locus in 37 maize inbred lines

Loci	Genetic parameter							
	N_a	N_e	H_o	H_e	I	F_{is}	Major allele frequency	PIC
Umc 2164	3.00	2.32	0.00	0.55	0.87	1.00	0.51	0.55
Umc 1506	3.00	2.20	0.00	0.53	0.89	1.00	0.62	0.48
Umc 1607	3.00	1.81	0.00	0.45	0.64	1.00	0.40	0.58
Umc 1137	3.00	2.26	0.17	0.55	0.92	0.70	0.55	0.52
Umc 2280	4.00	2.33	0.02	0.47	0.84	0.95	0.41	0.63
Umc 1363	3.00	1.94	0.00	0.47	0.78	1.00	0.60	0.50
Umc 1757	4.00	1.95	0.05	0.48	0.72	0.90	0.57	0.51
Umc 1272	6.00	2.56	0.00	0.59	1.060	1.00	0.54	0.61
Umc 1636	4.00	2.43	0.00	0.58	0.99	1.00	0.43	0.54
Umc 1857	3.00	2.18	0.00	0.54	0.88	1.00	0.62	0.48
Umc 1470	3.00	1.39	0.00	0.24	0.37	1.00	0.35	0.59
Umc 2278	3.00	1.48	0.00	0.30	0.52	1.00	0.81	0.30
Umc 1003	4.00	1.55	0.00	0.32	0.53	1.00	0.65	0.41
Umc 2085	3.00	1.83	0.00	0.45	0.69	1.00	0.51	0.41
Umc 1075	4.00	1.97	0.00	0.47	0.73	1.00	0.65	0.50
Blng 1063	3.00	1.11	0.03	0.09	0.19	0.72	0.95	0.10
Umc 2080	4.00	1.72	0.00	0.38	0.66	1.00	0.35	0.63
Umc 1415	2.00	1.18	0.00	0.15	0.29	1.00	0.92	0.14
Umc 2198	4.00	1.87	0.10	0.44	0.75	0.76	0.65	0.47
Umc 2319	3.00	1.66	0.02	0.33	0.54	0.93	0.41	0.56
Bnlg 1927	4.00	2.10	0.00	0.52	0.81	1.00	0.57	0.44
Umc 1066	6.00	2.79	0.05	0.54	1.01	0.91	0.51	0.64
Umc 1904	4.00	2.11	0.03	0.52	0.81	0.95	0.38	0.60
Umc 1639	3.00	2.06	0.00	0.46	0.77	1.00	0.54	0.50
Umc 2319	4.00	2.06	0.00	0.45	0.76	1.00	0.65	0.50
Umc 2294	4.00	2.60	0.00	0.59	0.98	1.00	0.40	0.62
Umc 2205	6.00	3.53	0.03	0.61	1.24	0.95	0.34	0.74
SSR 14	4.00	2.19	0.00	0.53	0.88	1.00	0.50	0.60
Mean	3.71	2.04	0.02	0.45	0.75	0.96	0.55	0.55

N_a: Number of observed alleles; N_e: Number of effective alleles; H_o: Observed Hetrozygosity; H_e: Expected hetrozygosity (Average gene diversity within genotypes); F_{is}: Fixation index; PIC: Polymorphic information content.

Fixation index (Fis) which measures the population differentiation due to genetic structure ranged from 0.70 to 1.00 (Table 3). The highest fixation index of 1 was obtained from 20 markers. These were 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. On the other hand, the lower fixation index of 0.70 was recorded from marker Umc 1137 with a mean value of 0.96. This indicates the presence of appreciable levels of inbreeding or substantial reduction of hetrozygosity among maize germplasm used in the current study. This was

expected considering the breeding history of the inbred lines, which were developed with successive self-pollination for at least five cycles (S_5) of inbreeding.

Major allele frequency ranged from 0.34 to 0.95 with a mean of 0.55. Markers Umc 1415 and Umc 2205 exhibited the highest and the lowest allele frequency, respectively, indicating the presence of high genetic variations between the tested materials. These substantial amount of variations are commonly attributed to influences of evolutionary factors such as mutation, migration, recombination, selection and genetic drift. In the context of the current study, however, cyclical inbreeding followed by selection in the process of inbred line development might have contributed significantly to higher genetic variation among the inbred lines.

Each marker generated polymorphic information content (PIC) which is used to measure the informativeness of a genetic marker for linkage studies. The values obtained in the current study ranged from 0.10 (Bnlg 1063) to 0.74 (Umc 2205) with a mean of 0.50 (Table 3). Lopes *et al*. (2015) stated SSR loci with PIC value $\geq$0.5 as highly informative, 0.25<PIC<0.5 moderately informative and those with PIC$\leq$0.25 uninformative. Hence, according to the above classification, markers with PIC value $\geq$ 0.5 are highly informative in detecting differences among the genotypes based on their genetic background. Among the 28 SSR markers used in the present study, 19 (67.8%) were highly informative, 6 (21.4%) moderately informative and 3 (10.7%) uninformative (Table 3). Analysis of molecular variance indicated that the highest percentage (77%) of the variation was attributed to genetic variability among the individuals (inbred lines) within populations (Table 4). While 20% variation was attributed to genetic variability among the population, only 3% was attributed to variation within individuals confirming discriminating ability of most of the SSR loci among the inbred lines. It also indicated the presence of high genetic variation among studied materials. Hence, majority of the SSR markers used in the current study can be recommended for further use in genetic diversity study of the Ethiopian maize germplasm being highly informative in detecting differences among maize genotypes. The three markers (Bnlg 1063, Umc 2278, and Umc 1415) with PIC value < 0.25 had less discrimination ability for detecting genetic variation among maize inbred lines. Thus, these markers are not recommended for use in maize genetic diversity study in the future. The average PIC value observed in the current study was higher than other values reported in previous study. For instance, Sharma *et al.* (2017) reported an average PIC value 0.36 in 33 maize inbred lines using 40 SSR markers. The difference in the PIC value between the current study and that of Sharma and colleagues (2017) might be due to differences in number of SSR markers, genotypes and number of genotypes used for investigation. According to Kassahun and Prassana (2003) the overall PIC

value could be influenced by several factors. These are the nature of germplasm used for the study, number of SSR loci as well as inbred lines analyzed, SSR loci assayed, in terms of the nature and type of repeats, and methodology employed for allele detection, for example, agarose vs. PAGE.

Table 4. Analysis of molecular variance (AMOVA) of 37 elite maize inbred lines studied using 28 SSR makers

Source	Df	SS	Variance components	% of total variance	P
Among Populations	2	110.47	55.24	20	>0.001
Among Individuals with in population	34	463.11	13.62	77	>0.001
Within Individuals	37	9.50	0.26	3	>0.001
Total	73	583.08	7.99	100	

Df: Degrees of freedom, SS: Sum of squares and MS: Mean squares

Another study reported that di-nucleotide repeat loci identified more number of alleles and had higher PIC values (Legesse *et al*., 2007). However, in the current study, this relationship was not maintained. All markers with di-, tri-, tetra-, and penta-nucleotide repeats had similar mean PIC value of 0.59. While the di-nucleotide repeat markers had higher number of alleles (4.2), the other repeat markers had a mean of 3.8 alleles. The number of alleles might be even more than those reported provided polyacrylamide gel was used for resolving the PCR products. This is mainly due to the robustness of polyacrylamide gel to detect the polymorphism among PCR products differing in as low as two nucleotides. On the other hand, the agarose gel has poor detection for few nucleotide polymorphisms (Sibov *et al*., 2003). In the current study, most SSR loci amplified single band in most of the genotypes except in few inbred lines where more than one bands were amplified (Figure 1). This was expected on the basis of the breeding history of the inbred lines which were developed through repeated cycles of inbreeding which lead to develop homozygous dominant loci. However, the rare multiple bands could also be expected because of residual heterozygous loci that could still be maintained and then amplified because of co-dominant nature of the SSR markers. Contamination with foreign pollen sources or seed during seed multiplication and maintenance of the inbred lines could also lead to the amplification of more than one bands in addition to mutation that could naturally happen at some loci.

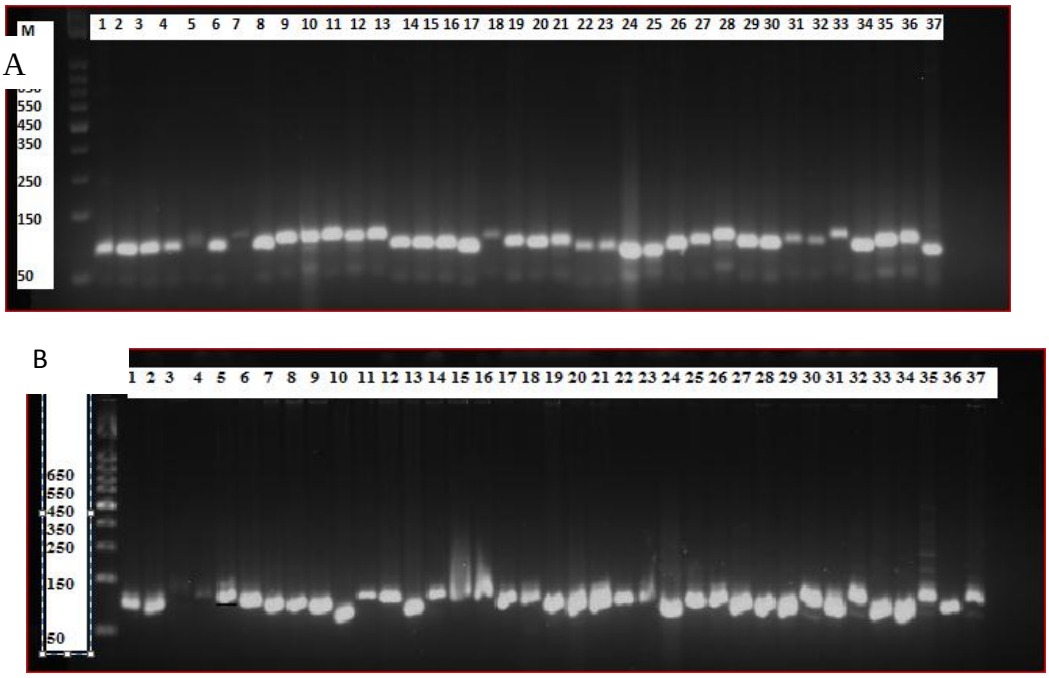

Figure 1. Microsatellite fingerprints amplified with SSR markers Umc 1857 (A) and Umc 2164 (B) on 37 elite maize inbred lines. **M**: A 100+50 bp ladder with sizes indicated in bp. Bands show the sizes of PCR products from the two SSR markers for 37 inbred lines indicated in Table 1.

Pattern of diversity and genetic relationship among populations

The mean number of alleles present in the populations ranged from 1.75 to 3.214 (Table 5). The highest number of alleles (3.214) were observed in the inbred lines obtained from Bako NMBRP (population 1) followed by soil acidity tolerant inbred lines (2.679) recently introduced from CIMMYT (population 3). Inbred lines previously introduced from CIMMYT and already used in the hybrid-breeding program (population 2) had the lowest number of alleles (1.75). This confirmed the presence of wider genetic variations among the populations. Inbred lines from Bako (population 1) showed the highest value for number of effective alleles (2.42), high value of Shannon information index (0.936), high value of observed hetrozygosity (0.028), and high value of expected hetrozygosity (0.538). On the other hand, inbred lines originally from CIMMYT (population 2) showed the lowest number of effective alleles (1.792), low Shannon information index (0.624), low observed hetrozygosity (0.010) and low expected hetrozygosity (0.390). These low values of CIMMYT lines for different parameters might be due to small number of inbred lines used in current study.

The overall mean of observed hetrozygosity (Ho) of 0.02 was lower than the corresponding average of expected hetrozygosity of 0.449, which indicated the presence of overall gain in homozygosity within the tested maize inbred lines.

This can be confirmed by visual observation on the agarose gels of the specific markers that revealed the existence of single band (Figure 1).

The fixation index (F) which indicates the population differentiation due to genetic structure ranged from 0.949 to 0.966. The highest fixation index was observed in acidity tolerant inbred lines (population 3). On the contrary, inbred lines from Bako (population 1) showed the lowest fixation index (0.949) with a mean of (0.957). The overall mean of fixation index indicates high level of inbreeding and substantial reduction of hetrozygosity. The homozygosity in the group was high due to selfing of the population. The degree of polymorphism varied from population to population. It ranged from 96.43% to 100%.

Table 5. Summary of genetic parameters for populations of elite inbred lines and 28 SSR markers

Population	Genetic parameters							
	N	N_a	N_e	I	H_o	H_e	F	%
Population 1 (Inbred lines from Bako NMBRP)	14	3.214	2.42	0.936	0.028	0.538	0.949	100
Population 2 (CIMMYT inbred lines already used in the breeding program)	10	2.321	1.792	0.624	0.010	0.390	0.957	92.86
Population 3 (soil acidity tolerant inbred lines introduced from CIMMYT in 2019)	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

Inbred lines sourced from Bako (population 1) had the highest level of polymorphism (100%), followed by inbred lines from population 3 (96.43 %), whereas the lowest amount of polymorphism was obtained from CIMMYT lines (population 2) (92.86%) with the average of 96.43%. The existence of high and wider range of polymorphism in the studied materials indicate the presence of high genetic variation and could be considered as a resource that can be exploited in future breeding programs.

Genetic distance and relationship among populations

The genetic relationship among populations can be evaluated using genetic distance measurements. Several methods were developed to examine the extent of 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 and relationships among populations and inbred lines. Genetic distance among the tested populations ranged from 0.27 to 0.42 (Table 6). The highest genetic distance (GD=0.42) was recorded between inbred lines from CIMMYT (population 2) and soil acidity tolerant inbred lines from the same source (population 3). Though the inbred lines contained in these

two populations were originally form CIMMYT, they might have been synthesized from germplasm of diverse genetic background. This might have contributed to wider genetic distance among the populations. Inbred lines in Population 3 were developed to resist soil acidity. The second largest genetic distance (GD=0.34) was observed between inbred lines from Bako (population 1) and inbred lines from CIMMYT (population 3), whereas the smallest genetic distance (GD=0.27) was observed between inbred lines from Bako (population 1) and CIMMYT (population 2). The existence of wider genetic distance among the inbred lines is important for future maize improvement program in terms of creating crosses which would yield wider genetic variation in segregation populations.

Principal coordinate analysis (PCoA)

The investigation of the Principal Co-ordinate Analysis (PCoA) showed that the first three most informative principal coordinates accounted for 69.37% of the genetic variations among maize inbred lines in the current study. The first, second, and third principal coordinates explained for 42.7%, 14.0% and 12.6%, respectively, of the total variations. The pattern of distribution of the inbred lines in the PCoA plot revealed three major clusters in the two-dimensional and three-dimensional coordinates (Figure 2). The principal coordinate analysis showed that inbred lines from diverse populations were grouped together; no separate group was formed for single population. The presence of seed exchange and gene flow between and within populations could be the probable reason for mixed clustering of inbred lines from different populations.

Table 6. Pair-wise population Nei's genetic distance showing the magnitude of genetic differentiation between the three populations of elite maize inbred lines

Population	Population 1 (Inbred lines from Bako NMBRP)	Population 2 (CIMMYT inbred lines already used in the breeding program)	Population 3 (soil acidity tolerant inbred lines introduced in 2019)
Population 1 (Inbred lines from Bako NMBRP)	0.00		
Population 2 (CIMMYT inbred lines already used in the breeding program)	0.27	0.00	
Population 3 (soil acidity tolerant inbred lines introduced in 2019)	0.34	0.42	0.00

The first principal coordinate clearly differentiated population 1 and 3 consisting of maize inbred lines of Ethiopian and acid tolerant lines of CIMMYT origin, respectively. However, considering the two-dimensional coordinate, there was no clear differentiation among inbred lines of the three populations, opposite to the clustering generated in Unweighted Neighbor Joining (NJ) tree. In the NJ tree, the three populations formed separate clusters except for few inbred lines that were

grouped in populations where they were not expected based on their genetic background. For example, CML144 and CML161 were grouped with inbred lines of Ethiopian origin (Population 1) and those from CIMMYT inbred lines (Population 2). Similarly four inbred lines of Ethiopian origin (F-7215, MBRC5BCF108-2-3-1-B-5-2-B-B-B-B, PO,OOE3-2-1-2-1 and SC22) and inbred line CML356 from Population 3 were grouped with CIMMYT inbred lines in Population 2 in cluster I. This might be due to gene movement among the inbred lines as a result of pollen contamination and seed mixture that could be encountered during seed multiplication and maintenance.

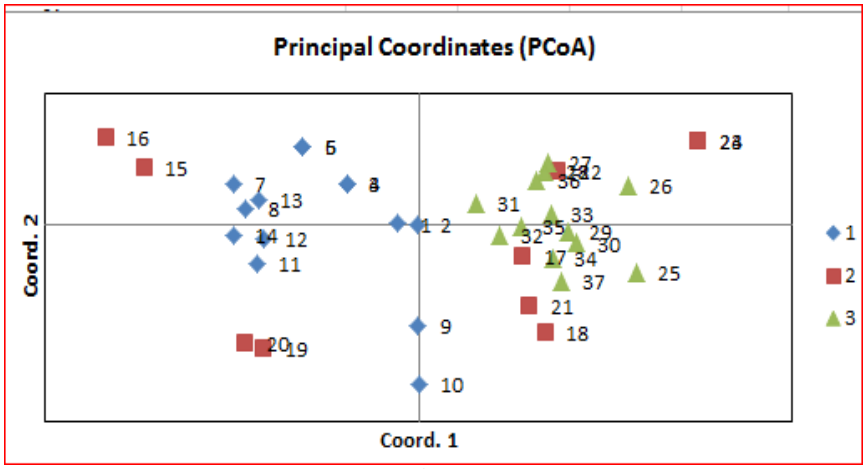

Figure 2. Principal coordinate analyses (PCoA) of the 37 maize inbred lines using SSR markers. Blue indicates inbred lines from Bako, Red refers to inbred lines introduced to Ethiopia longtime ago and were incorporated into the breeding program; and Green refers to lines recently introduced from CIMMYT. Names and descriptions of the 37 inbred lines are shown in Table 1.

CIMMYT inbred lines in Population 2 were introduced to Ethiopia long time ago and have been maintained at Bako Agricultural Research Center together with the Ethiopian inbred lines (Population 1). However, the acid tolerant inbred lines (Population 3) were introduced from CIMMYT in 2019. This recent introduction might be the reason for forming a separate cluster without mixing with inbred lines of the other two populations, which have been growing together for many years.

Cluster analysis and relationship among inbred lines

The NJ tree revealed that the 37 elite maize inbred lines were grouped into three major clusters consisting of 32.4%, 35.1%, and 32.5% of the total populations in clusters I, II and III, respectively, and forming different hierarchical sub-groups (Figure 3). The distribution of inbred lines in the three clusters followed clear pattern of grouping based on origin except few inbred lines that were grouped out of their populations forming mixture with inbred lines of other populations. Cluster I comprised mainly of inbred lines from CIMMYT (Population 2) and four inbred lines from Ethiopia (Bako) which were grouped under Population 1

considering their origin. Cluster II contained only inbred lines developed for tolerance to soil acidity (Population 3). This grouping confirms that the populations are genetically distant from each other as indicated by the Nei's genetic distance among the populations. Cluster III contained majority of inbred lines from Ethiopia, Bako (Population 1) mixed with two inbred lines of CIMMYT origin (Population 2). This is due to the presence of gene flow between these two populations as explained earlier. Cluster I was further divided into two sub-clusters. It comprised of seven inbred lines (CML 444, CML 536, CML 395, CML 165, CML 197, CML 202 and CML 312) of CIMMYT origin (Population 2) and three inbred lines (F-7215, MBRC5BCF108-2-3-1-B-5-2-B-B-B-B and PO,OOE3-2-1-1-2-1) of Ethiopian origin (Population 1) in sub-cluster I and two inbred lines (CML 334 and SC 22) of CIMMYT and Ethiopian origin in sub-cluster II. Cluster II contained inbred lines developed for tolerance to soil acidity (Population 3) alone. It comprised of 13 genotypes, namely 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. These inbred lines are genetically distinct from those in other populations. Hence are important in further genetic improvement in Ethiopia. Cluster III was also divided into two sub-clusters. It comprised of inbred lines from Bako (Population 1). In sub-cluster I, six lines, namely 124-b(109), 124-b(113), 142-1-e, 144-7-b, 35B-190-0510-2-1-2-2-1-2 and A-7033 were grouped while in the second sub-cluster another six CML 161, BKL 001, BKL 002, CML 144, BKL 003 and BKL 004 were grouped.

It is appreciable to see parental lines of some of the commercial hybrids to be grouped into different clusters. It is important to note that inbred lines CML161 and CML165 are parental lines of the Quality Protein Maize (QPM) hybrid BHQPY545, which was released as improved variety in Ethiopia in 2008, being assigned to different clusters. The grouping of these inbreds in separate clusters indicates the validity of the SSR markers in substantiating the heterotic grouping of the inbred lines based on phenotypic data. Similarly, the single cross commercial hybrid BH 540 was developed from inbred lines SC22 X 124-b (113) which were also grouped into different clusters yielding added evidence indicating validity of the SSR markers. On the other hand some of the inbred lines grouped together specially in sub-clusters are sister lines. This could be observed in the case of the four inbred lines, BKL 001, BKL 002, BKL 003 and BKL 004 which were in the same sub-clusters indicating validity of the SSR marker in grouping inbred lines based on their pedigree relationship. The two sister inbred lines 124-b(113) and 124-b(109) were also grouped together owning to closer genetic background.

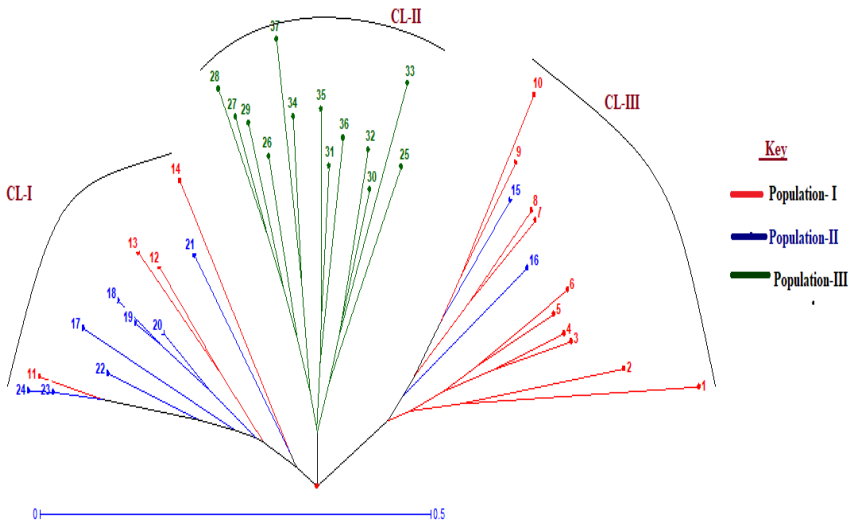

Figure 3. Unweighted Neighbor Joining (NJ) dendrogram showing genetic relationship of 37 maize inbreed lines using 28 microsatellites. Names and descriptions of the 37 inbred lines are shown in Table 1. Blue indicates inbred lines from Bako, Red refers to inbred lines introduced to Ethiopia longtime ago and incorporated into the breeding program; and Green refers to lines recently introduced from CIMMYT.

Genetic distance among the inbred lines

Nei's genetic distance among pairs of the 37 maize inbred lines varied from 0.16 to 0.98 for the inbred partners CML444 and CML536, and 124-b(109) and MBRC5BCF108-2-3-1-B-5-2-B-B-B-B, respectively (data not shown). The genetic distance obtained among the inbred lines was higher than reported in previous studies in tropical maize using single nucleotide polymorphic marker (Eritro *et al*., 2017 and Semagn *et al*., 2012) and genotyping by sequencing (Ogugo *et al*., 2015). In general, the result indicated broad genetic variation among the inbred lines. This could be explained in light of the inbred lines, which were used in the study. Most of the inbred lines were elite selected based on their diverse genetic background and used in hybrid breeding program and are currently parental lines of commercial hybrids in Ethiopia. Our prime interest was to look at genetic distance among inbred lines of Ethiopian and CIMMYT origin instead of among inbred lines of CIMMYT origin alone since our main target is to broaden the genetic basis of locally developed inbred lines. However, we also wished to know how locally developed inbred lines are also distantly related. When we look at the genetic distance among inbred lines in the different populations, the range was from 0.26 to 0.98 among inbred lines of Ethiopian origin and the highest genetic distance was in this population. Out of 84 pairs of inbred lines in the population 37(44%) had genetic distance of more than 0.7 and 19(22.6%) had genetic distances of more than 0.8; pairs with genetic distance of more than 0.9 were 124-b(109) and F-7215, 124-b(109) and MBRC5BCF108-2-3-1-B-5-2-B-B-

B-B, BKL004 and F-7215, BKL004 and MBRC5BCF108-2-3-1-B-5-2-B-B-B-B. The genetic distance among 124-b (109) and F-7215, and BKL004 and F-7215 could be expected owing to the diverse genetic background of the source germplasm from which the inbred lines were developed. Inbred lines 124-b(109) and BKL004 were developed from Ecuador 573 population while F-7215 was developed from Kitale synthetic II population. These two populations were from different heterotic groups and have been used widely as sources of inbred lines in the hybrid-breeding program of Ethiopia. The study also revealed higher genetic distances among inbred lines developed from the same source germplasm. This could be seen in the case of the two pairs, 124-b(109) and MBRC5BCF108-2-3-1-B-5-2-B-B-B-B, and BKL004 and MBRC5BCF108-2-3-1-B-5-2-B-B-B-B which were developed from Ecuador 573 population.

To have an idea of relevance of the genetic distance data generated with the SSR markers used in the study it would be of paramount importance to review the values of the genetic distance among inbred lines which are parental lines of released hybrids. Inbred pairs 124-b(109) and SC 22, and 124-b(113) and SC 22 are parents of the commercial hybrids BH543 and BH540, respectively. They had corresponding genetic distance of 0.79 and 0.76, respectively, substantiating the validity of the SSR markers in detecting genetic divergence among heterotic inbred lines. Genetic distances among the two pairs seemed to be equal in magnitude due to close genetic relations among the two inbred 124-b(109) and 124-b(113) which are sister lines. In hybrid maize breeding program the single cross female parents of a three-way cross hybrid are usually established among distantly related inbred lines to maximize seed production of the F_1 three way cross hybrids. The current study indicated that A-7033 and F-7215, and BKL002 and CML312, which are parents of the single cross female parents of BH660 and BH661, had genetic distances of 0.89 and 0.6, respectively. On the other hand, genetic distances among CML161 and CML165, and CML395 and CML202 that are, respectively, parents of BHQPY545 and the female parent of BH546 had lower genetic distances of 0.41 and 0.37 against our expectation. This could be due to low robustness of the SSR markers in identifying genetic distance among the lines. In such cases, other markers with wide genome coverage must be used.

In view of broadening the genetic basis of maize for tolerance to soil acidity in Ethiopia, genetic distances among the Ethiopian and CIMMYT inbred lines deserve attention. In this regard, pairs of Ethiopian and CIMMYT inbred lines developed for tolerance to soil acidity had genetic distance that varied from 0.42 to 0.95 and 41.6% of the pairs of inbred lines had genetic distance of more than 0.8 indicating their potential in broadening the genetic basis of maize for developing varieties tolerant to soil acidity stress. These pairs of inbred lines could be exploited in pedigree breeding through which segregating populations can be

selected for desirable traits in combination with tolerance to soil acidity. Among the Ethiopian maize inbred lines 124-b(109) and F-7215 had uniquely higher genetic distance with five soil acidity tolerant inbred lines. Specifically, 124-b(109) was distantly related with five inbred lines, viz., CML359, CML360, CML361, CML436 and CML438, and F-7215 was distantly related with CML361 with Nei`s genetic distance value of more than 0.9. The soil acidity tolerant inbred lines had moderate genetic distance with the CIMMYT inbred lines introduced previously and used in breeding program with value of more than 0.5 almost in all combinations. Compared to their combination with inbred lines of Ethiopian and CIMMYT origin, the soil acidity tolerant inbred lines had lower genetic distance among themselves. This could be expected based on the source germplasm from which they were developed. They were developed from South American populations that were genetically related. Hence, cross breeding among themselves may not yield hybrids with better yield performance.

In conclusion, the study has brought in new inbred lines, which could be exploited in broadening the genetic basis of local germplasm for developing varieties tolerant to soil acidity stress. However, these inbred lines have to be evaluated for adaptation to Ethiopian environment before embarking on exploiting their potential in breeding for tolerance to soil acidity.

References

Abate T, B Shiferaw, A Menkir, D Wegary, Y Kebede, K Tesfaye, M Kassie, G Bogale, B Tadesse, and T Keno. 2015. Factors that transformed maize productivity in Ethiopia. *Food Security* 7(5): 965-981.

Beyene Y, AM Botha, and AA Myburg. 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.

Cavalli- Sforza LL and AW Edwards. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 21(3): 550-570.

CSA (Central Statics Agency). 2018. The Federal Democratic Republic of Ethiopia Central

Demissew DA. 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.

Ertiro BT, K Semagn, B Das, M Olsen, M Labuschagne, M Worku, D Wegary, G Azmach, V Ogugo, T Keno, and B Abebe. 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.

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

FAOSTAT. 2018. Food and Agricultural Organization of the United Nations (FAO), FAO Statistical Database. URL: http://faostat.fao.org accessed December 7, 2019.

Gupta ASTHA and AK Singh. 2010. Studies on molecular diversity in maize inbred using SSR markers. Trends Biosci, **3(2):**106-109.

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

Legesse BW, AA Myburg, KV Pixley, and AM Botha. 2007. Genetic diversity of African inbred lines revealed by SSR markers. *Hereditas* 144: 10-17.

Liu K. and SV Muse. 2005. Power Marker: an integrated analysis environment for genetic

Lopes AD, CA Scapim, MFPS Machado, CA Mangolin, TA Silva, LB Cantagali, FF Teixeira,. and F Mora. 2015. Genetic diversity assessed by microsatellite markers in sweet corn cultivars. *Scientia Agricola* 72(6): 513-519.

Lu H and R Bernardo. 2001. Molecular markers diversity among current and historical maize inbred. *Theor. App. Genet.*100: 552-556.

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

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

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

Ogugo V, K Semagn, Y Beyene, S Runo, M Olsen, M Warburton. 2015, Parental genome contribution in maize DH lines derived from six backcross populations using genotyping by sequencing. Euphytica.5;202(1):129–39.

Pandit M, M Chakraborty, ZA Haider, A Pande, RP Sah, and K Sourav. 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.

Peakall R and PE Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics.28(19): 2537–2539.

Schlede H. 1989. Distribution of acid soils and liming materials in Ethiopia. *Ethiopian Institute of Geological Survey. Report Note* 326

Scott KD, P Eggler, P Seaton, G Rossetto, M Ablett, EM Lee, LS, and RJ Henry. 2000. Analysis of SSRs derived from grape ESTs. *Theoretical and Applied Genetics 100*(5): 723-726.

Semagn K, Y Beyene, D Makumbi, S Mugo, BM Prasanna, C Magorokosho, and G Atlin. 2012. Quality control genotyping for assessment of genetic identity and purity in diverse tropical maize inbred lines. *Theor Appl Genet*125(7): 1487-501.

Senior ML, JP Murphy, MM Goodman, and CW Stuber. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Science* 38(4): 1088-1098.

Sharma S, DP Mishra, B Kumar, H Kaur, and HR Jewlia. 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.

Sharopova N, R Velasco, D Richard, RD Thompson, and M Liscum. 2014. Development and mapping of SSR markers for maize. *Plant Molecular Biology* 48: 463–481.

Shiferaw B, BM Prasanna, J Hellin, and M Bänziger. 2011. Crops that feed the world 6. Past

Sibov ST, CL de Souz Jr, AAf Garica. 2003. Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. *Hereditas* 139: 96_106.

Statistical Agency. Report on Area and Production of Major Crops. CSA Statistical Bulletin Statistical Bulletin 589. Addis Ababa, Ethiopia.

Takezaki N and M Nei. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144(1): 389-399.

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

Tolessa B, K Mulatu, L Wolde, M Worku, and Leta Tulu. 1997. Reflections on the successful achievements of hybrid maize breeding in Ethiopia. *In:* Ransom JK, Palmer AFE, Zambezi, BT, Waddington SR, Pixley KV, Jewell DC. (Eds) Maize Productivity Gains Through Research and Technology Dissemination. Proceedings of the Fifth Eastern and Southern Africa Regional Maize Conference, 3-7 June 1996 Arusha, Tanzania. CIMMYT, Addis Ababa. pp. 67-71.

Tolessa B, T Gobezayehu, M Worku, Y Desalegne, K Mulatu, and G Bogale. 1993. Genetic improvement of maize in Ethiopia. *In*: B Tolessa and JK Ransom. (Eds) Proceedings of the First National Maize Workshop of Ethiopia; 5–7 May 1992, Addis Ababa, Ethiopia. IAR/CIMMYT, Addis Ababa.pp. 13–22.

Tulu L. 2004. Heterosis and Genetic Diversity in Crosses of Seven East African Maize (*Zea*

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

Vaz patto M, CZ Satovic, S Peg, and OP Fevereiro. 2004. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. *Euphytica*. 137: 63-67

Warburton ML, X Xianchun, J Crossa, J Franco, AE Melchinger, M Frisch, M Bohn and Hoisington, D. 2002. Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. *Crop Science 42*(6): 1832-1840.

Worku M, L Wolde, B Tolessa, K Mulatu, and L Tulu. 1996. Heterotic pattern of some intermediate maturing maize germplasm. *African Crop Science Journal* 4: 497-501.